

THESE

pour obtenir le titre de

Docteur en Sciences
de l'Université de Nice-Sophia Antipolis

Spécialité : Sciences de l'Univers - Géophysique

présentée et soutenue par

Romain BROSSIER

Imagerie sismique à deux dimensions des milieux visco-élastiques par inversion des formes d'ondes : développements méthodologiques et applications

Thèse dirigée par **Jean VIRIEUX** et **Stéphane OPERTO**
préparée au laboratoire Géoazur, Sophia-Antipolis
financée dans le cadre du Consortium SEISCOPE

soutenue le 30 Novembre 2009 devant le Jury

Michel DIETRICH	Professeur Institut Français du Pétrole	Examineur
Dimitri KOMATITSCH	Professeur Université de Pau	Examineur
Gilles LAMBARE	Chercheur Senior CGGVeritas	Examineur
Stéphane LANTERI	Directeur de Recherche INRIA Sophia-Antipolis	Membre invité
Guust NOLET	Professeur Université Nice-Sophia Antipolis	Membre invité
Stéphane OPERTO	Chargé de Recherche CNRS	Directeur de Thèse
René Edouard PLESSIX	Chercheur Senior SHELL	Rapporteur
Satish SINGH	Professeur Institut de Physique du Globe de Paris	Rapporteur
Jean VIRIEUX	Professeur Université Joseph Fourier	Directeur de Thèse

Remerciements

Résumé

La connaissance des structures internes de la Terre, à différentes échelles, présente des enjeux majeurs d'ordres économiques, humains, environnementaux et scientifiques. Diverses méthodes d'imagerie ont été développées en utilisant les informations contenues dans les ondes sismiques. La méthode d'inversion des formes d'ondes construit des images quantitative haute résolution des paramètres physiques du sous-sol, en exploitant le champ d'onde complet, sous la forme d'un problème d'optimisation. Dans ce travail de thèse, je présente l'application de l'inversion des formes d'ondes en domaine fréquentiel, pour imager les paramètres visco-élastiques dans des géométries à deux dimensions à grands offsets.

Dans un premier temps les développements méthodologiques et algorithmiques sont présentés. La modélisation de la propagation des ondes P-SV en domaine fréquentiel, le problème direct du processus d'imagerie, est assurée par une méthode d'éléments finis Galerkin discontinus, assurant une grande flexibilité dans le choix des ordres d'interpolation et dans l'utilisation de maillages triangulaires non-structurés. Le problème inverse est résolu sous une forme linéarisée, afin de limiter le nombre de simulations directes, et utilise l'algorithme quasi-Newton L-BFGS permettant de tirer bénéfice de l'estimation « économique » du Hessien. Le processus global d'imagerie est implémenté sous la forme d'un algorithme massivement parallèle destiné aux calculateurs modernes à mémoire distribuée.

Dans un deuxième temps, les algorithmes développés sont appliqués à des cas d'étude. Des applications sont menées dans des modèles synthétiques réalistes représentatifs d'environnements terrestres et marins. Ces études montrent les difficultés associées à la reconstruction des paramètres élastiques à partir de données mettant en jeu des phénomènes de propagations complexes (ondes converties, multiples, ondes de surfaces...). Des solutions sont proposées sous forme de processus hiérarchiques multi-échelles, afin de limiter les effets des non-linéarités du problème inverse et ainsi d'améliorer la convergence du processus vers le minimum global. Enfin, la sensibilité de différentes normes et critères de minimisation est analysée, à partir de données bruitées issues de modèles synthétiques réalistes, ainsi que sous l'approximation acoustique pour un jeu de données réelles pétrolière. Ces tests montrent certaines limites du formalisme classique basé sur la norme L_2 dans l'espace des données, tandis que la norme L_1 apparaît comme alternative robuste pour l'inversion de données décimées en domaine fréquentiel.

Abstract

The knowledge of Earth intern structures at different scales is of major interest for economy, humans, environment and science. Several methods have been developed for Earth imaging using seismic wave information. The full waveform inversion attempts to build quantitative high-resolution images of the subsurface physical parameters using the full wavefield, solved as an optimization procedure. In this thesis, I present application of two-dimensional frequency-domain full waveform inversion for imaging visco-elastic parameters from large offsets data.

In a first time, methodologies and algorithms are presented. The frequency-domain P-SV waves propagation modelling, the forward problem of the inversion process, is solved with a finite element discontinuous Galerkin method. This method allows a flexible choice of interpolation orders and the use of triangular unstructured meshes. The inverse problem is linearized in order to limit the number of forward problem simulations, and solved with the quasi-Newton L-BFGS algorithm in order to exploit information contained in the estimated Hessian matrix. The full imaging process is implemented in a massively-parallel algorithm for the distributed-memory architectures of modern computing centres.

In a second time, algorithms are applied to several case studies. Applications are performed in realistic synthetic models, representative of onshore and offshore environments. These studies show the difficulties associated to elastic parameters imaging from complex data including converted waves, multiples and surfaces waves. Several multi-scale hierarchical procedures are proposed in order to mitigate non-linearity of the inverse problem and to improve convergence toward the global minimum. Finally, a sensitivity study is performed to analyse the behaviour of several minimization norms and criteria, when noisy data are inverted. A first application in realistic synthetic models is presented before an acoustic application to field data. These tests show some limits of the classic L_2 norm in the data space, while the L_1 norm appears to be a robust alternative for frequency-domain inversion of decimated data.

Table des matières

Introduction générale	15
I Théorie et algorithmes	23
1 La modélisation de la propagation des ondes élastiques en 2D	25
1.1 Équations de l'élastodynamique linéaire	26
1.1.1 Équilibre des forces : conservation de la quantité de mouvement	27
1.1.2 Linéarité déplacement/déformation	27
1.1.3 Linéarité déformation/contrainte : loi de comportement	28
1.1.4 Milieux anisotropes <i>versus</i> isotropes, élastiques <i>versus</i> acoustiques... quelle approximation choisir ?	28
1.1.5 Les équations de propagation 3D	30
1.1.6 Quelles ondes ?	31
1.1.7 L'approximation 2D de la propagation des ondes	32
1.2 Méthodes de résolution des systèmes d'équations aux dérivées partielles	34
1.2.1 Introduction	34
1.2.2 Différences finies	34
1.2.3 Méthodes pseudo-spectrales	36
1.2.4 Éléments finis	36
1.2.5 Éléments spectraux	37
1.2.6 Galerkin discontinus	37
1.3 Conditions aux limites	38
1.3.1 Conditions initiales	38
1.3.2 Terme source	38
1.3.3 Conditions aux limites spatiales	39
1.4 Modélisation en domaine temporel ou fréquentiel ?	40
1.4.1 Problématique	40
1.4.2 Approche temporelle par schéma explicite d'intégration	40
1.4.3 Approche fréquentielle : solveur direct	41
1.4.4 Approche fréquentielle : solveur itératif	42
1.4.5 Approche fréquentielle : solveur hybride	43
1.4.6 Choix de l'approche fréquentielle + solveur direct	43
1.5 La discrétisation par éléments finis Galerkin discontinus en domaine fréquentiel	44
1.5.1 Système 2D P-SV en fréquence et conditions absorbantes PML	44
1.5.2 Transformation du système	45

TABLE DES MATIÈRES

1.5.3	Discrétisation spatiale	46
1.6	Application GD P_0 ou volume fini	53
1.6.1	Présentation de l'article	53
1.6.2	Résumé	54
1.6.3	Introduction	55
1.6.4	Formulation volume fini en domaine fréquentiel	57
1.6.5	Propriétés numériques	60
1.6.5.1	Analyse de dispersion numérique	60
1.6.5.2	Conditions absorbantes PML	61
1.6.5.3	Implémentation de la source	63
1.6.6	Résultats numériques	65
1.6.6.1	Comparaison avec solutions analytiques	65
1.6.6.2	Comparaison avec solutions numériques	68
1.6.6.3	Tests numériques en maillages non-structurés	73
1.6.7	Coût numérique de la méthode	74
1.6.8	Discussion et conclusion	75
1.6.9	Annexe A : équivalence VF P_0 et $O(\Delta x^2)$ DF sur grilles carrées régulières	77
1.6.10	Annexe B : développement volume fini	79
1.6.11	Conclusion	80
1.7	Extension aux ordres supérieurs P_1 et P_2	80
1.7.1	Introduction	80
1.7.2	Formulation	81
1.7.3	Approche <i>parsimonieuse</i>	82
1.7.4	Implémentation du terme de source	84
1.7.5	Application numérique	85
1.8	Conclusion	88
2	Le problème inverse : théorie, méthodes et problématiques	91
2.1	Problème aux moindres carrés linéarisé	92
2.1.1	La fonction coût	92
2.1.2	Linéarisation du problème inverse	93
2.1.3	Les équations normales	94
2.1.4	Estimation de la signature de la source	95
2.2	Gradient et Hessien	95
2.2.1	Interprétation et construction du gradient	95
2.2.2	Pouvoir de résolution spatiale du gradient	97
2.2.3	Non-linéarité et approche multi-échelles	98
2.2.4	Interprétation et construction du Hessien	100
2.3	Résolution numérique du problème d'optimisation locale	101
2.3.1	Méthode du Gradient	101
2.3.2	Le pas de descente α	101
2.3.3	Le Gradient Conjugué	102
2.3.4	Méthode de Newton et Gauss-Newton	103
2.3.5	Méthodes de Quasi-Newton	103
2.3.6	Préconditionnement	104
2.3.7	Régularisations	104
2.3.8	Exemple	105

2.4	Le modèle initial	106
2.4.1	Modèle cinématiquement compatible et saut de phase	106
2.4.2	Tomographie des temps de première arrivée	107
2.4.3	L'inversion dans le domaine de Laplace	108
2.4.4	Inversion de la phase de première arrivée	109
2.4.5	La stéréo-tomographie	110
2.4.6	Conclusions	110
2.5	Les alternatives à la norme L_2	110
2.5.1	Fonction Logarithmique : Amplitude vs Phase - lien avec Rytov	110
2.5.2	Fonctionnelles robustes	111
2.6	L'inversion multiparamètres	112
2.7	Conclusion	114
3	Algorithmes parallèles, optimisation et performance	117
3.1	Résumé	118
3.2	Introduction	118
3.3	Problème direct	120
3.3.1	Théorie et discrétisation	120
3.3.2	Quelle interpolation ?	122
3.4	Problème inverse	125
3.4.1	Revue de la théorie	125
3.4.2	Algorithme d'inversion	128
3.4.3	Calcul du gradient	128
3.4.4	L-BFGS et inversion multiparamètres	130
3.5	Parallélisation et performances	133
3.5.1	Problème direct	133
3.5.2	Problème inverse parallèle	134
3.5.3	Deux niveaux de parallélisme	134
3.5.4	Performance	135
3.6	Conclusion	136
II	Applications	139
4	Exemples synthétiques	141
4.1	Application terrestre : Le modèle SEG/EAGE Overthrust	142
4.1.1	Résumé	142
4.1.2	Introduction	143
4.1.3	Méthode et algorithme	145
4.1.3.1	Problème direct	145
4.1.3.2	Problème inverse	146
4.1.3.3	Implémentation parallèle	148
4.1.4	Préconditionnement des données et stratégies multi-échelles	148
4.1.5	Application à un modèle canonique	150
4.1.6	Application à un modèle synthétique terrestre	151
4.1.6.1	Modèle SEG/EAGE overthrust et configuration expérimentale	151
4.1.6.2	Inversion de données brutes	155

4.1.6.3	Inversion de données préconditionnées par fréquences discrètes successives	156
4.1.6.4	Inversion sans surface-libre	157
4.1.6.5	Inversion de données préconditionnées par groupes de fréquences	157
4.1.6.6	L-BFGS <i>versus</i> gradient conjugué préconditionné	160
4.1.6.7	Performances	161
4.1.7	Discussion	164
4.1.8	Conclusion	168
4.1.9	Test complémentaire : combien de composantes sont nécessaires?	169
4.2	Application marine : Le modèle Valhall	172
4.2.1	Présentation du modèle élastique de Valhall	172
4.2.2	Configuration de l'inversion	172
4.2.3	Inversion acoustique de données élastique	174
4.2.4	Inversion élastique hiérarchique	176
4.3	Conclusion	177
5	Sensibilité de l'inversion au bruit	181
5.1	Quelle norme pour une inversion de formes d'onde robuste?	182
5.1.1	Résumé	182
5.1.2	Introduction	182
5.1.3	Théorie et algorithme	184
5.1.3.1	La norme aux moindres carrés L_2	184
5.1.3.2	La norme L_1	185
5.1.3.3	Le critère de Huber	185
5.1.3.4	Le critère hybride L_1/L_2	186
5.1.3.5	Interprétation	186
5.1.3.6	Algorithme	187
5.1.4	Tests numériques : le modèle marin de Valhall	189
5.1.4.1	Configuration de l'inversion	189
5.1.4.2	Résultats	191
5.1.5	Tests numériques : le modèle terrestre SEG/EAGE Overthrust	194
5.1.5.1	Configuration de l'inversion	194
5.1.5.2	Résultats	195
5.1.6	Discussion	196
5.1.6.1	Le modèle marin de Valhall	196
5.1.6.2	Le modèle terrestre SEG/EAGE Overthrust	197
5.1.6.3	Implications pour l'inversion des formes d'ondes 3D	197
5.1.7	Conclusions	198
5.2	Application acoustique sur les données réelles de Valhall	200
5.2.1	Environnement, données et configuration de l'inversion	200
5.2.2	Résultats	203
5.2.2.1	Ligne 21	203
5.2.2.2	Ligne 29	204
5.2.3	Conclusion	205
5.3	Conclusion	207
	Conclusion et perspectives	211

Bibliographie	217
A Calcul des matrices de la formulation GD par intégration numérique	233
A.1 Rappels sur l'intégration numérique	233
A.2 Exemples sur interpolation P_1	234
B Publications complémentaires	237
B.1 Two-dimensional seismic imaging of the Valhall model from synthetic OBC data by frequency-domain elastic full-waveform inversion	238
B.2 Robust frequency-domain full-waveform inversion using the L_1 norm	243

Introduction générale

L'imagerie de la Terre : comment et pourquoi ?

La planète bleue sur laquelle nous vivons fascine l'humanité dans son long cheminement. Certains peuples en ont fait des dieux et des déesses, les philosophes de l'antiquité étudiaient déjà sa forme et son interaction avec les autres astres, Jules Verne imaginait un voyage dans ses profondeurs... Avec les avancées des connaissances et de la technologie, les scientifiques se sont intéressés plus récemment aux structures internes de la Terre en utilisant des informations indirectes : les ondes sismiques qui traversent le globe (fin du XIX^{ème} siècle).

En effet, la connaissance des structures du sous-sol par observation directe est très difficile. Elle nécessite des forages mettant en jeux des défis technologiques conséquents pour supporter les pressions et températures des profondeurs¹, ainsi que des moyens financiers considérables.

La connaissance de la Terre interne est donc limitée, en pratique, à l'utilisation de phénomènes physiques menant à des interprétations indirectes. En effet, la Terre peut être vue comme un filtre, une fonction de transfert qui modifie la propagation des phénomènes physiques. Ces derniers gardent donc une trace, une signature du chemin parcouru dans le sous-sol.

En fonction de l'échelle d'investigation, différentes physiques sont utilisées. Ainsi, pour l'imagerie de la proche surface (quelques mètres), les ondes électromagnétiques peuvent être utilisées. Cependant, leur forte atténuation dans la subsurface ne permet pas d'envisager de grandes distances de propagation. Au contraire, les ondes élastiques ou ondes sismiques, peuvent se propager sur plusieurs milliers de kilomètres si la source est suffisamment puissante. Les ondes sismiques, issues des gros séismes, ont historiquement été les premières à avoir été utilisées pour produire des cartes et des images des grandes structures profondes de la Terre : discontinuité de Mohorovičić, de Gutenberg, noyau liquide, graine...

Les motivations d'une connaissance approfondie des structures souterraines sont multiples et présentent des enjeux économiques, humains, environnementaux, scientifiques : la connaissance de la proche surface (quelques centimètres à quelques mètres) est indispensable pour la caractérisation des sols en génie civil et pour les applications géotechniques. L'imagerie des glissements de terrain est très utile pour l'estimation du risque gravitaire. Plus profondément (quelques mètres à quelques kilomètres), la détection et la caractérisation de réservoirs d'hydrocarbures présentent un enjeu économique majeur, à l'heure où les ressources sont de plus en plus difficiles à détecter et à extraire. Le suivi dans le temps de zones d'enfouissement de

¹Le forage le plus profond réalisé jusqu'à aujourd'hui est le forage SG3 près de Zapolyarny dans le péninsule de Kola en Russie. Ce forage, commencé en 1970, avait pour objectif 15 km de profondeur pour atteindre la discontinuité du Moho. Il fut arrêté en 1989 à 12.262 km de profondeur, officiellement pour des raisons politiques, mais les fortes contraintes techniques peuvent également avoir motivé ce choix.

déchets radioactifs ou d'injection de CO₂ représente également un challenge environnemental crucial pour les prochaines années. Avec le changement climatique global, la détection d'aquifères va se révéler un enjeu humain de premier ordre. La caractérisation des zones de failles actives est essentiel pour la compréhension des cycles sismiques et l'estimation du risque associé. A l'échelle crustale et lithosphérique, l'imagerie des géométries des structures tectoniques permet de comprendre les phénomènes mécaniques mis en jeux, et également d'améliorer notre connaissance des processus sismiques associés. Enfin, l'imagerie des zones profondes permet d'améliorer notre connaissance des phénomènes géodynamiques, marqueurs de l'histoire de la Terre.

Les sources sismiques

Les ondes sismiques utilisées pour l'imagerie du sous-sol peuvent avoir deux types d'origine : les sources naturelles (séismes) ou anthropiques.

Les séismes de magnitudes diverses procurent une grande quantité de sources sismiques. Les enregistrements des champs d'ondes issus de ces séismes, par des réseaux de capteurs déployés à diverses échelles, permettent d'établir des catalogues de données très souvent utilisés pour l'imagerie de la Terre. Notons cependant que pour les séismes, les paramètres de la source ne sont pas connus : la position de la source, son rayonnement spatial ainsi que sa signature temporelle sont inconnus et doivent donc être estimés.

Les sources d'origines anthropiques sont largement utilisées dans les campagnes de sismique active ayant pour but d'imager certaines régions du sous-sol à différentes échelles. Elles peuvent être explosives avec un rayonnement relativement isotrope (tir d'explosif, canon à air pour les acquisitions marines) ou directionnelles (camion vibreur, chute de poids, pot vibrant, marteau). Les expériences de sismique active permettent, d'une part, de maîtriser les densités et les localisations des dispositifs de sources et de capteurs, qui vont, par la suite, déterminer la robustesse des méthodes d'imagerie. D'autre part, elles permettent d'obtenir une estimation du rayonnement spatial des sources et de leurs signatures temporelles.

Notons que l'énergie des sources sismiques contraint directement le spectre fréquentiel des ondes émises. Ainsi, les gros séismes génèrent de très basses fréquences qu'il est impossible de reproduire avec des sources sismiques anthropiques conventionnelles. Cette considération est importante car les spectres des données sismiques dépendent en grande partie des sources utilisées, et vont par la suite contraindre la robustesse et la résolution des méthodes d'imagerie mises en jeux.

Pour les applications présentées dans ce travail, nous nous placerons dans un contexte de sismique active, même si les méthodes et outils développés ne sont pas limités à ce cas de figure.

La mesure des champs d'ondes

La mesure des champs d'ondes issus des sources sismiques peut s'effectuer par divers types de capteurs.

Dans le contexte d'acquisitions terrestres, des géophones mesurant les vitesses de déplacement des particules sont généralement disposés à la surface du globe. Des capteurs accélérométriques peuvent également être utilisés. Ces capteurs sont directionnels et mesurent la vitesse

ou l'accélération dans une direction spatiale. Des dispositifs multicomposantes sont souvent utilisés afin de mesurer les champs d'ondes dans les directions horizontales (parallèles à la surface de la Terre) et la direction verticale, permettant ainsi d'identifier les différents types d'ondes dont les signatures dépendent des directions. Ces dispositifs peuvent être utilisés pour des sources naturelles ou anthropiques.

Dans le contexte marin, les acquisitions les plus répandues pour les expériences de sismique active sont des hydrophones localisés dans l'eau : un bateau génère la source sismique par un canon à air et tracte, sous la surface de l'eau, une flûte d'hydrophones qui mesurent le champ de pression portant la signature des ondes propagées dans les structures du sous-sol. Des contraintes techniques limitent les longueurs possibles des flûtes. Pour des acquisitions à longs offsets, plusieurs bateaux sont donc nécessaires. Cependant, les mesures restent limitées au champ de pression qui ne porte pas la signature de tous les types d'ondes se propageant dans le sol (les ondes de cisaillement ne se propageant pas dans les liquides, leurs signatures sont donc très faibles dans les données de pression mesurées dans l'eau). Une alternative possible est l'utilisation de dispositifs en fond de mer. Les capteurs tels que les « sismomètres fond de mer » (*Ocean Bottom Seismometer* (OBS)) sont des dispositifs autonomes qui se fixent en fond de mer pour une durée déterminée (quelques semaines à quelques mois). Ces dispositifs sont généralement équipés de géophones multicomposantes permettant d'enregistrer les différents types d'ondes, ainsi que d'un hydrophone. Les OBS étant mobiles et autonomes, ils peuvent être disposés sur des géométries d'acquisition à grands offsets. Des dispositifs d'acquisition par câbles disposés en fond de mer (*Ocean Bottom Cable* (OBC)), équipés de capteurs multicomposantes, peuvent également être utilisés pour les acquisitions de sismique active. L'intérêt majeur de ces dispositifs est de pouvoir rester en place plusieurs années, permettant d'assurer le monitoring de certaines cibles ou réservoirs par des campagnes successives dans le temps.

La géométrie des dispositifs d'acquisition conditionne les types d'ondes potentiellement enregistrées. Pour les campagnes de sismique active, des dispositifs à courts offsets ont historiquement été utilisés. Ces dispositifs permettent d'enregistrer les ondes réfléchies et diffractées, avec des angles de diffraction faibles, dans les structures du sous-sol. Cependant, ces ondes utilisées seules posent certains problèmes de robustesse dans les méthodes d'imagerie. Les dispositifs à grands offsets ou grands angles sont actuellement des alternatives de plus en plus utilisées, permettant d'enregistrer les ondes réfléchies à incidences critiques et les ondes réfractées ou ondes tournantes. L'utilisation combinée des différents types d'ondes, enregistrés par des dispositifs grands offsets, par la méthode d'inversion des formes d'ondes fut la motivation principale de la création du projet SEISCOPE.

Les méthodes d'imagerie sismique « conventionnelles »

Les méthodes d'imagerie se classent en deux grandes familles : les méthodes quantitatives et les méthodes non-quantitatives. Les méthodes quantitatives permettent d'imager un ou des paramètres physiques représentatifs des matériaux du sous-sol de manière absolue. Les vitesses de propagation des ondes, la densité, l'atténuation intrinsèque, l'anisotropie peuvent ainsi être caractérisés par les méthodes quantitatives. L'utilisation des images quantitatives obtenues, ou de leurs combinaisons, permet d'interpréter en termes géologiques les milieux imagés. Au contraire, les méthodes non-quantitatives permettent d'obtenir des images de réflectivité, traduisant la capacité des points du milieu à diffracter les ondes. Ces diffractions sont occasionnées

par les variations spatiales des paramètres physiques. Les images permettent donc de caractériser les interfaces et les réflecteurs du milieu de manière relative dans l'espace. Chaque méthode ayant des limitations et des champs d'applications spécifiques, les méthodes quantitatives et non-quantitatives sont souvent utilisées conjointement dans les processus d'imagerie.

Historiquement, les données sismiques ont été « partiellement » utilisées pour imager le sous-sol. En effet, l'utilisation de la forme complète du signal sismique, la forme d'onde, est un processus non-linéaire par rapport aux paramètres physiques, rendant le problème difficile à résoudre². C'est pourquoi, afin de simplifier le problème d'imagerie, des approximations de la propagation des ondes furent faites, couplées à des utilisations incomplètes de l'information contenue dans les données sismiques, menant à des méthodes diverses. Par exemple, les temps d'arrivées des ondes sont souvent utilisés pour imager la Terre. Les tomographies des temps de premières arrivées, des ondes réfléchies, réfractées, de surface... sont massivement utilisées par les communautés académiques et industrielles pour obtenir des images quantitatives des vitesses des ondes à diverses échelles. La dispersion des ondes de surface, extraite de données d'acquisition terrestres, peut être utilisée pour caractériser quantitativement les vitesses de propagation d'ondes cisailantes dans la proche surface. En partant d'un milieu de référence lisse, les méthodes de migration permettent d'imager les réflecteurs du milieu. Des analyses d'amplitude des données élastiques en fonction de la distance horizontale à la source (Analyse AVO (*Amplitude-Variation-with-Offset*)) permettent de déterminer les coefficients élastiques des réflecteurs pour la caractérisation des réservoirs.

Pour des cibles à l'échelle de l'exploration pétrolière qui sont présentées dans ce travail, les procédures classiquement utilisées pour l'imagerie se décomposent en deux étapes appliquées de manière itérative, généralement effectuées sous l'approximation acoustique pour l'imagerie de la vitesse des ondes de compression :

- Construction du modèle de vitesse : la première étape consiste à trouver un milieu de propagation lisse (quantitatif) expliquant la cinématique des trains d'ondes. Cette construction peut être effectuée par analyse de vitesse, tomographies des temps.
- Migration : en utilisant le modèle de vitesse quantitatif, les données hautes fréquences sont migrées afin de localiser les réflecteurs dans le milieu (images généralement non-quantitatives). L'analyse des images migrées permet d'interpréter les structures géologiques et de déterminer si le processus doit être réitéré pour améliorer l'image.

Notons que cette procédure s'est révélée très efficace depuis plusieurs décennies. Cependant, elle souffre des approximations faite à chacune des étapes et de l'utilisation de géométries d'acquisition courts offsets, rendant le processus peu robuste pour imager des structures complexes et mettant en jeux des phénomènes de propagation complexes.

La méthode d'inversion des formes d'ondes

Principe

Contrairement aux méthodes « classiques », la méthode d'inversion des formes d'ondes tente de construire des modèles quantitatifs du sous-sol en exploitant le champ d'ondes complet sous la forme d'un problème d'optimisation. En utilisant la forme du signal, les informations de

²mais pas impossible, sinon ce travail n'existerait pas...

temps, de phase, d'amplitude sont prises en compte simultanément dans une méthode unique. Le principe est donc d'utiliser les différentes arrivées (ondes directes, réfléchies, réfractées, de surface, multiples) afin d'imager les différents paramètres physiques auxquels les ondes sismiques sont sensibles (vitesses de propagation, anisotropie, densité, atténuation). L'utilisation du signal complet signifie que théoriquement, cette méthode permet de produire des images quantitatives de haute résolution en exploitant à la fois les basses fréquences et les hautes fréquences contenues dans les données.

La méthode d'inversion des formes d'ondes définit un problème inverse. Les données sont les signaux portant les signatures des ondes sismiques tandis que le modèle, la solution du problème, est l'ensemble des paramètres physiques à reconstruire (Tarantola, 1987). Le problème direct consiste à calculer des données « synthétiques » dans un modèle numérique *a priori*. Le problème inverse consiste à modifier le modèle pour minimiser l'écart entre les données synthétiques calculées et les données observées, afin d'imager les structures géologiques présentes dans le sous-sol. La relation entre paramètres physiques et données est non-linéaire, le problème inverse doit donc être résolu comme un problème d'optimisation non-linéaire. Les sismogrammes sont généralement enregistrés à la surface, donnant une couverture incomplète (à part pour l'imagerie à l'échelle globale...) et contribuant à renforcer le caractère mal posé du problème inverse. Celui-ci est donc d'autant plus difficile à résoudre à cause des problèmes de fortes non-linéarités et de non-unicité.

Notons qu'un point majeur de la méthode d'inversion des formes d'onde réside dans le fait qu'elle utilise tout le champ d'ondes sans hypothèse particulière. Ainsi, la méthode peut être appliquée à toutes les échelles pour imager tous types de cibles. La méthode a principalement été développée jusqu'à aujourd'hui pour des applications de sismique active à l'échelle de l'exploration pétrolière (voir par exemple Ravaut *et al.*, 2004; Sirgue *et al.*, 2009; Brenders *et al.*, 2009). Cependant, la méthode peut sans restriction s'appliquer à la sismique pour des échelles de proche surface (Gelis *et al.*, 2007; Effelsen, 2009; Romdhane *et al.*, 2009), crustale ou lithosphérique (Operto *et al.*, 2006; Brenders & Pratt, 2007b), voir de Terre globale. Elle peut également s'appliquer à d'autres physiques comme les équations de Maxwell en électromagnétique (Lopez, 2009). Enfin, on pourrait voir des applications sur des objets divers en contrôle non-destructif ou en imagerie médicale.

Le problème direct

Le problème direct du processus d'inversion consiste à calculer les champs d'ondes dans un modèle numérique du sous-sol. En faisant l'hypothèse que le milieu peut être arbitrairement hétérogène dans toutes les directions de l'espace, le problème direct se ramène à calculer les solutions du système d'équations aux dérivées partielles qui régit la propagation des ondes. Notons que dans cet optique, le problème direct n'est pas linéarisé comme cela peut être le cas dans d'autres méthodes d'imagerie. La résolution du problème direct permet de calculer le champ d'ondes complet : ondes réfléchies, réfractées, multiplement réfléchies et diffractées.

En faisant l'hypothèse de la propagation acoustique, les ondes cisailantes sont négligées et seules les ondes de compression sont modélisées. Cette approximation est souvent faite car certaines configurations d'acquisition y sont favorables ; un prétraitement des données permet de supprimer une partie de l'empreinte des ondes cisailantes ; et c'est une solution plus économique en terme de calcul. En prenant en compte la propagation élastique des ondes, on se rapproche de la réalité en modélisant les différents types d'ondes sismiques : ondes converties, ondes de

surfaces... en payant le prix de calcul sous-jacent. Que ce soit pour l'approximation acoustique ou élastique, seuls les phénomènes de propagation linéaires³ sont généralement considérés.

Le problème inverse

Le problème inverse de l'inversion des formes d'ondes est, comme nous l'avons déjà souligné, fortement non-linéaire et mal posé. Dans le but de produire des images haute résolution du milieu, le nombre de paramètres à estimer dans le modèle est conséquent pour des géométries 2D ou 3D. Le coût numérique de résolution du problème direct impose, avec les moyens de calcul actuels, de ne pouvoir envisager de recherche globale ou semi-globale non-linéaire dans l'espace des modèles. En pratique, le problème inverse est donc résolu de manière linéarisée et itérative, sous la forme d'un problème d'optimisation locale autour du modèle de départ. L'approche inverse linéarisée permet de pouvoir envisager la résolution du problème avec les moyens de calcul actuels, mais souffre de plusieurs limitations liées à cette approximation : un modèle de départ suffisamment précis doit être disponible afin d'assurer la convergence du processus. De plus, l'approche inverse linéarisée ne permet d'exploiter que les simples diffractions présentes dans les données. Enfin, même si l'approche inverse linéarisée réduit considérablement le nombre de simulations, le problème inverse n'en reste pas moins fortement non-linéaire et nécessite différents niveaux de régularisation pour assurer la convergence vers le minimum global.

L'inversion élastique

Depuis une vingtaine d'année, les applications d'inversion des formes d'ondes ont principalement été basées sur l'approximation acoustique pour la reconstruction du paramètre de vitesse d'onde P (V_P). Ce choix est motivé par le coût moindre du problème direct acoustique, comparé à la modélisation élastique ; par le fait que la reconstruction d'un seul paramètre physique, en supposant que les autres paramètres sont connus ou empiriquement liés au paramètre imagé, donne un problème inverse moins non-linéaire et donc plus simple à résoudre ; par le fait que le modèle de départ pour le paramètre de vitesses d'onde S est souvent problématique à construire ; et par le fait que le paramètre V_P est généralement dominant dans les données. Ce choix est donc une approximation valide dans de nombreux cas.

Cependant, l'approximation acoustique est limitée en certains points : il est difficile d'envisager la reconstruction de paramètres secondaires, tel que la densité ou l'atténuation qui influent principalement sur les amplitudes des ondes, si le problème direct ne modélise pas fidèlement les amplitudes. Or, en négligeant le paramètre de vitesse S (V_S), paramètre non négligeable dans la réponse AVO des milieux à offsets intermédiaires et longs, la modélisation des amplitudes est biaisée. Dans certaines applications d'acquisitions terrestres, la signature du paramètre V_S est très marquée dans les données et les prétraitements pour supprimer sa signature ne sont pas toujours faciles à mettre oeuvre. En particulier pour l'imagerie de la proche surface, les ondes de surface, très énergétiques, sont proches temporellement des autres arrivées et sont donc délicates à supprimer. Enfin, la connaissance du paramètre V_S , combiné à V_P , permet de déterminer certains attributs des roches (porosité, saturation), améliorant ainsi la caractérisation de certaines zones géologiques (proche surface, réservoirs). L'utilisation de la

³Comme nous l'introduirons dans le chapitre 1 les phénomènes de propagation linéaires sont liés aux différentes linéarisations opérées dans les relations déplacements/déformations/contraintes.

méthode d'inversion des formes d'ondes basée sur une physique de propagation élastique est donc nécessaire dans de nombreuses applications.

Temps ou fréquence ?

Historiquement, la méthode d'inversion des formes d'ondes fut développée et utilisée dans les années 80 dans le domaine espace/temps (Tarantola, 1984, 1987; Gauthier *et al.*, 1986; Mora, 1989; Crase *et al.*, 1990, 1992). Dans ce domaine, les données sont représentées par des sismogrammes temporels. Les premières applications de la méthode souffraient d'une part d'un coût de calcul prohibitif, limitant le nombre d'itérations possibles, et d'autre part d'un choix inadapté des dispositifs d'acquisition à court offset. Ces configurations limitent la possibilité d'imager les grandes longueurs d'onde du milieu et rendent par conséquent le processus fortement non-linéaire.

Au début des années 90, Pratt & Worthington (1990); Pratt (1990) proposèrent d'effectuer l'inversion dans le domaine espace/fréquence pour des géométries d'acquisition à grands offsets⁴. Dans ce domaine, les données observées sont calculées par transformées de Fourier des données temporelles, et représentées par des nombres complexes. La configuration proposée est très astucieuse : les grands offsets permettent de reconstruire les grandes longueurs d'ondes du milieu, élément indispensable à la convergence du système itératif et qui manquait aux applications en domaine temporel de l'époque. Par ailleurs, la grande plage des angles d'ouverture, liée au choix de l'acquisition, couplée à l'échantillonnage des fréquences procurent une redondance dans le domaine des nombres d'ondes imagés. Ainsi, l'inversion peut se limiter à seulement quelques fréquences discrètes bien choisies sans risquer de sous-échantillonner les nombres d'ondes imagés (Sirgue & Pratt, 2004). De plus, le fait de commencer l'inversion depuis les basses fréquences, puis de monter progressivement vers les hautes fréquences, procure un caractère hiérarchique qui minimise la non-linéarité de l'inversion : les grandes structures du milieu sont d'abord imagées avant d'injecter des détails de plus en plus fins. Notons qu'une approche voisine fut développée en domaine temporel par Bunks *et al.* (1995). Ensuite, il apparaît que pour des géométries 2D, pour quelques fréquences et de nombreuses sources, il est moins coûteux, en temps de calcul, de résoudre le système d'équations d'onde en domaine espace/fréquence qu'en domaine espace/temps (Nihei & Li, 2007). Enfin, la prise en compte de l'atténuation intrinsèque du milieu se fait trivialement en domaine fréquentiel (Toksöz & Johnston, 1981), ce qui n'est pas le cas en domaine temporel. En utilisant la stratégie de Pratt, de nombreuses applications furent menées à bien durant les vingt dernières années. Par la suite, nous effectuerons l'inversion dans le domaine espace/fréquence.

Plan et enjeux

Afin de mener l'inversion des formes d'onde élastiques en domaine fréquentiel, différents aspects de développement d'outils et d'applications doivent être menés. Ce manuscrit est donc séparé en deux parties.

Dans une **première partie**, je développe les aspects liés aux méthodes et outils nécessaires à l'inversion des formes d'ondes élastiques. Le **chapitre 1** est consacré à la résolution du pro-

⁴Historiquement, Pratt & Worthington (1990); Pratt (1990) travaillaient sur des applications 2D avec des géométries d'acquisition en entre-puits.

blème direct du processus d'imagerie. Le problème direct requiert de calculer les solutions en domaine fréquentiel du champ d'ondes complet en milieu hétérogène. Afin de s'affranchir de certaines limites et contraintes des méthodes de différences finies et d'éléments finis continus, je développe une méthode de type éléments finis Galerkin discontinus permettant d'utiliser des maillages triangulaires. En utilisant différents ordres d'interpolation, le mélange des ordres et l'utilisation de maillages non-structurés, la méthode développée se révèle flexible et permet de s'adapter à des cas d'étude d'échelles et de complexités très variées. Dans le **chapitre 2**, j'introduis les éléments théoriques du problème inverse linéarisé. Les méthodes de résolution des problèmes d'optimisation locaux sont abordées ainsi que les méthodes d'obtention de modèles de départ. Le **chapitre 3** présente les aspects algorithmiques de l'outil d'inversion des formes d'ondes. La méthode d'inversion est par nature consommatrice de ressources informatiques pour la résolution des problèmes directs et inverses. Je discuterai des algorithmes, de l'implémentation et des performances de l'outil développé dans le contexte de calcul parallèle à hautes performances, permettant d'envisager des cibles d'imagerie d'échelles variées.

Dans la **deuxième partie** de ce manuscrit, je présente des applications d'inversion et les problématiques sous-jacentes. Le **chapitre 4** est consacré à des applications synthétiques réalistes en domaines terrestre et marin. Ces deux environnements permettent d'illustrer les problématiques spécifiques liées à l'inversion de données multicomposantes pour des reconstructions élastiques multiparamètres. Nous montrons que le schéma classique utilisé pour l'inversion monoparamètre de V_P sous l'approximation acoustique se révèle inefficace, et que des niveaux de hiérarchisation et de régularisation supplémentaires sont nécessaires pour converger vers le minimum global du problème. Le **chapitre 5** se focalise sur la sensibilité de l'inversion en domaine fréquentiel. En effet, seulement quelques fréquences discrètes sont généralement inversées en domaine fréquentiel. Une étude de sensibilité est menée avec l'inversion de données bruitées. Cette étude permet de mettre en évidence les limites de l'approche classique basée sur une fonctionnelle aux moindres carré et montre que des fonctionnelles plus robustes, comme la norme L_1 , sont des alternatives de choix dans le contexte d'inversion de données décimées. Enfin une application acoustique sur les données réelles du champ de Valhall est présentée pour l'application des normes L_2 et L_1 dans le contexte de données pétrolières de bonne qualité.

Première partie

Théorie et algorithmes

Chapitre 1

La modélisation de la propagation des ondes élastiques en 2D

Sommaire

1.1	Équations de l'élastodynamique linéaire	26
1.1.1	Équilibre des forces : conservation de la quantité de mouvement	27
1.1.2	Linéarité déplacement/déformation	27
1.1.3	Linéarité déformation/contrainte : loi de comportement	28
1.1.4	Milieux anisotropes <i>versus</i> isotropes, élastiques <i>versus</i> acoustiques... quelle approximation choisir ?	28
1.1.5	Les équations de propagation 3D	30
1.1.6	Quelles ondes ?	31
1.1.7	L'approximation 2D de la propagation des ondes	32
1.2	Méthodes de résolution des systèmes d'équations aux dérivées par- tielles	34
1.2.1	Introduction	34
1.2.2	Différences finies	34
1.2.3	Méthodes pseudo-spectrales	36
1.2.4	Éléments finis	36
1.2.5	Éléments spectraux	37
1.2.6	Galerkin discontinus	37
1.3	Conditions aux limites	38
1.3.1	Conditions initiales	38
1.3.2	Terme source	38
1.3.3	Conditions aux limites spatiales	39
1.4	Modélisation en domaine temporel ou fréquentiel ?	40
1.4.1	Problématique	40
1.4.2	Approche temporelle par schéma explicite d'intégration	40
1.4.3	Approche fréquentielle : solveur direct	41
1.4.4	Approche fréquentielle : solveur itératif	42
1.4.5	Approche fréquentielle : solveur hybride	43
1.4.6	Choix de l'approche fréquentielle + solveur direct	43
1.5	La discrétisation par éléments finis Galerkin discontinus en do- maine fréquentiel	44

1.5.1	Système 2D P-SV en fréquence et conditions absorbantes PML	44
1.5.2	Transformation du système	45
1.5.3	Discrétisation spatiale	46
1.6	Application GD P_0 ou volume fini	53
1.6.1	Présentation de l'article	53
1.6.2	Résumé	54
1.6.3	Introduction	55
1.6.4	Formulation volume fini en domaine fréquentiel	57
1.6.5	Propriétés numériques	60
1.6.6	Résultats numériques	65
1.6.7	Coût numérique de la méthode	74
1.6.8	Discussion et conclusion	75
1.6.9	Annexe A : équivalence VF P_0 et $O(\Delta x^2)$ DF sur grilles carrées régulières	77
1.6.10	Annexe B : développement volume fini	79
1.6.11	Conclusion	80
1.7	Extension aux ordres supérieurs P_1 et P_2	80
1.7.1	Introduction	80
1.7.2	Formulation	81
1.7.3	Approche <i>parsimonieuse</i>	82
1.7.4	Implémentation du terme de source	84
1.7.5	Application numérique	85
1.8	Conclusion	88

Dans ce chapitre, je vais tout d'abord introduire les équations de propagation des ondes dans les milieux élastiques. Ces équations doivent être résolues dans des milieux arbitrairement hétérogènes afin de constituer le moteur du problème direct de l'algorithme d'inversion des formes d'ondes. Afin d'envisager cette inversion dans le domaine espace/fréquence, je discuterai du choix du domaine de résolution du problème direct. Pour des applications à deux dimensions, nous verrons que le choix du domaine fréquentiel s'impose naturellement. Lorsque le problème est formulé en fréquence, un système linéaire creux de grande taille doit être résolu pour chaque fréquence considérée, système qui peut être résolu par diverses méthodes d'algèbre linéaire que j'introduirai. Pour résoudre les équations de l'élastodynamique, des méthodes numériques de discrétisation d'équations aux dérivées partielles seront introduites, en particulier la méthode d'éléments finis **Galerkin Discontinus** que nous avons choisie, avec des interpolations de bas ordres.

1.1 Équations de l'élastodynamique linéaire

La propagation des ondes dans les milieux élastiques, que nous allons considérer pour les applications de sismique, est gouvernée par les équations de l'élastodynamique linéaire. Ces équations découlent des principes de la mécanique des milieux continus linéaire, qui assurent d'une part la conservation de la quantité de mouvement et d'autre part, qui relie linéairement déplacements, déformations et contraintes dans la matière. Je rappellerai ici uniquement les principes généraux qui mènent aux équations de l'élastodynamique, le lecteur pouvant se référer à des ouvrages généraux de propagation d'onde pour de plus amples détails (Aki & Richards, 2002; Chapman, 2004).

1.1.1 Équilibre des forces : conservation de la quantité de mouvement

La première étape à considérer est l'équilibre des forces pour tout élément infinitésimal de matière en mouvement. Pour cela, considérons un référentiel fixe dans l'espace et le temps. Un point matériel de l'espace à trois dimensions est positionné en $\mathbf{x}_0 = (x_0, y_0, z_0)^t$ dans ce référentiel, à un temps de référence t_0 . Sa position à un temps ultérieur t étant \mathbf{x} , le vecteur déplacement du point est alors défini par $\mathbf{u} = \mathbf{x} - \mathbf{x}_0$. Son vecteur vitesse peut alors être écrit $\frac{\partial \mathbf{u}}{\partial t}$ et son vecteur accélération $\frac{\partial^2 \mathbf{u}}{\partial t^2}$.

Considérons un élément de volume V dans une représentation lagrangienne de l'espace en coordonnées cartésiennes. Ce volume est fermé et il est soumis à la loi de conservation de la masse $M(t)$:

$$\frac{dM(t)}{dt} = 0. \quad (1.1)$$

Nous pouvons alors appliquer le principe fondamental de la dynamique (deuxième loi de Newton) sur l'élément, avec invariance temporelle de la masse :

$$\iiint_V \rho \frac{\partial^2 \mathbf{u}}{\partial t^2} dV = \iiint_V \mathbf{F} dV + \iint_S \mathbf{T} dS. \quad (1.2)$$

Cette relation de conservation de la quantité de mouvement relie l'accélération de chaque élément de V et de masse volumique ρ , aux efforts extérieurs produits par des densités de forces de volume \mathbf{F} , et des forces de surface \mathbf{T} appliquées sur la surface extérieure S du volume V . Chaque composante i des forces de surface \mathbf{T} peut être reformulées comme :

$$T_i = \sigma_{ij} n_j, \quad (1.3)$$

où l'expression sous entend une sommation implicite sur les indices (convention de sommation d'Einstein), tout comme dans les expressions suivantes. σ_{ij} est le tenseur des contraintes et n_j est la j -ième composante du vecteur normal à S .

Le théorème de la divergence permet de transformer l'intégrale de surface des forces surfaciques en intégrale de volume :

$$\iint_S \sigma_{ij} n_j dS = \iiint_V \partial_j \sigma_{ij} dV, \quad (1.4)$$

où ∂_j représente la dérivée partielle selon la direction j . Nous pouvons ainsi déduire l'expression de la conservation de la quantité de mouvement sous forme locale

$$\rho \frac{\partial^2 u_i}{\partial t^2} = F_i + \partial_j \sigma_{ij}. \quad (1.5)$$

1.1.2 Linéarité déplacement/déformation

Une deuxième étape consiste à relier les déplacements des particules élémentaires de matière en déformation adimensionnelle. La mécanique des milieux continus permet de relier le tenseur des déformations aux déplacements par étude de l'allongement entre deux points d'un corps déformé. Nous pouvons ainsi exprimer le tenseur (symétrique) des déformations ϵ_{ij}

$$\epsilon_{ij} = \frac{1}{2} \left(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i} + \frac{\partial u_k}{\partial x_i} \frac{\partial u_k}{\partial x_j} \right). \quad (1.6)$$

Cette définition générale du tenseur ϵ_{ij} est non-linéaire par rapport aux déplacements. Cependant, pour la problématique de propagation d'ondes qui engendre des petits déplacements des particules autour de leurs positions d'équilibre (au moins pour les particules localisées loin de la source sismique...), l'hypothèse des petites déformations peut être appliquée. En supposant des petits déplacements et des petites déformations, les termes de second d'ordre peuvent être négligés, laissant place au tenseur des déformations linéarisées que nous utiliserons :

$$\epsilon_{ij} = \frac{1}{2} \left(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i} \right). \quad (1.7)$$

1.1.3 Linéarité déformation/contrainte : loi de comportement

La dernière étape permettant d'établir les équations de l'élastodynamique est le lien entre déformations et contraintes. Ce lien est généralement appelé loi de comportement et modélise la rhéologie du milieu considéré (modèle élastique, élasto-plastique, visco-élasto-plastique...).

Un développement limité de Taylor du tenseur de contraintes permet de lier contraintes et déformations par l'intermédiaire de tenseurs d'ordres élevés représentant la loi de comportement :

$$\sigma_{ij} = \sigma_{ij}^0 + C_{ijkl}\epsilon_{kl} + D_{ijklmn}\epsilon_{kl}\epsilon_{mn} + \mathcal{O}(\epsilon^3). \quad (1.8)$$

σ_{ij}^0 représente la précontrainte en l'absence de déformation lorsqu'il n'y a pas de chargement appliqué au matériau. C_{ijkl} et D_{ijklmn} sont des tenseurs d'ordre 4 et 6 représentant respectivement les comportements linéaire et quadratique des contraintes par rapport aux déformations.

Pour la problématique de propagation des ondes, nous considérons que le modèle de comportement élastique linéaire est suffisamment réaliste en champ lointain de la source. Nous verrons par la suite dans la partie 1.5.1 comment l'atténuation intrinsèque des milieux peut être introduite dans la propagation. Nous pouvons également considérer que la précontrainte est nulle¹.

Nous pouvons donc finalement considérer la loi de Hooke généralisée liant linéairement les tenseurs d'ordre 2 de contraintes et de déformations par le tenseur C d'ordre 4 :

$$\sigma_{ij} = C_{ijkl}\epsilon_{kl}. \quad (1.9)$$

1.1.4 Milieux anisotropes *versus* isotropes, élastiques *versus* acoustiques... quelle approximation choisir ?

Les symétries inhérentes aux tenseurs de déformations et de contraintes, ainsi que des considérations thermodynamiques sur le tenseur C , permettent de réécrire la loi de Hooke sous une forme vectorielle et matricielle faisant apparaître 21 coefficients élastiques indépendants

¹Le terme de précontrainte pourrait être conservé sans engendrer de problème dans la formulation. Il deviendrait nul dans les équations de l'élastodynamique qui s'intéressent uniquement aux dérivées temporelles des contraintes.

dans le cas le plus général de matériaux anisotropes linéaires tricliniques :

$$\begin{pmatrix} \sigma_1 \\ \sigma_2 \\ \sigma_3 \\ \sigma_4 \\ \sigma_5 \\ \sigma_6 \end{pmatrix} = \begin{pmatrix} C_{11} & C_{12} & C_{13} & C_{14} & C_{15} & C_{16} \\ C_{12} & C_{22} & C_{23} & C_{24} & C_{25} & C_{26} \\ C_{13} & C_{23} & C_{33} & C_{34} & C_{35} & C_{36} \\ C_{14} & C_{24} & C_{34} & C_{44} & C_{45} & C_{46} \\ C_{15} & C_{25} & C_{35} & C_{45} & C_{55} & C_{56} \\ C_{16} & C_{26} & C_{36} & C_{46} & C_{56} & C_{66} \end{pmatrix} \begin{pmatrix} \epsilon_1 \\ \epsilon_2 \\ \epsilon_3 \\ 2\epsilon_4 \\ 2\epsilon_5 \\ 2\epsilon_6 \end{pmatrix}, \quad (1.10)$$

les indices numériques étant relatifs à des couples de directions spatiales :

$$(1, 2, 3, 4, 5, 6)^t = (xx, yy, zz, yz, xz, xy)^t. \quad (1.11)$$

La Nature crée généralement des matériaux moins « bizarres » que les hommes². Les termes élastiques de couplage entre contraintes normales et déformations cisailantes, et inversement due à la symétrie du système, sont généralement nuls. Ainsi, 9 coefficients non nuls caractérisent les matériaux orthorhombiques, cas le plus générique de l'anisotropie rencontrée dans la nature (Tsvankin, 2005)(cas d'argiles fracturées par exemple). Cependant, des symétries isotropes sont généralement observées dans les roches permettant de caractériser les matériaux VTI (*Vertical Transverse Isotropic*), HTI (*Horizontal Transverse Isotropic*), TTI (*Tilted Transverse Isotropic*) généralement paramétrisés dans la communauté des sciences de la Terre par les 5 paramètres de Thomsen (1986) en 3D (4 en 2D).

Le cas le plus simple de matériaux élastiques est le cas isotrope, caractérisé par deux paramètres sous la forme des coefficients de Lamé : λ (pas de signification physique directe) et μ (module de cisaillement) :

$$\begin{pmatrix} \sigma_{xx} \\ \sigma_{yy} \\ \sigma_{zz} \\ \sigma_{yz} \\ \sigma_{xz} \\ \sigma_{xy} \end{pmatrix} = \begin{pmatrix} \lambda + 2\mu & \lambda & \lambda & 0 & 0 & 0 \\ \lambda & \lambda + 2\mu & \lambda & 0 & 0 & 0 \\ \lambda & \lambda & \lambda + 2\mu & 0 & 0 & 0 \\ 0 & 0 & 0 & \mu & 0 & 0 \\ 0 & 0 & 0 & 0 & \mu & 0 \\ 0 & 0 & 0 & 0 & 0 & \mu \end{pmatrix} \begin{pmatrix} \epsilon_{xx} \\ \epsilon_{yy} \\ \epsilon_{zz} \\ 2\epsilon_{yz} \\ 2\epsilon_{xz} \\ 2\epsilon_{xy} \end{pmatrix}, \quad (1.12)$$

ou sous forme du module d'Young E et du coefficient de Poisson ν , plus souvent utilisés en mécanique et définis par

$$E = \frac{\mu(3\lambda + 2\mu)}{\lambda + \mu} \quad \nu = \frac{\lambda}{2(\lambda + \mu)}. \quad (1.13)$$

Enfin, l'approximation acoustique de la propagation peut être faite (uniquement) à partir du système isotrope en posant $\mu = 0$ (ou $\nu = 0.5$), équivalent au fait que les fluides (et gaz) ne transmettent pas les contraintes cisailantes.

²J'ai travaillé pendant mon DEA sur des matériaux composites à base d'empilements de couches anisotropes : les matériaux naturels sont en général plus « sympathiques » que ce l'homme peut fabriquer...

1.1.5 Les équations de propagation 3D

En nous intéressant dorénavant aux milieux élastiques isotropes, nous pouvons formuler les équations de propagation des ondes dans ces milieux en déplacements/contraintes en reprenant l'équation (1.5) et en combinant les équations (1.7) et (1.12)

$$\begin{aligned}
 \rho \frac{\partial^2 u_x}{\partial t^2} &= \frac{\partial \sigma_{xx}}{\partial x} + \frac{\partial \sigma_{xy}}{\partial y} + \frac{\partial \sigma_{xz}}{\partial z} + F_x \\
 \rho \frac{\partial^2 u_y}{\partial t^2} &= \frac{\partial \sigma_{xy}}{\partial x} + \frac{\partial \sigma_{yy}}{\partial y} + \frac{\partial \sigma_{yz}}{\partial z} + F_y \\
 \rho \frac{\partial^2 u_z}{\partial t^2} &= \frac{\partial \sigma_{xz}}{\partial x} + \frac{\partial \sigma_{yz}}{\partial y} + \frac{\partial \sigma_{zz}}{\partial z} + F_z \\
 \sigma_{xx} &= (\lambda + 2\mu) \frac{\partial u_x}{\partial x} + \lambda \frac{\partial u_y}{\partial y} + \lambda \frac{\partial u_z}{\partial z} + \sigma_{xx0} \\
 \sigma_{yy} &= \lambda \frac{\partial u_x}{\partial x} + (\lambda + 2\mu) \frac{\partial u_y}{\partial y} + \lambda \frac{\partial u_z}{\partial z} + \sigma_{yy0} \\
 \sigma_{zz} &= \lambda \frac{\partial u_x}{\partial x} + \lambda \frac{\partial u_y}{\partial y} + (\lambda + 2\mu) \frac{\partial u_z}{\partial z} + \sigma_{zz0} \\
 \sigma_{xy} &= \mu \left(\frac{\partial u_x}{\partial y} + \frac{\partial u_y}{\partial x} \right) + \sigma_{xy0} \\
 \sigma_{yz} &= \mu \left(\frac{\partial u_y}{\partial z} + \frac{\partial u_z}{\partial y} \right) + \sigma_{yz0} \\
 \sigma_{xz} &= \mu \left(\frac{\partial u_x}{\partial z} + \frac{\partial u_z}{\partial x} \right) + \sigma_{xz0},
 \end{aligned} \tag{1.14}$$

dans lesquelles les termes d'excitation sur les contraintes σ_{ij_0} ont été ajoutés. Notons que les champs de déplacements u_i et de contraintes σ_{ij} sont dépendantes de l'espace et du temps tandis que les paramètres physiques (λ , μ et ρ) dépendent de l'espace et sont considérés invariants dans le temps (à l'échelle du temps de la propagation des ondes).

Nous pouvons alors reformuler le système (1.14) sous une forme plus classique de système hyperbolique d'ordre 1 mettant en jeu de façon homogène des dérivées premières sur le temps, en remplaçant le champ de déplacements par le champ de vitesses de déplacements ($V_i = \partial_t u_i$) (Virieux, 1986a)

$$\begin{aligned}
 \frac{\partial V_x}{\partial t} &= \frac{1}{\rho} \left(\frac{\partial \sigma_{xx}}{\partial x} + \frac{\partial \sigma_{xy}}{\partial y} + \frac{\partial \sigma_{xz}}{\partial z} \right) + f_x \\
 \frac{\partial V_y}{\partial t} &= \frac{1}{\rho} \left(\frac{\partial \sigma_{xy}}{\partial x} + \frac{\partial \sigma_{yy}}{\partial y} + \frac{\partial \sigma_{yz}}{\partial z} \right) + f_y \\
 \frac{\partial V_z}{\partial t} &= \frac{1}{\rho} \left(\frac{\partial \sigma_{xz}}{\partial x} + \frac{\partial \sigma_{yz}}{\partial y} + \frac{\partial \sigma_{zz}}{\partial z} \right) + f_z \\
 \frac{\partial \sigma_{xx}}{\partial t} &= (\lambda + 2\mu) \frac{\partial V_x}{\partial x} + \lambda \frac{\partial V_y}{\partial y} + \lambda \frac{\partial V_z}{\partial z} + \frac{\partial \sigma_{xx0}}{\partial t} \\
 \frac{\partial \sigma_{yy}}{\partial t} &= \lambda \frac{\partial V_x}{\partial x} + (\lambda + 2\mu) \frac{\partial V_y}{\partial y} + \lambda \frac{\partial V_z}{\partial z} + \frac{\partial \sigma_{yy0}}{\partial t} \\
 \frac{\partial \sigma_{zz}}{\partial t} &= \lambda \frac{\partial V_x}{\partial x} + \lambda \frac{\partial V_y}{\partial y} + (\lambda + 2\mu) \frac{\partial V_z}{\partial z} + \frac{\partial \sigma_{zz0}}{\partial t}
 \end{aligned}$$

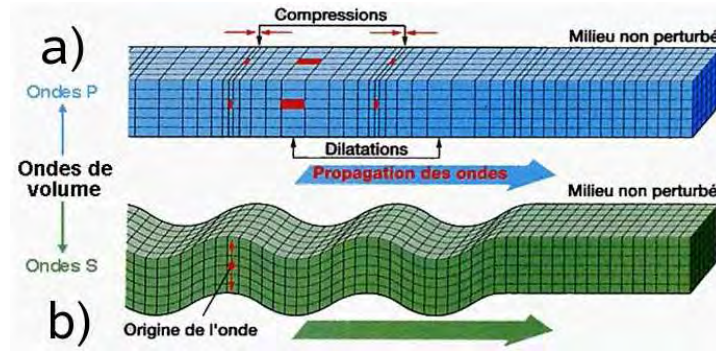


FIG. 1.1 – Schémas de propagation des ondes élastiques de volume (a) de compression (ondes P) et (b) de cisaillement (ondes S).

$$\begin{aligned}
 \frac{\partial \sigma_{xy}}{\partial t} &= \mu \left(\frac{\partial V_x}{\partial y} + \frac{\partial V_y}{\partial x} \right) + \frac{\partial \sigma_{xy_0}}{\partial t} \\
 \frac{\partial \sigma_{yz}}{\partial t} &= \mu \left(\frac{\partial V_y}{\partial z} + \frac{\partial V_z}{\partial y} \right) + \frac{\partial \sigma_{yz_0}}{\partial t} \\
 \frac{\partial \sigma_{xz}}{\partial t} &= \mu \left(\frac{\partial V_x}{\partial z} + \frac{\partial V_z}{\partial x} \right) + \frac{\partial \sigma_{xz_0}}{\partial t}.
 \end{aligned} \tag{1.15}$$

Notons qu'après la transformation du système, les forces extérieures f_i deviennent les dérivées temporelles des densités volumiques de forces F_i du système (1.14), divisées par la masse volumique ρ .

1.1.6 Quelles ondes ?

Le système (1.15) peut être écrit sous une forme compacte vectorielle

$$\frac{\partial \mathbf{U}}{\partial t} = A \frac{\partial \mathbf{U}}{\partial x} + B \frac{\partial \mathbf{U}}{\partial y} + C \frac{\partial \mathbf{U}}{\partial z} + \frac{\partial \mathbf{U}_0}{\partial t}, \tag{1.16}$$

où le vecteur $\mathbf{U} = (V_x, V_y, V_z, \sigma_{xx}, \sigma_{yy}, \sigma_{zz}, \sigma_{xy}, \sigma_{yz}, \sigma_{xz})^t$.

L'analyse des valeurs propres des matrices A , B et C montrent que deux types d'ondes se propagent dans le système d'équations élastodynamique :

1. des ondes de vitesse $V_P = \sqrt{\frac{\lambda+2\mu}{\rho}}$ appelées ondes de compression ou ondes P (Figure 1.1(a)). Elles vibrent dans la direction de propagation des ondes. La dénomination d'ondes P est relative à *pressure* ou *première*, car elles sont les plus rapides. Ce sont les seules ondes présentes dans les milieux fluides.
2. des ondes de vitesse $V_S = \sqrt{\frac{\mu}{\rho}}$ appelées ondes de cisaillement ou ondes S (Figure 1.1(b)). Elles vibrent dans le plan orthogonal à la direction de propagation. La physique des roches sur le coefficient de Poisson assure que $V_S < \frac{V_P}{\sqrt{2}}$. La dénomination d'ondes S est relative à *shear* ou *seconde*, car elles se propagent plus lentement que les ondes P.

1.1.7 L'approximation 2D de la propagation des ondes

Le système d'équation élastodynamique 3D (1.15) peut être réécrit sous des hypothèses de propagation à une ou deux dimensions.

Le cas 1D fait l'hypothèse de deux directions infinies d'invariance pour les propriétés physiques et les excitations (rendant par conséquent les champs d'ondes également invariants). L'équation résultante est la formulation classique de propagation des ondes sur une corde.

Le cas 2D fait l'hypothèse d'une seule direction infinie d'invariance. Cette direction est généralement la direction y en sciences de la Terre, z en mathématiques. Bien sûr, l'approximation 2D de la propagation, faite tout au long de ce travail, ne représente pas exactement la réalité car elle ne prend pas en compte la propagation 3D des ondes : les effets 3D du milieu de propagation et la vraie atténuation géométrique des champs d'ondes ne peuvent être pris fidèlement en compte. Cependant, cette approximation est généralement faite par simplicité et économie sans résulter sur des propagations irréalistes :

- on peut dans certains cas trouver une certaine invariance des propriétés physiques du milieu dans une des directions de l'espace, ce qui vérifie partiellement une des hypothèses de l'approximation.
- le biais occasionné sur l'atténuation géométrique des champs d'ondes, dû à la deuxième hypothèse d'une source linéique infiniment longue en 2D, peut être partiellement corrigé dans les champs d'ondes : correction de l'atténuation géométrique 3D ($1/r$) en 2D ($1/\sqrt{r}$) (Bleistein, 1986; Williamson & Pratt, 1995).
- la résolution du système (1.15), par une méthode numérique de résolution d'équations aux dérivées partielles (EDP), est beaucoup plus économique en terme de calcul lorsque le nombre de dimensions spatiales est minimum.

Ainsi, l'approximation 2D de la propagation des ondes classiquement faite en sciences de la Terre est obtenue en annulant toutes les dérivées spatiales selon y dans les équations 3D (1.15), et laisse place à deux systèmes indépendants.

Un premier système est relatif à la propagation des ondes P-SV, ondes de compression et de cisaillement polarisées uniquement dans le plan (x, z) :

$$\begin{aligned}
 \frac{\partial V_x}{\partial t} &= \frac{1}{\rho} \left(\frac{\partial \sigma_{xx}}{\partial x} + \frac{\partial \sigma_{xz}}{\partial z} \right) + f_x \\
 \frac{\partial V_z}{\partial t} &= \frac{1}{\rho} \left(\frac{\partial \sigma_{xz}}{\partial x} + \frac{\partial \sigma_{zz}}{\partial z} \right) + f_z \\
 \frac{\partial \sigma_{xx}}{\partial t} &= (\lambda + 2\mu) \frac{\partial V_x}{\partial x} + \lambda \frac{\partial V_z}{\partial z} + \frac{\partial \sigma_{xx0}}{\partial t} \\
 \frac{\partial \sigma_{zz}}{\partial t} &= \lambda \frac{\partial V_x}{\partial x} + (\lambda + 2\mu) \frac{\partial V_z}{\partial z} + \frac{\partial \sigma_{zz0}}{\partial t} \\
 \frac{\partial \sigma_{xz}}{\partial t} &= \mu \left(\frac{\partial V_x}{\partial z} + \frac{\partial V_z}{\partial x} \right) + \frac{\partial \sigma_{xz0}}{\partial t}.
 \end{aligned} \tag{1.17}$$

Le deuxième système est relatif à la propagation des ondes SH, ondes de cisaillement polarisées dans les plans (x, y) et (y, z) , invariantes dans la direction y et indépendantes des ondes P-SV :

$$\frac{\partial V_y}{\partial t} = \frac{1}{\rho} \left(\frac{\partial \sigma_{xy}}{\partial x} + \frac{\partial \sigma_{yz}}{\partial z} \right) + f_y$$

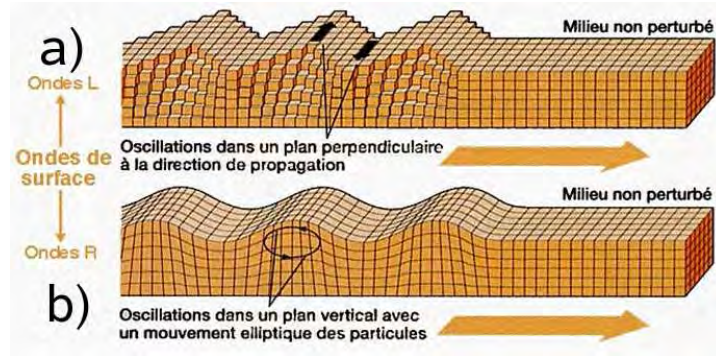


FIG. 1.2 – Schémas de propagation des ondes élastiques de surface (a) de Love (ondes L) et (b) de Rayleigh (ondes R).

$$\begin{aligned}\frac{\partial \sigma_{xy}}{\partial t} &= \mu \left(\frac{\partial V_y}{\partial x} \right) + \frac{\partial \sigma_{xy_0}}{\partial t} \\ \frac{\partial \sigma_{yz}}{\partial t} &= \mu \left(\frac{\partial V_y}{\partial z} \right) + \frac{\partial \sigma_{yz_0}}{\partial t}.\end{aligned}\quad (1.18)$$

Lorsque on considère la propagation des ondes dans un espace bordé d'une surface libre (approximation de la surface de la Terre au dessus de laquelle réside uniquement de l'air), des ondes de surface se propagent près de cette interface.

- Les ondes P et SV interfèrent constructivement et génèrent des ondes de Rayleigh (1887) donnant aux particules un mouvement elliptique rétrograde proche de la surface et prograde plus en profondeur (Figure 1.2(b)). Leur vitesse de propagation V_R dépend principalement de V_S mais également de V_P . On peut montrer que leur vitesse satisfait toujours la relation $V_R < V_S$ (Rayleigh, 1887). Les ondes de Rayleigh se propagent et explorent en profondeur proportionnellement à la longueur d'onde (et donc de la fréquence). Elles sont donc naturellement dispersives dès que le milieu devient hétérogène car la vitesse de chaque fréquence dépend de la profondeur du milieu sondé. Les ondes de Rayleigh se propagent selon différents modes discrets qui satisfont tous les équations de propagation et les conditions aux limites. Chaque mode propage un certain niveau d'énergie. Le mode fondamental (le plus lent) contient généralement la plus forte énergie, mais la répartition de cette énergie dépend de la fréquence et de l'hétérogénéité du milieu. La dispersion des différents modes des ondes de Rayleigh est couramment utilisée pour l'imagerie de la proche surface (Nazarian & Stokoe, 1984). Enfin, les ondes de Rayleigh portent généralement une grande énergie qui s'atténue moins que les ondes de volume (atténuation géométrique en $1/\sqrt{r}$ en 3D, pas d'atténuation en 2D), cause principale des dégâts occasionnés lors des séismes.
- Les ondes SH interfèrent entre elles et donnent naissance sous certaines conditions des propriétés de V_S dans le milieux, aux ondes de Love (1911) (Figure 1.2(a)). Ces ondes sont également dispersives. Je ne détaillerai pas plus la propagation des ondes de Love, propre aux ondes SH que nous n'utiliserons pas.

Par la suite, nous nous focaliserons uniquement sur les ondes P et SV régies par les équations (1.17).

1.2 Méthodes de résolution des systèmes d'équations aux dérivées partielles

1.2.1 Introduction

Le système 2D élastodynamique (1.17) des ondes P-SV (tout comme le système 3D) peut être résolu de façon **analytique** ou **pseudo-analytique** dans un certain nombre de configurations simples de milieux : milieu homogène, demi-espace homogène avec surface libre (Garvin, 1956), milieu à couches (Bouchon *et al.*, 1989). Cependant, dans la perspective d'imagerie par inversion des formes d'ondes, le système doit pouvoir être résolu dans des milieux arbitrairement hétérogènes. L'approche asymptotique du **tracé de rai** (Červený *et al.*, 1977) permet d'envisager des milieux hétérogènes sous l'approximation hautes fréquences de la propagation. Sous cette hypothèse, la longueur d'onde doit être beaucoup plus petite que la taille des hétérogénéités prises en compte dans le milieu. L'onde est exprimée par le produit d'un terme de phase fortement oscillant et d'un terme d'amplitude faiblement oscillant indépendant de la fréquence. Cette approximation laisse place à l'équation eikonal et l'équation de transport qui permettent respectivement de calculer le temps de trajet et l'amplitude d'une onde. Cependant, il apparaît qu'en pratique cette méthode n'est valide qu'en milieux « lisses » et peu complexes, la rendant non envisageable pour une imagerie haute résolution.

Pour notre application qui requiert la solution du champ d'onde complet en milieu fortement hétérogène, nous devons considérer des méthodes numériques de calcul des dérivées partielles dans un espace discret. Notons que les dérivées partielles spatiales et temporelles sont généralement discrétisées avec des méthodes dissociées. Je développe ci-dessous quelques méthodes utilisables et utilisées pour la propagation des ondes.

1.2.2 Différences finies

La méthode des différences finies (DF) se montre très naturelle car elle estime directement chaque dérivée par un développement de Taylor. Le lecteur peut se rapporter à l'ouvrage de Moczo *et al.* (2004) qui donne une très bonne introduction à cette méthode pour la résolution des équations d'ondes. Par exemple, si nous considérons deux points de l'espace i et $i + 1$ sur l'axe x , séparés par une distance h et portant respectivement les champs discrets u_i et u_{i+1} . Nous pouvons calculer la dérivée spatiale d'ordre 1 selon x au point $i + 1/2$:

$$\frac{\partial u_{i+1/2}}{\partial x} = \frac{u_{i+1} - u_i}{h} + \mathcal{O}(h^2). \quad (1.19)$$

Cet exemple illustre l'estimation de la dérivée par une *différence centrée d'ordre 2*.

Les DF sont généralement utilisées pour la discrétisation des dérivées temporelles des équations d'ondes, quelque soit le choix fait pour la discrétisation spatiale. Selon le schéma utilisé, on peut soit se ramener à une intégration temporelles explicite, où les champs d'ondes à un temps t sont calculés à partir des champs au(x) temps discret(s) précédent(s) ; soit à un schéma implicite où la résolution d'un système linéaire est nécessaire à chaque pas de temps. Les schémas explicites sont en principe préférés pour des applications de grandes tailles à cause du coût numérique important de résolution des schémas implicites.

Concernant la discrétisation spatiale, la méthode des différences finies est très utilisée en sciences de la Terre pour résoudre les équations de propagation d'ondes (Virieux, 1986a; Levander, 1988; Moczo *et al.*, 2004; Bohlen & Saenger, 2006). Cette utilisation massive est la conséquence de plusieurs phénomènes couplés : la grande simplicité de cette méthode, du point de vue théorique comme de l'implémentation, rend son utilisation attractive et très efficace d'un point de vue informatique. De plus, les objets d'études dans lesquels des ondes doivent être modélisées en sciences de la Terre peuvent souvent être représentés par des formes simples : rectangles, parallélépipèdes, cercles, boules. Ces formes sont facilement discrétisées avec des grilles régulières, discrétisations les plus courantes en DF.

Cependant, cette simplicité et l'utilisation de grilles régulières (carrées ou cubiques) donnent lieu à certaines limitations. En effet, le pas de la grille numérique de discrétisation est défini par rapport à la longueur d'onde minimale modélisée (dépendante de la fréquence maximale et de la vitesse de propagation minimale) pour limiter les phénomènes de dispersion numérique. Ainsi, pour une grille uniforme, la discrétisation spatiale est contrainte par la vitesse minimale dans le milieu même si cette dernière ne représente qu'une infime partie du modèle. Certains auteurs (Moczo *et al.*, 1997) ont développé des DF sur grilles déformées, raffinées proche de la surface libre par exemple, où les vitesses de propagation sont plus lentes. Cependant, les déformations doivent rester lisses pour ne pas introduire d'effets parasites dans les solutions, pouvant limiter les applications possibles.

Une deuxième limitation des DF est la représentation de la surface libre en cas de topographies complexes. En effet, la représentation cartésienne du milieu par DF approxime la topographie par une succession de marches d'escalier qui peut engendrer des diffractions parasites dans le champ d'ondes si le pas de la grille est trop grossier. Cette contrainte amène un suréchantillonnage de l'ensemble du milieu afin de représenter fidèlement la propagation des ondes de Rayleigh par exemple. Ainsi, Saenger & Bohlen (2004); Bohlen & Saenger (2006) ont montré qu'une discrétisation de plus de 60 points par longueur d'onde S était nécessaire avec un schéma numérique d'ordre 2 (stencil de Saenger *et al.* (2000)) pour modéliser précisément la propagation des ondes de Rayleigh dans un modèle avec une topographie gaussienne. Une discrétisation de 15 points par longueur d'onde S est suffisante pour la propagation des ondes de volume avec ce même schéma, signifiant qu'un suréchantillonnage d'un facteur 16 est nécessaire en 2D et 64 en 3D. Notons que des approches plus élaborées d'interfaces fictives immergées peuvent modéliser très précisément les topographies complexes en DF (Lombard *et al.*, 2008). Cependant, ces approches ne sont actuellement développées que pour des résolutions temporelles par schémas explicites d'intégration et pour des géométries 2D.

Enfin, une dernière limitation des DF, particulièrement pour la résolution des équations élastodynamiques en fréquence, est le choix de l'ordre du schéma. En effet, comme nous le verrons plus en détail dans la partie 1.4, pour une méthode de résolution par factorisation LU, le choix d'un schéma numérique à support compact en espace est primordial pour réduire le coût d'utilisation mémoire. Or, les schémas DF d'ordre élevés, les plus précis, sont les plus étendus spatialement, ce qui rend leur utilisation inenvisageable. L'utilisation de schémas à support compact, comparables aux schémas d'ordre 2 est donc primordiale (Hustedt *et al.*, 2004). Des schémas mixtes couplant plusieurs stencils d'ordre 2 et donnant une précision similaire à des schémas d'ordre 4 ont ainsi été efficacement conçus pour résoudre l'équation d'onde acoustique, avec une discrétisation de 4 points par longueur d'onde minimale (Jo *et al.*, 1996; Hustedt *et al.*, 2004; Operto *et al.*, 2007). Cependant, ce type de schéma est difficilement envisageable pour la propagation des ondes élastiques, le couplage optimal entre les différents schémas numériques

étant fortement dépendant du coefficient de Poisson dans le milieu (Stekl & Pratt, 1998), rendant leur utilisation délicate dans des milieux hétérogènes.

1.2.3 Méthodes pseudo-spectrales

Les méthodes pseudo-spectrales (Fornberg, 1996) résolvent les EDP en s'appuyant sur des bases d'interpolations exponentielles complexes par des allers et retours dans les domaines de Fourier spatiaux (domaines des nombres d'ondes). Ainsi, chaque dérivée partielle dans le domaine spatial se simplifie en un simple produit dans le domaine de Fourier :

$$\frac{\partial \mathbf{u}}{\partial x} = \mathcal{TF}^{-1}(-ik_x \mathcal{TF}(\mathbf{u}(x))) = \mathcal{TF}^{-1}(-ik_x \mathbf{U}(k_x)), \quad (1.20)$$

où \mathcal{TF} est la transformée de Fourier (TF), i le nombre imaginaire pur et k_x le vecteur nombre d'onde dans la direction x . Ces méthodes se révèlent très précises pour la résolution des équations d'ondes grâce à leurs bases d'interpolations très naturelles pour les solutions oscillantes. Des discrétisations très lâches sont autorisées, pouvant être à la limite du théorème d'échantillonnage (2 points par longueur d'onde). Pour résoudre les systèmes 2D ou 3D dans le domaine temporel avec un schéma d'intégration en temps explicite, des TF directes et inverses sont nécessaires à chaque pas de temps. Les applications 3D posent encore aujourd'hui des problèmes de performances informatiques pour les opérations de transposition de matrices 3D, nécessaires à l'application des TF sur la troisième dimension. Cependant, ces méthodes se montrent plus efficaces que des méthodes DF d'ordre 16 ou 20 tout en étant plus précises (Chu & Stoffa, 2008), grâce aux développements récents de bibliothèques informatiques de TF rapides multidimensionnelles très performantes (Frigo & Johnson, 2005). Ces méthodes restent néanmoins basées sur des grilles cartésiennes régulières posant des problèmes similaires au DF lorsque des topographies complexes et/ou des milieux à forts contrastes de vitesses sont envisagés. Enfin, pour résoudre la version fréquentielle du système élastodynamique, la grande étendue de l'opérateur de TF (sur tout le domaine) rend son application inenvisageable, des opérateurs compacts étant essentiels.

1.2.4 Éléments finis

La méthode des éléments finis (EF) (Zienkiewicz *et al.*, 2005) permet de résoudre les EDP sous forme variationnelle. La solution est approximée sur des bases de fonction, généralement polynomiales, définie sur des noeuds localisés dans des éléments discrétisant le milieu. Les éléments « communiquent » par les noeuds qu'ils partagent à leurs interfaces. La solution calculée est donc continue dans l'espace et nous pouvons qualifier cette méthode d'éléments finis continus par opposition aux méthodes discontinues introduites par la suite. Les EF permettent très naturellement de prendre en compte des conditions aux limites comme la surface libre, et sont très bien adaptés pour l'application à des maillages non-structurés, permettant de représenter fidèlement une topographie complexe ou des structures géologiques. Les équations élastodynamiques ont été résolues par EF par différents auteurs (Marfurt, 1984; Bielak *et al.*, 2003; Yoshimura *et al.*, 2003; Koketsu *et al.*, 2004), en utilisant des bases de polynômes d'ordre faibles. Cependant, comme pour les DF, la discrétisation nécessaire pour des solutions précises reste relativement fine et coûteuse.

1.2.5 Eléments spectraux

Pour palier le suréchantillonnage spatial des EF, les éléments finis d'ordres élevés, appelé (à tort) éléments spectraux (ES) (Faccioli *et al.*, 1997; Komatitsch & Vilotte, 1998; Chaljub *et al.*, 2003; Vilotte *et al.*, 2005), ont été développés sur des bases de polynômes de Legendre et des règles de quadrature numérique de Gauss-Lobatto-Legendre, donnant un taux de convergence « spectral » lors de la montée en ordre d'interpolation. De plus, cette formulation permet d'obtenir une matrice de masse naturellement diagonale, ne nécessitant pas d'inversion de matrice, contrairement aux EF « classiques ». Cette variante des éléments finis continus est aujourd'hui massivement utilisée en sismologie grâce à la discrétisation relativement grossière de l'espace et la grande précision des solutions qu'elle offre. Cependant, un frein majeur subsiste : le maillage. En effet, les ES « classiques » s'appuient sur des discrétisations de l'espace par quadrangles en 2D et parallélépipèdes en 3D, discrétisations difficiles à mettre en oeuvre dans de nombreuses applications, comparé à des maillages triangulaires (2D) et tétraédriques (3D). Actuellement, ces limitations tendent à être levées. Des équipes de recherche tentent activement d'étendre les ES sur des maillages triangulaires et tétraédriques (Dubiner, 1991; Pasquetti & Rapetti, 2004; Mercerat *et al.*, 2006; Pasquetti & Rapetti, 2006), sur lesquels les quadratures classiques de Gauss-Lobatto-Legendre ne sont pas définies, tandis que d'autres tentent de développer de performants mailleurs parallélépipédiques (Casarotti *et al.*, 2007). Notons enfin que la méthode d'ES tout comme les EF, ne permet pas d'envisager simplement le contact entre milieux solides et liquides pour lequel le paramètre μ est nul. Des conditions de continuité doivent être explicitement formulées entre les équations discrétisées de l'élastodynamique et de l'acoustique, rendant les applications de sismique marine plus contraignantes.

1.2.6 Galerkin discontinus

Dans les années 70, une autre variante des EF « classiques » a été développée sous une forme discontinue, appelée Galerkin discontinus (GD), pour les équations de transport de neutrons (Reed & Hill, 1973). Cockburn *et al.* (2000) présentent un ouvrage de synthèse sur ces méthodes. Les GD ont été utilisés pour résoudre les équations d'ondes électromagnétiques (Remaki, 2000) et ont récemment été introduits pour la propagation des ondes élastiques (Käser & Dumbser, 2006; Dumbser & Käser, 2006; Käser *et al.*, 2007; de la Puente *et al.*, 2007; Dumbser *et al.*, 2007) et la rupture sismique (BenJemaa *et al.*, 2009). Cette méthode peut être développée sur des maillages triangulaires et tétraédriques, leur donnant un certain avantage sur les ES. Les champs étant calculés de manière discontinue, les cellules du maillage « communiquent » via des flux numériques sans partager de noeuds communs. Cette caractéristique donne une grande flexibilité dans le choix des interpolations entre éléments voisins (mélange des ordres possibles entre éléments voisins) et la possibilité d'utiliser des maillages non-conformes (Fahs, 2007; Fahs & Lanteri, 2008). Notons le cas particulier d'utilisation de fonction de base polynomiale d'ordre 0 (interpolation P_0 , fonction constante par cellule) donnant la même formulation que la méthode de **volumes finis** de même ordre (LeVeque, 2002). Remarquons tout de même qu'à maillage et ordre d'interpolation équivalents, les méthodes GD considèrent plus de degrés de liberté que les EF ou ES, chaque élément possédant ses propres degrés de liberté parfois pour une même position physique des noeuds.

Pour les applications présentées dans ce travail, nous utiliserons la méthode **Galerkin discontinus** avec des ordres d'interpolation faibles comme je le justifierai par la suite. Ainsi,

nous pourrons tirer bénéfice de l'utilisation de maillages triangulaires non-structurés ; de l'ordre d'interpolation P_0 très efficaces pour certaines applications car très proche des méthodes de différences finies ; et de pouvoir envisager des contacts fluide/solide de géométries quelconques, sans condition d'interface particulière.

1.3 Conditions aux limites

Pour la résolution numérique d'EDP telle que l'élastodynamique, nous ne pouvons considérer les équations « brutes » valables seulement en domaine infini. Afin que le système d'équations soit consistant et donne une unique solution, des conditions aux limites spatiales et temporelles doivent être imposées.

1.3.1 Conditions initiales

Des conditions initiales doivent tout d'abord être imposées pour le temps « initial » de la propagation, spécifiant l'état des champs d'ondes. Le système élastodynamique étant hyperbolique d'ordre 1 et faisant intervenir deux champs vectoriels (de vitesses et de contraintes), les conditions initiales nécessaires peuvent être :

- des conditions de Dirichlet (champ imposé) au temps $t = 0$ pour les deux champs vectoriels des équations.
- un couple conditions de Dirichlet et conditions de Neumann (dérivée du champ imposée, ici dérivée temporelle) sur un des champs vectoriels, au temps $t = 0$.

Le choix de l'une ou l'autre des conditions permet de définir de façon unique l'état du système au temps $t = 0$.

Dans la pratique, la propagation des ondes à partir d'un point (ou zone) source, comme c'est le cas en sismique, s'effectue généralement dans un milieu initialement au repos et non perturbé. Une condition de type Dirichlet sur l'ensemble des champs tel que $\mathbf{U}(t = 0) = \mathbf{0}$ est donc généralement utilisées. Nous pourrions également imposer un état de champs non nul, pour la modélisation d'ondes planes par exemple, créant implicitement l'excitation du milieu.

1.3.2 Terme source

Lorsque le système est « initialisé » par des conditions initiales nulles, il doit être excité par une source d'énergie afin de générer des ondes. En sismique et sismologie, les sources sont généralement localisées dans une zone réduite de l'espace de propagation, créant un point source ou zone source. Dans un monde numérique discret, une zone source peut être vue comme un assemblage de points ou éléments sources. Le terme source doit être défini pour un ensemble de localisations de l'espace (au moins un point), imposant une force ponctuelle et/ou une contrainte externe, pour l'ensemble du temps de propagation et tel que l'énergie apportée soit non nulle.

1.3.3 Conditions aux limites spatiales

La modélisation de la propagation des ondes peut se faire soit dans un modèle représentant l'intégralité de l'objet d'étude (Terre globale par exemple), soit dans une zone limitée (subsurface alors que les ondes se propagent dans Terre entière). Des conditions aux limites spatiales doivent donc être imposées aux bords du modèle numérique de simulation, afin de reproduire les conditions réelles de propagation. On peut définir trois grands types de conditions aux limites :

- les conditions de Dirichlet spatiales imposent explicitement la valeur de certains champs aux bords du modèle. C'est le cas de la surface de la Terre, équivalente à une surface libre où les tractions normales (direction n) sont nulles : $\sigma_{ij}n = 0$
- les conditions de Neumann spatiales imposent explicitement la valeur de la dérivée de certains champs aux bords du modèle. C'est le cas de la surface libre où les vitesses de déplacement orthogonal des particules à la surface (direction n) doivent être continues de part et d'autre de l'interface : $\partial V_i n / \partial n = 0$
- les conditions absorbantes simulent un milieu infini. Ces conditions aux limites laissent passer les ondes sortantes du modèle (allant vers l'infini) mais sans laisser entrer d'énergie depuis l'extérieur (pas d'ondes entrantes). Ces conditions peuvent avoir un support très local (un seul noeud numérique), en s'appuyant sur des formulations analytiques de décomposition des champs d'ondes locaux en ondes planes entrantes et sortantes (Clayton & Engquist, 1980). Elles se révèlent très efficaces en terme de coût numérique, n'ajoutant pas de degrés de liberté supplémentaires à la simulation. Cependant, leur efficacité est optimale pour seulement quelques valeurs d'angles d'incidence des ondes. En dehors, des angles optimums, des réflexions ont lieu, occasionnant des champs d'ondes parasites non physiques entrant dans le domaine de simulation. D'autres méthodes de conditions absorbantes reposent sur des modifications locales des propriétés d'atténuation des milieux en ajoutant des couches supplémentaires autour du milieu de simulation. Ainsi, les « éponges » de Cerjan *et al.* (1985) amortissent les ondes sortantes de manière comparable à une atténuation intrinsèque progressive lorsque les ondes progressent dans ces couches. Cependant, là encore l'incidence normale est optimale et des réflexions parasites sont présentes pour d'autres angles d'incidence. Enfin, une dernière alternative est l'utilisation de *Perfectly Matched Layers* (PML), proposée par Berenger (1994). Ces couches reprennent l'idée des « éponges », mais avec une atténuation anisotrope, atténuant uniquement la composante orthogonale des ondes entrantes dans le PML. Théoriquement, dans la formulation continue des équations, cette méthode introduit un coefficient de réflexion nul pour toutes les directions d'incidence des ondes et toutes les fréquences. Dans un monde discret, des réflexions peuvent être observées pour des angles rasants, mais l'efficacité de cette méthode reste bien supérieure aux précédentes alternatives. La contrepartie de cette méthode est un surcoût numérique de calcul, dû aux degrés de liberté supplémentaires introduits par les couches absorbantes autour du modèle (comme pour les éponges) qui peuvent représenter une part non négligeable d'inconnues dans le système considéré. Par ailleurs, la formulation temporelle des équations nécessite des variables supplémentaires associées à chaque noeud situé dans le PML dû à la formulation décomposée des champs (Berenger, 1994) ou dû à l'utilisation de variables mémoires de CPML (Convolutionnal PML) (Komatitsch & Martin, 2007). Notons cependant que, dans une formulation fréquentielle, les PML n'ajoutent pas de surcoût numérique autre que celui engendré par l'incorporation des couches absorbantes.

1.4 Modélisation en domaine temporel ou fréquentiel ?

1.4.1 Problématique

Dans le cadre du problème d'imagerie, nous souhaitons formaliser le problème inverse dans le domaine espace/fréquence. Pour cela, les champs d'ondes doivent être représentés dans le domaine fréquentiel pour pouvoir être comparés aux données de référence également représentées dans ce domaine. Deux alternatives sont alors envisageables :

1. résoudre le système élastodynamique (1.17) dans le domaine espace/temps et transformer les champs d'ondes dans le domaine espace/fréquence (Nihei & Li, 2007; Sirgue *et al.*, 2008). La première résolution est généralement faite avec un schéma d'intégration en temps explicite, puis les champs sont convertis en fréquences par transformée de Fourier rapide (FFT) ou transformée de Fourier discrète (DFT).
2. transformer le système (1.17) en domaine espace/fréquence par transformée de Fourier, puis le résoudre dans ce domaine. Cette méthode se ramène à résoudre un système linéaire $\mathbf{Ax} = \mathbf{b}$ de grande taille par fréquence (ordre de matrice de plusieurs milliers à millions). La résolution de systèmes linéaires s'envisage généralement de manière directe ou itérative. Notons que la résolution en fréquence permet de s'affranchir de la condition de stabilité Courant-Friedrichs-Lewy (CFL), obligatoire pour les schémas explicites en temps.

Le choix d'une ou l'autre des alternatives doit se faire en fonction des complexités des temps de calcul et d'utilisation mémoire de chacune des approches. Considérons un problème à 2 dimensions dont les dérivées spatiales sont calculées par un schéma aux différences finies d'ordre 2 sur une grille carrée de $n \times n$ points (un schéma GD tel qu'utilisé par la suite étant local, il peut être assimilé à un schéma DF). Considérons un problème de sismique à N_s tirs et le calcul d'une seule fréquence. Notons que le raisonnement qui suit est valable pour un cas 2D. Pour une approche 3D, les complexités seraient différentes et les conclusions également. Les différentes alternatives en 3D sont discutées dans Virieux *et al.* (2009).

1.4.2 Approche temporelle par schéma explicite d'intégration

En suivant le raisonnement de Nihei & Li (2007), nous pouvons montrer que l'utilisation d'un schéma en temps explicite nécessite $\mathcal{O}(n^2)$ opérations par pas de temps dans la simulation. Pratiquement, le nombre de pas de temps N_t est généralement de l'ordre de grandeur de n . Ainsi, pour les N_s sources considérées, cette approche requiert $\mathcal{O}(n^2 N_s N_t) \approx \mathcal{O}(n^3 N_s)$ opérations et un stockage en mémoire vive de $\mathcal{O}(n^2)$ (en considérant que les différentes sources sont calculées de façon séquentielle). La phase de transformation des champs en fréquence n'est pas prise en compte dans ce calcul, bien qu'elle ne soit pas négligeable.

En effet, pour l'utilisation d'une FFT qui permet d'obtenir les champs pour toutes les fréquences du spectre, le stockage de tous les champs à tous les pas de temps est nécessaire (stockage de $\mathcal{O}(n^2 N_t)$). Ceci est difficilement envisageable sans recourir à du stockage sur disque, beaucoup plus lent.

Si seulement quelques fréquences sont nécessaires, comme c'est le cas pour l'inversion des formes d'ondes en fréquence, l'utilisation de DFT est plus économique. Elle permet de calculer

à la volée, pendant le calcul des champs temporels, les champs en fréquence par intégration numérique sans surcoût notable de calcul ni de stockage.

Dans une perspective de calculs hautes performances, ce type d'algorithme peut être facilement parallélisé en associant une source par processeur pour accélérer le calcul pour des acquisitions multisources. Dans le cas de gros modèles, un second niveau de parallélisme peut être introduit en découpant le modèle en sous-domaines et en associant chaque sous-domaine à un processeur. Ces deux types de parallélisations offrent de très bonnes performances car l'algorithme met en jeu peu de communications entre processeurs.

1.4.3 Approche fréquentielle : solveur direct

Les méthodes de résolution directe de systèmes linéaires s'appuient sur des manipulations algébriques des termes de la matrice \mathbf{A} , afin de la transformer sous une forme plus propice à la résolution du système, permettant d'obtenir directement la « vraie » solution du système (par opposition aux méthodes itératives).

Ainsi, nous pouvons penser à la résolution par méthode de Gauss qui transforme dans un premier temps la matrice initiale en une matrice triangulaire (avec modifications sur le second membre \mathbf{b}), puis dans un second temps qui calcule la solution \mathbf{x} par substitutions successives. Cette méthode classique est cependant peu robuste face à certains pivots très petits voire nuls qui posent problème numériquement. De plus, la matrice transformée est triangulaire pleine, nécessitant un stockage important pour de grandes matrices.

D'autres méthodes, comme la factorisation LU ou de Cholesky (LL^t), transforment la matrice \mathbf{A} en deux matrices triangulaires supérieures et inférieures (Press *et al.*, 2007). La méthode LU peut s'appliquer sur des matrices quelconques indéfinies tandis que la LL^t s'applique à des matrices symétriques définies positives. L'idée de ces méthodes est de transformer le système de départ $\mathbf{Ax} = \mathbf{b}$ en $\mathbf{LUx} = \mathbf{b}$ et de former ainsi deux systèmes couplés plus simples à résoudre : $\mathbf{Ly} = \mathbf{b}$ puis $\mathbf{Ux} = \mathbf{y}$. En effet, la résolution des deux systèmes est trivial et très économique numériquement, \mathbf{U} et \mathbf{L} étant triangulaires supérieure et inférieure. L'effort de calcul de ces méthodes est la factorisation initiale de la matrice sous forme \mathbf{LU} qui nécessite à la fois calcul et stockage. Notons que la factorisation est indépendante du second membre \mathbf{b} . Ainsi, pour des problèmes à plusieurs sources, la factorisation de la matrice n'est faite qu'une seule fois, les solutions étant ensuite calculées par de simples *forward* et *backward* substitutions dans les matrices triangulaires.

En considérant notre problème $n \times n$ à N_s sources, nous pouvons montrer que le système linéaire associé à une fréquence peut être résolu par factorisation \mathbf{LU} couplé avec un algorithme de réordonnancement des lignes et colonnes (George & Liu, 1989; Amestoy *et al.*, 1996; Karypis & Kumar, 1999) (minimisant le coût de calcul et de stockage) pour un coût de calcul de $\mathcal{O}(n^3)$ et de stockage de $\mathcal{O}(n^2 \log_2(n))$. Ces coûts sont indépendants du nombre de sources, les phases de substitutions de chaque source étant négligeables devant le coût de la factorisation.

Dans une optique de calculs hautes performances, ces algorithmes peuvent être parallélisés en utilisant des approches *multifrontales* (Duff & Reid, 1983; Amestoy *et al.*, 2000) qui permettent de composer les matrices \mathbf{L} et \mathbf{U} sous formes découpées en parallèle par différents processeurs. Cette technique de parallélisation permet d'accélérer les phases de factorisation et de substitution, mais surtout de stocker les matrices \mathbf{L} et \mathbf{U} de façon distribuée, utilisant ainsi la mémoire locale de chaque processeur. La scalabilité de ces algorithmes reste cependant très

limitée car ils mettent en jeu des volumes de communication entre processeurs conséquents, qui deviennent prépondérant lorsque les problèmes deviennent gros et que le nombre de processeurs devient important.

1.4.4 Approche fréquentielle : solveur itératif

Les méthodes de résolution itérative de systèmes linéaires s'appuient sur des approximations successives de la vraie solution et tentent de réduire itérativement l'erreur entre la solution courante et la vraie solution inconnue.

Les méthodes basiques de Jacobi ou Gauss-Siedel (Saad, 2003) permettent de calculer efficacement la solution d'un système linéaire réarrangé en appliquant itérativement l'opérateur matriciel à la solution précédente. Cependant, ces méthodes ne sont applicables que sous certaines conditions de rayon spectral de la matrice, conditions insatisfaites par l'opérateur discrétisé de l'équation d'onde.

Des méthodes plus élaborées telles que le GMRES ou BiCGSTAB (Saad, 2003) permettent de résoudre itérativement les systèmes linéaires indéfinis et mal conditionnés issus de la discrétisation des équations d'ondes, en utilisant des combinaisons de base orthogonales (base de Krylov) construites à partir des solutions des itérations précédentes. Ces méthodes permettent d'assurer la convergence du processus itératif en un nombre maximal d'itération dépendant de la taille du système. Cependant, cette limite est bien souvent inacceptable pratiquement et ces méthodes doivent être couplées avec l'utilisation de préconditionneurs qui accélèrent la convergence du système.

Le système préconditionné à résoudre devient $\mathbf{MA}\mathbf{x} = \mathbf{Mb}$ (préconditionneur à gauche), le but de la matrice \mathbf{M} étant d'être le plus proche possible de \mathbf{A}^{-1} à un coût numérique acceptable, afin procurer un meilleur conditionnement à la matrice \mathbf{MA} que celui de \mathbf{A} . Le développement de préconditionneurs efficaces est une branche de recherche spécifique des mathématiques appliquées, s'appuyant sur des considérations algébriques de la matrice initiale, physiques des phénomènes simulés par l'opérateur \mathbf{A} , ou couplant les deux. Des préconditionneurs efficaces ont pu être développés récemment pour l'équation d'Helmholtz (ondes acoustiques seulement) par Erlangga (2005); Erlangga & Herrmann (2008) assurant une convergence rapide du processus itératif. Ce préconditionneur repose sur l'utilisation de cycles de multigrilles couplée à l'utilisation d'une équation d'onde amortie.

Les complexités mémoire et temps des approches itératives sont très attractives si la convergence du système est rapide : $\mathcal{O}(n^2)N_{it}N_s$ pour le temps (N_{it} étant le nombre d'itération du processus itératif) et $\mathcal{O}(n^2)$ pour le stockage (Riyanti *et al.*, 2006). Pour l'équation d'Helmholtz, le préconditionneur de Erlangga & Herrmann (2008) permet d'obtenir $N_{it} = \mathcal{O}(1)$, rendant l'approche directement concurrente des solveurs directs en temps de calcul. Cependant, de tels préconditionneurs n'ont pas encore été développés pour les équations de l'élastodynamique, tout comme pour les équations de Maxwell (électromagnétiques).

Comme pour l'approche temporelle, la parallélisation de ces méthodes est très efficace avec une très bonne scalabilité et peut se faire par source et par décomposition en domaines physiques.

1.4.5 Approche fréquentielle : solveur hybride

Une méthode alternative hybride permet de combiner l'utilisation de solveur direct et itératif de manière naturellement parallèle. En effet, en découpant la matrice \mathbf{A} par blocs (relatifs à un découpage physique du domaine de calcul en sous-domaines), on peut se ramener à la résolution de plusieurs systèmes linéaires locaux, chacun associé à un sous-domaine physique, et liés à un dernier système linéaire d'interface entre sous-domaines : le système du complément de Schur. Un descriptif détaillé de cette méthode peut être trouvé dans Saad (2003). L'idée générale est d'appliquer un solveur direct, très efficace pour les petits systèmes, à chaque système local et un solveur itératif au problème d'interface. Cette combinaison de solveurs permet de tirer bénéfice des factorisations locales faites une seule fois pour les problèmes multisources donnant de bonnes performances pour les résolutions des domaines locaux, ainsi que du solveur itératif appliqué en parallèle au système du complément de Schur.

De bonnes performances ont pu être obtenues avec cette méthode sur des applications 2D et 3D avec l'équation d'Helmholtz (Sourbier *et al.*, 2008a,b) en utilisant un préconditionneur algébrique pour le système itératif. Cependant, comme pour les approches itératives pures, l'efficacité du préconditionneur est crucial et peu de développements ont été fait jusqu'à présent pour les équations élastodynamiques. J'ai implémenté un prototype de code élastodynamique basé sur une discrétisation GD d'ordre P_0 dans le squelette de l'algorithme de Sourbier *et al.* (2008b) dans le cadre de ce travail de thèse. Quelques tests ont clairement montré les difficultés de convergence du système itératif (convergence lente voire non convergence), comparé à l'équation d'Helmholtz, alors que les deux systèmes utilisaient un préconditionneur algébrique de type « Schwarz additif ». L'investigation plus approfondie de ce type de méthode avec une recherche de préconditionneur efficace pour l'élastodynamique reste donc une perspective intéressante de recherche.

1.4.6 Choix de l'approche fréquentielle + solveur direct

Les différentes alternatives au problème de résolution du système d'équation d'ondes ont montré que les approches reposant sur tout ou partie de solveur itératif n'ont pas atteint, aujourd'hui, un degré de maturité suffisant pour résoudre le système élastodynamique. Un travail d'optimisation des préconditionneurs reste nécessaire pour rivaliser avec les autres approches.

Les approches temporelles et fréquentielles par solveur direct ont chacune un avantage sur l'autre : l'approche temporelle coûte très peu en stockage et donc en mémoire vive, mais nécessite une coût de calcul conséquent. L'approche fréquentielle par solveur direct se montre très efficace en temps de calcul, surtout pour les problèmes multisources, mais nécessite une consommation mémoire plus importante.

Dans la perspective d'inversion de quelques fréquences seulement, et au vu de la complexité mémoire de l'approche fréquentielle par solveur direct, compatible avec les moyens des centres de calcul modernes, cette option sera choisie pour la suite de ce travail. Une version parallèle du solveur direct MUMPS (MUMPS-team, 2007), utilisant une factorisation \mathbf{LU} de la matrice \mathbf{A} (qui n'est pas symétrique à causes des conditions absorbantes PML), sera utilisée dans l'ensemble des tests de ce travail.

1.5 La discrétisation par éléments finis Galerkin discontinus en domaine fréquentiel

1.5.1 Système 2D P-SV en fréquence et conditions absorbantes PML

Afin de résoudre le système 2D P-SV en fréquence, nous devons tout d'abord appliquer une transformée de Fourier au système temporel (1.17) :

$$\begin{aligned}
 -i\omega V_x &= \frac{1}{\rho} \left(\frac{\partial \sigma_{xx}}{\partial x} + \frac{\partial \sigma_{xz}}{\partial z} \right) + f_x \\
 -i\omega V_z &= \frac{1}{\rho} \left(\frac{\partial \sigma_{xz}}{\partial x} + \frac{\partial \sigma_{zz}}{\partial z} \right) + f_z \\
 -i\omega \sigma_{xx} &= (\lambda + 2\mu) \frac{\partial V_x}{\partial x} + \lambda \frac{\partial V_z}{\partial z} - i\omega \sigma_{xx0} \\
 -i\omega \sigma_{zz} &= \lambda \frac{\partial V_x}{\partial x} + (\lambda + 2\mu) \frac{\partial V_z}{\partial z} - i\omega \sigma_{zz0} \\
 -i\omega \sigma_{xz} &= \mu \left(\frac{\partial V_x}{\partial z} + \frac{\partial V_z}{\partial x} \right) - i\omega \sigma_{xz0},
 \end{aligned} \tag{1.21}$$

où le terme de pulsation ω est évalué en *rad/s*. La transformée de Fourier suit la convention $f(\omega) = \int_{-\infty}^{+\infty} f(t)e^{-i\omega t} dt$.

Les conditions absorbantes de type PML peuvent être introduites dans le système fréquentiel en ajoutant les termes s_x et s_z sur les champs de vitesses et contraintes (Collino & Tsogka, 2001), agissant de façon anisotrope sur les dérivées partielles :

$$\begin{aligned}
 -i\omega V_x &= \frac{1}{\rho} \left(s_x \frac{\partial \sigma_{xx}}{\partial x} + s_z \frac{\partial \sigma_{xz}}{\partial z} \right) + f_x \\
 -i\omega V_z &= \frac{1}{\rho} \left(s_x \frac{\partial \sigma_{xz}}{\partial x} + s_z \frac{\partial \sigma_{zz}}{\partial z} \right) + f_z \\
 -i\omega \sigma_{xx} &= (\lambda + 2\mu) s_x \frac{\partial V_x}{\partial x} + \lambda s_z \frac{\partial V_z}{\partial z} - i\omega \sigma_{xx0} \\
 -i\omega \sigma_{zz} &= \lambda s_x \frac{\partial V_x}{\partial x} + (\lambda + 2\mu) s_z \frac{\partial V_z}{\partial z} - i\omega \sigma_{zz0} \\
 -i\omega \sigma_{xz} &= \mu \left(s_x \frac{\partial V_x}{\partial z} + s_z \frac{\partial V_z}{\partial x} \right) - i\omega \sigma_{xz0}.
 \end{aligned} \tag{1.22}$$

Les termes de PML étant définis classiquement par la fonction PML dans la direction r :

$$s_r = \frac{1}{1 + i\gamma_r/\omega}, \tag{1.23}$$

où le coefficient γ_r est nul dans le domaine de calcul, n'affectant pas la propagation, et non nul dans les couches PML, créant un amortissement des champs d'ondes. Ce coefficient suit généralement une fonction progressive définie par une branche de cosinus (Hustedt *et al.*, 2004) ou une fonction polynomiale (Komatitsch & Martin, 2007), qui vaut zéro proche du domaine de calcul et qui augmente à mesure que l'on pénètre dans la couche absorbante.

L'atténuation intrinsèque du milieu de propagation peut être facilement pris en compte dans le système fréquentiel en utilisant des valeurs complexes de paramètres physiques de vitesses

de propagation (Toksöz & Johnston, 1981). Différents modèles peuvent être pris en compte, tel que le modèle de Kolsky-Futterman (Kolsky, 1956; Futterman, 1962) :

$$\bar{c} = c \left[\left(1 + \frac{1}{\pi Q} |\log(\omega/\omega_r)| \right) + i \frac{\text{sign}(\omega)}{2Q} \right]^{-1}, \quad (1.24)$$

où c est la vitesse de l'onde, Q le facteur de qualité représentatif de l'atténuation, ω_r est une fréquence de référence du modèle et \bar{c} la vitesse de propagation complexe. Une version simplifiée négligeant les effets de dispersion des ondes est souvent utilisée en sismologie (Aki & Richards, 2002) :

$$\bar{c} = c \left[1 - \frac{i}{2Q} \right]. \quad (1.25)$$

1.5.2 Transformation du système

Dans le but de formuler le système (1.22) sous une forme pseudo-conservative, utile par la suite pour l'intégration du système, un changement de variables est appliqué :

$$\begin{aligned} -i\omega\rho V_x &= s_x \frac{\partial(T_1 + T_2)}{\partial x} + s_z \frac{\partial T_3}{\partial z} + \rho f_x \\ -i\omega\rho V_z &= s_x \frac{\partial T_3}{\partial x} + s_z \frac{\partial(T_1 - T_2)}{\partial z} + \rho f_z \\ \frac{-i\omega T_1}{\lambda + \mu} &= s_x \frac{\partial V_x}{\partial x} + s_z \frac{\partial V_z}{\partial z} - \frac{i\omega T_1^0}{\lambda + \mu} \\ \frac{-i\omega T_2}{\mu} &= s_x \frac{\partial V_x}{\partial x} - s_z \frac{\partial V_z}{\partial z} - \frac{i\omega T_2^0}{\mu} \\ \frac{-i\omega T_3}{\mu} &= s_x \frac{\partial V_z}{\partial x} + s_z \frac{\partial V_x}{\partial z} - \frac{i\omega T_3^0}{\mu}, \end{aligned} \quad (1.26)$$

où $T_1 = (\sigma_{xx} + \sigma_{zz})/2$, $T_2 = (\sigma_{xx} - \sigma_{zz})/2$ et $T_3 = \sigma_{xz}$. Notons que cette transformation est très naturelle pour considérer des milieux fluides : le terme T_1 équivaut à la pression hydrostatique et les équations 4 et 5 du système (1.26) sont nulles, laissant place à uniquement 3 équations à résoudre.

Ce système peut alors s'écrire sous une forme compacte vectorielle :

$$\begin{aligned} -i\omega\rho\vec{V} &= \overrightarrow{\text{div}(\vec{F}(\vec{T}))} - \frac{\partial s_x M_x}{\partial x} \vec{T} - \frac{\partial s_z M_z}{\partial z} \vec{T} + \rho\vec{f} \\ -i\omega\Lambda\vec{T} &= \overrightarrow{\text{div}(\vec{G}(\vec{V}))} - \frac{\partial s_x N_x}{\partial x} \vec{V} - \frac{\partial s_z N_z}{\partial z} \vec{V} - i\omega\Lambda\vec{T}^0, \end{aligned} \quad (1.27)$$

avec les matrices :

- $M_x = N_x^t = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$
- $M_z = N_z^t = \begin{bmatrix} 0 & 0 & 1 \\ 1 & -1 & 0 \end{bmatrix}$
- $\Lambda = \begin{bmatrix} \frac{1}{\lambda+\mu} & 0 & 0 \\ 0 & \frac{1}{\mu} & 0 \\ 0 & 0 & \frac{1}{\mu} \end{bmatrix}$

et les vecteurs :

- $\vec{V} = (V_x, V_z)^t$
- $\vec{T} = (T_1, T_2, T_3)^t$
- $\vec{f} = (f_x, f_z)^t$
- $\vec{T}^0 = (T_1^0, T_2^0, T_3^0)^t$.
- $\vec{F} = (s_x M_x \vec{T}, s_z M_z \vec{T})$
- $\vec{G} = (s_x N_x \vec{V}, s_z N_z \vec{V})$

1.5.3 Discrétisation spatiale

Nous allons suivre un développement proche de celui de BenJemaa (2007), formulé pour l'élastodynamique 3D temporel, adapté ici pour le cas 2D en domaine fréquentiel. L'idée générale des méthodes d'éléments finis consiste à rechercher une solution approchée du système sous la forme de combinaisons linéaires de fonctions de bases locales à chaque cellule définies aux degrés de liberté de la formulation. En considérant une formulation discontinue, les degrés de liberté sont propres à chaque cellule. Les échanges entre cellules sont définis par l'intermédiaire de flux numériques passant à travers les interfaces.

Considérons que le domaine de calcul Ω est discrétisé par des cellules polygonales. Chaque cellule \mathcal{T}_i est appelée cellule de contrôle. Les notations suivantes sont introduites afin d'alléger la suite des écritures :

1. $\partial\Omega$ est la frontière du domaine Ω
2. $\partial\mathcal{T}_i$ est la frontière de la cellule \mathcal{T}_i
3. $A_i = \int_{\mathcal{T}_i} dV$ est l'aire de la cellule \mathcal{T}_i
4. $V(i)$ est le voisinage de la cellule \mathcal{T}_i , et contient toutes les cellules partageant une arête (ou morceau d'arête) avec \mathcal{T}_i
5. $\mathcal{T}_{ik} = \mathcal{T}_i \cap \mathcal{T}_k$ est l'interface partagée entre les cellules \mathcal{T}_i et \mathcal{T}_k
6. \vec{n}_{ik} est le vecteur unitaire normal à l'interface \mathcal{T}_{ik} , sortant de \mathcal{T}_i vers \mathcal{T}_k

Dans chaque cellule \mathcal{T}_i , les champs de vitesses et de contraintes sont décrits dans une base locale de fonction $\vec{\varphi}_{ij}$ avec $1 \leq j \leq d_i$, où d_i est le nombre de degrés de liberté du champ dans la cellule \mathcal{T}_i . Pour assurer le caractère discontinu, $\vec{\varphi}_{ij}$ est nul en dehors de la cellule \mathcal{T}_i . Les champs sont ainsi représentés par :

$$\begin{aligned}\vec{V}_i &= \sum_{j=1}^{d_i} V_{ij} \vec{\varphi}_{ij} \\ \vec{T}_i &= \sum_{j=1}^{d_i} T_{ij} \vec{\varphi}_{ij},\end{aligned}\tag{1.28}$$

où V_{ij} et T_{ij} représentent respectivement le j-ième degré de liberté de \vec{V}_i et \vec{T}_i .

Dans la formulation développée ici, les propriétés physiques ainsi que les fonctions PML sont considérées constantes dans chaque cellule.

Multiplions le système (1.27) par une fonction test $\vec{\varphi}_{ij}$, choisie dans le même espace que les fonctions solutions, donnant la formulation Galerkin. Puis, intégrons ce système sur une cellule \mathcal{T}_i :

$$\begin{aligned} \int_{\mathcal{T}_i} -i\omega\rho\vec{\varphi}_{ij} \cdot \vec{V} &= \int_{\mathcal{T}_i} \vec{\varphi}_{ij} \cdot \overrightarrow{\text{div}(\vec{F}(\vec{T}))} + \int_{\mathcal{T}_i} \rho\vec{\varphi}_{ij} \cdot \vec{f} \\ \int_{\mathcal{T}_i} -i\omega\vec{\varphi}_{ij} \cdot \Lambda\vec{T} &= \int_{\mathcal{T}_i} \vec{\varphi}_{ij} \cdot \overrightarrow{\text{div}(\vec{G}(\vec{V}))} - \int_{\mathcal{T}_i} i\omega\vec{\varphi}_{ij} \cdot \Lambda\vec{T}^0. \end{aligned} \quad (1.29)$$

Notons que les deuxièmes et troisièmes termes du membre de droite des équations (1.27) sont nuls à cause des propriétés constantes des fonctions PML sur la cellule \mathcal{T}_i .

Une intégration par partie du premier terme du membre de droite du système (1.29) donne alors :

$$\begin{aligned} \int_{\mathcal{T}_i} -i\omega\rho\vec{\varphi}_{ij} \cdot \vec{V} &= - \int_{\mathcal{T}_i} \vec{\nabla}\vec{\varphi}_{ij} : \vec{F}(\vec{T}) + \int_{\partial\mathcal{T}_i} \vec{\varphi}_{ij} \cdot (\vec{F}(\vec{T})\vec{n}) + \int_{\mathcal{T}_i} \rho\vec{\varphi}_{ij} \cdot \vec{f} \\ \int_{\mathcal{T}_i} -i\omega\vec{\varphi}_{ij} \cdot \Lambda\vec{T} &= - \int_{\mathcal{T}_i} \vec{\nabla}\vec{\varphi}_{ij} : \vec{G}(\vec{V}) + \int_{\partial\mathcal{T}_i} \vec{\varphi}_{ij} \cdot (\vec{G}(\vec{V})\vec{n}) - \int_{\mathcal{T}_i} i\omega\vec{\varphi}_{ij} \cdot \Lambda\vec{T}^0, \end{aligned} \quad (1.30)$$

avec pour tous vecteur $\vec{w} = (w_1, \dots, w_d)^t$:

$$\vec{\nabla}\vec{w} = \begin{bmatrix} \partial_x w_1 & \partial_z w_1 \\ \vdots & \vdots \\ \partial_x w_d & \partial_z w_d \end{bmatrix}, \quad (1.31)$$

et pour toutes matrices $\mathcal{A} = (a_{ij})_{\substack{1 \leq i \leq n \\ 1 \leq j \leq m}} \in \mathcal{M}_{n,m}$ et $\mathcal{B} = (b_{ij})_{\substack{1 \leq i \leq n \\ 1 \leq j \leq m}} \in \mathcal{M}_{n,m}$:

$$\mathcal{A} : \mathcal{B} = \sum_{\substack{1 \leq i \leq n \\ 1 \leq j \leq m}} a_{i,j} b_{i,j}. \quad (1.32)$$

Le système (1.30) peut alors être approché par :

$$\begin{aligned} \int_{\mathcal{T}_i} -i\omega\rho_i\vec{\varphi}_{ij} \cdot \vec{V}_i &= - \int_{\mathcal{T}_i} \vec{\nabla}\vec{\varphi}_{ij} : \vec{F}(\vec{T}_i) + \int_{\partial\mathcal{T}_i} \vec{\varphi}_{ij} \cdot (\vec{F}(\vec{T}_{/\partial\mathcal{T}_i})\vec{n}) + \int_{\mathcal{T}_i} \rho_i\vec{\varphi}_{ij} \cdot \vec{f}_i \\ \int_{\mathcal{T}_i} -i\omega\vec{\varphi}_{ij} \cdot \Lambda_i\vec{T}_i &= - \int_{\mathcal{T}_i} \vec{\nabla}\vec{\varphi}_{ij} : \vec{G}(\vec{V}_i) + \int_{\partial\mathcal{T}_i} \vec{\varphi}_{ij} \cdot (\vec{G}(\vec{V}_{/\partial\mathcal{T}_i})\vec{n}) - \int_{\mathcal{T}_i} i\omega\vec{\varphi}_{ij} \cdot \Lambda_i\vec{T}_i^0 \end{aligned} \quad (1.33)$$

où $\vec{F}(\vec{T}_{/\partial\mathcal{T}_i})$ et $\vec{G}(\vec{V}_{/\partial\mathcal{T}_i})$ sont les approximations de $\vec{F}(\vec{T})$ et $\vec{G}(\vec{V})$ sur l'interface $\partial\mathcal{T}_i$.

Le flux centré (Remaki, 2000) est un choix possible d'approximation du flux numérique à l'interface. D'autres approximations comme les flux *upwind* ou *upwind partiellement pénalisés* peuvent également être utilisés. Käser & Dumbser (2006) utilisent par exemple des flux *upwind* pour la résolution du système élastodynamique en temps. Le flux centré permet de garder de bonnes propriétés de conservation de l'énergie discrète dans le système (BenJemaa, 2007; BenJemaa *et al.*, 2009). Cependant, comme nous le verrons dans les applications numériques, il introduit certains problèmes dans les applications. Pour notre développement, nous faisons le choix du flux centré exprimé par :

$$\vec{F}(\vec{T}_{/\mathcal{T}_{ik}}) = \vec{F}\left(\frac{\vec{T}_i + \vec{T}_k}{2}\right)$$

$$\vec{G}(\vec{V}_{/\mathcal{T}_{ik}}) = \vec{G}\left(\frac{\vec{V}_i + \vec{V}_k}{2}\right), \quad (1.34)$$

où \mathcal{T}_k est la cellule voisine de \mathcal{T}_i avec qui elle partage l'arête \mathcal{T}_{ik} . L'expression du flux peut se réécrire sous la forme :

$$\begin{aligned} \vec{F}(\vec{T}_{/\mathcal{T}_{ik}}) &= \frac{1}{2}(\vec{F}(\vec{T}_i) + \vec{F}(\vec{T}_k)) \\ \vec{G}(\vec{V}_{/\mathcal{T}_{ik}}) &= \frac{1}{2}(\vec{G}(\vec{V}_i) + \vec{G}(\vec{V}_k)). \end{aligned} \quad (1.35)$$

Le système (1.33) s'écrit alors :

$$\begin{aligned} \int_{\mathcal{T}_i} -i\omega\rho_i\vec{\varphi}_{ij}^t\vec{V}_i &= -\int_{\mathcal{T}_i} \vec{\nabla}\vec{\varphi}_{ij} : \vec{F}(\vec{T}_i) + \frac{1}{2} \sum_{k \in V(i)} \int_{\mathcal{T}_{ik}} \vec{\varphi}_{ij}^t (\vec{F}(\vec{T}_i) + \vec{F}(\vec{T}_k)) \vec{n}_{ik} + \int_{\mathcal{T}_i} \rho_i \vec{\varphi}_{ij}^t \vec{f}_i \\ \int_{\mathcal{T}_i} -i\omega\vec{\varphi}_{ij}^t \Lambda_i \vec{T}_i &= -\int_{\mathcal{T}_i} \vec{\nabla}\vec{\varphi}_{ij} : \vec{G}(\vec{V}_i) + \frac{1}{2} \sum_{k \in V(i)} \int_{\mathcal{T}_{ik}} \vec{\varphi}_{ij}^t (\vec{G}(\vec{V}_i) + \vec{G}(\vec{V}_k)) \vec{n}_{ik} \\ &\quad - \int_{\mathcal{T}_i} i\omega\vec{\varphi}_{ij}^t \Lambda_i \vec{T}_i^0, \end{aligned} \quad (1.36)$$

puis en remplaçant \vec{V}_i et \vec{T}_i par leur approximations (équation (1.28)) :

$$\begin{aligned} \sum_{k=1}^{d_i} -i\omega\rho_i V_{ik} \int_{\mathcal{T}_i} \vec{\varphi}_{ij}^t \vec{\varphi}_{ik} &= -\sum_{k=1}^{d_i} T_{ik} \int_{\mathcal{T}_i} \vec{\nabla}\vec{\varphi}_{ij} : \vec{F}(\vec{\varphi}_{ik}) \\ &\quad + \frac{1}{2} \sum_{k \in V(i)} \left[\sum_{r=1}^{d_i} T_{ir} \int_{\mathcal{T}_{ik}} \vec{\varphi}_{ij}^t \vec{F}(\vec{\varphi}_{ir}) \vec{n}_{ik} + \sum_{s=1}^{d_k} T_{ks} \int_{\mathcal{T}_{ik}} \vec{\varphi}_{ij}^t \vec{F}(\vec{\varphi}_{ks}) \vec{n}_{ik} \right] \\ &\quad + \sum_{k=1}^{d_i} \rho_i f_{ik} \int_{\mathcal{T}_i} \vec{\varphi}_{ij}^t \vec{\varphi}_{ik} \\ \sum_{k=1}^{d_i} -i\omega T_{ik} \int_{\mathcal{T}_i} \vec{\varphi}_{ij}^t \Lambda_i \vec{\varphi}_{ik} &= -\sum_{k=1}^{d_i} V_{ik} \int_{\mathcal{T}_i} \vec{\nabla}\vec{\varphi}_{ij} : \vec{G}(\vec{\varphi}_{ik}) \\ &\quad + \frac{1}{2} \sum_{k \in V(i)} \left[\sum_{r=1}^{d_i} V_{ir} \int_{\mathcal{T}_{ik}} \vec{\varphi}_{ij}^t \vec{G}(\vec{\varphi}_{ir}) \vec{n}_{ik} + \sum_{s=1}^{d_k} V_{ks} \int_{\mathcal{T}_{ik}} \vec{\varphi}_{ij}^t \vec{G}(\vec{\varphi}_{ks}) \vec{n}_{ik} \right] \\ &\quad - \sum_{k=1}^{d_i} i\omega T_{ij}^0 \int_{\mathcal{T}_i} \vec{\varphi}_{ij}^t \Lambda_i \vec{\varphi}_{ik}. \end{aligned} \quad (1.37)$$

Introduisons alors les notations matricielles suivantes pour chaque vecteur de fonction de base $\vec{\varphi}$:

$$\begin{aligned} \vec{\nabla}\vec{\varphi} : \vec{F}(\vec{\varphi}) &= \sum_{\alpha \in \{x,z\}} (\partial_\alpha \vec{\varphi})^t s_\alpha M_\alpha \vec{\varphi} \\ \vec{\nabla}\vec{\varphi} : \vec{G}(\vec{\varphi}) &= \sum_{\alpha \in \{x,z\}} (\partial_\alpha \vec{\varphi})^t s_\alpha N_\alpha \vec{\varphi}, \end{aligned} \quad (1.38)$$

ainsi que les matrices \mathcal{P} , \mathcal{Q} , \mathcal{P}' et \mathcal{Q}' :

$$\begin{aligned}
 \mathcal{P}_{ik} &= \sum_{\alpha \in \{x,z\}} n_{ik\alpha} s_{i\alpha} M_\alpha \\
 \mathcal{Q}_{ik} &= \sum_{\alpha \in \{x,z\}} n_{ik\alpha} s_{i\alpha} N_\alpha \\
 \mathcal{P}'_{ik} &= \sum_{\alpha \in \{x,z\}} n_{ik\alpha} s_{k\alpha} M_\alpha \\
 \mathcal{Q}'_{ik} &= \sum_{\alpha \in \{x,z\}} n_{ik\alpha} s_{k\alpha} N_\alpha.
 \end{aligned} \tag{1.39}$$

Nous avons ainsi les égalités suivantes :

$$\begin{aligned}
 \vec{\varphi}_{ij}^t \vec{F}(\vec{\varphi}_{ir}) \vec{n}_{ik} &= \vec{\varphi}_{ij}^t \mathcal{P}_{ik} \vec{\varphi}_{ir} \\
 \vec{\varphi}_{ij}^t \vec{G}(\vec{\varphi}_{ir}) \vec{n}_{ik} &= \vec{\varphi}_{ij}^t \mathcal{Q}_{ik} \vec{\varphi}_{ir} \\
 \vec{\varphi}_{ij}^t \vec{F}(\vec{\varphi}_{ks}) \vec{n}_{ik} &= \vec{\varphi}_{ij}^t \mathcal{P}'_{ik} \vec{\varphi}_{ks} \\
 \vec{\varphi}_{ij}^t \vec{G}(\vec{\varphi}_{ks}) \vec{n}_{ik} &= \vec{\varphi}_{ij}^t \mathcal{Q}'_{ik} \vec{\varphi}_{ks}.
 \end{aligned} \tag{1.40}$$

Remarque : Les propriétés des matrices $M_\alpha = N_\alpha^t$ et $\vec{n}_{ik} = -\vec{n}_{ki}$ permettent de définir les relation suivantes en dehors des zones PML :

$$\begin{aligned}
 \mathcal{Q}_{ik} &= \mathcal{P}_{ik}^t \\
 \mathcal{Q}'_{ik} &= \mathcal{P}'_{ik}{}^t \\
 \mathcal{P}_{ik} &= -\mathcal{P}'_{ki} \\
 \mathcal{Q}_{ik} &= -\mathcal{Q}'_{ki}.
 \end{aligned} \tag{1.41}$$

Notons $\tilde{\mathcal{K}}_i^1$ et $\tilde{\mathcal{K}}_i^2$ les matrices de masse locales définies par :

$$\begin{aligned}
 \left(\tilde{\mathcal{K}}_i^1\right)_{jk} &= \int_{\mathcal{I}_i} \vec{\varphi}_{ij}^t \vec{\varphi}_{ik} \quad 1 \leq j, k \leq d_i \\
 \left(\tilde{\mathcal{K}}_i^2\right)_{jk} &= \int_{\mathcal{I}_i} \vec{\varphi}_{ij}^t \Lambda \vec{\varphi}_{ik} \quad 1 \leq j, k \leq d_i,
 \end{aligned}$$

et si nous notons \mathbf{V}_i et \mathbf{T}_i les vecteurs colonnes $(V_{ij})_{1 \leq j \leq d_i}$ et $(T_{ij})_{1 \leq j \leq d_i}$, le système (1.37) s'écrit alors :

$$\begin{aligned}
 -i\omega \rho_i \left(\tilde{\mathcal{K}}_i^1 \mathbf{V}_i\right)_j &= - \sum_{\alpha \in \{x,z\}} \sum_{k=1}^{d_i} T_{ik} \int_{\mathcal{I}_i} (\partial_\alpha \vec{\varphi})^t s_\alpha M_\alpha \vec{\varphi} \\
 &+ \frac{1}{2} \sum_{k \in V(i)} \left[\sum_{r=1}^{d_i} T_{ir} \int_{\mathcal{I}_{ik}} \vec{\varphi}_{ij}^t \mathcal{P}_{ik} \vec{\varphi}_{ir} + \sum_{s=1}^{d_k} T_{ks} \int_{\mathcal{I}_{ik}} \vec{\varphi}_{ij}^t \mathcal{P}'_{ik} \vec{\varphi}_{ks} \right] \\
 &+ \rho_i \left(\tilde{\mathcal{K}}_i^1 \mathbf{f}_i\right)_j \\
 -i\omega \left(\tilde{\mathcal{K}}_i^2 \mathbf{T}_i\right)_j &= - \sum_{\alpha \in \{x,z\}} \sum_{k=1}^{d_i} V_{ik} \int_{\mathcal{I}_i} (\partial_\alpha \vec{\varphi})^t s_\alpha N_\alpha \vec{\varphi}
 \end{aligned}$$

$$\begin{aligned}
 & + \frac{1}{2} \sum_{k \in V(i)} \left[\sum_{r=1}^{d_i} V_{ir} \int_{\mathcal{T}_{ik}} \vec{\varphi}_{ij}^t \mathcal{Q}_{ik} \vec{\varphi}_{ir} + \sum_{s=1}^{d_k} V_{ks} \int_{\mathcal{T}_{ik}} \vec{\varphi}_{ij}^t \mathcal{Q}'_{ik} \vec{\varphi}_{ks} \right] \\
 & - i\omega \left(\tilde{\mathcal{K}}_i^2 \mathbf{T}_i^0 \right)_j.
 \end{aligned} \tag{1.42}$$

Nous pouvons alors finalement définir $\vec{\mathbf{V}}_i$ et $\vec{\mathbf{T}}_i$, les vecteurs décrivant les champs de vitesses et contraintes dans la cellule \mathcal{T}_i :

$$\begin{aligned}
 \vec{\mathbf{V}}_i &= (\mathbf{V}_{\mathbf{x}_i}, \mathbf{V}_{\mathbf{z}_i})^t \\
 \vec{\mathbf{T}}_i &= (\mathbf{T}_{\mathbf{1}_i}, \mathbf{T}_{\mathbf{2}_i}, \mathbf{T}_{\mathbf{3}_i})^t.
 \end{aligned} \tag{1.43}$$

Hypothèse : Nous supposons désormais que les fonctions de base ($\vec{\varphi}$) sont prises du même type ($\vec{\varphi}_{ij}, 1 \leq j \leq d_i$) pour représenter chaque composante des vecteurs $\vec{\mathbf{V}}_i$ et $\vec{\mathbf{T}}_i$.

Le système (1.42) peut alors être réécrit sous forme compacte :

$$\begin{aligned}
 -i\omega \rho_i \left(\mathcal{I}_3 \otimes \mathcal{K}_i \vec{\mathbf{V}}_i \right) &= - \sum_{\alpha \in \{x,z\}} \left(s_\alpha M_\alpha \otimes \mathcal{E}_{i_\alpha} \vec{\mathbf{T}}_i \right) \\
 &+ \frac{1}{2} \sum_{k \in V(i)} \left[(\mathcal{P}_{ik} \otimes \mathcal{F}_{ik} \vec{\mathbf{T}}_i) + (\mathcal{P}'_{ik} \otimes \mathcal{G}_{ik} \vec{\mathbf{T}}_k) \right] \\
 &+ \rho_i \left(\mathcal{I}_3 \otimes \mathcal{K}_i^1 \vec{\mathbf{T}}_i \right) \\
 -i\omega \left(\Lambda_i \otimes \mathcal{K}_i \vec{\mathbf{T}}_i \right) &= - \sum_{\alpha \in \{x,z\}} \left(s_\alpha N_\alpha \otimes \mathcal{E}_{i_\alpha} \vec{\mathbf{V}}_i \right) \\
 &+ \frac{1}{2} \sum_{k \in V(i)} \left[(\mathcal{Q}_{ik} \otimes \mathcal{F}_{ik} \vec{\mathbf{V}}_i) + (\mathcal{Q}'_{ik} \otimes \mathcal{G}_{ik} \vec{\mathbf{V}}_k) \right] \\
 &- i\omega \left(\Lambda_i \otimes \mathcal{K}_i \vec{\mathbf{T}}_i^0 \right),
 \end{aligned} \tag{1.44}$$

où nous avons l'expression des différentes matrices :

- Matrice de masse \mathcal{K}_i définie par intégration de surface (2D) sur la cellule \mathcal{T}_i :

$$\left(\mathcal{K}_i \right)_{jk} = \int_{\mathcal{T}_i} \varphi_{ij} \varphi_{ik} d\mathcal{T} \quad 1 \leq j, k \leq d_i \tag{1.45}$$

- Matrice \mathcal{E}_{i_α} définie par intégration de surface (2D) sur la cellule \mathcal{T}_i :

$$\left(\mathcal{E}_{i_\alpha} \right)_{jk} = \int_{\mathcal{T}_i} (\partial_\alpha \varphi_{ij}) \varphi_{ik} d\mathcal{T} \quad 1 \leq j, k \leq d_i \tag{1.46}$$

- Matrices \mathcal{F}_{ik} et \mathcal{G}_{ik} définie par intégration linéique (1D) sur l'interface \mathcal{T}_{ik} entre les cellules \mathcal{T}_i et \mathcal{T}_k :

$$\begin{aligned}
 \left(\mathcal{F}_{ik} \right)_{jr} &= \int_{\mathcal{T}_{ik}} \varphi_{ij} \varphi_{ir} d\mathcal{T} \quad 1 \leq j, r \leq d_i \\
 \left(\mathcal{G}_{ik} \right)_{js} &= \int_{\mathcal{T}_{ik}} \varphi_{ij} \varphi_{ks} d\mathcal{T} \quad 1 \leq j \leq d_i, 1 \leq s \leq d_k
 \end{aligned} \tag{1.47}$$

Le produit tensoriel \otimes s'exprime pour toutes matrices $\mathcal{A} = (a_{ij})_{\substack{1 \leq i \leq n \\ 1 \leq j \leq m}} \in \mathcal{M}_{n,m}$ et $\mathcal{B} = (b_{ij})_{\substack{1 \leq i \leq p \\ 1 \leq j \leq q}} \in \mathcal{M}_{p,q}$ sous la forme :

$$\mathcal{A} \otimes \mathcal{B} = \begin{bmatrix} a_{11}\mathcal{B} & \cdots & a_{1m}\mathcal{B} \\ \vdots & \vdots & \vdots \\ a_{n1}\mathcal{B} & \cdots & a_{nm}\mathcal{B} \end{bmatrix} \in \mathcal{M}_{np,mq}. \quad (1.48)$$

En supposant que la matrice de masse est inversible, nous pouvons introduire les matrices suivantes :

$$\begin{aligned} \tilde{\mathcal{E}}_{i_x} &= A_i \mathcal{K}_i^{-1} \mathcal{E}_{i_x} \\ \tilde{\mathcal{E}}_{i_z} &= A_i \mathcal{K}_i^{-1} \mathcal{E}_{i_z} \\ \tilde{\mathcal{F}}_{i1} &= A_i \mathcal{K}_i^{-1} \mathcal{F}_{i1} \\ \tilde{\mathcal{F}}_{i2} &= A_i \mathcal{K}_i^{-1} \mathcal{F}_{i2} \\ \tilde{\mathcal{F}}_{i3} &= A_i \mathcal{K}_i^{-1} \mathcal{F}_{i3} \\ \tilde{\mathcal{G}}_{i1} &= A_i \mathcal{K}_i^{-1} \mathcal{G}_{i1} \\ \tilde{\mathcal{G}}_{i2} &= A_i \mathcal{K}_i^{-1} \mathcal{G}_{i2} \\ \tilde{\mathcal{G}}_{i3} &= A_i \mathcal{K}_i^{-1} \mathcal{G}_{i3}, \end{aligned}$$

qui permettent d'écrire le système sous forme éclatée :

$$\begin{aligned} -i\omega \rho_i A_i \vec{V}_{x_i} &= - \left[s_{i_x} \tilde{\mathcal{E}}_{i_x} (\vec{T}_{1_i} + \vec{T}_{2_i}) + s_{i_z} \tilde{\mathcal{E}}_{i_z} \vec{T}_{3_i} \right] \\ &+ \frac{1}{2} \sum_{k \in V(i)} \tilde{\mathcal{F}}_{ik} \left[(\vec{T}_{1_i} + \vec{T}_{2_i}) s_{i_x} n_{ik_x} + \vec{T}_{3_i} s_{i_z} n_{ik_z} \right] \\ &+ \frac{1}{2} \sum_{k \in V(i)} \tilde{\mathcal{G}}_{ik} \left[(\vec{T}_{1_k} + \vec{T}_{2_k}) s_{k_x} n_{ik_x} + \vec{T}_{3_k} s_{k_z} n_{ik_z} \right] \\ &+ \rho_i A_i \vec{f}_{x_i} \\ -i\omega \rho_i A_i \vec{V}_{z_i} &= - \left[s_{i_x} \tilde{\mathcal{E}}_{i_x} \vec{T}_{3_i} + s_{i_z} \tilde{\mathcal{E}}_{i_z} (\vec{T}_{1_i} - \vec{T}_{2_i}) \right] \\ &+ \frac{1}{2} \sum_{k \in V(i)} \tilde{\mathcal{F}}_{ik} \left[\vec{T}_{3_i} s_{i_x} n_{ik_x} + (\vec{T}_{1_i} - \vec{T}_{2_i}) s_{i_z} n_{ik_z} \right] \\ &+ \frac{1}{2} \sum_{k \in V(i)} \tilde{\mathcal{G}}_{ik} \left[\vec{T}_{3_k} s_{k_x} n_{ik_x} + (\vec{T}_{1_k} - \vec{T}_{2_k}) s_{k_z} n_{ik_z} \right] \\ &+ \rho_i A_i \vec{f}_{z_i} \\ -i\omega A_i \vec{T}_{1_i} &= (\lambda_i + \mu_i) \left\{ - \left[s_{i_x} \tilde{\mathcal{E}}_{i_x} \vec{V}_{x_i} + s_{i_z} \tilde{\mathcal{E}}_{i_z} \vec{V}_{z_i} \right] \right. \\ &+ \frac{1}{2} \sum_{k \in V(i)} \tilde{\mathcal{F}}_{ik} \left[\vec{V}_{x_i} s_{i_x} n_{ik_x} + \vec{V}_{z_i} s_{i_z} n_{ik_z} \right] \\ &+ \left. \frac{1}{2} \sum_{k \in V(i)} \tilde{\mathcal{G}}_{ik} \left[\vec{V}_{x_k} s_{k_x} n_{ik_x} + \vec{V}_{z_k} s_{k_z} n_{ik_z} \right] \right\} \end{aligned}$$

$$\begin{aligned}
 & -i\omega A_i \vec{T}_{1_i}^0 \\
 -i\omega A_i \vec{T}_{2_i} &= \mu_i \left\{ - \left[s_{i_x} \tilde{\mathcal{E}}_{i_x} \vec{V}_{x_i} - s_{i_z} \tilde{\mathcal{E}}_{i_z} \vec{V}_{z_i} \right] \right. \\
 & + \frac{1}{2} \sum_{k \in V(i)} \tilde{\mathcal{F}}_{ik} \left[\vec{V}_{x_i} s_{i_x} n_{ik_x} - \vec{V}_{z_i} s_{i_z} n_{ik_z} \right] \\
 & + \left. \frac{1}{2} \sum_{k \in V(i)} \tilde{\mathcal{G}}_{ik} \left[\vec{V}_{x_k} s_{k_x} n_{ik_x} - \vec{V}_{z_k} s_{k_z} n_{ik_z} \right] \right\} \\
 & - i\omega A_i \vec{T}_{2_i}^0 \\
 -i\omega A_i \vec{T}_{3_i} &= \mu_i \left\{ - \left[s_{i_x} \tilde{\mathcal{E}}_{i_x} \vec{V}_{z_i} + s_{i_z} \tilde{\mathcal{E}}_{i_z} \vec{V}_{x_i} \right] \right. \\
 & + \frac{1}{2} \sum_{k \in V(i)} \tilde{\mathcal{F}}_{ik} \left[\vec{V}_{z_i} s_{i_x} n_{ik_x} + \vec{V}_{x_i} s_{i_z} n_{ik_z} \right] \\
 & + \left. \frac{1}{2} \sum_{k \in V(i)} \tilde{\mathcal{G}}_{ik} \left[\vec{V}_{z_k} s_{k_x} n_{ik_x} + \vec{V}_{x_k} s_{k_z} n_{ik_z} \right] \right\} \\
 & - i\omega A_i \vec{T}_{3_i}^0
 \end{aligned} \tag{1.49}$$

Le système d'équation générique (1.49) est développé pour des fonctions d'interpolation arbitraires. Un choix de fonction doit donc être fait. Les polynômes sont des fonctions très souvent utilisés en analyse numérique pour représenter des évolutions de phénomènes physiques grâce au large choix possible d'ordres d'interpolation. Les fonctions de base exponentielles complexes peuvent également être utilisées pour représenter les phénomènes périodiques, mais sont souvent associées à des discrétisations régulières de l'espace (*cf.* partie 1.2.3 sur les méthodes pseudo-spectrales) (Hamming, 1973).

Quel que soit le type d'interpolation choisi, le système (1.49) se ramène à résoudre le système linéaire

$$\mathbf{Ax} = \mathbf{b}, \tag{1.50}$$

pour chaque fréquence discrète considérée. \mathbf{A} est la matrice d'impédance, également nommée opérateur du problème direct, qui dépend du maillage, des propriétés physiques, des fonctions d'interpolation et de la fréquence. \mathbf{x} représente les champs d'ondes de contraintes et de vitesses. \mathbf{b} représente le terme source d'excitation du milieu.

Pour notre application, des fonctions de base polynomiales seront utilisées. Le choix des ordres d'interpolation est crucial, car il va déterminer à la fois la précision des solutions modélisées, la taille acceptable des cellules de discrétisation du milieu, mais également le coût de calcul qui doit être limité dans le contexte applicatif de l'inversion des formes d'ondes. La Figure 1.3 illustre la représentation d'un champ sinusoïdal 1D par des polynômes d'ordre 0 à 3, de manière continue. Notons que plus l'ordre d'interpolation est élevé, plus l'erreur de représentation est faible. Cependant, en vue de l'inversion des formes d'ondes, une part importante de l'information utile à l'inversion est contenue dans la phase. Une représentation de bas-ordre permet une estimation très raisonnable de la phase et explique pourquoi, dans la suite, des ordres d'interpolation faibles permettent d'obtenir des résultats satisfaisants.

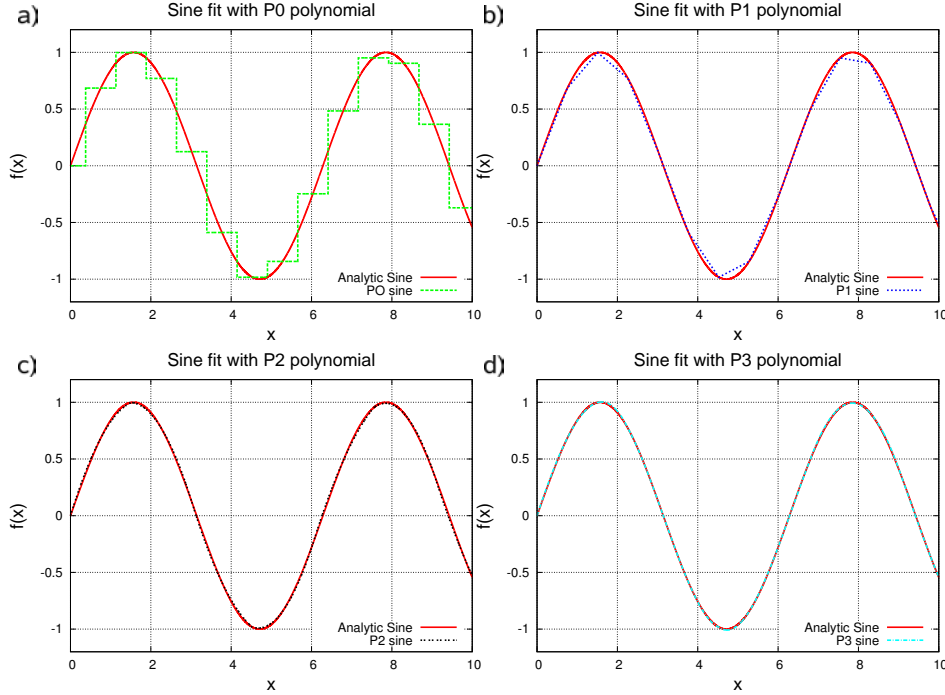


FIG. 1.3 – Représentation continue d'un champ sinusoïdal 1D par des polynômes d'ordres croissants de 0 à 3. Pour toutes les représentations, le nombre de degrés de liberté reste constant.

1.6 Application GD P_0 ou volume fini

1.6.1 Présentation de l'article

La représentation d'un champ périodique par des séries de fonctions constantes par morceau peut paraître quelque peu surprenante au premier abord. Cependant, c'est bien la représentation qu'offre le schéma d'interpolation P_0 , le plus bas ordre possible. Pour cette interpolation, chaque champ est représenté de façon constante pour chacune des cellules du milieu. Le nombre de degrés de liberté d_i du vecteur de fonction de base $\tilde{\varphi}_{ij}$ est donc égal à 1. Ce vecteur est un scalaire de valeur constante, indépendante de la position de l'espace.

Même si ce schéma peut paraître inadapté à la modélisation de la propagation des ondes, il se révèle être une alternative de choix. Ce schéma est équivalent à la formulation volumes finis P_0 et peut être facilement associé au schéma DF d'ordre 2 s'il est formulé sur un maillage carré (*cf.* partie 1.6.9). Son utilisation s'est révélée efficace et précise pour la résolution des équations de Maxwell (Remaki, 2000).

Grâce à la grande simplicité de la représentation des champs à travers des fonctions constantes par morceau, le formalisme générique de la partie 1.5.3 peut être occulté. Ainsi, les matrices $\tilde{\mathcal{F}}$ et $\tilde{\mathcal{G}}$ deviennent des scalaires et les matrices $\tilde{\mathcal{E}}$ sont nulles dans cette formulation.

Notons également qu'une formulation parsimonieuse du système (Luo & Schuster, 1990) est présentée afin de formuler de façon discrète le système élastodynamique du second ordre en vitesses comme cela a pu être fait en DF par Gelis *et al.* (2007). Nous reviendrons dans la

partie 1.7.3 sur l'intérêt et la faisabilité de cette formulation.

Dans la suite, je présente un article publié dans la revue *Geophysical Journal International* qui développe la théorie, l'implémentation et la validation numérique de la méthode GD P_0 . Cet article illustre également la non convergence de la méthode sur maillages non-structurés, ce qui motive l'extension de la méthode aux ordres d'interpolation supérieurs P_1 et P_2 (*cf.* partie 1.7).

**Parsimonious finite-volume frequency-domain method for 2D
P-SV-wave modelling**

R. Brossier, J. Virieux and S. Operto

Geophysical Journal International, 2008, **175**, 541-559

1.6.2 Summary

A new numerical technique for solving 2D elastodynamic equations based on a finite-volume frequency-domain approach is proposed. This method has been developed as a tool to perform two-dimensional (2D) elastic frequency-domain full-waveform inversion. In this context, the system of linear equations that results from the discretisation of the elastodynamic equations is solved with a direct solver, allowing efficient multiple-source simulations at the partial expense of the memory requirement. The discretisation of the finite-volume approach is through triangles. Only fluxes with the required quantities are shared between the cells, relaxing the meshing conditions, as compared to finite-element methods. The free surface is described along the edges of the triangles, which can have different slopes. By applying a parsimonious strategy, the stress components are eliminated from the discrete equations and only the velocities are left as unknowns in the triangles. Together with the local support of the P_0 finite-volume stencil, the parsimonious approach allows the minimising of core memory requirements for the simulation. Efficient perfectly matched layer absorbing conditions have been designed for damping the waves around the grid. The numerical dispersion of this FV formulation is similar to that of $O(\Delta x^2)$ staggered-grid finite-difference formulations when considering structured triangular meshes. The validation has been performed with analytical solutions of several canonical problems and with numerical solutions computed with a well-established finite-difference time-domain method in heterogeneous media. In the presence of a free surface, the finite-volume method requires ten triangles per wavelength for a flat topography, and fifteen triangles per wavelength for more complex shapes, well below the criteria required by the staircase approximation of $O(\Delta x^2)$ finite-difference methods. Comparisons between the frequency-domain finite-volume and the $O(\Delta x^2)$ rotated finite-difference methods also show that the former is faster and less memory demanding for a given accuracy level, an attractive feature for frequency-domain seismic inversion. We have thus developed an efficient method for 2D P-SV-wave modelling on structured triangular meshes as a tool for frequency-domain full-waveform inversion. Further work is required to improve the accuracy of the method on unstructured meshes.

Keywords

Seismic wave propagation, numerical modelling, finite volume approach, elastic waves, frequency domain

1.6.3 Introduction

Seismic wave propagation has been investigated by various numerical methods, such as finite-difference (FD), finite-element (FE) and boundary integral equations. When considering media with complex topography and a possible water layer, few approaches are available for efficient forward modelling, especially when the modelling scheme is dedicated to seismic full-waveform inversion applications that require thousands of forward modelling stages. The FD staggered-grid method proposed by Madariaga (1976) and Virieux (1986a) using the Yee scheme (Yee, 1966), which is based on a first-order velocity-stress hyperbolic system, is relatively popular and has been used intensively in the time domain for seismic imaging in spite of the limited solution accuracy when considering free surfaces with topography (Gauthier *et al.*, 1986). The solution has been shown to be stable when considering fluid zones inside the FD grid. The staircase approximation of the free surface, as proposed by Robertsson (1996), requires dense meshing. Recently, attempts have been proposed through mesh deformation (Hestholm & Ruud, 2002) and through immersed implicit boundary methods (Lombard *et al.*, 2008), although these need specific numerical developments near to the free surface.

The FE approaches, as proposed for elastic wave propagation by Marfurt (1984), have been reshaped recently with low-order interpolation for efficient 3D simulations (Bielak *et al.*, 2003; Yoshimura *et al.*, 2003; Koketsu *et al.*, 2004), while recent high-order interpolations have led to the so-called spectral element method (Faccioli *et al.*, 1997; Vilotte *et al.*, 2005; Komatitsch & Vilotte, 1998; Chaljub *et al.*, 2003). Both of these approaches with different mesh densities have allowed accurate modelling of free surface effects, as these FE methods adapt the mesh to the surface topography. Moczo *et al.* (1997) have proposed the combination of FD and FE methods to deal efficiently with complex topographies. Due to the weak formulation of these methods, the fluid/solid interface has to be tackled by explicit boundary conditions.

Other numerical methods have tried to avoid the necessary continuity of fields at shared nodes between elements, which has led to unconventional FE methods (Casadei *et al.*, 2002) and to discontinuous Galerkin (DG) methods (Cockburn *et al.*, 2000), which have been popularised in seismology by Dumbser, Käser and co-workers (Käser & Dumbser, 2006; Dumbser & Käser, 2006; Käser *et al.*, 2007; de la Puente *et al.*, 2007; Dumbser *et al.*, 2007). Both methods are particularly demanding of computer resources. Explicit boundary conditions have been worked out for fracture problems by BenJemaa *et al.* (2007) for a low-order interpolation of the DG method, which is nothing more than a FV method with constant value interpolation inside each element. These FV/DG methods work in the time domain using both velocity and stress fields, with very promising perspectives.

The aim of this study is to develop a frequency-domain modelling method that is suited to seismic imaging applications performed by frequency-domain full-waveform inversion. The above-mentioned methods may not be the optimal ones for this kind of application, for the reasons explained below. With the success of full-waveform inversion in the frequency domain (Pratt & Worthington, 1990; Pratt *et al.*, 1996, 1998), applications to real data using the acoustic approximation for 2D geometries have been performed for imaging complex structures (Ravaut *et al.*, 2004; Operto *et al.*, 2006), while the reconstruction of elastic parameters has been found to be a quite challenging problem (Gelis *et al.*, 2007). These approaches are based on a hierarchical multiscale inversion scheme that proceeds over a coarse subset of frequencies, from the low frequencies to the higher ones, and that requires a large number of forward simulations at each iteration of the multiscale reconstruction. To consider both onshore and

offshore applications of full-waveform inversion, the modelling method must be accurate in the case of complex media that incorporate either a free surface of arbitrary shape or a water layer, while remaining sufficiently fast to be able to perform multiple simulations in a tractable time. High-order accuracy methods that might perform well in the time domain turn out to be prohibitive in the frequency domain if the linear system that results from the discretisation of the frequency-domain wave equation is solved with a direct solver (Stekl & Pratt, 1998; Hustedt *et al.*, 2004). Direct solvers are generally used to perform 2D frequency-domain wave modelling because solutions for multiple sources can be efficiently computed by substitutions once the impedance matrix has been LU-factorized. A key feature for numerical efficiency is the compactness of the spatial operator that controls the numerical bandwidth of the impedance matrix, and therefore its fill-in during factorization. Therefore, we shall concentrate our comparisons on $O(\Delta x^2)$ FD methods that minimise the memory requirement of the linear system resolution for the frequency-domain formulation. The $O(\Delta x^4)$ FD method (Levander, 1988) has been shown to be an efficient compromise between memory saving and CPU demand for time-domain formulation. Unfortunately, even with a coarser meshing, this higher-order stencil dramatically increases the memory cost of the linear system resolution in the frequency-domain formulation (Hustedt *et al.*, 2004).

Optimal compact FD stencils based on the so-called mixed-grid method that combines the Cartesian FD stencil and the rotated FD stencil (Saenger *et al.*, 2000) and anti-lumped mass features have been designed for frequency-domain wave-propagation modelling. This approach of combining stencils has been shown to be very efficient for the acoustic wave equation (Jo *et al.*, 1996; Hustedt *et al.*, 2004), for which only four grid points per wavelength can be used. Stekl & Pratt (1998) have shown limitations in the elastic case, when a liquid-solid interface is involved; here, the Cartesian stencil has to be removed. The same limitation applies when considering a complex free surface, leaving second-order FD stencils. For the rotated FD stencil (Saenger *et al.*, 2000), the free surface condition is verified simply through a vacuum approach, with a minimum of 25 nodes per shear wavelength for a flat topography, and up to 60 nodes per shear wavelength with a complex topography (Saenger & Bohlen, 2004; Bohlen & Saenger, 2006) in the time domain. Gelis *et al.* (2007) have also shown that this is similar in the frequency domain, leading to the solving of rather significant sparse linear systems. The method we propose here will significantly reduce this dense sampling near to the free surface, a key issue for frequency formulation.

We thus present a FV P_0 method based on a first-order hyperbolic elastodynamic system in the frequency domain, as is usual for FV formulations (Remaki, 2000). We will deduce the discretised system of linear equations to be solved only for velocity components, using the parsimonious strategy (Luo & Schuster, 1990). The numerical dispersion behaviour of this scheme will be analysed before discussing the perfectly matched layer (PML) conditions (Berenger, 1994). The implementation of the source will require specific attention before going on to numerical validations against both the analytical solutions and the numerical solutions obtained by other numerical methods. For these examples, we will show the respective numerical costs of frequency-domain FV and $O(\Delta x^2)$ FD methods. We will conclude with the potential of this FV approach in the frequency domain for the modelling of seismic 2D P-SV waves and the perspectives for inversion.

1.6.4 Finite volume formulation in the frequency domain

We consider a first-order hyperbolic elastodynamic system for 2D P-SV waves in isotropic medium in the frequency domain where both velocities (V_x, V_z) and stress $(\sigma_{xx}, \sigma_{zz}, \sigma_{xz})$ are unknown quantities as described by the following differential system :

$$\begin{aligned}
-\iota\omega V_x &= \frac{1}{\rho(\mathbf{x})} \left\{ \frac{\partial \sigma_{xx}}{\partial x} + \frac{\partial \sigma_{xz}}{\partial z} \right\} + F_x \\
-\iota\omega V_z &= \frac{1}{\rho(\mathbf{x})} \left\{ \frac{\partial \sigma_{xz}}{\partial x} + \frac{\partial \sigma_{zz}}{\partial z} \right\} + F_z \\
-\iota\omega \sigma_{xx} &= (\lambda(\mathbf{x}) + 2\mu(\mathbf{x})) \frac{\partial V_x}{\partial x} + \lambda(\mathbf{x}) \frac{\partial V_z}{\partial z} - \iota\omega \sigma_{xx0} \\
-\iota\omega \sigma_{zz} &= \lambda(\mathbf{x}) \frac{\partial V_x}{\partial x} + (\lambda(\mathbf{x}) + 2\mu(\mathbf{x})) \frac{\partial V_z}{\partial z} - \iota\omega \sigma_{zz0} \\
-\iota\omega \sigma_{xz} &= \mu(\mathbf{x}) \left\{ \frac{\partial V_x}{\partial z} + \frac{\partial V_z}{\partial x} \right\} - \iota\omega \sigma_{xz0},
\end{aligned} \tag{1.51}$$

where the Lamé coefficients that describe the medium are denoted by λ , μ , the density by ρ and the angular frequency by ω . Source terms are either point forces (F_x, F_z) or applied stresses $(\sigma_{xx0}, \sigma_{zz0}, \sigma_{xz0})$ as introduced in system (1.51). The pure imaginary number defined as $x^2 = -1$ is denoted as ι . The Fourier transform follows the usual convention as $f(\omega) = \int_{-\infty}^{+\infty} f(t)e^{-i\omega t} dt$. To develop a pseudo-conservative formulation that will be useful for integration over a surface in 2D, we will consider the following new vector with three components $\vec{T}^t = (T_1, T_2, T_3) = ((\sigma_{xx} + \sigma_{zz})/2, (\sigma_{xx} - \sigma_{zz})/2, \sigma_{xz})$. Moreover, we must consider a finite domain, and therefore we apply PML absorbing conditions (Berenger, 1994) through the functions s_x, s_z for the velocity equations and the functions s'_x, s'_z for the stress equations. More details on the expression of these damping functions s_x, s_z, s'_x, s'_z are given below. The new differential system equivalent to system (1.51) can be written as :

$$\begin{aligned}
-\iota\omega\rho V_x &= s_x \frac{\partial(T_1 + T_2)}{\partial x} + s_z \frac{\partial T_3}{\partial z} + \rho F_x \\
-\iota\omega\rho V_z &= s_x \frac{\partial T_3}{\partial x} + s_z \frac{\partial(T_1 - T_2)}{\partial z} + \rho F_z \\
\frac{-\iota\omega T_1}{\lambda + \mu} &= s'_x \frac{\partial V_x}{\partial x} + s'_z \frac{\partial V_z}{\partial z} - \frac{\iota\omega T_1^0}{\lambda + \mu} \\
\frac{-\iota\omega T_2}{\mu} &= s'_x \frac{\partial V_x}{\partial x} - s'_z \frac{\partial V_z}{\partial z} - \frac{\iota\omega T_2^0}{\mu} \\
\frac{-\iota\omega T_3}{\mu} &= s'_x \frac{\partial V_z}{\partial x} + s'_z \frac{\partial V_x}{\partial z} - \frac{\iota\omega T_3^0}{\mu}
\end{aligned} \tag{1.52}$$

We apply a surface integration over a control cell identified by the index i . For practical reasons of meshing, control cells are often taken as triangles, but the formulation still stands for any polygonal cells as quadrangles, for example. The geometrical description of a medium depends on the meshing tool we are using, and the filling of 2D space with triangles or 3D space with tetrahedra is often provided by mesh designers. Appendix 1.6.9 illustrates the development of the FV method in a regular Cartesian mesh, which is equivalent to a second-order FD formulation in a Cartesian grid without the staggered grid structure. Therefore, we can compare the

numerical performances of the FV method we have developed and the staggered-grid $O(\Delta x^2)$ FD method, which is less intensive than the $O(\Delta x^2)$ FD method considering all of the unknowns at each node of the grid. The formulation presented here is based on triangular cells in a conformal mesh, which imposes three edges and neighbours for each cell considered. The quantities are constant inside each cell, an assumption known as the P_0 approximation. Higher-order interpolations of P_k are often referred to as discontinuous Galerkin methods (Käser & Dumbser, 2006), which should be compared with the new high-order FD schemes, a task that we have not tackled here since these approaches cannot be adapted to our frequency-domain formulation as tackled with a direct solver. According to Green's theorem, we end up with the discrete system written in a vectorial form (see Appendix 1.6.10 for the complete derivation) :

$$\begin{aligned} -\omega A_i \rho_i \vec{V}_i &= \sum_{j \in \partial K_i} l_{ij} G_{ij} + A_i \rho_i \vec{F}_i \\ -\omega A_i \Lambda_i \vec{T}_i &= \sum_{j \in \partial K_i} l_{ij} H_{ij} - \omega A_i \Lambda_i \vec{T}_i^0 \end{aligned} \quad (1.53)$$

The surface of the i cell is denoted by $A_i = \int_{K_i} dS$. The index $j \in \partial K_i$ labels the three neighbouring cells that have a joint edge with the i cell. The length of the edge between cells i and j is denoted by l_{ij} . The numerical approximation of fluxes is denoted by $l_{ij} H_{ij}$ and $l_{ij} G_{ij}$. The matrix Λ_i is the diagonal matrix defined by $\Lambda_i = \text{diag}(1/(\lambda_i + \mu_i), 1/\mu_i, 1/\mu_i)$. Finally, source vectors applied inside the i cell are denoted by \vec{F}_i and \vec{T}_i^0 .

Centred numerical fluxes of the velocity and stress components between two cells are introduced because they preserve a discrete energy inside the entire zone away from the PML. These were first proposed by Remaki (2000) and used by BenJemaa *et al.* (2007) for the elastodynamics in the time domain, and Dolean *et al.* (2006) for Maxwell equations in the frequency domain. This gives us the following estimation of the quantities G_{ij} and H_{ij} :

$$\begin{aligned} G_{ij} &= \sum_{r \in \{x,z\}} n_{ijr} N_k \frac{s_{r_i} \vec{T}_i + s_{r_j} \vec{T}_j}{2} \\ H_{ij} &= \sum_{r \in \{x,z\}} n_{ijr} M_k \frac{s'_{r_i} \vec{V}_i + s'_{r_j} \vec{V}_j}{2}, \end{aligned} \quad (1.54)$$

where the normal vector component r is oriented for each edge of cell i towards cell j and is denoted by n_{ijr} . Projector matrices defined for a vectorial formulation are denoted by M_k and N_k .

The geometrical properties of triangles give

$$\sum_{j \in \partial K_i} l_{ij} P_{ij} = 0 \quad (1.55)$$

where $P_{ij} = \sum_{r \in \{x,z\}} n_{ijr}$. These geometrical properties ensure that for the first-order system, the unknowns in a given cell i depend only on the unknowns of the surrounding cells, and not on the unknowns of the cell i that are cancelled out by the construction as it is built by centred fluxes.

Discrete equations expressing the stress components, as the second vectorial equation of system (1.53), can be eliminated by inserting them into the fluxes that are required in the

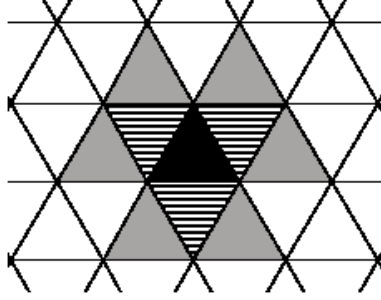


FIG. 1.4 – Illustration of the numerical scheme. In black, the central cell for which the solution is computed. The neighbouring cells (hatched) have unknowns that do not influence the scheme. In grey, the neighbours of these neighbouring cells with unknowns that are involved.

velocity equations, the first vectorial equation of system (1.53), thus reducing the memory requirement for wave propagation. Luo & Schuster (1990) proposed this parsimonious strategy for the FD method, and it turns out that it also works well for this FV method. After elimination of the stress components, we end up with the following two algebraic equations for the two unknown velocity components :

$$\begin{aligned}
 \omega^2 V_{x_i} &= \omega F_{x_i} + \frac{\omega}{A_i \rho_i} \sum_{j \in \partial K_i} \frac{l_{ij}}{2} \left\{ n_{ij_x} s_{x_j} \left[\frac{\iota(\lambda_j + \mu_j)}{\omega A_j} \sum_{k \in \partial K_j} \frac{l_{jk}}{2} (n_{jk_x} s'_{x_k} V_{x_k} + n_{jk_z} s'_{z_k} V_{z_k}) + T_{1_j}^0 \right] \right. \\
 &+ n_{ij_x} s_{x_j} \left[\frac{\iota \mu_j}{\omega A_j} \sum_{k \in \partial K_j} \frac{l_{jk}}{2} (n_{jk_x} s'_{x_k} V_{x_k} - n_{jk_z} s'_{z_k} V_{z_k}) + T_{2_j}^0 \right] \\
 &+ \left. n_{ij_z} s_{z_j} \left[\frac{\iota \mu_j}{\omega A_j} \sum_{k \in \partial K_j} \frac{l_{jk}}{2} (n_{jk_x} s'_{x_k} V_{z_k} + n_{jk_z} s'_{z_k} V_{x_k}) + T_{3_j}^0 \right] \right\} \\
 \omega^2 V_{z_i} &= \omega F_{z_i} + \frac{\omega}{A_i \rho_i} \sum_{j \in \partial K_i} \frac{l_{ij}}{2} \left\{ n_{ij_z} s_{z_j} \left[\frac{\iota(\lambda_j + \mu_j)}{\omega A_j} \sum_{k \in \partial K_j} \frac{l_{jk}}{2} (n_{jk_x} s'_{x_k} V_{x_k} + n_{jk_z} s'_{z_k} V_{z_k}) + T_{1_j}^0 \right] \right. \\
 &- n_{ij_z} s_{z_j} \left[\frac{\iota \mu_j}{\omega A_j} \sum_{k \in \partial K_j} \frac{l_{jk}}{2} (n_{jk_x} s'_{x_k} V_{x_k} - n_{jk_z} s'_{z_k} V_{z_k}) + T_{2_j}^0 \right] \\
 &+ \left. n_{ij_x} s_{x_j} \left[\frac{\iota \mu_j}{\omega A_j} \sum_{k \in \partial K_j} \frac{l_{jk}}{2} (n_{jk_x} s'_{x_k} V_{z_k} + n_{jk_z} s'_{z_k} V_{x_k}) + T_{3_j}^0 \right] \right\}, \tag{1.56}
 \end{aligned}$$

where $k \in \partial K_j$ labels the index of the cells j , the neighbours of the cell i . Due to properties of the triangle given by expression (1.55), and due to the centred flux estimations and the parsimonious formulation that lead to this algebraic system (1.56), the velocity unknowns only depend on the velocities of the neighbours of the neighbours of the cell : neighbouring velocity unknowns are not directly involved in the numerical scheme for a given element. Figure 1.4 illustrates this configuration on a regular mesh, where the numerical scheme centred on the black central cell depends on the unknowns at this cell and on the unknowns belonging to the neighbours of the neighbouring cells (grey cells). No dependency with hatched-cell unknowns is seen, as previously noted by LeVeque (2007) as the black/red pattern of the centred numerical schemes.

Equations (1.56) can be recast in matrix form as $\mathcal{A}V = \mathcal{B}$, where the sparse impedance

matrix \mathcal{A} contains 14 non-zero coefficients per row in the general case (i.e., without any regular structure), due to the expected irregular numbering of the cells inside the mesh. We should also stress that the corresponding matrix for the parsimonious $O(\Delta x^2)$ rotated FD stencil (Gelis *et al.*, 2007) has 18 non-zero elements.

The free-surface condition is explicitly expressed in the numerical scheme by considering a ghost cell above the free surface that has the same velocity and the opposite stress components to those below the free surface, in order to fulfil the zero stress at the free surface while keeping the correct numerical estimation of the particle velocity at the free surface. Using these velocities and stresses in the ghost cell, the stress flux across the free surface interface vanishes, while the velocity flux is twice the value that would have been obtained by neglecting the flux contribution above the free surface. This boundary condition has been implemented by modifying the impedance matrix accordingly without introducing any new unknown quantities. Similarly, fluid-solid interface modelling requires discontinuities of particle tangential velocities at the boundary. The specific interface conditions that follow the same strategy as that described by BenJemaa *et al.* (2007) for crack simulations, have also led to a modification of the impedance matrix for both cells that share an interface segment, without increasing its complexity. The same equations with identical numerical schemes are used for both fluid and solid media, where the μ value is set to zero inside the fluid medium.

1.6.5 Numerical properties

Discretisation leads to numerical dispersion in the particle velocity wavefields. For unstructured meshes, the dispersion could not be estimated analytically, while a regular distribution of equilateral triangles will allow such an investigation. Moreover, we must consider specific properties at the edges of the grid for the extension of the medium to infinity. Finally, the source implementation has to excite the entire grid without exciting a specific sub-grid that is related to the centred pattern of our system.

1.6.5.1 Numerical dispersion analysis

The numerical dispersion can be estimated for such a discrete system of a regular distribution of equilateral triangles (see the mesh configuration in Figure 1.4 for this pattern). We here consider an incident-plane wave propagating inside an infinite and homogeneous medium away from the PML and source zones. Due to the regular mesh, the Hermitian structure of the matrix makes eigenvalues real. They are computed numerically for different Poisson ratios from 0.0 to 0.5, for different number of cells per wavelength, and with incidence angles ranging between 0° and 180° with steps of 15° . The dispersion curves are quite similar to those obtained by Virieux (1986b) using an accurate $O(\Delta x^2)$ FD scheme, as shown in Figure 1.5. The rule of thumb of ten grid points per wavelength appears to provide acceptable propagation dispersion, whatever the value of the shear wave velocity, which can decrease to zero without any numerical instability. The $O(\Delta x^2)$ rotated staggered-grid FD stencil provides similar results with a grid length that is higher by a factor of $\sqrt{2}$ (Saenger *et al.*, 2000). We should also note that we have considered the triangle edge length for the FV approach and the grid step length for the FD approach, as our discrete reference values.

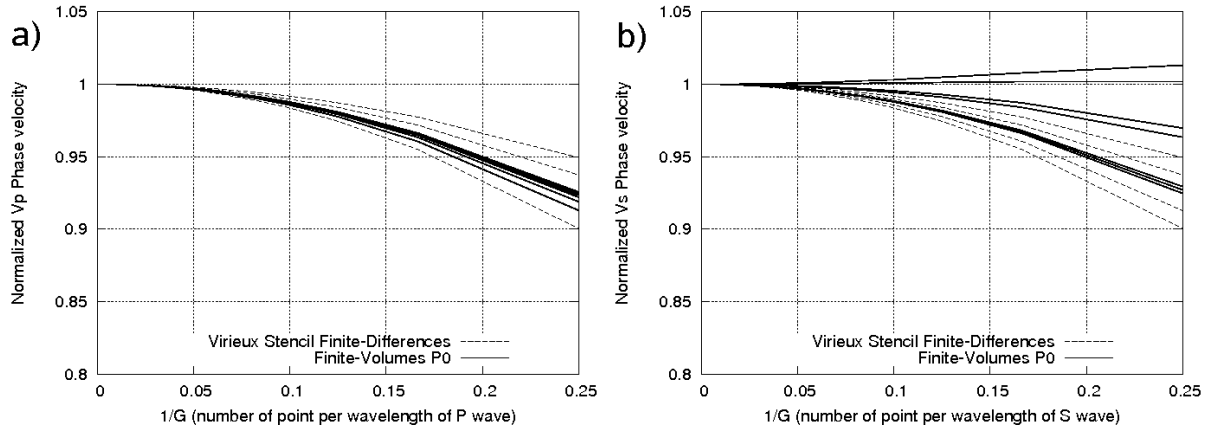


FIG. 1.5 – (a) P-wave and (b) S-wave normalised phase-velocity dispersion curves for different plane waves with various incident angles for both the FV (continuous lines) and the $O(\Delta x^2)$ Cartesian FD approaches (dashed lines).

1.6.5.2 PML absorbing boundary conditions

The frequency domain allows the straightforward numerical implementation of PML conditions, without any splitting of particle velocity components or additional integration of memory variables, as for the time-domain formulation, due to the complex coordinate change (Chew & Liu, 1996). Numerical tests show that the PML efficiency strongly depends on the mesh structure in the PMLs. The PML absorbing boundary condition requires that PML-PML interfaces be oriented along the Cartesian directions (Berenger, 1994). This condition is not satisfied if triangles of arbitrary orientations are used in the PML layers. In such a case, we saw poor absorption, as illustrated in Figure 1.6-(a) for a distribution of non-constrained triangles in arbitrary orientations in the PML layers. Therefore, a constrained mesh in PML zones with multiple layer structures parallel to the Cartesian directions (see Figure 1.7 for the discretisation of the lower left quarter of the medium) provides efficient absorption of the elastic waves. A transition is, of course, performed between the main central zone of the grid and the PML constrained zones. With this constraint on the mesh construction, at each parallel interface of the PML layers, the major part of the energy is contained along the damped direction, and the numerical flux energy is globally damped by the variations in the PML functions s_x, s_z, s'_x, s'_z as we move deeper into the PML zone. We can here note that use in the PML of quadrangle cells oriented along the Cartesian axes should provide a very efficient behaviour as absorbing boundaries, although this is less easy to implement with triangular mesh generators.

The standard frequency PML function is defined by :

$$s_r = \frac{1}{1 + \nu \gamma_r / \omega}, \quad (1.57)$$

where the index r can be x or z . A similar expression is obtained for the s'_r function. Both of these functions are only used inside the PML zones. Outside of the damping zone, the values of the s_r and s'_r functions are simply equal to 1. The value γ_r is typically used as a polynomial or cosine function for progressive energy damping. Numerical tests have shown better behaviour with a modified PML function derived from the developments indicated by Drossaert & Gian-

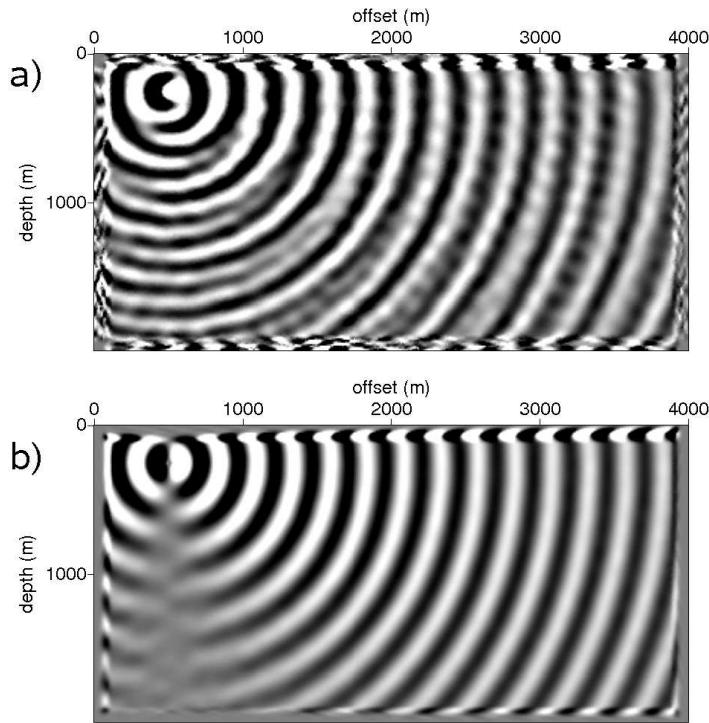


FIG. 1.6 – Frequency map solutions where the real part of the horizontal velocity for different PML configurations is displayed. (a) No mesh constraints applied, and (b) with mesh constraints applied.

nopoulos (2007). The definition of the damping functions s_r and s'_r are extended through the following equation :

$$s_r = \frac{1}{\kappa_r + \nu\gamma_r/\omega}, \quad (1.58)$$

with the expressions κ_r and γ_r as cosine functions. A linear dependency of the term γ_r with frequency has shown good damping behaviour that is almost independent of the frequency. These functions are defined as :

$$\begin{aligned} \gamma_r(l) &= \omega B \cos\left(\frac{l\pi}{2l_{pml}}\right) \\ \kappa_r(l) &= 1 + C \cos\left(\frac{l\pi}{2l_{pml}}\right), \end{aligned} \quad (1.59)$$

along the perpendicular direction, where l_{pml} is the size of the PML zone, which is taken as fifteen cells in the examples that we have selected here. Efficient damping has been obtained with the values of $B = 25$ and $C = 2$ for equations (1.59). For an illustration of the PML mesh structure and of the numerical implementation of the wave-absorbing effects, a test was performed in a homogeneous infinite medium with a P-wave velocity of 2500 m/s, an S-wave

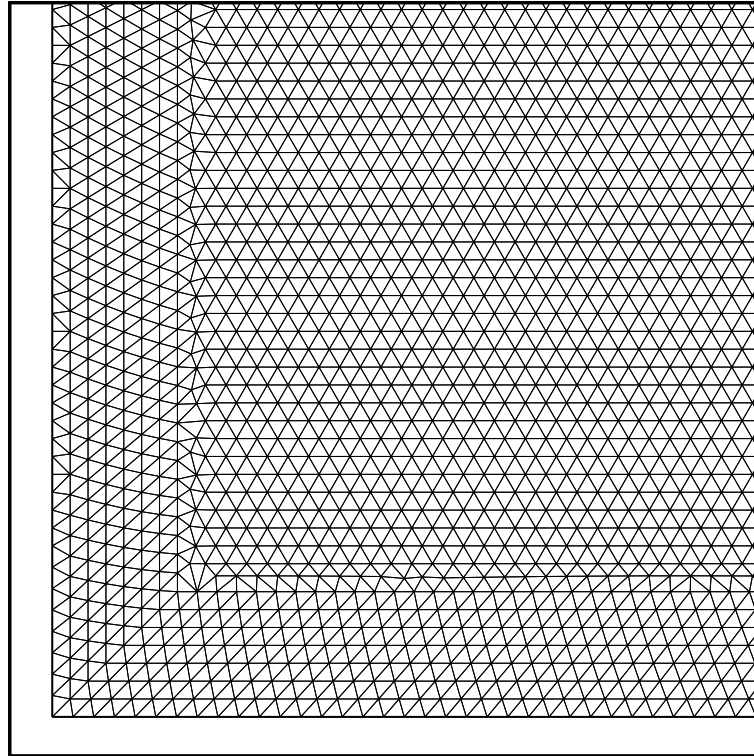


FIG. 1.7 – PML construction with multiple layer structures parallel to the Cartesian directions for the lower left quarter of the medium.

velocity of 1558 m/s, and a density of 1500 kg/m³ inside a finite grid. The simulation was performed at 4 Hz with a unitary explosive source. Figure 1.6 shows the frequency map solution for (a) a non-constrained mesh, and (b) a constrained mesh. Only the real part of the horizontal velocity is shown, although the behaviour is similar for the other components of the solutions. The mesh structure clearly has a significant influence on the PML efficiency.

1.6.5.3 Source implementation

The introduction of both point forces and point excitation stresses in the first-order system of equations (1.51) allows us to develop various point excitations : an impulsive force along the Cartesian directions or an explosive source can be applied easily using forces and stresses, respectively. A staggered behaviour over the mesh occurs in discrete equations, where the unknowns of one cell do not depend directly on the unknowns of its neighbours, but on the unknowns of the neighbours of the neighbouring cells. For particular mesh configurations, such as a regular equilateral mesh, if the source excitation is applied to one cell, simulations show that in the mesh only one cell out of two is excited. To avoiding this checker-board pattern, we spread the source over several cells using a Gaussian function. Of course, the numerical dispersion will express its properties based on twice the coarseness of the grid, as this is the price to be paid when considering centred fluxes. A similar observation can be made for the $O(\Delta x^2)$ rotated FD scheme. This has been the main reason for different authors moving to a staggered grid formulation. Figure 1.8 shows this pattern for an infinite homogeneous model

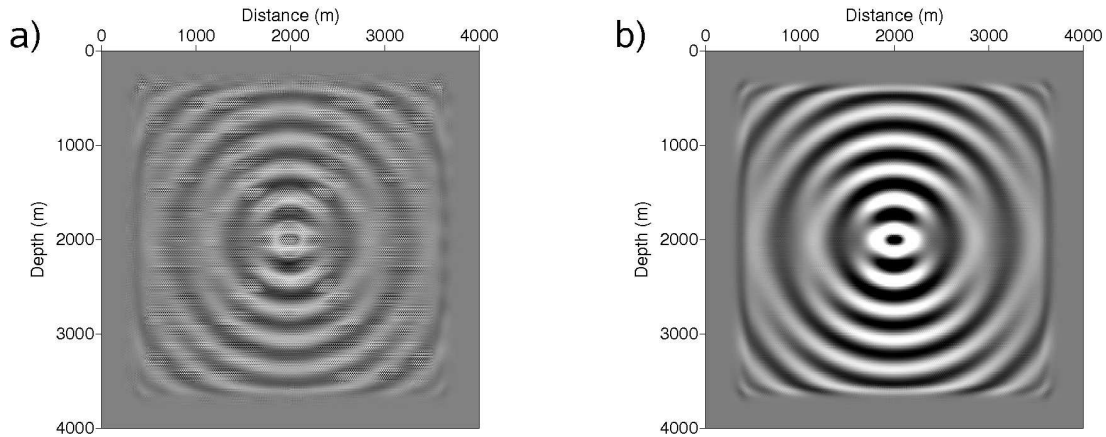


FIG. 1.8 – Frequency map solutions with the real parts of the horizontal velocity for a horizontal point force in an infinite medium. (a) Single cell excitation, and (b) smooth source excitation on several cells (29).

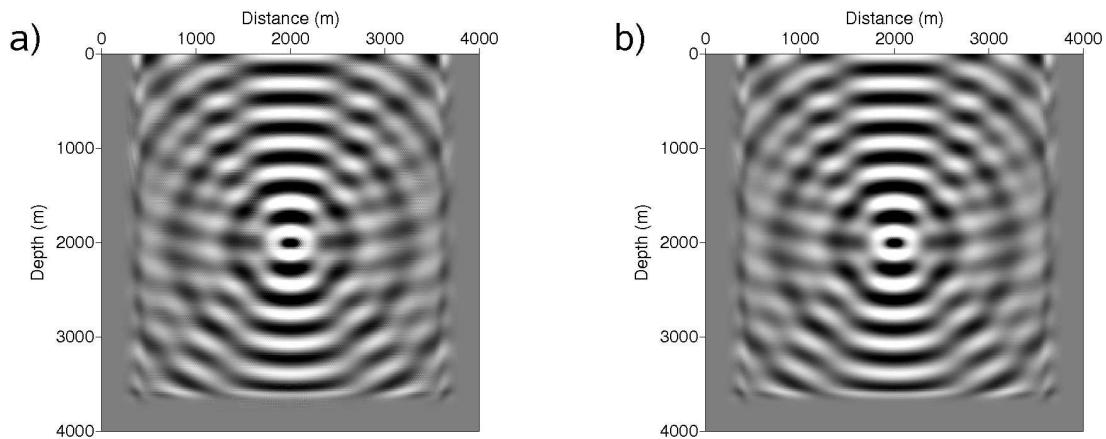


FIG. 1.9 – Frequency map solution with real parts of the horizontal velocity for a horizontal point force in a flat topography model. (a) Single cell excitation, and (b) smooth source excitation on several cells (29). The mesh pattern is less visible than for Figure 1.8, although it is still present in (a).

with a central horizontal force. A single cell is excited in Figure 1.8-(a), whereas a Gaussian source is applied to 29 cells in Figure 1.8-(b). It can also be noted that for the model with a free surface, the explicit formulation of such a boundary recouples the two decoupled submeshes, as shown in Figure 1.9 with the same source configuration as in the previous test. We still need to implement smooth source excitation for propagation inside the medium.

1.6.6 Numerical results

Several benchmarks will be presented to assess the accuracy of this new method, and we will focus more specifically on the influence of the mesh structure. Indeed, numerical tests with regular equilateral meshes and arbitrary unstructured meshes have shown the influence of the mesh regularity on solution accuracy : in spite of mesh refining, the solutions of the simulations do not converge to the true solution when considering unstructured meshes, whereas convergence is achieved with regular meshes. A sensitivity study on perturbed regular meshes has confirmed this behaviour. Accuracy problems in unstructured meshes typically involve travel-time shifts in seismograms. Dolean *et al.* (2006) observed the same behaviour of convergence dependency for P_0 interpolation when applied to Maxwell equations in the frequency domain, while convergence of a linear P_1 interpolation appears to be less dependent on the regularity of the mesh. Moreover, the theoretical evaluation of the second-order accuracy in space of the numerical scheme has been demonstrated for regular structured meshes (Remaki, 2000). In our different comparisons, we will specify when we consider regular equilateral meshes or unstructured meshes. In the following, we first compare the FV results with the analytical solutions for different canonical problems. The solutions to more complex models are then compared with FD solutions.

Analytical solutions and reference FD codes have been constructed in the time domain, and comparisons are performed in this domain. FV solutions are, of course, computed for several frequencies, spanning over the source wavelet bandwidth. To avoid the wrap-around effect in seismograms, complex frequencies (Mallick & Frazer, 1987) are used in frequency-domain simulations. An inverse fast Fourier transformation will provide us with seismograms in the time domain for the comparison of the solutions. The size of the cells is chosen with respect to the maximal frequency of the source bandwidth. The results for the unstructured meshes are finally illustrated.

1.6.6.1 Comparison with analytical solutions

An infinite homogeneous model with an explosive source : the acoustic case

Acoustic propagation can be modelled inside an infinite homogeneous elastic medium using an explosive source that generates only a P-wave pulse. The numerical grid is bounded by four absorbing layers on the edges. For the homogeneous case, the analytical solutions are build up for the Helmholtz equation. A comparison of the radial velocity with the analytical solution was performed for a 2500 m/s P-wave velocity medium with a regular equilateral mesh of size 1/10 of the P-wavelength (no S-wave is generated). Figure 1.10 shows the seismogram of the radial velocity at a receiver at a distance of 500 m from the source, and it illustrates the good agreement between the FV (crosses) and analytical (continuous line) seismograms. The tangential velocity is not strictly equal to zero, due to the smooth source and the numerical errors, although it still remains negligible.

Flat free-surface medium : the Garvin problem

The Garvin analytical solution deals with the propagation of elastic waves in a homogeneous half space with a flat free surface and an explosive source (Garvin, 1956). Comparisons with a

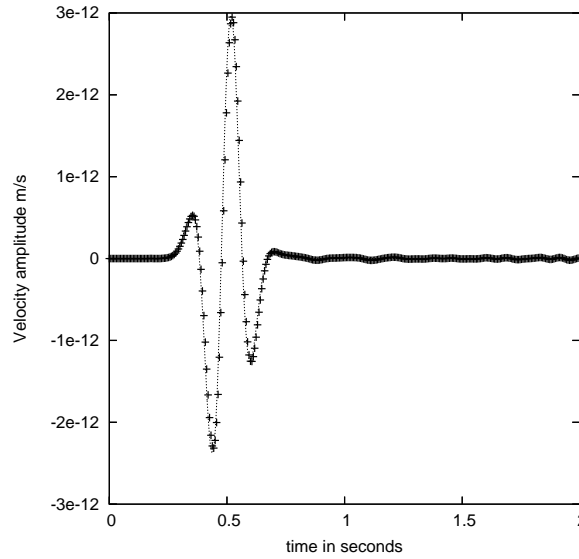


FIG. 1.10 – Comparison between analytical (continuous line) and numerical (crosses) seismograms for the acoustic case.

Garvin solution are quite a challenging problem because the FV method must model Rayleigh waves with good accuracy near to the free surface, a critical issue for an efficient full waveform inversion algorithm. A homogeneous medium with a P-wave velocity of 3464 m/s, an S-wave velocity of 2000 m/s, and a density of 2000 kg/m³ was considered. A Gaussian explosive source was considered at 150 m in depth, with 15 m of correlation length and with a line of receivers set on the surface with an offset of 200 m to 4000 m, and a space step of 200 m. A Ricker wavelet of central frequency 4 Hz was chosen. The simulation was performed over the source bandwidth, ranging from 0 Hz to 14 Hz. A regular equilateral mesh was taken, with a cell edge size of 15 m, which represents 1/10 of the S-wavelength.

The horizontal and vertical particle-velocity seismograms computed with the FV method are compared with the analytical ones in Figure 1.11. The direct and the Rayleigh waves are both modelled with very good levels of accuracy in shape and amplitude for the whole range of offsets for the two components. No time shift appears with the offsets, which confirms the small dispersion of the scheme when considering 10 cells per S-wavelength. Since our seismograms have relative simple shapes, we have considered the relative RMS residuals of the particle velocity field in the time domain for a quantitative estimation of the accuracy of the numerical solutions. The relative residual at one observer is the L2 norm of one component of the differential seismograms over the L2 norm of the same component of the reference seismogram. Moreover, we consider the total relative residual, named the TRR value, as the average of the relative residuals over the receivers. The TRR values are $4.1 \cdot 10^{-2}$ and $2.6 \cdot 10^{-2}$ for the horizontal and vertical components, respectively.

The two-layer model with a horizontal interface

An analytical solution can be constructed when considering two homogeneous half spaces. A compressional point source will act in the upper layer. The software code known as EX2DELEL

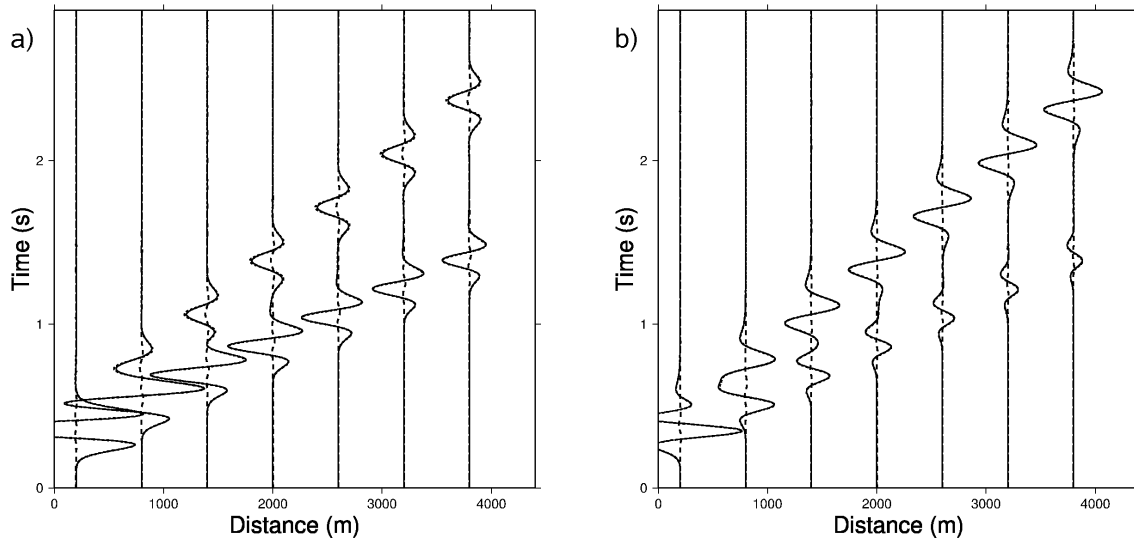


FIG. 1.11 – Seismograms for the Garvin problem. (a) Horizontal and (b) vertical components of the velocity at the receivers. The analytical solution is represented by dotted lines, FV by continuous lines, and the differences by dashed lines.

and provided by the Spice consortium (<http://www.spice-rtn.org>) was used to compute these solutions. Green's functions were first computed by the Cagniard-De Hoop technique, and a numerical convolution with the source wavelet gave the total response. The FV simulation was performed with the PML conditions on the four edges of the model for considering an infinite medium. The model dimensions were 12×2.5 km. The interface between the two layers was at a depth of 1150 m. The receivers were placed on a line at a depth of 280 m, with a space step of 200 m, from 0 m to 12000 m in distance, leading to an array of 201 sensors. The explosive source was placed at a distance of 500 m and at a depth of 370 m, with a correlation length of 30 m. The source wavelet was a Ricker wavelet with a central frequency of 4 Hz. Two tests were performed in regular equilateral meshes to evaluate the fluid-solid and solid-solid interfaces.

Fluid-solid interface

Fluid-solid interface modelling is quite challenging for marine acquisition or simulation in reservoirs, as discontinuities must exist at the boundary. It is important to verify the accuracy of the solution, and we considered a simple planar interface. The upper medium was considered as an acoustic one with a P-wave velocity of 1500 m/s, an S-wave velocity of 0 m/s, and a density of $1,000 \text{ kg/m}^3$, while the lower medium had a P-wave velocity of 3400 m/s, an S-wave velocity of 1963 m/s, and a density of 2400 kg/m^3 . A discretisation of 13 cells per P-wavelength in the fluid domain was chosen to keep the numerical dispersion negligible. Figure 1.12 shows the horizontal and vertical components of the particle velocity for the analytical (dotted), FV (continuous) and residual (dashed) solutions. There is good agreement between the analytical and FV solutions. The TRR values are $15.2 \cdot 10^{-2}$ for the horizontal component, and $17.9 \cdot 10^{-2}$ for the vertical one. However, we can see parasite reflections from PML zones exactly where the acoustic/elastic interface penetrates the PML layer. Figure 1.13 illustrates such reflections for a shorter offset model.

Solid-solid interface

The solid-solid interface test was performed with values of 2500 m/s, 1558 m/s and 1500

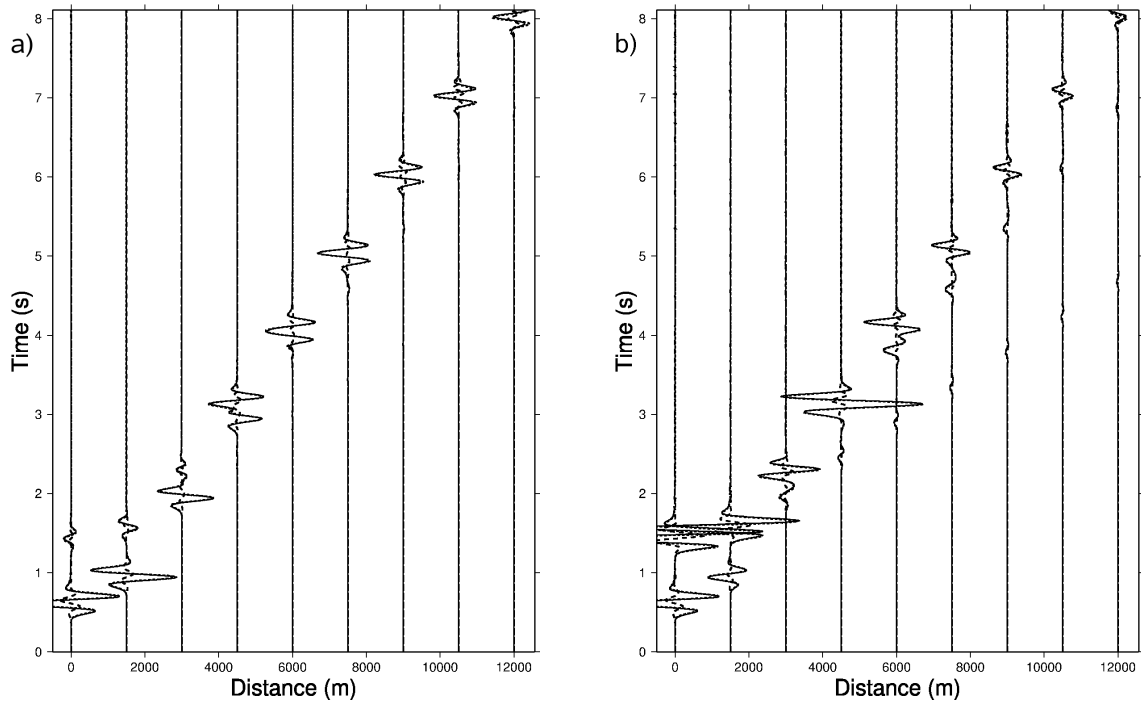


FIG. 1.12 – Seismograms computed in the two-layer model with a liquid-solid interface. (a) Horizontal and (b) vertical components of the velocity at the receivers. The analytical solution is represented by dotted lines, FV by continuous lines, and the differences by dashed lines.

kg/m³ for the P-wave velocity, the S-wave velocity, and the density for the upper half-space, and values of 3400 m/s, 1963 m/s and 2400 kg/m³, respectively, for the lower half-space. The FV seismograms were computed in a regular mesh, with ten cells per shear wavelength. There is good agreement between the analytical and FV seismograms (Figure 1.14). The TRR values are $10.0 \cdot 10^{-2}$ and $12.3 \cdot 10^{-2}$ for the horizontal and vertical components, respectively. Note the efficient absorption of the PML in the case of the elastic-elastic interface.

1.6.6.2 Comparison with numerical solutions

The FV method needs to be benchmarked according to the other numerical techniques that can be applied to a more complex medium, which should be more representative of the realistic applications of full waveform inversion. The FV solutions were validated against seismograms computed with a time-domain $O(\Delta x^2, \Delta t^2)$ rotated staggered-grid FD method (Saenger *et al.*, 2000), for three complex media : the corner-edge model ; a homogeneous hill model for considering complex topography ; and a realistic heterogeneous model corresponding to a subset of the so-called Marmousi II model.

The Corner-Edge Model

A synthetic model, known as the corner-edge model (Figure 1.15), is defined by a flat free surface and a corner with a sharp velocity contrast that introduces multiple reflections and

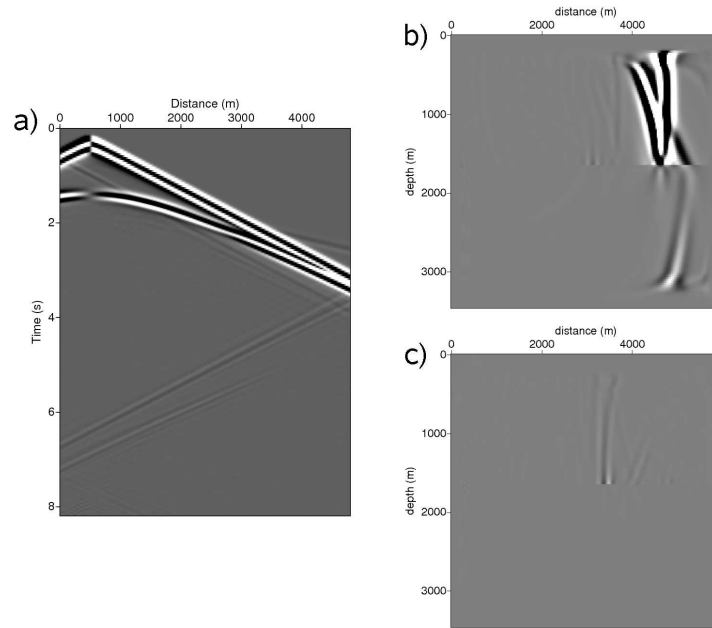


FIG. 1.13 – Seismograms at (a) receiver positions, and (b, c) snapshots for the horizontal component of the particle velocity for the two-layer problem with a liquid-solid interface and short offset geometry. Note the parasite reflections from the PML with a liquid-solid interface. The snapshots illustrate the incident wavefield at 2.8 s (b) and the reflected waves from the PML at 4.8 s (c).

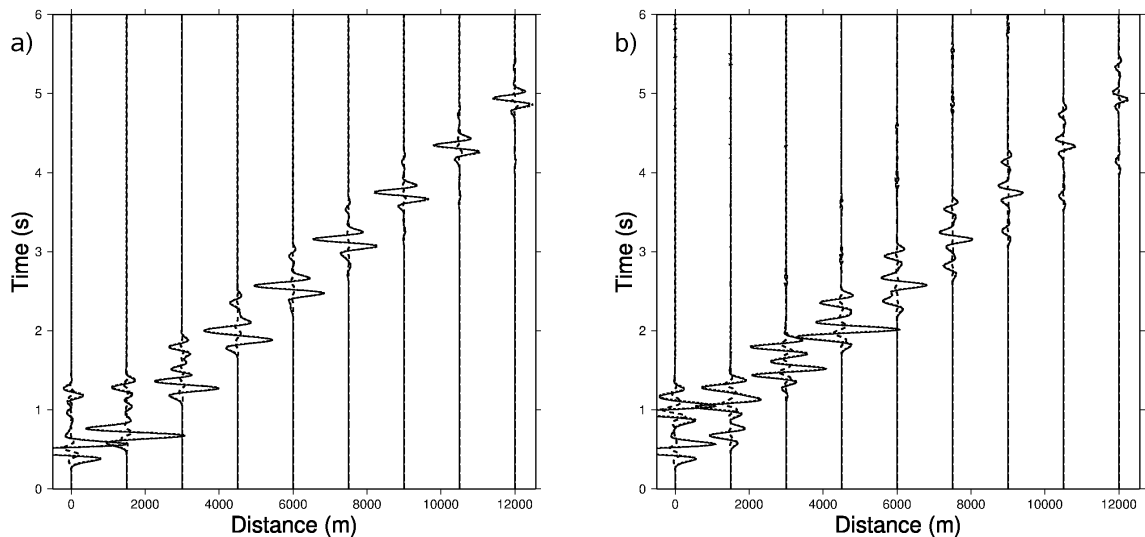


FIG. 1.14 – Seismograms computed for the two elastic-layer model. (a) Horizontal and (b) vertical components of the velocity at the receivers. The analytical solution is represented by dotted lines, FV by continuous lines, and the differences by dashed lines.

diffractions for both body and surface waves (Virieux, 1986a). The upper medium had a P-

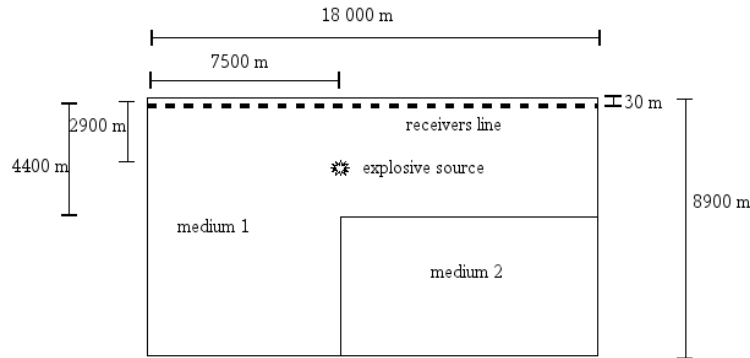


FIG. 1.15 – Geometry of the Corner-Edge model.

wave velocity of 6000 m/s, whereas the lower medium had a P-wave velocity of 9000 m/s. The S-wave velocity was computed from the P-wave velocity, with a ratio of $\sqrt{3}$. The model had a homogeneous density of 2500 kg/m³. The explosive source was located at (x=7500 m, z=2900 m), and it had a Ricker wavelet of central frequency of 4 Hz as the time function. The receiver line was placed below the topography at a depth of 30 m, with a receiver spacing of 50 m, from 0 m to 18000 m. An equilateral mesh was constructed with an edge length of 26.6 m, corresponding to the discretisation rule of ten cells per minimum shear wavelength. The agreement between the FV and FD solutions is illustrated in Figure 1.16. The TRR values are $10.1 \cdot 10^{-2}$ for the horizontal and $9.8 \cdot 10^{-2}$ for the vertical components of the velocity.

Complex topography Model

A homogeneous model with a hill-shaped topography was used to assess the accuracy of the FV method with a non-flat free surface. A homogeneous medium was used (4000 m/s and 2309 m/s, for P-wave and S-wave velocities, respectively, and 2,000 kg/m³ for density). An explosive source was set in the middle of the hill, 25 m below the topography, and it had a Ricker wavelet of central frequency of 4 Hz as a time dependence. The receiver line was located at a depth of 5 m below the free surface. Figure 1.17 shows the real part of a 10-Hz monochromatic wavefield for the horizontal velocity component. An equilateral mesh allows the modelling of the topography by straight lines (Figure 1.18), without the stair-case description of the FD methods. This description is not perfect, as it should be with an unstructured mesh, but numerical simulations have shown quite accurate results with 15 cells per shear wavelength with this topography, whereas the second-order rotated staggered-grid FD stencil requires 60 points. The FV and FD seismograms computed with these two above-mentioned discretisation rules (15 and 60 cells per shear-wavelength, respectively) are compared in Figure 1.19, and they show good agreement. The TRR values are $11.7 \cdot 10^{-2}$ and $12.4 \cdot 10^{-2}$ for the horizontal and vertical components of the velocity, respectively. The surface waves are well modelled, and no numerical dispersion occurs. Simulations with a finer mesh led to comparable seismograms for both the FV and the FD methods.

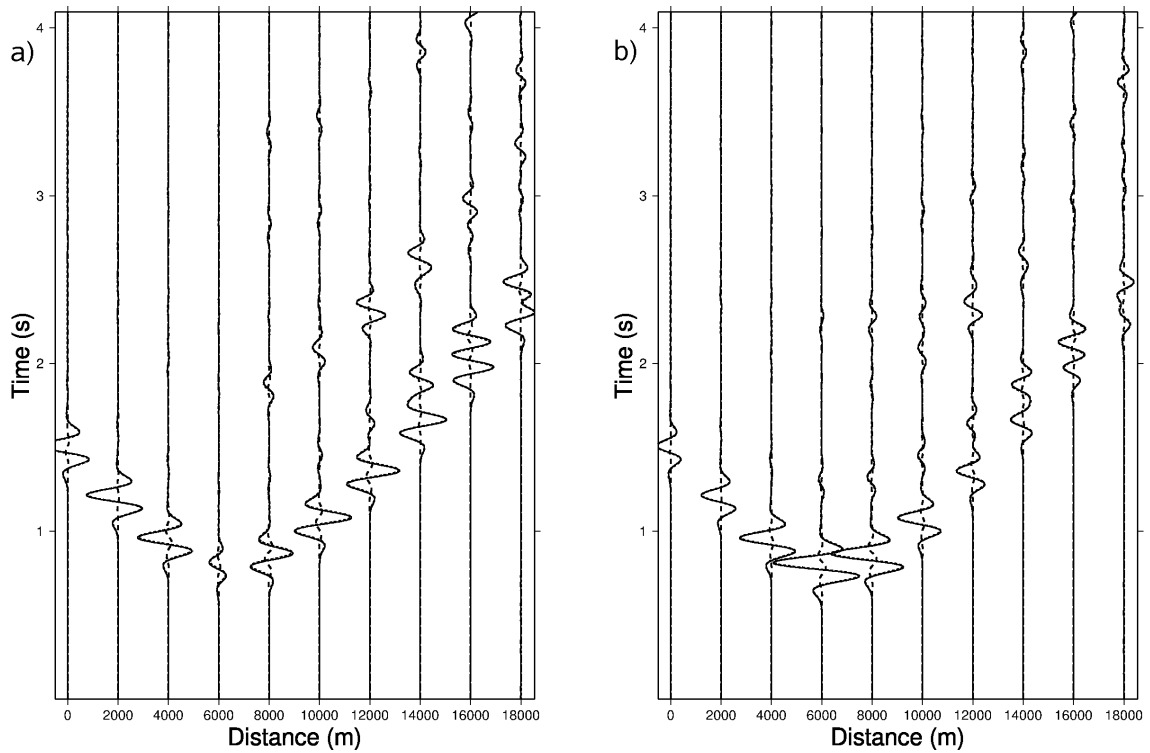


FIG. 1.16 – Seismograms computed for the Corner-Edge model. (a) Horizontal and (b) vertical components of the velocity at the receivers. The reference solution computed with the FD method is represented by dotted lines, FV by continuous lines, and the residuals by dashed lines. Both of these solutions are very similar across the entire time window.

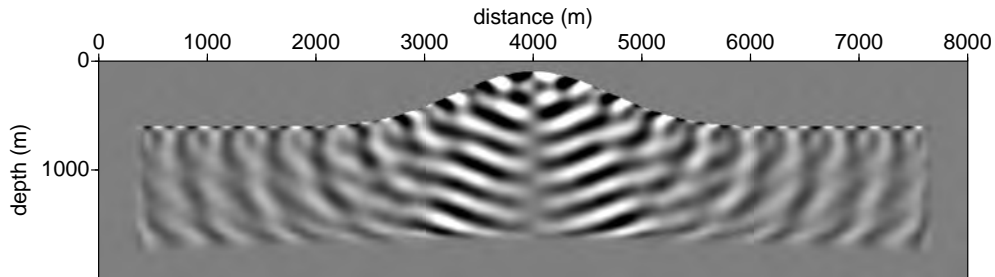


FIG. 1.17 – Monochromatic wavefield of the horizontal velocity for the hill model. The real part of the wavefield is illustrated for a 10-Hz simulation.

Realistic model : a subset of the Marmousi II model

The Marmousi II synthetic model represents a complex elastic medium, which makes it suitable for testing the FV method we propose here. A limited target of the model with multiple interfaces was chosen to limit the core memory requested by the frequency-domain formulation

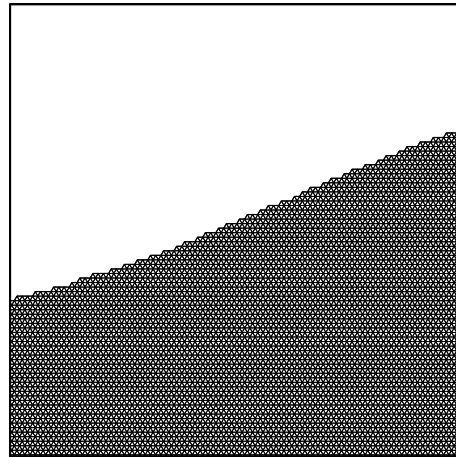


FIG. 1.18 – Description of a complex topography with regular equilateral triangles.

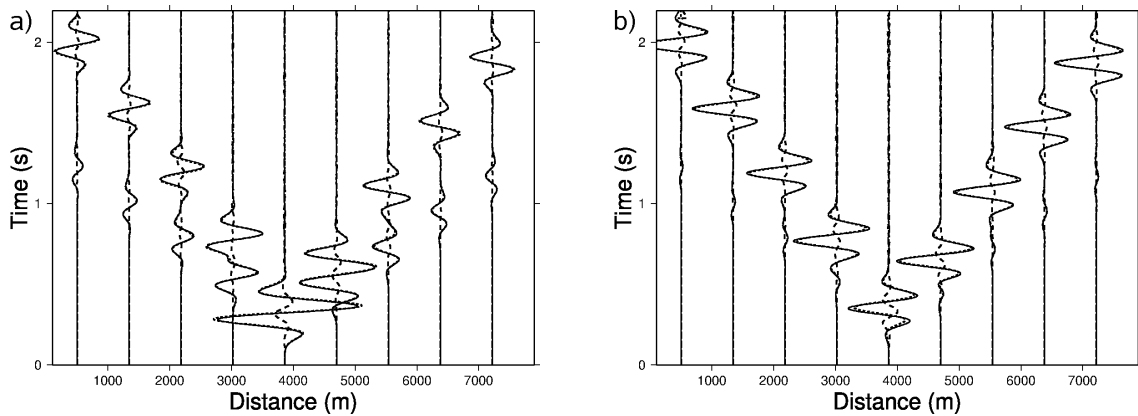


FIG. 1.19 – Seismograms computed for the hill model. (a) Horizontal and (b) vertical components of the velocity. The reference solution computed with the FD method is represented by dotted lines, the FV solution by continuous lines, and the difference between the two solutions with dashed lines.

used for building the time-domain seismograms. This target, the dimensions of which were $5000 \text{ m} \times 2000 \text{ m}$ ($6000 \text{ m} \times 2500 \text{ m}$ with the PML layers), is illustrated in Figure 1.20 for the P-wave velocities. An explosive source was located at $(x=1000 \text{ m}, z=100 \text{ m})$, and a Ricker wavelet of central frequency of 4 Hz was considered. The receiver line was set at a depth of 25 m below the topography. The edges of the regular triangular mesh had lengths of 7.1 m, corresponding to 10 cells per shear wavelength. Figure 1.21 illustrates the seismograms at the receivers. Comparisons with the FD method show quite similar results. The small differences that occur for the horizontal velocity can be attributed to the model description, which is slightly different for the square and the triangle parametrisation. The FV simulations in finer triangular meshes led to similar seismograms, hence providing additional validation of the discretisation rule of ten cells per S-wavelength for heterogeneous media, which is quite encouraging for future work.

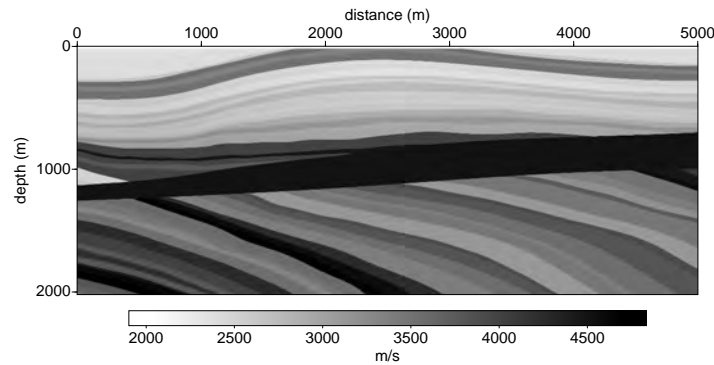


FIG. 1.20 – P-wave velocity distribution of the realistic model taken from the Marmousi II model.

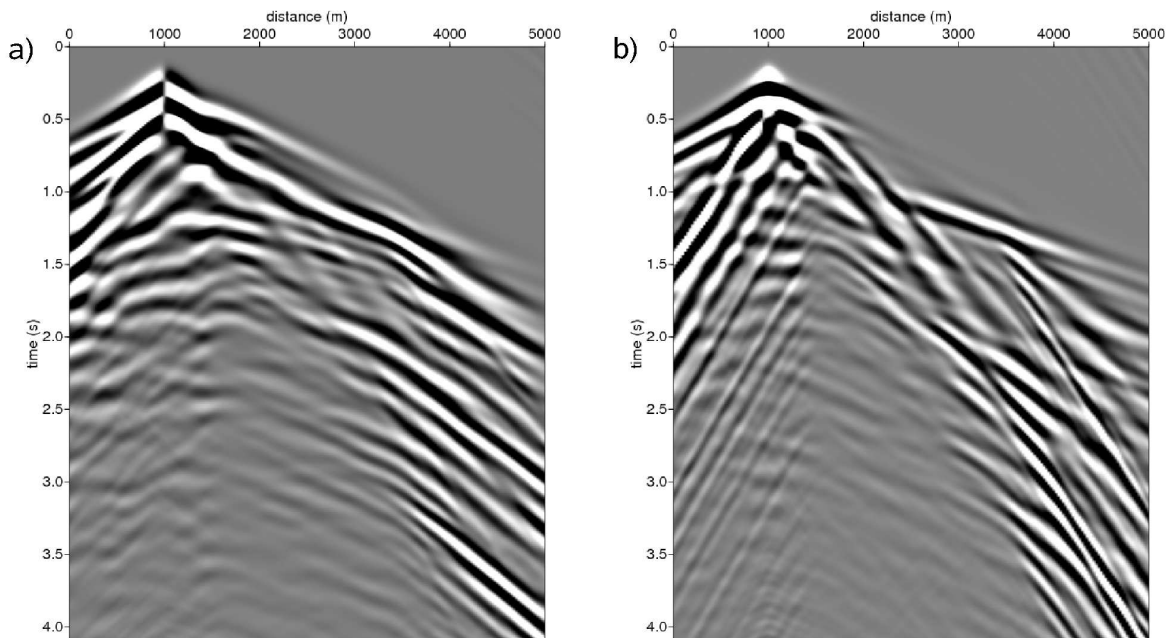


FIG. 1.21 – Seismograms computed in the target of the Marmousi II model with the FV method, using a regular equilateral mesh. (a) Horizontal and (b) vertical particle velocities at the receiver positions.

1.6.6.3 Numerical tests with unstructured meshes

This FV method is now analysed in unstructured meshes. The P_0 interpolation should provide solutions with a given level of accuracy in such meshes. Previous hill models and Marmousi II models will be considered. The FV solutions computed for an equilateral mesh will be used as the reference solutions. Unstructured meshes allow very precise modelling of the free surface when a complex topography is considered. No constraints on triangle angles have been applied

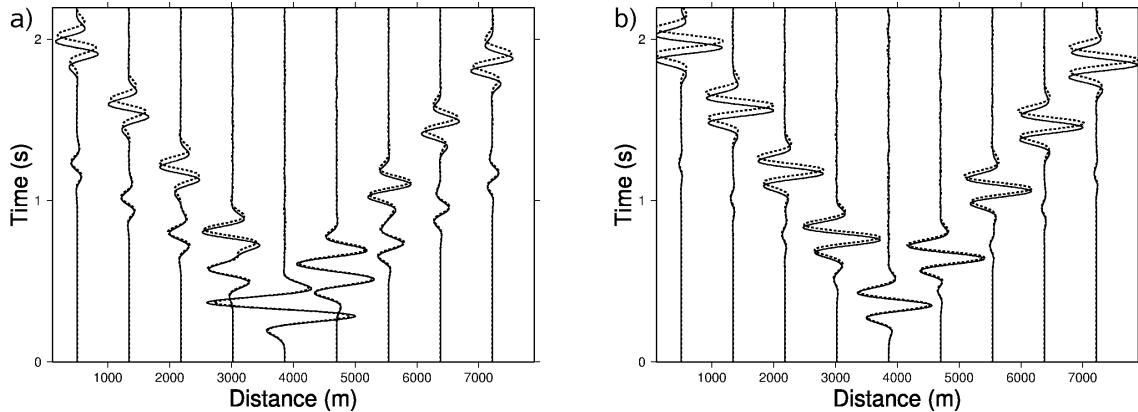


FIG. 1.22 – Seismograms computed in the hill model using equilateral and unstructured meshes. (a) Horizontal and (b) vertical components of the particle velocity at the receivers. The reference solution computed with FV in regular equilateral mesh is plotted with dotted lines, and the solution in an unstructured mesh with continuous lines. Note the advance of the solution computed in the unstructured mesh, as it increases with propagation time.

for either model. Moreover, the triangle sizes can be adapted to the propagated wavelengths locally, to minimise the number of cells, and therefore the number of unknowns to be solved in the linear system, a very appealing feature when performing factorisation of the impedance matrix. The hill model simulation was performed with an unstructured mesh by considering the discretisation rule of fifteen cells per minimum shear wavelength. The comparison between seismograms computed for an equilateral structured and for unstructured meshes is shown in Figure 1.22. Of course, there is a good match of the amplitudes, whereas there is a negative time shift of the phases that increases with the propagation time in seismograms computed in the unstructured mesh. The Marmousi II model simulation was performed in an unstructured mesh adapted to the local shear-wave velocity, with at least ten cells per shear wavelength. The seismograms look similar to those computed for an equilateral mesh (compare Figures 1.21 and 1.23). However, a direct comparison between the seismograms computed in the equilateral and unstructured meshes (Figure 1.24) shows the same slight negative time delay as is seen for the hill model. The simulations in finer unstructured meshes for both the hill and Marmousi II models have not shown better convergence of the numerical solution, as expected for the P_0 interpolation for unstructured meshes.

1.6.7 Numerical cost of method

Computational costs are always a difficult question, and we will focus here on a comparison between the FV and FD methods for the same order of accuracy in the frequency domain. We will focus on the specific example of the Marmousi II model (Figure 1.20). The source was an explosion. The modelled frequency was 13 Hz. We can illustrate the CPU time and memory requirements of the FV method with the second-order parsimonious rotated frequency-domain FD method (Gelis *et al.*, 2007). Both of these methods make use of the direct solver MUMPS (MUMPS-team, 2007), which performs the resolution of the linear system by LU decomposition of the sparse matrix through a multifrontal approach. The medium is discretised with 10 cells per minimum S-wavelength for the FV approach in a regular equilateral mesh, with 10 cells

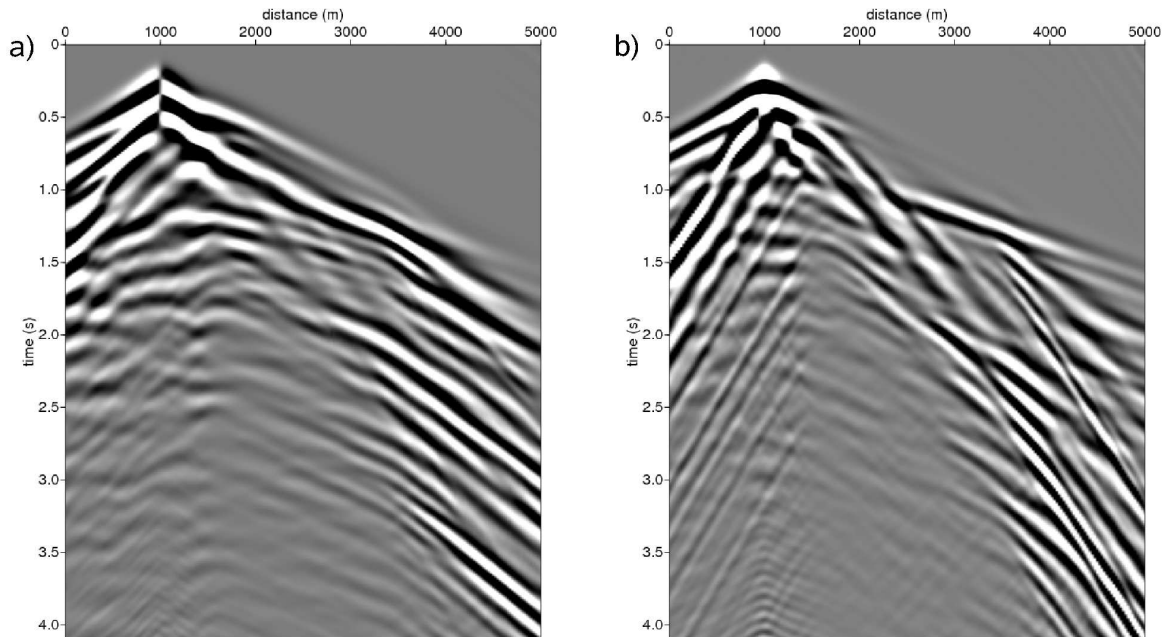


FIG. 1.23 – Seismograms computed in the target of the Marmousi II model using an unstructured mesh. (a) Horizontal and (b) vertical components of the velocity at the receivers. These seismograms can be compared with that computed in the same model parametrised with an equilateral mesh (Figure 1.21).

per local S-wavelength for the FV method in an unstructured mesh, and with 28 points for the FD method in order to have an acceptable numerical dispersion.

Table 1.1 illustrates the requirements of both of these methods for sequential execution on a single processor. The coarser parameterisation of the FV naturally leads to less unknowns to be computed and a less expensive estimation in terms of CPU time and core memory for all of the MUMPS phases : a factor of 2.5 can be noted in this example, between the FV method in a regular mesh and the FD method. Moreover, an adaptive unstructured mesh allows for a significant decrease in the numerical resources when a heterogeneous medium is considered, although we must be aware of the approximate precision of the solution when performing the seismic inversion procedure. However, the mesh description of the medium introduces several additional table build-ups and manipulations for the matrix construction, which are more time consuming than the simple implicit construction of the FD techniques on a regular grid. Extra CPU time costs of the FV methods will remain small for the different MUMPS numerical procedures, and will occur only once when considering a full waveform inversion algorithm ; they will not hamper the benefit of using a coarser grid for the FV method.

1.6.8 Discussion and conclusion

A FV method dedicated to full waveform inversion has been formulated in the space-frequency domain for 2D P-SV wave propagation. By using the parsimonious approach, only

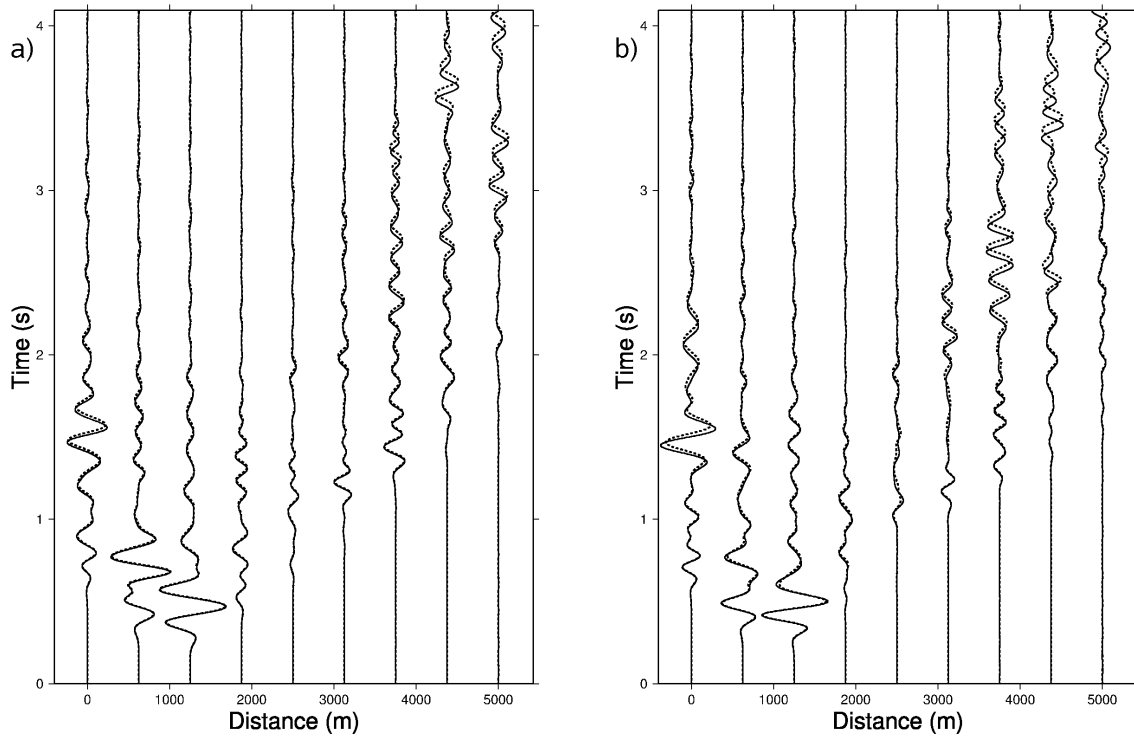


FIG. 1.24 – Comparison between the FV seismograms computed in the target of the Marmousi II model with equilateral and unstructured meshes. (a) Horizontal and (b) vertical components of the velocity at the receivers. The reference solution computed with FV in a regular equilateral mesh is represented by dotted lines, and the solution in an unstructured mesh by continuous lines.

TAB. 1.1 – Overview of the numerical cost for the FV method in regular and unstructured meshes and the $O(\Delta x^2)$ FD method in the frequency domain for a realistic model at 13 Hz.

Numerical method	regular FV	unstructured FV	FD
Number of unknowns to solve	1 421 364	549 638	4 850 020
Time to prepare data for matrix building (s)	54.4	40.1	1.0
Time for matrix building (s)	1.6	0.80	10.9
Time for factorisation (s)	272.3	79.5	999.4
Memory use for factorisation (Mb)	3448	1333	12061
Time for resolution of 1 shot (s)	3.8	1.7	13.5

particle velocity unknowns are used in the build-up of the impedance matrix, which incorporates various medium properties, including a free surface and possible liquid-solid interfaces. Comparisons between numerical solutions computed with analytic and numerical reference solutions for canonical and realistic configurations have shown that a structured equilateral mesh provides accurate results for a discretisation of 10 cells per shear wavelength, even if the topography and surface waves are considered. A complex topography should require a finer description of 15 cells, coarser than $O(\Delta x^2)$ FD due to the triangular meshing. Unstructured meshes are easily taken into account in the P_0 FV formulation, although they suffer from a

lack of precision even when fine meshes are considered. The CPU/memory requirements are naturally less expensive than for $O(\Delta x^2)$ FD, in spite of the complex table manipulations due to the mesh description of medium. Finally, the FV method with regular meshes appears to be very efficient when compared with $O(\Delta x^2)$ FD methods, especially when a realistic topography is considered. Considering unstructured meshes allows a significant decrease in the numerical resources; however, with lower accuracy of the wavefield approximation. Future work will focus on the accuracy needed in forward modelling for the application of full waveform inversion. We will investigate further the usefulness of unstructured meshes and the correlated accuracy for such an imaging strategy. As an alternative, moving to higher order for the interpolation in the discontinuous Galerkin approach will be a possibility. By considering the P_1 interpolation, we may find the best compromise between accuracy for the wavefield estimation and computational efficiency required by the inversion.

Acknowledgments

This study was funded by the SEISCOPE consortium <http://seiscope.unice.fr>, which is sponsored by BP, CGG-VERITAS, EXXON-MOBIL, SHELL and TOTAL, and by Agence Nationale de la Recherche (ANR) under project ANR-05-NT05-2_42427. The LU factorisation of the impedance matrix was performed with the MUMPS package, available on <http://graal.ens-lyon.fr/MUMPS/index.html>. The mesh generation was performed with the help of TRIANGLE, available on <http://www.cs.cmu.edu/~quake/triangle.html>. The Garvin problem analytical solutions were computed with code provided by U. Iturrarán and F. J. Sánchez-Sesma. The two-layer analytic solutions were computed with the EX2DELEL code, available on <http://www.spicertrn.org/library/software/EX2DELEL>. Many thanks go to Mondher Ben Jemaa, Nathalie Glinsky, Stéphane Lantéri and Serge Piperno of the CAIMAN project for interesting and stimulating discussions on the DG P_k approach. Access to the high-performance computing facilities of the MESOCENTRE SIGAMM computer centre provided the required computer resources, and we gratefully acknowledge both this facility and the support of the staff, and in particular of A. Miniussi. Reviews by P. Moczo and J. Robertsson significantly improved the first version of this manuscript.

1.6.9 Appendix A : Equivalence FV P_0 and $O(\Delta x^2)$ FD in regular square grids

The proposed FV method can be developed in any polygonal mesh structure. In a mesh of regular squares, the FV formulation is equivalent to the second-order FD Cartesian stencil, as we now show. Let us start with the first-order discrete system in velocity and stress :

$$\begin{aligned}\omega^2 \vec{V}_k &= \frac{\iota\omega}{A_k \rho_k} \sum_{j \in \partial K_k} l_{kj} (n_{kjx} N_x s_{x_j} + n_{kjz} N_z s_{z_j}) \frac{\vec{T}_j}{2} + \iota\omega \vec{F}_k \\ \omega^2 \vec{T}_k &= \frac{\iota\omega \Lambda_k^{-1}}{A_k} \sum_{j \in \partial K_k} l_{kj} (n_{kjx} M_x s'_{x_j} + n_{kjz} M_z s'_{z_j}) \frac{\vec{V}_j}{2} + \omega^2 \vec{T}_k^0,\end{aligned}\quad (1.60)$$

where the cell numbering is denoted by the global index k . We modify the numbering of the cells towards a double index system (i,j) that is more suitable for a mesh of squares (Figure 22) where the cell (i,j) has four neighbours located in positions $(i,j-1)$, $(i-1,j)$, $(i+1,j)$

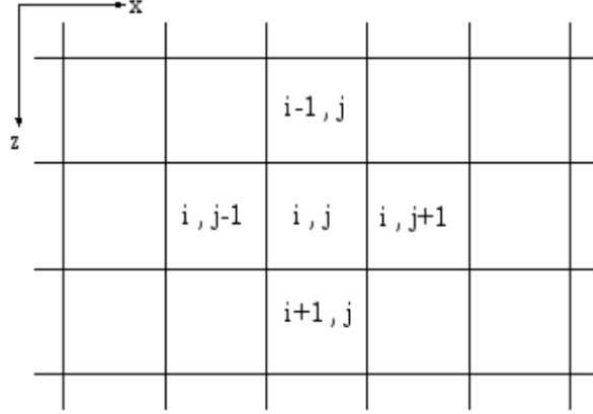


FIG. 1.25 – The regular Cartesian mesh for the development of the FV method equivalent to $O(\Delta x^2)$ order FD.

and $(i, j + 1)$. The quantity h denotes the constant step length of the grid. The system (1.60) can now be written as a vectorial expression with two subscript indices :

$$\begin{aligned}
 \omega^2 \overrightarrow{V}_{i,j} &= \frac{\omega}{h^2 \rho_{i,j}} \left(-h N_x s_{x_{i,j-1}} \frac{\overrightarrow{T}_{i,j-1}}{2} + h N_x s_{x_{i,j+1}} \frac{\overrightarrow{T}_{i,j+1}}{2} - h N_z s_{z_{i-1,j}} \frac{\overrightarrow{T}_{i-1,j}}{2} \right. \\
 &\quad \left. + h N_z s_{z_{i+1,j}} \frac{\overrightarrow{T}_{i+1,j}}{2} \right) + \omega \overrightarrow{F}_{i,j} \\
 \omega^2 \overrightarrow{T}_{i,j} &= \frac{\omega \Lambda_{i,j}^{-1}}{h^2} \left(-h M_x s'_{x_{i,j-1}} \frac{\overrightarrow{V}_{i,j-1}}{2} + h M_x s'_{x_{i,j+1}} \frac{\overrightarrow{V}_{i,j+1}}{2} - h M_z s'_{z_{i-1,j}} \frac{\overrightarrow{V}_{i-1,j}}{2} \right. \\
 &\quad \left. + h M_z s'_{z_{i+1,j}} \frac{\overrightarrow{V}_{i+1,j}}{2} \right) + \omega^2 \overrightarrow{T}_{i,j}^0,
 \end{aligned} \tag{1.61}$$

which gives the following system of equations for scalar quantities :

$$\begin{aligned}
 \omega^2 V_{x_{i,j}} &= \frac{\omega}{2h\rho_{i,j}} \left(-s_{x_{i,j-1}} T_{1_{i,j-1}} - s_{x_{i,j-1}} T_{2_{i,j-1}} + s_{x_{i,j+1}} T_{1_{i,j+1}} + s_{x_{i,j+1}} T_{2_{i,j+1}} \right. \\
 &\quad \left. - s_{z_{i-1,j}} T_{3_{i-1,j}} + s_{z_{i+1,j}} T_{3_{i+1,j}} \right) + \omega F_{x_{i,j}} \\
 \omega^2 V_{z_{i,j}} &= \frac{\omega}{2h\rho_{i,j}} \left(-s_{x_{i,j-1}} T_{3_{i,j-1}} + s_{x_{i,j+1}} T_{3_{i,j+1}} - s_{z_{i-1,j}} T_{1_{i-1,j}} + s_{z_{i-1,j}} T_{2_{i-1,j}} \right. \\
 &\quad \left. + s_{z_{i+1,j}} T_{1_{i+1,j}} - s_{z_{i+1,j}} T_{2_{i+1,j}} \right) + \omega F_{z_{i,j}} \\
 \omega^2 T_{1_{i,j}} &= \frac{\omega(\lambda_{i,j} + \mu_{i,j})}{2h} \left(-s'_{x_{i,j-1}} V_{x_{i,j-1}} + s'_{x_{i,j+1}} V_{x_{i,j+1}} - s'_{z_{i-1,j}} V_{z_{i-1,j}} + s'_{z_{i+1,j}} V_{z_{i+1,j}} \right) + \omega^2 T_{1_{i,j}}^0 \\
 \omega^2 T_{2_{i,j}} &= \frac{\omega\mu_{i,j}}{2h} \left(-s'_{x_{i,j-1}} V_{x_{i,j-1}} + s'_{x_{i,j+1}} V_{x_{i,j+1}} + s'_{z_{i-1,j}} V_{z_{i-1,j}} - s'_{z_{i+1,j}} V_{z_{i+1,j}} \right) + \omega^2 T_{2_{i,j}}^0 \\
 \omega^2 T_{3_{i,j}} &= \frac{\omega\mu_{i,j}}{2h} \left(-s'_{x_{i,j-1}} V_{z_{i,j-1}} + s'_{x_{i,j+1}} V_{z_{i,j+1}} - s'_{z_{i-1,j}} V_{x_{i-1,j}} + s'_{z_{i+1,j}} V_{x_{i+1,j}} \right) + \omega^2 T_{3_{i,j}}^0,
 \end{aligned} \tag{1.62}$$

which are equivalent to the system of discrete equations of a full second-order Cartesian FD stencil inside which two staggered sub-systems propagate the seismic wavefield in an uncoupled way.

1.6.10 Appendix B : Finite Volume Development

The FV method is applied to first-order elastodynamic systems described by equations (1.52). By introducing the projector matrices M_x , N_x , M_z and N_z defined by :

$$\begin{aligned} N_x &= M_x^t = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \\ N_z &= M_z^t = \begin{bmatrix} 0 & 0 & 1 \\ 1 & -1 & 0 \end{bmatrix} \end{aligned} \quad (1.63)$$

and the diagonal matrix $\Lambda = \text{diag}(\frac{1}{\lambda+\mu}, \frac{1}{\mu}, \frac{1}{\mu})$, system (1.52) can be written in a vectorial form with a divergence expression :

$$\begin{aligned} -\omega\rho\vec{V} &= \overrightarrow{\text{div}(s_x N_x \vec{T}, s_z N_z \vec{T})} - \frac{\partial s_x N_x \vec{T}}{\partial x} - \frac{\partial s_z N_z \vec{T}}{\partial z} + \rho\vec{F} \\ -\omega\Lambda\vec{T} &= \overrightarrow{\text{div}(s'_x M_x \vec{V}, s'_z M_z \vec{V})} - \frac{\partial s'_x M_x \vec{V}}{\partial x} - \frac{\partial s'_z M_z \vec{V}}{\partial z} - \omega\Lambda\vec{T}^0 \end{aligned} \quad (1.64)$$

We then introduce vectorial forms : $\vec{G}(\vec{T}) = (s_x N_x \vec{T}, s_z N_z \vec{T})$ and $\vec{H}(\vec{V}) = (s'_x M_x \vec{V}, s'_z M_z \vec{V})$, and apply a surface integration over a control cell identified by the index i .

$$\begin{aligned} \int_{K_i} -\omega\rho\vec{V}dS &= \int_{K_i} \overrightarrow{\text{div}(\vec{G}(\vec{T}))}dS - \int_{K_i} \frac{\partial s_x N_x \vec{T}}{\partial x}dS - \int_{K_i} \frac{\partial s_z N_z \vec{T}}{\partial z}dS + \int_{K_i} \rho\vec{F}dS \\ \int_{K_i} -\omega\Lambda\vec{T}dS &= \int_{K_i} \overrightarrow{\text{div}(\vec{H}(\vec{V}))}dS - \int_{K_i} \frac{\partial s'_x M_x \vec{V}}{\partial x}dS - \int_{K_i} \frac{\partial s'_z M_z \vec{V}}{\partial z}dS - \int_{K_i} \omega\Lambda\vec{T}^0dS \end{aligned} \quad (1.65)$$

Through Green's theorem, the surface integration of divergence terms allows flux integrals to appear :

$$\begin{aligned} \int_{K_i} -\omega\rho\vec{V}dS &= \int_{\partial K_i} \vec{G}(\vec{T})\vec{n}dL - \int_{K_i} \frac{\partial s_x N_x \vec{T}}{\partial x}dS - \int_{K_i} \frac{\partial s_z N_z \vec{T}}{\partial z}dS + \int_{K_i} \rho\vec{F}dS \\ \int_{K_i} -\omega\Lambda\vec{T}dS &= \int_{\partial K_i} \vec{H}(\vec{V})\vec{n}dL - \int_{K_i} \frac{\partial s'_x M_x \vec{V}}{\partial x}dS - \int_{K_i} \frac{\partial s'_z M_z \vec{V}}{\partial z}dS - \int_{K_i} \omega\Lambda\vec{T}^0dS \end{aligned} \quad (1.66)$$

where ∂K_i is the boundaries of cell K_i and \vec{n} is the external normal vector of ∂K_i .

We end up with the already explained discrete system written in a vectorial form. The partial derivatives of all of the PML functions are cancelled by the P_0 assumption :

$$\begin{aligned} -\omega A_i \rho_i \vec{V}_i &= \sum_{j \in \partial K_i} l_{ij} G_{ij} + A_i \rho_i \vec{F}_i \\ -\omega A_i \Lambda_i \vec{T}_i &= \sum_{j \in \partial K_i} l_{ij} H_{ij} - \omega A_i \Lambda_i \vec{T}_i^0 \end{aligned} \quad (1.67)$$

The integration of property (1.55) and flux formulation (1.54) into the discrete system (1.67) gives the first-order discrete system where the parsimonious strategy can be applied :

$$\begin{aligned} \omega^2 \vec{V}_i &= \frac{\omega}{A_i \rho_i} \sum_{j \in \partial K_i} l_{ij} (n_{ijx} N_x s_{xj} + n_{ijz} N_z s_{zj}) \frac{\vec{T}_j}{2} + \omega \vec{F}_i \\ \omega^2 \vec{T}_i &= \frac{\omega \Lambda_i^{-1}}{A_i} \sum_{j \in \partial K_i} l_{ij} (n_{ijx} M_x s'_{xj} + n_{ijz} M_z s'_{zj}) \frac{\vec{V}_j}{2} + \omega^2 \vec{T}_i^0 \end{aligned} \quad (1.68)$$

1.6.11 Conclusion

L'interpolation de bas ordre P_0 se révèle être une méthode de choix pour palier un certain nombre de difficultés occasionnées par les méthodes classiques de type DF, tout en étant très proche. L'utilisation de maillages réguliers triangulaires de type équilatéral permet entre autre de représenter les topographies complexes par des successions de triangles réguliers, engendrant moins de diffractions parasites que la représentation en marches d'escalier des DF, et permettant ainsi une discrétisation plus grossière du milieu.

Cependant, cette formulation ne permet pas d'appréhender facilement toutes les applications possibles d'imagerie. La première limitation est liée au type de maillage régulier équilatéral, pour garder une consistance du schéma et une précision acceptable des solutions. En effet, la taille des cellules est uniforme dans tout le domaine de calcul et est contrainte par la plus petite longueur d'onde du milieu, engendrant un suréchantillonnage spatial pour des milieux fortement hétérogènes avec des larges gammes de vitesses de propagations. De même, la modélisation de topographies complexes est limitée à des formes relativement lisses. Enfin, le contact liquide/solide pose certains problèmes, rendant difficile la modélisation en milieu marin.

1.7 Extension aux ordres supérieurs P_1 et P_2

1.7.1 Introduction

Les interpolations d'ordre supérieurs à P_0 sont indispensables pour l'utilisation de maillages triangulaires non-structurés, clé de la représentation fine des topographies et bathymétries complexes, et de l'adaptation aux propriétés locales du milieu afin d'éviter le suréchantillonnage spatial.

Les schémas GD tirent généralement bénéfice des ordres élevés d'interpolation, pour une discrétisation grossière du milieu tout en obtenant de grands niveaux de précision dans les solutions (Käser & Dumbser, 2006). Dans la formulation GD telle que développée dans la partie 1.5.3, les propriétés physiques du milieu sont considérées constantes par cellule. Pour l'imagerie de milieux très hétérogènes par inversion de formes d'ondes, le pouvoir de résolution spatiale optimal est de l'ordre de $\lambda/2$ où λ est la longueur d'onde locale (*cf.* partie 2.2.2). Pour l'inversion, une discrétisation de l'ordre de $\lambda/4$ est donc une représentation optimale du milieu. En basant l'inversion sur le même maillage que le problème direct de modélisation, la taille des cellules du maillage doit donc être le plus cohérent possible avec la résolution attendue de l'imagerie, et la taille nécessaire à une solution assez précise en GD, tout en évitant tout suréchantillonnage spatial coûteux en temps de calcul et en mémoire. Käser & Dumbser (2006) montrent que même s'ils fournissent de grands niveaux de précision dans les solutions, les ordres d'interpolation supérieurs à 2 sont relativement coûteux en temps de calcul s'ils sont appliqués pour des maillages aussi fins que ceux requis par l'imagerie, ces ordres étant optimaux pour des discrétisations grossières de l'espace. C'est pour cela que nous nous intéresserons uniquement aux ordres P_1 et P_2 , donnant une représentation linéaire et quadratique des champs au sein de chaque cellule.

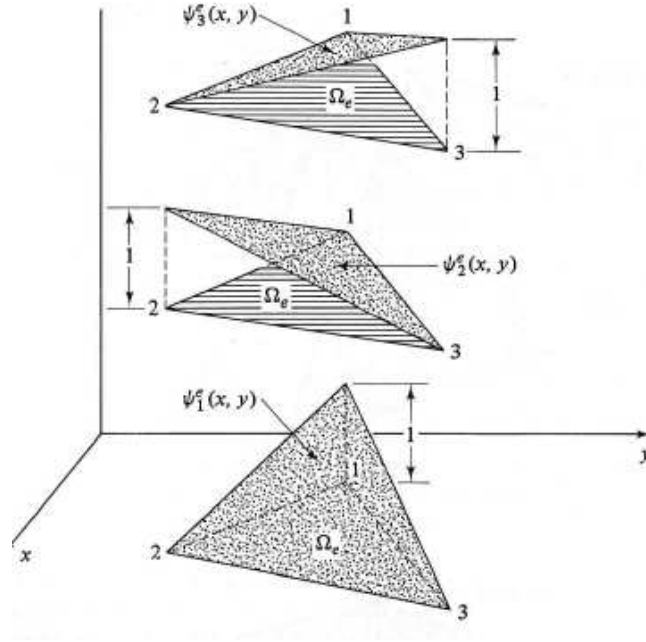


FIG. 1.26 – Schémas des fonctions de base P_1 basé sur les coordonnées barycentriques du triangle (d'après Becker *et al.*, 2001).

1.7.2 Formulation

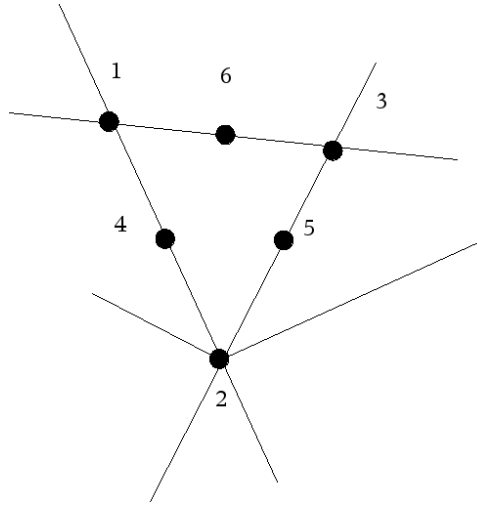
Les ordres d'interpolations P_1 et P_2 que nous utilisons sont basés sur des polynômes de Lagrange donnant respectivement des représentations linéaires et quadratiques des solutions.

L'interpolation P_1 considère 3 degrés de liberté par champs et par cellule. Ces degrés de liberté sont localisés sur les sommets de chaque triangle (Figure 1.26), permettant de formaliser les fonctions de base P_1 ($\vec{\varphi}_i = (\psi_1, \psi_2, \psi_3)^t$) basées sur les coordonnées barycentriques $((L_1, L_2, L_3)^t$) du triangle :

$$\begin{aligned}
 \psi_1 &= L_1 = \frac{1}{2A} (x_2 z_3 - x_3 z_2 + (z_2 - z_3)x + (x_3 - x_2)z) \\
 \psi_2 &= L_2 = \frac{1}{2A} (x_3 z_1 - x_1 z_3 + (z_3 - z_1)x + (x_1 - x_3)z) \\
 \psi_3 &= L_3 = \frac{1}{2A} (x_1 z_2 - x_2 z_1 + (z_1 - z_2)x + (x_2 - x_1)z)
 \end{aligned} \tag{1.69}$$

L'interpolation P_2 considère 6 degrés de liberté par champs et par cellules. Ces degrés de liberté sont localisés sur les sommets de chaque triangle ainsi qu'au milieu de chaque arête (Figure 1.27). En utilisant les coordonnées barycentriques, nous pouvons écrire les polynômes P_2 :

$$\begin{aligned}
 \psi_1 &= 2(L_1 - 1)L_1 \\
 \psi_2 &= 2(L_2 - 1)L_2 \\
 \psi_3 &= 2(L_3 - 1)L_3
 \end{aligned}$$


 FIG. 1.27 – Schémas des noeuds des fonctions de base P_2 .

$$\begin{aligned}
 \psi_4 &= 4L_1L_2 \\
 \psi_5 &= 4L_2L_3 \\
 \psi_6 &= 4L_3L_1
 \end{aligned} \tag{1.70}$$

A partir de la formulation des vecteurs des fonctions de base, les matrices $\tilde{\mathcal{F}}$, $\tilde{\mathcal{G}}$ et $\tilde{\mathcal{E}}$ sont obtenues par intégration numérique des combinaisons des fonctions de base (équations (1.45), (1.46) et (1.47)) sur chaque triangle. L'annexe A donne un rappel sur l'intégration numérique et un exemple pour l'application GD P_1 .

Enfin, le mélange des ordres d'interpolation peut être envisagé très facilement dans la formulation en développant les matrices $\tilde{\mathcal{F}}$ et $\tilde{\mathcal{G}}$ pour les ordres considérés.

1.7.3 Approche *parsimonieuse*

Comme cela a été illustrée pour la formulation GD P_0 dans la partie 1.6.4, l'approche parsimonieuse (Luo & Schuster, 1990) permet d'obtenir le système discrétisé relatif aux équations élastodynamiques du second ordre à partir des équations du premier ordre. Cette méthode permet de discrétiser les opérateurs du second ordre en DF à partir des schémas construits pour les opérateurs différentiels du premier ordre (Hustedt *et al.*, 2004; Operto *et al.*, 2007). La substitution des champs de contraintes dans les équations par leur expression fonction de la vitesse, permet de limiter le nombre de degrés de liberté en ne calculant que les champs de vitesses au lieu des champs de vitesses et de contraintes. Pour une résolution du système linéaire avec un solveur direct, cette méthode permet de réduire l'ordre de la matrice à résoudre d'un facteur 2.5 (2 champs de vitesses contre 5 champs de vitesses/contraintes). Cependant, la matrice d'impédance obtenue par l'approche parsimonieuse est plus dense et possède une largeur de bande numérique plus grande, occasionnant un plus grand remplissage lors de la factorisation. Ainsi, les tests numériques ont montré que l'approche parsimonieuse permettait d'économiser 25 à 30 % de mémoire, comparativement à l'approche fondée sur la résolution du système d'ordre 1 pour le schéma P_0 .

Cependant, la plus grande complexité des formulations P_1 , P_2 et mixtes, comparé au P_0 , rend difficile l'écriture analytique et générique du schéma parsimonieux pour ces ordres d'interpolation.

Une alternative possible est le passage en formulation parsimonieuse de façon numérique après la création de la matrice discrète du premier ordre. Je présente ci-dessous un exemple simple de cette manipulation sur une matrice 4×4 :

Considérons le système représentant la matrice du premier ordre discrétisé :

$$\begin{bmatrix} a_{11} & a_{12} & 0 & 0 \\ a_{21} & a_{22} & a_{23} & 0 \\ 0 & a_{32} & a_{33} & a_{34} \\ 0 & 0 & a_{43} & a_{44} \end{bmatrix} \begin{bmatrix} V_1 \\ P_2 \\ V_3 \\ P_4 \end{bmatrix} = \begin{bmatrix} s_1 \\ s_2 \\ s_3 \\ s_4 \end{bmatrix}, \quad (1.71)$$

où V_1 et V_3 sont les éléments du vecteur vitesse, P_2 et P_4 sont les éléments du vecteur contrainte et les termes s_i sont les termes sources. La structure particulière de la matrice, dans laquelle les lignes d'inconnues de contraintes ne sont dépendantes que d'inconnues de vitesses, et inversement, permet d'écrire un système uniquement en vitesses de manière équivalente au système (1.71) :

$$\begin{bmatrix} a_{11} - \frac{a_{12}a_{21}}{a_{22}} & -\frac{a_{12}a_{23}}{a_{22}} \\ -\frac{a_{32}a_{21}}{a_{22}} & a_{33} - \frac{a_{32}a_{21}}{a_{22}} - \frac{a_{34}a_{43}}{a_{44}} \end{bmatrix} \begin{bmatrix} V_1 \\ V_3 \end{bmatrix} = \begin{bmatrix} s_1 - \frac{a_{12}s_2}{a_{22}} \\ s_3 - \frac{a_{32}s_2}{a_{22}} - \frac{a_{34}s_4}{a_{44}} \end{bmatrix}, \quad (1.72)$$

en supposant que les termes de collocation a_{22} et a_{44} sont non nuls, ce qui est vrai en pratique dans la formulation GD si le module du terme de pulsation ω n'est pas nul (il pourrait avoir une partie réelle nulle et imaginaire non nulle).

Je propose l'algorithme 1.1 qui détaille cette manipulation de la matrice en posant :

1. A la matrice du premier ordre de termes $a_{i,j}$ et B la matrice du second ordre de termes $b_{i,j}$.
2. s le vecteur second membre du premier ordre et \bar{s} le vecteur second membre du second ordre.
3. ns l'ordre de la matrice du premier ordre.
4. vel un vecteur de taille ns dont la valeur est 1 si la ligne correspondante est une inconnue de vitesse, et 0 si c'est une inconnue de contrainte.
5. $nztab$ un vecteur donnant, pour chaque ligne de la matrice, le nombre de terme non nul moins un (moins un pour enlever le terme de collocation).
6. $indexcolumn$ un tableau 2D qui pour chaque ligne i de la matrice et chaque index de terme non nul j , donne l'index jc de la colonne du terme j dans la ligne i (le terme de collocation n'est pas présent dans ce tableau).
7. $pass$ un vecteur donnant l'index $inew$ dans la matrice B , d'une ligne i de la matrice A .

Notons cependant que cette technique nécessite de stocker simultanément en mémoire les deux matrices A et B , pouvant causer des limitations pour les problèmes de grandes tailles. De plus, l'implémentation numérique de cet algorithme de façon séquentielle est relativement coûteuse en temps, avec un temps nécessaire à la manipulation des matrices, parfois supérieur au temps de factorisation. Cette approche n'est donc pas viable pour les multiples applications dans le cadre d'un algorithme d'inversion de formes d'ondes. Enfin, la parallélisation de cet

Algorithm 1.1 Numerical parsimonious transfert from 1^{st} order matrix to 2^{nd} order matrix

```

1: for  $i = 1$  to  $ns$  do
2:   if  $vel(i) = 1$  then
3:      $inew = pass(i)$ 
4:      $b_{inew,inew} = a_{i,i}$ 
5:      $\bar{s}_{inew} = s_i$ 
6:     for  $j = 1$  to  $nztab(i)$  do
7:        $jc = indexcolumn(i, j)$ 
8:        $jr = jc$ 
9:       for  $k = 1$  to  $nztab(jr)$  do
10:         $kc = indexcolumn(jr, k)$ 
11:         $knew = pass(kc)$ 
12:         $b_{inew,knew} = b_{inew,knew} - \frac{a_{kc,jr}a_{jr,kc}}{a_{jr,jr}}$ 
13:       end for
14:        $\bar{s}_{inew} = \bar{s}_{inew} - \frac{s_{jr}a_{i,jr}}{a_{jr,jr}}$ 
15:     end for
16:   end if
17: end for

```

algorithme sur une architecture à mémoires distribuées n'est pas triviale car elle nécessite d'une part une décomposition en domaine spatiaux avec recouvrement associé chacun à un CPU, et d'autre part un *mapping* astucieux des degrés de liberté de la matrice sur ces sous-domaines.

Finalement, pour tirer bénéfice du caractère générique du choix des ordres d'interpolation, nous résoudrons par la suite uniquement le système discrétisé issu de la formulation du premier ordre. Le développement d'une approche parsimonieuse de façon analytique ou numérique après discrétisation reste une perspective potentielle d'amélioration de la formulation.

1.7.4 Implémentation du terme de source

L'intérêt principal des ordres d'interpolation élevés réside dans l'utilisation de maillages non-structurés. Contrairement aux maillages réguliers équilatéraux, les maillages non-structurés n'ont, par définition, pas de symétries et n'engendrent donc pas de problème de *red-black pattern*, comme nous avons pu le voir pour l'interpolation P_0 dans la partie 1.6.5, dû à la combinaison de maillages réguliers et de schémas de flux centrés. Ainsi, l'implémentation numérique du terme de source peut se faire sans étalement.

Grâce aux fonctions d'interpolation des schémas P_k , la source peut être localisée spatialement dans la cellule qui la contient. Cette localisation reste, bien sûr, approximée par le degré d'interpolation utilisé. En fonction de la position spatiale de la source physique, les termes de sources numériques aux degrés de liberté de la cellule se verront attribuer un poids correspondant à la projection de la position physique, sur la base de fonction d'interpolation. L'énergie injectée dans le système linéaire est égale à la somme des énergies de chaque degrés de liberté. Pour chaque degrés de liberté, l'énergie est proportionnelle à l'amplitude de source associée, multipliée par l'intégrale surfacique de sa fonction sur la cellule.

Le problème P_2

L'interpolation P_2 présente un cas pathologique pour ce type de représentation de source lorsque la position physique de la source tombe à proximité d'un sommet de triangle (le terme proximité étant utilisé pour une position proche, voire sur le sommet). Dans ce cas, les degrés de liberté du triangle, sauf le degré de liberté du sommet de la source, se voient attribuer des amplitudes de sources très faible. Or, l'énergie est principalement portée par la fonction de base du sommet dont l'intégrale est nulle en 2D et 3D... Ainsi, le système n'est excité que par du bruit numérique dans ce cas particulier.

Les solutions pour corriger ce problème sont multiples mais aucune n'est exempte de défaut :

- Décaler la source dans la cellule pour impliquer d'autres degrés de liberté, mais induit une erreur due au décalage spatial.
- Utiliser une autre interpolation pour la cellule de la source, mais engendre une erreur de précision si l'ordre utilisé est plus faible.
- Etaler la source sur plusieurs cellules, mais induit une erreur due la modélisation de la source « moins » ponctuelle, surtout si les cellules sont grosses.
- Modifier le maillage localement pour que chaque source soit localisée dans une position non problématique, mais nécessite une maîtrise fine des outils de maillage.
- Changer le type de base d'interpolation pour éviter ce problème en P_2 , en prenant une représentation modale par exemple (Käser & Dumbser, 2006).

Par la suite, pour le P_2 , nous utiliserons un raffinement local du maillage dans la zone des sources ainsi qu'un décalage arbitraire de la source au barycentre de la cellule. Cette combinaison permet d'assurer la qualité de la source tout en minimisant l'erreur de décalage spatial grâce à l'utilisation de cellules de faibles tailles.

1.7.5 Application numérique

Je vais présenter quelques applications numériques des ordres d'interpolation P_1 , P_2 et mixte P_0 - P_1 afin de quantifier la précision et les règles de discrétisation, ainsi que d'illustrer les applications possibles de ces schémas.

Une première application concerne le cas de Garvin (1956), simulant un demi-espace élastique homogène au toit duquel se situe une surface libre plane. Une source explosive localisée en profondeur excite le milieu. Les Figures 1.28 et 1.29 montrent des comparaisons de sismogrammes en temps des vitesses de déplacement, mesurées à la surface libre. Les interpolations P_1 et P_2 sont utilisées dans des maillages non-structurés, avec des discrétisations moyennes respectives de 12 et 3 cellules par longueur d'onde minimale. Ces comparaisons montrent un bon accord entre solutions analytiques et numériques pour ces discrétisations, aussi bien pour l'onde directe, que pour l'onde de surface, plus délicate à modéliser.

La courbe de la Figure 1.30 montre l'erreur L_2 obtenue pour un cas de Garvin simulé en domaine fréquentiel lorsque la taille du maillage évolue pour les interpolations P_0 , P_1 et P_2 . Ainsi, une erreur équivalente est obtenue entre l'interpolation P_0 sur maillage équilatéral à 10 cellules par longueur d'onde, et le P_2 sur maillage non-structuré à 3 cellules par longueur d'onde. L'interpolation P_1 nécessite un maillage plus raffiné de l'ordre de 14 cellules par longueur d'onde, mais permet d'utiliser des maillages non-structurés. Ces conclusions en terme de précision sont

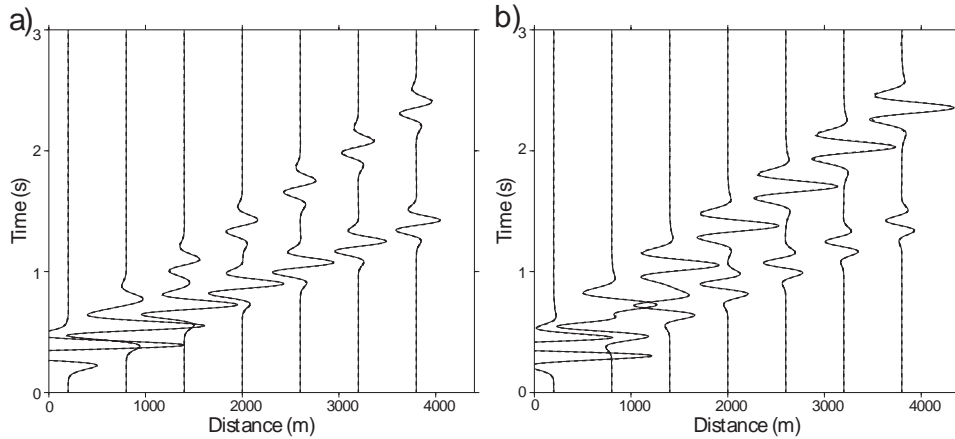


FIG. 1.28 – Comparaison en domaine temporel de la solution analytique (courbe continue) et du schéma GD P_1 (courbe discontinue) pour le cas Garvin discrétisé à 12 cellules par longueur d’onde en maillage non-structuré. Données des vitesses de déplacement (a) horizontal et (b) vertical à la surface.

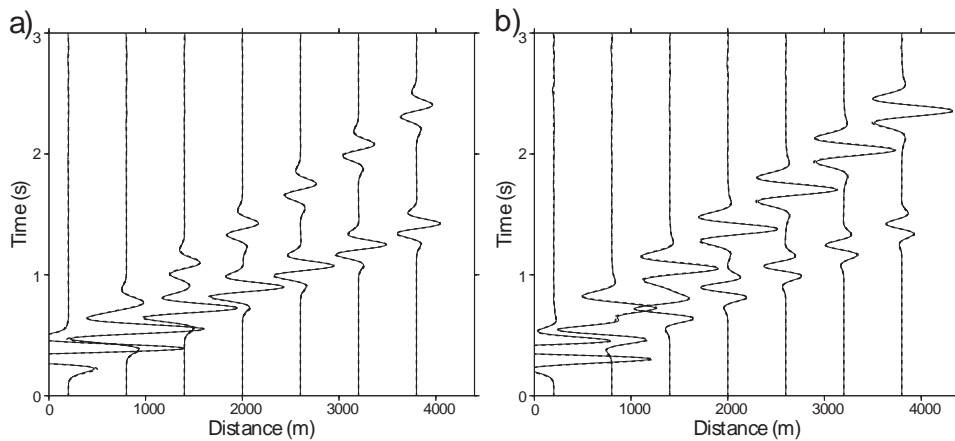


FIG. 1.29 – Comparaison en domaine temporel de la solution analytique (courbe continue) et du schéma GD P_2 (courbe discontinue) pour le cas Garvin discrétisé à 3 cellules par longueur d’onde en maillage non-structuré. Données des vitesses de déplacement (a) horizontal et (b) vertical à la surface.

conformes avec les études de convergence de Käser & Dumbser (2006); Delcourte *et al.* (2009); Etienne *et al.* (2009) : nous utiliserons donc ces critères de discrétisation pour l’application à l’inversion des formes d’ondes.

Une seconde application dans un cas de modèle *Corner Edge* (*cf.* Figure 1.15) montre une application mixte P_0 - P_1 . La Figure 1.31 représente les champs d’ondes fréquentiels à 5 Hz, des vitesses de déplacement horizontale et verticale dans ce modèle pour une source superficielle. Une couche de cellules non-structurée dans les premiers 300 m du modèle est calculée par interpolation P_1 , tandis que la partie profonde est calculée par une interpolation P_0 dans un maillage régulier équilatéral. Notons que l’utilisation combinée des schémas P_0 et P_1 , intro-

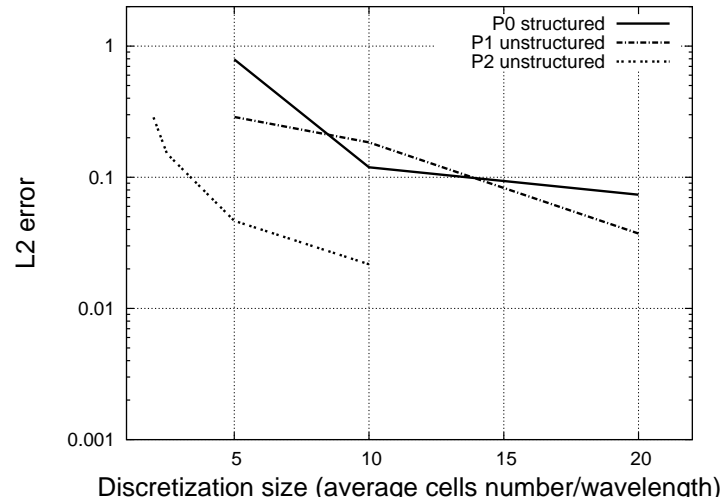


FIG. 1.30 – Etude d’erreur des schémas GD P_0 , P_1 et P_2 en fonction de la discrétisation pour un cas de Garvin.

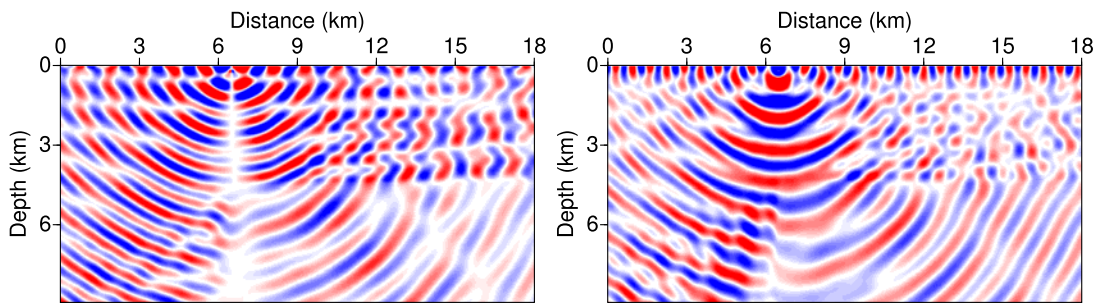


FIG. 1.31 – Partie réelle des champs monochromatiques de vitesses horizontales (gauche) et verticales (droite) à 5 Hz pour un cas *Corner Edge* simulé avec une interpolation mixte GD P_0 - P_1 . Notons que l’interface entre les deux interpolations à 300 m de profondeur n’introduit pas d’artefact dans les champs.

duit très naturellement dans la formulation GD, n’engendre pas d’artefacts dans les champs d’ondes. L’utilisation de ce mélange d’ordre est donc particulièrement intéressante pour combiner l’intérêt du maillage non-structuré, calculé par interpolation P_1 , pour la représentation de la topographie ou de la bathymétrie, la localisation des sources et des récepteurs sans étalement de source, tout en utilisant un maillage régulier, calculé par interpolation P_0 , si le milieu est peu contrasté.

Une troisième application dans un cas marin se focalise sur le contact eau/solide. Le modèle réaliste marin de Valhall (Figure 1.32) qui présente une faible couche d’eau (70 m) est utilisé. Les récepteurs sont localisés en fond de mer afin de simuler une acquisition par câble. La source explosive est localisé 5 m sous la surface de l’eau à la distance 7.4 km. La Figure 1.33 montre des données en domaine temporel de vitesses de déplacement, calculés à partir d’une simulation en domaine fréquentiel, pour une source ricker de fréquence centrale 4 Hz. Pour ce test, une

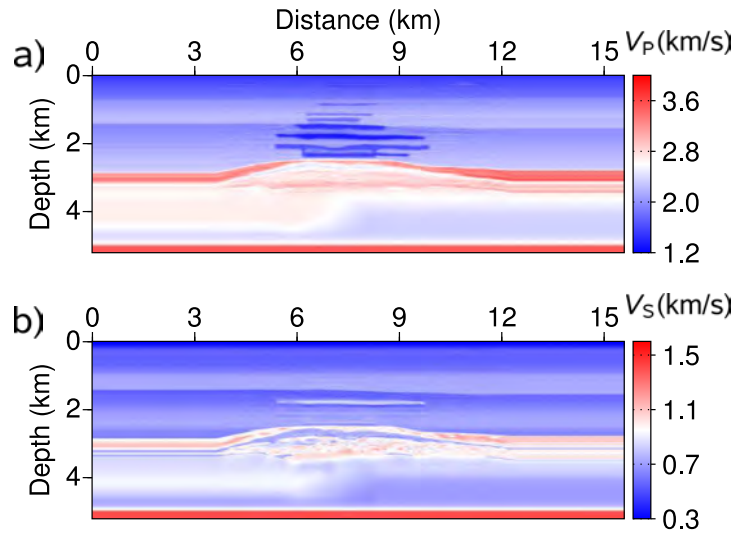


FIG. 1.32 – Le modèle synthétique de Valhall pour les vitesses d’ondes (a) P et (b) S.

interpolation mixte est choisie. Une couche de cellules non-structurée de 130 m d’épaisseur est calculée par DG P_1 pour modéliser le contact eau/solide. Un maillage équilatéral régulier, calculé par DG P_0 , prolonge cette couche. Nous reviendrons plus en détail sur ce modèle et le maillage utilisé dans la partie 3.3.2. Notons que grâce à l’utilisation du maillage non-structuré dans la zone de contact eau/solide, le phénomène de *red-black pattern* à l’origine des problèmes avec l’interpolation P_0 (cf. partie 1.6.6.1 et Figures 1.13) n’apparaît pas dans cette simulation : les données ne présentent ni instabilité ni ondes dispersives pénétrant dans le modèle après le contact des fronts d’ondes avec les zones PML.

Une dernière application est présentée dans le modèle synthétique réaliste IFP/TOTAL FOOTHILL (Figure 1.34). Ce modèle d’acquisition terrestre à topographie complexe (les zones bleues foncées des modèles représentent l’air au dessus de la surface libre) reprend les structures géologiques de zone de piémont. Notons qu’une zone altérée de 50 à 100 mètres d’épaisseur, dont les vitesses de propagation sont très lentes ($V_P = 750 \text{ m/s}$, $V_S = 350 \text{ m/s}$), est présente en subsurface. Ce type de configurations géologiques est très contraignant pour les méthodes basées sur des maillages ou grilles réguliers. Un maillage non-structuré adapté aux propriétés de vitesses locales a été créé pour une simulation GD P_2 , permettant d’une part de représenter très finement la topographie, et d’autre part de raffiner la densité des cellules dans la zone altérée de subsurface. La Figure 1.35 représente les parties réelles de champs monochromatiques de vitesses à 5 Hz.

Nous reviendrons plus en détails dans le chapitre 3 sur le coût numérique des différentes approches ainsi que sur les critères de choix des ordres d’interpolation et des maillages sur des cas réalistes.

1.8 Conclusion

Dans ce chapitre, j’ai présenté le problème direct de l’algorithme d’inversion des formes d’ondes : la résolution des équations de l’élastodynamique linéaire 2D pour les ondes P-SV.

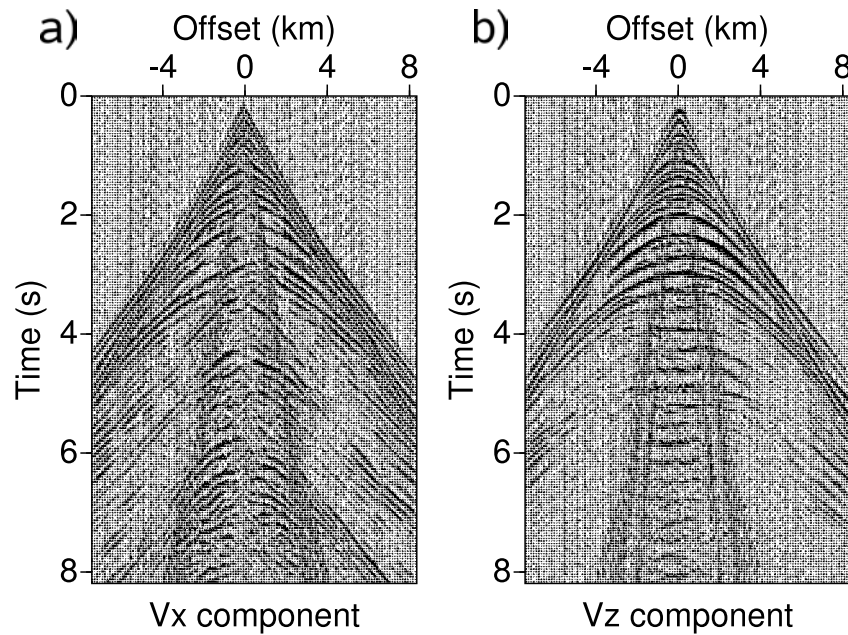


FIG. 1.33 – Données de vitesses de déplacement (a) horizontal et (b) vertical enregistrées en fond de mer dans le modèle Valhall. Une interpolation mixte P_0 - P_1 est utilisée dans cette simulation. L'utilisation de l'interpolation P_1 dans un maillage non-structuré permet de modéliser le contact eau/solide sans instabilité.

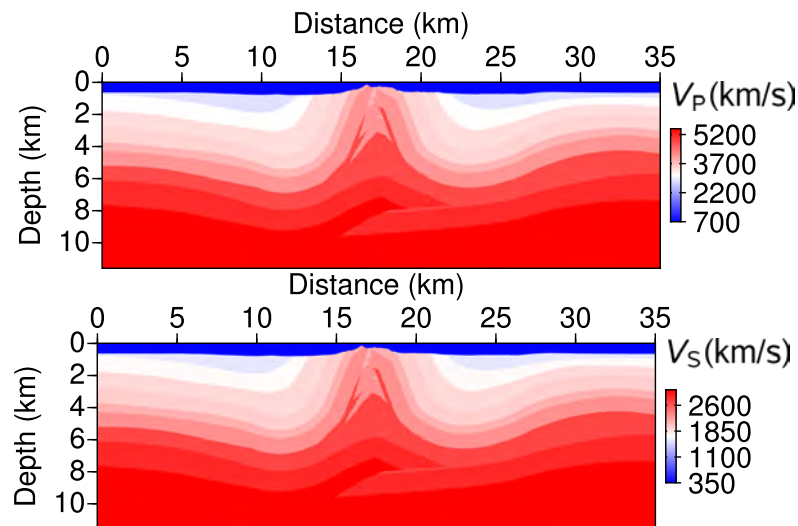


FIG. 1.34 – Modèle synthétique réaliste FOOTHILL. Les modèles de vitesse d'ondes P et S sont représentés respectivement en haut et en bas.

Ces équations sont résolues par une méthode d'éléments finis Galerkin discontinus qui combine un certain nombre d'avantages comparé aux méthodes plus classiques de différences finies ou d'éléments finis continus. Ainsi, les GD permettent d'utiliser des maillages triangulaires non-

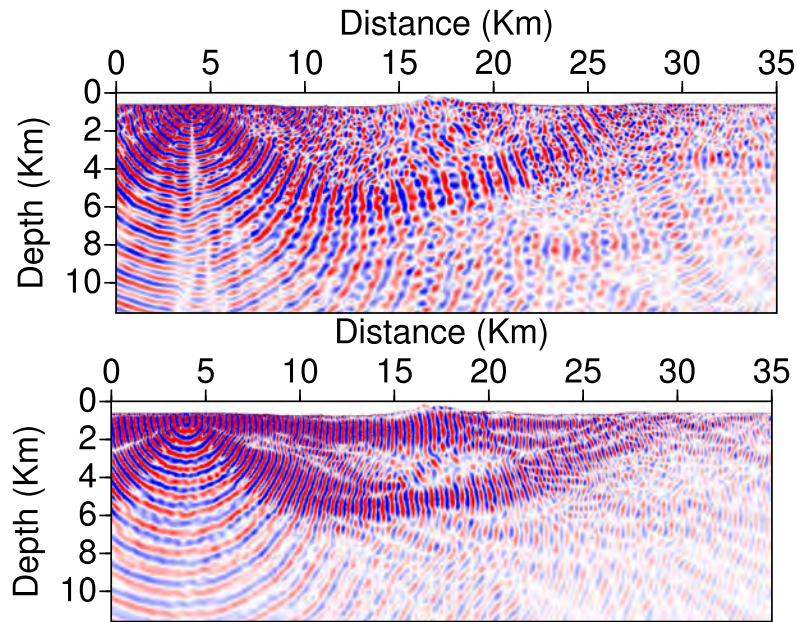


FIG. 1.35 – Partie réelle des champs monochromatique à 5 Hz pour le modèle FOOTHILL simulé avec une interpolation GD P_2 . Données des vitesses de déplacement horizontal (haut) et vertical (bas) à la surface.

structurés (pour les ordres supérieurs à P_0), représentant fidèlement des topographies complexes et permettant une adaptation locale de la taille des cellules aux paramètres physiques locaux. Cette adaptation se révèle très bénéfique pour des milieux à gradients de vitesse ou contenant des zones altérées de très faibles vitesses engendrant un suréchantillonnage spatial lors de l'utilisation de grilles régulières. L'utilisation combinée des maillages non-structurés et des ordres supérieurs à P_0 permet de modéliser le contact liquide/solide de forme arbitraire sans aucune condition explicite à l'interface ni de changement d'équations d'ondes pour la partie liquide. L'interpolation P_0 se révèle également très intéressante car basée sur des maillages réguliers équilatéraux, sa formulation est très proche des DF tout en permettant des discrétisations plus grossières. C'est donc une méthode de choix pour les milieux faiblement contrastés.

Chapitre 2

Le problème inverse : théorie, méthodes et problématiques

Sommaire

2.1	Problème aux moindres carrés linéarisé	92
2.1.1	La fonction coût	92
2.1.2	Linéarisation du problème inverse	93
2.1.3	Les équations normales	94
2.1.4	Estimation de la signature de la source	95
2.2	Gradient et Hessien	95
2.2.1	Interprétation et construction du gradient	95
2.2.2	Pouvoir de résolution spatiale du gradient	97
2.2.3	Non-linéarité et approche multi-échelles	98
2.2.4	Interprétation et construction du Hessien	100
2.3	Résolution numérique du problème d'optimisation locale	101
2.3.1	Méthode du Gradient	101
2.3.2	Le pas de descente α	101
2.3.3	Le Gradient Conjugué	102
2.3.4	Méthode de Newton et Gauss-Newton	103
2.3.5	Méthodes de Quasi-Newton	103
2.3.6	Préconditionnement	104
2.3.7	Régularisations	104
2.3.8	Exemple	105
2.4	Le modèle initial	106
2.4.1	Modèle cinématiquement compatible et saut de phase	106
2.4.2	Tomographie des temps de première arrivée	107
2.4.3	L'inversion dans le domaine de Laplace	108
2.4.4	Inversion de la phase de première arrivée	109
2.4.5	La stéréo-tomographie	110
2.4.6	Conclusions	110
2.5	Les alternatives à la norme L_2	110
2.5.1	Fonction Logarithmique : Amplitude vs Phase - lien avec Rytov	110
2.5.2	Fonctionnelles robustes	111

2.6	L'inversion multiparamètres	112
2.7	Conclusion	114

Dans ce chapitre, j'introduis la formulation du problème inverse de l'inversion des formes d'ondes dans le domaine espace/fréquence. Celui-ci est généralement posé sous la forme d'un problème d'optimisation locale aux moindres carrés, fondé sur une linéarisation. Le gradient et le Hessien de la fonction coût sont développés avec la méthode de rétropropagation des résidus (formalisme de l'état adjoint), puis sont interprétés physiquement avant de présenter les alternatives possibles pour résoudre le problème d'optimisation locale. La construction du modèle de départ est ensuite abordée puis les alternatives de la norme L_2 . Enfin, je présente les enjeux et difficultés liés à l'inversion de plusieurs paramètres.

Le lecteur peut se référer aux ouvrages de Menke (1984); Tarantola (1987) qui discutent et présentent plus en détails le formalisme, les hypothèses et les problématiques des problèmes inverses.

2.1 Problème aux moindres carrés linéarisé

2.1.1 La fonction coût

La méthode d'imagerie par inversion de formes d'ondes est basée sur la minimisation de l'écart entre des données observées et des données calculées dans un modèle \mathbf{m} . Appliquée aux ondes sismiques, les données sont représentées dans le domaine temporel par des sismogrammes, et en domaine fréquentiel par des données complexes pour chaque fréquence. Pour la suite de cette étude, nous nous placerons dans le domaine fréquentiel.

On peut définir le vecteur résidu $\Delta \mathbf{d}$, comme la différence entre le vecteur de données observées \mathbf{d}_{obs} et le vecteur de données calculées $\mathbf{d}_{calc}(\mathbf{m})$. Considérons une source sismique et une fréquence. Les trois vecteurs sont de taille n_{data} , équivalent au nombre de données pour la source considérée. Pour n_{rec} récepteurs à une seule composante, nous avons $n_{data} = n_{rec}$. Dans un cas multicomposantes, nous pouvons avoir $n_{data} = 2 n_{rec}$ ou $n_{data} = 3 n_{rec}$.

Le vecteur des données calculées $\mathbf{d}_{calc}(\mathbf{m})$ est obtenu par résolution du problème direct $\mathbf{Ax} = \mathbf{b}$, en appliquant un opérateur de projection \mathcal{P}_{data} au vecteur solution \mathbf{x} du domaine complet.

L'objectif de l'inversion des formes d'ondes est de minimiser chacun des termes du vecteur résidu $\Delta \mathbf{d}$. Pour ce faire, un estimateur scalaire, appelé fonction coût, représentatif du vecteur $\Delta \mathbf{d}$ doit être introduit afin d'être minimisé. Dans l'hypothèse d'une distribution gaussienne des erreurs dans les données, Tarantola (1987) a montré que le maximum de la fonction densité de probabilité des modèles est obtenu par le modèle minimisant la fonctionnelle aux moindres carrés. Ainsi, la fonction coût L_2 , classiquement utilisée pour l'inversion de formes d'ondes en domaine fréquentiel (Pratt & Worthington, 1990; Pratt, 1990), est donnée par

$$\mathcal{C}(\mathbf{m}) = \frac{1}{2} \Delta \mathbf{d}^\dagger \Delta \mathbf{d}, \tag{2.1}$$

où l'exposant \dagger est l'opérateur adjoint, équivalent à l'application des opérateurs transposé (t) et conjugué (*). Cette équation, comme les suivantes, est développée pour une source et une

fréquence. Dans un cas multisources et/ou multifréquences, une sommation sur les sources et/ou les fréquences aurait lieu.

Afin de prendre en compte des informations *a priori* sur les données, un opérateur matriciel \mathbf{S}_d est appliqué au vecteur $\Delta \mathbf{d}$, donnant la fonction coût généralement utilisée (*weighted least-squares*) :

$$\mathcal{C}(\mathbf{m}) = \frac{1}{2} \Delta \mathbf{d}^\dagger \mathbf{S}_d^\dagger \mathbf{S}_d \Delta \mathbf{d}. \quad (2.2)$$

Par la suite nous noterons $\mathbf{S}_d^\dagger \mathbf{S}_d = \mathbf{W}_d$.

L'objectif de la méthode est donc de minimiser cette fonction coût (2.2), représentative des n_{data} résidus de l'espace des données, en modifiant le vecteur modèle \mathbf{m} de taille n_{mod} . En pratique, pour des applications réalistes nous avons $n_{mod} \sim \mathcal{O}(10^4) - \mathcal{O}(10^6)$.

2.1.2 Linéarisation du problème inverse

La minimisation de la fonction coût (2.2) peut être abordée par deux grands types de méthodes d'exploration de l'espace des modèles :

1. Les méthodes de recherche globale et semi-globale (Sen & Stoffa, 1995) : méthode de recherche sur grille ou de Monte-carlo (Metropolis & Ulam, 1949), de recuit simulé (Kirkpatrick *et al.*, 1983) par exemple. Ces méthodes permettent d'explorer, sans *a priori* fort, l'espace des modèles afin de trouver le minimum absolu de la fonction coût. Cependant, elles nécessitent un grand nombre de simulations directes. Elles sont donc généralement destinées pour des problèmes d'optimisation à peu de paramètres ($\mathcal{O}(10^1)$ à $\mathcal{O}(10^2)$) et/ou pour lesquels les solutions du problème direct sont calculables à un coût très faible.
2. Les méthodes de recherche locale : en partant d'un modèle de départ *a priori*, localisé dans le bassin d'attraction du minimum global de la fonction coût pour des fonctions non-convexes, ces méthodes permettent de converger vers le modèle du minimum absolu. Ces méthodes permettent d'envisager des problèmes à grand nombre d'inconnues et/ou pour lesquels les solutions du problème direct sont relativement coûteuses à obtenir.

En raison du grand nombre de paramètres à estimer et au coût important du calcul des solutions du problème direct, la méthode d'inversion des formes d'ondes est généralement formulée de manière linéarisée et itérative pour utiliser des méthodes d'optimisation locale. A chaque itération k , nous considérons que le minimum de la fonction $\mathcal{C}(\mathbf{m}^{(k)})$ est recherché au voisinage d'un modèle de départ $\mathbf{m}^{(k-1)}$ issu de l'itération précédente. Bien sûr, un modèle \mathbf{m}_0 est nécessaire à la première itération du processus et devra être déterminé au préalable (*cf.* partie 2.4). Ainsi, à chaque itération, nous pouvons écrire le modèle recherché comme la somme du modèle de départ $\mathbf{m}^{(k-1)}$ et du vecteur perturbation $\Delta \mathbf{m}$: $\mathbf{m}^{(k)} = \mathbf{m}^{(k-1)} + \Delta \mathbf{m}^{(k)}$. En considérant que les amplitudes du vecteur perturbation sont relativement petites face au vecteur modèle, un développement de Taylor à l'ordre 2 nous donne :

$$\begin{aligned} \mathcal{C}(\mathbf{m}^{(k)}) = \mathcal{C}(\mathbf{m}^{(k-1)} + \Delta \mathbf{m}^{(k)}) &= \mathcal{C}(\mathbf{m}^{(k-1)}) + \sum_{j=1}^{n_{mod}} \frac{\partial \mathcal{C}(\mathbf{m}^{(k-1)})}{\partial m_j} \Delta m_j^{(k)} \\ &+ \frac{1}{2} \sum_{j=1}^{n_{mod}} \sum_{k=1}^{n_{mod}} \frac{\partial^2 \mathcal{C}(\mathbf{m}^{(k-1)})}{\partial m_j \partial m_k} \Delta m_j^{(k)} \Delta m_k^{(k)} + \mathcal{O}(\mathbf{m}^3). \end{aligned} \quad (2.3)$$

La dérivée de cette équation par rapport au paramètre du model m_l donne alors :

$$\frac{\partial \mathcal{C}(\mathbf{m}^{(k)})}{\partial m_l} = \frac{\partial \mathcal{C}(\mathbf{m}^{(k-1)})}{\partial m_l} + \sum_{j=1}^{n_{mod}} \frac{\partial^2 \mathcal{C}(\mathbf{m}^{(k-1)})}{\partial m_j \partial m_l} \Delta m_j^{(k)}. \quad (2.4)$$

En annulant la dérivée de la fonction $\mathcal{C}(\mathbf{m}^{(k)})$, l'équation de Newton permet d'obtenir le minimum de la fonction, considérée localement parabolique, et donnant le vecteur perturbation :

$$\Delta \mathbf{m}^{(k)} = - \left[\frac{\partial^2 \mathcal{C}(\mathbf{m}^{(k-1)})}{\partial \mathbf{m}^2} \right]^{-1} \frac{\partial \mathcal{C}(\mathbf{m}^{(k-1)})}{\partial \mathbf{m}}. \quad (2.5)$$

Notons que l'équation de Newton (2.5) donne le minimum de la fonction en une itération si la fonction est quadratique ($\mathcal{O}(\mathbf{m}^3) = 0$), ce qui est le cas des problèmes inverses linéaires. Pour le problème d'inversion des formes d'ondes, le problème inverse est non-linéaire, il doit donc être résolu itérativement en le linéarisant à chaque itération.

Le terme de premier ordre $\frac{\partial \mathcal{C}(\mathbf{m}^{(k-1)})}{\partial \mathbf{m}}$ est le vecteur gradient de la fonction. Le terme de second ordre $\frac{\partial^2 \mathcal{C}(\mathbf{m}^{(k-1)})}{\partial \mathbf{m}^2}$ est la matrice Hessienne ou Hessian.

2.1.3 Les équations normales

L'équation de Newton (2.5) permet d'obtenir la perturbation du modèle appliquée à chaque itération à partir du gradient et du Hessian qu'il faut exprimer en fonction des données (Pratt *et al.*, 1998).

Dérivons tout d'abord la fonction $\mathcal{C}(\mathbf{m})$ par rapport au vecteur paramètre \mathbf{m} , afin d'obtenir le gradient $\mathcal{G}_{\mathbf{m}}$:

$$\begin{aligned} \mathcal{G}_{\mathbf{m}} &= \frac{\partial \mathcal{C}(\mathbf{m})}{\partial \mathbf{m}} \\ &= - \left[\left(\frac{\partial \mathbf{d}_{cal}(\mathbf{m})}{\partial \mathbf{m}} \right) \mathbf{W}_d (\mathbf{d}_{obs} - \mathbf{d}_{cal}(\mathbf{m}))^* + (\mathbf{d}_{obs} - \mathbf{d}_{cal}(\mathbf{m})) \mathbf{W}_d \left(\frac{\partial \mathbf{d}_{cal}(\mathbf{m})}{\partial \mathbf{m}} \right)^* \right] \\ &= -\Re \left[\left(\frac{\partial \mathbf{d}_{cal}(\mathbf{m})}{\partial \mathbf{m}} \right)^\dagger \mathbf{W}_d (\mathbf{d}_{obs} - \mathbf{d}_{cal}(\mathbf{m})) \right] \\ &= -\Re \left[\mathbf{J}^\dagger \mathbf{W}_d \Delta \mathbf{d} \right], \end{aligned} \quad (2.6)$$

où \Re représente la partie réelle d'un nombre complexe et \mathbf{J} la matrice de sensibilité ou matrice des dérivées de Fréchet. Notons que \mathbf{J} représente l'opérateur discret de Born si on considère le problème direct linéarisé $\Delta \mathbf{d} = \mathbf{J} \Delta \mathbf{m}$ (Tarantola, 1987; Pratt *et al.*, 1996).

La dérivée du gradient par rapport au vecteur paramètre \mathbf{m} permet d'obtenir l'expression du Hessian

$$\mathbf{B} = \frac{\partial^2 \mathcal{C}(\mathbf{m})}{\partial \mathbf{m}^2} = \Re \left[\mathbf{J}^\dagger \mathbf{W}_d \mathbf{J} \right] + \Re \left[\frac{\partial \mathbf{J}^t}{\partial \mathbf{m}^t} \mathbf{W}_d (\Delta \mathbf{d}^* \dots \Delta \mathbf{d}^*) \right]. \quad (2.7)$$

En remplaçant le gradient et le Hessian de l'équation de Newton (2.5) par leurs expressions (2.6) et (2.7), nous obtenons le système d'équations normales :

$$\begin{aligned} \mathbf{B}^{(k-1)} \Delta \mathbf{m}^{(k)} &= -\mathcal{G}^{(k-1)} \\ \Re \left[\mathbf{J}^\dagger \mathbf{W}_d \mathbf{J} + \frac{\partial \mathbf{J}^t}{\partial \mathbf{m}^t} \mathbf{W}_d (\Delta \mathbf{d}^* \dots \Delta \mathbf{d}^*) \right]^{(k-1)} \Delta \mathbf{m}^{(k)} &= \Re \left[\mathbf{J}^\dagger \mathbf{W}_d \Delta \mathbf{d} \right]^{(k-1)}. \end{aligned} \quad (2.8)$$

2.1.4 Estimation de la signature de la source

Pour un problème d'inversion de données réelles, la signature temporelle de la source sismique est généralement inconnue. Elle doit donc être estimée comme une inconnue du problème d'inversion afin de pouvoir calculer le vecteur résidu pour des données issues d'une « même source » : même amplitude et même phase. Pratt (1999) propose une estimation de la signature de la source pour chaque fréquence discrète au sein même du processus d'inversion itératif, en considérant que la signature de la source est le résultat d'un problème d'inversion linéaire. En effet les champs d'ondes \mathbf{x} du système linéaire $\mathbf{Ax} = \mathbf{b}$ du problème direct sont linéairement reliés au terme de source \mathbf{b} . La solution du problème inverse linéaire est donnée par :

$$s = \frac{\mathbf{d}_{cal}(\mathbf{m})^\dagger \mathbf{d}_{obs}}{\mathbf{d}_{cal}(\mathbf{m})^\dagger \mathbf{d}_{cal}(\mathbf{m})}, \quad (2.9)$$

où l'estimation s est le scalaire complexe de correction à appliquer aux champs calculés $\mathbf{d}_{cal}(\mathbf{m})$. Notons que l'estimation de la source peut être indifféremment considérée pour l'ensemble des tirs d'une seule fréquence si la source réelle est répétitive, ou pour chaque source individuelle si ce n'est pas le cas, ce qui augmente le nombre de degrés de liberté pour l'espace des paramètres des sources.

2.2 Gradient et Hessien

Pratt *et al.* (1998) donnent des interprétations physiques du gradient et Hessien. Je donne ici un résumé de ces interprétations et les liens sous-jacent qu'elles produisent.

2.2.1 Interprétation et construction du gradient

Le gradient $\mathcal{G}_{\mathbf{m}}$ de la fonction coût $\mathcal{C}(\mathbf{m})$, est un vecteur de taille n_{mod} qui lie les dérivées des données par rapport aux paramètres. Nous pouvons tout d'abord tirer une interprétation physique en repartant de l'expression du problème direct $\mathbf{Ax} = \mathbf{b}$ (équation (1.50)). En différenciant le problème direct par rapport à un élément m_l du vecteur model \mathbf{m} , nous obtenons :

$$\mathbf{A} \frac{\partial \mathbf{x}}{\partial m_l} = -\frac{\partial \mathbf{A}}{\partial m_l} \mathbf{x}. \quad (2.10)$$

L'équation (2.10) montre que le champ aux dérivées partielles ($\frac{\partial \mathbf{x}}{\partial m_l}$), relatif à la matrice des dérivées de Fréchet \mathbf{J} , est équivalent à un champ d'ondes propagé par l'opérateur \mathbf{A} , ayant pour source le terme $-\frac{\partial \mathbf{A}}{\partial m_l} \mathbf{x}$.

Ce terme source est le produit de $\partial \mathbf{A} / \partial m_l$ par le champ incident \mathbf{x} . La matrice $\partial \mathbf{A} / \partial m_l$ est construite par différenciation de l'opérateur direct \mathbf{A} par rapport au paramètre m_l . C'est donc une matrice extrêmement creuse dont les termes non nuls sont localisés à proximité du terme diagonal de la position du paramètre m_l dans la matrice \mathbf{A} , qui porte la signature de la diffraction occasionnée par un point localisé à la position physique de m_l et ayant une perturbation du paramètre m_l . Cette signature, appelée diagramme de rayonnement, peut être visualisée pour différents paramètres afin d'étudier les couplages et découplages entre paramètres, en fonction des angles d'incidence des champs d'ondes (Tarantola, 1986; Forgues,

1996). Le terme source de l'équation (2.10) peut donc être interprété comme le champ émis par la source incidente \mathbf{b} et diffracté par le paramètre m_l .

Le gradient (équation (2.6)) est donc le produit en fréquence (corrélation à décalage nul en temps) entre le champ aux dérivées partielles $\frac{\partial \mathbf{x}}{\partial m_l}$ et le vecteur résidu $\Delta \mathbf{d}^*$. Cette corrélation permet de tester la similarité des deux champs et d'obtenir une image des points diffractants manquants dans le modèle $\mathbf{m}^{(k-1)}$, donnant lieu à des résidus dans $\Delta \mathbf{d}^*$.

Du point de vue de l'implémentation, l'équation (2.10) ne serait pas efficace car n_{mod} problèmes directs devraient être résolus par source du système d'acquisition pour pouvoir assembler le gradient. En remplaçant l'expression du champ aux dérivées partielles (2.10) dans l'expression du gradient (2.6), nous obtenons :

$$\mathcal{G}_{m_l} = \Re \left[\mathbf{x}^t \left[\frac{\partial \mathbf{A}}{\partial m_l} \right]^t \mathbf{A}^{-1t} \tilde{\mathcal{P}}_{data} \mathbf{W}_d \Delta \mathbf{d}^* \right], \quad (2.11)$$

où l'opérateur $\tilde{\mathcal{P}}_{data}$ projette le vecteur résidu $\Delta \mathbf{d}^*$ dans l'espace du problème direct. L'opérateur matriciel \mathbf{A}^{-1} est l'inverse de l'opérateur du problème direct et contient les fonctions de Green pour des sources issues de chacun des éléments de l'espace du problème direct. En vertu du principe de réciprocité spatiale (Claerbout, 1976; Aki & Richards, 2002), nous avons l'égalité :

$$\mathbf{A}^{-1t} = \mathbf{A}^{-1}, \quad (2.12)$$

qui permet de réécrire le gradient :

$$\mathcal{G}_{\mathbf{m}} = \sum_{l=1}^{n_{mod}} \mathcal{G}_{m_l}, \quad (2.13)$$

avec

$$\begin{aligned} \mathcal{G}_{m_l} &= \Re \left[\mathbf{b}^t \mathbf{A}^{-1t} \left[\frac{\partial \mathbf{A}}{\partial m_l} \right]^t \mathbf{A}^{-1} \tilde{\mathcal{P}}_{data} \mathbf{W}_d \Delta \mathbf{d}^* \right] \\ &= \Re \left[\mathbf{x}^t \left[\frac{\partial \mathbf{A}}{\partial m_l} \right]^t \mathbf{A}^{-1} \tilde{\mathcal{P}}_{data} \mathbf{W}_d \Delta \mathbf{d}^* \right] \\ &= \Re \left[\mathbf{x}^t \left[\frac{\partial \mathbf{A}}{\partial m_l} \right]^t \mathbf{r} \right]. \end{aligned} \quad (2.14)$$

Les expressions (2.13) et (2.14) montrent clairement que le gradient résulte du produit entre le champ incident \mathbf{x} , issu de la source \mathbf{b} , et du champ rétropropagé \mathbf{r} , issu des résidus \mathbf{d}^* aux positions des récepteurs, et pondéré par le diagramme de rayonnement $\frac{\partial \mathbf{A}}{\partial m_l}$ du paramètre imagé m_l .

Notons que l'expression (2.14) peut également être obtenue en utilisant le formalisme de l'état-adjoint utilisant une fonctionnelle augmentée et des multiplicateurs de Lagrange, issue de la communauté de l'optimisation (Lions, 1972). Une revue de cette méthode ainsi que l'application au problème d'inversion de formes d'onde est présenté dans Plessix (2006). Le formalisme des systèmes Karush–Kuhn–Tucker (KKT) (Nocedal & Wright, 1999), proche de celui de l'état-adjoint, est présenté dans Akcelik (2002); Askan *et al.* (2007), et mène également aux mêmes formulations.

Finalement, l'équation (2.14) montre que la construction du vecteur gradient nécessite de calculer deux champs d'ondes, par source du dispositif d'acquisition, pour en faire le produit :

1. le champ incident \mathbf{x} ,
2. le champ adjoint rétropropagé \mathbf{r} ,

donnant un algorithme beaucoup plus économique que celui utilisant l'équation (2.10).

Plusieurs connexions avec le principe de construction du gradient peuvent être faites :

- Le gradient peut être relié au principe d'imagerie en diffraction, où l'image du modèle de perturbation est construite, en vertu du principe de Huygens, par sommation des images élémentaires associées à chaque point diffractant (McMechan & Fuis, 1987).
- Le gradient peut être lié au principe du renversement temporel (Fink, 1993). L'opérateur conjugué (*) appliqué au résidu fréquentiel peut être assimilé à une propagation en sens inverse du temps dans le domaine temporel. Dans le principe du renversement temporel, on propage en sens inverse du temps un ensemble de données issues d'une source primaire, mesurées en des points de mesure et émises depuis ces points. L'ensemble des champs d'ondes propagés depuis les points de mesure se combinent pour donner une focalisation maximum au point physique de la source primaire. Dans le but d'imager le milieu avec l'inversion des formes d'ondes, on ne propage dans le milieu que les résidus des données, afin non pas de focaliser à la source, mais de focaliser sur les positions physiques des points diffractant manquants.
- D'une manière très liée au renversement temporelle, le gradient peut être comparé à l'opérateur du principe d'imagerie utilisé en migration (Claerbout, 1985).

Notons enfin que le gradient représente la dérivée première des données par rapport à chacun des paramètres m_l du vecteur modèle \mathbf{m} . Les éléments du gradient sont donc indépendants les uns des autres, par construction. Les hétérogénéités sont reconstruites par sommation de points diffractants. Alors que le problème direct de l'inversion des formes d'ondes simule l'ensemble du champ d'ondes (multiples réfléchies à la surface, multiples dans la couche d'eau, ondes de surfaces...), le gradient ne considère que les diffractions simples et ne permet pas, par construction, de prendre en compte de phénomènes de propagation plus complexes. Cette conséquence forte de l'approche linéarisée du problème inverse permet d'imager les grandes longueurs d'ondes du milieu. Cependant, il faut garder à l'esprit cette hypothèse forte pour envisager d'imager des hautes fréquences, plus sensibles aux petites hétérogénéités du milieu qui génèrent des diffractions multiples. La prise en compte du champ complet dans le problème direct et l'approche itérative du problème inverse permettent néanmoins de minimiser l'impact des diffractions multiples, mais la limite haute fréquence de l'inversion linéarisée, en terme de validité et de robustesse, reste encore peu définie dans la littérature.

2.2.2 Pouvoir de résolution spatiale du gradient

Une estimation de la résolution spatiale du gradient (2.14) peut être obtenue à partir d'une analyse en onde plane (Sirgue & Pratt, 2004). Considérons un milieu homogène de vitesse c_0 , une source, un récepteur et un point diffractant (Figure 2.1). Nous supposons que nous sommes en champ lointain pour lequel les champs d'ondes peuvent être considérés comme des ondes planes, et nous négligerons les effets de l'amplitude. Une onde plane se propage selon la direction $\vec{\mathbf{s}}$, de la source vers le point diffractant, et une autre dans la direction $\vec{\mathbf{r}}$, du récepteur vers le point diffractant. Nous pouvons alors écrire l'expression de ces deux ondes planes comme :

$$G_0(\mathbf{x}, \mathbf{s}) = \exp(ik_0\vec{\mathbf{s}}.\mathbf{x})$$

$$G_0(\mathbf{x}, \mathbf{r}) = \exp(ik_0\vec{\mathbf{r}}.\mathbf{x}), \quad (2.15)$$

où $k_0 = \omega/c_0$. Ces formes d'ondes planes peuvent être introduites dans l'expression du gradient (2.14) comme champ incident et champ adjoint, donnant l'équation :

$$\mathcal{G} = -\omega^2 \Re \{ \exp(-ik_0(\vec{\mathbf{s}} + \vec{\mathbf{r}}).\mathbf{x}) \Delta d \}. \quad (2.16)$$

L'équation (2.16) peut être vue comme un terme d'une série de Fourier tronquée. Ainsi, l'inversion des formes d'ondes peut être assimilée à une transformée de Fourier inverse dont les fonctions de bases sont les nombres d'ondes et les coefficients sont les termes du vecteur résidu Δd . Les bornes et l'échantillonnage de la transformée de Fourier inverse sont contrôlés par la géométrie du dispositif d'acquisition (Sirgue & Pratt, 2004).

Le vecteur nombre d'onde diffracté $k_0(\vec{\mathbf{s}} + \vec{\mathbf{r}})$, caractéristique de la résolution spatiale du gradient, peut alors s'exprimer dans le cas simple d'un réflecteur 1D pour une paire source/récepteur (Figure 2.1), en fonction de la fréquence et de l'angle de diffraction :

$$\mathbf{k} = \frac{2\omega}{c_0} \cos(\theta/2) \vec{\mathbf{n}}, \quad (2.17)$$

où $\vec{\mathbf{n}}$ est le vecteur unitaire dans direction $\vec{\mathbf{s}} + \vec{\mathbf{r}}$.

L'équation (2.17) permet ainsi de tirer plusieurs conclusions :

- Pour des dispositifs d'acquisition de surface, la résolution optimum, pour une fréquence donnée, est obtenue à offset nul pour lequel la résolution verticale est de $\lambda/2$
- Les basses fréquences et les grands offsets permettent d'échantillonner les petits nombres d'ondes, donnant l'information des grandes structures du milieu. Cependant, notons que les grands offsets mettent en jeu des distances de propagation plus grandes et plus susceptibles de contenir des déphasages importants.
- La couverture en nombre d'onde du modèle est déterminée à la fois par la fréquence et l'angle de diffraction, ce dernier pouvant en relié à l'offset source/récepteur.

2.2.3 Non-linéarité et approche multi-échelles

L'analyse du gradient et la redondance dans le domaine des nombres d'onde occasionné par la couverture des angles de diffraction et les fréquences ont permis de concevoir une procédure d'inversion hiérarchique et « économique » dans le domaine fréquentiel.

En utilisant des dispositifs d'acquisition grand offset, la couverture continue des angles de diffraction dans le milieu permet de choisir quelques fréquences discrètes à inverser dans le spectre, tout en assurant la continuité d'échantillonnage dans le domaine des nombres d'onde (Sirgue & Pratt, 2004). Des règles de choix de fréquences ont ainsi pu être établies par Sirgue & Pratt (2004); Mulder & Plessix (2008).

En commençant l'inversion par les basses fréquences, le processus d'imagerie reconstruit d'abord les grandes structures du milieu. En montant progressivement en fréquence, on échantillonne des nombres d'ondes de plus en plus grands, ajoutant à l'image des détails de plus en plus fins. Cette procédure hiérarchique des basses fréquences vers les hautes fréquences permet d'atténuer la non-linéarité du problème inverse et ainsi d'améliorer la convergence vers le minimum global de la fonction coût (si un modèle de départ suffisamment précis est utilisé, cf. partie 2.4).

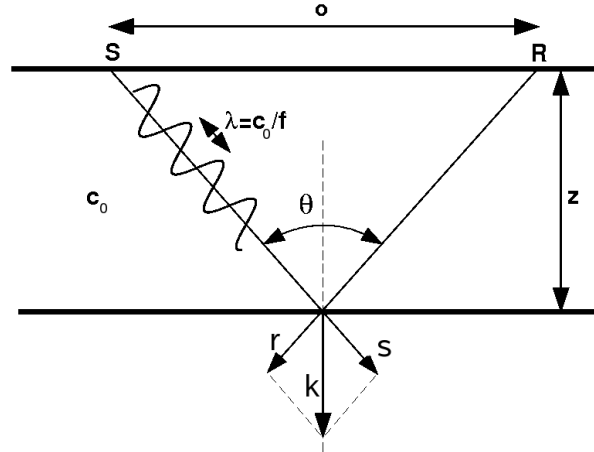


FIG. 2.1 – Configuration 1D simple comprenant une source S, un récepteur R et un point diffractant. La résolution locale au point diffractant de l'inversion de formes d'ondes est portée par le vecteur nombre d'onde \mathbf{k} .

L'utilisation combinée de la procédure hiérarchique des basses fréquences vers les hautes fréquences, et la décimation des données résultant de l'inversion de seulement quelques fréquences donne lieu à l'Algorithme 2.1, simple et économique en calcul comparé à des approches d'inversion en domaine temporel (Sirgue & Pratt, 2004; Brenders & Pratt, 2007a). Cette approche a permis une « renaissance » de la méthode d'inversion des formes d'ondes au début des années 90, après les résultats mitigés des années 80 effectués en domaine temporel. De nombreuses applications ont été effectuées depuis une vingtaine d'année avec cette méthodologie, principalement avec l'approximation acoustique.

En travaillant dans le domaine espace/fréquence, les données sont prises en compte sur toute la plage des temps de manière simultanée. Pour une fréquence et un récepteur, toutes les arrivées des différents trains d'ondes sont considérées dans un champ scalaire unique de valeur complexe. Il est toutefois possible de sélectionner des arrivées particulières dans les données temporelles observées par fenêtrage (Operto *et al.*, 2006). Cependant, la modélisation directe en domaine fréquentiel ne permet pas d'obtenir cette souplesse pour fenêtrer les données calculées. Le seul recours possible en domaine fréquentiel, si seulement quelques fréquences sont considérées, est d'utiliser des fréquences complexes. L'utilisation d'un terme de pulsation complexe $\bar{\omega} = \omega + i\gamma$ dans le problème direct est équivalent à modéliser des données temporelles amorties par une exponentielle décroissante (Shin *et al.*, 2002; Brenders & Pratt, 2007b). On a en effet la relation :

$$F(\omega + i\gamma)exp^{\gamma t_0} = \int_{-\infty}^{+\infty} f(t)exp^{-\gamma(t-t_0)}exp^{-i\omega t}dt, \quad (2.18)$$

où la fonction en domaine fréquentiel $F(\omega + i\gamma)exp^{\gamma t_0}$ est la transformée de Fourier du signal temporel $f(t)$ amorti par l'exponentielle décroissante $exp^{-\gamma(t-t_0)}$. Le terme t_0 est l'origine temporelle de la fonction exponentielle et peut être pris en pratique comme le temps de première arrivée ou le temps initial de la simulation. L'utilisation des fréquences complexes permet de sélectionner le volume de données modélisé dans le domaine temporel. L'amplitude des arrivées tardives, relatives aux ondes converties P-S et aux ondes de surfaces, peut ainsi être atténuée dans la modélisation en domaine fréquentiel.

Algorithm 2.1 Classic algorithm for frequency-domain full waveform inversion

```

1: for frequency = low to high do
2:   while (NOT convergence AND  $k < niter_{max}$ ) do
3:     Build gradient vector  $\mathcal{G}_{\mathbf{m}}^{(k-1)}$ 
4:     Build perturbation vector  $\delta\mathbf{m}$ 
5:     Define optimal step length  $\alpha$ 
6:     Update model  $\mathbf{m}^{(k+1)} = \mathbf{m}^{(k)} + \alpha\delta\mathbf{m}$ 
7:   end while
8: end for

```

Nous verrons dans les prochains chapitres 3, 4 et 5, que l'utilisation des fréquences complexes se révèle être un paramètre de régularisation qui permet, en plus de l'approche multi-échelle sur les fréquences croissantes, de minimiser l'impact des non-linéarités du problème inverse.

2.2.4 Interprétation et construction du Hessian

Le Hessian \mathbf{B} est la matrice de taille $n_{mod} \times n_{mod}$ des dérivées secondes de la fonction coût $\mathcal{C}(\mathbf{m})$ par rapport au vecteur modèle \mathbf{m} . Elle est formée par la somme de deux termes (équation (2.7)).

Le deuxième terme de (2.7) d'ordre 2 : $\Re \left[\frac{\partial \mathbf{J}^t}{\partial \mathbf{m}^t} \mathbf{W}_d (\Delta \mathbf{d}^* \dots \Delta \mathbf{d}^*) \right]$ est généralement considéré comme petit et négligeable lorsque les résidus sont faibles et/ou que le problème est quasi-linéaire (Tarantola, 1987; Pratt *et al.*, 1998). De plus, alors que le premier terme $\Re [\mathbf{J}^\dagger \mathbf{W}_d \mathbf{J}]$ est relativement aisé à calculer, le deuxième fait intervenir les dérivées de la matrice des dérivées de Fréchet, occasionnant un certain nombre de résolutions du problème direct supplémentaires. Notons tout de même que grâce aux dérivées d'ordre 2 sur les données, ce terme permet de prendre en compte les doubles diffractions.

Le premier terme du Hessian, généralement appelé Hessian approximé, \mathbf{B}_a , lorsque le deuxième terme est négligé, est formé par un produit en domaine fréquentiel, soit une corrélation à décalage nul en temps, entre les champs aux dérivées partielles de la matrice \mathbf{J} . La diagonale représente l'autocorrélation à décalage nul et contient donc le carré des amplitudes des champs aux dérivées partielles. Ce terme intègre l'effet de l'atténuation géométrique des champs aux dérivées partielles, et permet de corriger cet effet du gradient lorsque l'inverse du Hessian est appliqué dans les équations normales (2.8)¹. Les termes hors-diagonaux sont issus de l'intercorrélacion entre les champs aux dérivées partielles des différents paramètres m_l . Ces termes renseignent sur les corrélations entre les paramètres et leur implication dans les équations normales peut être considérée comme un opérateur de déconvolution du gradient, permettant d'atténuer les effets de largeur de bande limitée de la source et du dispositif d'acquisition et les défauts d'éclairage. Cette déconvolution du gradient permet ainsi d'améliorer la focalisation de la perturbation à appliquer (Pratt *et al.*, 1998). De plus, pour les problèmes multiparamètres, l'application du Hessian au gradient permet de mettre à l'échelle les gradients de chaque classe de paramètres.

¹Le terme est au carré car il faut enlever l'effet de l'atténuation géométrique à la fois dans les champs aux dérivées partielles de \mathbf{J} et dans les résidus $\Delta \mathbf{d}$ qui sont multipliés dans le gradient.

La construction du Hessien approximé \mathbf{B}_a requiert la résolution de problèmes directs supplémentaires. Le terme (i,j) de cette matrice peut s'écrire :

$$\mathbf{B}_{a_{m_i, m_j}} = \Re \left[\mathbf{x}^\dagger \left[\frac{\partial \mathbf{A}}{\partial m_i} \right]^\dagger \mathbf{A}^{-1} \tilde{\mathcal{P}}_{data}^t \mathbf{W}_d \mathbf{W}_d \tilde{\mathcal{P}}_{data} \mathbf{A}^{-1} \left[\frac{\partial \mathbf{A}}{\partial m_j} \right] \mathbf{x} \right], \quad (2.19)$$

où nous pouvons voir que la matrice ($\mathbf{A}^{-1} \tilde{\mathcal{P}}_{data}$) est requise, équivalente aux fonctions de Green issues des positions de mesure des données (et éventuellement des directions de mesure pour des capteurs directionnels). La construction de \mathbf{B}_a requiert donc la résolution de n_{data} problèmes directs supplémentaires.

Une estimation du Hessien approximé, nommé Pseudo-Hessien \mathbf{B}_p , fut proposée par Shin *et al.* (2001) afin de calculer le Hessien sans surcoût par rapport au gradient. Son expression est donnée par :

$$\mathbf{B}_{p_{m_i, m_j}} = \Re \left[\mathbf{x}^\dagger \left[\frac{\partial \mathbf{A}}{\partial m_i} \right]^\dagger \left[\frac{\partial \mathbf{A}}{\partial m_j} \right] \mathbf{x} \right]. \quad (2.20)$$

2.3 Résolution numérique du problème d'optimisation locale

En considérant qu'un modèle *a priori* permet de localiser le modèle \mathbf{m} dans le bassin d'attraction du minimum global de la fonction coût $\mathcal{C}(\mathbf{m})$, une méthode d'optimisation locale doit être mise en oeuvre afin de converger vers ce minimum en un nombre minimal d'itérations. Les méthodes locales utilisent toujours l'information de la direction du gradient pour converger. En effet, la direction opposée au gradient donne la direction de plus grande pente de la fonction coût, à la position locale \mathbf{m} .

2.3.1 Méthode du Gradient

La méthode la plus simple est la méthode du gradient. Cette méthode fait l'hypothèse que le Hessien \mathbf{B} des équations normales (2.8) est un scalaire, permettant de résoudre facilement le système :

$$\Delta \mathbf{m} = -\alpha \mathcal{G}. \quad (2.21)$$

Cette méthode ne requiert aucune estimation du Hessien, n'entraînant pas de surcoût de calcul par rapport à la construction du gradient. Seul le scalaire α , appelé *pas de descente*, est à déterminer (*cf.* partie 2.3.2). Cette méthode converge théoriquement de façon linéaire. En pratique, il apparaît que, pour le problème d'inversion des formes d'onde, cette méthode ne permet pas toujours de converger dans le minimum global pour l'ensemble des paramètres. En particulier pour les acquisitions de surface, les paramètres localisés en profondeur souffrent des effets de l'atténuation géométrique, présents dans le gradient et non corrigés (Pratt *et al.*, 1998).

2.3.2 Le pas de descente α

Le pas de descente α est généralement déterminé par recherche linéaire afin d'obtenir le modèle $\mathbf{m}^{(k)}$ qui minimise ou qui assure une décroissance suffisante de la fonction coût. On

peut formaliser mathématiquement les propriétés que doit assurer le pas de descente α choisi, sous formes des « conditions de Wolf » ou des « conditions de Golstein » (Nocedal & Wright, 1999).

On peut obtenir la valeur du pas de descente à partir d'une approximation linéaire (Tarantola, 1987) :

$$\alpha = \frac{\mathcal{G}^t \mathcal{G}}{\mathcal{G}^t \mathcal{G} + \mathcal{G}^t \mathbf{J}^t \mathbf{J} \mathcal{G}}, \quad (2.22)$$

où le terme $\mathbf{J}\mathcal{G}$ est généralement calculé par différences finies d'ordre 1 en faisant une résolution du problème direct supplémentaire en perturbant le modèle d'un petit terme ϵ :

$$\mathbf{J}\mathcal{G} = \frac{1}{\epsilon} (\mathbf{d}_{calc}(\mathbf{m} + \epsilon \mathcal{G}) - \mathbf{d}_{calc}(\mathbf{m})). \quad (2.23)$$

Cette approximation de α nécessite tout de même de vérifier que le modèle $\mathbf{m}^{(k)}$ obtenu minimise bien la fonction coût par rapport à $\mathbf{m}^{(k-1)}$.

Pour l'inversion de plusieurs classes de paramètres de natures différentes, Sambridge *et al.* (1991) ont développé une méthode de sous-espaces qui permet de déterminer les pas de descente à appliquer à chacune des classes.

D'autres méthodes de recherche linéaire sont basées sur des interpolations quadratiques ou cubiques de la fonction coût dans la direction de recherche (Nocedal & Wright, 1999). Ainsi, pour une interpolation quadratique, on cherchera un couple de pas de descente (α_1, α_2) tel que :

$$\begin{aligned} \mathcal{C}(\mathbf{m}^{(k-1)} + \alpha_1 \Delta \mathbf{m}) &< \mathcal{C}(\mathbf{m}^{(k-1)}) \\ \mathcal{C}(\mathbf{m}^{(k-1)} + \alpha_2 \Delta \mathbf{m}) &> \mathcal{C}(\mathbf{m}^{(k-1)} + \alpha_1 \Delta \mathbf{m}). \end{aligned}$$

L'interpolation parabolique passant par les trois couples $(0, \mathcal{C}(\mathbf{m}^{(k-1)}))$, $(\alpha_1, \mathcal{C}(\mathbf{m}^{(k-1)} + \alpha_1 \Delta \mathbf{m}))$ et $(\alpha_2, \mathcal{C}(\mathbf{m}^{(k-1)} + \alpha_2 \Delta \mathbf{m}))$ permet ainsi d'obtenir le minimum de la parabole de l'intervalle $[0, \alpha_2]$, donnant le pas de descente optimum α_{opt} .

2.3.3 Le Gradient Conjugué

La méthode du gradient conjugué est une variante de la méthode du gradient. Cette méthode se révèle très efficace pour la résolution de systèmes linéaires. Pour notre cas d'inversion non-linéaire, la variante non-linéaire du gradient conjugué est utilisée.

L'idée générale de la méthode est la suivante : pour une itération k , au lieu de se diriger dans la direction de plus forte descente donnée par $-\mathcal{G}_{\mathbf{m}}^{(k-1)}$, le gradient conjugué combine la direction courante $-\mathcal{G}_{\mathbf{m}}^{(k-1)}$ et les directions des itérations précédentes $\mathbf{p}^{(k-2)}$. On peut montrer que les directions de descente successives dans la méthode du gradient sont orthogonales, entraînant une convergence lente (Figure 2.2(a)). Le gradient conjugué permet de combiner ces directions et accélère souvent la convergence de façon notable (Figure 2.2(b)). La combinaison linéaire des directions de descente est donnée par :

$$\mathbf{p}^{(k)} = \mathcal{G}_{\mathbf{m}}^{(k-1)} + \beta^{(k)} \mathbf{p}^{(k-2)}. \quad (2.24)$$

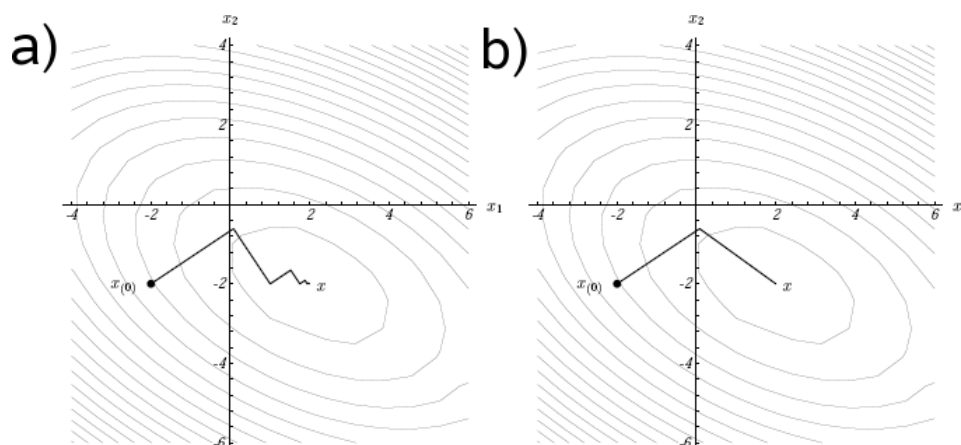


FIG. 2.2 – Schémas de convergence pour (a) la méthode du gradient descendant dans la direction de plus grande pente, et (b) la méthode du gradient conjugué combinant les directions de descente et accélérant la convergence.

Il existe diverses manières de combiner les directions de descente et donc de déterminer le coefficient $\beta^{(k)}$. Pour le problème d'inversion de formes d'ondes, le coefficient de Polak & Ribière (1969), exprimé par

$$\beta^{(k)} = \frac{\left(\mathcal{G}_{\mathbf{m}}^{(k-1)} - \mathcal{G}_{\mathbf{m}}^{(k-2)}\right)^t \mathcal{G}_{\mathbf{m}}^{(k-1)}}{\mathcal{G}_{\mathbf{m}}^{(k-2)t} \mathcal{G}_{\mathbf{m}}^{(k-2)}}, \quad (2.25)$$

est généralement utilisé (Mora, 1987; Tarantola, 1987; Crase *et al.*, 1990).

2.3.4 Méthode de Newton et Gauss-Newton

La méthode de Newton utilise le Hessien complet \mathbf{B} (équation (2.7)) tandis que la méthode de Gauss-Newton ne prend en compte que le Hessien approximé \mathbf{B}_a (équation (2.19)). Ces méthodes ont des convergences théoriques quadratiques, mais requièrent de calculer les problèmes directs supplémentaires qu'elles nécessitent. Par ailleurs, elles doivent former et résoudre le système linéaire de grande taille des équations normale (2.8). A cause de leur coût de calcul prohibitif, ces méthodes ne sont généralement pas considérées pour des problèmes réalistes de grande taille, malgré les informations importantes pour la convergence contenues dans le Hessien.

Notons tout de même que ces méthodes ont été décrites par Akcelik (2002); Askan *et al.* (2007); Epanomeritakis *et al.* (2008) et appliquées pour l'imagerie d'un cas 2D synthétique représentatif de la San Fernando Valley à partir d'ondes SH. Le système d'équations normales est résolu, à chaque itération non-linéaire d'inversion des formes d'ondes, par une méthode de gradient-conjugué linéaire préconditionnée, sans construction explicite de la matrice \mathbf{B} .

2.3.5 Méthodes de Quasi-Newton

Les méthodes de quasi-Newton reposent sur des estimations « économiques » du Hessien ou de son inverse, qui permettent à la fois de tirer bénéfice des informations contenues

dans le Hessien, sans payer le prix d'une véritable itération de Newton ou de Gauss-Newton. Ainsi, nous pouvons citer l'algorithme Broyden-Fletcher-Goldfarb-Shanno (BFGS) (Nocedal & Wright, 1999) qui permet d'estimer le Hessien à partir des gradients et des modèles des itérations précédentes, augmentant ainsi la précision du Hessien estimé avec les itérations. Nocedal (1980) créa une version moins coûteuse avec l'algorithme L-BFGS, en ne gardant en mémoire qu'un nombre limité de vecteurs. L'algorithme, sous forme de deux boucles récursives, permet de ne pas calculer explicitement l'estimation du Hessien, qui serait coûteux à stocker en mémoire, mais calcule directement le produit $\mathbf{B}^{-1}\mathcal{G}$ à partir d'additions, de soustractions et de produits scalaires de vecteurs. Ainsi, L-BFGS est très compétitif en temps de calcul et mémoire utilisée, comparé aux méthodes de gradient de gradient conjugué.

La méthode L-BFGS sera utilisée dans les applications des chapitres 3, 4 et 5, et sera plus amplement détaillée dans les parties 3.4.4 et 4.1.6.6.

2.3.6 Préconditionnement

Afin de tirer bénéfice du coût de calcul attractif des méthodes de gradient et gradient conjugué, tout en prenant en compte certaines informations du Hessien, certains auteurs ont proposé de prendre en compte uniquement la diagonale du Hessien approximé (Pratt *et al.*, 1998; Operto *et al.*, 2006) ou du pseudo-Hessien (Shin *et al.*, 2001). Le Hessien est effectivement à structure diagonale dominante et l'approximation diagonale se révèle judicieuse car sous cette forme, son inverse est calculé de façon triviale. Ainsi, les termes diagonaux du Hessien peuvent être pris en compte pour corriger efficacement le gradient des effets de l'atténuation géométrique présents dans les champs aux dérivées partielles, à un coût numérique abordable. De même, l'estimation diagonale du Hessien peut servir d'estimation initiale dans l'algorithme L-BFGS.

Certains auteurs (Hu *et al.*, 2009a) préconisent de prendre en compte les termes diagonaux du Hessien, mais également les termes les plus proches afin de tirer bénéfice de ces termes hors-diagonaux. Cependant, cette approche nécessite de résoudre le système d'équations normales avec ce Hessien incomplet qui peut être coûteux, surtout en 3D.

2.3.7 Régularisations

Le problème d'inversion des formes d'ondes est un problème inverse mal-posé et mal-contraint. Ainsi, il arrive très souvent que des artefacts (généralement hautes fréquences) apparaissent dans le vecteur perturbation. Il est donc indispensable d'appliquer des régularisations au problème afin de minimiser l'impact de ces artefacts.

On peut définir une fonctionnelle augmentée d'un terme de régularisation sur le modèle (Tarantola, 1987) :

$$\mathcal{C}(\mathbf{m}) = \frac{1}{2}\Delta\mathbf{d}^\dagger\mathbf{W}_d\Delta\mathbf{d} + \frac{1}{2}\epsilon(\mathbf{m} - \mathbf{m}_{prior})^\dagger\mathbf{W}_m(\mathbf{m} - \mathbf{m}_{prior}), \quad (2.26)$$

où la matrice \mathbf{W}_m est un opérateur de pondération dans l'espace des modèles (inverse de la matrice de covariance dans l'espace des modèles) (Tarantola, 1987), \mathbf{m}_{prior} est un modèle de référence à déterminer et le terme ϵ permet de pondérer les poids des deux termes de la fonction.

En géophysique, où les structures à imager sont relativement lisses, on injecte généralement dans la matrice \mathbf{W}_m un opérateur amplifiant les contrastes dans les structures tel qu'un opérateur différentiel en espace d'ordre 1 ou 2 (Press *et al.*, 2007). Ce type de régularisation est

généralement appelé *régularisation de Tikhonov*. En minimisant la fonction coût, on minimise ainsi les contrastes et cet opérateur maximise le lissage du modèle lorsque il est dérivé dans le gradient et le Hessien.

De manière plus pragmatique, de opérateurs de lissage peuvent être appliqués au vecteur perturbation $\Delta \mathbf{m}$ sous forme de filtres spatiaux Gaussiens adaptatifs (Ravaut *et al.*, 2004; Operto *et al.*, 2006) ou de filtres dans le domaine des nombres d'ondes spatiaux (Sirgue, 2003).

Certains auteurs ont également introduit des opérateurs de régularisation visant au contraire à préserver les contrastes dans les images, principalement basés sur les outils développés en traitement d'image et par la communauté électromagnétique qui image des structures très contrastées. Ainsi, Akcelik (2002) utilise une régularisation de type *Total Variation* basée sur une norme L_1 dans l'espace des modèles pour imager des milieux synthétiques simples. Cependant, Hu *et al.* (2009b) montrent que, sur des milieux complexes assez lisses, ces régularisations visant à accentuer les contrastes donnent de moins bon résultats que celles basées sur Tikhonov.

Notons enfin que le terme ϵ de l'équation (2.26) peut être relié à la méthode d'optimisation de Levenberg-Marquard qui modifie le Hessien tel que :

$$\mathbf{B}_{lm} = \mathbf{B} + \epsilon \mathbf{I}, \quad (2.27)$$

où \mathbf{B}_{lm} est le Hessien de Levenberg-Marquard et \mathbf{I} est la matrice identité. Cette méthode a pour but de minimiser le poids des petites valeurs du Hessien quand celui-ci est symétrique défini positif, ou plus radicalement de le rendre symétrique défini positif lorsque certaines valeurs du Hessien deviennent trop faibles ou nulles. Cette méthode peut être également liée aux variantes simples des méthodes de régions de confiance (Nocedal & Wright, 1999) pour lesquelles le terme ϵ est adapté pour chaque paramètre en fonction de la perturbation maximale que l'on peut admettre. En diminuant le terme ϵ , on augmente le poids du Hessien estimé, augmentant la confiance que l'on donne au paramètre. En augmentant ϵ , on diminue le poids du Hessien estimé, ce qui contraint la perturbation à rester faible. Ce terme ϵ peut être comparé à un facteur de préblanchiment pour la déconvolution en traitement du signal.

2.3.8 Exemple

Afin d'illustrer les performances des méthodes d'optimisation présentées précédemment, une application est présentée sur la fonction analytique non-convexe de Rosenbrock, très souvent utilisée pour ce type de test. La Figure 2.3 représente cette fonction pour la version à deux variables dont le minimum global est localisé en $(1, 1)$. Pour la fonction à 2 variables, la fonction est non-convexe mais ne possède qu'un seul minimum global. Pour le cas à plus de 2 variables, des minimums secondaires existent. Le Tableau 2.1 illustre le nombre d'itérations nécessaires aux différents algorithmes pour converger vers le minimum global de la fonction (tolérance de convergence fixée à 10^{-6}) prise pour 2 et 50 variables. Le point de départ pour l'optimisation est fixé pour les deux tests au point $(4, \dots, 4)$. L'estimation du pas de descente est assuré par une recherche d'optimum par interpolation parabolique. Les résultats montrent clairement la hiérarchie des performances des différents algorithmes. En particulier, le gradient conjugué et surtout le L-BFGS montrent leur intérêt face à la méthode du gradient, alors que le coût de calcul est sensiblement équivalent dans le contexte d'inversion des formes d'ondes. La méthode de Newton, très performante, requiert cependant le calcul explicite du Hessien et de résoudre les équations normales (2.8).

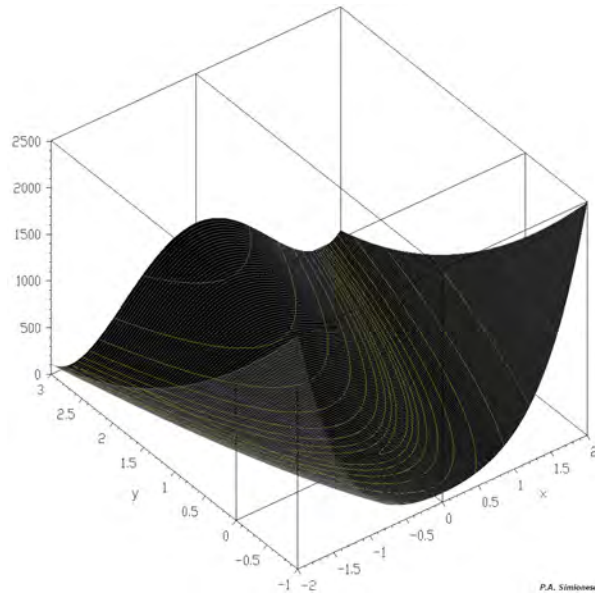


FIG. 2.3 – Fonction non-convexe de Rosenbrock à 2 variables. Le minimum global est localisé en (1, 1).

TAB. 2.1 – Performances des algorithmes d’optimisation pour les fonctions de Rosenbrock à 2 et 50 variables

Test	Algorithme	Nombre d’itérations
Rosenbrock 2 variables	Gradient	1 934
	Gradient Conjugué	458
	L-BFGS (3 mémoires)	29
	Newton	14
Rosenbrock 50 variables	Gradient	11 809
	Gradient Conjugué	3 479
	L-BFGS (3 mémoires)	168
	Newton	19

2.4 Le modèle initial

2.4.1 Modèle cinématiquement compatible et saut de phase

Ecrite sous une forme linéarisée, la méthode d’inversion des formes d’ondes nécessite de partir d’un modèle initial \mathbf{m}_0 localisé dans le bassin d’attraction de la fonction coût. Ce bassin est d’autant plus large que la fréquence inversée est petite (Mulder & Plessix, 2008). Pour les acquisitions de sismique active, les fréquences minimales générées et enregistrées avec un rapport signal sur bruit acceptable, se situent de nos jours vers 2-4 Hz, rendant impossible de commencer une inversion de formes d’ondes sans modèle suffisamment précis. On peut relier le bassin d’attraction de la fonction coût au concept de modèle cinématiquement compatible. Ainsi, si le modèle initial procure un décalage de phase inférieure à une demi-période de signal,

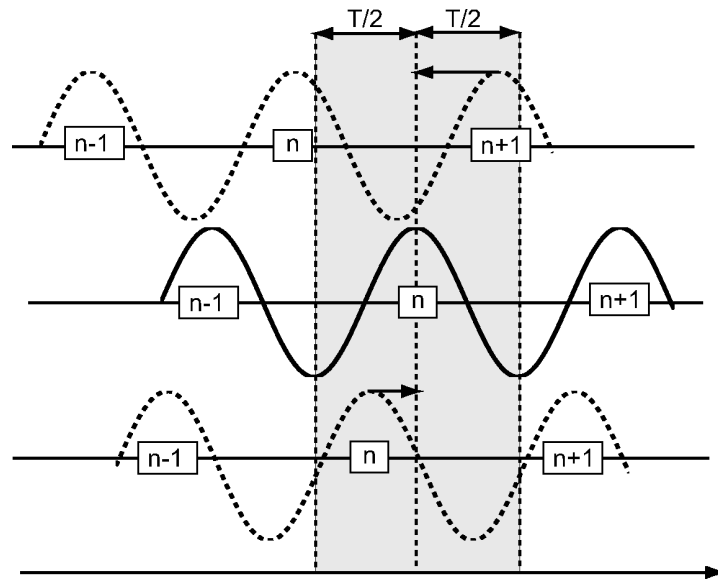


FIG. 2.4 – Schéma du principe du saut de phase. Si le déphasage entre les signaux reste inférieur à la demi-période $T/2$, l’algorithme linéarisé ajustera les bonnes phases. Dans le cas contraire, l’algorithme ajustera des arches de signaux déphasées et divergera dans un minimum secondaire.

pour la fréquence considérée, alors l’approche linéarisée pourra ajuster les signaux observés et calculés sans ambiguïté de phase (Figure 2.4). Si le décalage est supérieur, l’algorithme cherchera à ajuster des signaux avec un déphasage d’une ou plusieurs périodes, entraînant la convergence de l’algorithme vers un minimum secondaire (Sirgue, 2003). Ce phénomène s’applique pour toutes les fréquences considérées dans l’inversion et conditionne des règles de choix de fréquence (*cf.* partie 2.2.3). Concernant le modèle initial, il doit être suffisamment précis pour respecter cette compatibilité cinématique par rapport à la première fréquence que l’on peut considérer dans les données. Je présente ci-dessous des méthodes d’obtention d’un modèle initial pour des applications d’exploration.

2.4.2 Tomographie des temps de première arrivée

La tomographie des temps de première arrivée est très largement utilisée en sismologie pour produire des images lisses de la Terre (Nolet, 1987; Hole, 1992; Zelt & Barton, 1998). En utilisant les temps pointés des premières arrivées des ondes, le problème inverse se formalise de manière non-linéaire dans lequel la matrice des dérivées de Fréchet est généralement explicitement construite, même si le formalisme de l’état-adjoint permet de s’en affranchir au profit de l’efficacité (Taillandier *et al.*, 2009).

La résolution de la tomographie des temps de première arrivée est de l’ordre de la première zone de Fresnel (Williamson, 1991). Comme cette approche n’exploite que les temps de première arrivée, de très grands offsets sont nécessaires en sismique de surface pour garantir une couverture spatiale des rais dans la partie profonde du modèle. Pour une application sur le modèle synthétique du BENCHMARK BP 2004, Brenders & Pratt (2007c) montrent que

l'utilisation d'une tomographie des temps de première arrivée avant une inversion de formes d'ondes acoustique nécessite de basses fréquences (0.5 Hz) et un dispositif d'acquisition de 16 km. Malgré ces conditions, seule la partie superficielle du modèle peut être imagée. Les applications sur données réelles présentées dans Ravaut *et al.* (2004); Operto *et al.* (2006); Malinowsky & Operto (2008) montrent le processus d'imagerie commençant avec la tomographie des temps de première arrivée puis l'inversion de formes d'ondes acoustiques.

Les limitations de cette méthode sont les suivantes : la nécessité de pointer les temps, ce qui peut se révéler délicat lorsque le modèle présente des anomalies de vitesses lentes (Prioux *et al.*, 2009); l'obtention du modèle V_P seul dans le cas de sources explosives; la nécessité d'utiliser des sources générant des ondes S pour imager V_S ; et la nécessité d'utiliser des dispositifs très grands offsets pour obtenir une couverture des rais suffisante dans la partie profonde du modèle.

2.4.3 L'inversion dans le domaine de Laplace

Récemment proposée par Shin & Cha (2008), l'inversion dans le domaine de Laplace se veut une méthode de choix pour la construction de modèles de départ pour l'inversion des formes d'ondes. La méthode utilise comme données, les champs d'ondes aux récepteurs transformés dans le domaine de Laplace. Cette transformation est équivalente à la fréquence nulle de la transformée de Fourier d'un signal temporel amorti par une exponentielle décroissante :

$$\begin{aligned} L(\gamma) &= \int_0^{+\infty} f(t) \exp^{-\gamma(t)} dt \\ \Leftrightarrow F(\omega + i\gamma) &= \int_{-\infty}^{+\infty} f(t) \exp^{-\gamma(t)} \exp^{-i\omega t} dt \quad \text{avec} \begin{cases} \omega = 0 \\ f(t) = 0 \quad \forall t < 0, \end{cases} \end{aligned} \quad (2.28)$$

où $L(\gamma)$ est la transformée de Laplace du signal temporel $f(t)$ pour le coefficient de Laplace γ . Le champ d'ondes $L(\gamma)$ dans le domaine de Laplace est à valeur réelle.

Cette méthode permet de reconstruire les grandes longueurs d'ondes du milieu en utilisant les formes d'ondes (théoriquement les nombres d'ondes nuls). Cependant, les forts coefficients d'amortissement γ des exponentielles mis en jeu pour la stabilité de transformée de Laplace engendrent une focalisation des données utilisées proche de la première arrivée (Shin & Cha, 2008). Utilisant une information très proche de la tomographie des temps de première arrivée, on ne peut donc pas s'attendre à une qualité d'image très différente, d'autant que, si la méthode n'utilise pas explicitement de pointés de la première arrivée, elle nécessite tout de même de la localiser précisément pour appliquer un *mute* avant cette arrivée, afin de supprimer le bruit pour le calcul de la transformée de Laplace de manière stable.

Cette méthode possède l'avantage de pouvoir s'utiliser très facilement à partir d'un algorithme d'inversion de formes d'ondes en fréquence, en utilisant une fréquence à partie réelle nulle et à partie imaginaire non-nulle.

Des résultats impressionnants ont pu être obtenus avec cette méthode pour la reconstruction du paramètre V_P sous l'approximation acoustique, pour des applications synthétiques par Shin & Cha (2008); Lee *et al.* (2008); Shin & Ha (2008); Pyun *et al.* (2008a) et sur des données réelles du Golf de Mexico par Shin & Cha (2008); Shin & Ha (2008) dans une zone avec un corps de sel. Pyun *et al.* (2008b) ont montré une application 3D élastique de l'inversion en domaine de Laplace sur le modèle synthétique SEG/EAGE Salt. Cependant, peu de détails ont

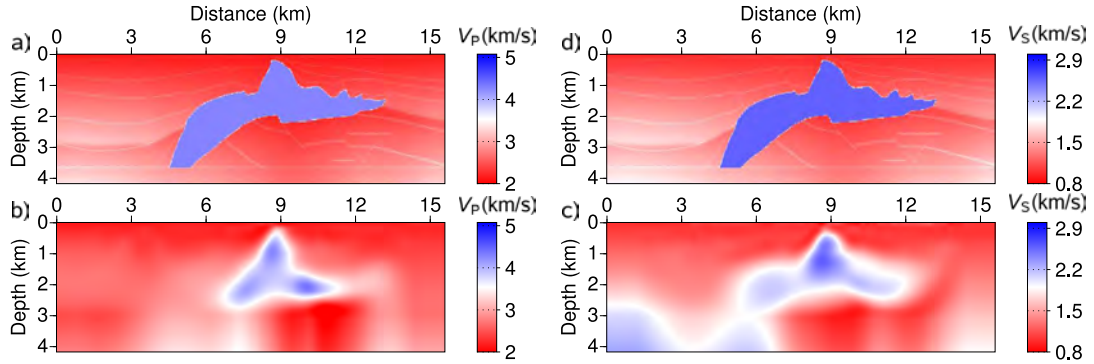


FIG. 2.5 – Résultats d’inversion élastique dans le domaine de Laplace pour une section 2D du modèle SEG/EAGE Salt. Les panneaux (a) et (c) représentent les vrais modèles V_P et V_S . Les panneaux (b) et (d) représentent les modèles reconstruits pour V_P et V_S .

été publiés sur la robustesse de l’approche pour la reconstruction de deux paramètres de vitesse en utilisant uniquement une information d’amplitude localisée proche de la première arrivée.

La Figure 2.5 montre les résultats d’un test d’inversion élastique dans le domaine de Laplace que j’ai mené dans une coupe 2D du modèle SEG/EAGE Salt. Le modèle V_S a été généré avec un coefficient de Poisson constant de 0.24 (modèle terrestre sans eau). Les données de référence ont été calculées avec le même algorithme que celui utilisé pour l’inversion. L’inversion est menée en supposant que la signature de la source est connue. Les modèles de départ sont des gradients de vitesses pour les deux modèles. 88 points sources verticaux sont enregistrés par 88 capteurs deux composantes sur une distance de 15 km. La surface libre est considérée dans ce test. 27 coefficients de Laplace γ ont été inversés séquentiellement de façon décroissante entre 10 et 0.4. Les champs d’ondes sont très vite amortis en amplitude avec l’offset dans le domaine de Laplace, c’est pourquoi Shin & Cha (2008) utilisent une fonction coût logarithmique (*cf.* partie 2.5.1) qui compense cet effet de décroissance. Dans ce test, j’ai utilisé une fonction coût L_2 normalisée par les données observées (Mulder & Plessix, 2008) :

$$\mathcal{C}(\mathbf{m}) = \frac{1}{2} \sum_{i=1}^{n_{data}} \frac{\Delta d_i^2}{d_{obs_i}^2}, \quad (2.29)$$

qui compense l’effet de la forte décroissance d’amplitude, comme le fait la fonction logarithmique. Les résultats montrent des modèles encourageants pour ce cas synthétique idéal. Cependant, d’autres tests ont montré une grande sensibilité de la méthode au choix des coefficients de Laplace et au type de source. La robustesse de l’approche pour des applications plus complexes et des données réelles déterminera si cette méthode peut être envisagée pour construire des modèles de départ pour plusieurs paramètres.

2.4.4 Inversion de la phase de première arrivée

En utilisant toujours l’information de la première arrivée, un algorithme d’inversion de formes d’ondes en fréquence peut être utilisé pour inverser la phase de la première arrivée en utilisant des fréquences complexes. La partie réelle représente la fréquence de propagation

tandis que la partie imaginaire représente le coefficient de l'exponentielle décroissante appliquée en temps (Min & Shin, 2006; Effelsen, 2009). Cette approche est très proche de la tomographie des temps de première arrivée, utilisant une information quasi-identique. Notons tout de même que la prise en compte d'une fréquence réelle permet de lier cette méthode à la tomographie par fréquence finie (Montelli *et al.*, 2004).

2.4.5 La stéréo-tomographie

L'imagerie par stéréo-tomographie est fondée sur l'inversion des temps d'arrivée d'événements localement cohérents, pointés sur des collections de sismogrammes triés en sources et récepteurs communs (Lambaré, 2008). Cette méthode permet de prendre en compte les arrivées réfléchies, réfractées, qu'elles soient de polarisations P-P ou P-S. Ainsi, des images des modèles de vitesse d'ondes P et S peuvent être construits. Des applications sur données synthétiques et réelles sont présentées dans Billette & Lambaré (1998); Alerini *et al.* (2002); Billette *et al.* (2003).

2.4.6 Conclusions

La construction d'un modèle de départ de qualité pour l'inversion des formes d'onde est un problème crucial qui détermine la réussite ou l'échec du processus d'imagerie. Les méthodes mentionnées ci-dessus sont des outils possibles, ayant chacun des avantages et des inconvénients. Notons que, dans la perspective d'inversion de formes d'ondes élastiques, seules une tomographie des premières arrivées d'ondes S (si la source génère des ondes S), l'inversion dans le domaine de Laplace (si elle se révèle stable et robuste) et la stéréo-tomographie permettent d'obtenir un modèle de départ pour les vitesses V_S .

La problématique de la détermination du modèle de départ est un problème ouvert qui motive, à l'heure actuelle, de nombreux travaux dans les équipes de recherche académiques et industrielles.

2.5 Les alternatives à la norme L_2

La fonction coût classiquement et majoritairement utilisée en inversion des formes d'ondes est basée sur la minimisation de la norme L_2 de la différence des données (équation (2.2)). Comme mentionné auparavant dans la partie 2.1.1, ce choix est motivé par des hypothèses statistiques sur les données. Il existe toutefois des alternatives à cette fonctionnelle que je vais présenter.

2.5.1 Fonction Logarithmique : Amplitude vs Phase - lien avec Rytov

Shin & Min (2006) ont introduit la fonction L_2 logarithmique basée sur la minimisation de la fonctionnelle :

$$\mathcal{C}(\mathbf{m}) = \frac{1}{2} \sum_{i=1}^{n_{data}} \left[\ln \frac{d_{calc_i}}{d_{obs_i}} \right]^2. \quad (2.30)$$

Cette fonction permet de séparer facilement les empreintes de l'amplitude et de la phase dans les données car le logarithme d'un nombre complexe $z = |z| \exp^{i\phi}$ est défini par :

$$\ln(z) = \ln(|z|) + i\phi. \quad (2.31)$$

En prenant en compte uniquement la partie réelle du logarithme complexe, on ne considère que l'amplitude. En prenant en compte uniquement la partie imaginaire, seule la phase est considérée.

Le gradient de cette fonction est donné par :

$$\begin{aligned} \mathcal{G}_{m_l} &= \Re \left[\mathbf{x}^t \left[\frac{\partial \mathbf{A}}{\partial m_l} \right]^t \mathbf{A}^{-1} \tilde{\mathcal{P}}_{data} \mathbf{r} \right] \\ \text{avec } r_i &= \frac{\ln \left(\frac{d_{calc_i}}{d_{obs_i}} \right)^*}{d_{calc_i}}. \end{aligned} \quad (2.32)$$

Une étude complète de la sensibilité de l'inversion au champ d'ondes complet, à la phase seulement et à l'amplitude seulement fut présentée par Shin *et al.* (2007); Bednar *et al.* (2007); Pyun *et al.* (2007). Cette étude montre que l'inversion du champ complet et de la phase donne les meilleurs résultats, tandis que l'amplitude, lorsqu'elle est inversée seule, reste mal contrainte.

Notons tout de même que pour rester stable, la fonction logarithmique requiert quelques précautions lors de son implémentation à cause de l'utilisation de la fonction logarithme qui peut engendrer des *overflows/underflows* numériques (Kim *et al.*, 2008). De plus, des problèmes de déroulement de phase peuvent avoir lieu lors du calcul de la phase du résidu à partir des phases des données observées et calculées. Afin d'éviter ces problèmes, la fonction logarithme fait l'hypothèse que les phases des données observées et calculées ne sont pas déphasées de plus de 2π (Shin & Min, 2006). Cette hypothèse peut être reliée aux limites de la linéarisation qui imposent la compatibilité cinématique du modèle initial.

On peut faire le lien entre la fonction logarithmique et l'approximation de Rytov (Woodwards, 1992) qui sépare l'amplitude et la phase du champ. L'approximation de Rytov, équivalente à la fonction logarithmique de Shin & Min (2006), fut développée pour l'inversion des formes d'ondes en fréquence par Gelis *et al.* (2007). Les résultats de Gelis *et al.* (2007) ne montrèrent pas de supériorité notable de cette approximation face à la formulation classique, dénommée (à tort ?) par approximation de Born. Notons que la matrice des dérivées de Fréchet de chaque couple source/récepteur est équivalente pour la formulation classique (équation (2.14)) et pour la formulation logarithmique/Rytov (équation (2.32)). Cette équivalence fut également notée par Woodward (1992) pour l'analyse des *wavepaths* calculés par équation d'onde asymptotique, construits par les approximations de Born et Rytov au premier ordre. La différence, essentielle, entre les deux formulations de gradient réside dans la façon de combiner les champs aux dérivées partielles des différents couples source/récepteur à travers le terme $\Delta \mathbf{d}^*$ pour la formulation classique et $\ln(d_{calc_i}/d_{obs_i})^*/d_{calc_i}$ pour la formulation logarithmique, résultant sur des gradients différents.

2.5.2 Fonctionnelles robustes

L'hypothèse de statistiques gaussiennes dans les données peut ne pas être validée dans tous les cas. Si les données ne respectent plus les hypothèses qui ont conduit au choix de la norme

L_2 , alors cette dernière peut souffrir d'un manque de robustesse, n'étant plus représentative du vecteur résidu $\Delta \mathbf{d}$. Nous pouvons penser à des erreurs isolées de fortes amplitudes dans les données par exemple. Si ces larges erreurs ne sont pas supprimées avant l'inversion, alors elle engendreront de forts résidus qui seront amplifiés dans la fonctionnelle L_2 et son gradient. Des normes et critères de minimisation alternatifs à la norme L_2 ont donc été introduits pour palier ces inconvénients.

La norme L_1 , non basée sur des hypothèses statistiques gaussiennes, fut ainsi introduite par Tarantola (1987); Crase *et al.* (1990) pour l'inversion des formes d'ondes en domaine temporel, et se révéla très robuste au bruit dans les données. De même, nous pouvons noter les critères de Cauchy (Crase *et al.*, 1990), *sech* (Crase *et al.*, 1990; Monteiller *et al.*, 2005), de Huber (Huber, 1973; Guitton & Symes, 2003; Ha *et al.*, 2009), Hybride L_1/L_2 (Bube & Langan, 1997; Bube & Nemeth, 2007), qui cherchent tous à combiner le comportement de la norme L_2 pour les petits résidus, et celui de la norme L_1 pour les grands résidus, séparés par une zone de transition définie par un seuil à déterminer.

Ces différents critères et normes ont été marginalement testés pour le problème d'inversion des formes d'ondes en fréquence. Nous développerons dans le chapitre 5 une étude de sensibilité au bruit en utilisant différents critères de minimisation pour l'inversion des formes d'ondes élastiques.

2.6 L'inversion multiparamètres

Depuis les années 90, la méthode d'inversion des formes d'ondes a principalement été développée et utilisée pour imager le modèle de vitesse d'ondes P en considérant l'approximation acoustique pour la propagation des ondes. En effet, ce paramètre possède la signature la plus dominante dans les données pour de nombreuses configurations d'acquisitions, et l'approximation acoustique, moins coûteuse à modéliser, se révèle efficace lorsque les données sont correctement prétraitées. Ainsi les travaux de Pratt & Worthington (1990); Pratt *et al.* (1998); Pratt (1999); Ravaut *et al.* (2004); Sirgue & Pratt (2004); Operto *et al.* (2006); Brenders & Pratt (2007a,b); Sirgue *et al.* (2007); Ben-Hadj-Ali *et al.* (2008); Brenders *et al.* (2009); Plessix (2009); Sirgue *et al.* (2009) par exemple, s'attachent à reconstruire le modèle V_P sous l'approximation acoustique, à partir de données synthétiques ou réelles, en 2D ou 3D.

Cependant, même si son intérêt est majeur, le modèle V_P utilisé seul ne permet pas toujours l'interprétation fine et fiable des structures complexes du sous-sol. Les modèles de densité, de vitesse de propagation d'ondes S, et d'atténuation, permettent de confirmer ou d'infirmer certaines hypothèses émises à partir de V_P .

Par exemple, le paramètre de densité, couplé au paramètre V_P donne une information primordiale sur l'impédance du milieu, largement utilisé et interprété pour la détection de réservoirs dans l'industrie pétrolière. Le modèle de vitesse d'onde S, combiné à celui de V_P permet d'obtenir une estimation de la porosité et de saturation des milieux, et permet de calculer le coefficient de Poisson qui très influencé par la présence de liquides et de gaz. Ces informations permettent de caractériser des réservoirs pour l'exploration, d'en faire le monitoring dans le temps.

L'imagerie de plusieurs paramètres physiques par inversion des formes d'ondes est cependant peu répandue. Le frein majeur réside dans la grande difficulté associée au problème inverse

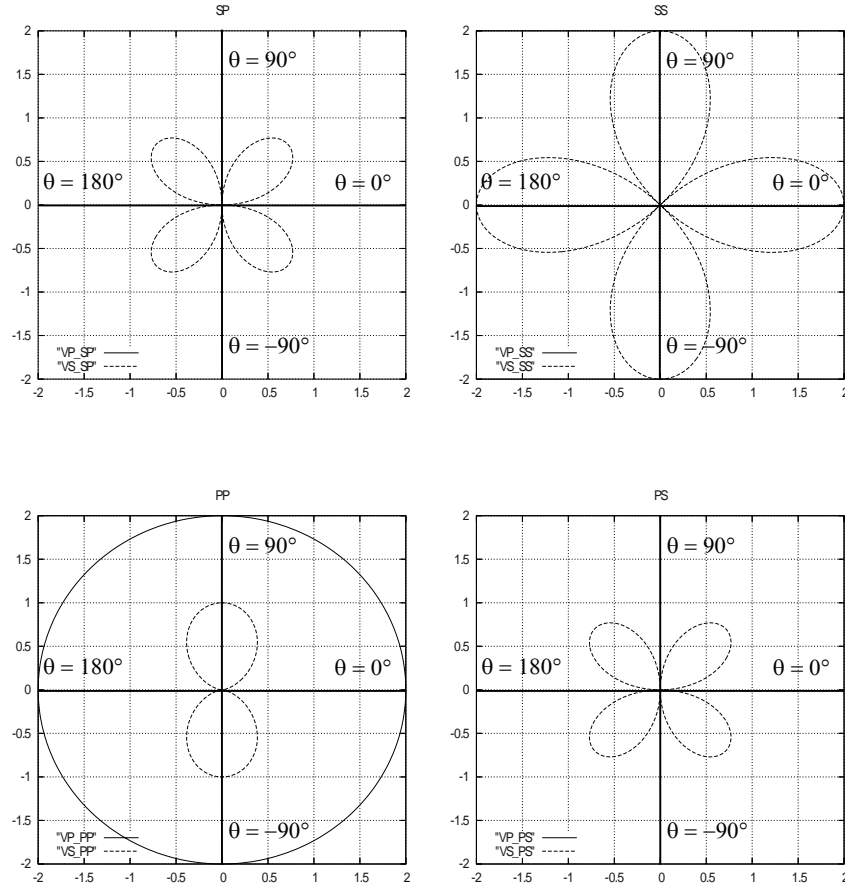


FIG. 2.6 – Diagrammes de rayonnement des diffractions P-P, P-S, S-P et S-S pour des anomalies de V_P (courbes continues) et de V_S (courbes discontinues) en fonction de l'angle de diffraction θ . Courbes calculées d'après Forgues (1996) pour $V_P = 2V_S$.

mal-posé. L'imagerie de V_P seul, sous l'approximation acoustique, rend le problème plus linéaire et plus facile à résoudre en supposant que les autres paramètres sont connus et invariants, ou directement liés à V_P . En considérant l'inversion simultanée et indépendante de plusieurs paramètres et/ou avec une physique de propagation plus complexe, on accentue la non-linéarité du problème inverse résultant (1) du couplage entre paramètres, (2) de la signature des paramètres dans les données, (3) des phénomènes de propagation plus complexes et (4) de la pondération des différentes classes de paramètres dans l'algorithme d'optimisation.

L'analyse des diagrammes de diffraction des différents paramètres (Tarantola, 1986; Forgues, 1996) montre ainsi la grande difficulté de discerner la densité et V_P pour des angles de diffraction faibles. De même, le facteur de qualité est très fortement lié à la vitesse de propagation rendant son inversion mal contrainte si le dispositif d'acquisition n'est pas adéquat (Ribodetti & Virieux, 1996; Ribodetti *et al.*, 2000; Mulder & Hak, 2009). Pour les paramètres élastiques, la géométrie de l'acquisition oriente le choix de la paramétrisation vers les impédances élastiques I_P et I_S pour les courts offsets, ou les vitesses V_P et V_S pour des grands offsets, afin d'obtenir le découplage maximal entre les paramètres (Tarantola, 1986; Forgues, 1996) (Figure 2.6).

En fonction des structures à imager et des données, le poids des différents paramètres dans

les données peut varier. Ainsi, pour l'inversion acoustique de V_P et de l'atténuation, Kamei & Pratt (2008) recommandent d'inverser d'abord V_P sans mettre à jour l'atténuation, la signature de V_P dominant souvent les données. Dans un second temps, l'inversion simultanée des deux paramètres permet de tirer bénéfices d'information plus « fines » d'amplitude dans les résidus une fois que le modèle V_P explique en grande partie la cinématique. Cette approche permet également de séparer les effets de l'atténuation extrinsèque relative à l'hétérogénéité du milieu et de l'atténuation intrinsèque. Pour l'inversion élastique, Tarantola (1986) recommande d'imager d'abord V_P ou I_P , le paramètre dominant, puis le paramètre lié aux vitesses d'ondes S (V_S ou I_S). Sears *et al.* (2008) ont appliqué judicieusement cette stratégie, couplé à des sélections de données par fenêtrage en temps de certaines phases, pour l'inversion de formes d'ondes temporelle en milieu marin synthétique. Cette stratégie s'est également révélée efficace pour le milieu marin Valhall présenté dans la partie 4.2. Par contre, comme nous le verrons dans la partie 4.1, lorsque des données terrestres sont considérées avec inversion des ondes de volume et des ondes de surface, les deux paramètres dominent les données et doivent être inversés simultanément.

La complexité des données accentue le caractère non-linéaire de l'inversion. Pour l'imagerie élastique de données synthétiques en milieu marin, Sears *et al.* (2008) montrent la forte non-linéarité du problème et utilise différents niveaux de régularisation sur les données et les paramètres imagés, afin d'utiliser les différentes conversions P-S. Gelis *et al.* (2007) montrent sur des exemples synthétiques canoniques, que la reconstruction de V_P et V_S , en utilisant de façon combinée les ondes de volume et les ondes de surface, donne des résultats assez mitigés. La cause est associée à la forte énergie propagée par les ondes de surface qui ne sondent qu'une profondeur limitée du milieu et qui domine la fonction coût et le gradient.

La pondération des différentes classes de paramètres reconstruits simultanément dans le processus d'optimisation accentue la non-linéarité du problème inverse. La mise à l'échelle du gradient est théoriquement assurée par la prise en compte du Hessien complet dans l'équation de Newton (si le problème est localement parabolique). Or, les méthodes de Newton et de Gauss-Newton ne sont généralement pas utilisées à cause de leur coût numérique (*cf.* partie 2.3.4). Pour palier ce problème, Sambridge *et al.* (1991) ont développé une méthode de sous-espaces qui permet de déterminer les pas de descente à appliquer à chaque gradient des différentes classes de paramètres. Cette méthode repose sur une méthode de Newton, dont les paramètres sont les pas de descentes de chaque classe. Le Hessien explicitement formé est une petite matrice carrée dont la taille est le nombre de classes de paramètres, elle est donc facilement inversible. Cependant, l'approche nécessite de calculer un champ direct supplémentaire par source et par classe de paramètre à chaque itération du processus itératif non-linéaire. Nous verrons dans la partie 3.4.4 que la méthode quasi-Newton L-BFGS, qui estime économiquement l'inverse du Hessien, permet de mettre à l'échelle les différentes classes de paramètres sans nécessiter de champs directs supplémentaires.

2.7 Conclusion

Dans ce chapitre, le formalisme du problème d'inversion linéarisé en fréquence a été présenté à partir de la minimisation de la norme aux moindres carrés du vecteur résidu. Les rôles respectifs du gradient et du Hessien ont été introduits, montrant d'une part la résolution attendue de la méthode et l'algorithme classiquement utilisé, basé sur un volume de données limité, et d'autre

part le rôle important du Hessien dans la convergence de la méthode. Les méthodes de résolution du problème d'optimisation locale ainsi que les méthodes de construction du modèle de départ ont été présentées. Enfin les critères alternatifs à la norme L_2 ont été introduit avant d'envisager les enjeux et les problèmes associés à l'inversion multiparamètres. Ce chapitre permet donc de définir une base de méthodes pour l'investigation des problématiques liées à l'inversion des formes d'ondes élastiques qui seront présentés dans les chapitres suivants.

Chapitre 3

Algorithmes parallèles, optimisation et performance

Sommaire

3.1	Résumé	118
3.2	Introduction	118
3.3	Problème direct	120
3.3.1	Théorie et discrétisation	120
3.3.2	Quelle interpolation ?	122
3.4	Problème inverse	125
3.4.1	Revue de la théorie	125
3.4.2	Algorithme d'inversion	128
3.4.3	Calcul du gradient	128
3.4.4	L-BFGS et inversion multiparamètres	130
3.5	Parallélisation et performances	133
3.5.1	Problème direct	133
3.5.2	Problème inverse parallèle	134
3.5.3	Deux niveaux de parallélisme	134
3.5.4	Performance	135
3.6	Conclusion	136

L'inversion des formes d'ondes est une méthode d'imagerie relativement coûteuse en moyens informatiques. Dans ce chapitre, je présente les aspects liés aux algorithmes et l'implémentation de la méthode sous la forme d'un article soumis pour publication dans la revue *Computers & Geosciences*. La résolution du problème direct sera d'abord introduite avec les coûts numériques respectifs des différentes interpolations possibles de la méthode GD. Nous verrons que la complexité et la configuration des cas d'études dirigent le choix de l'interpolation et du type de maillage. Ensuite, nous aborderons le problème inverse en nous focalisant sur l'implémentation de plusieurs niveaux hiérarchiques dans un schéma d'inversion multirésolutions, l'optimisation du calcul du gradient et la résolution du problème d'optimisation par la méthode L-BFGS permettant d'estimer économiquement le Hessien. Enfin, l'implémentation parallèle et les performances de l'algorithme global seront discutées. Nous nous attacherons particulièrement à

l'intérêt d'une implémentation avec deux niveaux de parallélisme, permettant pour certaines applications, d'économiser à la fois du temps de calcul et de la mémoire vive.

Two-dimensional frequency-domain visco-elastic full waveform inversion : parallel algorithms, optimization and performance

R. Brossier

article soumis pour publication dans la revue *Computers & Geosciences*

3.1 Abstract

Full waveform inversion (FWI) is an appealing seismic data-fitting procedure for the derivation of high-resolution quantitative models of the subsurface at various scales. Full modeling and inversion of visco-elastic waves from multiple seismic sources allows the recovering of different physical parameters, although it remains computationally challenging. An efficient massively parallel, frequency-domain FWI algorithm is implemented here on large-scale distributed-memory platforms for imaging two-dimensional visco-elastic media. The resolution of the elastodynamic equations, as the forward problem of the inversion, is performed in the frequency domain on unstructured triangular meshes, using a low-order finite element discontinuous Galerkin method. The linear system resulting from discretization of the forward problem is solved with a parallel direct solver. The inverse problem, which is presented as a linearized least-squares optimization problem, is solved in parallel with a quasi-Newton method, and this allows reliable estimation of multiple classes of visco-elastic parameters. Two levels of parallelism are implemented in the algorithm, based on message passing interfaces and multi threading, for optimal use of computational time and the core-memory resources available on modern distributed-memory multi-core computational platforms. The algorithm allows for imaging of realistic targets at various scales, ranging from near-surface geotechnic applications to crustal scale exploration.

Keywords

Seismic wave modeling ; finite element discontinuous Galerkin ; multi-parameter seismic imaging ; quasi-Newton optimization ; massively parallel computing

3.2 Introduction

Quantitative imaging of the subsurface is essential for geotechnic and civil engineering applications, for seismic hazard and natural resources exploration, and for our geodynamics knowledge. The physical properties of the subsurface are often estimated through seismic-wave analysis, with wave travel-times generally used to perform tomography of the Earth interior. In the 1980s, the full waveform inversion (FWI) method was introduced by Tarantola (1984) to exploit the complete information contained in the waveforms of seismograms, and to infer high-resolution models of the subsurface. FWI was originally developed in the time domain (Tarantola, 1984, 1987), and it has become tractable and popular in the frequency domain since the pioneering work of Pratt & Worthington (1990); Pratt (1990). FWI is currently a

field of active research, and particularly for active seismic surveys at various scales. Computationally efficient frequency-domain FWI was designed by limiting the inversion to a few discrete frequencies, taking advantage of the redundant wavenumber cover provided by wide-aperture surveys (Sirgue & Pratt, 2004; Brenders & Pratt, 2007a). Frequency-domain FWI potentially provides high-resolution quantitative images of physical parameters, although it suffers from two main difficulties.

The first of these difficulties relates to the computational cost of the forward problem; namely, the numerical resolution of the wave equation in heterogeneous media for multiple sources. In two dimensions, the forward problem can be solved efficiently in the frequency domain (Nihei & Li, 2007), leading to the resolution of a linear system per frequency, the right-hand side of which is the source. This system is solved efficiently for multiple sources with a direct solver, which performs one LU factorization of the matrix, followed by a forward and a backward substitution per source. For the acoustic wave equation, optimal finite-difference (FD) schemes have been designed for optimal use of the CPU core memory when used with direct solvers (Jo *et al.*, 1996; Hustedt *et al.*, 2004). However, these optimal schemes are difficult to design for the elastic wave equation, because the numerical scheme directly depends on the Poisson ratio of the medium, which leads to stability issues when considering free-surface boundary conditions and liquid/solid interfaces (Stekl & Pratt, 1998). Moreover, FD suffers from the usual limitation linked to Cartesian regular grids: the grid interval is defined by the smallest propagated wavelength to avoid numerical dispersion, or by complex interface geometries, which require fine discretization to avoid parasite diffractions from the corners of the Cartesian grid. These two constraints generally lead to local oversampling of the medium, and therefore a waste of computational resources (Saenger & Bohlen, 2004; Bohlen & Saenger, 2006). To deal with non-regular grids, finite element (FE) methods can be considered with either low interpolation orders (Marfurt, 1984; Bielak *et al.*, 2003), or high interpolation orders, using the so-called spectral element method (Komatitsch & Vilotte, 1998; Chaljub *et al.*, 2003). These methods require, however, explicit boundary conditions for modeling the water/solid contact, and are generally developed for quadrangle grids, which leads to meshing issues. Another alternative is the finite element discontinuous Galerkin (DG) method (Käser & Dumbser, 2006; Dumbser & Käser, 2006; Brossier *et al.*, 2008). This DG method allows the use of triangular/tetrahedral meshes, and it is suitable for the handling of strong physical contrasts in the medium, including liquid/solid contact. The DG method is chosen in the present study to perform the FWI forward problem in the frequency domain.

The second difficulty of FWI is related to the ill-posedness and non-linearity of the inverse problem, which is generally formulated as a least-squares local optimization (Tarantola, 1987; Pratt & Worthington, 1990), so as to manage the numerical cost of the forward problem. The ill-posedness of FWI arises mainly from: the limited accuracy of the starting model in the framework of local optimization; the incomplete illumination of the subsurface provided by conventional seismic surveys; the reconstruction of multiple classes of parameters with different sensitivities and signatures in the data; the lack of low frequencies in the source bandwidth; and the presence of noise. Several hierarchical multiscale strategies that proceed from low frequencies to higher frequencies have been proposed to mitigate the non-linearity of the inverse problem (Pratt & Worthington, 1990; Bunks *et al.*, 1995; Sirgue & Pratt, 2004; Brossier *et al.*, 2009a). Subsets of data and parameter classes can also be hierarchically inverted to mitigate the nonlinearity and nonuniqueness issues in the framework of elastic inversion (Sears *et al.*, 2008; Brossier *et al.*, 2009b).

To tackle targets of realistic size and complexity with FWI, efficient algorithms must be implemented on large-scale distributed-memory computational platforms for both the forward and the inverse problems. An example of such algorithms is presented in Sourbier *et al.* (2009a), who implemented a parallel FWI algorithm in the framework of the two-dimensional (2D) acoustic monoparameter approximation, with a FD forward problem and steepest-descent-based optimization scheme.

The present study presents the implementation of a massively parallel frequency-domain FWI algorithm for imaging 2D visco-elastic media. The first section will focus on the forward problem resolution, as performed with a low-order DG method. In the second section, the inversion algorithm problem is presented, with special emphasis on the reconstruction of multi-parameter classes from vectorial wavefields using a quasi-Newton method. The last section will review the parallel implementation of the algorithm, with an analysis of its performance on a distributed-memory platform.

3.3 Forward Problem

3.3.1 Theory and discretization

Two-dimensional, elastic, frequency-domain FWI requires the frequency-domain solutions of the 2D P-SV wave equations for heterogeneous media and multiple sources. We consider the first-order hyperbolic system where both particle velocities (V_x, V_z) and stresses ($\sigma_{xx}, \sigma_{zz}, \sigma_{xz}$) are unknown quantities, as described by the system :

$$\begin{aligned}
 -i\omega V_x &= \frac{1}{\rho(\mathbf{x})} \left\{ \frac{\partial \sigma_{xx}}{\partial x} + \frac{\partial \sigma_{xz}}{\partial z} \right\} + f_x \\
 -i\omega V_z &= \frac{1}{\rho(\mathbf{x})} \left\{ \frac{\partial \sigma_{xz}}{\partial x} + \frac{\partial \sigma_{zz}}{\partial z} \right\} + f_z \\
 -i\omega \sigma_{xx} &= (\lambda(\mathbf{x}) + 2\mu(\mathbf{x})) \frac{\partial V_x}{\partial x} + \lambda(\mathbf{x}) \frac{\partial V_z}{\partial z} - i\omega \sigma_{xx0} \\
 -i\omega \sigma_{zz} &= \lambda(\mathbf{x}) \frac{\partial V_x}{\partial x} + (\lambda(\mathbf{x}) + 2\mu(\mathbf{x})) \frac{\partial V_z}{\partial z} - i\omega \sigma_{zz0} \\
 -i\omega \sigma_{xz} &= \mu(\mathbf{x}) \left\{ \frac{\partial V_x}{\partial z} + \frac{\partial V_z}{\partial x} \right\} - i\omega \sigma_{xz0},
 \end{aligned} \tag{3.1}$$

where λ and μ are the Lamé coefficients, ρ is the density, and ω is the angular frequency. Source terms are either point forces (f_x, f_z) or applied stresses ($\sigma_{xx0}, \sigma_{zz0}, \sigma_{xz0}$). i is the purely imaginary term defined by $i = \sqrt{-1}$. Only isotropic media are considered in this analysis. The intrinsic attenuation of the medium is easily taken into account in the frequency domain using complex-valued wave velocities. In the present study, the Kolsky-Futterman model (Kolsky, 1956; Futterman, 1962) without the dispersion term is used (Toksöz & Johnston, 1981).

A change in variable is applied to the system of equations (3.1) to develop a pseudo-conservative formulation that is useful for the derivation of the DG formulation. We now consider the following vector with three components $(T_1, T_2, T_3) = ((\sigma_{xx} + \sigma_{zz})/2, (\sigma_{xx} - \sigma_{zz})/2, \sigma_{xz})$. Note that, when considering this variable change, T_1 represents the hydrostatic pressure measured in liquids. In the case of liquid/solid propagation, such as in marine seismic experiments,

the T_1 variable provides direct access to the pressure measured by hydrophones, while T_2 and T_3 are zero in liquid, relative to the deviatoric part of the stress tensor.

Moreover, we must consider a finite domain, and therefore, we apply Perfectly Matched Layer (PML) absorbing conditions (Berenger, 1994) through the functions s_x, s_z , defined in the direction r as $s_r = 1/(1 + i\gamma_r/\omega)$, where γ_r is zero inside the computational domain and nonzero inside the absorbing layers. The new differential system that includes the PML functions, which is equivalent to the system of equations (3.1), can be written as :

$$\begin{aligned}
 -i\omega\rho V_x &= s_x \frac{\partial(T_1 + T_2)}{\partial x} + s_z \frac{\partial T_3}{\partial z} + \rho f_x \\
 -i\omega\rho V_z &= s_x \frac{\partial T_3}{\partial x} + s_z \frac{\partial(T_1 - T_2)}{\partial z} + \rho f_z \\
 \frac{-i\omega T_1}{\lambda + \mu} &= s'_x \frac{\partial V_x}{\partial x} + s'_z \frac{\partial V_z}{\partial z} - \frac{i\omega T_1^0}{\lambda + \mu} \\
 \frac{-i\omega T_2}{\mu} &= s'_x \frac{\partial V_x}{\partial x} - s'_z \frac{\partial V_z}{\partial z} - \frac{i\omega T_2^0}{\mu} \\
 \frac{-i\omega T_3}{\mu} &= s'_x \frac{\partial V_z}{\partial x} + s'_z \frac{\partial V_x}{\partial z} - \frac{i\omega T_3^0}{\mu}.
 \end{aligned} \tag{3.2}$$

To derive the DG formulation, the model is first discretized with polygonal cells. For each cell, the system of equations (3.2) is multiplied by a test function, corresponding to a k^{th} -order polynomial. The test function is nonzero only in the polygonal cell that ensures the discontinuous property of the scheme. In the present study, Lagrange polynomials are used for the test functions. The system of equations (3.2) is then integrated over each cell, which leads to the so-called weak formulation of the system. We assume that the physical properties of the medium are constant inside each cell, and we use centered numerical fluxes to exchange energy between cells (Remaki, 2000). The reader is referred to Käser & Dumbser (2006); Dumbser & Käser (2006); Brossier *et al.* (2008); Delcourte *et al.* (2009) for a detailed description of the DG formulation applied to the elastodynamic equations.

Discretization of equations (3.2) with the DG method leads to a linear system for each frequency considered :

$$\mathbf{A}\mathbf{u} = \mathbf{b}, \tag{3.3}$$

where \mathbf{A} is the sparse impedance matrix, which is also known as the forward problem operator, and which depends on the frequency, the mesh geometry, the DG interpolation order of each cell, and the physical properties. \mathbf{b} represents the external excitation. \mathbf{u} is the unknown vector, which represents the particle velocity and the stress wavefields in the entire computational domain, and is based on a local polynomial representation inside each cell. In the present study, the system of equations (3.3) is solved with the massively parallel direct solver MUMPS (Amestoy *et al.*, 2006). More details on this algorithm and its performance will be discussed in section 3.5.1.

3.3.2 Which interpolation ?

The theoretical resolution of the FWI is $\lambda/2$ if normal-incidence data are recorded, where λ is the local propagated wavelength (Sirgue & Pratt, 2004). A suitable discretization for the inversion mesh should therefore be close to $\lambda/4$ to ensure a correct sampling of the medium. Since the physical properties are assumed to be constant inside each cell, a suitable discretization of the forward problem based on the same mesh as that of the inversion should also be close to $\lambda/4$ (Sourbier *et al.*, 2009a).

DG methods generally take advantage of high orders of interpolation to coarsely discretize the medium and to obtain a high order of accuracy (Käser & Dumbser, 2006; Dumbser & Käser, 2006). In contrast, I will focus here on low interpolation orders for FWI applications because these low orders of interpolation will provide sufficiently accurate solutions to the forward problem for meshes that are consistent with the expected resolution of FWI.

The lowest interpolation order, P_0 , is a possible choice that represents the wavefield by a series of piece-wise constant functions inside each cell, and it gives sufficiently accurate solutions for ten to fifteen cells per wavelength in structured equilateral triangular meshes (Brossier *et al.*, 2008). P_0 interpolation cannot, however, be used on unstructured meshes, where it does not converge to the correct solution, whatever the size of the cell. Moreover, the liquid/solid interface causes some troubles for structured meshes, and the P_0 interpolation cannot, therefore, be used in this case. Finally, the P_0 scheme can be related to the finite volume P_0 scheme, and is very close to $O(\Delta x^2)$ FD schemes (Brossier *et al.*, 2008).

The higher orders of interpolation, P_1 and P_2 , allow the use of unstructured meshes, which is an appealing feature for dealing with complex topographies and heterogeneous media, which should require a local adaptation of the mesh size to the local properties (*h*-adaptive meshes). Sufficiently accurate solutions can be obtained in unstructured meshes with discretization of ten to fifteen cells per wavelength for P_1 , and three to three and a half cells per wavelength for P_2 .

Seismograms computed with the DG P_1 and DG P_2 in a homogeneous halfspace are compared with the analytical solution of the Garvin (1956) problem (Figures 3.1 and 3.2). The source is an explosion at depth and the receivers are on the free surface. The comparison is performed in the time domain, from DG solutions computed in the frequency domain. To assess the accuracy of the surface wave and of the direct wave, I have here used a discretization of twelve cells and three cells per minimum wavelength in unstructured triangular meshes for P_1 and P_2 interpolations, respectively. The comparison between the DG and analytical seismograms shows that sufficiently accurate solutions for both wave types are achieved for these discretizations. For the Garvin test computed in the frequency domain at 5 Hz, Figure 3.3 reviews the accuracy of the three interpolation orders as functions of the mesh size.

The P_2 interpolation provides the best consistency between the discretizations required by the forward problem to compute sufficiently accurate solutions (Figures 3.2 and 3.3), and the spatial resolution of FWI (i.e., the mesh interval corresponding to a quarter of a wavelength). Higher orders of interpolation would allow the use of coarser meshes for the same level of accuracy, but these coarse meshes would lead to an undersampling of the medium for FWI. Therefore, high orders of interpolation should be applied on unnecessarily fine meshes, at the expense of the computational efficiency, and should, therefore, be rejected for FWI applications. As I will illustrate later, lower orders of interpolation (i.e., P_0 and P_1) can, however, be useful as an alternative to P_2 , depending on the medium properties.

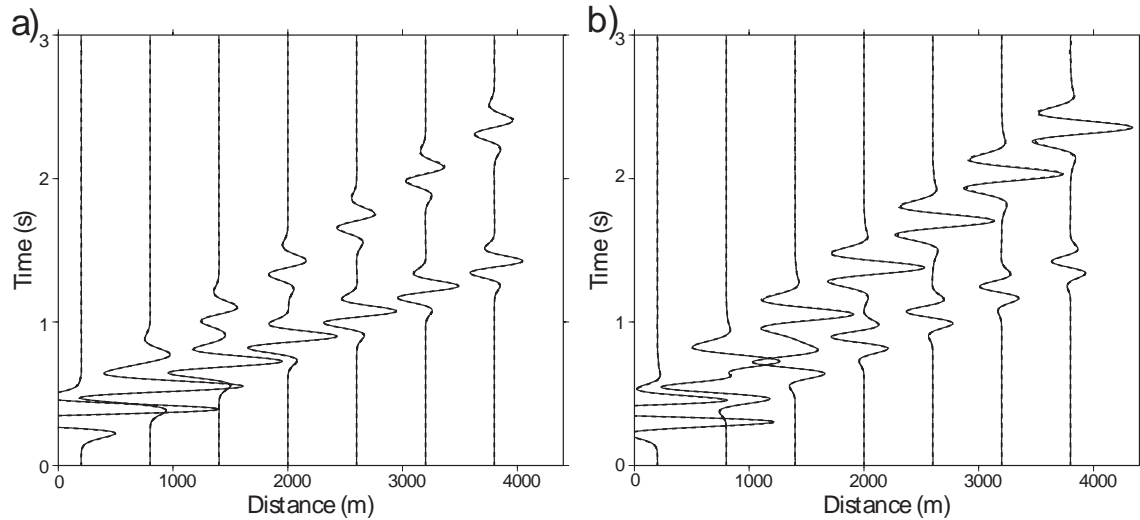


FIG. 3.1 – Time-domain comparison between DG P_1 (solid lines) and analytical (dashed lines) solutions for the Garvin test. The seismograms represent : (a) the horizontal; and (b) the vertical components of the particle velocity at the free surface.

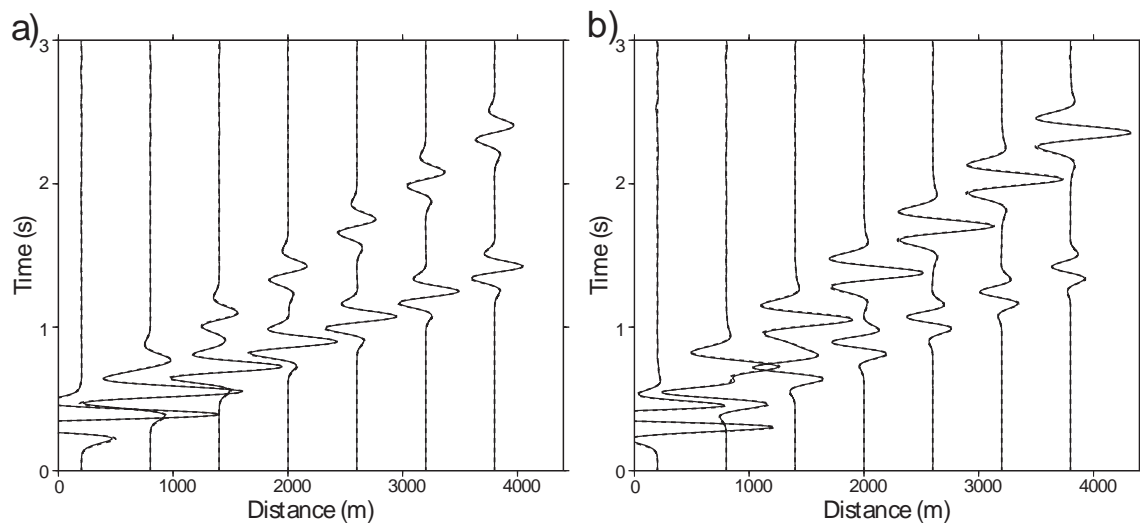


FIG. 3.2 – Time-domain comparison between DG P_2 (solid lines) and analytical (dashed lines) solutions for the Garvin test. The seismograms represent : (a) the horizontal; and (b) the vertical components of the particle velocity at the free surface.

For computational efficiency, an interesting feature of the DG method is that it allows for mixing of the interpolation orders (p -adaptivity), because the degrees of freedom are not shared by neighboring cells in the DG formalism. An efficient mix is seen in P_0 - P_1 , because both interpolations share the same discretization rule (i.e., ten cells per wavelength), and therefore, they facilitate the meshing at the transition zones between the P_0 and P_1 interpolations.

Table 3.1 summarizes the computational time and memory requirements for the DG me-

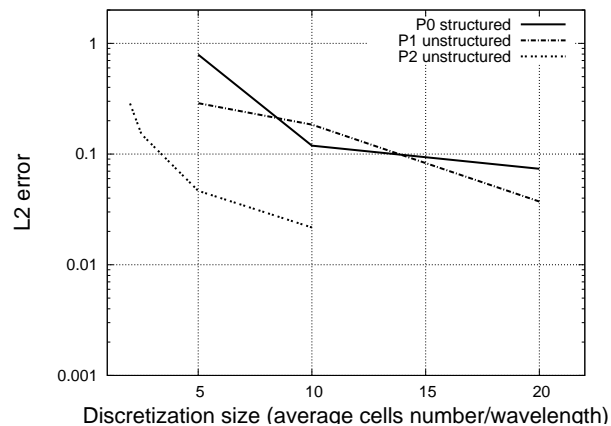


FIG. 3.3 – L_2 error at the free surface between the analytical solution and the numerical solutions computed with the DG P_0 , P_1 and P_2 interpolation orders as a function of the mesh size. The L_2 error is computed for the 5 Hz frequency, and for the horizontal and vertical geophone components.

thod for two simple models, where the P_0 , P_1 , P_0 - P_1 and P_2 orders of interpolation are used. These have, respectively, a discretization of : ten cells per minimum wavelength in a structured equilateral mesh ; ten cells per local wavelength in an unstructured h -adaptive mesh ; ten cells per local wavelength with a structured/unstructured hp -adaptive mesh ; and three cells per local wavelength in an unstructured h -adaptive mesh. It needs to be remembered that these four combinations of interpolation order and discretization give similar accuracies (Figure 3.3). Simulations are performed in parallel using 16 processors. The first model is a homogeneous half space, and the second one is a two-layer half space. The first simulation clearly shows that the P_0 interpolation on the regular mesh provides the best time and memory performances when the model contains a limited range of velocity. In the second model, the thickness of the upper layer is one tenth of the thickness of the half space, and the velocity in the upper layer is one tenth of the velocity of the lower layer. The upper layer mimics a complex weathering zone in the near surface, which requires a very fine mesh and, therefore, a dramatic increase in the number of cells when regular structured meshes are used (as it should also be with FD). In this case, the P_2 interpolation on the h -adaptive unstructured mesh provides the best time and memory performances, due to the local refinement of the mesh (Table 3.1). The P_0 - P_1 mesh uses h -adaptive unstructured cells (P_1) in the upper layer and structured equilateral cells (P_0) in the lower layer. These two extreme canonical cases illustrate how the heterogeneity and the velocity range of the subsurface drive the choice of the mesh type and interpolation order. In the case of a broad velocity bandwidth in the model, P_2 interpolation used with h -adaptive unstructured meshes should give the best parameterization. In the case of a narrow velocity bandwidth, the P_0 or the P_0 - P_1 interpolations should be more relevant. If a flat free surface is present without liquid, the P_0 interpolation can be used, while the P_0 - P_1 interpolation should be used when the model contains a liquid/solid interface in a shallow water environment or a complex topography above a shallow weathered layer. For example, a layer of unstructured cells computed with P_1 allows discretization of a complex topography or water/solid interface of arbitrary shape, above a P_0 regular equilateral mesh, discretizing a more homogeneous medium. The P_0 - P_1 approach is close in spirit to the hybrid FE/FD method proposed by Moczo

TAB. 3.1 – Computational resources required for the forward problem solved with DGs P_0 , P_1 , P_0 - P_1 and P_2 , in two simples cases, on 16 processors

Test	Resource	P_0	P_1	P_0 - P_1	P_2
Homogenous	Cell numbers	113 097	136 724	116 363	12 222
	Degrees of freedom	565 485	2 050 860	695 795	366 660
	Time for factorization (s)	2.1	59.2	4.1	9.4
	Memory for factorization (Gb)	2.39	19.80	3.09	4.57
	Time to solve 116 RHS (s)	6.2	28.3	8.1	6.1
Two-layers	Cell numbers	2 804 850	291 577	247 303	32 664
	Degrees of freedom	14 024 250	4 373 655	2 360 405	979 920
	Time for factorization (s)	333.2	109.5	31.8	22.1
	Memory for factorization (Gb)	80.50	33.26	14.37	10.84
	Time to solve 116 RHS (s)	190.4	72.4	30.2	16.2

et al. (1997). We can, however, note that the P_1 interpolation used alone never appears to be the optimal choice.

Figures 3.4, 3.5 and 3.6 further illustrate the mesh design for three realistic offshore and onshore models. The first represents a velocity model of the Valhall gas field in a shallow-water environment in the North Sea (Munns, 1985). The sea bed is at 70-m depth. The narrow velocity range in most parts of the model requires the use of a regular mesh as much as possible for computational efficiency. To correctly discretize the shallow-water layer, a p -adaptivity implemented with a mixed P_0 - P_1 interpolation is chosen : a refined unstructured P_1 layer of cells is used for the first 130 m of the subsurface for accurate modeling of the interface waves at the liquid/solid interface, and for accurate positioning of the sources, which are located 5 m below the surface, and of the receivers, which are on the sea bed. The second example represents an onshore foothill model that contains a complex topography, several dipping structures, and a thin weathered layer at the surface ($V_P=700$ m/s, $V_S=350$ m/s) : a dramatic configuration for regular grid-based methods. An h -adaptive unstructured mesh that is suitable for the P_2 interpolation order was designed for accurate modeling of the free-surface effects of the topography, and for the local adaptation of the cell size to the local wave speeds. The third example corresponds to a crustal-scale model, which represents the Nankai subduction zone, offshore of Japan (Operto *et al.*, 2006). We use a P_2 h -adaptive unstructured mesh to deal with the broad velocity bandwidth spanned by the crustal model.

3.4 Inverse Problem

3.4.1 Theory review

We start with a brief review of FWI. A more extensive overview is presented in Virieux & Operto (2009). FWI is an optimization problem that is generally recast as a linearized least-squares problem that attempts to minimize the misfit between the recorded and the modeled wavefields (Tarantola, 1987). The inverse problem can be formulated in the frequency domain (Pratt & Worthington, 1990), and the associated objective function to be minimized is defined by :

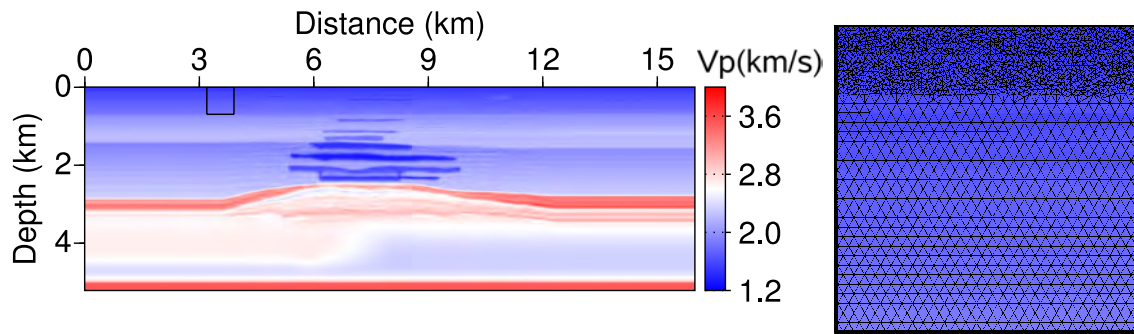


FIG. 3.4 – The synthetic Valhall P-wave velocity model representative of oil and gas shallow water targets, and a close-up of the mesh used for mixed DG P_0 - P_1 simulations.

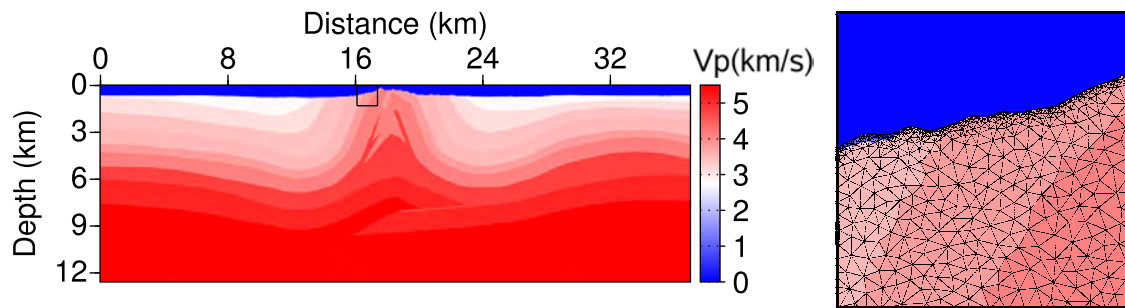


FIG. 3.5 – The synthetic IFP/TOTAL Foothill P-wave velocity model representative of onshore targets with complex topographies and weathering zones, and a close-up of the mesh for DG P_2 simulations.

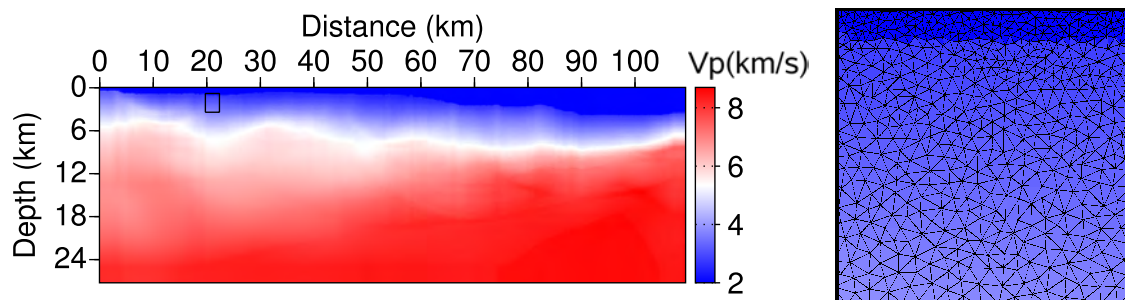


FIG. 3.6 – The crustal P-wave velocity model representative of the Japanese Nankai trough, and a close-up of the h -adaptive unstructured mesh for DG P_2 simulations.

$$\mathcal{C}^{(n)} = \sum_{if=1}^{nf} \sum_{is=1}^{ns} \frac{1}{2} \Delta \mathbf{d}^{(n)\dagger} \mathbf{W}_d \Delta \mathbf{d}^{(n)}, \quad (3.4)$$

where $\Delta \mathbf{d}^{(n)} = \mathbf{d}_{obs} - \mathbf{d}_{calc}^{(n)}$ is the data misfit vector, the difference between the observed data \mathbf{d}_{obs} and the modeled data $\mathbf{d}_{calc}^{(n)}$ computed in the model $\mathbf{m}^{(n)}$ at the iteration n of the inversion. Superscript \dagger indicates the adjoint (transposed conjugate), and \mathbf{W}_d is a diagonal weighting matrix that is applied to the misfit vector to scale the relative contributions of each of its components. The summations in equation (3.4) are performed over the ns sources and a group of nf simultaneously inverted frequencies.

With the objective function assumed to be locally parabolic, its minimization with the Newton method provides the following expression of the perturbation model $\delta \mathbf{m}^{(n)}$:

$$\mathbf{B}^{(n)} \delta \mathbf{m}^{(n)} = -\mathcal{G}^{(n)}, \quad (3.5)$$

where $\mathbf{B}^{(n)}$ and $\mathcal{G}^{(n)}$ denote the Hessian matrix and the gradient of the objective function at iteration n , respectively.

The gradient \mathcal{G} of the objective function with respect to the model parameters $\mathbf{m} = \{m_i\}_{i=1,N}$, where N denotes the number of unknowns, can be derived from the adjoint-state formalism using the back-propagation technique (Plessix, 2006).

This gives for the i^{th} component of the gradient \mathcal{G} :

$$\mathcal{G}_{m_i} = \sum_{if=1}^{nf} \sum_{is=1}^{ns} \Re \left\{ \mathbf{u}^t \frac{\partial \mathbf{A}^t}{\partial m_i} \mathbf{A}^{-1} \mathcal{P}_{data} \mathbf{W}_d \Delta \mathbf{d}^* \right\}, \quad (3.6)$$

where t and * denote the transpose and conjugate operators, respectively, \Re denotes the real part of a complex number, and \mathcal{P}_{data} is an operator that extends with zeroes the residual vector represented in the data space to the forward problem space (Pratt *et al.*, 1998). The gradient is computed by a zero-lag convolution in time, between the incident wavefield \mathbf{u} from the source, and the adjoint back-propagated wavefield $\mathbf{A}^{-1} \mathcal{P}_{data} \mathbf{W}_d \Delta \mathbf{d}^*$, using residuals at receiver positions as a composite source. Therefore, only two forward problems per shot are required for gradient building. The radiation pattern of the diffraction by the model parameter m_i is denoted by the sparse matrix $\partial \mathbf{A} / \partial m_i$.

Due to the computational cost of building the Hessian operator and solving equation (3.5), Newton and Gauss-Newton methods are generally not considered for realistically sized problems (Pratt *et al.*, 1998). Steepest-descent or conjugate-gradient methods that are preconditioned by the diagonal terms of an approximate Hessian are more conventionally used to compute $\delta \mathbf{m}^{(n)}$ (Pratt *et al.*, 1998; Operto *et al.*, 2006; Shin *et al.*, 2001).

Because the Hessian is generally not taken into account exactly, the model is then updated with an optimal step-length $\alpha^{(n)}$, computed by a linear inversion (Tarantola, 1987) or by parabola fitting (Soubier *et al.*, 2009a)

$$\mathbf{m}^{(n+1)} = \mathbf{m}^{(n)} + \alpha^{(n)} \delta \mathbf{m}^{(n)}. \quad (3.7)$$

3.4.2 Inversion algorithm

CPU-efficient frequency-domain FWI is generally carried out by successive inversions of single frequencies, by proceeding from the low frequencies to the higher ones (Pratt & Worthington, 1990; Sirgue & Pratt, 2004). This defines a multiresolution framework that mitigates of the non-linearity of the inverse problem associated with high frequency cycle-skipping artifacts. CPU-efficient algorithms can be designed by selecting a few coarsely sampled frequencies, such that the wavenumber redundancy, which results from dense sampling of frequencies and aperture angles, is decimated. This strategy has proven to be effective for several applications of acoustic FWI (Ravaut *et al.*, 2004; Operto *et al.*, 2006; Brenders & Pratt, 2007a). However, it might lack robustness when complex wave phenomena are present, such as P-to-S conversions, multiples and surface waves.

A two-level hierarchical procedure (Algorithm 3.1) to mitigate the FWI non-linearities in the case of elastic onshore problems was proposed by Brossier *et al.* (2009a). This algorithm embeds three main loops :

1. The outer loop is over the frequency groups ; namely, a set of frequencies that are simultaneously inverted. In the case of complex wave phenomena, the use of simultaneous frequencies in the inversion procedure constrains better the optimization for convergence towards the global minimum, taking into account the more redundant information contained in the simultaneous frequencies.
2. The second loop is over the time-damping factors that control the amount of information over time that is preserved in the seismograms for inversion. Time damping is performed in frequency-domain modeling through the use of complex-valued frequencies, which is equivalent to damping seismograms in time by an exponential decay (Shin *et al.*, 2002; Brenders & Pratt, 2007b). Moreover, this data preconditioning provides a significant improvement in the signal-to-noise ratio with real data (Brenders *et al.*, 2009).
3. The third loop is over iterations of the non-linear problem for one frequency group inversion.

This algorithm was applied successfully for imaging the elastic parameters of onshore and offshore synthetic data in Brossier *et al.* (2009a,b).

3.4.3 Gradient computation

The gradient computation with the adjoint-state method is described formally by equation (3.6). Without loss of sense and using some notations that are more adapted to the DG formalism, this expression can be re-written as :

$$\mathcal{G}_{m_i} = \sum_{if=1}^{nf} \sum_{is=1}^{ns} \Re \left\{ \mathbf{s}_{p_k}^t \mathbf{A} \mathbf{1}_{p_k}^{-1t} \frac{\partial \mathbf{A} \mathbf{1}_{p_k}^t}{\partial m_i} \mathbf{A} \mathbf{1}_{p_k}^{-1} \mathcal{P}_{data} \mathbf{W}_d \Delta \mathbf{d}_{p_k}^* \right\}, \quad (3.8)$$

where subscript p_k denotes the k order of interpolation used to discretize the wavefields and source vectors, and the forward operator matrix. The $\mathbf{A} \mathbf{1}$ matrix results from the discretization of the first-order velocity/stress formulation of the elastodynamic equations, equation (3.2), used to perform the forward problem. \mathbf{s}_{p_k} denotes the source vector.

Algorithm 3.1 Two-level hierarchical frequency-domain FWI algorithm

```

1: for  $ig = 1$  to  $ng$  do
2:   for  $id = 1$  to  $nd$  do
3:     while (NOT convergence AND  $n < n_{max}$ ) do
4:       for  $if = 1$  to  $nf$  do
5:         Compute incident wavefields  $\mathbf{u}$  from sources
6:         Compute residual vectors  $\Delta \mathbf{d}$  and cost function  $\mathcal{C}^{(n)}$ 
7:         Compute adjoint back-propagated wavefields  $\mathbf{A}^{-1} \mathcal{P}_{data} \mathbf{W}_d \Delta \mathbf{d}^*$  from receivers
8:         Build gradient vector  $\mathcal{G}^{(n)}$  (with Algorithm 3.2)
9:       end for
10:      Compute perturbation vector  $\delta \mathbf{m}^{(n)}$  (with Algorithm 3.3)
11:      Define optimal step length  $\alpha^{(n)}$  by parabola fitting
12:      Update model  $\mathbf{m}^{(n+1)} = \mathbf{m}^{(n)} + \alpha^{(n)} \delta \mathbf{m}^{(n)}$ 
13:    end while
14:  end for
15: end for

```

For computational time and memory use efficiency, equation (3.8) can be re-written using several tricks.

As a first step, the wavefield solutions computed at each node of the p_k interpolation polynomials are interpolated at the barycentre of each cell of the triangular mesh, and are subsequently assumed to be piece-wise constant in the cell. After this interpolation, the discretizations of the wavefield solutions and the medium properties are identical, and are consistent with those of a P_0 interpolation. The radiation pattern matrix $\partial \mathbf{A} \mathbf{1}_{P_0} / \partial m_i$ is built using a P_0 interpolation order, to be consistent with the discretization of the projected solutions and medium properties. This does not affect the model reconstruction because the radiation pattern operator used in the inverse problem does not need to be discretized to the same order of accuracy as the impedance matrix used in the forward problem. Let us introduce the projection matrix \mathbf{P}_{P_0} that interpolates the p_k wavefields at the barycentre of each cell. Expression (3.8) can be re-written as :

$$\begin{aligned}
\mathcal{G}_{m_i} &= \sum_{if=1}^{nf} \sum_{is=1}^{ns} \Re \left\{ \mathbf{s}_{p_k}^t \mathbf{A} \mathbf{1}_{p_k}^{-1t} \mathbf{P}_{P_0}^t \frac{\partial \mathbf{A} \mathbf{1}_{P_0}^t}{\partial m_i} \mathbf{P}_{P_0} \mathbf{A} \mathbf{1}_{p_k}^{-1} \mathcal{P}_{data} \mathbf{W}_d \Delta \mathbf{d}_{p_k}^* \right\} \\
&= \sum_{if=1}^{nf} \sum_{is=1}^{ns} \Re \left\{ \mathbf{u}_{P_0}^t \frac{\partial \mathbf{A} \mathbf{1}_{P_0}^t}{\partial m_i} \mathbf{r}_{P_0} \right\},
\end{aligned} \tag{3.9}$$

where \mathbf{u}_{P_0} and \mathbf{r}_{P_0} are the projected incident and adjoint velocity/stress wavefields. Note that this first projection allows the saving of a factor of three and six, both in computational time to build the gradient and core-memory use to store the wavefields, when the forward problem is solved with DG P_1 and P_2 , respectively. If the forward problem is computed with P_0 , this step does not provide any saving.

A second improvement can also be applied to expression (3.9). The 2D-PSV elastodynamic equations can be written either as a first-order velocity/stress system that involves five

Algorithm 3.2 Algorithm of gradient computation

```

1: for  $m_i = 1$  to  $m_N$  do
2:   Compute  $\frac{\partial \mathbf{A}\mathbf{2}_{P_0}^t}{\partial m_i}$ 
3:   for  $s = 1$  to  $nsrc$  do
4:     Compute  $\beta \leftarrow \frac{\partial \mathbf{A}\mathbf{2}_{P_0}^t}{\partial m_i} \mathbf{u}_{P_0/V}$  through full vector/sparse matrix product routine
5:     Compute  $\mathcal{G}_{m_i} \leftarrow \mathcal{G}_{m_i} + \Re(\beta^t \mathbf{r}_{P_0/V})$  with sparse vector/full vector dot product routine
6:   end for
7: end for
    
```

velocity/stress wavefields, or as a second-order velocity system that involves two velocity wavefields. The P_0 forward problem operator $\mathbf{A}\mathbf{2}_{P_0}$ for the second-order velocity system can be assembled, starting from the P_0 -discretized first-order velocity/stress system and eliminating the auxiliary stress wavefields through a *parsimonious* formulation (Luo & Schuster, 1990), developed in Brossier *et al.* (2008). The forward problem operator $\mathbf{A}\mathbf{2}_{P_0}$ can subsequently be used to build the radiation pattern matrix that is required by the gradient computation. Let us define the operator \mathbf{P}_V , which restricts the velocity/stress wavefield vector to the velocity wavefield vector. Expression (3.9) can be written as :

$$\begin{aligned}
 \mathcal{G}_{m_i} &= \sum_{if=1}^{nf} \sum_{is=1}^{ns} \Re \left\{ \mathbf{u}_{P_0}^t \mathbf{P}_V^t \frac{\partial \mathbf{A}\mathbf{2}_{P_0}^t}{\partial m_i} \mathbf{P}_V \mathbf{r}_{P_0} \right\} \\
 &= \sum_{if=1}^{nf} \sum_{is=1}^{ns} \Re \left\{ \mathbf{u}_{P_0/V}^t \frac{\partial \mathbf{A}\mathbf{2}_{P_0}^t}{\partial m_i} \mathbf{r}_{P_0/V} \right\}, \tag{3.10}
 \end{aligned}$$

where $\mathbf{u}_{P_0/V}$ and $\mathbf{r}_{P_0/V}$ are the projected incident and adjoint velocity wavefields. This second improvement does not lead to significant computational time saving compared to the formulation (3.9) because $\frac{\partial \mathbf{A}\mathbf{2}_{P_0}^t}{\partial m_i}$ is denser than $\frac{\partial \mathbf{A}\mathbf{1}_{P_0}^t}{\partial m_i}$. However, it allows a reduction of the memory storage of the wavefields by a factor two and a half. Of note, the inverted data can be either the particle velocity or the hydrostatic pressure when the expression of the gradient, equation (3.10), is used, although the gradient is built from the velocity wavefields, $\mathbf{u}_{P_0/V}$ and $\mathbf{r}_{P_0/V}$. There is only the need to set a residual source in pressure for $\Delta \mathbf{d}$, back-propagate these pressure residuals with the velocity/stress forward problem operator through the matrix $\mathbf{A}\mathbf{1}_{P_k}$, and extract the modeled adjoint velocity wavefields $\mathbf{r}_{P_0/V}$ for gradient computation. Note also that visco-acoustic FWI can be performed using the elastic algorithm : the fourth and fifth lines of system (3.2) are not considered in the forward problem ; the T_2 and T_3 wavefields are forced to zero ; and V_S is set to zero.

Finally, the gradient vector is computed through Equation (3.10) with Algorithm 3.2.

3.4.4 L-BFGS and multi-parameter inversion

FWI is generally solved with the steepest-descent or conjugate-gradient methods, preconditioned by a diagonal approximation of the Hessian (Pratt *et al.*, 1998; Operto *et al.*, 2006; Shin *et al.*, 2001). Limited-memory, quasi-Newton methods, such as the limited-memory Broyden-Fletcher-Goldfarb-Shanno (L-BFGS) method (Nocedal, 1980), can however provide an efficient

alternative to classic steepest-descent or conjugate-gradient methods for FWI problems (Brosier *et al.*, 2009a).

The Hessian matrix allows for : correction of the gradient from the geometrical amplitude spreading of the incident and adjoint wavefields; improvement of the imaging resolution by correctly deconvolving the gradient from limited-bandwidth effects, taking into account the off-diagonal terms of the Hessian; suitable scaling of the model perturbation vector for different parameter classes; and accounting for double scattering effects (Pratt *et al.*, 1998; Nocedal & Wright, 1999). L-BFGS provides an approximation of the inverse \mathbf{H} of the Hessian matrix from few gradient and model difference vectors, stored from previous iterations. Moreover, the double-loop recursive algorithm designed by Nocedal (1980) does not explicitly build and store $\mathbf{H}^{(k)}$, but directly computes the perturbation vector $\delta\mathbf{m}^{(k)} = -\mathbf{H}^{(k)}\mathcal{G}^{(k)}$ with additions, differences and inner products of vectors. A diagonal approximation of the inverse Hessian computed from the diagonal terms of the approximate Hessian (Pratt *et al.*, 1998; Operto *et al.*, 2006) or pseudo-Hessian (Shin *et al.*, 2001), can be provided to the L-BFGS algorithm for a better and faster estimation of $\mathbf{H}^{(k)}\mathcal{G}^{(k)}$.

In the framework of the inversion of multiple classes of parameter with different units and amplitudes (for example, reconstruction of wave speeds, density, attenuation factors), the L-BFGS algorithm requires some normalizations for consistent dimensionless computation of the model perturbation vector. Algorithms 3.3 and 3.4 review the computation of the perturbation model with appropriate normalizations that account for the different classes of parameters in the L-BFGS algorithm (Nocedal, 1980), modified following S. Gratton (2008, personal communication). The normalization of the gradient, the diagonal Hessian, and the perturbation model are performed according to a characteristic value m_{i0} that is defined for each parameter class, taken as the mean of the initial model of each parameter class. This normalization is equivalent to computation of the gradient and the diagonal Hessian for normalized parameters \mathbf{m}_i/m_{i0} , where i denotes the parameter class.

Figure 3.7 shows the FWI results in a canonical configuration : 113 horizontal point force sources are located all around the target each 50 m, and each source wavefield is recorded by 112 two-component geophones all around the target each 50 m. Absorbing boundary conditions are used on the four sides of the domain to simulate an infinite medium. The true models contain one circular heterogeneity of 200 m diameter at different locations for each of the five parameters : P-wave and S-wave velocities (V_P and V_S), density (ρ), and P-wave and S-wave quality factors (Q_p and Q_s) (Figure 3.7). Table 3.2 summarizes the background and heterogeneity values for each parameter. The inversion is performed for ten frequencies between 4 Hz and 13 Hz each 1 Hz, starting from homogeneous background models. The inverted parameters are λ , μ (the Lamé parameters), ρ , Q_p and Q_s , to deliberately invert parameters with contrasted amplitudes (Q_p and $Q_s \approx 10^1 - 10^2$ while λ and $\mu \approx 10^8 - 10^9$). Despite these differences, the normalization procedure allows for a correct estimation of the Hessian through L-BFGS, and allows for a convergence towards acceptable models without any need for heuristic weighting of each parameter class. Note that the weighting provided by the inverse Hessian estimation does not require the performing of extra forward problems, whereas the subspace method of Sambridge *et al.* (1991), which is generally used in multiparameter inversions, requires the performing of one extra forward problem per parameter class to estimate the steplength. Note, however, that in Figure 3.7 there are some coupling effects between some parameters that result from the physics of the diffraction pattern of each parameter class, and not from the choice of the optimization method.

Algorithm 3.3 Perturbation computation : outer part

```

1: if  $n = 1$  then
2:   Compute diagonal of pseudo-Hessian  $\mathbf{B}_0$  (Shin et al., 2001)
3:   for  $i = 1$  to  $n_{parameter\_class}$  do
4:     Normalize diagonal of Hessian  $\widetilde{\mathbf{B}}_{0\mathbf{m}_i} = \mathbf{B}_{0\mathbf{m}_i} m_{i0}^2$ 
5:   end for
6:   Apply damping to the diagonal of Hessian (Levenberg-Marquardt method)  $\widetilde{\mathbf{B}}_0 = \widetilde{\mathbf{B}}_0 + \epsilon \mathbf{I}$ 
7: end if
8: for  $i = 1$  to  $n_{parameter\_class}$  do
9:   Normalize gradient  $\widetilde{\mathcal{G}}_{\mathbf{m}_i}^{(k)} = \mathcal{G}_{\mathbf{m}_i}^{(k)} m_{i0}$ 
10: end for
11: Compute  $\widetilde{\mathcal{D}}^{(k)} = \widetilde{\mathbf{B}}_0^{-1} \widetilde{\mathcal{G}}^{(k)}$ 
12: if  $n > k$  then
13:   Discard the vector pair  $(\mathbf{y}_{n-k-1}, \mathbf{s}_{n-k-1})$ 
14: end if
15: if  $n > 1$  then
16:   Compute and store  $\mathbf{s}_{n-1} \leftarrow \mathbf{m}^{(n)} - \mathbf{m}^{(n-1)}$ 
17:   Compute and store  $\mathbf{y}_{n-1} \leftarrow \widetilde{\mathcal{D}}^{(n)} - \widetilde{\mathcal{D}}^{(n-1)}$ 
18: end if
19: Compute  $\widetilde{\delta\mathbf{m}}^{(n)}$  with Algorithm 3.4
20: for  $i = 1$  to  $n_{parameter\_class}$  do
21:   Denormalize perturbation vector  $\delta\mathbf{m}^{(n)} = \widetilde{\delta\mathbf{m}}^{(n)} m_{i0}$ 
22: end for
    
```

Algorithm 3.4 Perturbation computation : L-BFGS algorithm (Nocedal & Wright, 1999, page 178)

```

1:  $\mathbf{q} \leftarrow \widetilde{\mathcal{D}}^{(n)}$ 
2: for  $i = n - 1$  to  $n - k$  do
3:    $\rho_i \leftarrow \frac{1}{\mathbf{y}_i^t \mathbf{s}_i}$ 
4:    $\alpha_i \leftarrow \rho_i \mathbf{s}_i^t \mathbf{q}$ 
5:    $\mathbf{q} \leftarrow \mathbf{q} - \alpha_i \mathbf{y}_i$ 
6: end for
7:  $\gamma_n \leftarrow \frac{\mathbf{s}_{n-1}^t \mathbf{y}_{n-1}}{\mathbf{y}_{n-1}^t \mathbf{y}_{n-1}}$ 
8:  $\mathbf{r} \leftarrow \gamma_n \mathbf{q}$ 
9: for  $i = n - k$  to  $n - 1$  do
10:   $\beta \leftarrow \rho_i \mathbf{y}_i^t \mathbf{r}$ 
11:   $\mathbf{r} \leftarrow \mathbf{r} + \mathbf{s}_i (\alpha_i - \beta)$ 
12: end for
13:  $\widetilde{\delta\mathbf{m}}^{(k)} = -\mathbf{r}$ 
    
```

The reader can refer to Brossier *et al.* (2009a,b) for realistic applications of this strategy to onshore and offshore synthetic data from the imaging of P-wave and S-wave velocity models.

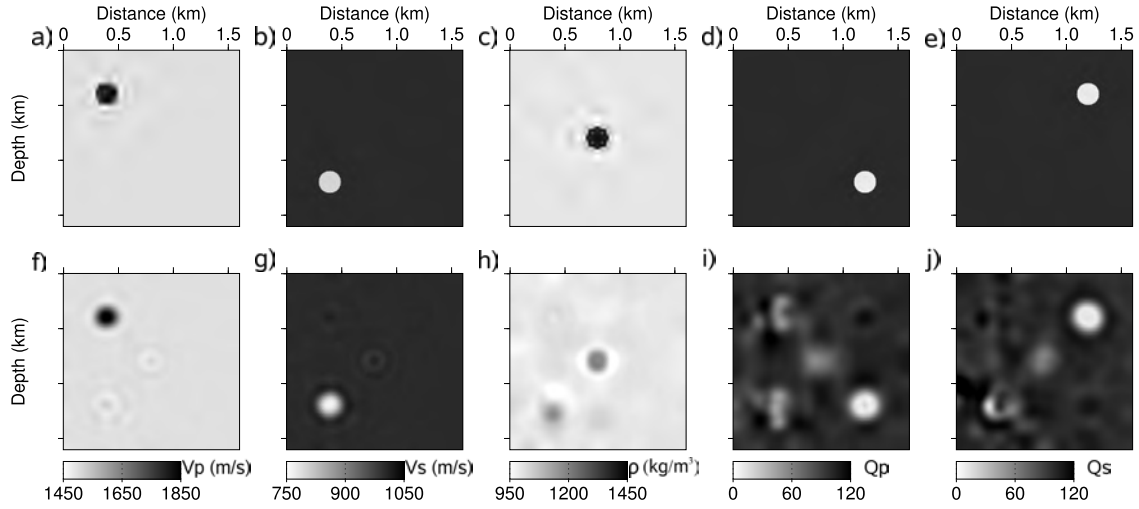


FIG. 3.7 – Example of multi-parameter inversion. (a-e) True models : (a) V_P , (b) V_S , (c) ρ , (d) Q_P and (e) Q_S models. (f-j) Reconstructed FWI models : (f) λ , (g) μ , (h) ρ , (i) Q_P and (j) Q_S .

TAB. 3.2 – Physical parameters in the background and heterogeneities for the canonical inversion test.

Parameter	Background value	Heterogeneity value
V_P (m/s)	1500	1800
V_S (m/s)	1000	800
ρ (kg/m ³)	1000	1400
Q_P	100	10
Q_S	100	10

3.5 Parallelization and performance

To tackle imaging of realistically sized targets, the computationally demanding FWI method needs to be implemented in parallel to exploit multiple CPU facilities, also with a large amount of available core-memory, compared to sequential execution.

3.5.1 Forward problem

The forward problem of FWI, as the resolution of the linear system of equation (3.3) for multiple sources (right-hand-side (RHS)), is performed with the massively parallel direct solver MUMPS (Amestoy *et al.*, 2006), which performs a LU decomposition of the matrix \mathbf{A} , based on a multi-frontal approach (Duff & Reid, 1983). The parallelization allows for the speeding up of the factorization by more than one order of magnitude, compared to sequential execution, and for the distribution of the LU factors in the core memory over all of the processors (Sourbier *et al.*, 2009b), which makes this quite efficient for solving large problems without intensive I/O resources. The MUMPS software is based on Message Passing Interface (MPI) standards for

applications with distributed memory architectures. The workflow of the MUMPS algorithm can be summarized as : (1) analysis of the \mathbf{A} matrix pattern, followed by reordering of the matrix to limit the fill-in during factorization. The analysis step is performed sequentially on the master processor with the current MUMPS version (but does not modify the workflow if done in parallel) ; (2) parallel factorization of the \mathbf{A} matrix with dynamic pivoting ; (3) parallel forward and backward substitutions to obtain solution vectors \mathbf{u} from RHS vectors \mathbf{b} .

3.5.2 Overview of the parallel inverse problem

In the framework of iterative FWI, the physical properties of the model change at each iteration, but the mesh geometry is not changed. Therefore, the \mathbf{A} matrix keeps the same pattern, allowing a single analysis phase in the whole inversion procedure. This feature is interesting because the matrix analysis is the only time-demanding step performed sequentially on the master process. After the substitution step, the solution vectors \mathbf{u} are returned by MUMPS in a distributed form over the in-core memory of the processors. However, the distribution of the solutions performed by MUMPS does not always provide a well-balanced load for the processors, because the MUMPS strategy is to optimize LU decomposition with accurate dynamic pivoting and reduced fill in. The distribution is therefore not suitable for the application of the projection operator $\mathbf{P}_V \mathbf{P}_{P_0}$ to the local in-core solutions (i.e., the solutions at the different nodes for one wavefield of a triangular cell are not systematically stored in-core on the same processor). To overcome this limitation, the solutions are re-ordered following well-balanced mesh partitioning, performed with the METIS software (Karypis & Kumar, 1999). The MUMPS-distributed solutions are mapped to the METIS decomposition with MPI point-to-point communications, before the application of operator $\mathbf{P}_V \mathbf{P}_{P_0}$. This mapping/projection step represents about 25 % of the MUMPS substitution time. The forward and back-propagated wavefields are stored in the in-core memory for all sources. The gradient and the diagonal of the Hessian are therefore efficiently computed in parallel with local in-core wavefields for the local subdomain associated with each process, before being centralized on the master processor for perturbation-model building. Of note, the assemblage and storage of the sparse \mathbf{A} matrix are performed by the master process, as the storage of the mesh-related tables, the wavefield solutions at the receiver positions for computing the objective function, and the composite residual sources for adjoint-wavefield computation. To avoid prohibitive memory allocation on the master processor, the master process is not involved in the factorization and substitution/projection steps, the wavefield storage, and the gradient/Hessian computation.

3.5.3 Two parallelism levels

The scalability of parallel direct solvers is intrinsically limited because of the large amount of communications between the MPI processes, and the memory overhead generated by the number of MPI process involved in the factorization (Sourbier *et al.*, 2009b). Therefore, the single parallelization level with MPI can be not adapted for large-scale applications. A second parallelism level, based on multi-threading, can, however, be efficiently implemented on multi-core architectures. The main computationally demanding parallel tasks of the FWI algorithm are the forward problem resolutions (factorization and substitutions/projection) and the building of the gradient and diagonal Hessian.

The MUMPS solver uses the Basic Linear Algebra Subroutines (BLAS) library mainly for level 2 and 3 operations (matrix/vector, and matrix/matrix operations). Several BLAS libraries are available in multi-thread distributions, which allows an efficient low-level of parallelization for multi-core architectures with shared memory.

The gradient and diagonal Hessian computations have several nested loops embedded (Algorithm 3.2). The outer loop over the model parameters can be easily parallelized with shared-memory thread technology (as with standard OpenMP) to speed up the loop on the available CPUs without duplicating or exchanging data.

An alternative to classical parallelization using one MPI process per CPU is, therefore, to run one MPI process on n CPUs, sharing the same memory, and to launch n threads per MPI-process during the BLAS operations, and the gradient and diagonal Hessian computations. These two levels of parallelism can mitigate the memory overhead of the factorization, the network usage for intensive communications, and the computational time for LU factors.

3.5.4 Performance

Simulations are performed on a distributed-memory architecture composed of 18 nodes, each of which includes two quad-core 2.7 GHz Opteron processors with 64 Gbytes of shared memory. The nodes are interconnected by an Infiniband ConnectX network. This cluster is located at Geoazur Institute, and it provides 144 CPUs and 1.125 Tbytes of memory.

A first test corresponds to a middle-scale application that is focused on the elastic Vallhall model (Figure 3.4). A 1 042 720-cell mesh was designed with the Triangle software (Shewchuk, 1996). The mesh is composed of a thin unstructured layer on top for P_1 interpolation, and a large zone of equilateral structured triangles for P_0 in depth. This mesh allows frequencies of up to 7 Hz to be modeled in the elastic case. The substitution/projection phase is performed for 315 RHSs, corresponding to 315 explosive sources. Figure 3.8 shows the evolution of the computational times and the memory usage for the main inversion steps as a function of the number of CPUs, using one- and two-level parallelism.

We clearly see a decrease in the computational time when the CPU number increases, although the efficiency is limited by the intrinsic scalability of the MUMPS direct solver. The best trade-off between performance and computational resources is obtained with 24 CPUs, with one MPI-process per CPU used. When two levels of parallelism are used with shared-memory threads, the in-core memory required by the factorization decreases, an appealing feature if the memory available is limited. However, the memory saving is at the expense of the computational time, particularly for the substitution/projection step. Here, I postulate that for this middle-scale application, the speed-up provided by the multi-thread computation is not significant with respect to the time overheads due to thread launching, leading to a limited time performance for this application.

A second test focuses on a large-scale elastic crustal model (Figure 3.6). An h -adaptive unstructured mesh was designed with Triangle for DG P_2 modeling up to 6.5 Hz, leading to 528 080 cells. To mimic the SFJ-OBS experiment (Operto *et al.*, 2006), 93 sources were considered, corresponding to 93 ocean-bottom seismometers, for the substitution/projection phase. Figure 3.9 summarises the computational times and memory usage for this test. Note that the time for gradient building is negligible compared to that of the MUMPS steps, due to the small number of RHSs and the cell number in the mesh. For this large-scale application that

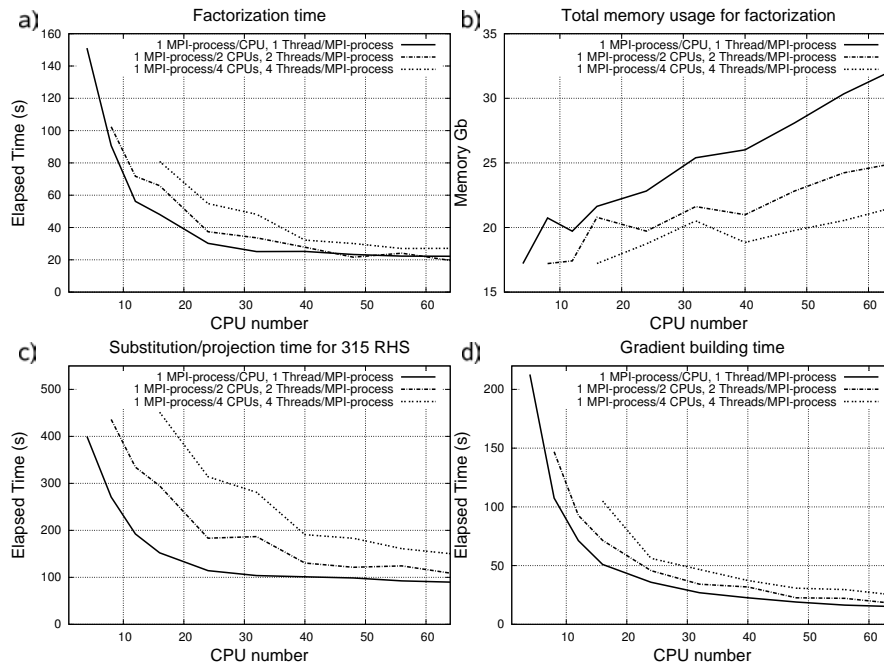


FIG. 3.8 – Computational time and memory use for the main inversion tasks for a middle-scale application, as a function of the number of CPUs, and the parallel strategy, as indicated

requires a large amount of distributed memory to store the 200 Gbytes of LU factors, shared-memory-threads technology is the alternative choice to the one-level parallelism, to minimize the number of MPI processes, and therefore the memory overheads during LU factorization. The data obtained with multiple threads show that both the computational time for the factorization and the substitution/projection steps can be decreased, while saving a significant amount of core memory.

The results of these two applications suggest that the two-level parallelism over MPI processes and shared-memory threads that are embedded in our FWI algorithm provide an efficient and flexible tool to tackle realistic size target imaging. Depending on the computing facilities available and the size of the case study, a judicious combination of the two levels of parallelism can be found for optimal performances in terms of computational time and core-memory usage.

3.6 Conclusion

A massively parallel algorithm was developed to perform frequency-domain full-waveform modeling and inversion for imaging 2D visco-elastic media. The forward problem of inversion, as the resolution of the elastodynamics equation in the frequency-domain, is solved using a low-order finite element discontinuous Galerkin method. This allows complex topographies and high velocity-contrasts to be taken into account, and to use unstructured h -adaptive triangular meshes combined with p -adaptive interpolations. The linear system, resulting from this discretization is solved with the parallel MUMPS direct solver, allowing the LU factors to be stored in a distributed form over the processors. The wavefield solutions are also stored in a

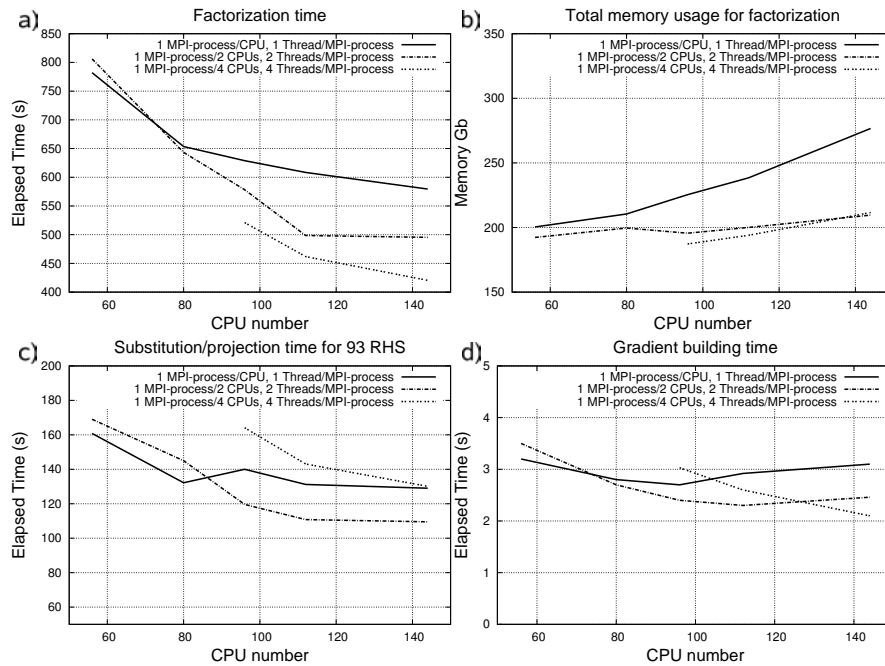


FIG. 3.9 – Computational time and memory use for the main inversion tasks for a large-scale application, as a function of the number of CPU, and the parallel strategy, as indicated.

distributed form, following mesh partitioning for a well-balanced workload over the processors. The inverse problem of the FWI is solved with a linearized least-squares optimization scheme. The implementation of the gradient of the objective function has been optimized for computational time and memory saving. A quasi-Newton L-BFGS optimization has been implemented to estimate the inverse of the Hessian matrix at a negligible computational cost, improving the reconstruction of several classes of parameter (P-wave and S-wave speeds, density, and attenuation factors Q_P and Q_S). A two-parallelism level through MPI-processes and shared-memory threads has been implemented for optimal use of the computational time and core-memory resources, which will depend on the computer facilities.

This FWI algorithm will be used to tackle the imaging of realistically sized targets at various scales, from near-surface geotechnic applications to crustal-scale exploration. Ongoing studies include the extension to anisotropic (vertically transverse isotropic) media and the adaptation to teleseismic configurations for lithospheric imaging.

Acknowledgments

This study was funded by the SEISCOPE consortium (<http://seiscope.unice.fr>), which is sponsored by BP, CGG-VERITAS, EXXON-MOBIL, SHELL and TOTAL, and by Agence Nationale de la Recherche (ANR) under project ANR-05-NT05-2-42427. The author is grateful to S. Operto (Géoazur, CNRS) and J. Virieux (LGIT, Université Joseph Fourier) for fruitful and stimulating discussions on full waveform modeling and inversion, on parallel implementation, and for manuscript review. Many thanks also go to Serge Gratton (CERFACS and CNES) for

interesting discussions on quasi-Newton optimization methods. The LU factorization of the impedance matrix was performed with MUMPS (<http://graal.ens-lyon.fr/MUMPS/index.html>). The mesh generation was performed with the help of Triangle (<http://www.cs.cmu.edu/~quake/triangle.html>). The Garvin problem analytical solutions were computed with code provided by U. Iturrarán and F. J. Sánchez-Sesma. Access to the high-performance computing facilities of the Mesocentre SIGAMM computer center, the IDRIS national center (project 082280) and the GEOAZUR computing center, provided the required computer resources. Many thanks go to J. Kommendal and L. Sirgue from BP, for providing the elastic synthetic models of Valhall.

Conclusion

Dans ce chapitre, j'ai illustré l'implémentation des algorithmes de résolution des problèmes direct et inverse pour l'imagerie des paramètres visco-élastiques en deux dimensions, dans un contexte de calcul parallèle.

Le système élastodynamique 2D en domaine fréquentiel est discrétisé par une méthode d'éléments finis Galerkin discontinus. Le choix des ordres d'interpolation utilisés, ainsi que le type de maillage, sont directement liés à la complexité structurale des applications d'inversion, permettant de trouver le meilleur compromis précision/coût de calcul. La résolution du système linéaire issu de la discrétisation est assurée par le solveur direct parallèle MUMPS. Cette parallélisation permet à la fois de stocker les facteurs LU de la matrice sous une forme distribuée en mémoire vive des différents processeurs, et d'autre part, d'accélérer les phases de factorisation et de substitution, diminuant ainsi le temps de calcul.

Pour assurer la résolution du problème inverse, une procédure d'inversion hiérarchique à deux niveaux sur des groupes de fréquences et des facteurs de régularisations de fréquences complexes est proposée. Une optimisation de la formulation du gradient permet de limiter le coût de stockage en mémoire des champs directs et adjoints, ainsi que le temps de calcul du gradient. L'utilisation de la méthode quasi-Newton L-BFGS estime économiquement l'inverse du Hessien, permettant de mettre à l'échelle les différentes classes de paramètres imagées simultanément, sans surcoût de calcul.

Le processus d'inversion est mené en parallèle. Les solutions des champs d'ondes du problème direct sont tout d'abord stockées sous une forme distribuée. Ces solutions sont ensuite réordonnées selon un découpage en domaines physiques du maillage afin d'optimiser la charge des processeurs. Le gradient et l'estimation du Hessien sont assemblés en parallèles à partir des solutions locales de chaque sous-domaine.

L'implémentation d'un double niveau de parallélisme par des processus MPI et des threads à mémoire partagée permet d'optimiser le temps de calcul et l'utilisation mémoire en fonction des applications et des ressources informatiques disponibles. La flexibilité de cet algorithme permet d'envisager des applications d'inversion visco-élastiques à différentes échelles, de la proche surface aux échelles crustales.

Deuxième partie

Applications

Chapitre 4

Exemples synthétiques

Sommaire

4.1	Application terrestre : Le modèle SEG/EAGE Overthrust	142
4.1.1	Résumé	142
4.1.2	Introduction	143
4.1.3	Méthode et algorithme	145
4.1.4	Préconditionnement des données et stratégies multi-échelles	148
4.1.5	Application à un modèle canonique	150
4.1.6	Application à un modèle synthétique terrestre	151
4.1.7	Discussion	164
4.1.8	Conclusion	168
4.1.9	Test complémentaire : combien de composantes sont nécessaires?	169
4.2	Application marine : Le modèle Valhall	172
4.2.1	Présentation du modèle élastique de Valhall	172
4.2.2	Configuration de l'inversion	172
4.2.3	Inversion acoustique de données élastique	174
4.2.4	Inversion élastique hiérarchique	176
4.3	Conclusion	177

Dans ce premier chapitre applicatif, nous allons nous intéresser aux particularités et difficultés de l'inversion des formes d'onde pour la reconstruction des paramètres élastiques. A partir de modèles synthétiques réalistes, dans leurs tailles et leurs complexités structurales, nous allons mener des études de sensibilité afin de quantifier les difficultés associées à l'imagerie multiparamètres de données complexes.

Une première étude s'attachera à un cas d'imagerie en milieu terrestre. Gelis *et al.* (2007) montrent que l'inversion en domaine fréquentiel de données terrestres, combinant ondes de volume et ondes de surface, est relativement instable et largement dominée par la forte énergie des ondes de surface. Nous allons nous attacher à quantifier cette dépendance aux ondes de surface qui détermine le succès ou l'échec de l'inversion. Par ailleurs, nous proposerons une approche hiérarchique d'inversion à travers plusieurs niveaux de régularisation, afin de minimiser l'impact des phénomènes de propagations complexes donnant son caractère très non-linéaire au problème.

Dans une deuxième étude, un cas d'imagerie en milieu marin sera abordé. Les configurations marines permettent de s'affranchir des problèmes liés aux ondes de surfaces. Cependant, le cas d'étude abordé est caractérisé par une profondeur d'eau très faible et un contraste des propriétés physiques très faible au fond de mer. Ces deux propriétés confèrent aux données des caractéristiques particulières, notamment une signature faible des ondes S dans le champ d'ondes, qu'il est nécessaire de prendre en compte pour une reconstruction robuste des paramètres élastiques.

4.1 Application terrestre : Le modèle SEG/EAGE Overthrust

Cette partie d'application terrestre est présentée sous la forme d'un article publié dans la session spéciale « New Advances on Seismic Imaging and Inversion » (Novembre-Décembre 2008) de la revue *Geophysics*.

<p>Seismic imaging of complex onshore structures by two-dimensional elastic frequency-domain full-waveform inversion</p>

<p>R. Brossier, S. Operto and J. Virieux <i>Geophysics</i>, 2009, 74, 6, in press</p>
--

4.1.1 Abstract

Quantitative imaging of the elastic properties of the subsurface at depth is essential for civil engineering applications and for oil and gas reservoir characterization. A realistic synthetic example provides for an assessment of the potential and limits of two-dimensional elastic full-waveform inversion of wide-aperture seismic data, for recovering high resolution P-wave and S-wave velocity models of complex onshore structures. Full-waveform inversion of land data is challenging, because of the increased non-linearity introduced by free-surface effects, such as the propagation of surface waves in the heterogeneous near-surface. Moreover, the short wavelengths of the shear wavefield require an accurate S-wave velocity starting model if low frequencies are not available in the data. Different multiscale strategies with the aim of mitigating these non-linearities are here evaluated. Massively parallel full-waveform inversion is implemented in the frequency domain. The numerical optimization relies on a limited-memory quasi-Newton algorithm that out-performs the more classic preconditioned conjugate-gradient algorithm. The forward problem is based upon a discontinuous Galerkin method on triangular mesh, which allows accurate modeling of free-surface effects. Sequential inversions of increasing frequencies define the most natural level of hierarchy in the multiscale imaging. In the case of land data involving surface waves, the regularization introduced by hierarchical frequency inversions is not enough for adequate convergence of the inversion. A second level of hierarchy implemented with complex-valued frequencies is necessary, and provides convergence of the inversion towards acceptable P-wave and S-wave velocity models. Among the possible strategies for sampling frequencies in the inversion, successive inversions of slightly overlapping frequency groups proves to be the most reliable when compared with the more standard sequential inversion of single frequencies. This suggests that simultaneous inversion of multiple frequencies is critical when considering complex wave phenomena.

4.1.2 Introduction

Quantitative imaging of the elastic properties of the subsurface is essential for oil and gas reservoir characterization, and for the monitoring of CO₂ sequestration with time-lapsed acquisitions. Indeed, fluids and gas have significant effects on the elastic properties of the subsurface in terms of the Poisson-ratio anomalies. This quantitative imaging is also required for near-surface imaging in the framework of civil engineering applications, because the shear properties of the shallow weathered layers have a strong impact on the elastic wavefield. Moreover, at the near-surface scale, short propagation times do not easily allow the separation in time of the body waves and the surface waves. In this case, filtering or muting surface waves is not easy, and both types of waves need to be involved in the imaging, which requires solving the elastic-wave equation. Evolution of acquisition systems towards wide-aperture/wide-azimuth geometries and multi-component recordings is another motivation behind the development of elastic-imaging methods, because the occurrence of P-to-S mode conversion is dominant at wide-aperture angles.

Reservoir characterization is classically performed by amplitude-versus-offset analysis in the prestack domain (e.g., Jin *et al.*, 2000; Buland & Omre, 2003). An alternative is the full-waveform inversion (FWI) of the elastic wavefield recorded by multiple components with the aim of reconstructing the P-wave and S-wave velocity (V_P and V_S) models of the subsurface (or other related parameters, such as impedances, if density is involved in the inversion) with a resolution limit of the order of half a wavelength. The misfit between the recorded and the modeled wavefields is minimized through resolution of a numerical optimization problem (Tarantola, 1987; Nocedal & Wright, 1999). The FWI forward problem is based on the complete solution of the full (two-way) wave equation.

One drawback is that elastic FWI is a computationally challenging problem. Recent advances in high performance computing on large-scale distributed memory platforms allow two-dimensional (2D) problems of representative sizes to be tackled, while the 3D problems start to be investigated only in the acoustic case (Sirgue *et al.*, 2007; Ben-Hadj-Ali *et al.*, 2008; Vigh & Starr, 2008; Warner *et al.*, 2008). Moreover, the complexity of the elastic wavefield at wide aperture makes the inverse problem highly non-linear and ill-posed. This is even more dramatic for onshore applications where surface waves with high energy as well as body waves can be used in the optimization process. FWI is conventionally solved with local optimization (linearized) approaches, which makes the inversion sensitive to the limited accuracy of the starting model. As such, it is essential to use realistic synthetic case studies to investigate under which conditions the non-linearity of the elastic FWI can be mitigated.

Most of the recent applications of FWI to real data have been performed under acoustic approximation (Pratt & Shipp, 1999; Hicks & Pratt, 2001; Ravaut *et al.*, 2004; Gao *et al.*, 2006; Operto *et al.*, 2006; Bleibinhaus *et al.*, 2007). Although reliable results can be obtained with acoustic approximation if judicious data pre-processing and inversion preconditioning is applied (Brenders & Pratt, 2007b), elastic FWI is desirable for applications that involve the detection of fluids and gas, and for CO₂ sequestration. Moreover, acoustic FWI can lead to erroneous models when applied to elastic data when the velocity models show high velocity contrasts and when no specific FWI pre-processing and tuning is applied to the data (Barnes & Charara, 2008).

Only a few applications of elastic FWI have been presented in the literature recently. In the early applications, Crase *et al.* (1990, 1992) applied elastic FWI to real land and marine

reflection data using limited offsets. In this framework, the FWI was applied as quantitative migration processing for imaging P and S impedances. With the benefit provided by wide-aperture data to build the large and intermediate wavelengths of the subsurface recognized by Mora (1987, 1988), acoustic and elastic FWI evolved as an attempt to build high resolution velocity models. Shipp & Singh (2002) performed 2D time-domain FWI of a small wide-aperture marine data subset recorded by a long streamer (12-km long). Although the forward problem was solved using the elastic wave equation, only the V_P parameter was reconstructed during FWI, assuming an empirical relationship between V_P and V_S and between V_P and density. They designed a hierarchical multi-step approach based on layer-stripping, and offset and time windowing, where the aim was to mitigate the non-linearity of the inverse problem. Sears *et al.* (2008) designed a similar multi-step strategy to perform elastic time-domain FWI of multi-component ocean-bottom-cable data. Their strategy, was based on selection of data component, parameter class and arrival type (by time windowing), and it revealed itself especially useful when the amount of P-to-S conversion was small at the sea bottom, which makes the reconstruction of the V_S model particularly ill-posed. Choi & Shin (2008) and Choi *et al.* (2008) applied elastic frequency-domain FWI to onshore and offshore versions of the synthetic Marmousi2 model (Martin *et al.*, 2006), respectively. They successfully imaged the model using a velocity-gradient starting model and a very low starting frequency (0.16 Hz). Shi *et al.* (2007) applied elastic time-domain FWI to marine data collected from a gas field in western China. They successfully imaged a zone of Poisson-ratio anomalies associated with gas layers. Accurate starting V_P and V_S models were built from the P-wave and P-SV-wave velocity analysis and from *a priori* information of several well logs along the profile. Gelis *et al.* (2007) implemented a 2D elastic frequency-domain FWI using the Born and the Rytov approximations for the linearization of the inverse problem. They highlighted the dramatic footprint of the surface waves on the imaging of small inclusions in homogeneous background models. To mitigate this footprint, they only involved body waves during the early stages of the inversion, by only selecting short-offset traces.

The present study presents a 2D massively parallel elastic frequency-domain FWI algorithm based on a discontinuous Galerkin (DG) forward problem (Brossier *et al.*, 2008), and its application to a realistic synthetic onshore case study. The aim of this application was to assess whether surface waves and body waves recorded by wide-aperture acquisition geometries can be jointly inverted to build high resolution V_P and V_S of complex onshore structures. We implement FWI in the frequency domain, which presents some distinct advantages with respect to the time-domain formulation (Pratt & Worthington, 1990; Pratt, 1999; Sirgue & Pratt, 2004). The inverse problem can be solved with a local optimization approach using either a conjugate gradient or a quasi-Newton method (Nocedal & Wright, 1999). The gradient of the objective function is computed with the adjoint-state technique (Plessix, 2006). Successive inversions of increasing frequency provide a natural framework for multiscale imaging algorithms. Moreover, by sacrificing the data redundancy of multifold acquisitions, the inversion of a limited number of frequencies can be enough to build reliable velocity models, provided the acquisition geometry spans over sufficiently long offsets. This limited number of frequencies can be efficiently modeled in the 2D case for multiple shots once the impedance matrix that results from discretization of the frequency-domain wave equation has been factorized through a LU decomposition (Marfurt, 1984; Nihei & Li, 2007). Finally, attenuation can be implemented in the forward problem in a straightforward way, and without extra computational cost, by using complex velocities. The main drawback of the frequency-domain FWI formulation arises from the difficulty of time

windowing of the modeled data when inverting one or a few sparsely sampled frequencies at a time. Time windowing allows a selection of specific arrivals to be included in the various stages of the inversion. A last resort is the use of complex-valued frequencies, which damp arrivals starting at a given travelttime (Shin *et al.*, 2002).

In the next section of the present study, we briefly review the theory of frequency-domain elastic full-wavefield modeling and inversion. In the following section, we review several multiscale strategies to mitigate the non-linearity of the elastic inverse problem. These strategies involve two nested levels of hierarchy over frequencies and aperture angles in the inversion algorithm. The effectiveness of these strategies is first illustrated with a simple two-parameter elastic problem with a free surface. Then we have applied the elastic frequency-domain FWI algorithm to a realistic synthetic example, for the reconstruction of a dip section of the SEG/EAGE Overthrust model, assuming a constant Poisson ratio. The results of the different analyses show that simultaneous inversions of multiple frequencies and data preconditioning by time damping are critical to obtain reliable results when surface waves propagating in a heterogeneous near-surface are present in the elastic wavefield. We will also illustrate the improvements provided by quasi-Newton algorithms, compared to more conventional conjugate-gradient approaches.

4.1.3 Method and algorithm

4.1.3.1 Forward problem

Two-dimensional elastic frequency-domain FWI requires computation of the frequency solution of the elastic P-SV equations in heterogeneous media. We present a short review of the P_0 DG method used in this study. The reader is referred to Brossier *et al.* (2008) for more details.

We consider the first-order hyperbolic system where both particle velocities (V_x, V_z) and stresses ($\sigma_{xx}, \sigma_{zz}, \sigma_{xz}$) are unknown quantities, as described by the system :

$$\begin{aligned}
 -i\omega V_x &= \frac{1}{\rho(\mathbf{x})} \left\{ \frac{\partial \sigma_{xx}}{\partial x} + \frac{\partial \sigma_{xz}}{\partial z} \right\} + f_x \\
 -i\omega V_z &= \frac{1}{\rho(\mathbf{x})} \left\{ \frac{\partial \sigma_{xz}}{\partial x} + \frac{\partial \sigma_{zz}}{\partial z} \right\} + f_z \\
 -i\omega \sigma_{xx} &= (\lambda(\mathbf{x}) + 2\mu(\mathbf{x})) \frac{\partial V_x}{\partial x} + \lambda(\mathbf{x}) \frac{\partial V_z}{\partial z} - i\omega \sigma_{xx_0} \\
 -i\omega \sigma_{zz} &= \lambda(\mathbf{x}) \frac{\partial V_x}{\partial x} + (\lambda(\mathbf{x}) + 2\mu(\mathbf{x})) \frac{\partial V_z}{\partial z} - i\omega \sigma_{zz_0} \\
 -i\omega \sigma_{xz} &= \mu(\mathbf{x}) \left\{ \frac{\partial V_x}{\partial z} + \frac{\partial V_z}{\partial x} \right\} - i\omega \sigma_{xz_0},
 \end{aligned} \tag{4.1}$$

where the Lamé coefficients that describe the medium are denoted by λ and μ , the density by ρ , and the angular frequency by ω . Source terms are either point forces (f_x, f_z) or applied stresses ($\sigma_{xx_0}, \sigma_{zz_0}, \sigma_{xz_0}$). i is the purely imaginary term defined by $i = \sqrt{-1}$. Only isotropic media are considered in this study.

Equations (4.1) are discretized and solved with the DG method (Käser & Dumbser, 2006). We used low-order P_0 interpolation, which corresponds to piecewise constant velocity and stress fields, and physical parameters in each cell. The DG method is applied to the weak formulation of the system (4.1), and the partial derivatives are computed through numerical fluxes

exchanged at interfaces between cells. The perfectly-matched layers (PML) method are used for absorbing boundary conditions along the edges of the model (Berenger, 1994). An explicit free-surface boundary condition for arbitrary complex topographies is implemented on top of the model by canceling fluxes of normal stresses along the boundary that consist of edges of triangles. A sufficient level of accuracy for FWI can be obtained using the DG P_0 method with ten to fifteen cells per minimum shear wavelength in regular equilateral meshes (Brossier *et al.*, 2008). These regular meshes will be used for the simulations presented below. Note that extension of the DG method to higher order interpolations is required for modeling in unstructured meshes and for arbitrarily combining different interpolations during one simulation (e.g., Dumbser & Käser, 2006).

Discretization of equations (4.1) leads to solving the linear system :

$$\mathbf{A}\mathbf{v} = \mathbf{s}, \quad (4.2)$$

where the coefficients of the impedance matrix \mathbf{A} , namely, the forward modeling operator, depends on the modeled frequency and the medium properties. Vector \mathbf{s} represents the source term. Vector \mathbf{v} represents the unknowns for particle velocities and stress in each cell. Note that only the vertical and horizontal particle velocity wavefields will be inverted in this study.

4.1.3.2 Inverse Problem

FWI is an optimization problem which can be recast as a linearized least-squares problem that attempts to minimize the misfit between the recorded and the modeled wavefields (Tarantola, 1987). The inverse problem can be formulated in the frequency domain (Pratt & Worthington, 1990) and the associated objective function to be minimized is defined by :

$$\mathcal{C}^{(k)} = \frac{1}{2}\Delta\mathbf{d}^\dagger\mathbf{W}_d\Delta\mathbf{d} = \frac{1}{2}\Delta\mathbf{d}^\dagger\mathbf{S}_d^\dagger\mathbf{S}_d\Delta\mathbf{d}, \quad (4.3)$$

where $\Delta\mathbf{d} = d_{obs} - d_{calc}^{(k)}$ is the data misfit vector, the difference between the observed data d_{obs} and the modeled data $d_{calc}^{(k)}$ computed in the model $\mathbf{m}^{(k)}$. Superscript \dagger indicates the adjoint (transposed conjugate) and \mathbf{S}_d is a diagonal weighting matrix applied to the misfit vector to scale the relative contributions of each of its components. k is the iteration number.

The gradient $\mathcal{G}^{(k)}$ of the objective function is given by :

$$\mathcal{G}^{(k)} = \mathcal{R}\left\{\mathbf{J}^t\mathbf{W}_d\Delta\mathbf{d}^*\right\} \quad (4.4)$$

where \mathbf{J} is the Fréchet derivative matrix. Adjoint-state formalism allows efficient computation of the gradient with the back-propagation technique, without explicit computation of \mathbf{J} (Plessix, 2006). This leads to the following expression of the gradient with respect to the parameter m_i (Pratt *et al.*, 1998; Gelis *et al.*, 2007) :

$$\mathcal{G}_{m_i}^{(k)} = \mathcal{R}\left\{\mathbf{v}^t\left[\frac{\partial\mathbf{A}^t}{\partial m_i}\right]\mathbf{A}^{-1}\mathbf{W}_d\Delta\mathbf{d}^*\right\}, \quad (4.5)$$

which shows that the gradient can be recast as a product between the incident wavefields \mathbf{v} and the back-propagated wavefields $\mathbf{A}^{-1}\mathbf{W}_d\Delta\mathbf{d}^*$, using residuals at receiver positions as a composite source. Therefore, only two forward problems per shot are required for gradient building. In

equation (4.5), we exploited the reciprocity of the Green's functions to remove the transpose operator in the expression of the back-propagated residuals ($\mathbf{A}^{-1t}\mathbf{W}_d\Delta\mathbf{d}^* = \mathbf{A}^{-1}\mathbf{W}_d\Delta\mathbf{d}^*$).

The radiation pattern of the diffraction by the model parameter m_i is denoted by $\partial\mathbf{A}/\partial m_i$. The analysis of these radiation patterns suggests that V_P and V_S parameterization is that which is optimal for large diffraction angles (*i.e.*, for wide-angle reflections), while the I_p and I_s impedances should provide a better decoupling between the two classes of parameters for small diffraction angles (*i.e.*, short-angle reflections) (Tarantola, 1986).

A second-order Taylor expansion of the objective function provides the perturbation model $\delta\mathbf{m}$, which minimizes the objective function assumed to be locally parabolic, expressed as :

$$\mathbf{B}^{(k)}\delta\mathbf{m} = -\mathcal{G}^{(k)}, \quad (4.6)$$

where $\mathbf{B}^{(k)}$ is the Hessian of the objective function. Due to the cost of the computation of $\mathbf{B}^{(k)}$, Newton and Gauss-Newton methods are generally not considered for realistic size problems (Pratt *et al.*, 1998). Steepest-descent or conjugate-gradient methods preconditioned by the diagonal terms of an approximate Hessian are more conventionally used (Pratt *et al.*, 1998; Operto *et al.*, 2006). Shin *et al.* (2001) used the diagonal part of the pseudo-Hessian, a less computationally demanding approximation of the truncated Hessian. Accounting for the Hessian accelerates the convergence of the inversion and improves the resolution of the imaging by correctly scaling and deconvolving the gradient by geometrical amplitude and limited-bandwidth effects.

In the following, we use a quasi-Newton method for the FWI problem. The limited-memory Broyden-Fletcher-Goldfarb-Shanno (L-BFGS) method is commonly used in numerical optimization to solve large-scale, non-linear problems (Nocedal, 1980). As indicated, L-BFGS is a limited memory version of the BFGS method, and it appears to be one of the most robust and efficient limited-memory quasi-Newton algorithms (Nocedal & Wright, 1999). The quasi-Newton L-BFGS(n) method estimates curvature information contained in the Hessian matrix from a limited number (n) of gradient difference vectors and model difference vectors associated with the n most recent iterations (n is usually chosen between 3 and 20). The iterative process is preconditioned by an initial guess of the Hessian, which is typically the diagonal terms of an approximate Hessian. Using the Sherman-Morrison-Woodbury formula (Nocedal & Wright, 1999), at each iteration the L-BFGS algorithm computes an improved estimation of the inverse $\mathbf{H}^{(k)}$ of the Hessian matrix $\mathbf{B}^{(k)}$. Therefore, resolution of the linear system (4.6) is avoided and the perturbation model is directly inferred from :

$$\delta\mathbf{m} = -\mathbf{H}^{(k)}\mathcal{G}^{(k)}. \quad (4.7)$$

The double-loop recursive algorithm designed by Nocedal (1980) does not explicitly build and store $\mathbf{H}^{(k)}$, but directly computes the right-hand side of equation (4.7) with additions, differences and inner products of the vectors.

Finally, the model is updated with the perturbation vector :

$$\mathbf{m}^{(k+1)} = \mathbf{m}^{(k)} + \alpha\delta\mathbf{m}, \quad (4.8)$$

where α denotes the step length, which minimizes the objective function. In this study, we estimate α by parabola fitting. When the Hessian matrix is taken into account in the inversion, perturbation models associated with each parameter class computed with equation (4.7) are

correctly scaled with respect to each other (Nocedal & Wright, 1999). Therefore, an estimation of only one step length α is necessary to minimize the objective function for the multi-parameter classes.

A subspace method (Sambridge *et al.*, 1991) has also been tested that led to similar results ; however, this was more computationally intensive because it required extra forward problem simulations.

4.1.3.3 Parallel implementation

The massively parallel direct solver MUMPS (Amestoy *et al.*, 2006) is based on a multi-frontal method (Duff & Reid, 1983) and it is used to solve the linear system that results from the discretization of the forward problem (Equation (4.2)). Parallel LU factorization of the \mathbf{A} matrix allows the speeding up of the factorization by more than one order of magnitude, compared to sequential execution. Moreover, LU factors are stored in a distributed form over the in-core memory of the processors (Sourbier *et al.*, 2009b), making this quite efficient for solving larger problems without intensive I/O resources. After the substitution phase, the multiple solutions are distributed over the processors following the domain decomposition driven by the distribution of the LU factors. Equation (4.5) shows that the gradient computation is a weighted product of the forward problem solutions ; namely, the incident wavefields and the back-propagated residual wavefields. This product can be performed easily in parallel by assigning one processor to each subdomain. To improve the load balancing over the processors, MUMPS-distributed solutions are re-ordered with message passing interface (MPI) point-to-point communications before gradient computation, using a well-balanced mesh partitioning that is performed with METIS software (Karypis & Kumar, 1999). The gradient and the initial estimation of Hessian are therefore efficiently computed in parallel, before being centralized on the master processor for perturbation-model building.

4.1.4 Full-waveform inversion data preconditioning and multiscale strategies

CPU-efficient frequency-domain FWI is generally carried out by successive inversions of single frequencies, by proceeding from the low frequencies to the higher ones (Pratt & Worthington, 1990; Sirgue & Pratt, 2004). This defines a multiresolution framework that helps mitigation of the non-linearity of the inverse problem associated with high frequency cycle-skipping artifacts. CPU-efficient algorithms can be designed by selecting a few coarsely sampled frequencies such that the wavenumber redundancy that results from dense sampling of frequencies and aperture angles is decimated. This strategy, which is referred to as the sequential inversion approach in the following, has proven to be effective for several applications of acoustic FWI (Ravaut *et al.*, 2004; Operto *et al.*, 2006; Brenders & Pratt, 2007a). However, it might lack robustness when complex wave phenomena are present. For example, reconstruction of a low velocity layer in the Overthrust model was improved when several frequencies were inverted simultaneously, rather than successively, during acoustic FWI (Sourbier *et al.*, 2009b). More significant wave propagation effects are expected in elastic FWI because of conversions, dispersive surface waves, and frequency-dependent attenuation.

More robust, but more computationally expensive, multiscale FWI schemes should be designed by partially preserving the redundancy of multifold seismic data. The first scheme, referred to as the Bunks approach in the following, is an adaptation in the frequency domain of the

time-domain approach of Bunks *et al.* (1995). It consists of successive inversions of overlapping frequency groups. The first group contains only the starting frequency, and one higher frequency is added from one group to the next. The second approach, which is referred to as the simultaneous inversion approach in the following, consists of successive inversions of slightly overlapping frequency groups. The main difference from the Bunks approach is that several frequencies are simultaneously inverted at each inversion stage, and that overlapping between two next-frequency groups is minimized. Definition of the frequency bandwidth of each group should be driven by the best compromise between the need to avoid high frequency cycle-skipping artifacts, and the need to simultaneously invert as many frequencies as possible to stabilize the inversion.

Non-linearity of FWI can also be efficiently mitigated by selecting a subset of specific arrivals (*i.e.*, early arrivals, reflected phases) in the data by time windowing (e.g., Sheng *et al.*, 2006; Sears *et al.*, 2008). Frequency-domain wave modeling is not as flexible as the time-domain system for the preconditioning of the data by time windowing, since a limited number of frequencies is conventionally processed at a given step of the inversion. Data preconditioning can, however, be applied in the frequency domain by means of complex frequencies ($\omega + i\gamma$), which is equivalent to damp seismograms in the time domain (Shin *et al.*, 2002; Brenders & Pratt, 2007b). The Fourier transform of a signal $f(t)$ damped in time by $\exp^{-\gamma(t-t_0)}$ is given by :

$$F(\omega + i\gamma)\exp^{\gamma t_0} = \int_{-\infty}^{+\infty} f(t)\exp^{-\gamma(t-t_0)}\exp^{-i\omega t}dt, \quad (4.9)$$

where the damping can be applied from an arbitrarily arrival time t_0 , which usually corresponds to the first-arrival time.

Time damping applied from the first-arrival time can be viewed as a heuristic way to select aperture angles of P-waves in the data. Arrivals located in time close to the first-arrival times correspond to wide-aperture P-wave events, while the later arriving phase from the first-arrival traveltimes corresponds to shorter aperture P-wave events and converted waves. Since aperture angle is an additional parameter to frequency in the control of the spatial resolution of FWI (Miller *et al.*, 1987; Wu & Toksöz, 1987; Sirgue & Pratt, 2004), aperture selection can be exploited to implement a second level of hierarchy in FWI in addition to that naturally introduced by frequency selection. Another benefit expected from complex-valued frequencies is the damping during the early FWI iterations of converted P-S waves, free-surface multiples and surface waves, which introduces additional non-linearities into the inversion. In practice, this second level of hierarchy can be implemented by progressively relaxing the time damping during each frequency-group inversion. Note that when complex-valued frequencies are used in FWI, the amplitude term $\exp^{\gamma t_0}$ must be introduced in the weighting matrix \mathbf{S}_d (equation (4.3)) to apply the same damping to the partial derivative wavefields and to the data residuals, which are cross-correlated during gradient building.

We designed the elastic frequency-domain FWI algorithm such that each of the above-mentioned strategies can be easily tested. The FWI algorithm implements the two above-mentioned nested hierarchical levels through an outer loop over the frequency groups and an inner loop over the damping terms (*i.e.*, the imaginary part of the complex-valued frequencies). Here, we should note that a frequency group is a set of real frequencies that are simultaneously inverted. A third loop is over inversion iteration, and a fourth is over the frequencies of the group. An updated model is produced after one iteration of the inversion involving one frequency

Algorithm 4.1 Two-level hierarchical FWI

```

1: for frequency_group = group_1 to group_n do
2:   for data_damping = high_damping to low_damping do
3:     while (NOT convergence AND iter < niter_max) do
4:       for frequency = frequency_1 to frequency_n do
5:         Propagate wavefields from sources
6:         Compute Residuals  $\Delta \mathbf{d}$  and Cost function  $\mathcal{C}_{\mathbf{m}}^{(k)}$ 
7:         Backpropagate Residuals from receivers
8:       end for
9:       Build gradient vector  $\mathcal{G}_{\mathbf{m}}^{(k)}$ 
10:      if iter = 1 then
11:        Compute diagonal of Pseudo-Hessian
12:      end if
13:      Compute perturbation vector  $\delta \mathbf{m}$  with Quasi-Newton L-BFGS method
14:      Define optimal step length  $\alpha$  by parabola fitting
15:      Update model  $\mathbf{m}^{(k+1)} = \mathbf{m}^{(k)} + \alpha \delta \mathbf{m}$ 
16:    end while
17:  end for
18: end for

```

group and one damping term. The major steps of the frequency-domain FWI algorithm can be summarized in algorithm 4.1.

4.1.5 Application to a canonical model

FWI is an ill-posed, non-linear problem even for apparently simple models involving few parameters. Mulder & Plessix (2008) illustrated analytically the non-linearity of 3D acoustic FWI with a 1D velocity gradient model defined by two parameters. The objective function showed multiple local minima around the true global minimum. We consider here a similar two-parameter problem for a 2D elastic model with a free surface on top of it.

A 1D V_P gradient model defined by $V_P(z) = V_0 + \eta z$ is considered, where V_0 is the P-wave velocity at the surface and η is the vertical velocity gradient. The S-wave velocity is inferred from the P-wave velocity by considering a constant Poisson ratio of 0.24. A vertical point-force is located just below the free surface, and 350 receivers record horizontal and vertical particle velocities on the free surface along a 17-km-long profile.

The objective function is plotted as a function of V_0 and η for the 5.8 Hz frequency in Figure 4.1a. The global minimum is located at the coordinates of the true model ($V_0 = 4$ km/s, $\eta = 0.35$ s⁻¹). Cross sections along the η axis for $V_0 = 4$ km/s and along the V_0 axis for $\eta = 0.35$ s⁻¹ show the non-convex shape of the objective function (Figure 4.1). Multiple local minima are present, particularly on the η section, even for this simple gradient model and the low frequency content in the data.

We repeated the same simulations for damped data using $\gamma = 3.33$ (equation (4.9)). Using this data preconditioning, the objective function is now convex (Figures 4.1b and 4.1c). The convex shape of the objective function should ensure the convergence of the inversion towards

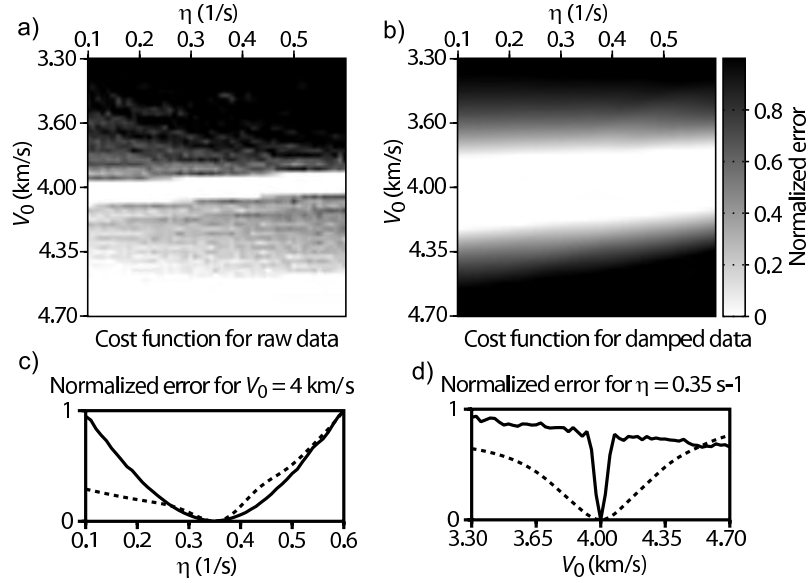


FIG. 4.1 – Objective function as a function of the two model parameters for a) full data set and b) damped data set. c) Cross-sections of maps shown above for $V_0=4$ km/s. The solid line and the dashed line correspond to the full data-set and the damped data-set, respectively. d) Cross-sections of maps shown above for $\eta=0.35$ s $^{-1}$. The solid line and the dashed line correspond to the full data-set and the damped data set respectively.

the global minimum of the objective function for all of the starting models sampled in Figure 4.1.

4.1.6 Onshore synthetic case study

4.1.6.1 SEG/EAGE Overthrust model and experimental setup

We considered a $20 \text{ km} \times 4.65 \text{ km}$ dip section of the SEG/EAGE Overthrust model to assess the potential of 2D elastic frequency-domain FWI for imaging complex onshore structures (Aminzadeh *et al.*, 1997)(Figure 4.2a). The Overthrust model was a $20 \text{ km} \times 20 \text{ km} \times 4.65 \text{ km}$ 3D acoustic model that represents an onshore complex thrust sedimentary succession constructed on top of a basement block. Several faults and channels were present in the model, as well as a complex weathering zone on the surface. For elastic FWI, a V_S model was built from the V_P model using a constant Poisson ratio of 0.24. A uniform density of 1000 kg.m^{-3} was considered and assumed to be known during the inversion. A free surface was set on top of the model.

The onshore, wide aperture survey consisted of 199 explosive sources spaced every 100 m and located 25 m below the free surface. All of the shots were recorded by 198 vertical and horizontal geophones, which were spaced every 100 m on the surface. The vertical and horizontal components of particle velocity were used as the dataset for the elastic FWI. The data were computed with the same algorithm for both observed and computed data in inversion. The

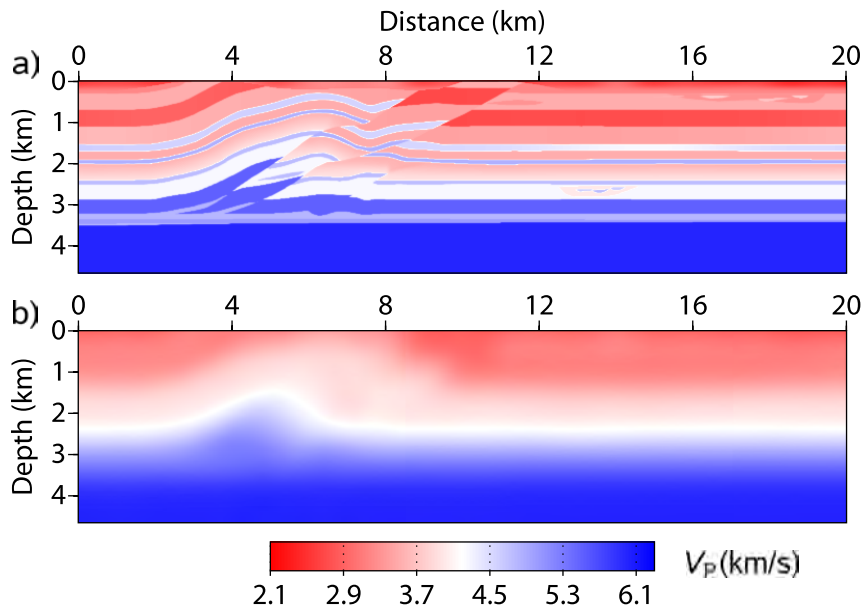


FIG. 4.2 – a) Dip section of the synthetic SEG/EAGE Overthrust model. P-wave velocity is depicted. b) Starting model used for elastic FWI.

source signature was assumed to be known in FWI. An elastic shot gather is shown in Figure 4.3(a-b) for the horizontal and vertical components of particle velocity. The corresponding shot gather computed when an absorbing boundary condition is set on top of the model is shown for comparison in Figure 4.3(c-d), to highlight the additional wave complexities introduced by free-surface effects (*i.e.*, surface waves and body-wave reflexion from the free surface). Of note, the high amplitudes of the surface waves dominate the wavefield especially on the vertical component (Figure 4.3(a-b)). Starting V_P and V_S models for FWI were computed by smoothing the true velocity models with a 2D Gaussian function of vertical and horizontal correlation ranges of 500 m (Figure 4.2b). This starting model was proven to be accurate enough to image the Overthrust model by 2D acoustic frequency-domain FWI using a realistic starting frequency of 3.5 Hz (Soubier *et al.*, 2009b). For elastic inversions presented hereafter, we used a lower starting frequency of 1.7 Hz. Using a starting frequency of 3.5 Hz for elastic FWI led to a deficit of long wavelengths in the V_S models, which made the inversion converge towards a local minimum.

The different behavior of acoustic and elastic FWI for the Overthrust case study highlights the increased sensitivity of elastic FWI with respect to the limited accuracy of the starting models. Five discrete frequencies (1.7, 2.5, 3.5, 4.7 and 7.2 Hz) were used for the elastic FWI. This frequency sampling should allow continuous sampling of the wavenumber spectrum according to the criterion of Sirgue & Pratt (2004). In the following, we shall consider the three different strategies to manage frequencies described in the section “Full-waveform inversion data preconditioning and multiscale strategies” : successive inversion of single frequencies, successive inversion of frequency groups of increasing bandwidth, and successive inversion of slightly overlapping frequency groups. For each frequency group, the inversion is subdivided into two steps. In the first step, no offset-dependent gain was applied to the data. Although we scale the

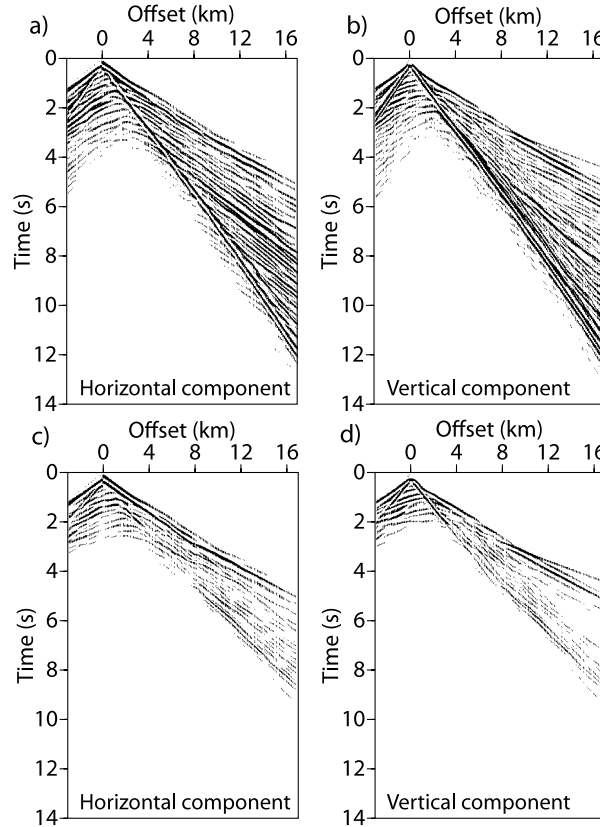


FIG. 4.3 – Seismograms computed in Overthrust model for (a) horizontal and (b) vertical components of particle velocity. The shot is located at a horizontal distance of 3 km. A free-surface was set on top of the model. (c-d) As for Figure 4.3(a-b), except that an absorbing boundary condition was implemented on top of the model.

gradient by the diagonal terms of the Hessian, we observed some lack of reconstruction in the deep part of the model, suggesting that the near-offset traces have a dominant contribution in the objective function. This layer-stripping effect may provide additional regularization of the inversion, in addition to that provided by the frequency and aperture angle selections. In the second step, we applied a quadratic gain with offset to the data, to strengthen the contribution of long-offset data in the inversion, and hence to improve the imaging of the deeper part of the model.

During the two-step inversion, the coefficients of the diagonal weighting operator \mathbf{S}_d were respectively given by :

$$\begin{aligned} \mathbf{S}_d &= \exp^{\gamma t_0} \\ \mathbf{S}_d &= \exp^{\gamma t_0} |offset|^2, \end{aligned} \quad (4.10)$$

with the reminder that the coefficients $\exp^{\gamma t_0}$ account for the offset-dependent time damping ; equation (4.9). For all of the tests presented below, except for the first, we used five damping factors per frequency to precondition the data ($\gamma = 1.5, 1.0, 0.5, 0.1, 0.033$). A shot gather computed for the first four damping factors is shown in Figure 4.4, to illustrate the amount

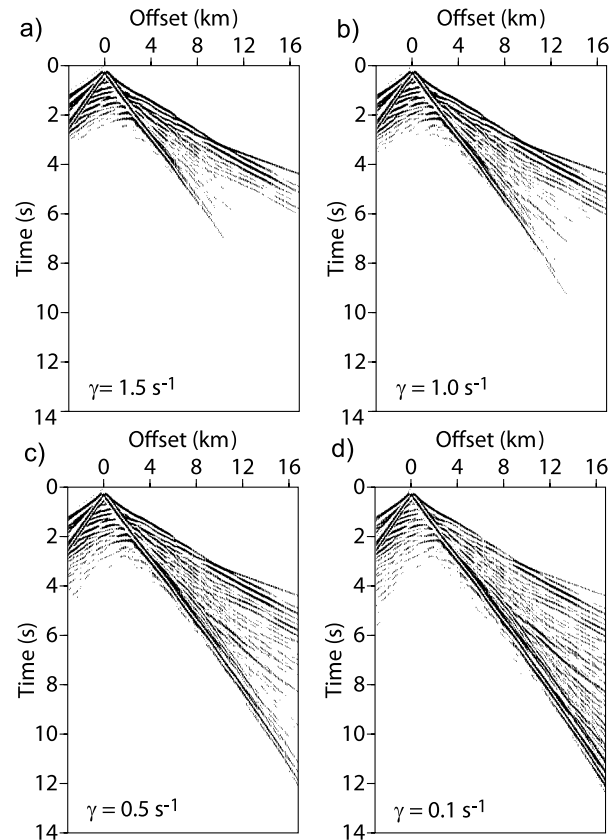


FIG. 4.4 – Seismograms for vertical component of particle velocity computed in the dip section of the Overthrust model using four values of imaginary frequency. a) $\gamma = 1.5 \text{ s}^{-1}$, b) $\gamma = 1.0 \text{ s}^{-1}$, c) $\gamma = 0.5 \text{ s}^{-1}$, d) $\gamma = 0.1 \text{ s}^{-1}$. Time-damping was applied from the first-arrival traveltimes to preserve long-offset information.

of information preserved in the data. Note how the high damping limits the offset range over which surface waves are seen. Inversion was also regularized by Gaussian smoothing of the perturbation model, the aim of which was to cancel high frequency artifacts in the gradient. The diagonal terms of the pseudo-Hessian matrix (Shin *et al.*, 2001) provided an initial guess of the Hessian for the L-BFGS algorithm without introducing extra computational costs during gradient building. Five differences of gradients and models vectors were used for the L-BFGS algorithm. The model parameters for inversion were V_P and V_S , which are suitable for wide aperture acquisition geometries (Tarantola, 1986). The loop over the inversion iteration of one complex-valued frequency group was stopped when a maximum iteration number of 45 was reached or when the convergence criterion was reached (relative decrease of two successive cost functions lower than $5 \cdot 10^{-5}$). The schedule of the frequencies and damping terms used in the sequential approach, the Bunks approach and the simultaneous approach are outlined in Table 4.1.

In the following, we shall quantify the data misfit for each test with the normalized misfit

TAB. 4.1 – Inversion parameters for the sequential, Bunks, and simultaneous approaches. FG : frequency group number ; F (Hz) : frequencies within a frequency group ; γ (1/s) : damping factors (imaginary part of frequency).

FG	Sequential - F(Hz)	Bunks - F(Hz)	Simultaneous - F(Hz)	γ (1/s)
1	1.7	1.7	1.7, 2.5, 3.5	1.5, 1, 0.5, 0.1, 0.033
2	2.5	1.7, 2.5	3.5, 4.7, 7.2	1.5, 1, 0.5, 0.1, 0.033
3	3.5	1.7, 2.5, 3.5		1.5, 1, 0.5, 0.1, 0.033
4	4.7	1.7, 2.5, 3.5, 4.7		1.5, 1, 0.5, 0.1, 0.033
5	7.2	1.7, 2.5, 3.5, 4.7, 7.2		1.5, 1, 0.5, 0.1, 0.033

TAB. 4.2 – Final $L2$ misfit and model quality mq for the different reconstructed models. Seq : Sequential approach ; Bunks : Bunks approach ; Sim : Simultaneous approach ; without FS : Sequential approach without free-surface effects ; PCG : Sequential approach computed with PCG optimization.

Test	Data misfit \bar{C}	V_P mq	V_S mq
Seq	$4.12 \cdot 10^{-1}$	$5.54 \cdot 10^{-2}$	$6.47 \cdot 10^{-2}$
Bunks	$1.54 \cdot 10^{-1}$	$5.22 \cdot 10^{-2}$	$5.33 \cdot 10^{-2}$
Sim	$1.46 \cdot 10^{-1}$	$5.03 \cdot 10^{-2}$	$5.39 \cdot 10^{-2}$
Without FS	$9.03 \cdot 10^{-2}$	$4.09 \cdot 10^{-2}$	$3.89 \cdot 10^{-2}$
PCG	$6.71 \cdot 10^{-1}$	$5.56 \cdot 10^{-2}$	$6.89 \cdot 10^{-2}$

\bar{C} , defined by :

$$\bar{C} = \frac{\sum_{i=1}^5 \|\Delta \mathbf{d}_i(\mathbf{m}_f)\|_2}{\sum_{i=1}^5 \|\Delta \mathbf{d}_i(\mathbf{m}_0)\|_2} \quad (4.11)$$

where $\Delta \mathbf{d}_i(\mathbf{m}_f)$ denotes the data misfit vector for the i^{th} frequency and for the final FWI model \mathbf{m}_f , and \mathbf{m}_0 denotes the starting model shown in Figure 4.2b.

The FWI model quality will be quantified by :

$$mq = \frac{1}{N} \left\| \frac{\mathbf{m}_f - \mathbf{m}_{true}}{\mathbf{m}_{true}} \right\|_2 \quad (4.12)$$

where \mathbf{m}_{true} denotes the exact model either for V_P or V_S , and N is the number of grid points in the computational domain. The normalized misfit and the model quality for the different tests presented hereafter are outlined in Table 4.2.

4.1.6.2 Raw data inversion

A first inversion test was performed without data damping (*i.e.*, without using complex-valued frequencies), which implies that all of the arrivals were involved in the inversion. The five frequencies (Table 4.1) were successively inverted with the sequential approach. The FWI V_P

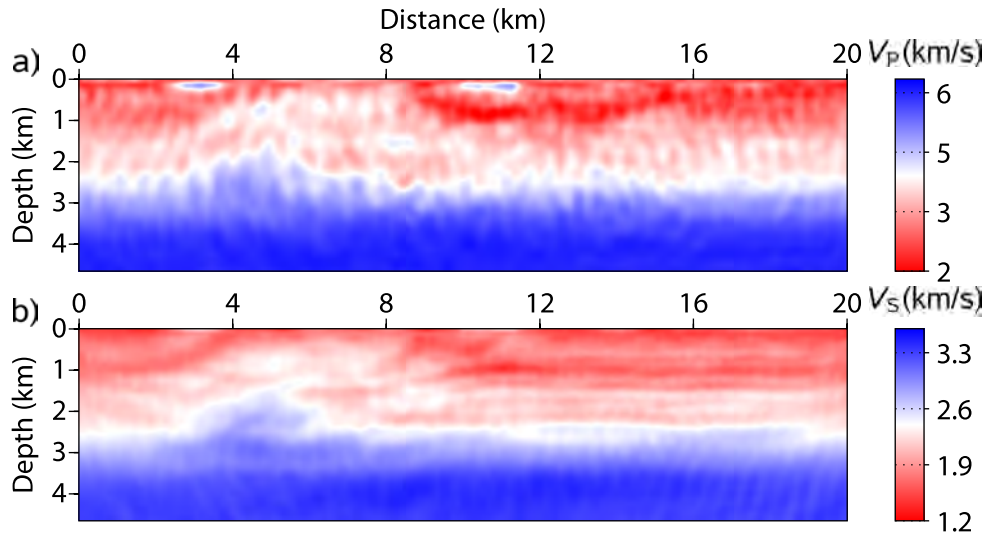


FIG. 4.5 – Sequential inversion of raw data - (a) V_P and (b) V_S models after frequency of 7.2 Hz.

and V_S models after inversion are shown in Figure 4.5. The inversion clearly failed to converge towards the true models for both of the V_P and V_S parameters, even at low frequencies.

4.1.6.3 Successive single-frequency inversions of damped data

We repeated the previous experiment, except that the five damping terms ($\gamma = 1.5, 1.0, 0.5, 0.1, 0.033$) were used to stabilize the inversion (Table 4.1). The final FWI V_P and V_S models are shown in Figure 4.6. Contrary to the previous experiment, most of the layers were now successfully reconstructed. Comparison between 1D vertical profiles extracted from the true model, the starting model and the FWI models shows a reliable estimate of velocity amplitudes despite a low maximum frequency of 7.2 Hz (Figure 4.7). To take into account the limited bandwidth effect of the source in the FWI model appraisal, we also plotted the vertical profiles of the true models after low-pass filtering at the theoretical resolution of FWI for a maximum frequency of 7.2 Hz : the true models were converted from depth to time using the velocities of the starting model, and low-pass filtered with a cut-off frequency of 7 Hz, before conversion back to the depth domain. We noted that the V_S model is more affected by spurious artifacts than the V_P model, especially in the deep part of the model. This may be due to a deficit of small wavenumbers in the V_S models, resulting from the shorter propagated wavelengths, which makes the reconstruction of the V_S parameter more non-linear. Secondly, we saw some inaccuracies in the reconstruction of both the V_P and V_S parameters in the shallowest parts of the models (Figure 4.7). The resulting residuals of the surface waves and reflections from the free surface may have been erroneously back-projected in the deeper part of the model, leading to the above-mentioned noise (Figure 4.8). The final normalized L_2 misfit computed for the five frequencies was $4.12 \cdot 10^{-1}$. The V_P and V_S model qualities are $5.54 \cdot 10^{-2}$ and $6.47 \cdot 10^{-2}$, respectively.

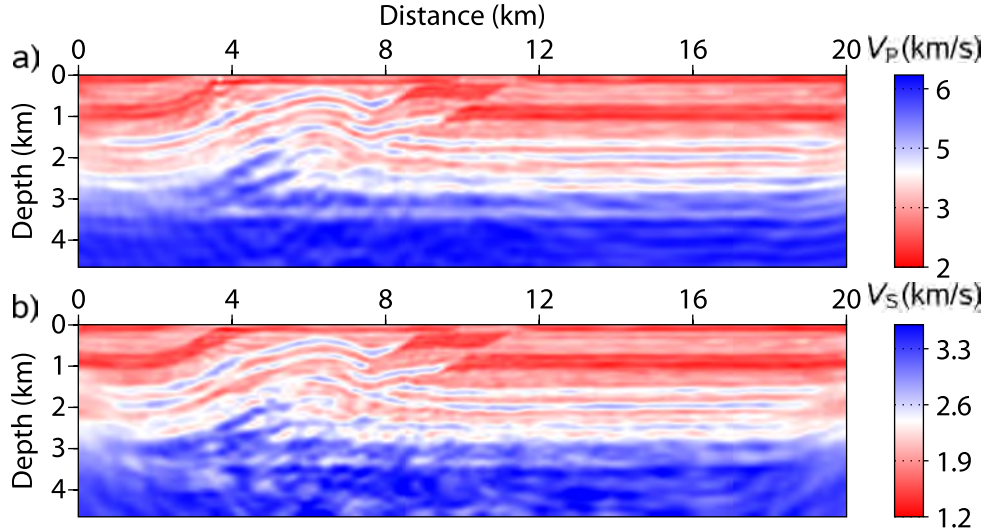


FIG. 4.6 – Sequential inversion of damped data - (a) V_P and (b) V_S models after inversion. The L-BFGS algorithm was used for optimization. Five frequency components were inverted successively. Five damping coefficients were successively used for data preconditioning during each mono-frequency inversion.

4.1.6.4 Full-waveform inversion without free-surface effects

To assess the footprint of surface waves and free-surface reflections on elastic FWI, we inverted the data computed in the Overthrust model with an absorbing boundary condition on top of it, instead of a free surface. The same inversion process was used as in the previous section (sequential approach with damped data). The final FWI V_P and V_S models were very close to the low-pass filtered versions of the true models, and they were not affected by any spurious artifacts (Figure 4.9). The V_P and V_S model qualities are $4.09 \cdot 10^{-2}$ and $3.89 \cdot 10^{-2}$, respectively. Comparison with the previous results (Figure 4.6) illustrates the substantial increase of non-linearity introduced by surface waves and free-surface reflections in elastic FWI.

4.1.6.5 Simultaneous multi-frequency inversion of damped data

We now investigated the influence of simultaneous multiple-frequency inversion strategies for FWI improvement. We first considered the Bunks approach for FWI. The different frequency groups and damping terms are outlined in Table 4.1. Each mono-frequency dataset belonging to a frequency group was computed with a unit Dirac source wavelet, which implies that each frequency of a group has a similar weight in the inversion. This is equivalent to inversion of deconvolved data (i.e., data with a flat amplitude spectrum).

The final FWI V_P and V_S models shown in Figure 4.10 are slightly improved compared to those of the sequential approach (Figure 4.6). The improvements are more obvious on the vertical profiles of the final FWI models (Figure 4.11). Near-surface instabilities in the V_P and V_S models were mitigated, although not fully removed, and the estimations of the velocity amplitudes were improved in most parts of the model (compare Figures 4.7 and 4.11). Mitigation of

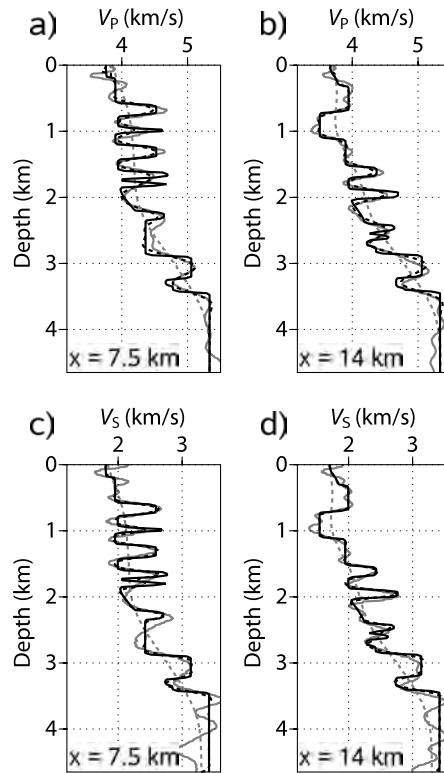


FIG. 4.7 – Sequential inversion of damped data - Vertical profiles for the V_P (a-b) and V_S (c-d) parameters. Profiles (a)-(c) and (b)-(d) are at horizontal distances of 7.5 km and 14 km respectively. Profiles of the starting and the true models are plotted with dashed gray lines and solid black lines, respectively. A low-pass filtered version of the true model at the theoretical resolution of FWI is plotted with a dashed black line for comparison with the FWI results. The profiles of the FWI models of Figure 4.6 are plotted with solid gray lines.

the near-surface instabilities translated into a significant misfit reduction for the surface waves and free-surface reflections (Figure 4.12). The final normalized $L2$ misfit decreased in this case to $1.54 \cdot 10^{-1}$. The V_P and V_S qualities are $5.22 \cdot 10^{-2}$ and $5.33 \cdot 10^{-2}$, respectively.

In a second step, we considered the simultaneous approach implemented by successive inversions of two overlapping frequency groups (Table 4.1). The frequency bandwidth of each group was chosen such that cycle-skipping artifacts were avoided. For example, we tried to simultaneously invert the five frequencies listed in Table 4.1. In this case, the inversion failed to converge towards acceptable models. Note that the computational cost of the simultaneous approach is similar to that of the sequential one if the same convergence rate for the two approaches is assumed. The total number of iterations in the simultaneous approach is less important, because the iterations are performed over fewer frequency groups at the expense of more factorization and substitution phases per frequency group. The extra cost of the simultaneous approach is only due to the overlap between the frequency groups.

The final FWI V_P and V_S velocity models shown in Figure 4.13 were improved with respect to the velocity models produced by the sequential approach (Figure 4.6) and the Bunks approach (Figure 4.10), especially for the V_S velocity model in the thrust zone. The vertical

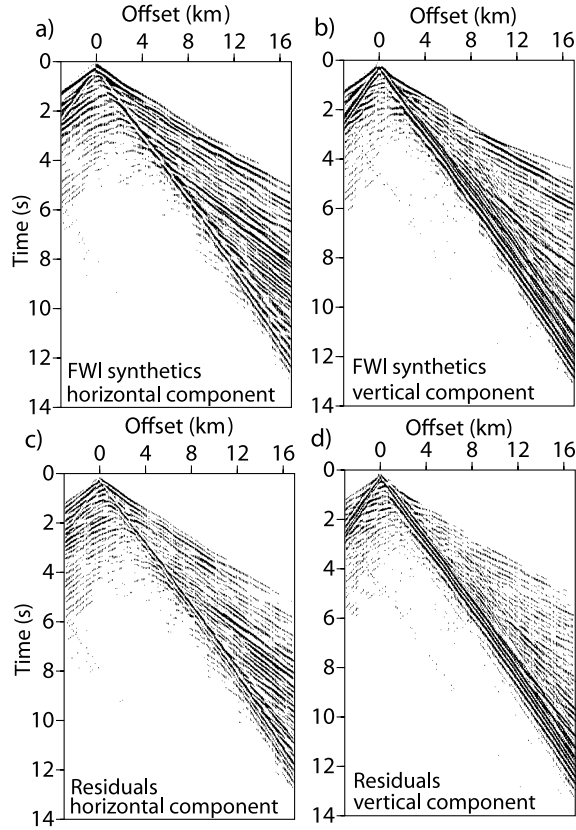


FIG. 4.8 – Sequential inversion of damped data - Seismograms computed in the FWI models of Figure 4.6 for shot located at a horizontal distance of 3 km. a) Horizontal component. b) Vertical component. (c-d) Residuals between seismograms computed in the true V_P and V_S models (Figure 4.3a) and in the FWI models of Figure 4.6. c) Vertical component. d) Horizontal component.

profiles extracted from the final FWI models do not show near-surface instabilities anymore (Figure 4.14), which allowed a significant data misfit reduction for the surface waves and free-surface reflections (Figure 4.15). A significant misfit reduction of the wide aperture arrivals recorded at large offsets was also seen. The final $L2$ misfit is $1.46 \cdot 10^{-1}$. The V_P and V_S model qualities are $5.03 \cdot 10^{-2}$ and $5.39 \cdot 10^{-2}$, respectively. We note, however, slightly underestimated velocity amplitudes in the deep part of the V_P and V_S models at the thrust location (see below 3 km depth in Figure 4.14a, c). We attribute this amplitude deficit to a slower convergence of the simultaneous approach when compared to that of the sequential one, which results from the fact that more information is simultaneously inverted in the simultaneous approach. The imaging was further improved by decreasing the frequency interval by a factor of 2 within each frequency group (five frequencies instead of three frequencies per group). This resampling contributes to the strengthening of the spectral redundancy in the imaging. Close-ups of the V_P and V_S models centered on the thrust zone show how the resolution and the signal-to-noise ratio of the velocity models were still improved by involving more frequencies in one inversion iteration (Figure 4.16). Note that using five frequencies instead of three in each group leads to a factor of $5/3$ in computational costs.

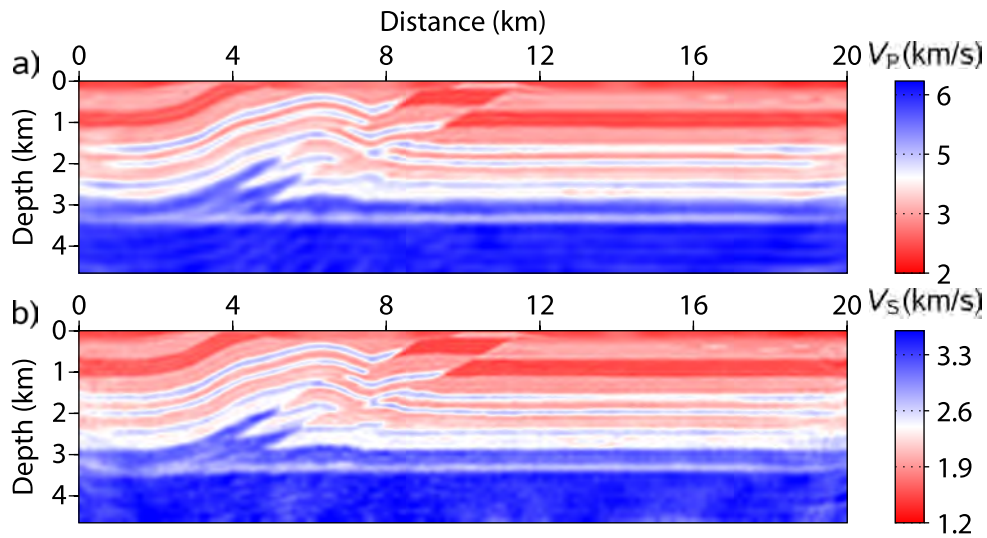


FIG. 4.9 – Sequential inversion without free-surface effects - (a) V_P and (b) V_S models. The models can be compared with that of Figures 4.6 to assess the footprint of free-surface effects on elastic FWI.

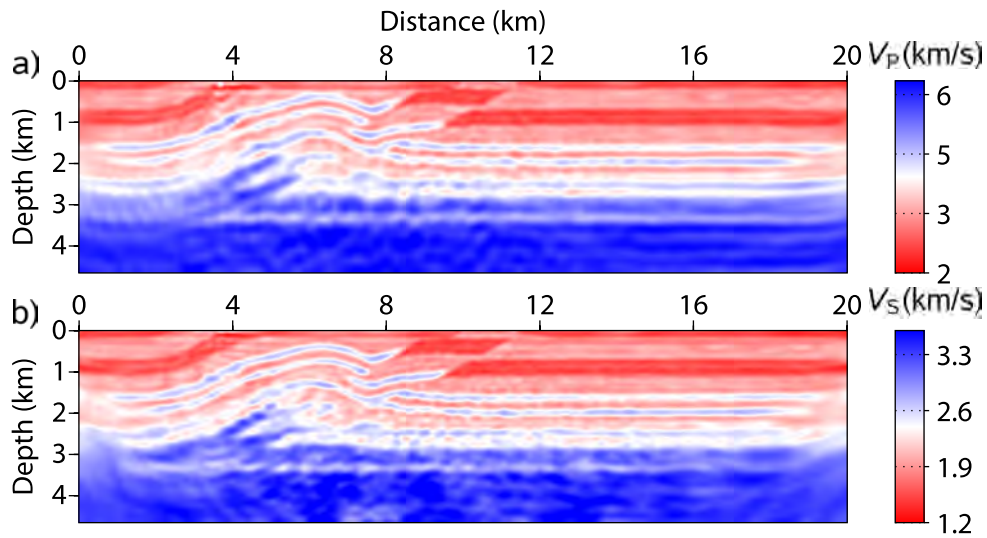


FIG. 4.10 – Bunks inversion - (a) V_P and (b) V_S models obtained with the frequency-domain adaptation of the multiscale approach of Bunks *et al.* (1995).

4.1.6.6 L-BFGS versus preconditioned conjugate-gradient optimizations

The sequential approach was applied using a preconditioned conjugate-gradient (PCG) algorithm for numerical optimization. Preconditioning is performed by scaling the gradient by the diagonal terms of the pseudo-Hessian matrix. The same frequencies and damping terms were used as for the L-BFGS run based on the sequential approach (Table 4.1). The velocity models recovered with the PCG algorithm are shown in Figure 4.17 and can be compared with

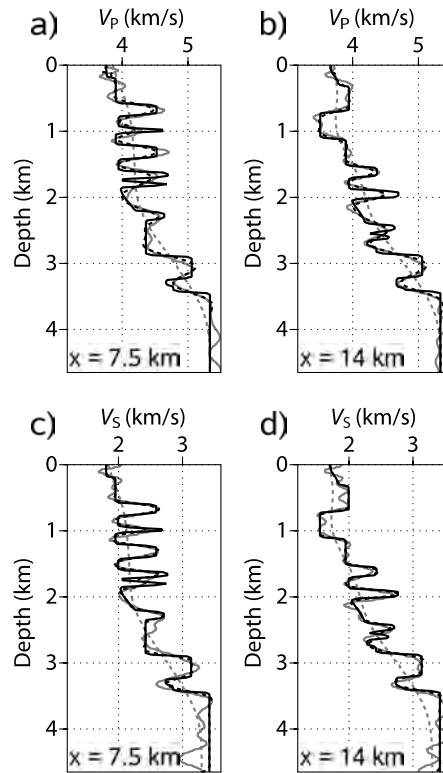


FIG. 4.11 – Same as Figure 4.7, but for the profiles extracted from the models recovered by the Bunks approach (Figure 4.10).

the corresponding L-BFGS ones (Figure 4.6). Amplitude estimations and focusing of structures were improved by the L-BFGS algorithm (Figure 4.18), leading to sharper models. The improvements in the model resolution and the quantitative estimate of the model parameters can be attributed to the approximate estimation of the off-diagonal terms of the Hessian performed by the L-BFGS algorithm. These off-diagonal elements help to deconvolve the models from limited-bandwidth effects resulting from the limited source bandwidth and the limited extent of the acquisition geometry (e.g., Pratt *et al.*, 1998).

Figure 4.19 shows the objective functions as a function of iteration number for the L-BFGS and PCG algorithms. L-BFGS provides accelerated and improved convergence when compared to PCG. The PCG convergence level (*i.e.*, the minimum value of the objective function reached during optimization) cannot reach that of L-BFGS because the off-diagonal information of Hessian estimated by L-BFGS cannot be retrieved by more iterations of PCG. Note that the final L_2 misfit of PCG is $6.71 \cdot 10^{-1}$, whereas that of L-BFGS is $4.12 \cdot 10^{-1}$. These two issues open promising applications of L-BFGS for computationally challenging problems, such as for 3D FWI.

4.1.6.7 Computational aspect

All of the simulations were performed on the cluster of the SIGAMM computer center, which is composed of a 48-node cluster, with each node comprising 2 dual-core 2.4-GHz Op-

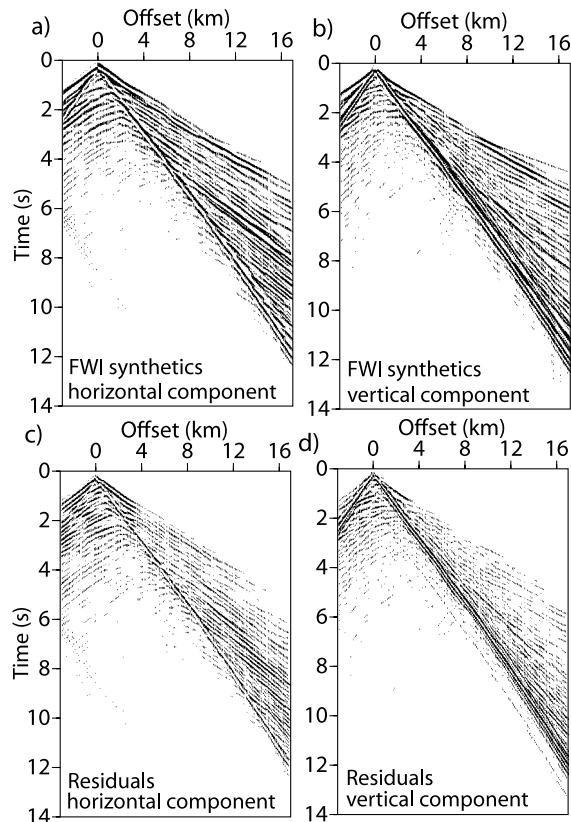


FIG. 4.12 – Same as Figure 4.8, but for seismograms computed in the FWI models recovered by the Bunks approach (Figure 4.10).

teron processor, providing 19.2 Gflops peak performance per node. This computer has a distributed memory architecture, where each node has 8 GBytes of RAM. The interconnection network between processors is Infiniband 4X. Twenty-four processors were used for each simulation, leading to the best compromise between execution time and numerical resources used. A single regular equilateral mesh composed of 265675 cells was designed for the simulations. Although the mesh could have been adapted to the inverted frequency, we did not consider this strategy here, and the mesh was kept constant whatever the inverted frequency. Table 4.3 outlines the memory requirements and computational time of the major tasks performed by the parallel FWI algorithm. Of note, most of the memory and computational time were dedicated to the LU factorization and substitution phases performed during the multi-source forward problem. Computation of the gradient had a negligible computational cost, due to the domain-decomposition parallelism. The L-BFGS algorithm required a negligible extra amount of memory and computational time compared to a steepest-descent or PCG algorithm, suggesting that this optimization scheme can be efficiently used for realistic 2D and 3D FWI applications.

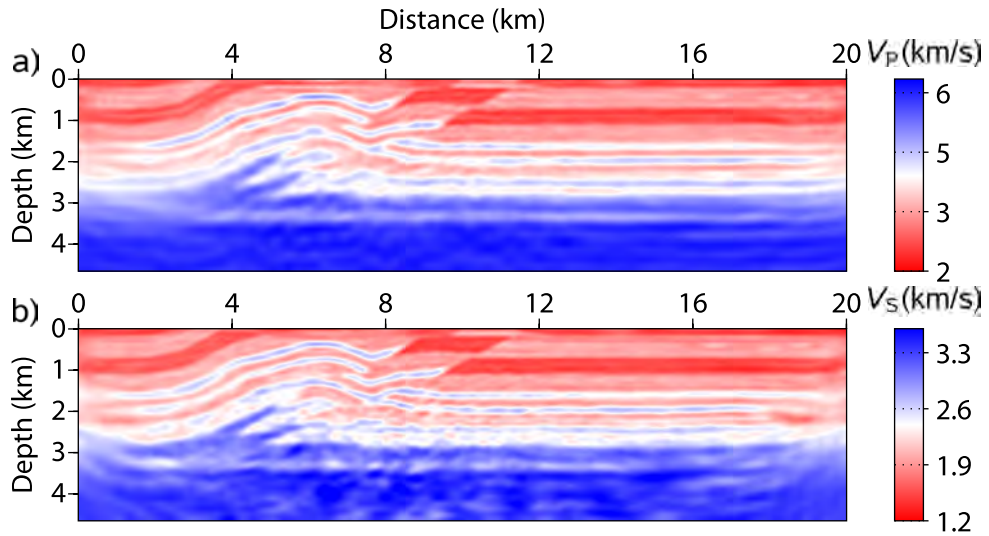


FIG. 4.13 – Simultaneous inversion - Final models obtained by successive inversion of two overlapping frequency groups composed of three frequencies each. (a) V_P model. (b) V_S model.

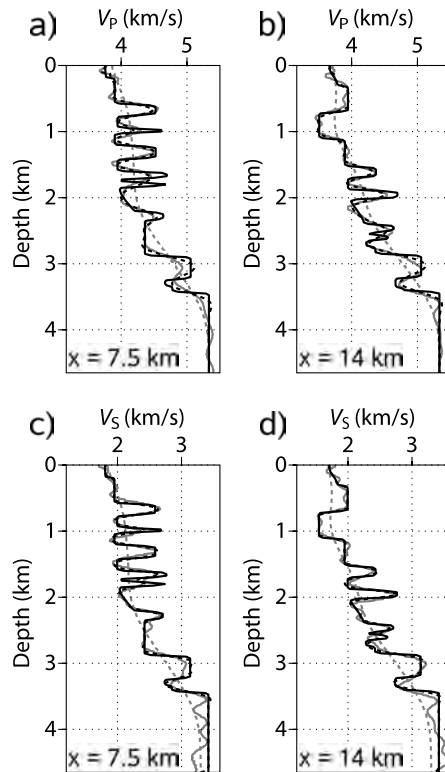


FIG. 4.14 – Same as Figure 4.7, but for the profiles extracted from the models recovered by the simultaneous approach (Figure 4.13).

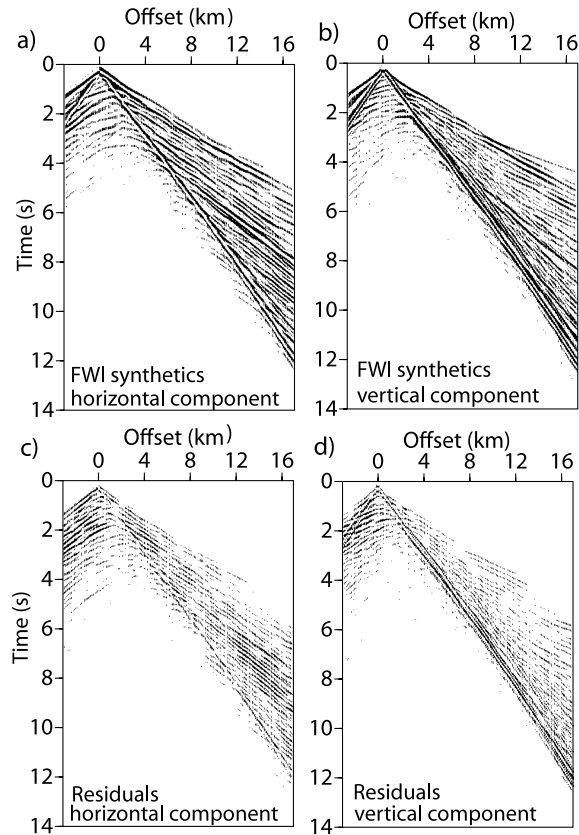


FIG. 4.15 – Same as Figure 4.8, but for seismograms computed in the FWI models recovered by successive inversion of 2 overlapping frequency groups (Figure 4.13).

TAB. 4.3 – Computational cost of the main tasks performed by FWI. Time estimations were averaged over several iterations for 24 processors.

Time for factorization (s)	21.3
MUMPS total memory for factorization (Gbytes)	4.8
Time for substitutions (199 shots) (s)	22.9
MUMPS total memory for substitution (Gbytes)	20.2
Time for gradient build up (s)	13.8
Time for L-BFGS perturbation computation (s)	1.2
Memory for L-BFGS(5) history (Mbytes)	21.2

4.1.7 Discussion

Application of elastic FWI to the Overthrust model has highlighted the strong non-linearity of the inversion resulting from free-surface effects. The impact of these effects on FWI can

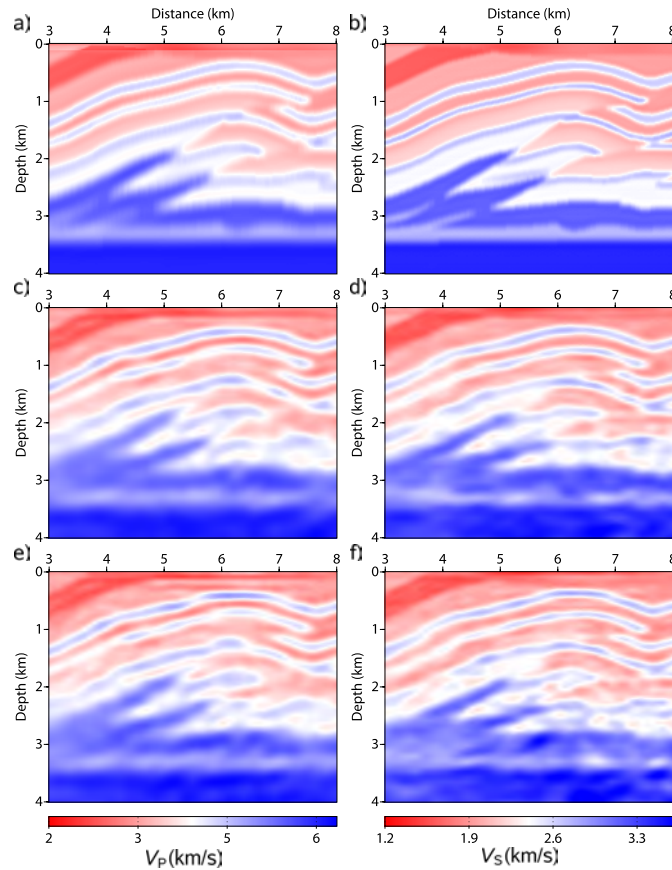


FIG. 4.16 – Comparison between FWI models obtained by successive inversion of two overlapping groups of frequencies (simultaneous approach) when three and five frequencies per group are used in the inversion, respectively. The frequency bandwidth is the same for each experiment but the frequency interval differs. a) Close-up of the true V_P model after low-pass filtering at the theoretical resolution of FWI. (c)-(e) Close-up of the FWI V_P model when five frequencies (c) and three frequencies (e) per group are used respectively. (b)-(d)-(f) Same as (a)-(c)-(e) for the V_S model.

be assessed by comparing the FWI results inferred from the data including or without the free-surface effects (compare Figures 4.6 and 4.9). As the best models were obtained when free-surface effects are not considered, this shows that for this case study, inversion of the surface waves was not useful towards an improvement of the reconstruction of the near-surface structure.

We interpret the failure of the raw-data inversion as the footprint of surface waves, the amplitudes of which dominate the wavefield and carry no information of the deep part of the model (Figure 4.5). Similar effects of surface waves on elastic FWI were also seen on a smaller scale by Gelis *et al.* (2007). Comparison between the sequential FWI results obtained with the raw data and the preconditioned data illustrates how the time damping helps to mitigate the non-linearities of FWI by injecting progressively more complex wave phenomena in the inversion (compare Figures 4.5 and 4.6).

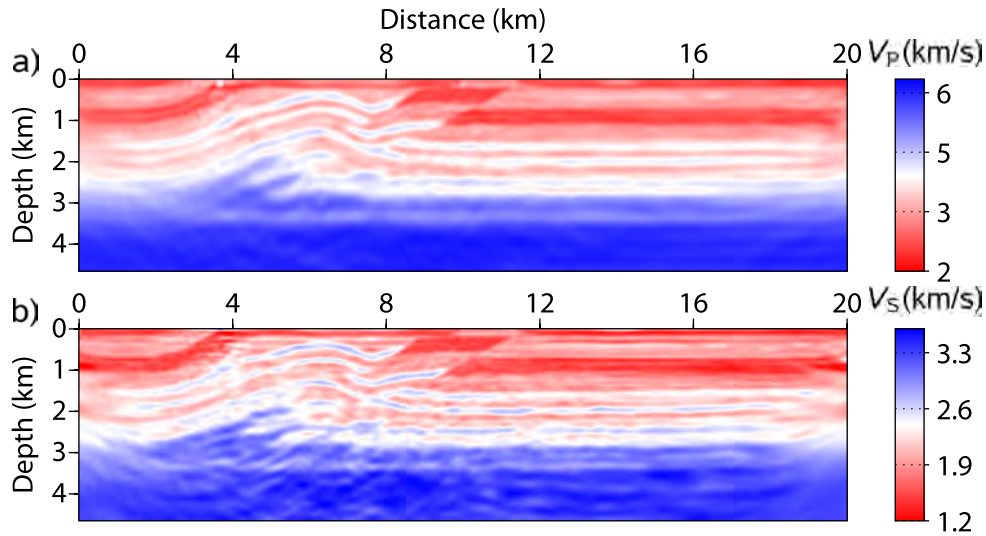


FIG. 4.17 – FWI velocity models obtained with the PCG algorithm. a) V_P parameter. b) V_S parameter. The sequential approach with five damping terms was used. The velocity models can be compared with those recovered by the L-BFGS algorithm (Figure 4.6).

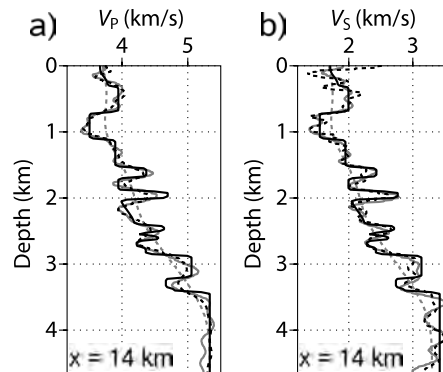


FIG. 4.18 – Comparison between velocity profiles extracted from the FWI models recovered by the L-BFGS (solid gray line) and the PCG algorithm (dashed black line). a) V_P models. b) V_S models. Starting and true models are depicted with dashed gray line and solid black line, respectively.

The further improvements obtained by simultaneous inversion of multiple frequencies combined with hierarchical inversions of damped data show that preserving some wavenumber redundancy in elastic FWI is critical to mitigate the non-linearity of the inversion associated with the propagation of surface waves in weathered near-surface layers and free-surface reflections. The more stable results obtained with the simultaneous approach compared to the Bunks approach, especially in the near-surface, suggest that several frequencies must be simultaneously inverted from the early stage of the inversion (compare Figures 4.10 and 4.13). Strengthening the wavenumber redundancy by decreasing the frequency interval in each frequency group further improved the imaging (Figure 4.16).

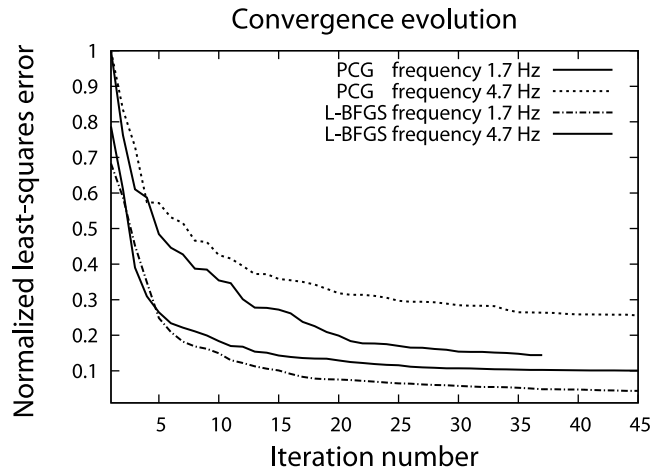


FIG. 4.19 – L-BFGS and PCG objective functions plotted as a function of iteration number for the inversion of the complex frequencies $(1.7+i1.5)$ Hz and $(4.7+i1.5)$ Hz. The curves associated with one frequency are normalized by the PCG objective functions at the first iteration. Other frequencies and damping factors show similar trends.

Another factor that increased the non-linearity of elastic FWI was the short S wavelengths that may require more accurate starting models or lower frequencies to converge towards an acceptable model. The maximum frequency of the starting frequency group must be chosen such that it prevents cycle-skipping artifacts that can result from the limited accuracy of the starting S-wave velocity model. Laplace-domain waveform inversion, which has been recently proposed as a reliable approach to build smooth initial elastic models of the subsurface (Pyun *et al.*, 2008b; Shin & Cha, 2008), may represent an approach to tackle the issue of building the starting model. An alternative approach is PP-PS stereotomography (Alerini *et al.*, 2002), including the joint inversion of refraction and reflection traveltimes of wide-aperture data.

As mentioned above, with this case study, we were not able to illustrate the usefulness of the surface waves for the reconstruction of the near-surface structure, since the most accurate FWI models were inferred without involving free-surface effects in the data (Figure 4.9). Instead, we have shown how to manage the non-linearities introduced by the surface waves, by means of judicious data preconditioning and FWI tuning. Alternatively, the surface waves in the recorded and modeled data can be filtered out or muted. We did not investigate this approach at this stage because efficient filtering of the modeled surface waves in the frequency domain is not straightforward. Note that the surface waves must be filtered out not only at the receiver positions, but also at each position in the computational domain where they have significant amplitudes, in order to remove their footprint from the gradient of the objective function. This investigation still requires further work.

More realistic applications of elastic FWI in more complex models still need to be investigated. Areas of complex topography, such as foot-hills, will lead to conversions from surface waves to body waves and *vice versa*, which may carry additional information on the near-surface. The robustness of elastic FWI for imaging models with heterogeneous Poisson ratios is a second field of investigation, especially in areas of soft seabed where the P-S-converted wavefield may have a limited signature in the data (Sears *et al.*, 2008). In the present study, we inverted data

computed with a constant density that was assumed to be known. For real data inversion, estimation of the density is required for a more reliable amplitude match. Reliable estimation of the density by FWI is difficult, because the P-wave velocity and density have similar radiation patterns at short apertures (Forgues & Lambaré, 1997). The benefit provided by wide apertures to uncouple these two parameters needs to be investigated. Other extensions of isotropic elastic FWI may relate to reconstruction of attenuation factors and some anisotropic parameters. Vertical transversely isotropic elastic FWI should be easily implemented from isotropic elastic FWI, since only the expression of the coefficients of the P-SV elastodynamic system need to be modified compared to the isotropic case (e.g., Carcione *et al.*, 1988).

Application of 2D elastic FWI to real data will require additional data preprocessing that was not addressed in this study, such as source estimation (Pratt, 1999) and 3D to 2D amplitude corrections (Bleistein, 1986; Williamson & Pratt, 1995). The sensitivity of the elastic FWI to the approximations and errors underlying this processing will need to be determined.

4.1.8 Conclusion

This study presents a new massively parallel 2D elastic frequency-domain FWI algorithm, with an application to a dip section of the SEG/EAGE onshore Overthrust model. Strong non-linearities of elastic FWI arise both from the presence of converted and surface waves, and from the limited accuracy of the V_S starting model. These two factors prevent convergence of FWI on the global minimum of the objective function if no specific preconditioning is applied to the data and no low starting frequency is available. Data preconditioning performed by time damping is necessary to converge towards acceptable velocity models, whatever the frequency sampling strategy is. Secondly, successive inversions of overlapping frequency groups out-perform successive inversions of single frequencies for the removal of instabilities in the near-surface of the FWI models. The bandwidth of the frequency groups must be chosen such that cycle-skipping artifacts are avoided, while injecting a maximum amount of redundant information into the frequency groups. The quasi-Newton algorithm of L-BFGS outperforms the most popular preconditioned conjugate-gradient algorithm in terms of convergence rate and convergence level, without significant extra computational costs, and hence is shown to be very useful for this application.

Acknowledgments

This study was funded by the SEISCOPE consortium <http://seiscope.unice.fr>, sponsored by BP, CGG-VERITAS, EXXON-MOBIL, SHELL and TOTAL, and by Agence Nationale de la Recherche (ANR) under project ANR-05-NT05-2-42427. The linear systems were solved with MUMPS package, available on <http://graal.ens-lyon.fr/MUMPS/index.html>. The mesh generation was performed with help of TRIANGLE, available on <http://www.cs.cmu.edu/~quake/triangle.html>. Access to the high-performance computing facilities of the SIGAMM (Observatoire de la Côte d'Azur) and IDRIS computer center (project 082280) provided the required computer resources, and we gratefully acknowledge both of these facilities and the support of their staff. Many thanks go to Serge Gratton (CERFACS and CNES, Toulouse, FRANCE) for interesting and stimulating discussions on quasi-Newton optimization methods. We would like to thank Isabelle Lecomte, Christophe Barnes, Hansruedi Maurer and Changsoo Shin for their fruitful comments on the manuscript.

4.1.9 Test complémentaire : combien de composantes sont nécessaires ?

Les tests présentés précédemment considèrent une inversion multiparamètres de données de géophone multicomposantes. Pour l'imagerie des paramètres élastiques par inversion des formes d'ondes avec un problème direct calculé par une approche asymptotique, Jin *et al.* (1992) concluaient qu'en milieu marin, pour des acquisitions à courts offsets, deux composantes sont nécessaires pour l'imagerie multiparamètres de I_P et I_S . Cette conclusion est basée sur l'hypothèse que seules les phases P-P sont enregistrées à court offset dans une acquisition marine monocomposante. L'inversion d'une seule composante enregistrant uniquement des conversions P-P ne permet donc pas de contraindre l'inversion de deux paramètres. Le nombre de composantes est un critère primordial dans la conception d'une acquisition sismique, aussi bien d'un point de vue économique que logistique.

A partir du cas terrestre précédemment utilisé, j'essaie de donner quelques éléments de réponse au nombre de composantes nécessaires dans le contexte d'acquisition terrestre. Je reprends la configuration optimale utilisée dans les tests multicomposantes précédents, c'est à dire basée sur l'inversion simultanée de groupes de fréquences avec recouvrement entre groupes, et en utilisant des données préconditionnées. Les tests sont conduits en utilisant uniquement la composante verticale, généralement utilisée dans les acquisitions monocomposante.

Un premier test reprend le cas de 5 fréquences inversées en deux groupes de trois fréquences : (1.7, 2.0, 3.5) et (3.5, 4.8, 7.2) Hz. La Figure 4.20 représente les modèles V_P et V_S imagés. Les grandes structures du milieu sont bien localisées dans les deux modèles à partir de la composante verticale uniquement. Le modèle V_P ne montre pas d'artefact majeur mais souffre d'un déficit d'amplitude dans la reconstruction, comparé au test à deux composantes (*cf.* Figures 4.13 et 4.14). Le modèle V_S , plus difficile à imager à cause de sa forte sensibilité, montre plus d'artefacts, en particulier dans la zone de chevauchement et dans la partie profonde avec là encore un déficit d'amplitude (Figure 4.21).

Un deuxième test vise à inverser plus de fréquences simultanément afin de compenser le déficit de données de la composante horizontale par une redondance de la composante verticale. La Figure 4.22 montre un test d'inversion de 17 fréquences réparties en deux groupes : (1.7, 1.8, 2.0, 2.3, 2.5, 2.7, 3.0, 3.2, 3.5) et (3.5, 3.8, 4.1, 4.4, 4.8, 5.2, 6.0, 6.5, 7.2) Hz. Ces deux groupes reprennent les mêmes bornes inférieures et supérieures que le cas précédant afin de pouvoir comparer l'influence de l'ajout de fréquences intermédiaires, sans modifier les bornes du spectre de nombres d'onde imagés. Les résultats montrent d'une part que l'ajout d'information redondante par l'inversion de fréquences multiples permet de mieux contraindre l'inversion, sans artefacts majeurs dans les deux modèles. Cependant, les amplitudes des vitesses restent encore sous-estimées par rapport au test à deux composantes, malgré la redondance apportée par le nombre conséquent de fréquences prises en compte (Figure 4.23).

Ces tests illustrent, pour un cas terrestre dans lequel les deux paramètres de vitesse V_P et V_S ont de fortes signatures dans les données des deux géophones, que l'inversion de la données verticale seule permet de reconstruire le milieu pour les deux paramètres de vitesse. La configuration d'acquisition utilisée et la conclusion sont différentes de celles de Jin *et al.* (1992). Notons que l'information de la composante horizontale manquante dans l'inversion nécessite de prendre une densité de fréquences inversées plus importante afin de minimiser les artefacts de reconstruction, au détriment du coût de calcul. De plus, les amplitudes des structures imagées restent sous-évaluées par rapport aux résultats précédents, et montrent la complémentarité des informations contenues dans les deux composantes.

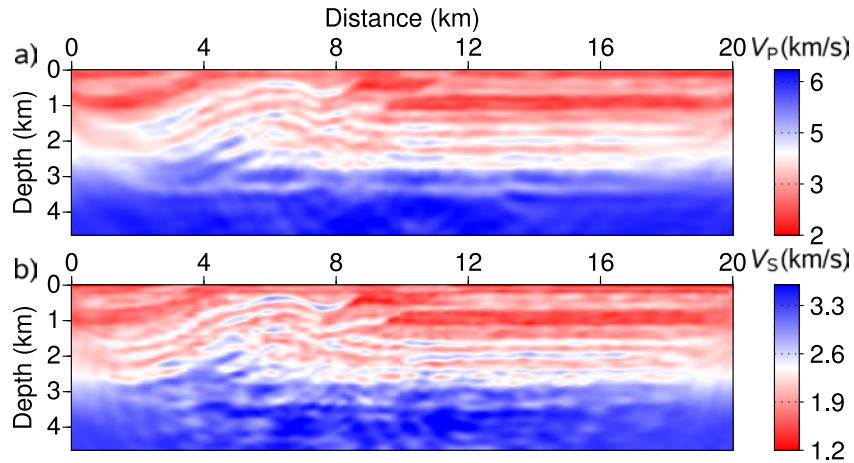


FIG. 4.20 – Modèles Overthrust a) V_P et b) V_S , reconstruits après inversion de données pré-conditionnées pour groupes de 3 fréquences. Seule la composante verticale est utilisée.

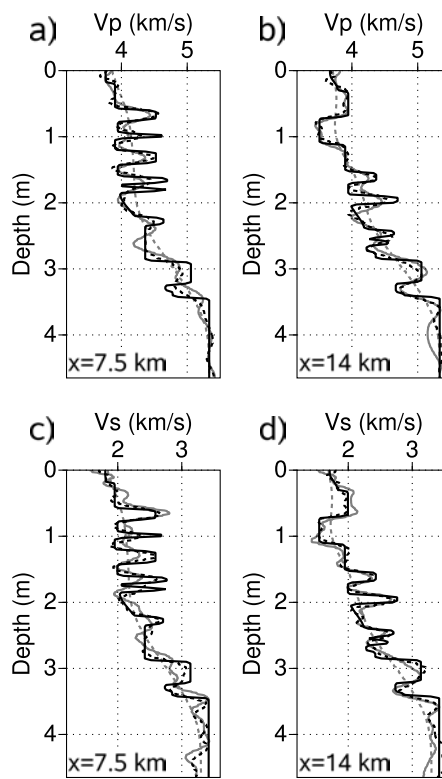


FIG. 4.21 – Coupes verticales des modèles Overthrust imagés par inversion de 2 groupes de 3 fréquences à partir de la composante verticale du géophone uniquement. (a)-(b) modèle V_P et (c)-(d) modèle V_S aux distances 7.5 et 14 km. Les modèles de départ et les vrais modèles sont représentés respectivement en lignes grises pointillées et noires continues. Le modèle issu de l'inversion de la donnée de géophone vertical est en gris continu tandis que la référence à 2 géophones et 5 fréquences (2 groupes) est en noire pointillée.

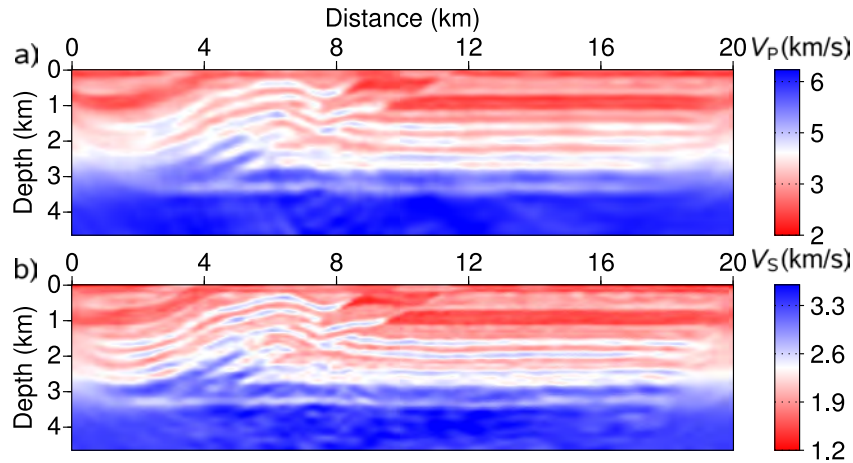


FIG. 4.22 – Modèles Overthrust a) V_P et b) V_S , reconstruits après inversion de données pré-conditionnées pour groupes de 9 fréquences. Seule la composante verticale est utilisée.

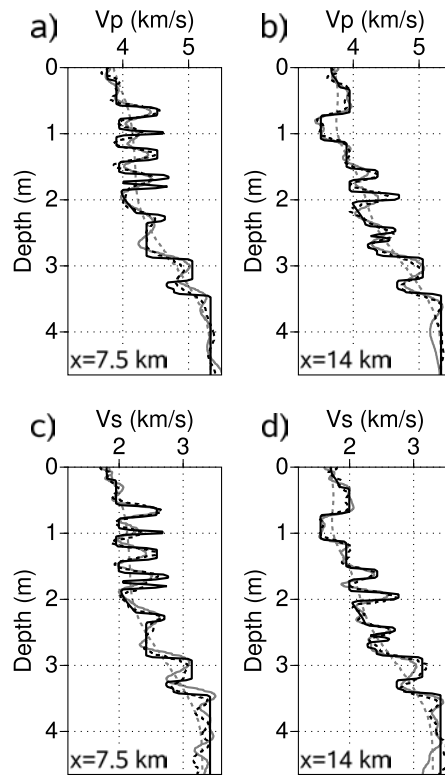


FIG. 4.23 – Coupes verticales des modèles Overthrust imagés par inversion de 2 groupes de 9 fréquence à partir de la composante verticale du géophone uniquement. (a)-(b) modèle V_P et (c)-(d) modèle V_S aux distances 7.5 et 14 km. Les modèles de départ et les vrais modèles sont représentés respectivement en lignes grises pointillées et noires continues. Le modèle issu de l'inversion de la donnée de géophone vertical est en gris continu tandis que la référence à 2 géophones et 5 fréquences (2 groupes) est en noire pointillée.

4.2 Application marine : Le modèle Valhall

Cette application a été partiellement présentée dans « Two-dimensional seismic imaging of the Valhall model from synthetic OBC data by frequency-domain elastic full-waveform inversion », 2009, R. Brossier, S. Operto and J. Virieux, SEG Technical Program Expanded Abstracts (*cf.* annexe B.1).

4.2.1 Présentation du modèle élastique de Valhall

Le modèle synthétique de Valhall (Figure 4.24) est représentatif des champs de gaz et pétrole, en eau peu profonde de Mer du Nord (Munns, 1985). Le fond de mer est plat à une profondeur de seulement 70 m. Les cibles principales sont : le nuage gaz dans l'épaisse couche de sédiments et plus profondément le pétrole piégé sous le toit rocheux de l'anticlinal formé de craie. Le gaz est facilement identifiable dans le modèle de vitesse V_P par ses anomalies de vitesses lentes tandis qu'il influe peu le modèle V_S , donnant de fortes anomalies du coefficient de Poisson.

L'acquisition choisie pour les tests représente l'acquisition permanente réellement mise en place dans les champs : des câbles en fond de mer (*Ocean Bottom Cable* (OBC)) sont équipés de capteurs 4-composantes (1 hydrophone, 1 géophone vertical et 2 géophones horizontaux) espacés de 50 m (Kommedal *et al.*, 2004). 315 sources explosives sont disposées le long du profil, 5 m sous la surface de l'eau. Des données synthétiques en domaine temporel sont calculées pour une source localisée à la distance 3 km pour ce modèle élastique (Figure 4.25), ainsi que sous l'approximation acoustique (Figure 4.26). Ces données sont calculées par un algorithme de différences finies avec un schéma d'intégration temporelle explicite. Les données montrent clairement les ondes guidées dans la couche d'eau pour la simulation élastique (associée à de la dispersion numérique), donnant peu d'informations sur le milieu. Cependant, la comparaison entre simulation acoustique et élastique montre la très faible signature du paramètre V_S dans ce modèle peu contrasté, avec des vitesses d'ondes S très faibles dans la couche sédimentaire superficielle générant peu de conversions P-S à l'interface eau/solide. Ce type de configuration donne une inversion fortement non-linéaire et mal contrainte pour la reconstruction du paramètre V_S faiblement représenté dans les données (Sears *et al.*, 2008).

4.2.2 Configuration de l'inversion

Pour l'inversion des formes d'ondes, des données de référence sont calculées dans le vrai modèle synthétique (Figure 4.24) par la méthode GD en domaine fréquentiel avec une interpolation mixte P_0 - P_1 . Une couche de cellules non-structurée de 160 m est calculée par l'interpolation P_1 afin de modéliser le contact eau/solide (Figure 3.4). Un maillage régulier équilatéral, calculé avec une interpolation P_0 prolonge cette couche en profondeur.

Pour le processus d'inversion, 5 fréquences sont considérées successivement (2, 3, 4, 5 et 6 Hz) et 3 coefficients d'amortissement (partie imaginaire de la fréquence) sont appliqués pour chacune des fréquences ($\gamma = 2, 0.33, 0.1$). Le modèle de départ utilisé (Figure 4.27) est une version lissée du vrai modèle avec une longueur de corrélation horizontale fixe de 500 m et une longueur de corrélation verticale augmentant linéairement avec la profondeur de 25 m à 1000 m. Ce modèle de départ pourrait représenter idéalement un modèle obtenu en combinant

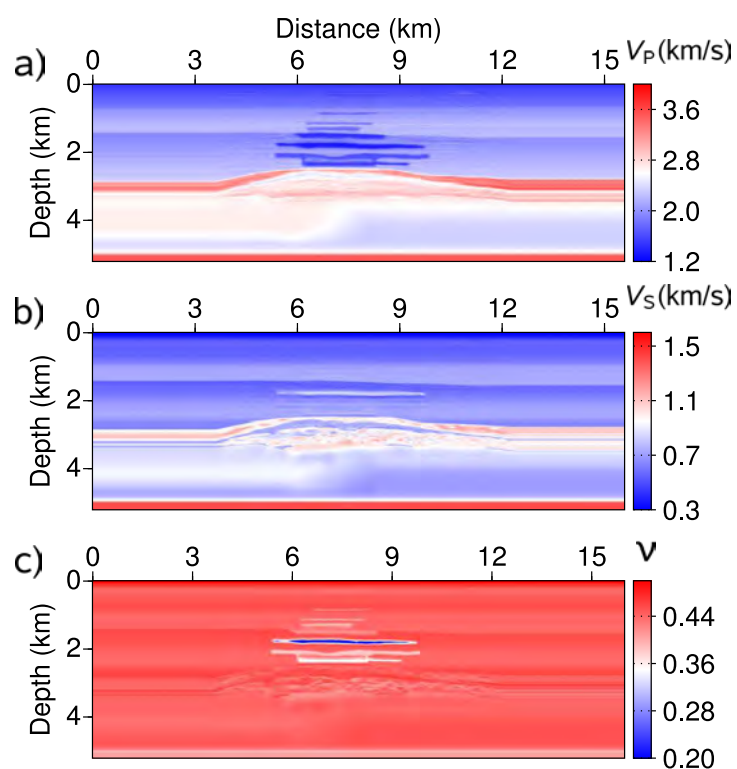


FIG. 4.24 – Le modèle synthétique de Valhall pour les vitesses d'ondes (a) P et (b) S. (c) le modèle de coefficient de Poisson associé.

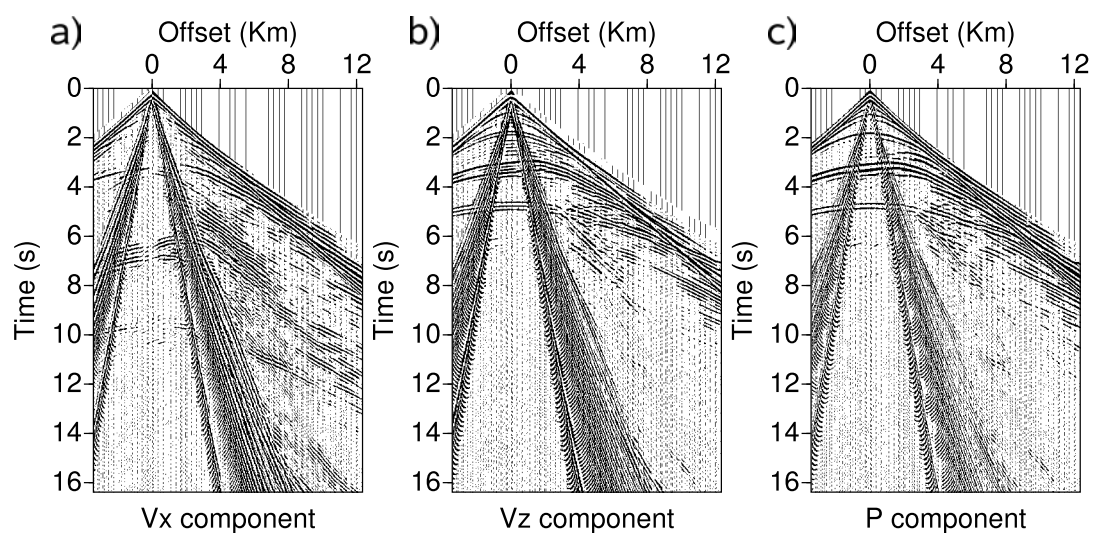


FIG. 4.25 – Exemple de collection en tir commun pour une source localisée à la distance 3 km dans le modèle élastique de Valhall, pour les composantes (a) horizontales et (b) verticales des vitesses de déplacement, et (c) la pression hydrostatique.

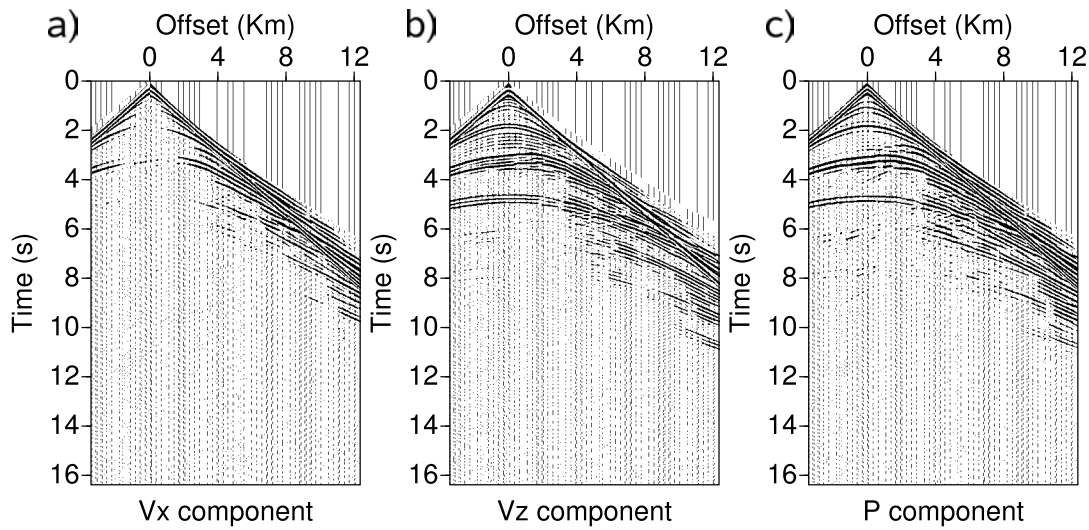


FIG. 4.26 – Exemple de collection en tir commun pour une source localisée à la distance 3 km dans le modèle acoustique de Valhall, pour les composantes (a) horizontales et (b) verticales des vitesses de déplacement, et (c) la pression hydrostatique.

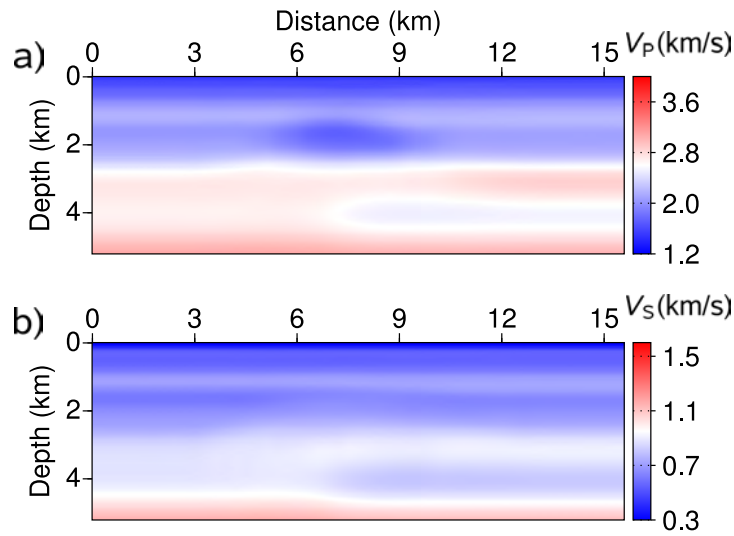


FIG. 4.27 – Imagerie du modèle de Valhall : modèles de départ (a) V_P et (b) V_S .

tomographie des temps d'arrivées des ondes réfractées et réfléchies (Prioux *et al.*, 2009). Pour tous les tests, la signature de la source, considérée inconnue, est estimée par inversion linéaire (*cf.* partie 2.1.4).

4.2.3 Inversion acoustique de données élastique

Comme nous l'avons remarqué dans l'observation des données, les arrivées d'ondes P dominent les champs d'ondes, particulièrement sur les composantes de l'hydrophone et du géo-

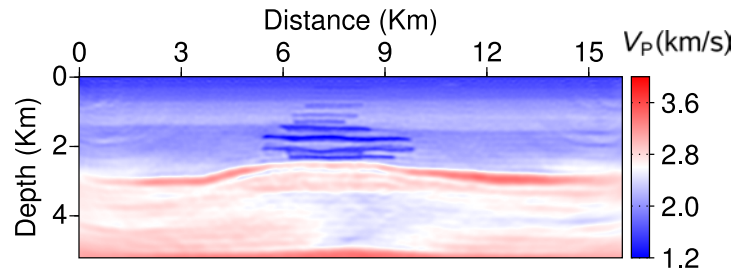


FIG. 4.28 – Imagerie du modèle de Valhall : modèle V_P imagé avec l’approximation acoustique en utilisant la composante hydrophone du champ élastique.

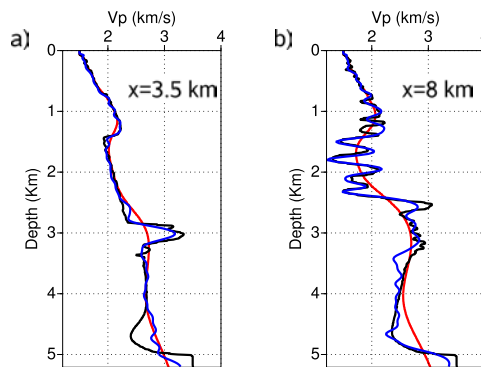


FIG. 4.29 – Imagerie du modèle de Valhall : profils verticaux du modèle V_P imagé avec l’approximation acoustique en utilisant la composante hydrophone du champ élastique. Les courbes noires, rouges et bleues représentent respectivement les modèles vrais, de départ et imagés.

phone verticale, rendant ce cas d’étude favorable à l’approximation acoustique pour l’inversion (Barnes & Charara, 2008).

Un premier test d’imagerie avec l’approximation acoustique est mené avec l’algorithme FWT2D (Sourbier *et al.*, 2009a) basé sur un moteur de problème direct en différences finies et une optimisation du problème inverse non-linéaire par une méthode de gradient préconditionné par les termes diagonaux du Hessien approximé (Operto *et al.*, 2006). La composante hydrophone des données élastiques est inversée sans aucun prétraitement. Quinze itérations sont effectuées par facteur de régularisation conduisant à 45 itérations par fréquence. La Figure 4.28 montre le modèle V_P reconstruit par l’inversion acoustique qui n’introduit pas d’artefact majeur dans le modèle. Les profils verticaux (Figure 4.29) montrent la bonne qualité de la reconstruction, malgré les basses fréquences inversées.

Ce test confirme que l’approximation acoustique est envisageable dans certains cas d’environnements marins où les données sont clairement dominées par le paramètre V_P . Notons que, dans un cas comme celui-ci, le modèle obtenu par inversion acoustique pourrait servir de modèle de départ pour l’inversion de paramètres secondaires tels que V_S , ou les paramètres de densité et d’atténuation.

4.2.4 Inversion élastique hiérarchique

L'inversion élastique est effectuée avec un problème direct basé sur la méthode GD. L'algorithme L-BFGS est mis en oeuvre pour l'optimisation et la diagonale du pseudo-Hessien (Shin *et al.*, 2001) est utilisée comme estimation initiale du Hessien. Dix itérations sont effectuées par coefficient de régularisation, conduisant à 30 itérations par fréquence.

Sears *et al.* (2008) ont montré, avec une inversion des formes d'ondes en domaine temporel, que la reconstruction du modèle V_S est particulièrement non-linéaire lorsque les modèles de vitesses ont de faibles contrastes, ne générant pas d'ondes converties P-S suffisamment énergétiques pour garantir des contraintes fortes pendant l'inversion. Afin de minimiser l'impact de cette non-linéarité, nous décomposons l'inversion élastique en deux phases hiérarchiques :

1. Dans une première étape, seul le modèle V_P est imagé à partir de la composante hydrophone. Le problème direct est modélisé avec l'approximation élastique mais le modèle V_S n'est ni inversé, ni mis à jour. Le but de cette première étape est de minimiser le poids des résidus d'ondes P, les plus énergétiques dans les données. Afin de minimiser le coût de calcul, un maillage grossier adapté au modèle V_P est utilisé, générant une dispersion des ondes S, mais sans affecter significativement l'imagerie.
2. Dans une deuxième étape, les paramètres V_P et V_S sont imagés simultanément à partir des composantes horizontales et verticales des géophones. Afin d'exploiter efficacement les arrivées P-S à moyens et longs offsets, un gain d'amplitude croissant en fonction de l'offset est appliqué aux données par l'intermédiaire de la matrice \mathbf{S}_d .

Les tests ont montré que, si les deux paramètres sont reconstruits depuis le début de l'inversion sans cette stratégie hiérarchique en deux étapes, le processus converge vers un minimum local à cause de l'inversion de V_S très mal contrainte et instable lorsque le modèle V_P n'est pas assez précis.

Les Figures 4.30 et 4.31 montrent les modèles reconstruits avec l'approche hiérarchique et les profils correspondants. L'inversion permet clairement de reconstruire les structures géologiques principales comme l'anticlinal et le nuage de gaz dans les sédiments. Le modèle V_P reconstruit est proche de celui obtenu par inversion acoustique (Figure 4.28), tout en étant de meilleure qualité dans la partie profonde (supérieur à 2.5 km). Le modèle V_S souffre d'un déficit des bas nombres d'ondes, en particulier dans la zone profonde où le modèle de départ est moins précis. Malgré les basses fréquences utilisées (2 Hz), ce déficit est expliqué par les vitesses de propagation d'ondes S très faibles, limitant la capacité de l'inversion à résoudre les grandes longueurs d'ondes (*cf.* expression (2.17) sur le pouvoir de résolution spatiale du gradient). Ce cas illustre la problématique associée au modèle de départ de V_S qui doit contenir un spectre de nombres de d'onde suffisamment large, en particulier lorsque les vitesses de propagation sont faibles. Cependant, la partie superficielle permet d'identifier le nuage de gaz grâce à la combinaison des paramètres élastiques.

Notons que, dans ce cas où les différentes classes de paramètres ont des empreintes contrastées dans les données, l'utilisation de la composante horizontale est indispensable pour la reconstruction de V_S . Les tests menés avec la même configuration d'inversion mais sans la composante horizontale ont révélé une grande difficulté à imaginer le paramètre V_S (Figure 4.32 et 4.33) qui est très peu représenté sur la composante géophone vertical, malgré les grands offsets utilisés. Des tests d'inversion de groupes de fréquences denses n'ont pas permis de compenser le déficit d'information dans les données sensibles aux paramètres V_S . Malgré le dispositif grand-angle

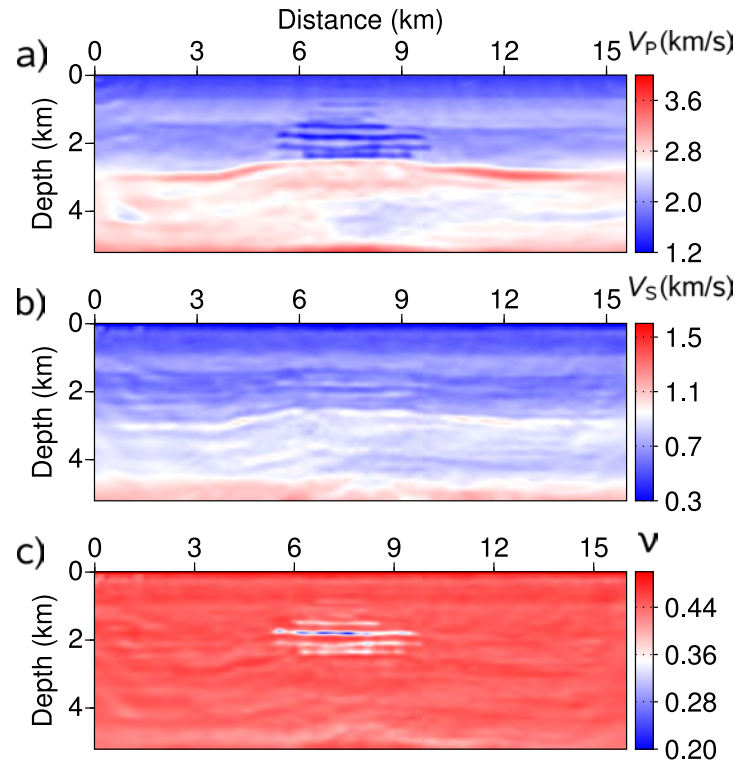


FIG. 4.30 – Imagerie du modèle de Valhall : modèles (a) V_P et (b) V_S de l'inversion élastique hiérarchique. (c) le modèle associé du coefficient de Poisson.

utilisé dans cette étude, cette conclusion est en accord avec les travaux de Jin *et al.* (1992), dans lesquels une inversion fondée sur l'utilisation exclusive des événements P-P ne permet pas de reconstruire I_S .

4.3 Conclusion

Ce chapitre d'application a permis de mettre en évidence les difficultés liées à l'inversion multiparamètres de données élastiques synthétiques sans bruit.

D'une part, lorsque les données portent une forte signature des deux paramètres V_P et V_S , comme c'est le cas pour des applications terrestres possédant des structures géologiques aux vitesses de propagation très contrastées, l'inversion simultanée des deux paramètres est requise dès le début de l'inversion. Le préconditionnement des données implémenté avec les fréquences complexes permet de régulariser efficacement le problème et ainsi de converger vers le minimum global de la fonction coût. Lorsque les données possèdent un niveau de complexité élevé (ondes de surfaces, multiples), l'inversion simultanée de plusieurs fréquences permet d'apporter une redondance essentielle à la convergence du système d'optimisation. A coût numérique équivalent, il est donc essentiel de trouver le meilleur compromis entre quantité d'information injectée pour une itération à travers les fréquences multiples, tout en évitant le phénomène de saut de phase dû à l'utilisation de fréquences trop élevées.

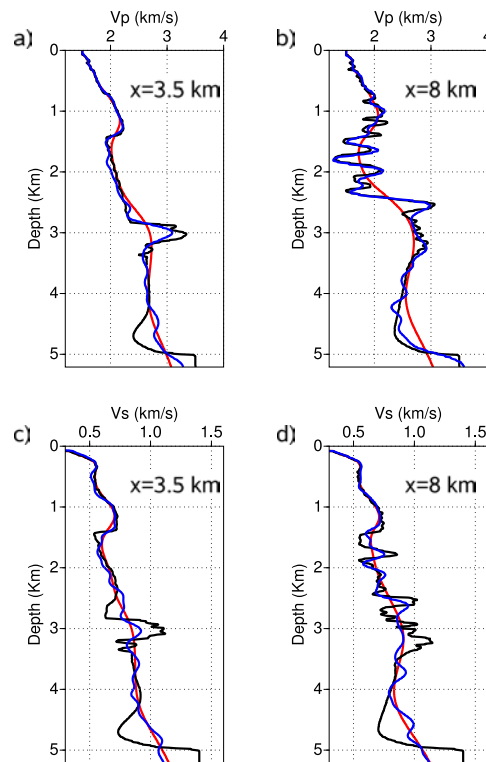


FIG. 4.31 – Imagerie du modèle de Valhall : profils verticaux des modèles V_P et V_S reconstruits par inversion élastique hiérarchique. Les courbes noire, rouges et bleues représentent respectivement les modèles vrais, de départ et imagés.

D'autre part, lorsque la signature des paramètres dans les données est inégale, l'inversion hiérarchique des différentes classes de paramètres et de différents types d'arrivées (ondes de compression, ondes converties P-S) se révèle nécessaire. Dans un premier temps, l'inversion du paramètre dominant permet de réduire les résidus associés aux ondes dominantes. Dans un second temps, l'inversion peut utiliser des phases moins significatives en terme d'énergie, afin de reconstruire les paramètres « secondaires ».

Enfin le choix du type de données à inverser se révèle être dépendant des données. Lorsque les signatures des deux modèles de vitesses sont fortement présentes dans les données des deux géophones, l'utilisation de la composante verticale seule peut permettre d'imager les deux paramètres en suréchantillonnant les fréquences inversées par rapport au cas deux composantes. Cependant, si la signature du paramètre V_S est moins marquée, et se manifeste principalement sur la composante horizontale, comme c'est le cas pour le modèle marin Valhall, l'utilisation des deux composantes est requise pour l'inversion des deux paramètres, comme le concluaient Jin *et al.* (1992) à partir d'approches asymptotiques.

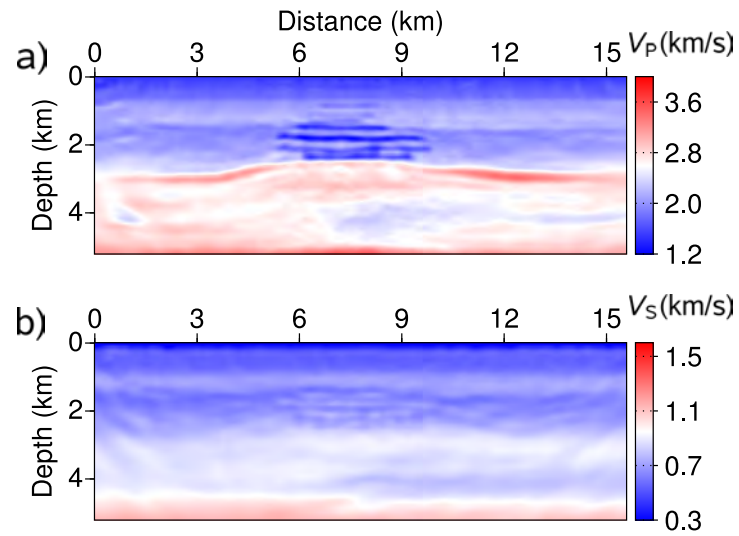


FIG. 4.32 – Imagerie du modèle de Valhall : modèles (a) V_P et (b) V_S de l'inversion élastique hiérarchique sans la composante de géophone horizontal portant la signature de V_S . Notons la bonne reconstruction de V_P mais V_S reste très peu modifié car peu représenté dans les données.

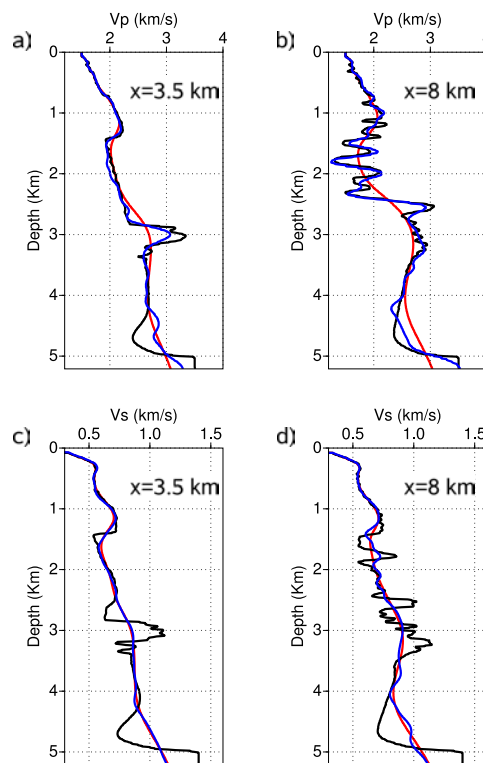


FIG. 4.33 – Imagerie du modèle de Valhall : profils verticaux des modèles V_P et V_S reconstruits par inversion élastique hiérarchique à partir de la composante verticale uniquement. Les courbes noires, rouges et bleues représentent respectivement les modèles vrais, de départ et imagés.

Chapitre 5

Sensibilité de l'inversion au bruit

Sommaire

5.1	Quelle norme pour une inversion de formes d'onde robuste? . . .	182
5.1.1	Résumé	182
5.1.2	Introduction	182
5.1.3	Théorie et algorithme	184
5.1.4	Tests numériques : le modèle marin de Valhall	189
5.1.5	Tests numériques : le modèle terrestre SEG/EAGE Overthrust	194
5.1.6	Discussion	196
5.1.7	Conclusions	198
5.2	Application acoustique sur les données réelles de Valhall	200
5.2.1	Environnement, données et configuration de l'inversion	200
5.2.2	Résultats	203
5.2.3	Conclusion	205
5.3	Conclusion	207

Dans ce chapitre, nous posons le problème de la sensibilité et de la robustesse de l'inversion des formes d'ondes en domaine fréquentiel au bruit dans les données. En effet, comme nous l'avons vu auparavant aux chapitres 2 et 4, l'inversion en domaine fréquentiel exploite la redondance dans le domaine des nombres d'ondes imagés pour n'inverser que quelques fréquences discrètes. Ces fréquences ne représentent qu'une faible quantité d'information, comparativement au volume de données disponibles. Le choix de n'inverser qu'une quantité limitée de données est directement motivé par le coût de calcul de l'approche qui doit être minimal. Cependant, le principe d'utiliser des données décimées est opposé aux méthodes généralement utilisées en imagerie sismique dans lesquelles la qualité des images est apportée par la sommation de données fortement redondantes. Nous nous intéressons donc, dans un premier temps, à la robustesse de l'inversion des formes d'ondes en domaine fréquentiel de données synthétiques décimées et bruitées issues de milieux réalistes. La méthode d'imagerie « classique » basée sur les moindres carrés est appliquée ainsi que des formalismes alternatifs plus robustes, afin d'évaluer les potentialités et limites des différentes approches dans le contexte contrôlé des tests synthétiques. Dans un deuxième temps, une application sur un jeu de données réelles pétrolières de la zone de Valhall en Mer du Nord, est présentée dans le cas de l'approximation acoustique de la propagation et de l'inversion.

5.1 Quelle norme pour une inversion de formes d'onde robuste ?

Cette partie présente des applications synthétiques d'inversion de formes d'ondes en domaine fréquentiel de données bruitées. La sensibilité et la robustesse aux bruits de différentes normes et critères de minimisation sont testés, afin de montrer les limites de l'approche classique basée sur la minimisation de la norme aux moindres carrés L_2 , et de montrer que des fonctionnelles alternatives plus robustes peuvent être utilisées. Cette partie est présentée sous la forme d'un article soumis à publication dans la revue *Geophysics*. Ce sujet a également fait l'objet d'une note publiée dans la revue *Geophysical Research Letters* et présentée dans l'annexe B.2.

Which data residual norm for robust elastic frequency-domain Full Waveform Inversion ?

R. Brossier, S. Operto and J. Virieux

article soumis pour publication dans la revue *Geophysics*

5.1.1 Abstract

Elastic full-waveform inversion is an ill-posed data-fitting procedure that is sensitive to noise, inaccuracies of the starting model, the definition of multi-parameter classes, and inaccurate modeling of wavefield amplitudes. We have here investigated the performances of different minimization functionals, as the least-square norm (L_2), the least-absolute-values norm (L_1), and some combinations of both (the Huber and the so-called Hybrid criteria), with reference to two noisy offshore (Valhall model) and onshore (Overthrust model) synthetic datasets. The four minimization functionals are implemented in two-dimensional, elastic, frequency-domain full-waveform inversion, where efficient multiscale strategies are designed by successive inversions of a few increasing frequencies. For both the offshore and the onshore case studies, the L_1 norm provides the most reliable V_P and V_S models, even when strongly decimated datasets that correspond to few frequencies are used in the inversion, and when outliers pollute the data. The L_2 norm can provide reliable results in the presence of uniform white noise for both V_P and V_S if the data redundancy is increased by refining the frequency sampling interval in the inversion, at the expense of the computational efficiency. Unlike the L_2 norm, the L_1 norm, the Huber and the Hybrid criteria allow for successful imaging of V_S model from noisy data in soft-seabed environment, where the P-to-S waves have a small footprint in the data. The Huber and the Hybrid criteria are, however, shown to be sensitive to a threshold criterion, which controls the transition between the two criteria and which requires tedious trial-and-error investigations for reliable estimation. The L_1 norm provides a robust alternative to the L_2 norm for the inversion of decimated datasets in the framework of efficient frequency-domain full-waveform inversion.

5.1.2 Introduction

Quantitative imaging of the Earth subsurface is essential for reservoir characterization, for the monitoring of CO₂ sequestration, and for civil engineering applications. Full-waveform inversion (FWI) is a data-fitting procedure used to derive high-resolution quantitative models

of the subsurface, through the exploitation of the full information content of the data (Tarantola, 1984). When applied in the frequency domain, computationally efficient FWI algorithms (known as efficient frequency-domain FWI in the following) can be designed by sacrificing the redundancy of the wave-number coverage provided by wide-aperture acquisition surveys (Pratt & Worthington, 1990; Pratt, 1999). The decimation of the wave-number redundancy can be implemented by limiting the inversion to a few judiciously chosen discrete frequencies at the expense of the signal-to-noise ratio of the reconstructed models (Sirgue & Pratt, 2004; Brenders & Pratt, 2007a).

FWI potentially provides high-resolution models of the subsurface, but it suffers from two main difficulties. The first is related to the computational cost of the forward problem ; namely, the numerical resolution of the two-way wave equation in heterogeneous media for multiple sources. In the frequency domain, computationally efficient approaches based on direct solvers have been developed for two-dimensional (2D) acoustic and elastic wave propagation (Jo *et al.*, 1996; Hustedt *et al.*, 2004; Stekl & Pratt, 1998; Brossier *et al.*, 2009a). For the 3D problems, the computational burden of the forward problem has motivated many efforts to develop efficient 3D modeling engines for frequency-domain FWI. These have been based on direct solvers, iterative solvers, hybrid direct/iterative solvers, and time-domain approaches (Nihei & Li, 2007; Operto *et al.*, 2007; Plessix, 2007; Warner *et al.*, 2007; Sirgue *et al.*, 2008; Sourbier *et al.*, 2008a; Etienne *et al.*, 2009). The second difficulty is related to the ill-posedness and the non-linearity of the inverse problem, which is generally formulated as a least-squares local optimization, so as to manage the numerical cost of the forward problem (Tarantola & Valette, 1982). The ill-posedness of FWI mainly arises from the lack of low frequencies in the source bandwidth and the incomplete illumination of the subsurface provided by conventional seismic surveys. Consequently, the problem is highly non-linear and the results strongly depend on the accuracy of the starting model in the framework of local optimization and on the presence of noise. Several hierarchical multiscale strategies that proceed from low frequencies to higher frequencies have been proposed to mitigate the non-linearity of the inverse problem (Pratt & Worthington, 1990; Bunks *et al.*, 1995; Sirgue & Pratt, 2004; Brossier *et al.*, 2009a).

The noise footprint in seismic imaging is conventionally mitigated by stacking highly redundant multifold data. However, improving our understanding of the inversion sensitivity to noise is a key issue, and in particular when the data redundancy is decimated in the framework of efficient frequency-domain FWI. The least-squares objective function remains the most commonly used criterion in FWI, although it theoretically suffers from poor robustness in the presence of large isolated and non-Gaussian errors. Other norms can therefore be considered. The least-absolute-values norm (L_1) is not based on Gaussian statistics in the data space, and it was introduced into time-domain FWI by Tarantola (1987); Crase *et al.* (1990); it has been shown to be weakly sensitive to noise. Djikpéssé & Tarantola (1999) used the L_1 norm successfully to invert field data from the Gulf of Mexico with time-domain FWI. Surprisingly, this norm has been marginally used during recent applications of FWI. Pyun *et al.* (2009) used a L_1 -like norm for frequency-domain FWI : the L_1 norm is applied independently to the real and imaginary parts of the complex-valued wavefield. The resulting functional does not rigorously define a norm from a mathematical viewpoint, because the functional does not satisfy the scalar multiplication property of norms ($\|\alpha\mathbf{x}\| = |\alpha| \cdot \|\mathbf{x}\|$ for complex-valued scalar α and vector \mathbf{x}). The violation of the norm property makes the value of the functional to vary with the phase of the residuals, when the residual amplitude is kept constant. Although this mathematical approximation, quite robust results were obtained. Alternative functionals, such

as the Huber criterion (Huber, 1973; Guitton & Symes, 2003) and the Hybrid L_1/L_2 criterion (Bube & Langan, 1997) can also be considered. Ha *et al.* (2009) applied the Huber criterion for frequency-domain FWI and illustrated its robust behavior compared to the L_2 norm when considering a dense frequency sampling in inversion. All of these criteria behave as the L_2 norm for small residuals and as the L_1 norm for large residuals. A threshold, which needs to be defined, controls where the transition between these two different behaviors takes place, with a more or less smooth shape depending on the criteria. These Hybrid criteria are efficient for dealing with outliers in data. However, they assume Gaussian statistics as soon as the L_2 norm is used, leading to the difficult issue of the estimation of the threshold.

This study presents applications of 2D elastic frequency-domain FWI for imaging realistic complex offshore and onshore structures in the presence of noisy synthetic data. A special emphasis is put on the performances of different minimization criteria in the framework of efficient, elastic frequency-domain FWI.

In the next section, we briefly review the theoretical aspects of different possible norms and criteria that can be applied to frequency-domain FWI. Then we apply these objective functions to two synthetic datasets that are contaminated by ambient, random, white noise with and without outliers. We assess the sensitivity of the inversion to noise in the case of decimated noisy data. We show that the L_2 norm is highly sensitive to non-Gaussian errors and requires the consideration of denser frequency sampling to improve the signal-to-noise ratio of the model. The L_1 norm shows very robust behavior, even in the case of highly decimated data, and provides, therefore, an interesting alternative to the L_2 norm for the design of efficient, frequency-domain algorithms. Investigations of other functionals, i.e., the Huber and the Hybrid criteria, highlight the difficulties for finding the best threshold, which will require some tedious trial-and-error investigations.

5.1.3 Theory and algorithm

5.1.3.1 The least-squares norm

The least-squares formalism provides the most usual framework for frequency-domain FWI (Pratt & Worthington, 1990; Pratt, 1990). The L_2 functional is usually written in the following form :

$$\mathcal{C}_{L_2}^{(k)} = \frac{1}{2} \Delta \mathbf{d}^\dagger \mathbf{S}_d^\dagger \mathbf{S}_d \Delta \mathbf{d}, \quad (5.1)$$

where $\Delta \mathbf{d} = \mathbf{d}_{obs} - \mathbf{d}_{calc}^{(k)}$ is the data misfit vector for one source and one frequency, the difference between the observed data \mathbf{d}_{obs} and the modeled data $\mathbf{d}_{calc}^{(k)}$ computed in the model $\mathbf{m}^{(k)}$. k is the iteration number of the non-linear iterative inversion. Superscript \dagger indicates the adjoint operator, and \mathbf{S}_d is a diagonal weighting matrix that is applied to the misfit vector to scale the relative contributions of each of its components.

Differentiation of $\mathcal{C}_{L_2}^{(k)}$ with respect to the model parameters gives the following expression of the gradient :

$$\mathcal{G}_{L_2}^{(k)} = \mathcal{R} \left\{ \mathbf{J}^t \mathbf{S}_d^\dagger \mathbf{S}_d \Delta \mathbf{d}^* \right\}, \quad (5.2)$$

where \mathbf{J} is the Fréchet derivative matrix, t and * denote the transpose and conjugate operators, respectively, and \mathcal{R} denotes the real part of a complex number. The gradient of the misfit

function, equation (5.2), can be efficiently computed without explicitly forming \mathbf{J} , using the adjoint-state method (Plessix, 2006). This gives the following expression :

$$\mathcal{G}_{m_i}^{(k)} = \mathcal{R} \left\{ \mathbf{v}^t \frac{\partial \mathbf{A}^t}{\partial m_i} \mathbf{A}^{-1} \mathbf{S}_d^\dagger \mathbf{S}_d \Delta \mathbf{d}^* \right\}, \quad (5.3)$$

where \mathbf{A} is the forward-problem operator, which linearly relates the source \mathbf{s} to the wavefield, $\mathbf{v} : \mathbf{A} \mathbf{v} = \mathbf{s}$. The modeled seismic data used in FWI, \mathbf{d}_{calc} , is related to the seismic wavefield, \mathbf{v} , by a projection operator, which extracts the values of the seismic wavefield at the receiver positions. The sparse matrix $\partial \mathbf{A} / \partial m_i$ represents the radiation pattern of the diffraction by the model parameter m_i , and, hence, gives some insight on the sensitivity of the data to a specific class of parameter as a function of the aperture angle. The gradient of the misfit function can be interpreted as a weighted zero-lag convolution between the incident wavefield \mathbf{v} and the adjoint residual wavefield back-propagated from the receiver positions, $\mathbf{A}^{-1} \mathbf{S}_d^\dagger \mathbf{S}_d \Delta \mathbf{d}^*$.

The misfit function and its gradient, equations (5.1) and (5.3), are given for one source and one frequency. For multiple sources and frequencies, the expressions are obtained by summing the contribution of each source and frequency.

5.1.3.2 The least-absolute-value norm

We can extend the L_1 norm developed by Tarantola (1987); Crase *et al.* (1990) for real arithmetic numbers to the complex arithmetic required by frequency-domain data through the misfit function :

$$\mathcal{C}_{L_1}^{(k)} = \sum_{i=1, N} |s_{d_i} \Delta d_i|, \quad (5.4)$$

where $|x| = (xx^*)^{1/2}$, N is the number of elements in the misfit vector for one source and one frequency, and s_{d_i} are the elements of the diagonal \mathbf{S}_d . The gradient of the misfit function is given by :

$$\mathcal{G}_{L_1}^{(k)} = \mathcal{R} \left\{ \mathbf{J}^t \mathbf{S}_d \mathbf{r} \right\} \quad \text{with } r_i = \frac{\Delta d_i^*}{|\Delta d_i|} \quad \text{for } 1 \leq i \leq N, \quad (5.5)$$

where we assume that $|\Delta d_i| > 0$, considering the machine precision used. For all of the tests that we have performed, we never met any case where $|\Delta d_i| = 0$. In the case of real arithmetic numbers, the term $\Delta d_i^* / |\Delta d_i|$ of expression (5.5) corresponds to the *sign* function (Tarantola, 1987; Crase *et al.*, 1990).

5.1.3.3 The Huber criterion

The Huber (1973) criterion can be defined for complex arithmetic numbers as :

$$\mathcal{C}_{L_{Huber}}^{(k)} = \sum_{i=1, N} c_i \quad \text{with } \begin{cases} c_i = \frac{|s_{d_i} \Delta d_i|^2}{2\epsilon} & \text{for } |s_{d_i} \Delta d_i| \leq \epsilon, \\ c_i = |s_{d_i} \Delta d_i| - \frac{\epsilon}{2} & \text{for } |s_{d_i} \Delta d_i| > \epsilon, \end{cases} \quad (5.6)$$

where ϵ is a threshold that controls the transition between the L_1 and the L_2 norms. In equation (5.6), the Huber criterion is continuous for all Δd_i , and particularly for the value that satisfies $|s_{d_i} \Delta d_i| = \epsilon$.

The gradient of the Huber functional is given by :

$$\mathcal{G}_{L_{Huber}}^{(k)} = \mathcal{R}\left\{\mathbf{J}^t \mathbf{S}_d \mathbf{r}\right\} \quad \text{with} \quad \begin{cases} r_i = \frac{s_{d_i} \Delta d_i^*}{|\Delta d_i|} & \text{for } |s_{d_i} \Delta d_i| \leq \epsilon \text{ and } 1 \leq i \leq N, \\ r_i = \frac{\Delta d_i^*}{|\Delta d_i|} & \text{for } |s_{d_i} \Delta d_i| > \epsilon \text{ and } 1 \leq i \leq N. \end{cases} \quad (5.7)$$

5.1.3.4 The Hybrid L_1/L_2 criterion

Bube & Langan (1997) introduced a Hybrid L_1/L_2 criterion in order to overcome some limitations of the Huber criterion that introduces artificial nonuniqueness in full-rank linear problems (Bube & Nemeth, 2007).

The Hybrid L_1/L_2 functional can be written for complex arithmetic numbers as :

$$\mathcal{C}_{L_{Hybrid}}^{(k)} = \sum_{i=1, N} c_i \quad \text{with } c_i = \left(1 + \frac{|s_{d_i} \Delta d_i|^2}{\epsilon^2}\right)^{1/2} - 1, \quad (5.8)$$

where ϵ is the threshold between the L_1 and L_2 behaviors. Note that we have the following properties :

$$c_i \approx \begin{cases} \frac{|s_{d_i} \Delta d_i|^2}{2\epsilon^2} & \text{for small } \Delta d_i, \\ \frac{|s_{d_i} \Delta d_i|}{\epsilon} - 1 & \text{for large } \Delta d_i, \end{cases} \quad (5.9)$$

which show that the Hybrid functional is asymptotically equivalent to the L_2 and L_1 norms for small-amplitude and large-amplitude residuals, respectively.

The gradient of the Hybrid L_1/L_2 criterion is given by :

$$\mathcal{G}_{L_{Hybrid}}^{(k)} = \mathcal{R}\left\{\mathbf{J}^t \mathbf{S}_d^\dagger \mathbf{S}_d \mathbf{r}\right\} \quad \text{with } r_i = \frac{\Delta d_i^*}{\epsilon^2 \left(1 + \frac{|s_{d_i} \Delta d_i|^2}{\epsilon^2}\right)^{1/2}} \quad \text{for } 1 \leq i \leq N. \quad (5.10)$$

5.1.3.5 Interpretation

Equations (5.2), (5.5), (5.7) and (5.10) clearly show that the gradients of the misfit functions have similar forms, but with different source terms for the back-propagated adjoint wavefield. This implies that the same FWI algorithm can be used to compute the gradients of the different misfit functions with the same computational cost, provided that the source term of the adjoint back-propagated wavefield and the misfit function can be computed for each functional.

Figure 5.1 shows the misfit function and the source term of the back-propagated wavefield as functions of the real arithmetic unweighted misfit Δd for the four minimization criteria. The L_2 norm naturally gives a high weight to large residual, which leads to a lack of robustness for this approach in the case of incoherent large errors in data. For the L_1 norm, the data residuals are normalized according to their amplitudes, which gives clear insights into why this is expected to be less sensitive to large residuals. The Huber and Hybrid criteria follow the L_2 and L_1 behaviors for small and large residuals, respectively, defined by the threshold ϵ , with a different transition shape.

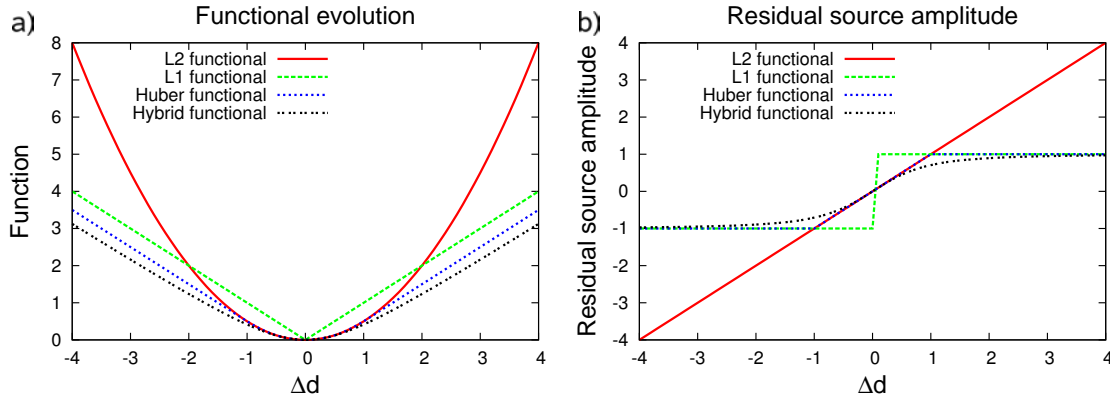


FIG. 5.1 – (a) The values of the criteria as functions of an unweighted real arithmetic misfit Δd ; and (b) the associated residual source amplitude in the gradient expression. L_2 , L_1 , and the Huber and Hybrid functionals are shown in the red, green, blue and black lines, respectively. The last two criteria are plotted for $\epsilon = 1$.

5.1.3.6 Algorithm

The 2D elastic, frequency-domain FWI algorithm used in this study is described in Brossier *et al.* (2009a), where the reader is thus referred to for a complete description of the algorithm. The algorithm embeds three main loops :

1. The outer loop is over the frequency groups, namely, a set of frequencies simultaneously inverted. In the case of complex wave phenomena, such as P-to-S conversions, and multiples and surface waves, the simultaneous inversion of multiple frequencies better constrains the optimization for convergence towards the global minimum, taking into account more redundant information. The frequency interval within the groups is driven by the necessity to use as much redundant data as possible during one FWI iteration, while keeping the numerical cost affordable. Moreover, the highest frequency of one frequency group is chosen by trial-and-error such that cycle-skipping artifacts are avoided.
2. The second loop is over time-damping factors that control the amount of information that is preserved over time in the seismograms during FWI. Time damping is applied in frequency-domain modeling by using complex-valued frequencies, given by $f_c = f + i\gamma$ where f is the real frequency, and γ is the damping factor. This is equivalent to the damping of seismograms in time by $e^{-\gamma t}$ (Shin *et al.*, 2002; Brenders & Pratt, 2007b). This damping can be applied from an arbitrary traveltime t_0 , which can be defined as the first-arrival traveltime. Moreover, this data preconditioning provides significant improvement to the signal-to-noise ratio in the case of real data (Brenders *et al.*, 2009). The choice of the damping factors is heuristic, and is driven by the data complexity. We generally start inversion with a weak amount of data close to the first arrival using high γ values, before progressively including later arrivals by decreasing the γ factor. The reader is referred to Brossier *et al.* (2009a) for an illustration of the significant impact of the damping factors in elastic FWI of land data where body waves and surface waves are jointly inverted.
3. The third loop is over iterations of the non-linear inversion of one frequency group and one damping factor.

The two outer nested loops define two hierarchical multiscale levels in the inversion that are helpful to mitigate the non-linearity of the inversion, particularly in the case of elastic multi-parameter inversion of complex data.

The forward problem is performed with a finite-element discontinuous Galerkin method for solving the elastodynamic equations in the frequency domain (Brossier *et al.*, 2008). The linear system resulting from this discretization is solved in parallel using the MUMPS LU solver (Amestoy *et al.*, 2006).

The optimization problem is solved with the quasi-Newton L-BFGS algorithm (Nocedal, 1980). In quasi-Newton algorithm, the updated model is given by :

$$\mathbf{m}^{(k+1)} = \mathbf{m}^{(k)} + \alpha^{(k)} H^{(k)} \mathcal{G}^{(k)}, \quad (5.11)$$

where $H^{(k)}$ is an approximation of the inverse of the Hessian, where the Hessian operator is formed by the second derivative of the misfit function with respect to the model parameters, and $\alpha^{(k)}$ is the step length. The L-BFGS algorithm estimates the product $H^{(k)} \mathcal{G}^{(k)}$ at a negligible computational cost, from few gradients and models of the previous iterations. Estimation of the off-diagonal terms of the Hessian improves the resolution of the imaging by correctly scaling and deconvolving the gradient, compared to a more classic preconditioned conjugate gradient (Brossier *et al.*, 2009a). Newton or quasi-Newton algorithms are generally applied to quadratic or locally-quadratic misfit functions. In contrast, the L_1 norm and the Huber and Hybrid criteria are not twice continuously differentiable, a condition for the convergence of the L-BFGS algorithm. Nonetheless, the numerical examples performed in this study show that the violation of the convergence conditions did not significantly affect, in practice, the convergence of the L-BFGS algorithm. Similar conclusions were derived by Guitton & Symes (2003), who applied the L-BFGS algorithm with the Huber criterion for velocity analysis. The convergence of the L-BFGS algorithm for functional which are not twice continuously differentiable can be interpreted by the fact that only an approximation of the inverse of the Hessian, not the exact one, is computed in the L-BFGS algorithm from only the first derivative (i.e., gradient) of the misfit functions at previous iterations (Guitton & Symes, 2003). For all the applications presented in this study, the informations of the five previous iterations were used for L-BFGS. The diagonal of the L_2 pseudo-Hessian matrix (Shin *et al.*, 2001) is used as an initial guess for the L-BFGS algorithm for each minimization criterion investigated in this study. The aim of the diagonal pseudo-Hessian is to remove the geometrical spreading signature of the incident and backpropagated wavefields from the gradient amplitudes.

The perturbation model estimated at each iteration is regularized with an adaptive Gaussian smoothing operator, to filter out high frequencies artifacts that are not constrained by the current group of inverted frequencies. The local vertical correlation length of the Gaussian filter is defined according to the inverted frequency and the local wavespeeds such that wavelengths significantly smaller than half the propagated wavelength (i.e., the maximum vertical resolution of FWI) are filtered out in the gradient (Sirgue & Pratt, 2004). The horizontal correlation length of the Gaussian smoother is estimated by trial-and-error, and should be driven by the lateral heterogeneity of the structure as well as the dip illumination provided by the available source-receiver offsets.

The line search in the perturbation direction is performed by parabolic fitting, to compute the optimal steplength $\alpha^{(k)}$, and to guarantee the decrease of the misfit function.

The source is estimated in the FWI algorithm by a linear inversion (Pratt, 1999, his equation 17). For one frequency, this gives for the scalar source :

$$s = \frac{\mathbf{g}^t \mathbf{d}_{obs}^*}{\mathbf{g}^t \mathbf{g}^*}, \quad (5.12)$$

where \mathbf{g} denotes the incident Green's functions recorded at receiver positions, computed in the starting model of the current iteration. The source is estimated in the FWI algorithm with equation (5.12), once the Green's functions \mathbf{g} are computed. The algorithm proceeds with the computation of the data residual $\Delta \mathbf{d}$, the backpropagation of the residuals, the computation of the gradient and pseudo-Hessian, the estimation of the step length, and the model update. The alternate estimation of the source and the model is repeated at each iteration. Of note, the source is estimated with the least-squares norm even in presence of outliers, whatever the norm used for the model update is. We found out that the source estimation was quite robust in presence of noise for the examples shown hereafter : the error does not exceed 5 % in amplitude and 1% in phase. The robustness of the source estimation can be attributed to the strong redundancy of the data for the mono-frequency scalar source estimation.

For all numerical tests of this study, the threshold value ϵ for the Huber and Hybrid criteria was fixed to $\epsilon = 0.2 \text{ mean}(|d_{obs_i}|)$. This value was shown practically to be less sensitive to outliers in the data than the one indicated by Guitton & Symes (2003), based on $\text{max}(|d_{obs_i}|)$.

5.1.4 Numerical tests : the offshore Valhall model

5.1.4.1 Inversion set-up

A first numerical example is based on the synthetic Valhall model (Figure 5.2), which is representative of oil and gas fields in shallow water environments of the North Sea (Munns, 1985). The main targets are a gas cloud in the large sediment layer, and in a deeper part of the model, the trapped oil underneath the cap rock, which is formed of chalk. Gas clouds are easily identified by the low P-wave velocities, whereas their signature is much weaker in the V_S model. The selected acquisition mimics a four-component ocean-bottom cable survey (Kommedal *et al.*, 2004), with a line of 315 explosive sources positioned 5 m below water surface, and 315 3C sensors on the sea bed. This geological setting leads to a particularly ill-posed problem for S-wave velocity reconstruction, due to the relatively small shear-wave velocity contrast at the sea bed, which prevents recording of significant P-to-S converted waves. A successful inversion requires a multi-step hierarchical strategy in the manner of Sears *et al.* (2008), as developed in Brossier *et al.* (2009b) for noise-free data. In this study, we assess the same approach for noisy data :

1. In the first step, the P-wave velocity is reconstructed from the hydrophone data. The forward problem is performed with the elastic discontinuous Galerkin method, but the V_S model is left unchanged during the FWI. The aim of this first stage is to improve the V_P model so as to significantly decrease the P-wave residuals. During this first step, a coarse mesh that is adapted to the V_P wavelength is designed for computational efficiency. In this case, S-wave modeling is affected by numerical dispersion that, however, does not significantly impact the reconstruction of the V_P model. This first stage is justified by the fact that the P-to-S converted waves have a minor footprint in the hydrophone component. This negligible sensitivity of the hydrophone data to the V_S structure has

allowed us for the successful acoustic inversion of the elastic data computed in the Valhall model. (Brossier *et al.*, 2009b).

2. In the second step, the V_P and V_S models are reconstructed simultaneously from the horizontal and vertical components of the geophones. An amplification, with a gain given by the power of 2 of the source/receiver offset, is applied to the data through the matrix \mathbf{S}_d . This weighting increases the weight of the intermediate-to-long-offset data at which the converted P-to-S arrivals are recorded.

Five frequencies were inverted successively (2, 3, 4, 5 and 6 Hz). The starting frequency (2 Hz) is lower than the one available in the real OBC hydrophone-data of Valhall (3.5 Hz) (Sirgue *et al.*, 2009). Note, however, that a starting frequency as small as 2 Hz was recently used to perform acoustic FWI of OBS data (Plessix, 2009). The use of such low starting frequency is required by the fact the V_S has a higher resolution power than V_P because of shorter propagated wavelength, and, hence, requires lower starting frequency or more accurate starting model. Our main concern in this study is to tune the elastic FWI with a reasonably-realistic experimental setup such that differences in the behaviors of the different functionals are highlighted. In the following, we shall not address anymore the impact of the starting frequency and the starting model in FWI, and will focus on the comparative performances of different data residual functionals for a given pair of starting frequency and model.

During each frequency inversion, we used 3 time-damping factors ($\gamma=2, 0.33, 0.1 \text{ s}^{-1}$) applied in cascade to the monochromatic data. For the smaller damping factor, the entire wavefield, including converted waves and free-surface multiples, is involved in the inversion. Starting models were built by smoothing the true models with a Gaussian filter, the vertical correlation length of which increased linearly from 25 m to 1,000 m with depth, while the horizontal one was fixed at 500 m (Figure 5.3). This smoothing should reasonably mimic the spatial resolution of a velocity model developed by refraction traveltime tomography (Prioux *et al.*, 2009). The deep part of this starting model is clearly smoother than a velocity model inferred from reflection traveltime tomography as the one shown in Sirgue *et al.* (2009).

Ten iterations were performed per damping factor, leading to 30 iterations per frequency inversion. V_P and V_S are the reconstructed parameters. The density is constant and assumed to be known in the inversion.

Two tests were performed, with and without outliers in the data. For both tests, a random uniform white noise was introduced into the observed data, computed using the forward-problem engine implemented in the inversion code (the so-called inverse crime). The observed data were computed using a (Dirac) delta function for the source wavelet. Therefore, each frequency component of the data has the same signal-to-noise ratio (SNR), since white noise was considered. Note however that the source wavelet spectrum has a negligible influence in frequency-domain FWI, where single frequencies or groups of frequencies of narrow bandwidth are generally inverted sequentially, namely, independently, in the framework of multiscale approaches.

The signal-to-noise ratio was set to 10 dB, based on the power value of the signal defined for one frequency by :

$$SNR(dB) = 10 \log_{10} \frac{P_{data}}{P_{noise}} \quad \text{with } P = \sum_{ishot=1}^{nshot} \sum_{irec=1}^{nreceiver} d_{ishot,irec} d_{ishot,irec}^* \quad (5.13)$$

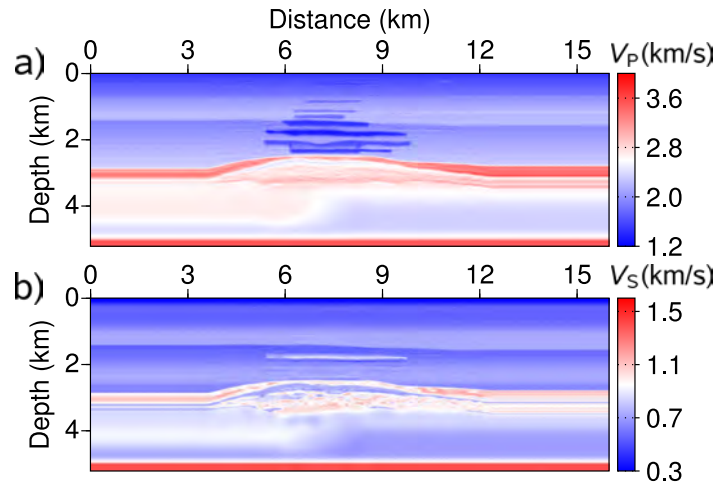


FIG. 5.2 – The true synthetic Valhall model for (a) P-wave and (b) S-wave velocities.

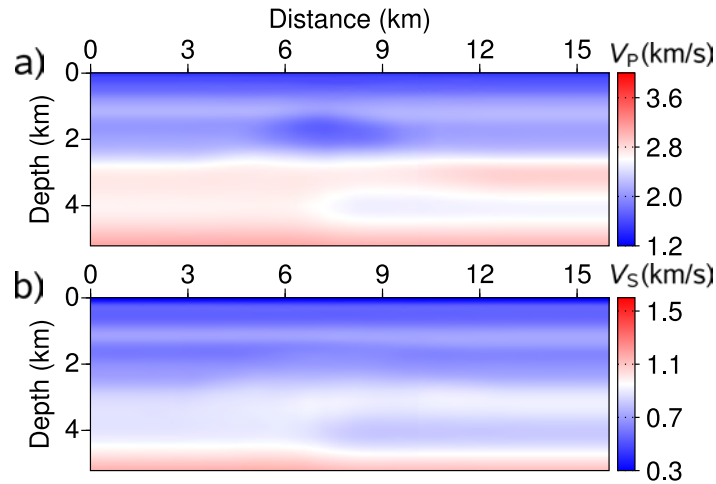


FIG. 5.3 – (a) V_P and (b) V_S starting models for FWI, as inferred by Gaussian smoothing of the true models (Figure 5.2).

Figure 5.4 shows the 4-Hz, noise-free and noisy data in the source/receiver domain for the hydrophone data.

5.1.4.2 Results

During the first test, we considered only the ambient noise. The V_P and V_S models inferred from the four minimization criteria (L_2 , L_1 , Huber, Hybrid) after the second step of inversion are shown in Figure 5.5. These reveal very good results for V_P models for all functionals, whereas only the robust L_1 norm, the Huber and Hybrid criteria provide acceptable V_S models.

In a second test, we introduced outliers into the data : large errors (i.e., the noise was locally multiplied by 20) were introduced randomly in one trace out of a hundred, to simulate

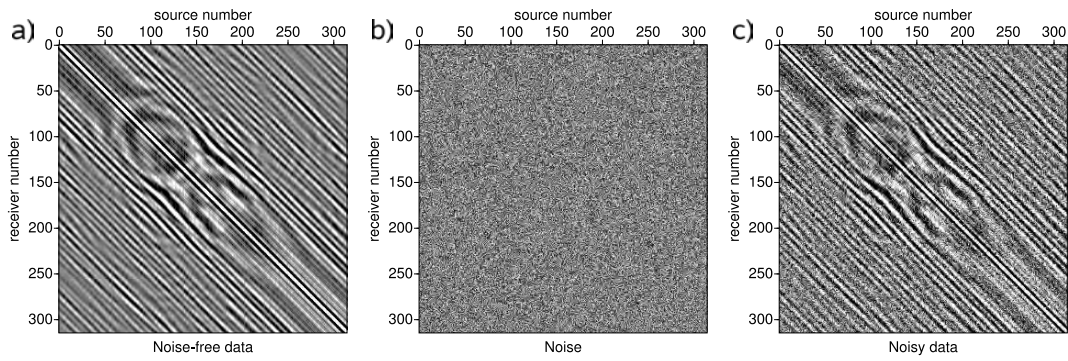


FIG. 5.4 – Real part of the 4-Hz frequency-domain data in the source/receiver domain for the Valhall model. (a) Noise-free hydrophone data ; (b) added noise ; and (c) resulting contaminated data used for FWI.

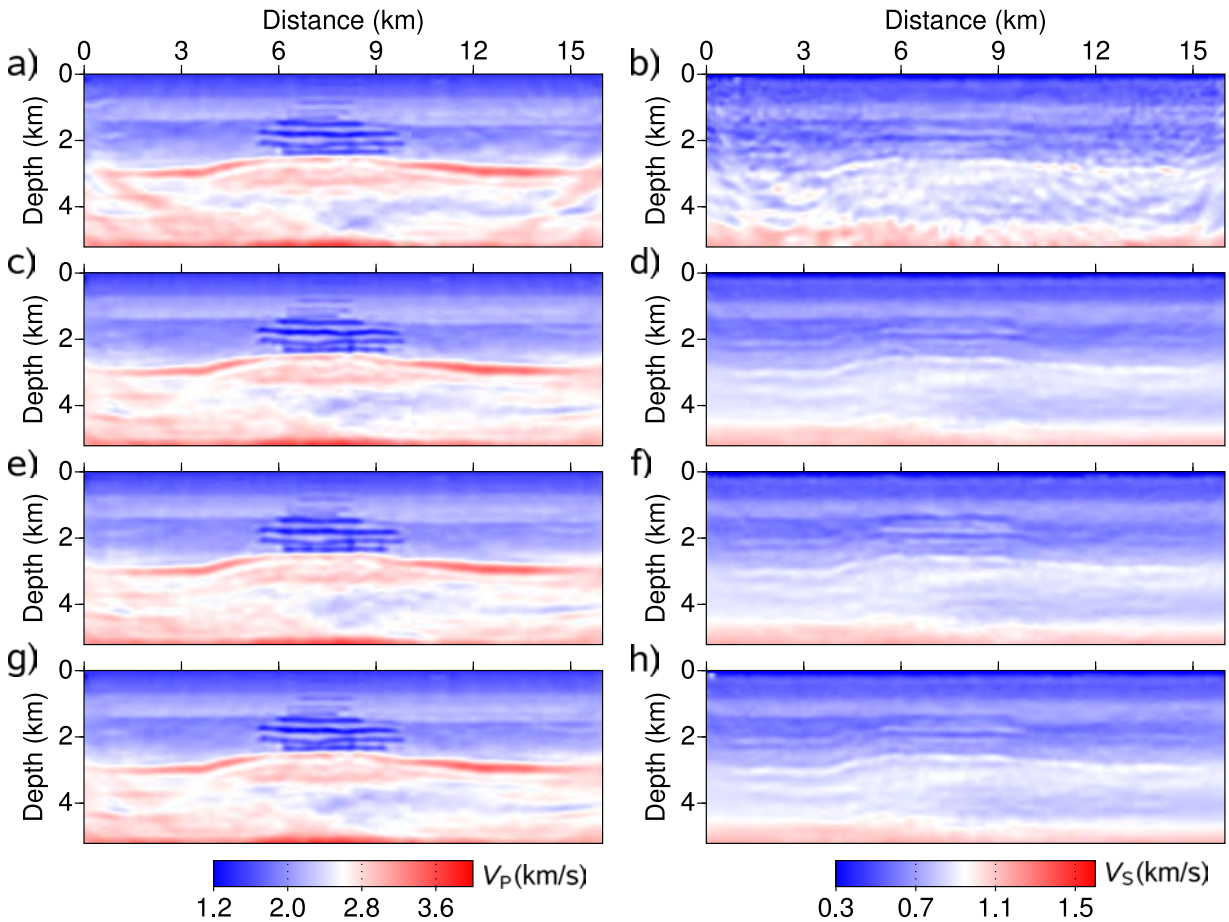


FIG. 5.5 – Reconstructed V_P (left panels) and V_S (right panels) models for the first Valhall test with the noisy data, after the two FWI steps. (a-b) L_2 norm ; (c-d) L_1 norm ; (e-f) Huber criterion ; and (g-h) Hybrid criterion.

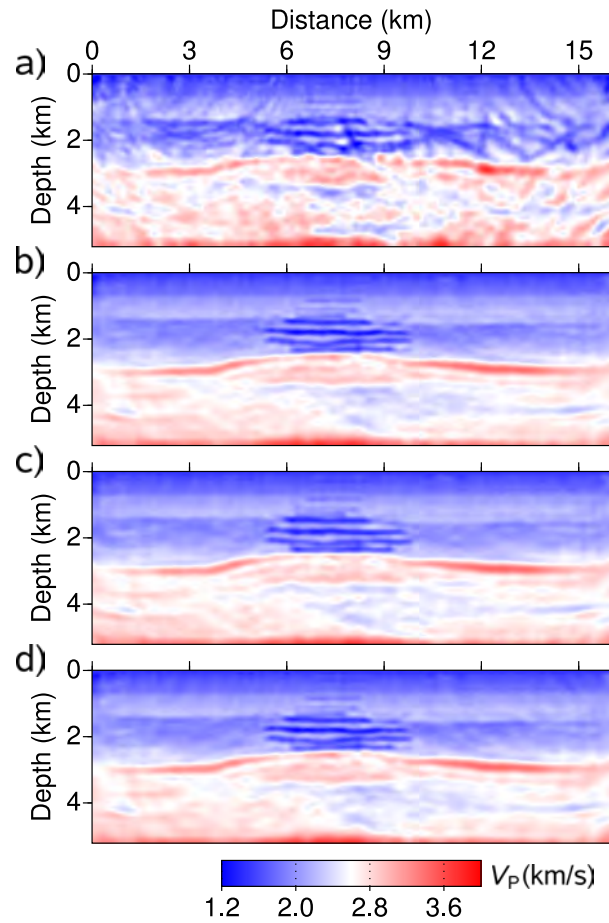


FIG. 5.6 – Reconstructed V_P models for the second Valhall test with the noisy data containing outliers, after the first FWI step. (a) L_2 norm; (b) L_1 norm; (c) Huber criterion; and (d) Hybrid criterion.

a poorly preprocessed dataset. The resulting noise is consequently no longer uniform for this test. The V_P models obtained after the first inversion step with the four functionals are shown in Figure 5.6. The L_1 norm, and the Huber and Hybrid criteria provide accurate V_P models, whereas the inversion rapidly converged towards a local minimum when the L_2 norm was used. For the L_2 norm, the inversion was stopped close to the first step because of the convergence towards a local minimum. The results obtained after the second inversion step performed with the L_1 norm showed reliable reconstruction of both the V_P and V_S models (Figure 5.7), which are close to those obtained from data without outliers (Figure 5.5(c-d)). This highlights the limited sensitivity of the L_1 norm to outliers even for the V_S reconstruction.

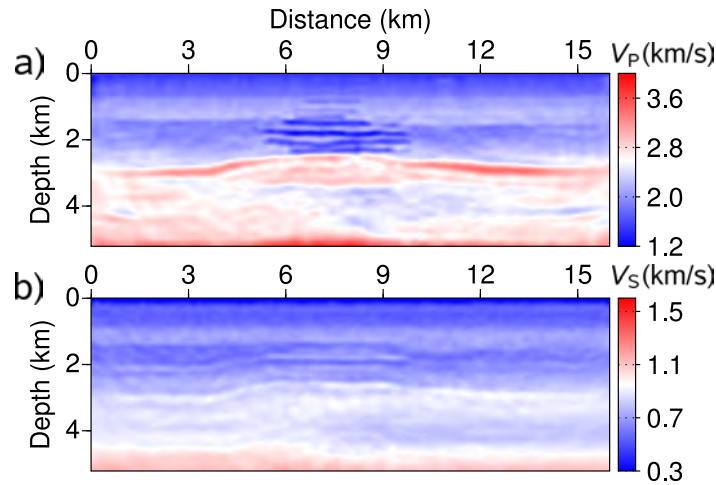


FIG. 5.7 – Reconstructed (a) V_P and (b) V_S models for the second Valhall test with the noisy data containing outliers, after the two FWI steps with the L_1 norm. Note that the models are very close to those of Figure 5.5(c)(d).

5.1.5 Numerical tests : The onshore SEG/EAGE overthrust model

5.1.5.1 Inversion set-up

A second numerical example focuses on the SEG/EAGE overthrust model. High-amplitude surface waves are present in the data of this onshore model, and need to be taken into account during inversion. Brossier *et al.* (2009a) designed a hierarchical scheme to invert the body and surface waves jointly, by using simultaneous frequency inversion. We followed this approach in this study. The acquisition geometry was composed of 199 explosive sources 20 m below the surface, and 198 vertical and horizontal geophones recording wavefields were located on the free surface.

Five discrete frequencies, distributed among two slightly overlapping frequency groups, were inverted : (1.7, 2.0, 3.5) and (3.5, 4.8, 7.2) Hz. The choice of these two groups of frequencies were shown to be efficient for inversion while limiting the numerical cost (Brossier *et al.*, 2009a). Five time-damping factors were used in cascade for each frequency group ($\gamma = 1.5, 1.0, 0.5, 0.1, 0.033$). We used a higher number of damping factors and a smaller interval between damping factors than for the Valhall example, because the overthrust case study is more non linear than the Valhall one, and, hence, requires injecting more progressively increasing amount of data. This increasing non linearity results from the presence of high-amplitude surface waves in the onshore overthrust case study.

Forty-five non-linear iterations were performed per damping factor. Figure 5.8 shows the true V_P model of the overthrust and the 500 m Gaussian smoothed version used as the starting model. A constant Poisson's ratio of 0.24 was fixed to build the true and the starting V_S models. The density was constant and assumed to be known during the inversion. V_P and V_S are the reconstructed parameters for the inversion of the horizontal and vertical components of the particle velocity.

The inverted data were computed with a Dirac source wavelet. Random uniform white noise

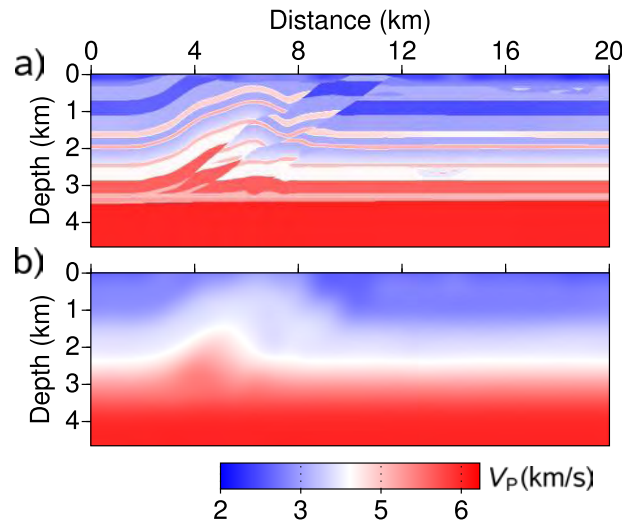


FIG. 5.8 – (a) True and (b) starting V_P models for the synthetic SEG/EAGE overthrust tests. The V_S models are derived from the V_P models using a constant Poisson ratio of 0.24.

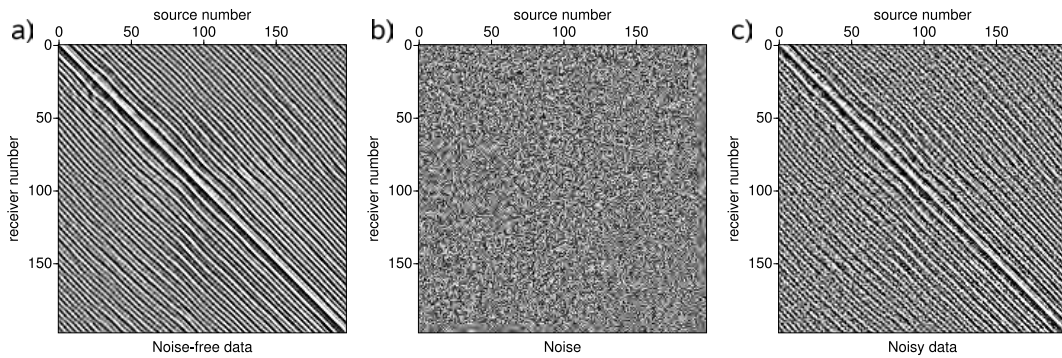


FIG. 5.9 – Real part of the 3.5-Hz frequency-domain data in the source/receiver domain for the overthrust model. (a) Noise-free horizontal component of particle velocity recorded by geophones; (b) added noise; and (c) resulting contaminated data used for FWI.

was introduced into the observed data, with a signal-to-noise ratio of 7 dB for each frequency component. Figure 5.9 shows the 3.5 Hz, noise-free and noisy data in the source/receiver domain for the horizontal component of particle velocity.

5.1.5.2 Results

The V_P and V_S models obtained with the different minimization criteria are shown in Figure 5.10. The L_1 norm again provides the most reliable results. The Huber and Hybrid criteria show quite robust results, in particular in the shallow part of the models, even if the Hybrid criterion suffers from high frequency artifacts despite the smoothing regularization operator applied to the perturbation model at each iteration. The models obtained with the L_2 norm are polluted by strong artifacts, particularly in the thrust area and in the deep structure.

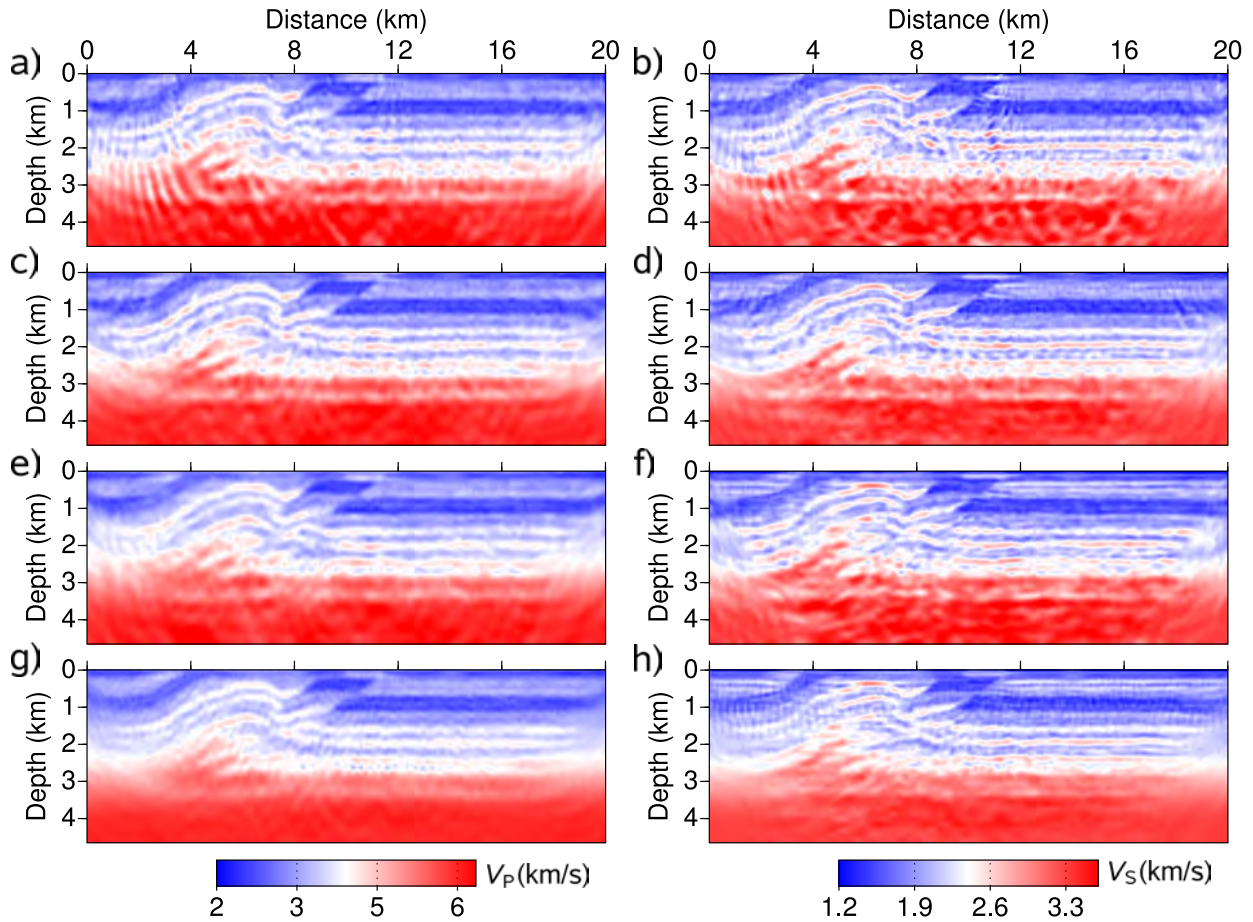


FIG. 5.10 – Reconstructed V_P (left panels) and V_S (right panels) models for the overthrust test, obtained by FWI. (a-b) L_2 norm ; (c-d) L_1 norm ; (e-f) Huber criterion ; and (g-h) Hybrid criterion.

5.1.6 Discussion

5.1.6.1 The offshore Valhall model

The results of the first test, where only ambient noise without outliers was considered, show reliable reconstruction of the V_P model for the four norms, whereas only the L_1 , the Huber and the Hybrid functionals provide reliable reconstruction of the V_S model.

In this shallow-water environment with low velocity contrasts at the sea bed, the V_P imaging is more linear than the V_S imaging for two main reasons. First, the larger P-wavelengths are less resolving than their S counterparts, and are therefore less sensitive to the inaccuracies of the starting model in the framework of a multi-scale reconstruction (Brossier *et al.*, 2009a). Secondly, the P-waves dominate the seismic wavefield, whereas the P-to-S waves have a weaker footprint in the data. The limited signature of the S-waves in the data makes the inversion poorly conditioned for the S-wave parameter class, even with noise-free data. Brossier *et al.* (2009b) showed how the hierarchical two-step strategy allows us to increase the sensitivity of the inversion to the V_S parameter during the second step, and hence the successful reconstruction

of the V_S model with the L_2 norm in the case of noise-free data. However, adding noise to the data still contributes to the weakening of the sensitivity of FWI to the P-to-S arrivals. In this case, the two-step strategy implemented with the L_2 norm failed to reconstruct the V_S model. In contrast, the L_1 norm, the Huber and the Hybrid criteria successfully converge towards acceptable V_S models by mitigating the contribution of the residual amplitude in the reconstruction.

The second test, where outliers were introduced into the data, illustrates first the expected failure of the L_2 norm in the presence of high-amplitude isolated noise. The L_2 norm intrinsically amplifies the weight of the high-amplitude residuals during inversion, hence causing divergence of the optimization if the residuals do not correspond to useful seismic arrivals. The L_1 norm, as well as the Huber and Hybrid criteria, shows stable behavior for V_P imaging in this unfavorable context, because the isolated, high-amplitude outliers have a negligible contribution in these functionals. The strong robustness of the L_1 norm with respect to noise is illustrated by its ability to reconstruct the V_S model from low-amplitude P-to-S converted waves, even in the presence of outliers.

5.1.6.2 The onshore SEG/EAGE Overthrust Model

In an onshore context, where body waves and surface waves are jointly inverted, the data are very sensitive to both V_P and V_S parameters. For example, the V_S velocities on the near surface have a significant impact on the high-amplitude surface waves. If the starting V_P and V_S models for FWI are not sufficiently accurate, then high-amplitude, surface-wave residuals can direct the inversion towards a local minimum of the misfit function. In this context, data redundancy provided by the simultaneous inversion of multiple frequencies is essential to converge towards acceptable models (Brossier *et al.*, 2009a). In the case of noisy data, it might however be necessary to strengthen this redundancy, by decreasing the frequency interval within each frequency group when the L_2 norm is used. Figure 5.11 shows the V_P and V_S models obtained with the L_2 norm, when the number of frequencies per group was increased from three to nine : i.e., (1.7, 1.8, 2.0, 2.3, 2.5, 2.7, 3.0, 3.2, 3.5) and (3.5, 3.8, 4.1, 4.4, 4.8, 5.2, 6.0, 6.5, 7.2) Hz. Improvements of both the V_P and V_S models, compared to those of Figure 5.10(a-b), clearly show that increasing the data redundancy improves the signal-to-noise ratio of the models, at the expense of the computational cost. On the contrary, the L_1 norm shows more stable behavior in the presence of noise, and therefore it provides more reliable results with efficient frequency-domain FWI, where only a few frequencies are involved in the inversion. The Huber and Hybrid criteria show less convincing results compared to the Valhall experiment, highlighting the difficult tuning of the threshold ϵ .

5.1.6.3 Implications for 3D FWI

Some implications for 3D FWI can be derived from the results of the synthetic experiments. Historically, the two main motivations behind the use of 2D frequency-domain modeling and inversion were : (1) to design computationally efficient algorithms by limiting the modeling and the inversion to a few discrete frequencies ; and (2) to design multiscale strategies by proceeding from low frequencies to higher ones (Pratt & Worthington, 1990). For 3D problems, comparative analyses of modeling methods has shown that time-domain, explicit-scheme modeling provides a computationally efficient alternative to frequency-domain modeling methods based either on

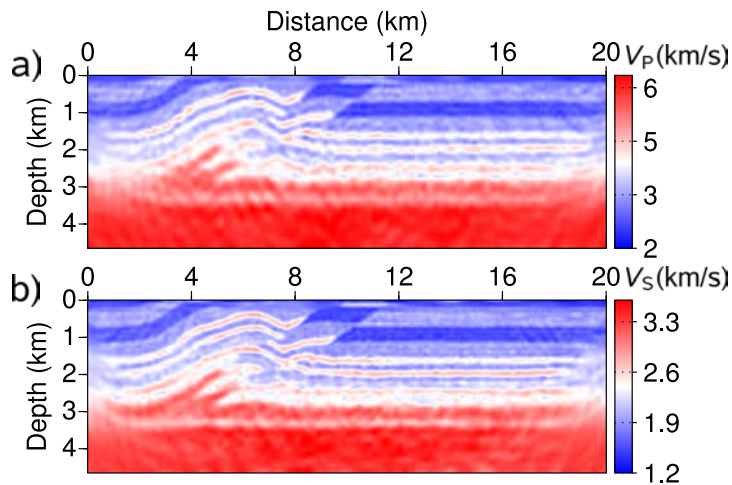


FIG. 5.11 – Reconstructed (a) V_P and (b) V_S overthrust models after FWI using the L_2 norm applied to a dense dataset of two slightly overlapping groups of nine frequencies each. Note the improvement in the signal-to-noise ratio compared to Figure 5.10(a-b).

direct or iterative solvers to perform frequency-domain FWI (Nihei & Li, 2007; Plessix, 2007; Sirgue *et al.*, 2008; Virieux *et al.*, 2009).

Our results have highlighted the benefits provided by the data redundancy for performing robust elastic FWI, where both the V_P and the V_S parameters are reconstructed and the L_2 norm is used. Therefore, is it worth leaving the inversion in the frequency domain if the time-domain modeling engine provides the opportunity to fully exploit the data redundancy by mean of time-domain FWI? A more quantitative analysis of the computational burden that results from implementation of the inversion in the time domain compared to the frequency domain might be needed to answer this question. The computational burden resulting from implementation of the inversion in the time domain might be due to disk access or extra forward-problem resolutions that depend on the implemented numerical strategy (Akcelik, 2002; Liu & Tromp, 2006; Symes, 2007). Multiscale strategies in time-domain FWI can be implemented using the approach of Bunks *et al.* (1995), where successive inversions of datasets of increasing high-frequency content are performed. If efficient 3D elastic frequency-domain FWI must be performed, the L_1 norm definitively provides the more robust criterion for the noise levels investigated in this study.

5.1.7 Conclusions

Application of elastic FWI to offshore and onshore data shows the strong sensitivity of the L_2 norm to non-Gaussian errors in the data when decimated discrete frequencies in FWI are considered for computational efficiency. The marginally used L_1 norm appears to be weakly sensitive to noise even in the presence of outliers, and provides stable results for onshore and offshore FWI applications. In particular, the L_1 norm allows successful inversion of low-amplitude P-to-S arrivals for reliable V_S model building in an offshore environment, even in the presence of noise. Alternative functionals, such as the Huber and Hybrid criteria, which combine the L_1 and L_2 norms, can provide stable results if the threshold controlling the transition between

the L_1 and the L_2 behaviors is well chosen. The judicious estimation of this threshold by trial-and-error tests is a clear drawback of the Huber and Hybrid criteria, even if these functionals should be as good as the L_1 norm if the threshold is well chosen. More automatic functionals, such as the L_1 or L_2 norms, should therefore be recommended for inversion of field data. The L_1 norm reveals an interesting alternative to the L_2 norm, especially when decimated datasets are processed by efficient frequency-domain FWI. The results obtained with the L_2 norm can, however, be improved in the presence of ambient noise, by increasing the data redundancy through simultaneous inversion of multiple frequencies at the expense of the computational cost. Some implications for 3D FWI can be derived from these conclusions. On the one hand, the L_1 norm might be an appealing alternative to the L_2 norm for FWI based on frequency-domain seismic modeling, which requires consideration of a limited number of frequencies for computational efficiency. On the other hand, explicit time-marching algorithms are competitive with frequency-domain seismic modeling methods for 3-D problems. If the wavefields are computed fully in the time-domain for computational efficiency, the inversion of decimated data in the frequency domain may not provide a significant computational saving compared to a time-domain inversion. In this case, the inversion may be left in the time domain, to fully take advantage of the data redundancy.

Acknowledgments

This study is funded by the SEISCOPE consortium (<http://seiscope.unice.fr>), which is sponsored by BP, CGG-VERITAS, EXXON-MOBIL, SHELL and TOTAL, and by Agence Nationale de la Recherche (ANR) under project ANR-05-NT05-2-42427. The LU factorization of the impedance matrix was performed with MUMPS (<http://graal.ens-lyon.fr/MUMPS/index.html>). The mesh generation was performed with the help of TRIANGLE (<http://www.cs.cmu.edu/~quake/triangle.html>). Access to the high-performance computing facilities of the MESOCENTRE SIGAMM computer centre and IDRIS national (project 082280) centers provided the required computer resources, and we gratefully acknowledge both of these facilities and the support of their staff. Many thanks go to J. Kommendal and L. Sirgue from BP, for providing the elastic synthetic models of Valhall. We would like to thank Bill Harlan, Partha Routh and three anonymous reviewers for their fruitful comments on the manuscript.

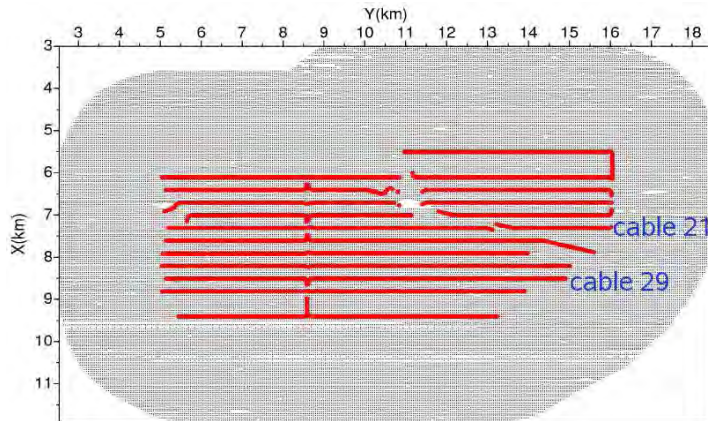


FIG. 5.12 – Dispositif d’acquisition 3D de la zone de Valhall. Les câbles en fond de mer sont localisés en rouge. Les points noirs représentent les sources. Seuls les câbles de récepteurs 21 et 29, et les lignes de sources associées, seront utilisés dans les tests 2D. (Modifié d’après Sirgue *et al.*, 2009).

5.2 Application acoustique sur les données réelles de Valhall

5.2.1 Environnement, données et configuration de l’inversion

Je présente une application d’inversion de formes d’ondes pour l’imagerie de V_P à partir d’un jeu de données pétrolières du champs de Valhall en Mer du Nord, fourni par BP. L’environnement est proche du modèle synthétique (*cf.* partie 4.2) avec une faible profondeur d’eau (environ 70 m) (Munns, 1985). La zone est caractérisée par une épaisse couche de sédiments superficielle, piégeant diverses poches de gaz. Plus profondément, un anticlinal formé de calcaire crayeux forme le toit rocheux d’un réservoir de gaz et de pétrole. Cette zone est relativement difficile à imager car les poches de gaz produisent des anomalies lentes de vitesses d’ondes P, fortement atténuantes, qui diminuent le pouvoir de pénétration des composantes hautes fréquences des sources.

Un dispositif d’acquisition multicomposantes permanent par câble (OBC) est installé en fond de mer au dessus des champs (Kommedal *et al.*, 2004). La Figure 5.12 montre la géométrie 3D du dispositif. Les câbles sont équipés de capteurs tous les 50 m. Les coupes de la Figure 5.13 montrent les structures imagées par inversion des formes d’ondes acoustique 3D de Sirgue *et al.* (2009). La partie superficielles du nuage de gaz dans l’épaisse couche de sédiments est bien visualisée par l’anomalie de vitesses lente à 1050 m de profondeur sur la Figure 5.13.

Cette application est menée dans le cadre de l’approximation acoustique pour la modélisation et l’inversion, avec l’algorithme décrit dans le chapitre 3 en mode acoustique. Cette approximation permet de réduire le coût du problème direct. De plus, les tests menés sur le modèle synthétique ont montré que cette approximation était acceptable dans ce contexte marin.

Deux lignes de données sont considérées :

- le câble 21 de longueur 10.8 km localisé à proximité du nuage de gaz. La ligne de 320 tirs s’étend sur 16 km au dessus du câble. (*cf.* Figures 5.12 et 5.13)

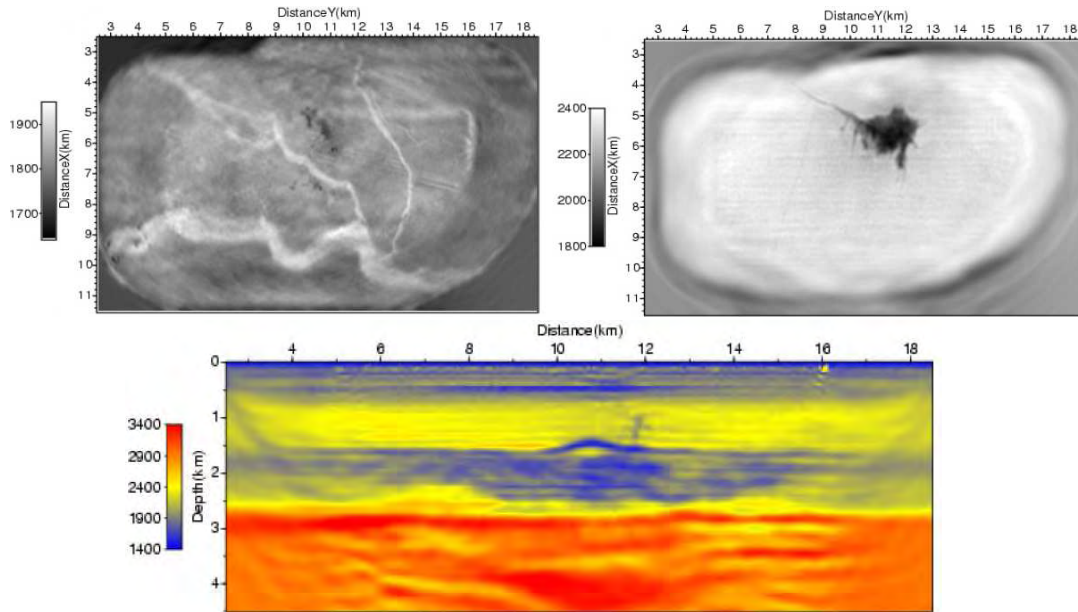


FIG. 5.13 – Modèle de vitesse 3D imagé par inversion de formes d’ondes 3D acoustique, par Sirgue *et al.* (2009). (haut) Coupes horizontales localisées à 150 m (gauche) et 1050 m (droite) de profondeur. La coupe superficielle montre les différents chenaux présents dans la couche sédimentaire superficielle, tandis que la coupe plus profonde montre le toit du nuage de gaz donnant l’anomalie de vitesses lentes. (bas) Coupe verticale à la distance $X=6500$ m.

- le câble 29 de longueur 9.75 km localisé en dehors du nuage de gaz. La ligne de 314 tirs s’étend sur 15.7 km au dessus du câble. (*cf.* Figures 5.12 et 5.13)

Afin de minimiser l’impact des ondes converties P-S, les données d’hydrophone sont utilisées. Sous l’approximation acoustique de la propagation, la réciprocité spatiale entre sources et capteurs est utilisée afin de minimiser le nombre de simulations directes dans l’inversion. La Figure 5.14 montre des exemples de données temporelles triées en récepteur commun, pour chacune des deux lignes qui vont être utilisées pour l’inversion. Le spectre des données nous permet de commencer l’inversion à partir de 4 Hz. Quatre fréquences discrètes sont considérées : 4, 5, 6 et 7 Hz. Un facteur de régularisation $\gamma = 1$ est utilisé pour minimiser le poids des arrivées tardives. Les prétraitements des données dédiées à l’inversion ont consisté à :

- Déconvolution (blanchiment spectral à phase minimum).
- Contrôle qualité.
- Filtrage dans la bande de fréquences utilisée (filtre Butterworth).
- Filtre de cohérence par matrice spectrale (Mari *et al.*, 1997) (avec application d’un gain automatique avant le filtrage puis suppression du gain après)
- Filtre en k pour supprimer les arrivées lentes des ondes guidées dans la couche d’eau.
- Mute avant la première arrivée
- Application de l’exponentielle décroissante en domaine temporel pour $\gamma = 1$
- Transformée de Fourier avec extraction des fréquences 4, 5, 6 et 7 Hz

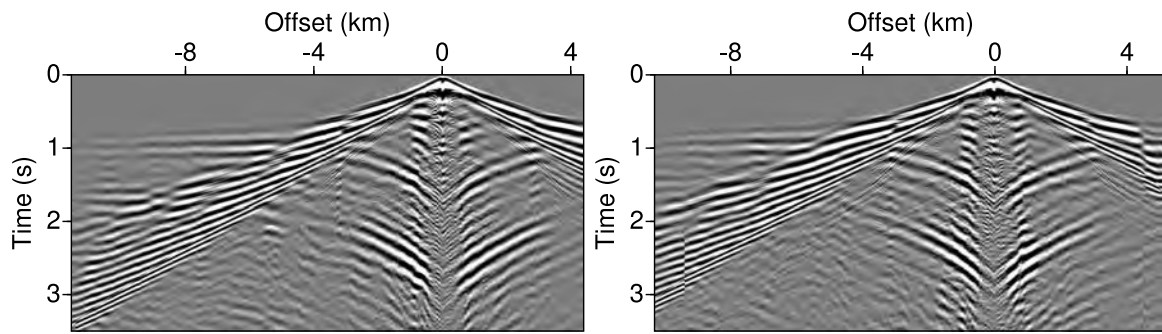


FIG. 5.14 – Exemples de traces triées en récepteur commun prétraités pour les lignes 21 (gauche) et 29 (droite).

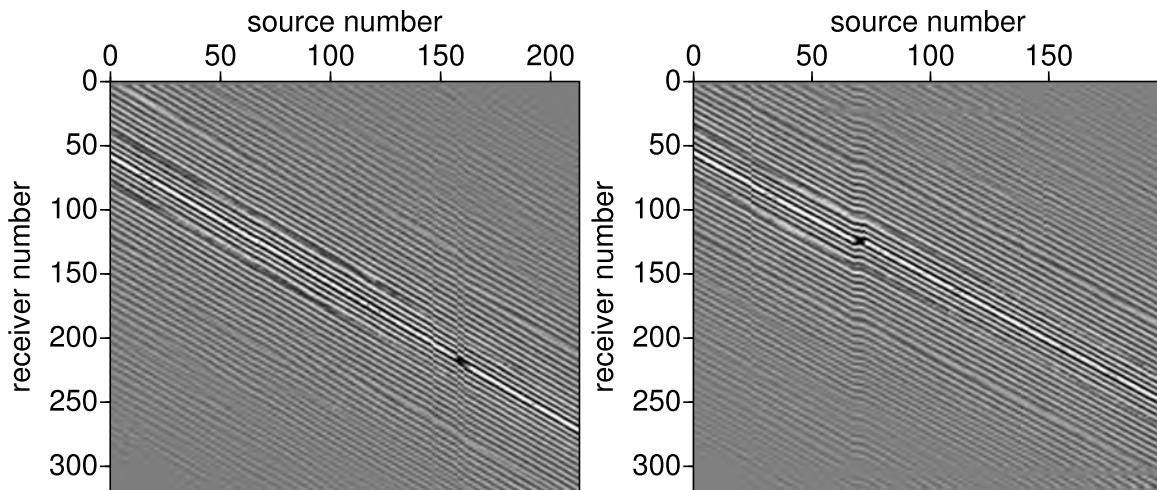


FIG. 5.15 – Exemples de données prétraitées en domaine fréquentiel. Partie réelle des données dans l'espace récepteur/source à 7 Hz pour les lignes 21 (gauche) et 29 (droite).

La Figure 5.15 montre des données monochromatiques prétraitées dans l'espace source/récepteur à 7 Hz.

Les modèles de départ utilisés pour l'inversion sont issus d'une tomographie en réflexion, fournis par BP, donnant une bonne estimation des vitesses de propagation verticales. Ce modèle est ensuite corrigé cinématiquement par une tomographie des premières arrivées plus représentative des vitesses de propagation horizontales suivies par les ondes réfractées. Il est cependant clair que les effets d'anisotropie présents dans cette région seraient mieux pris en compte en utilisant des paramètres anisotropes dans l'outil de modélisation. La Figure 5.16 montre les modèles initiaux utilisés.

L'inversion est menée avec l'algorithme L-BFGS avec la diagonale du pseudo-Hessien pour estimation initiale. Vingt itérations sont effectuées par fréquence discrète. Le but de ces tests est de quantifier la sensibilité des normes L_2 et L_1 pour l'inversion de données réelles. Ces deux normes sont donc appliquées pour chacun des jeux de données. La source sismique est estimée pendant l'inversion par inversion linéaire.

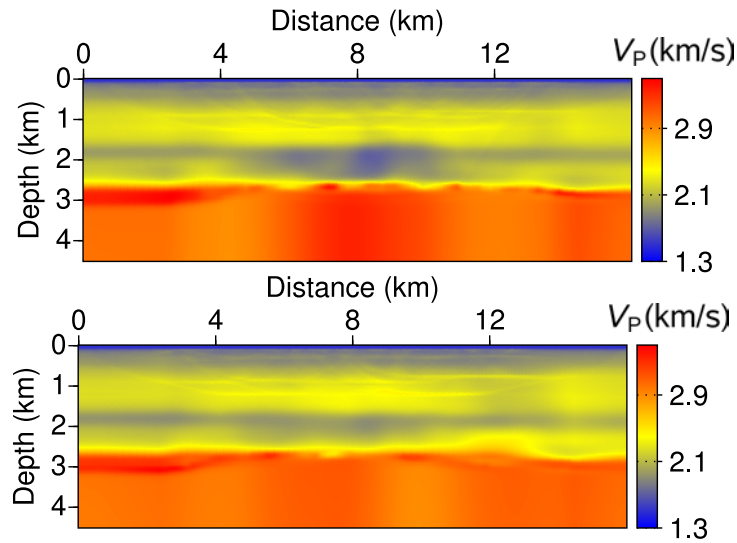


FIG. 5.16 – Modèles initiaux utilisés pour l’inversion des formes d’onde pour les lignes 21 (haut) et 29 (bas).

5.2.2 Résultats

5.2.2.1 Ligne 21

Les Figures 5.17 et 5.18 montrent les résultats d’inversion pour la ligne 21 obtenus avec les normes L_1 et L_2 . Les résultats des deux normes montrent une convergence vers des modèles très semblables, soulignant la bonne qualité du jeu de données et la redondance suffisante apportée par le choix des fréquences. Cette ligne, localisée à proximité du nuage de gaz met en évidence de fortes anomalies de vitesses lentes caractéristiques du gaz, sous 2 km de profondeur. Plus profondément, la base du réservoir est marquée par un réflecteur hautes vitesses vers 2600 m de profondeur. La Figure 5.19 montre des données en domaine temporel, illustrant les trains d’ondes produits par le modèle imagé. On peut noter que l’inversion a pu exploiter les différentes réflexions présentes dans les données, malgré des défauts d’amplitude. Ces défauts peuvent être expliqués par l’inversion basse fréquence (7 Hz maximum) et par la physique approximée utilisée : propagation acoustique avec densité constante et sans atténuation. Les deux réflexions dominantes sont associées au toit et à la base de la zone à faible vitesse. Les trains d’ondes réfléchies entre ces deux arrivées sont associés aux structures alternant anomalies de vitesses lentes et rapides, centrée sur 2 km de profondeur (profils 1D des Figures 5.17 et 5.18). Enfin, l’interface superficielle imagée vers 0.4 km de profondeur reproduit l’onde réfractée présente dans les données réelles. Cette interface est d’ailleurs présente sur le modèle de Sirgue *et al.* (2009) (Figure 5.13).

Les Figures 5.20, 5.21 et 5.22 montrent les données en domaine fréquentiel à l’issue de l’inversion ainsi que l’accord avec les données observées, pour la dernière itération de la dernière fréquence (7 Hz). Les deux normes donnent un ajustement très similaire entre les données observées et données calculées dans les modèles finaux. Pour les deux normes, notons des instabilités de reconstruction dans les zones superficielles non couvertes par les tirs.

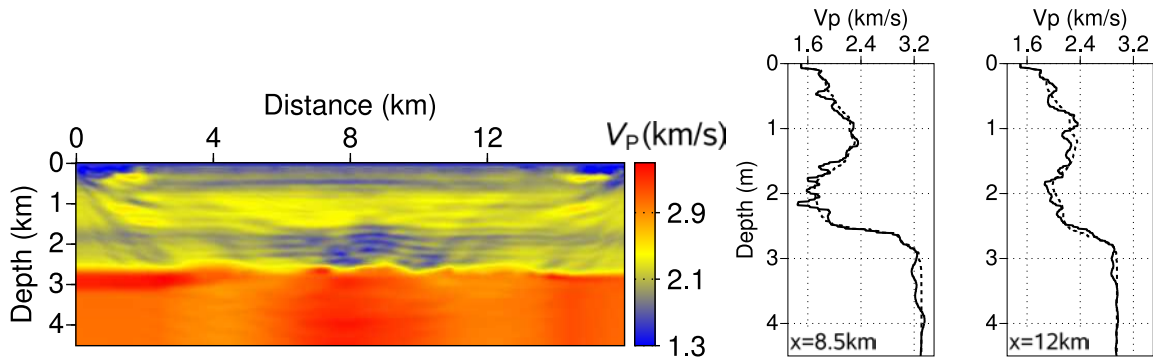


FIG. 5.17 – Modèles V_P imagés pour la ligne 21 avec la norme L_1 . Les coupes 1D verticales sont prises à 8.5 et 12 km. Les courbes discontinues et continues représentent respectivement le modèle initial et final de l'inversion.

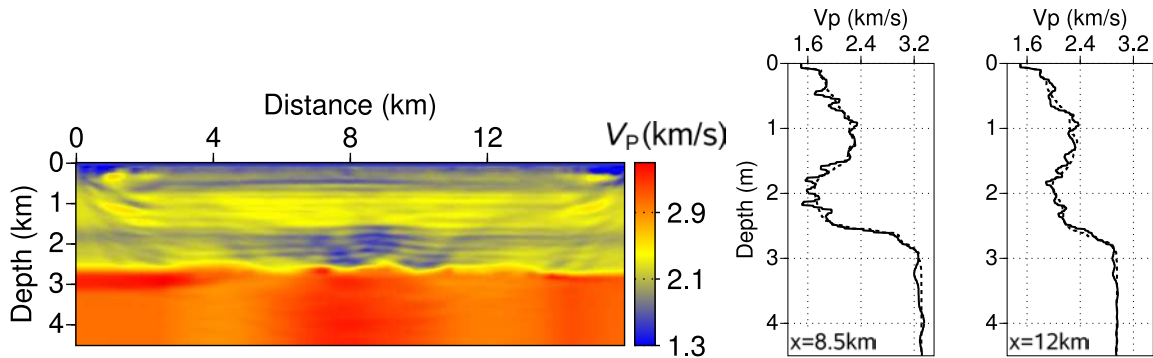


FIG. 5.18 – Modèles V_P imagés pour la ligne 21 avec la norme L_2 . Les coupes 1D verticales sont prises à 8.5 et 12 km. Les courbes discontinues et continues représentent respectivement le modèle initial et final de l'inversion. Notons que les résultats sont très semblables à la Figure 5.17 obtenue par norme L_1 .

5.2.2.2 Ligne 29

Les Figures 5.23 et 5.24 montrent les résultats d'inversion pour la ligne 29 obtenus avec les normes L_1 et L_2 . Les résultats des deux normes montrent cette fois des modèles différents. Contrairement au modèle de la norme L_2 qui a peu évolué, la norme L_1 a pu modifier le modèle de vitesse. Ces différences de modèles se traduisent par un bon accord des données pour la norme L_1 (Figures 5.25 et 5.27) dans toute la gamme d'offsets. Les données observées et calculées du modèle L_2 montrent clairement de plus gros résidus d'amplitude et de phase qui n'ont pu être exploités dans l'inversion (Figures 5.26 et 5.27). Cette faible convergence de la norme L_2 peut être partiellement expliquée par une plus grande quantité de traces inexploitable, systématiquement supprimées lors de l'étape de contrôle qualité. Cela diminue par conséquent la redondance des données et augmente le risque de laisser des traces de mauvaise qualité dans les données. Pour la norme L_1 , notons des instabilités de reconstruction dans les zones superficielles non couvertes par les tirs. La Figure 5.19 montre des données en domaine temporel, illustrant les trains d'ondes produits par le modèle imagé. Notons là encore des défauts d'amplitudes

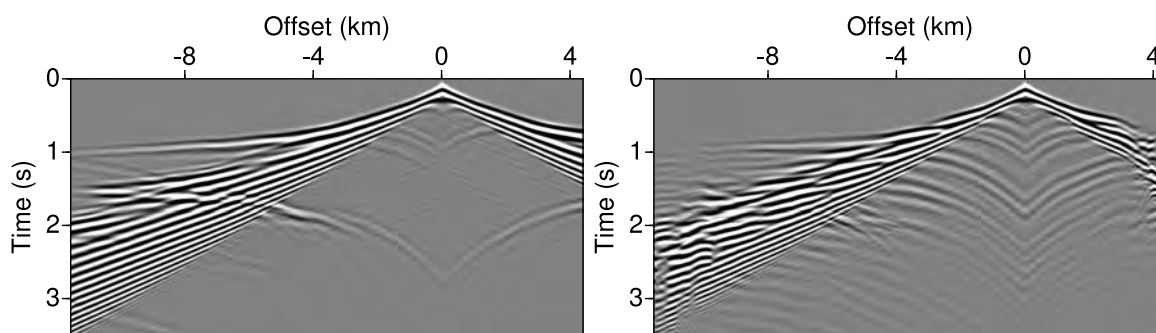


FIG. 5.19 – Exemple de données en domaine temporel calculées dans le modèle initial (gauche) et final (droite) de la ligne 21 (Modèle obtenu par norme L_1). Notons que les premières et dernières traces sont affectées par la reconstruction instable de la subsurface dans les zones non couvertes par les tirs.

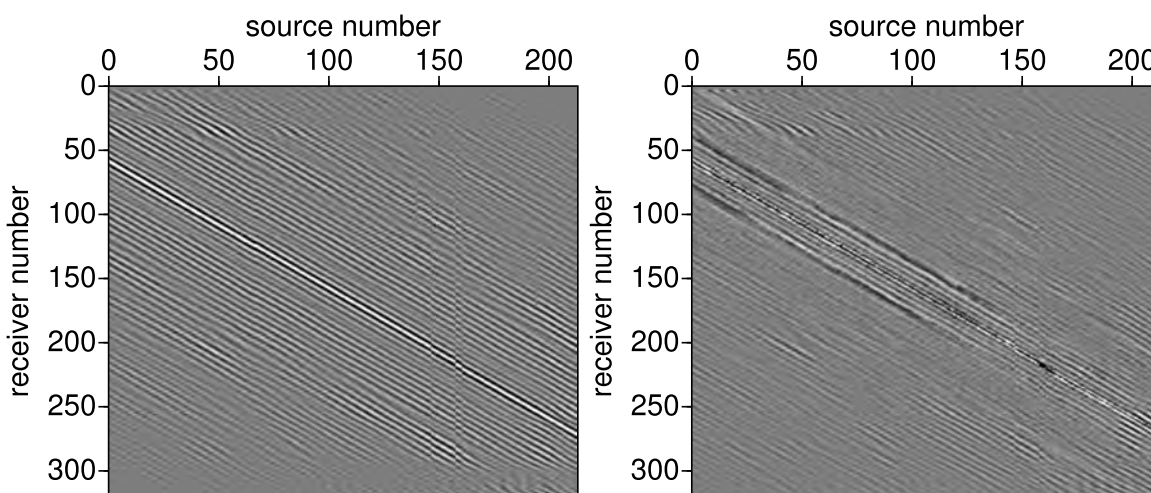


FIG. 5.20 – Partie réelle des données en fréquence à 7 Hz dans l'espace source/récepteur. Données calculées à l'issue de l'inversion de la ligne 21 avec la norme L_1 (gauche) et résidus associés (droite).

dans les données synthétiques calculées.

5.2.3 Conclusion

Cette application acoustique d'inversion des formes d'ondes a permis de mettre en oeuvre les normes L_1 et L_2 dans le contexte de données réelles du champ de Valhall. Pour la ligne 21, les deux formalismes ont convergé vers des modèles très similaires, comme c'est le cas pour les applications synthétiques lorsque les données utiles sont suffisamment redondantes. Les modèles obtenus sont cohérents avec les structures imagées par Sirgue *et al.* (2009) à partir d'une couverture 3D. En particulier, les images obtenues font apparaître l'interface lente à 0.4 km de profondeur et les anomalies lentes et rapide successives vers 2 km, marquant la présence du

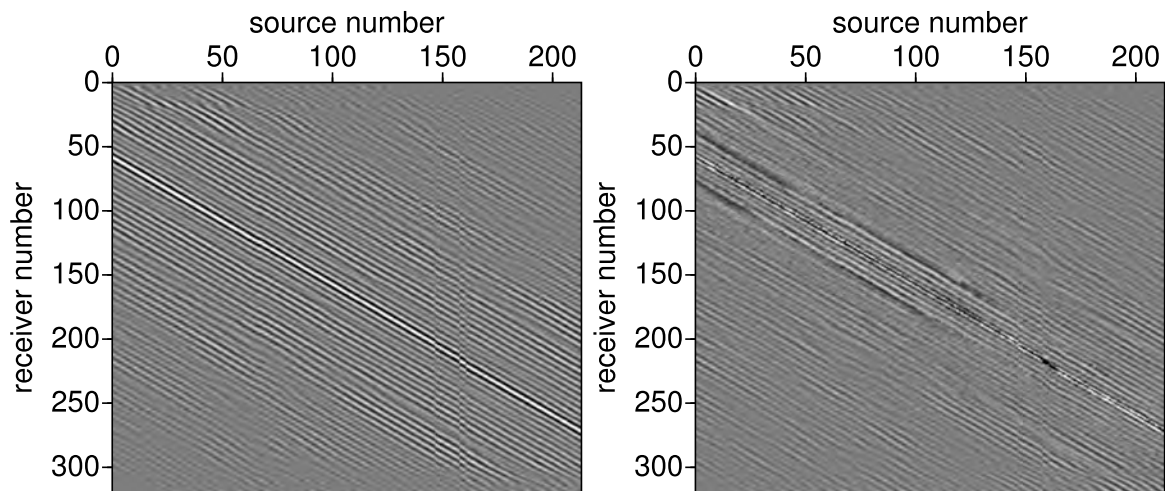


FIG. 5.21 – Partie réelle des données en fréquence à 7 Hz dans l'espace source/récepteur. Données calculées à l'issue de l'inversion de la ligne 21 avec la norme L_2 (gauche) et résidus associés (droite).

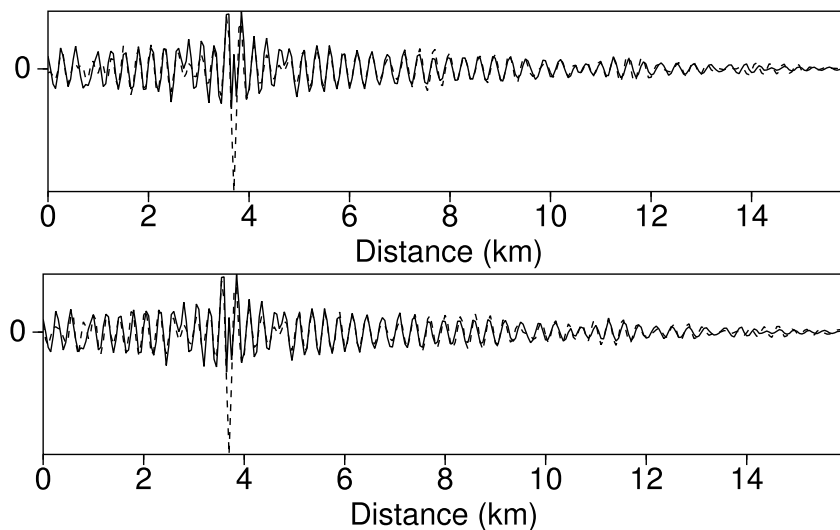


FIG. 5.22 – Accord des données en fréquence pour la ligne 21. Partie réelle des données à 7 Hz de la source 20 localisée à 3.6 km, pour les modèles de la norme L_1 (haut) et L_2 (bas).

nuage de gaz. Notons que les modèles obtenus sont moins perturbés que les modèles 3D dans la partie profonde, qui peut être lié à la différence entre couverture 2D et 3D, ainsi qu'à l'utilisation d'un seul facteur de régularisation γ dans notre cas 2D, qui limite le poids des arrivées tardives dans l'inversion. Pour la ligne 29, la norme L_1 a su exploiter les informations des résidus pour modifier le modèle tandis que la norme L_2 a faiblement convergé. Cette application illustre donc la complémentarité des deux formalismes et l'intérêt majeur de la norme L_1 comme alternative à la méthode classique. L'application élastique reste naturellement une perspective de travail primordiale pour exploiter ce jeu de données multicomposantes. Notons pour les deux jeux

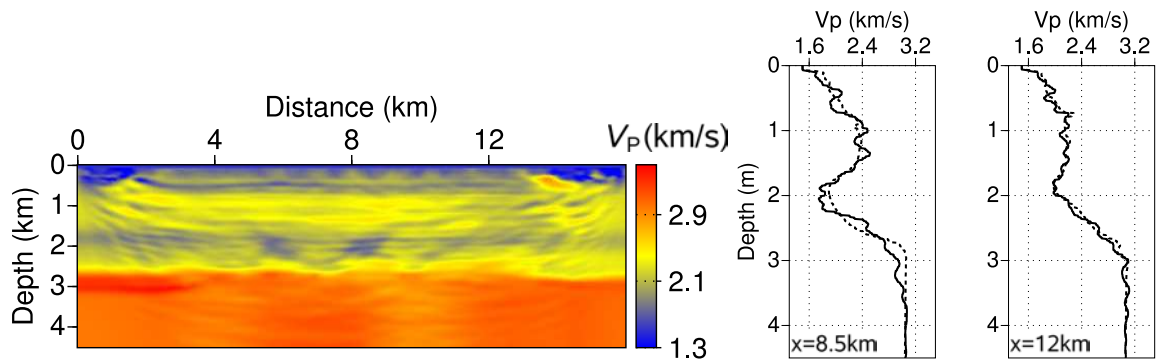


FIG. 5.23 – Modèles V_P imagés pour la ligne 29 avec la norme L_1 . Les coupes 1D verticales sont prises à 8.5 et 12 km. Les courbes discontinues et continues représentent respectivement le modèle initial et final de l'inversion.

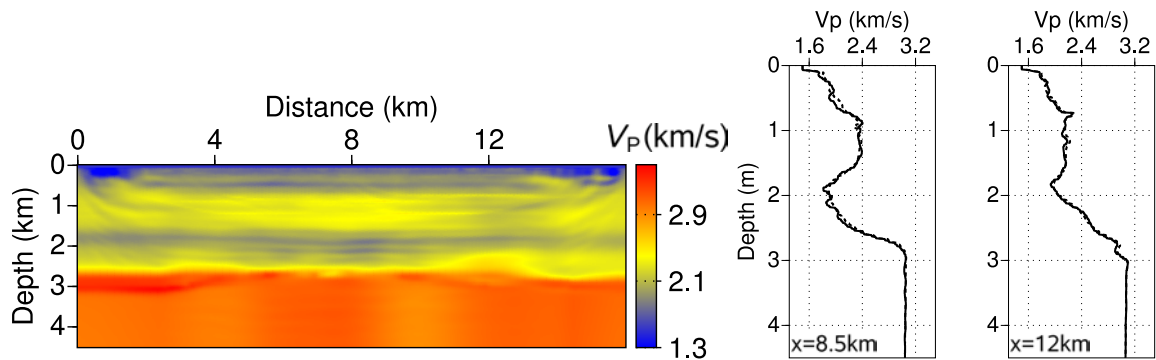


FIG. 5.24 – Modèles V_P imagés pour la ligne 29 avec la norme L_2 . Les coupes 1D verticales sont prises à 8.5 et 12 km. Les courbes discontinues et continues représentent respectivement le modèle initial et final de l'inversion.

de données que l'utilisation de groupes de fréquences et d'un nombre plus grand d'itérations permettraient certainement d'améliorer la qualité des images obtenues.

5.3 Conclusion

Ce chapitre applicatif s'est focalisé sur la sensibilité de l'inversion des formes d'ondes en domaine fréquentiel, de données décimées et bruitées. Les applications synthétiques ont mis en évidence les limites de l'approche classique basée sur la norme L_2 , lorsque du bruit ambiant ou de grandes erreurs contaminent les données. Au contraire, l'utilisation de la norme L_1 se révèle robuste pour ces applications, en étant très peu sensible aux bruits. Les critères combinant comportements L_2 et L_1 sont des alternatives possibles, mais la transition entre les deux comportements reste un paramètre de « réglage » en plus à déterminer... L'application acoustique sur les données réelles a montré des résultats très similaires pour les normes L_1 et L_2 pour une des applications. Pour l'autre, seule la norme L_1 a pu exploiter les résidus pour converger vers un modèle réduisant l'erreur sur l'ajustement des données, illustrant l'intérêt

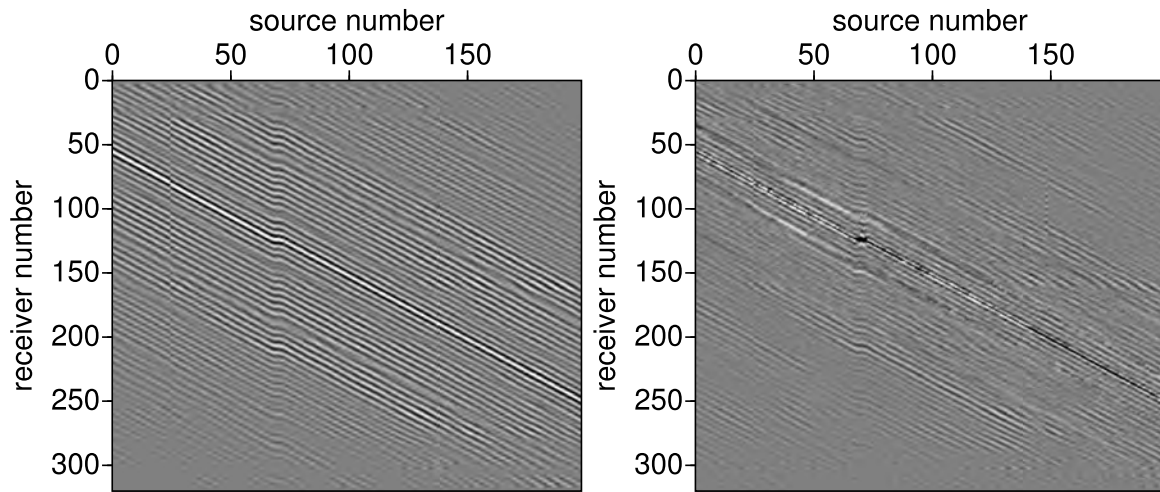


FIG. 5.25 – Partie réelle des données en fréquence à 7 Hz dans l'espace source/récepteur. Données calculées à l'issue de l'inversion de la ligne 29 avec la norme L_1 (gauche) et résidus associés (droite).

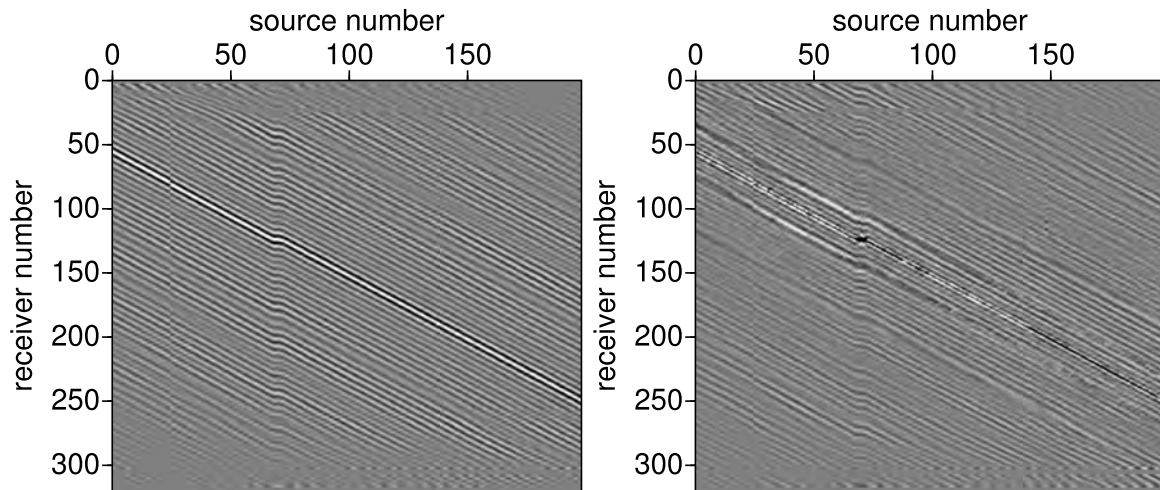


FIG. 5.26 – Partie réelle des données en fréquence à 7 Hz dans l'espace source/récepteur. Données calculées à l'issue de l'inversion de la ligne 29 avec la norme L_2 (gauche) et résidus associés (droite). Notons que les résidus sont plus forts que pour la norme L_1 (Figure 5.25)

majeur de ce formalisme.

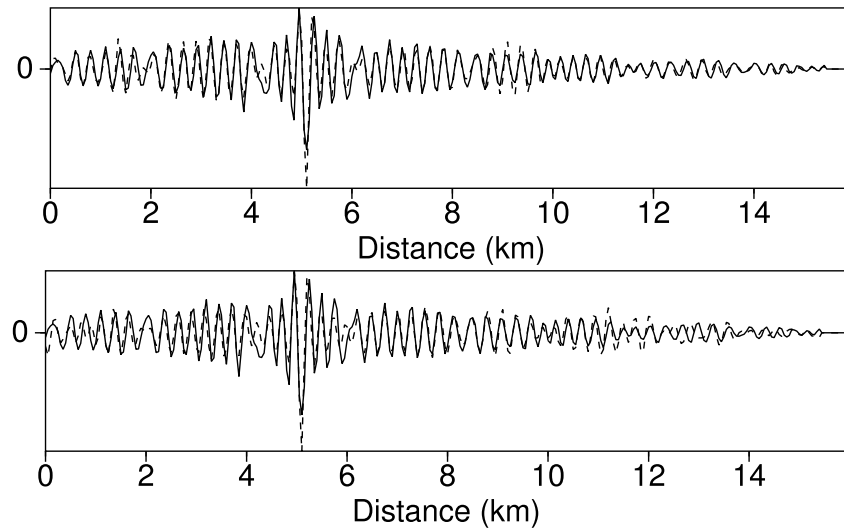


FIG. 5.27 – Accord des données en fréquence pour la ligne 29. Partie réelle des données à 7 Hz de la source 50 localisée à 5.1 km, pour les modèles de la norme L_1 (haut) et L_2 (bas).

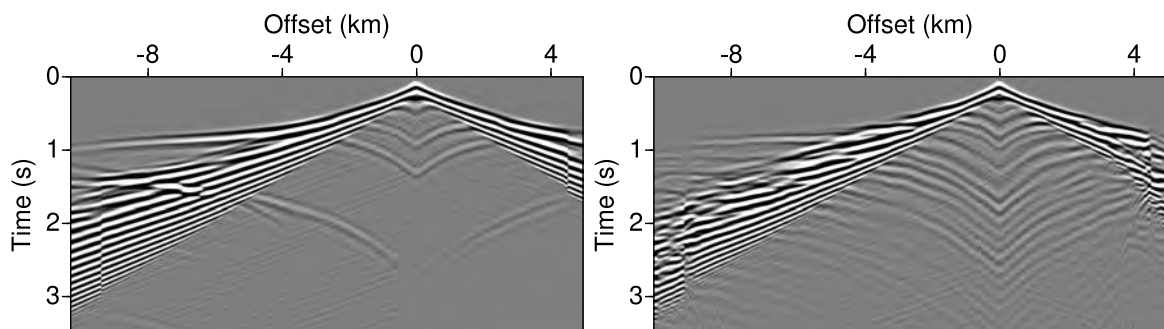


FIG. 5.28 – Exemple de données en domaine temporel calculées dans le modèle initial (gauche) et final (droite) de la ligne 29 (Modèle obtenu par norme L_1). Notons que les premières et dernières traces sont affectées par la reconstruction instable de la subsurface dans les zones non couvertes par les tirs.

Conclusion et perspectives

Conclusions générales de ce travail

J'ai présenté un travail de développements méthodologiques et d'applications de la méthode d'inversion des formes d'ondes pour l'imagerie des paramètres viscoélastiques dans des géométries à deux dimensions.

Problème direct

Dans le chapitre 1 j'ai introduit les équations du problème direct du processus d'inversion : les équations de l'élastodynamique linéaire. Les différentes alternatives pour résoudre ces équations ont été présentées, afin d'obtenir des solutions en domaine espace/fréquence, avant de développer une méthode d'éléments finis Galerkin Discontinus en domaine fréquentiel. Cette méthode permet de tirer bénéfice de l'utilisation de maillages triangulaires non-structurés, plus facilement manipulables que les maillages quadrangulaires. Par ailleurs, la flexibilité des choix d'ordres d'interpolation permet, selon les configurations d'applications, de trouver le meilleur compromis entre précision des solutions numériques et coût de calcul (temps et mémoire). Ainsi, l'ordre d'interpolation P_0 , utilisable uniquement en maillages équilatéraux réguliers, est très proche des méthodes de différences finies d'ordre 2 dans la formulation, la précision et le coût numérique. Le choix de cet ordre d'interpolation rend la méthode très efficace dans les milieux faiblement contrastés, sans topographie complexe ni couche d'eau, comme le serait une méthode de différences finies. Pour s'attaquer à des milieux plus complexes, on peut tirer bénéfice des ordres d'interpolation plus élevés (P_1 , P_2 et mixte) en utilisant des maillages non-structurés, adaptés localement aux propriétés physiques. Ce choix permet d'optimiser l'utilisation des ressources informatiques en adaptant au mieux la discrétisation tout en gardant une précision satisfaisante des solutions.

Problème inverse

Dans le chapitre 2, le formalisme du problème inverse est présenté. Le problème inverse étant fortement non-linéaire, mal posé et faisant intervenir un grand nombre de degrés de liberté, il est résolu de manière linéarisée. Cette linéarisation impose alors de partir d'un modèle de départ suffisamment précis pour ne pas converger dans un minimum secondaire. La détermination de ce modèle de départ n'a pas été analysée dans le cadre de ce travail mais constitue un point majeur qui contrôle le succès ou l'échec de la méthode. Ensuite, des algorithmes d'optimisation locale efficaces et économiques doivent être utilisés afin de converger rapidement vers le minimum

global. Enfin, le problème inverse doit être résolu hiérarchiquement afin de minimiser l'impact des non-linéarités du processus.

Algorithme

Le processus global d'inversion des formes d'ondes étant consommateur de calcul informatique, les implémentations des algorithmes des problèmes directs et inverses doivent être optimisées et tirer bénéfice de la puissance des calculateurs parallèles modernes. J'ai développé un algorithme massivement parallèle pour des architectures de calculateurs à mémoire distribuée. Le système linéaire creux de grande taille du problème direct est résolu grâce au solveur direct parallèle MUMPS. Les différentes étapes du processus d'inversion sont effectuées en parallèle selon une méthode de décomposition en domaines physiques. L'algorithme d'optimisation quasi-Newton L-BFGS est implémenté afin de tirer bénéfice de l'estimation du Hessien. En effet, ce dernier permet d'accélérer la convergence du problème inverse non-linéaire, d'améliorer la focalisation des images et de pouvoir facilement envisager la reconstruction de plusieurs classes de paramètres d'ordre de grandeurs différents, comparativement aux méthodes plus classiquement utilisées de gradient ou de gradient conjugué préconditionnés. Enfin, un double niveau de parallélisme, à travers des processus MPI (Message Passing Interface standard) et des threads à mémoire partagée, permet de tirer bénéfice des calculateurs modernes à mémoire distribuée basés sur des technologies multicœurs. Ainsi, les coûts de calcul (temps et mémoire) sont minimisés en fonction des applications et des ressources disponibles. L'implémentation de cet algorithme permet d'envisager des applications à des échelles très variées, de la proche surface aux échelles crustales ou lithosphériques.

Applications

En utilisant les outils développés pour l'inversion des formes d'ondes élastiques, plusieurs applications ont permis de mettre en évidence les problématiques spécifiques de l'imagerie élastique multiparamètres à partir de données multicomposantes, et de proposer des solutions.

L'application synthétique en milieu terrestre, nécessitant l'inversion combinée des ondes de volume et de surface pour l'imagerie simultanée de V_P et V_S , a permis d'illustrer différents points : la surface libre présente dans les simulations (et dans la réalité) et une source importante de non-linéarité dans l'inversion. En effet, elle introduit des ondes de surface très énergétiques et des ondes multiples non exploitables directement dans l'inversion linéarisée. L'inversion en domaine fréquentiel « classique » basée sur l'inversion séquentielle de fréquences discrètes, partant des basses fréquences vers les hautes fréquences, s'est révélée inefficace dans cette application trop non-linéaire. Un niveau de hiérarchie supplémentaire a donc été introduit, basé sur l'utilisation des fréquences complexes. Cet ultime recours en domaine fréquentiel permet de sélectionner, par rapport au temps de propagation, la quantité de données impliquée dans l'inversion. Un processus hiérarchique a été développé en commençant l'inversion avec peu de données, proches de la première arrivée, puis en augmentant progressivement la quantité de données afin de mieux contraindre l'inversion. Pour des propagations d'ondes complexes, l'inversion séquentielle fréquence par fréquence révéla que la quantité de données simultanément inversée ne permettait pas une reconstruction stable dans l'intégralité du modèle. L'utilisation de groupes de fréquences simultanément inversées, avec recouvrement entre groupes, procura la redondance d'information nécessaire à une meilleure convergence vers le minimum global.

Le choix des fréquences au sein d'un groupe reste un point critique : la fréquence maximale doit être suffisamment basse pour ne pas introduire de phénomènes de saut de phase introduit par la précision limitée du modèle initial ; le nombre de fréquences doit assurer une redondance suffisante pour contraindre l'inversion ; et le nombre de fréquence doit rester faible pour minimiser le coût de calcul.

L'application sur le modèle synthétique de Valhall a montré les problématiques différentes inhérentes aux environnements marins : lorsque les contrastes de vitesse sont faibles avec des vitesses d'ondes S très faibles dans la couche sédimentaire superficielle, les données enregistrées en fond de mer portent une très faible signature des ondes converties, rendant l'inversion de V_S fortement non-linéaire et mal contrainte. Une inversion hiérarchique doit donc être menée. L'utilisation de fréquences discrètes croissantes, couplée à l'utilisation des fréquences complexes pour sélectionner le volume de données inversé sont requises pour minimiser la non-linéarité du problème. Dans ce contexte où les différents paramètres ont des empreintes plus ou moins fortes dans les données, une hiérarchisation sur les données inversées et les paramètres reconstruits est également nécessaire. Dans un premier temps, la reconstruction du paramètre dominant V_P , en utilisant des données faiblement marquées par la signature des ondes S (hydrophone), permet de reconstruire un modèle de V_P et de minimiser significativement les résidus des ondes P. Dans un second temps, le paramètre secondaire V_S peut être reconstruit conjointement avec V_P , en utilisant les données de géophone, judicieusement pondérées avec l'offset, dans lesquelles la signature d'onde S est plus marquée.

Les applications sur données synthétiques bruitées ont illustré la problématique de la robustesse de l'inversion en domaine fréquentiel de données décimées. La redondance des données, largement utilisée dans les méthodes sismiques « classiques » et sacrifiée pour l'inversion des formes d'onde en fréquence, se montre en effet nécessaire lorsque les données sont affectées par du bruit ambiant. De plus, le formalisme classique basé sur la norme aux moindres carrés se révèle très sensible aux larges erreurs présentes dans les données. La méthode alternative basée sur la minimisation de la norme L_1 se montre au contraire très robuste pour l'inversion de données bruitées et décimées en fréquence, et très peu sensible aux larges erreurs. Cette norme a effet montré une grande robustesse pour imager le paramètre V_S dans le contexte marin du modèle synthétique de Valhall dont les ondes converties P-S ont des amplitudes faibles, rendant leur exploitation très sensible au rapport signal sur bruit. Cette norme devrait donc être sérieusement considérée comme une alternative aux méthodes classiques basées sur la norme L_2 . L'application sur les données réelles du champ pétrolier de Valhall a montré l'équivalence des résultats des normes L_1 et L_2 pour la ligne d'acquisition 21, tandis que seule la norme L_1 a pu exploiter les données de l'inversion de la ligne d'acquisition 29.

Perspectives

Problème direct et algorithme

La résolution du problème direct pourrait évoluer de diverses manières. Le développement de GD d'ordres plus élevés, théoriquement moins coûteux en nombre de degrés de liberté (Käser & Dumbser, 2006), en basant la formulation sur une hypothèse de milieu hétérogène au sein de chaque cellule, pourrait permettre de diminuer le coût numérique du problème direct. Ce constat fait par Käser & Dumbser (2006) pour une résolution temporelle, devra être vérifiée

dans le contexte de résolution fréquentielle avec solveur direct. Ce changement majeur, fait au profit du coût de résolution, nécessiterait par contre de profonds changements dans l'algorithme d'inversion.

Quel que soit l'ordre de GD utilisé, l'évolution vers une formulation parsimonieuse pourrait être envisagée pour diminuer le coût mémoire de la factorisation. L'algorithme proposé pour effectuer le passage à partir de la matrice du premier ordre devra être optimisé pour une implémentation parallèle si la formulation générique GD parsimonieux se révèle trop délicate à mettre en oeuvre.

L'extension vers des matériaux anisotropes VTI, HTI ou TTI pourra être envisagée afin de prendre en compte l'anisotropie naturelle présente dans certaines structures géologiques. Cette extension peut être implémentée sans surcoût de calcul dans l'outil de modélisation, en remplaçant les coefficients isotropes de Lamé par les coefficients élastiques des milieux anisotropes ($C_{11}, C_{13}, C_{33}, C_{55}$ en 2D).

Pour certaines applications, la prise en compte d'une physique plus complexe, prenant en compte les effets poro-élastiques, pourra être développée. Cette physique nécessite cependant de considérer un système d'équations aux dérivées partielles contenant plus de champs et donc plus coûteux à résoudre.

L'algorithme pourra être étendu pour d'autres types de sources. Seulement des sources artificielles ont été considérées dans ce travail de thèse. On pourrait prendre en compte des mécanismes et des rayonnements de sources plus complexes afin d'utiliser des données de sismique passive issues de petits séismes locaux. L'imagerie lithosphérique pourrait également être envisagée à partir d'ondes planes provenant de télé-séismes (Bostock *et al.*, 2001).

L'extension du solveur direct MUMPS avec une phase d'analyse parallèle, seul chaînon séquentiel consommateur de temps de calcul manquant actuellement dans le processus d'imagerie, permettra d'envisager des applications de plus grandes tailles, en équilibrant les charges des différents processeurs.

Concernant l'extension directe à trois dimensions, l'approche fréquentielle par solveur direct ou hybride apparaît inenvisageable à cause du coût mémoire induit par la troisième dimension (Etienne *et al.*, 2008). Une approche GD temporelle se révèle donc plus réaliste pour calculer les champs en fréquence (Etienne *et al.*, 2009). Une limitation réside tout de même dans la non-convergence du schéma P_0 , des maillages réguliers et structurés n'existant pas en 3D. Les ordres P_1 ou P_2 sont donc les ordres d'interpolation possibles, engendrant un surcoût notable par rapport aux DF pour des milieux peu contrastés. L'alternative DF n'est donc pas exclue pour certaines applications 3D.

Selon les potentialités et limites des méthodes d'éléments spectraux triangulaires et tétraédriques actuellement en développement, ces méthodes émergentes pourront être investiguées pour quantifier leurs précisions et performances, lorsqu'elles auront atteint un niveau de maturité suffisant.

Enfin, on peut espérer que les travaux actuels et futurs sur les préconditionneurs permettront d'envisager la résolution des équations élastodynamiques avec des solveurs itératifs ou hybrides performants, rivalisant ou dépassant les performances des approches temporelles. Ces différentes approches, contrairement aux méthodes directes, sont en effet plus adéquates pour envisager des applications 3D réalistes massivement parallèles à l'aube de l'aire du *petascale computing*.

Problème inverse et algorithme

Les contraintes liées au coût du calcul du problème direct et aux moyens de calcul actuels et disponibles dans les prochaines années, imposent pour le moment de ne considérer que des approches locales du problème inverse. Il est donc clair que les perspectives associées à ce travail doivent se focaliser sur les méthodes de construction de modèles de départ pour différents paramètres. Le modèle initial étant essentiel à la convergence du processus, c'est un point majeur qui va conditionner l'utilisation massive ou non de la méthode d'inversion des formes d'ondes dans l'industrie par exemple. L'autre point, qui découle de la linéarisation, est la robustesse de l'approche à hautes fréquences, le problème étant moins linéaire. La limite haute fréquence de validité de l'approche linéarisée reste un point à déterminer dans le futur.

On pourra envisager d'étendre l'outil d'inversion pour imager les paramètres anisotropes, si ils sont pris en compte dans le problème direct, afin de paramétrer plus fidèlement les milieux dans certaines structures géologiques.

Les méthodes d'optimisation et les processus de hiérarchisation développés et utilisés en 2D n'ont pas de frein majeur pour être étendus en 3D. Notons tout de même que des approches visant à minimiser le nombre de simulations de problèmes directs en sommant les sources (Ben-Hadj-Ali *et al.*, 2009a,b; Herrmann *et al.*, 2009; Krebs *et al.*, 2009) devront certainement être utilisées pour rendre l'inversion 3D (élastiques) envisageable dans les prochaines années.

L'étude de sensibilité de l'inversion en fréquence de données bruitées décimées a montré certaines limitations et pose le problème du choix du domaine fréquentiel ou temporel pour les applications 3D. En effet, si les données fréquentielles sont calculées initialement en domaine temporel pour la performance, est-t-il raisonnable de décimer une partie des champs temporels que l'on a calculé? Cependant, un nombre limité de fréquences peut sûrement être suffisant pour contraindre l'inversion de données réelles bruitées, sans aller jusqu'à la limite de Nyquist, en utilisant la norme L_2 , ou l'alternative proposée par la norme L_1 . Seule l'étude des choix d'implémentation et les applications pourront déterminer la meilleure stratégie.

Des variantes du problème inverse pourraient être développés pour l'imagerie élastique, afin d'améliorer la robustesse du processus. En effet, la méthode « classique » repose sur les résidus et la norme de l'erreur dans le domaine des données. Or les méthodes de migration se sont intéressées depuis longtemps à la cohérence dans le domaine des images migrées. La méthode dite *Shot-Based Differential Semblance Optimization* fut proposée pour combiner les deux points de vue et vise à la fois à minimiser l'erreur des données et maximiser la semblance des images migrées (Symes & Carazzone, 1991; Plessix, 2006).

Enfin, on pourrait imaginer, à plus ou moins long terme, d'envisager l'inversion de manière totalement non-linéaire par des méthodes d'exploration globale ou semi-globale, afin de s'affranchir des limitations de l'approche linéarisée. Ces méthodes pourraient être couplées avec des méthodes de sommation de sources pour réduire de nombre de simulations. Ces méthodes pourraient être investiguées avec peu de paramètres par exemple, pour la construction de modèles de départ. Enfin, seules ces méthodes pourront, à mon sens, réellement exploiter les formes d'ondes pour générer des images quantitatives très haute résolution en exploitant les hautes fréquences des données.

Place de la méthode

Le rôle et la place de la méthode d'inversion des formes d'onde dans la chaîne de traitement des données sismique de diverses échelles seront également des points majeurs à déterminer dans le futur. En effet si elle n'est pas linéarisée, la méthode permet théoriquement d'être autosuffisante pour construire des images quantitatives très haute résolution du milieu en partant d'aucun *a priori*. En pratique, l'approche linéarisée impose de connaître un modèle de départ suffisamment précis pour traiter les basses fréquences, tandis que les hautes fréquences sont fortement non-linéaires, car beaucoup plus sensibles aux effets de diffractions multiples non pris en compte. Les limites du contenu fréquentiel utilisable seront à déterminer précisément pour utiliser la méthode au mieux et la coupler avec d'autres approches pour imager efficacement notre sous-sol.

Applications

Enfin, l'application sur des données synthétiques et surtout réelles sera certainement le point majeur des perspectives directes de ce travail. Les méthodes, outils et procédures élaborés et validés sur des exemples synthétiques réalistes dans le cadre de ce travail permettront d'aborder l'imagerie des paramètres élastiques de données réelles à diverses échelles.

Les outils développés ont ainsi d'ores et déjà été appliqués par Romdhane *et al.* (2009) sur des données synthétiques pour l'exploration de la très proche surface de milieux à topographie très complexe. Des applications sur données synthétiques et réelles, acquises sur un banc de mesure en laboratoire, ont également été menées par Bretaudeau *et al.* (2009) pour l'imagerie de cavité vide en utilisant de manière combinée ondes de volume et ondes de surface. Des applications sur des données de sismiques actives pour l'exploration pétrolière et crustale pourront être envisagées, avec les données multicomposantes de la zone de Valhall et de la fosse de Nakaï (Japon), qui a déjà fait l'objet d'une application d'inversion acoustique (Operto *et al.*, 2006). Des applications de monitoring de réservoirs pourraient également tirer bénéfice des informations des paramètres élastiques. L'imagerie des failles et fractures pourra enfin être envisagée.

Les extensions des outils actuels permettront d'envisager des applications prenant en compte des effets d'anisotropie, des effets poro-élastiques, des applications à l'échelle lithosphérique à partir de données télésismiques, à l'échelles régionale à partir de séismes locaux ou à l'échelle globale du globe.

Bibliographie

- Akcelik, V. (2002). *Multiscale Newton-Krylov methods for inverse acoustic wave propagation*. Thèse de doctorat, Carnegie Mellon University, Pittsburgh, Pennsylvania.
- Aki, K. & Richards, P. G. (2002). *Quantitative seismology, theory and methods, second edition*. University Science Books, Sausalito, California.
- Alerini, M., Bégat, S. L., Lambaré, G. & Baina, R. (2002). 2d pp- and ps-stereotomography for a multicomponent dataset. *SEG Technical Program Expanded Abstracts*, 21(1):838–841.
- Amestoy, P., Davis, T. A. & Duff, I. S. (1996). An approximate minimum degree ordering algorithm. *SIAM J. Matrix. Anal. and Applics.*, 17:886–905.
- Amestoy, P. R., Duff, I. S. & L’Excellent, J. Y. (2000). Multifrontal parallel distributed symmetric and unsymmetric solvers. *Comput. Methods in Appl. Mech. Eng.*, 184:501–520.
- Amestoy, P. R., Guermouche, A., L’Excellent, J. Y. & Pralet, S. (2006). Hybrid scheduling for the parallel solution of linear systems. *Parallel Computing*, 32:136–156.
- Aminzadeh, F., Brac, J. & Kunz, T. (1997). *3-D Salt and Overthrust models*. Society of Exploration Geophysicists.
- Askan, A., Akcelik, V., Bielak, J. & Ghattas, O. (2007). Full waveform inversion for seismic velocity and anelastic losses in heterogeneous structures. *Bulletin of the Seismological Society of America*, 97(6):1990–2008.
- Barnes, C. & Charara, M. (2008). Full-waveform inversion results when using acoustic approximation instead of elastic medium. *SEG Technical Program Expanded Abstracts*, 27(1):1895–1899.
- Becker, A. B., Carey, G. F. & Oden, J. T. (2001). *Finite Element, An Introduction, Volume 1*. Prentice-Hall.
- Bednar, J. B., Shin, C. & Pyun, S. (2007). Comparison of waveform inversion, part 2 : phase approach. *Geophysical Prospecting*, 55(4):465–475.
- Ben-Hadj-Ali, H., Operto, S. & Virieux, J. (2008). Velocity model building by 3D frequency-domain, full-waveform inversion of wide-aperture seismic data. *Geophysics*, 73(5):VE101–VE117.
- Ben-Hadj-Ali, H., Operto, S. & Virieux, J. (2009a). Efficient 3D frequency-domain full waveform inversion (FWI) with phase encoding. *In Expanded Abstracts*, page 5812. European Association of Geoscientists & Engineers.

BIBLIOGRAPHIE

- Ben-Hadj-Ali, H., Operto, S. & Virieux, J. (2009b). Three-dimensional frequency-domain full waveform inversion with phase encoding. *In 79th Annual SEG Conference & Exhibition, Houston*. Society of Exploration Geophysicists.
- BenJemaa, M. (2007). *Etude et simulation numérique de la rupture dynamique des séismes par des méthodes d'éléments finis discontinus*. Thèse de doctorat, Université Nice-Sophia Antipolis.
- BenJemaa, M., Glinsky-Olivier, N., Cruz-Atienza, V. M. & Virieux, J. (2009). 3D Dynamic rupture simulations by a finite volume method. *Geophysical Journal International*, 178:541–560.
- BenJemaa, M., Glinsky-Olivier, N., Cruz-Atienza, V. M., Virieux, J. & Piperno, S. (2007). Dynamic non-planar crack rupture by a finite volume method. *Geophysical Journal International*, 171:271–285.
- Berenger, J.-P. (1994). A perfectly matched layer for absorption of electromagnetic waves. *Journal of Computational Physics*, 114:185–200.
- Bielak, J., Loukakis, K., Hisada, Y. & Yoshimura, C. (2003). Domain Reduction Method for Three-Dimensional Earthquake Modeling in Localized Regions, Part I : Theory. *Bulletin of the Seismological Society of America*, 93(2):817–824.
- Billette, F. & Lambaré, G. (1998). Velocity macro-model estimation from seismic reflection data by stereotomography. *Geophys. J. Int.*, 135(2):671–680.
- Billette, F., Le Bégat, S., Podvin, P. & Lambaré, G. (2003). Practical aspects and applications of 2D stereotomography. *Geophysics*, 68:1008–1021.
- Bleibinhaus, F., Hole, J. A., Ryberg, T. & Fuis, G. S. (2007). Structure of the California Coast Ranges and San Andreas Fault at SAFOD from seismic waveform inversion and reflection imaging . *Journal of Geophysical Research*, 112(B06315):doi :10.1029/2006JB004611.
- Bleinstein, N. (1986). Two-and-one-half dimensional in-plane wave-propagation. *Geophysical Prospecting*, 34:686–703.
- Bohlen, T. & Saenger, E. H. (2006). Accuracy of heterogeneous staggered-grid finite-difference modeling of Rayleigh waves. *Geophysics*, 71:109–115.
- Bostock, M. G., Rondenay, S. & Shragge, J. (2001). Multiparameter two-dimensional inversion of scattered teleseismic body waves 1. theory for oblique incidence. *Journal of Geophysical Research*, 106(12):30771–30782.
- Bouchon, M., Campillo, M. & Gaffet, S. (1989). A boundary integral equation - discrete wavenumber representation method to study wave propagation in multilayered media having irregular interfaces. *Geophysics*, 54:1134–1140.
- Brenders, A., Pratt, R. & Charles, S. (2009). Waveform tomography of 2-d seismic data in the canadian foothills - data preconditioning by exponential time-damping. *In Expanded Abstracts*, page U041. EAGE.

- Brenders, A. J. & Pratt, R. G. (2007a). Efficient waveform tomography for lithospheric imaging : implications for realistic 2D acquisition geometries and low frequency data. *Geophysical Journal International*, 168:152–170.
- Brenders, A. J. & Pratt, R. G. (2007b). Full waveform tomography for lithospheric imaging : results from a blind test in a realistic crustal model. *Geophysical Journal International*, 168(1):133–151.
- Brenders, A. J. & Pratt, R. G. (2007c). Waveform tomography of marine seismic data : what can limited offset offer ? In *Extended Abstracts*, pages 3024–3029.
- Bretaudeau, F., Leparoux, D., Brossier, R. & Abraham, O. (2009). Small scale modeling, a tool to assess subsurface imaging methods : Application to seismic full waveform inversion. In *Expanded Abstracts*, page P041. EAGE.
- Brossier, R., Operto, S. & Virieux, J. (2009a). Seismic imaging of complex onshore structures by two-dimensional elastic frequency-domain full-waveform inversion. *Geophysics*, 74(6):in press.
- Brossier, R., Operto, S. & Virieux, J. (2009b). Two-dimensional seismic imaging of the valhall model from synthetic OBC data by frequency-domain elastic full-waveform inversion. *SEG Technical Program Expanded Abstracts*, 28(1):2293–2297.
- Brossier, R., Virieux, J. & Operto, S. (2008). Parsimonious finite-volume frequency-domain method for 2-D P-SV-wave modelling. *Geophysical Journal International*, 175(2):541–559.
- Bube, K. P. & Langan, R. T. (1997). Hybrid l_1/l_2 minimization with applications to tomography. *Geophysics*, 62(4):1183–1195.
- Bube, K. P. & Nemeth, T. (2007). Fast line searches for the robust solution of linear systems in the hybrid l_1/l_2 and huber norms. *Geophysics*, 72(2):A13–A17.
- Buland, A. & Omre, H. (2003). Bayesian linearized AVO inversion. *Geophysics*, 68:185–198.
- Bunks, C., abd S. Zaleski, F. M. S. & Chavent, G. (1995). Multiscale seismic waveform inversion. *Geophysics*, 60(5):1457–1473.
- Carcione, J. M., Kosloff, D. & Kosloff, R. (1988). Wave-propagation simulation in an elastic anisotropic (transversely isotropic) solid. *Q. J. Mech. appl. Math.*, 41(3):319–345.
- Casadei, F., Gabellini, E., Fotia, G., Maggio, F. & Quarteroni, A. (2002). A mortar spectral/finite element method for complex 2D and 3D elastodynamic problems. *Comput. Method Appl. Mech. Engrg*, 191:5119–5148.
- Casarotti, E., Stupazzini, M., Lee, S.-J., Komatitsch, D., Piersanti, A. & Tromp, J. (2007). CUBIT and seismic wave propagation based upon the spectral-element method : An advanced unstructured mesher for complex 3D geological media. pages 579–597. Springer.
- Cerjan, C. D., Kosloff, D. & Reshef, M. (1985). A non-reflecting boundary condition for direct acoustic and elastic wave equations. *Geophysics*, 50:705–708.
- Červený, V., Molotkov, I. A. & Pšenčík, I. (1977). *Ray theory in Seismology*. Charles University Press, Praha.

- Chaljub, E., Capdeville, Y. & Vilotte, J. (2003). Solving elastodynamics in a fluid-solid heterogeneous sphere : a parallel spectral element approximation on non-conforming grids. *J. Compu. Phys.*, 187:457–491.
- Chapman, C. (2004). *Fundamentals of seismic waves propagation*. Cambridge University Press, Cambridge, England.
- Chew, W. C. & Liu, Q. H. (1996). Perfectly matched layers for elastodynamics : a new absorbing boundary condition. *J. Compu. Acous.*, 4:341–359.
- Choi, Y., Min, D. & Shin, C. (2008). Two-dimensional waveform inversion of multi-component data in acoustic-elastic coupled media. *Geophysical Prospecting*, 56(6):863–881.
- Choi, Y. & Shin, C. (2008). Frequency-Domain Elastic Full Waveform Inversion Using the New Pseudo-Hessian Matrix : Experience Of Elastic Marmousi 2 Synthetic Data. *Bulletin of the Seismological Society of America*, 98(5):2402–2415.
- Chu, C. & Stoffa, P. L. (2008). A pseudospectral-finite difference hybrid approach for large-scale seismic modeling and rtm on parallel computers. *SEG Technical Program Expanded Abstracts*, 27(1):2087–2091.
- Claerbout, J. F. (1976). *Fundamentals of Geophysical Data Processing*. McGraw-Hill Book Co.
- Claerbout, J. F. (1985). *Imaging the Earth's interior*. Blackwell Scientific Publications.
- Clayton, R. W. & Engquist, B. (1980). Absorbing boundary conditions for wave equation migration. *Geophysics*, 45(5):895–904.
- Cockburn, B., Karniadakis, G. & Shu, C. W. (2000). *Discontinuous Galerkin Method : Theory, Computation and Application*. Springer - Lecture Notes in Computational Science Engineering.
- Collino, F. & Tsogka, C. (2001). Application of the perfectly matched absorbing layer model to the linear elastodynamic problem in anisotropic heterogeneous media. *Geophysics*, 66:294–307.
- Cruse, E., Pica, A., Noble, M., McDonald, J. & Tarantola, A. (1990). Robust elastic non-linear waveform inversion : application to real data. *Geophysics*, 55:527–538.
- Cruse, E., Wideman, C., Noble, M. & Tarantola, A. (1992). Nonlinear elastic inversion of land seismic reflection data. *Journal of Geophysical Research*, 97:4685–4705.
- de la Puente, J., Käser, M., Dumbser, M. & Igel, H. (2007). An Arbitrary High Order Discontinuous Galerkin Method for Elastic Waves on Unstructured Meshes IV : Anisotropy. *Geophysical Journal International*, 169(3):1210–1228.
- Delcourte, S., Fezoui, L. & Glinsky-Olivier, N. (2009). A high-order discontinuous galerkin method for the seismic wave propagation. *ESAIM : Proc.*, 27:70–89.
- Djikpéssé, H. A. & Tarantola, A. (1999). Multiparameter l_1 norm waveform fitting : Interpretation of gulf of mexico reflection seismograms. *Geophysics*, 64(4):1023–1035.

- Dolean, V., Fol, H., Lanteri, S. & Piperno, S. (2006). Méthode de type galerkin discontinu pour la résolution numérique des équations de Maxwell en régime fréquentiel. Rapport de Recherche 5904, INRIA Sophia-Antipolis.
- Drossaert, F. H. & Giannopoulos, A. (2007). A nonsplit complex frequency-shifted PML based on recursive integration for FDTD modeling of elastic waves. *Geophysics*, 72(2):T9–T17.
- Dubiner, M. (1991). Spectral methods on triangles and other domains. *J. Sci. Comput.*, 6(4):345–390.
- Duff, I. S. & Reid, J. K. (1983). The multifrontal solution of indefinite sparse symmetric linear systems. *ACM Transactions on Mathematical Software*, 9:302–325.
- Dumbser, M. & Käser, M. (2006). An Arbitrary High Order Discontinuous Galerkin Method for Elastic Waves on Unstructured Meshes II : The Three-Dimensional Isotropic Case. *Geophysical Journal International*, 167(1):319–336.
- Dumbser, M., Käser, M. & Toro, E. (2007). An Arbitrary High Order Discontinuous Galerkin Method for Elastic Waves on Unstructured Meshes V : Local Time Stepping and p-Adaptivity. *Geophysical Journal International*, 171(2):695–717.
- Effelsen, K. (2009). A comparison of phase inversion and traveltime tomography for processing of near-surface refraction traveltimes. *Geophysics*, in-press.
- Epanomeritakis, I., Akçelik, V., Ghattas, O. & Bielak, J. (2008). A Newton-CG method for large-scale three-dimensional elastic full waveform seismic inversion. *Inverse Problems*, 24:1–26.
- Erlangga, Y. A. (2005). *A robust and efficient iterative method for the numerical solution of the Helmholtz equation*. Thèse de doctorat, Delft University of Technology.
- Erlangga, Y. A. & Herrmann, F. J. (2008). An iterative multilevel method for computing wavefields in frequency-domain seismic inversion. *SEG Technical Program Expanded Abstracts*, 27(1):1956–1960.
- Etienne, V., Brossier, R., Operto, S. & Virieux, J. (2008). A 3D Parsimonious Finite-Volume Frequency-Domain Method for Elastic Wave Modelling. *In Expanded Abstracts*. EAGE.
- Etienne, V., Virieux, J., Glinsky, N. & Operto, S. (2009). Seismic modelling with discontinuous galerkin finite-element method - application to large scale 3D elastic media. *In Expanded Abstracts*, page P131. EAGE.
- Faccioli, E. F., Paolucci, R. & Quarteroni, A. (1997). 2D and 3D elastic wave propagation by a pseudo-spectral domain decomposition method. *J. Seismol.*, 1:237–251.
- Fahs, H. (2007). Numerical evaluation of a non-conforming discontinuous Galerkin method on triangular meshes for solving the time-domain Maxwell equations on non-conforming meshes. Rapport de Recherche 6311, INRIA Sophia-Antipolis.
- Fahs, H. & Lanteri, S. (2008). Convergence and stability of a high-order leap-frop based discontinuous Galerkin method for the Maxwell equations on non-conforming meshes. Rapport de Recherche 6699, INRIA Sophia-Antipolis.

- Fink, M. (1993). Time-reversal mirrors. *Journal of Physics D : Applied Physics*, 26(9):1333–1350.
- Forgues, E. (1996). *Inversion linearisée multi-paramètres via la théorie des rais*. Thèse de doctorat, Institut Français du Pétrole - University Paris VII.
- Forgues, E. & Lambaré, G. (1997). Parameterization study for acoustic and elastic ray+born inversion. *Journal of Seismic Exploration*, 6:253–278.
- Fornberg, B. (1996). *A Practical Guide to Pseudospectral Methods*. Cambridge Monographs on Applied and Computational Mathematics (No. 1).
- Frigo, M. & Johnson, S. G. (2005). The design and implementation of FFTW3. *Proceedings of the IEEE*, 93(2):216–231. Special issue on “Program Generation, Optimization, and Platform Adaptation”.
- Futterman, W. (1962). Dispersive body waves. *Journal of Geophysics Research*, 67:5279–5291.
- Gao, F., Levander, A. R., Pratt, R. G., Zelt, C. A. & Fradelizio, G. L. (2006). Waveform tomography at a groundwater contamination site : Vsp-surface data set. *Geophysics*, 71(1): H1–H11.
- Garvin, W. W. (1956). Exact transient solution of the buried line source problem. *Proc. Roy. Soc. London*, 234:528–541.
- Gauthier, O., Virieux, J. & Tarantola, A. (1986). Two-dimensional nonlinear inversion of seismic waveforms : numerical results. *Geophysics*, 51(7):1387–1403.
- Gelis, C., Virieux, J. & Grandjean, G. (2007). 2D elastic waveform inversion using Born and Rytov approximations in the frequency domain. *Geophysical Journal International*, 168:605–633.
- George, A. & Liu, J. W. H. (1989). The evolution of the minimum degree algorithm. *SIAM review*, 31:1–19.
- Guitton, A. & Symes, W. W. (2003). Robust inversion of seismic data using the Huber norm. *Geophysics*, 68(4):1310–1319.
- Ha, T., Chung, W. & Shin, C. (2009). Waveform inversion using a back-propagation algorithm and a Huber function norm. *Geophysics*, 74(3):R15–R24.
- Hamming, R. W. (1973). *Numerical Methods for Scientists and Engineers*. McGraw-Hill, New York, 2 édition.
- Herrmann, F. J., Erlangga, Y. A. & Lin, T. T. Y. (2009). Compressive simultaneous full-waveform simulation. *Geophysics*, 74(4):A35–A40.
- Hestholm, S. & Ruud, B. (2002). 3D free-boundary conditions for coordinate-transform finite-difference seismic modelling. *Geophysical Prospecting*, 50:463–474.
- Hicks, G. J. & Pratt, R. G. (2001). Reflection waveform inversion using local descent methods : estimating attenuation and velocity over a gas-sand deposit. *Geophysics*, 66(2):598–612.

- Hole, J. A. (1992). Nonlinear high-resolution three-dimensional seismic travel time tomography. *Journal of Geophysical Research*, 97:6553–6562.
- Hu, W., Abubakar, A. & Habashy, T. (2009a). Preconditioned non-linear conjugate gradient method for seismic full-waveform inversion. *In Expanded Abstracts*. EAGE.
- Hu, W., Abubakar, A. & Habashy, T. M. (2009b). Simultaneous multifrequency inversion of full-waveform seismic data. *Geophysics*, 74(2):R1–R14.
- Huber, P. J. (1973). Robust regression : Asymptotics, conjectures, and Monte Carlo. *Ann. Statist.*, 1(5):799–821.
- Hustedt, B., Operto, S. & Virieux, J. (2004). Mixed-grid and staggered-grid finite difference methods for frequency domain acoustic wave modelling. *Geophysical Journal International*, 157:1269–1296.
- Jin, S., Cambois, G. & Vuillermoz, C. (2000). Shear-wave velocity and density estimation from PS-wave AVO analysis : Application to an OBS dataset from the North Sea. *Geophysics*, 65(5):1446–1454.
- Jin, S., Madariaga, R., Virieux, J. & Lambaré, G. (1992). Two-dimensional asymptotic iterative elastic inversion. *Geophysical Journal International*, 108:575–588.
- Jo, C. H., Shin, C. & Suh, J. H. (1996). An optimal 9-point, finite-difference, frequency-space 2D scalar extrapolator. *Geophysics*, 61:529–537.
- Kamei, R. & Pratt, R. G. (2008). Waveform tomography strategies for imaging attenuation structure for cross-hole data. *In Extended Abstracts, Roma*, page F019.
- Karypis, G. & Kumar, V. (1999). A fast and high quality multilevel scheme for partitioning irregular graphs. *SIAM Journal on Scientific Computing*, 20(1):359 – 392.
- Käser, M. & Dumbser, M. (2006). An Arbitrary High Order Discontinuous Galerkin Method for Elastic Waves on Unstructured Meshes I : The Two-Dimensional Isotropic Case with External Source Terms. *Geophysical Journal International*, 166:855–877.
- Käser, M., Dumbser, M., de la Puente, J. & Igel, H. (2007). An Arbitrary High Order Discontinuous Galerkin Method for Elastic Waves on Unstructured Meshes III : Viscoelastic Attenuation. *Geophysical Journal International*, 168(1):224–242.
- Kim, Y., Cha, Y. H., Shin, C., Ko, S. & Seo, Y. (2008). Improved logarithmic waveform inversion considering the power-spectrum of the wavefield. *SEG Technical Program Expanded Abstracts*, 27(1):1966–1970.
- Kirkpatrick, S., Gelatt, C. D. & Vecchi, M. P. (1983). Optimization by simulated annealing. *Science*, 220:671–680.
- Koketsu, K., Fujiwara, H. & Ikegami, Y. (2004). Finite-element Simulation of Seismic Ground Motion with a Voxel Mesh. *Pure and Applied Geophysics*, 161:2183–2198.
- Kolsky, H. (1956). The propagation of stress pulses in viscoelastic solids. *Philosophical Magazine*, 1:693–710.

- Komatitsch, D. & Martin, R. (2007). An unsplit convolutional perfectly matched layer improved at grazing incidence for the seismic wave equation. *Geophysics*, 72(5):SM155–SM167.
- Komatitsch, D. & Vilotte, J. P. (1998). The spectral element method : an efficient tool to simulate the seismic response of 2D and 3D geological structures. *Bull. Seismol. Soc. Am.*, 88:368–392.
- Kommedal, J. H., Barkved, O. I. & Howe, D. J. (2004). Initial experience operating a permanent 4C seabed array for reservoir monitoring at Valhall. *SEG Technical Program Expanded Abstracts*, 23(1):2239–2242.
- Krebs, J., Anderson, J., Hinkley, D., Neelamani, R., Lee, S., Baumstein, A. & Lacasse, M. D. (2009). Fast full-wavefield seismic inversion using encoded sources. *Geophysics*.
- Lambaré, G. (2008). Stereotomography. *Geophysics*, 73(5):VE25–VE34.
- Lee, D., Cha, Y. H. & Shin, C. (2008). The direct-removal method of waveform inversion in the laplace inversion for deep-sea environments. *SEG Technical Program Expanded Abstracts*, 27(1):1981–1985.
- Levander, A. R. (1988). Fourth-order finite-difference P-SV seismograms. *Geophysics*, 53(11):1425–1436.
- LeVeque, R. J. (2002). *Finite Volume Methods for Hyperbolic Problems*. Cambridge Texts in Applied Mathematics.
- LeVeque, R. J. (2007). *Finite Difference Methods for Ordinary and Partial Differential Equations, Steady State and Time Dependent Problems*. SIAM.
- Lions, J. (1972). *Nonhomogeneous boundary value problems and applications*. Springer Verlag, Berlin.
- Liu, Q. & Tromp, J. (2006). Finite-frequency kernels based on adjoint methods. *Bulletin of the Seismological Society of America*, 96(6):2383–2397.
- Lombard, B., Piraux, J., Gelis, C. & Virieux, J. (2008). Free and smooth boundaries in 2-D finite-difference schemes for transient elastic waves. *Geophysical Journal International*, 172:252–261.
- Lopez, F. (2009). *Inversion des formes d’ondes électromagnétiques de données radar (GPR) multioffsets*. Thèse de doctorat, Université Denis Diderot Paris 7.
- Love, A. E. H. (1911). Theory of the propagation of seismic waves. *Some problemes of Geodynamics (in Geophysics reprint series (vol 24) : Classics of Elastic Wave Theory)*, pages 144–179.
- Luo, Y. & Schuster, G. T. (1990). Parsimonious staggered grid finite-differencing of the wave equation. *Geophysical Research Letters*, 17(2):155–158.
- Madariaga, R. (1976). Dynamics of an expanding circular fault. *Bull. Seismol. Soc. Am.*, 66:639–666.

- Malinowsky, M. & Operto, S. (2008). Quantitative imaging of the Permo-Mesozoic complex and its basement by frequency domain waveform tomography of wide-aperture seismic data from the Polish basin. *Geophysical Prospecting*, 56:805–825.
- Mallick, S. & Frazer, L. N. (1987). Practical aspects of reflectivity modeling. *Geophysics*, 52:1355–1364.
- Marfurt, K. (1984). Accuracy of finite-difference and finite-elements modeling of the scalar and elastic wave equation. *Geophysics*, 49:533–549.
- Mari, J.-L., F. Glangeaud & Coppens, F. (1997). *Traitement du signal pour géologues et géophysiciens*. Editions Technip.
- Martin, G. S., Wiley, R. & Marfurt, K. J. (2006). Marmousi2 : An elastic upgrade for marmousi. *The Leading Edge*, 25(2):156–166.
- McMechan, G. A. & Fuis, G. S. (1987). Ray equation migration of wide-angle reflections from southern alaska. *Journal of Geophysical Research*, 92(1):407–420.
- Menke, W. (1984). *Geophysical Data Analysis : Discrete Inverse Theory*. Academic Press, Inc., Orlando, USA.
- Mercerat, E., Vilotte, J. & Sanchez-Sesma, F. (2006). Triangular spectral element simulation of two-dimensional elastic wave propagation using unstructured triangular grids. *Journal of Scientific Computing*, 166:679–698.
- Metropolis, N. & Ulam, S. (1949). The Monte Carlo Method. *Journal of the American Statistical Association*, 44(247):335–341.
- Miller, D., Oristaglio, M. & Beylkin, G. (1987). A new slant on seismic imaging : Migration and integral geometry. *Geophysics*, 52(7):943–964.
- Min, D. J. & Shin, C. (2006). Refraction tomography using a waveform-inversion back-propagation technique. *Geophysics*, 71(3):R21–R30.
- Moczo, P., Bystricky, E., Carcione, J. M. & Bouchon, M. (1997). Hybrid modeling of P-SV seismic motion at inhomogeneous viscoelastic topographic structures. *Bull. Seismol. Soc. Am.*, 87:1305–1323.
- Moczo, P., Kristek, J. & Halada, L. (2004). *The Finite-Difference Method for Seismologists An Introduction*. Comenius University Bratislava.
- Monteiller, V., Got, J.-L., Virieux, J. & Okubo, P. (2005). An efficient algorithm for double-difference tomography and location in heterogeneous media, with an application to the Kilauea volcano. *Journal of Geophysical Research*, 110(B12306):doi :10.1029/2004JB00346.
- Montelli, R., Nolet, G., Dahlen, F. A., Masters, G., Engdahl, E. R. & Hung, S. H. (2004). Finite-frequency tomography reveals a variety of plumes in the mantle. *Science*, 303:338–343.
- Mora, P. (1988). Elastic wavefield inversion of reflection and transmission data. *Geophysics*, 53:750–759.

BIBLIOGRAPHIE

- Mora, P. (1989). Inversion = migration + tomography. *Geophysics*, 54(12):1575–1586.
- Mora, P. R. (1987). Nonlinear two-dimensional elastic inversion of multi-offset seismic data. *Geophysics*, 52:1211–1228.
- Mulder, W. & Plessix, R.-E. (2008). Exploring some issues in acoustic full waveform inversion. *Geophysical Prospecting*, 56(6):827–841.
- Mulder, W. A. & Hak, B. (2009). Simultaneous imaging of velocity and attenuation perturbations from seismic data is nearly impossible. In *Expanded Abstracts*, page S043.
- MUMPS-team (2007). *MUMPS - MULTifrontal Massively Parallel Solver users' guide - version 4.7 (April 2007)*. ENSEEIHT-ENS Lyon, <http://www.enseeiht.fr/apo/MUMPS/> or <http://graal.ens-lyon.fr/MUMPS>.
- Munns, J. W. (1985). The Valhall field : a geological overview. *Marine and Petroleum Geology*, 2:23–43,.
- Nazarian, S. & Stokoe, K. H. (1984). In situ shear wave velocities from spectral analysis of surface waves. *Proc. 8th Conf. on Earthquake Eng., San Francisco*, 3:31–38.
- Nihei, K. T. & Li, X. (2007). Frequency response modelling of seismic waves using finite difference time domain with phase sensitive detection (TD-PSD). *Geophysical Journal International*, 169:1069–1078.
- Nocedal, J. (1980). Updating Quasi-Newton Matrices With Limited Storage. *Mathematics of Computation*, 35(151):773–782.
- Nocedal, J. & Wright, S. J. (1999). *Numerical Optimization*. New York, US : Springer.
- Nolet, G. (1987). *Seismic tomography with applications in global seismology and exploration geophysics*. D. Reidel publishing Company.
- Operto, S., Virieux, J., Amestoy, P., L'Écellent, J.-Y., Giraud, L. & Ben-Hadj-Ali, H. (2007). 3D finite-difference frequency-domain modeling of visco-acoustic wave propagation using a massively parallel direct solver : A feasibility study. *Geophysics*, 72(5):SM195–SM211.
- Operto, S., Virieux, J., Dessa, J. X. & Pascal, G. (2006). Crustal imaging from multi-fold ocean bottom seismometers data by frequency-domain full-waveform tomography : application to the eastern Nankai trough. *Journal of Geophysical Research*, 111(B09306): doi :10.1029/2005JB003835.
- Pasquetti, R. & Rapetti, F. (2004). Spectral element methods on triangles and quadrilaterals : comparisons and applications. *J. Comput. Phys.*, 198(1):349–362.
- Pasquetti, R. & Rapetti, F. (2006). Spectral element methods on unstructured meshes : Comparisons and recent advances. *Journal of Scientific Computing*, 27:377–387.
- Plessix, R.-E. (2006). A review of the adjoint-state method for computing the gradient of a functional with geophysical applications. *Geophysical Journal International*, 167(2):495–503.
- Plessix, R. E. (2007). A helmholtz iterative solver for 3D seismic-imaging problems. *Geophysics*, 72(5):SM185–SM194.

- Plessix, R. E. (2009). 3-D frequency-domain full-waveform inversion with an iterative solver. *Geophysics*, in-press.
- Polak, E. & Ribière, G. (1969). Note sur la convergence de méthodes de directions conjuguées. *Revue Française d'Informatique et de Recherche Opérationnelle*, 16:35–43.
- Pratt, R. G. (1990). Inverse theory applied to multi-source cross-hole tomography. part II : elastic wave-equation method. *Geophysical Prospecting*, 38:311–330.
- Pratt, R. G. (1999). Seismic waveform inversion in the frequency domain, part I : theory and verification in a physic scale model. *Geophysical Journal International*, 64:888–901.
- Pratt, R. G., Shin, C. & Hicks, G. J. (1998). Gauss-Newton and full Newton methods in frequency-space seismic waveform inversion. *Geophysical Journal International*, 133:341–362.
- Pratt, R. G. & Shipp, R. M. (1999). Seismic waveform inversion in the frequency domain, part II : Fault delineation in sediments using crosshole data. *Geophysics*, 64:902–914.
- Pratt, R. G., Song, Z. M. & Warner, M. (1996). Two-dimensional velocity models from wide-angle seismic data by wavefield inversion. *Geophysical Journal International*, 124:323–340.
- Pratt, R. G. & Worthington, M. H. (1990). Inverse theory applied to multi-source cross-hole tomography. Part I : acoustic wave-equation method. *Geophysical Prospecting*, 38:287–310.
- Press, W. H., Teukolsky, S. A., Vetterling, W. T. & Flannery, B. P. (2007). *Numerical Recipes 3rd Edition : The Art of Scientific Computing*. Cambridge University Press, 3 édition.
- Prieux, V., Operto, S., Brossier, R. & Virieux, J. (2009). Application of acoustic full waveform inversion to the synthetic Valhall model. *SEG Technical Program Expanded Abstracts*, 28(1): 2268–2272.
- Pyun, S., Shin, C. & Bednar, J. B. (2007). Comparison of waveform inversion, part 3 : amplitude approach. *Geophysical Prospecting*, 55(4):477–485.
- Pyun, S., Shin, C., Lee, H. & Yang, D. (2008a). 3D elastic full waveform inversion in the laplace domain. *SEG Technical Program Expanded Abstracts*, 27(1):1976–1980.
- Pyun, S., Shin, C., Lee, H. & Yang, D. (2008b). 3D elastic full waveform inversion in the laplace domain. *SEG Technical Program Expanded Abstracts*, 27(1):1976–1980.
- Pyun, S., Shin, C. & Son, W. (2009). Frequency-domain waveform inversion using an L1-norm objective function. In *Expanded Abstracts*, page P005. EAGE.
- Ravaut, C., Operto, S., Improta, L., Virieux, J., Herrero, A. & dell'Aversana, P. (2004). Multi-scale imaging of complex structures from multi-fold wide-aperture seismic data by frequency-domain full-wavefield inversions : application to a thrust belt. *Geophysical Journal International*, 159:1032–1056.
- Rayleigh, L. (1887). On waves propagated along the plane surface of an elastic solid. *Proceedings of the London Mathematical Society (in Geophysics reprint series (vol 24) : Classics of Elastic Wave Theory)*, 17:4–11.

- Reed, W. & Hill, T. (1973). Triangular mesh methods for the neutron transport equation. Technical Report LA-UR-73-479, Los Alamos Scientific Laboratory.
- Remaki, M. (2000). A new finite volume scheme for solving Maxwell's system. *COMPEL*, 19(3):913–931.
- Ribodetti, A., Operto, S., Virieux, J., Lambaré, G., Valéro, H.-P. & Gibert, D. (2000). Asymptotic viscoacoustic diffraction tomography of ultrasonic laboratory data : a tool for rock properties analysis. *Geophysical Journal International*, 140:324–340.
- Ribodetti, A. & Virieux, J. (1996). Asymptotic theory for imaging the attenuation factors Q_p and Q_s . In *Inverse Problems of Wave Propagation and Diffraction, Proceedings, Aix-les-Bains, France 1996*, pages 334–353. Springer-Verlag.
- Riyanti, C. D., Erlangga, Y. A., Plessix, R. E., Mulder, W. A., Vuik, C. & Oosterlee, C. (2006). A new iterative solver for the time-harmonic wave equation. *Geophysics*, 71(E):57–63.
- Robertsson, J. O. A. (1996). A numerical free-surface condition for elastic/viscoelastic finite-difference modeling in the presence of topography. *Geophysics*, 61:1921–1934.
- Romdhane, A., Grandjean, G., Brossier, R., Operto, S. & Réjiba, F. (2009). Full waveform inversion of seismic data for 2d shallow structures imagery : Limitations and perspectives. In *Expanded Abstracts*, page P002. EAGE.
- Saad, Y. (2003). *Iterative methods for sparse linear systems*, volume second edition. SIAM, Philadelphia.
- Saenger, E. H. & Bohlen, T. (2004). Finite-difference modelling of viscoelastic and anisotropic wave propagation using the rotated staggered grid. *Geophysics*, 69:583–591.
- Saenger, E. H., Gold, N. & Shapiro, S. A. (2000). Modeling the propagation of elastic waves using a modified finite-difference grid. *Wave motion*, 31:77–92.
- Sambridge, M. S., Tarantola, A. & Kennett, B. L. (1991). An alternative strategy for non-linear inversion of seismic waveforms. *Geophysical Prospecting*, 39:723–736.
- Sears, T., Singh, S. & Barton, P. (2008). Elastic full waveform inversion of multi-component OBC seismic data. *Geophysical Prospecting*, 56(6):843–862.
- Sen, M. K. & Stoffa, P. L. (1995). *Global Optimization Methods in Geophysical Inversion*. Elsevier Science Publishing Co.
- Sheng, J., Leeds, A., Buddensiek, M. & Schuster, G. T. (2006). Early arrival waveform tomography on near-surface refraction data. *Geophysics*, 71(4):U47–U57.
- Shewchuk, J. R. (1996). Triangle : Engineering a 2D Quality Mesh Generator and Delaunay Triangulator. In Lin, M. C. & Manocha, D., éditeurs : *Applied Computational Geometry : Towards Geometric Engineering*, volume 1148 de *Lecture Notes in Computer Science*, pages 203–222. Springer-Verlag. From the First ACM Workshop on Applied Computational Geometry.
- Shi, Y., Zhao, W. & Cao, H. (2007). Nonlinear process control of wave-equation inversion and its application in the detection of gas. *Geophysics*, 72(1):R9–R18.

- Shin, C. & Cha, Y. H. (2008). Waveform inversion in the Laplace domain. *Geophysical Journal International*, 173(3):922–931.
- Shin, C. & Ha, W. (2008). A comparison between the behavior of objective functions for waveform inversion in the frequency and laplace domains. *Geophysics*, 73(5):VE119–VE133.
- Shin, C., Jang, S. & Min, D. J. (2001). Improved amplitude preservation for prestack depth migration by inverse scattering theory. *Geophysical Prospecting*, 49:592–606.
- Shin, C. & Min, D.-J. (2006). Waveform inversion using a logarithmic wavefield. *Geophysics*, 71(3):R31–R42.
- Shin, C., Min, D.-J., Marfurt, K. J., Lim, H. Y., Yang, D., Cha, Y., Ko, S., Yoon, K., Ha, T. & Hong, S. (2002). Traveltime and amplitude calculations using the damped wave solution. *Geophysics*, 67:1637–1647.
- Shin, C., Pyun, S. & Bednar, J. B. (2007). Comparison of waveform inversion, part 1 : conventional wavefield vs logarithmic wavefield. *Geophysical Prospecting*, 55(4):449–464.
- Shipp, R. M. & Singh, S. C. (2002). Two-dimensional full wavefield inversion of wide-aperture marine seismic streamer data. *Geophysical Journal International*, 151:325–344.
- Sirgue, L. (2003). *Inversion de la forme d'onde dans le domaine fréquentiel de données sismiques grand offset*. Thèse de doctorat, Université Paris 11, France - Queen's University, Canada.
- Sirgue, L., Barkved, O. I., Gestel, J. P. V., Askim, O. J. & Kommedal, J. H. . (2009). 3D waveform inversion on Vallhall wide-azimuth OBC. In *Expanded Abstracts*, page U038.
- Sirgue, L., Etgen, J. & Albertin, U. (2007). 3D full-waveform inversion : Wide- versus narrow-azimuth acquisitions. *SEG Technical Program Expanded Abstracts*, 26(1):1760–1764.
- Sirgue, L., Etgen, J. T. & Albertin, U. (2008). 3D Frequency Domain Waveform Inversion using Time Domain Finite Difference Methods. In *Proceedings 70th European Association of Geoscientists and Engineers, Conference and Exhibition, Roma, Italy*, page F022.
- Sirgue, L. & Pratt, R. G. (2004). Efficient waveform inversion and imaging : a strategy for selecting temporal frequencies. *Geophysics*, 69(1):231–248.
- Sourbier, F., Haidar, A., Giraud, L., Operto, S. & Virieux, J. (2008a). Frequency-domain full-waveform modeling using a hybrid direct-iterative solver based on a parallel domain decomposition method : A tool for 3D full-waveform inversion? *SEG Technical Program Expanded Abstracts*, 27(1):2147–2151.
- Sourbier, F., Haiddar, A., Giraud, L., Brossier, R., Operto, S. & Virieux, J. (2008b). Frequency-domain full-waveform modeling using a hybrid direct-iterative solver based on a parallel domain decomposition method. In *70th Annual EAGE Conference & Exhibition, Roma*. Eur. Ass. Expl.Geophys.
- Sourbier, F., Operto, S., Virieux, J., Amestoy, P. & L'Excellent, J.-Y. (2009a). Fwt2d : A massively parallel program for frequency-domain full-waveform tomography of wide-aperture seismic data—part 1 : Algorithm. *Computers & Geosciences*, 35(3):487 – 495.

- Sourbier, F., Operto, S., Virieux, J., Amestoy, P. & L'Excellent, J.-Y. (2009b). Fwt2d : A massively parallel program for frequency-domain full-waveform tomography of wide-aperture seismic data—part 2 : Numerical examples and scalability analysis. *Computers & Geosciences*, 35(3):496 – 514.
- Stekl, I. & Pratt, R. G. (1998). Accurate viscoelastic modeling by frequency-domain finite difference using rotated operators. *Geophysics*, 63:1779–1794.
- Symes, W. W. (2007). Reverse time migration with optimal checkpointing. *Geophysics*, 72(5): SM213–SM221.
- Symes, W. W. & Carazzone, J. J. (1991). Velocity inversion by differential semblance optimization. *Geophysics*, 56:654–663.
- Taillandier, C., Noble, M., Chauris, H. & Calandra, H. (2009). First-arrival travel time tomography based on the adjoint state method. *Geophysics*, in-press.
- Tarantola, A. (1984). Inversion of seismic reflection data in the acoustic approximation. *Geophysics*, 49(8):1259–1266.
- Tarantola, A. (1986). A strategy for non linear inversion of seismic reflection data. *Geophysics*, 51(10):1893–1903.
- Tarantola, A. (1987). *Inverse problem theory : methods for data fitting and model parameter estimation*. Elsevier, New York.
- Tarantola, A. & Valette, B. (1982). Generalized nonlinear inverse problems solved using the least square criterion. *Reviews of Geophys. and Space Phys.*, 20:219–232.
- Thomsen, L. A. (1986). Weak elastic anisotropy. *Geophysics*, 51:1954–1966.
- Toksöz, M. N. & Johnston, D. H. (1981). *Geophysics reprint series, No. 2 : Seismic wave attenuation*. Society of exploration geophysicists, Tulsa, OK.
- Tsvankin, I. (2005). *Seismic signatures and analysis of reflection data in anisotropic media*. ELSEVIER, 2 édition.
- Vigh, D. & Starr, E. W. (2008). 3D prestack plane-wave, full-waveform inversion. *Geophysics*, 73(5):VE135–VE144.
- Vilotte, J., Festa, G. & Madariaga, R. (2005). Spectral element simulations of rupture dynamics along kinked faults. In *EOS Transactions*, volume 86. American Geophysical Union, San Francisco, USA.
- Virieux, J. (1986a). P-SV wave propagation in heterogeneous media, velocity-stress finite difference method. *Geophysics*, 51:889–901.
- Virieux, J. (1986b). *Séismes : rupture et onde (volumes I et II)*. Thèse de doctorat, Université Paris VII.
- Virieux, J. & Operto, S. (2009). An overview of full waveform inversion in exploration geophysics. *Geophysics*, in-press.

- Virieux, J., Operto, S., Ben-Hadj-Ali, H., Brossier, R., Etienne, V., Sourbier, F., Giraud, L. & Haidar, A. (2009). Seismic wave modeling for seismic imaging. *The Leading Edge*, 28(5):538–544.
- Warner, M., Stekl, I. & Umpleby, A. (2007). Full wavefield seismic tomography - iterative forward modelling in 3D. *In Expanded Abstracts*, page C025. EAGE.
- Warner, M., Stekl, I. & Umpleby, A. (2008). 3D wavefield tomography : synthetic and field data examples. *SEG Technical Program Expanded Abstracts*, 27(1):3330–3334.
- Williamson, P. (1991). A guide to the limits of resolution imposed by scattering in ray tomography. *Geophysics*, 56:202–207.
- Williamson, P. & Pratt, G. (1995). A critical review of 2.5D acoustic wave modeling procedures. *Geophysics*, 60:591–595.
- Woodwards, M. J. (1992). Wave-equation tomography. *Geophysics*, 57:15–26.
- Wu, R.-S. & Toksöz, M. N. (1987). Diffraction tomography and multisource holography applied to seismic imaging. *Geophysics*, 52:11–25.
- Yee, K. S. (1966). Numerical solution of initial boundary value problems involving Maxwell's equations in isotropic media. *IEEE Trans. Antennas and Propagation*, 14:302–307.
- Yoshimura, C., Bielak, J., Hisada, Y. & Fernandez, A. (2003). Domain Reduction Method for Three-Dimensional Earthquake Modeling in Localized Regions, Part II : Verification and Applications. *Bulletin of the Seismological Society of America*, 93(2):825–841.
- Zelt, C. & Barton, P. J. (1998). Three-dimensional seismic refraction tomography : a comparison of two methods applied to data from the faeroe basin. *Journal of Geophysical Research*, 103(B4):7187–7210.
- Zienkiewicz, O. C., Taylor, R. L. & Zhu, J. Z. (2005). *The finite element method its basis and fundamentals*. Elsevier. six edition.

Annexe A

Calcul des matrices de la formulation GD par intégration numérique

Sommaire

A.1 Rappels sur l'intégration numérique	233
A.2 Exemples sur interpolation P_1	234

A.1 Rappels sur l'intégration numérique

Le calcul de l'intégrale d'une fonction polynomiale \mathcal{F} sur une forme géométrique simple (1D, 2D, 3D) peut être faite de manière exacte et peu coûteuse par intégration numérique de Gauss.

Pour l'application à la méthode GD, pour laquelle les matrices $\tilde{\mathcal{F}}$, $\tilde{\mathcal{G}}$ et $\tilde{\mathcal{E}}$ doivent être calculées pour chaque cellule du maillage, l'idée générale de l'intégration par quadrature de Gauss est de transformer l'intégration d'une fonction simple sur un triangle dont la géométrie est propre à chaque cellule, en une intégration d'une fonction complexe sur un triangle simple et unique. Pour cela, on définit une cellule de référence (cellule mère) dans les coordonnées parents (ξ, η) et une cellule fille dans les coordonnées cartésienne (x, y) (Figure A.1).

L'intégration numérique de la fonction $\mathcal{F}(x, z)$ sur la cellule fille de géométrie arbitraire va être effectuée en appliquant une transformation géométrique à la fonction pour obtenir $\mathcal{F}(\xi, \eta)$ qui sera intégrée facilement sur la cellule mère de géométrie simple. En posant J la matrice Jacobienne de la transformation de la cellule fille en cellule mère, alors l'intégration peut s'écrire :

$$\int_{\Omega_e} \mathcal{F}(x, z) dV = \int_{\tilde{\Omega}} \det(J) \mathcal{F}(\xi, \eta) dV. \quad (\text{A.1})$$

L'intégration sur une cellule mère triangulaire se fait en évaluant la fonction $\mathcal{F}(\psi_1, \psi_2, \psi_3)$ aux points de Hammer, les points de Gauss étant définies pour des cellules quadrangulaires en 2D.

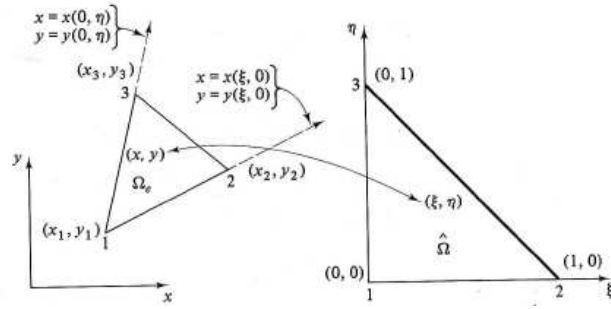


FIG. A.1 – Cellule fille en coordonnées (x, y) et cellule mère en coordonnées parent (ξ, η) (from Becker *et al.* (2001))

TAB. A.1 – Points de Hammer (trois triplets de positions) pour l'intégration numérique sur cellule triangulaire.

w_i	ξ_1^i	ξ_2^i	ξ_3^i
1/3	0.5	0.5	0
1/3	0.5	0	0.5
1/3	0	0.5	0.5

Des tables de positions des points de Gauss et de Hammer sont disponibles dans différents ouvrages d'éléments finis (par exemple Becker *et al.*, 2001; Zienkiewicz *et al.*, 2005). Ainsi nous avons :

$$\int_{\Omega_e} \mathcal{F}(\psi_1, \psi_2, \psi_3) dV = A \sum_{i=1}^{nG} w_i \mathcal{F}(\xi_1^i, \xi_2^i, \xi_3^i), \quad (\text{A.2})$$

où A est l'aire de la cellule Ω_e . Le nombre nG de points de Hammer $(\xi_1^i, \xi_2^i, \xi_3^i)$ et de poids w_i sont déterminés par l'ordre du polynôme à intégrer.

A.2 Exemples sur interpolation P_1

En utilisant trois triplets de points de Hammer (Table A.1), nous pouvons calculer la matrice \mathcal{K}_i définie par l'équation (1.45). Par exemple, pour le terme (1, 1) de la matrice, nous avons $\mathcal{F}(\psi_1, \psi_2, \psi_3) = \psi_1^2$, donnant :

$$\int_{\Omega_e} \mathcal{F}(\psi_1, \psi_2, \psi_3) dV = \int_{\Omega_e} \psi_1^2 dV = A_i \sum_{l=1}^3 w_l \psi_{l_1}^2 = A_i \left(\frac{1}{3} 0.5^2 + \frac{1}{3} 0.5^2 \right) = \frac{A_i}{6},$$

et pour le terme (1, 2), nous avons $\mathcal{F}(\psi_1, \psi_2, \psi_3) = \psi_1 \psi_2$, donnant :

$$\int_{\Omega_e} \psi_1 \psi_2 dV = \frac{A_i}{12}. \quad (\text{A.3})$$

En appliquant la même procédure pour la matrice entière, nous obtenons :

$$\mathcal{K}_i = \frac{A_i}{12} \begin{bmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{bmatrix}. \quad (\text{A.4})$$

De même, nous pouvons calculer la matrice \mathcal{E} par intégration sur le triangle (équation (1.46)), faisant intervenir les dérivées spatiales cartésiennes des fonctions de base. En utilisant l'expression cartésienne des fonctions de base (équation (1.69)), nous avons pour le terme (1, 1) de la matrice \mathcal{E}_x calculé avec $\mathcal{F}(\psi_1, \psi_2, \psi_3) = (\partial_x \psi_1) \psi_1$:

$$\begin{aligned} \int_{\mathcal{T}_i} \mathcal{F}(\psi_1, \psi_2, \psi_3) dV &= \int_{\mathcal{T}_i} (\partial_x \psi_1) \psi_1 dV = \int_{\mathcal{T}_i} \left(\frac{z_2 - z_3}{2A} \right) \psi_1 dV = \frac{A_i(z_2 - z_3)}{2A_i} \sum_{i=1}^3 w_i \psi_{i_1} \\ &= \frac{l_1 n_{1x}}{2} \left(\frac{1}{3} 0.5 + \frac{1}{3} 0.5 \right) = \frac{l_1 n_{1x}}{6}, \end{aligned} \quad (\text{A.5})$$

où l_1 est la longueur de l'arrête [2, 3] et n_{1x} est la composante horizontale du vecteur unitaire \vec{n}_1 normal au segment [2, 3]. Nous obtenons ainsi la matrice \mathcal{E}_{i_x} :

$$\mathcal{E}_{i_x} = \frac{1}{6} \begin{bmatrix} l_1 n_{1x} & l_1 n_{1x} & l_1 n_{1x} \\ l_2 n_{2x} & l_2 n_{2x} & l_2 n_{2x} \\ l_3 n_{3x} & l_3 n_{3x} & l_3 n_{3x} \end{bmatrix}, \quad (\text{A.6})$$

et la matrice \mathcal{E}_{i_z} :

$$\mathcal{E}_{i_z} = \frac{1}{6} \begin{bmatrix} l_1 n_{1z} & l_1 n_{1z} & l_1 n_{1z} \\ l_2 n_{2z} & l_2 n_{2z} & l_2 n_{2z} \\ l_3 n_{3z} & l_3 n_{3z} & l_3 n_{3z} \end{bmatrix}. \quad (\text{A.7})$$

Enfin, nous pouvons calculer les matrices \mathcal{F}_{ik} et \mathcal{G}_{ik} par intégration linéique par points de Gauss sur les arrêtes frontières \mathcal{T}_{ik} entre la cellule \mathcal{T}_i et les cellules \mathcal{T}_k . Nous supposons que la numérotation des points, des segments et des voisins suit la convention de la figure A.2 afin de garder une notation générique.

Ainsi nous avons par exemple le terme (1, 1) de la matrice \mathcal{F}_{i1} :

$$\int_{\mathcal{T}_{i1}} \psi_1^2 dL = \frac{l_1}{2} \int_{-1}^{+1} 0 dL = 0, \quad (\text{A.8})$$

à cause de la projection de la fonction ψ_1 sur \mathcal{T}_{i1} qui est nulle (*cf.* figure 1.26). Le terme (2, 2) donne :

$$\int_{\mathcal{T}_{i1}} \psi_2^2 dL = \frac{l_1}{2} \int_{-1}^{+1} \frac{-(\xi - 1)^2}{2} dL = \frac{2l_1}{6}, \quad (\text{A.9})$$

et le terme (2, 3) :

$$\int_{\mathcal{T}_{i1}} \psi_2 \psi_3 dL = \frac{l_1}{2} \int_{-1}^{+1} \frac{-(\xi - 1)}{2} \frac{-(-\xi - 1)}{2} dL = \frac{l_1}{6}. \quad (\text{A.10})$$

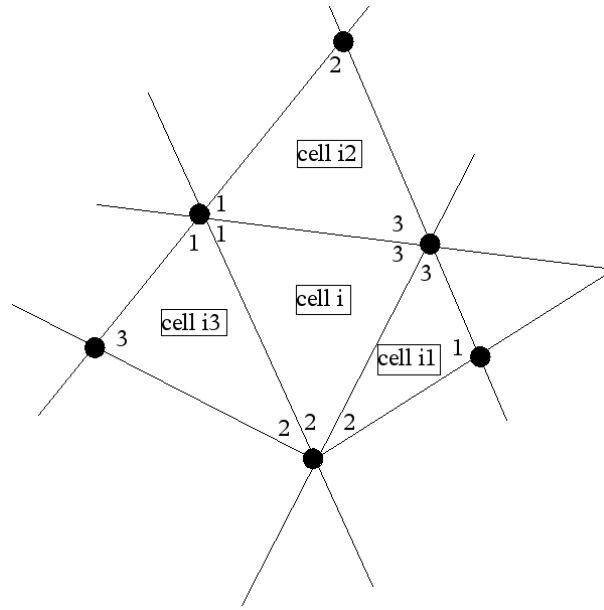


FIG. A.2 – Configuration des numérotations locales des points, arrêtes et voisins

Les trois matrices sont obtenues de manières similaires :

$$\mathcal{F}_{i1} = \frac{l_1}{6} \begin{bmatrix} 0 & 0 & 0 \\ 0 & 2 & 1 \\ 0 & 1 & 2 \end{bmatrix} \quad (\text{A.11})$$

$$\mathcal{F}_{i2} = \frac{l_2}{6} \begin{bmatrix} 2 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 2 \end{bmatrix} \quad (\text{A.12})$$

$$\mathcal{F}_{i3} = \frac{l_3}{6} \begin{bmatrix} 2 & 1 & 0 \\ 1 & 2 & 0 \\ 0 & 0 & 0 \end{bmatrix} \quad (\text{A.13})$$

Notons que ce cas P_1 est illustratif car les intégrations peuvent être faites terme par terme. Pour les ordres supérieurs ou des combinaisons d'ordres, des routines de calcul automatiques pourraient être utilisées.

Annexe B

Publications complémentaires

Sommaire

- B.1 Two-dimensional seismic imaging of the Valhall model from synthetic OBC data by frequency-domain elastic full-waveform inversion**238
 - B.2 Robust frequency-domain full-waveform inversion using the L_1 norm**243
-

Cette annexe comporte deux publications complémentaires aux articles présentés dans le corps de la thèse.

La première publication est un résumé étendu (*Expended Abstract*) de la conférence SEG 2009, qui présente les applications d'inversion acoustique et élastique des données synthétiques élastiques du modèle marin de Valhall.

La deuxième publication est une note publiée dans la revue *Geophysical Research Letters*, qui traite des limites de la norme aux moindres carrés pour l'inversion de données bruitées et de l'intérêt de la norme L_1 comme alternative au formalisme classique.

Two-dimensional seismic imaging of the Valhall model from synthetic OBC data by frequency-domain elastic full-waveform inversion

Romain Brossier ^{*}, Stéphane Operto ^{*} and Jean Virieux [†]

^{*} Géoazur - Université Nice Sophia-Antipolis - CNRS, [†] Laboratoire de Géophysique Interne et Tectonophysique, Université Joseph Fourier

SUMMARY

Quantitative imaging of the elastic properties of the subsurface is essential for reservoir characterization. We apply two-dimensional frequency-domain full-waveform inversion (FWI) to image shallow-water synthetic V_P and V_S models of the Valhall oil and gas field from 3-C ocean-bottom-cable data. In soft-seabed environment where a small amount of P-to-S mode conversion occurs at the sea bottom, the seismic wavefield is dominated by the P waves whereas the S waves have a weaker signature. We first show that acoustic FWI of elastic data provides accurate V_P model in such environment. Elastic FWI, that is desired for reservoir characterization, is a more difficult task. Hierarchical processing of the different parameter classes and data components is required, in addition to low-frequency data and robust multi-scale algorithm, to converge toward acceptable models.

INTRODUCTION

Quantitative seismic imaging of earth parameters is one of the main challenge for reservoir characterization. Frequency-domain full-waveform inversion (FWI) allows to build accurate models of complex structures from wide-aperture acquisition geometries (Pratt and Worthington, 1990). Most of the applications of FWI to real data at different exploration scales were performed under the acoustic approximation (Ravaut et al., 2004; Operto et al., 2006; Bleibinhaus et al., 2007) while only few papers recently tackled the elastic problem (Shi et al., 2007). Elastic parameters reconstruction is indeed a difficult and highly non-linear problem. The initial model building, that directly drives the convergence of the inversion, is a difficult issue for the S-wave velocity parameter (or linked parameters such as the Poisson's ratio): indeed, imaging of the V_S parameter has potentially a high-resolution power due to the short S wavelengths, that makes the inversion more sensitive to the inaccuracies of the starting model (Brossier et al., 2009). Free-surface effects, such as surface-waves and free-surface multiples, add additional non linearities in the inversion of land data, that require different hierarchical inversion regularizations (Gelis et al., 2007; Brossier et al., 2009). Weak P-to-S conversions in marine data with low S-wave velocity contrasts at the sea bed require the hierarchical reconstruction of the different parameter classes, and the hierarchical inversions of the different data components including careful phase selections for proper inversion convergence, even for simple canonical models (Sears et al., 2008). However, reconstruction of elastic parameters is useful in reservoir characterization: for example, Poisson's ratio anomalies are good marker of gas because the V_S is weakly affected by the presence of gas unlike the V_P one.

In this study, we present an application of frequency-domain elastic FWI to the synthetic Valhall model, representative of oil and gas fields in shallow water environment of North Sea. After a short review of elastic wave modeling and FWI, we first show that acoustic FWI of elastic data provided reliable V_P model in the Valhall geological environment. Second, we propose a multi-step strategy for elastic FWI, that allowed reliable reconstruction of V_S parameter in addition to that of the V_P model.

METHOD

Forward problem

Two-dimensional frequency-domain elastic FWI requires the solution of the two-way wave equation for P-SV waves in 2D heterogeneous media. We shall limit ourselves to the isotropic case in this study. In the frequency domain, one may write the elastic wave equation as:

$$\begin{aligned} -i\omega\rho\mathbf{V} &= \nabla \cdot \boldsymbol{\sigma} + \rho f \\ -i\omega\boldsymbol{\sigma} &= \mathbf{c} : \nabla\mathbf{V} - i\omega\boldsymbol{\sigma}_0, \end{aligned} \quad (1)$$

where unknown quantities of the forward problem are horizontal and vertical particle velocities $\mathbf{V} = (V_X, V_Z)$ and stress $\boldsymbol{\sigma}$ fields. The system 1 is discretized with a Discontinuous Galerkin method (Brossier et al., 2008). Low-order P_0 and P_1 interpolations are used for wavefield parameterization (figure 1). The P_1 interpolation is used on top of the model (in the first 160 meters) in an unstructured triangular mesh. This allows the accurate positioning of the sources and receivers in the mesh, by mean of the linear interpolation basis functions and the locally-refined mesh. The water-solid interface at a depth of 70 m is embedded inside the P_1 zone, because the P_0 interpolation is unstable at the water-solid interface (Brossier et al., 2008). For time and memory computational savings, the P_1 interpolation order is limited to the shallow part of the model, whereas the P_0 interpolation is used below 160 m on a regular equilateral mesh. The sparse linear system resulting from the discretization of system 1 is solved efficiently for multiple sources using the parallel direct solver MUMPS (Amestoy et al., 2006).

Inverse problem

Frequency-domain FWI (Pratt et al., 1998) is an optimization problem, which can be recast as a linearized least-squares problem that minimizes the misfit between the frequency-domain recorded data (d_{obs}) and modeled data ($d_{calc}^{(k)}$) at iteration k , defined by the objective function:

$$\mathcal{E}^{(k)} = \frac{1}{2} (d_{obs} - d_{calc}^{(k)})^\dagger \mathbf{S}_d^\dagger \mathbf{S}_d (d_{obs} - d_{calc}^{(k)}). \quad (2)$$

Superscript [†] indicates the adjoint (transposed conjugate) and \mathbf{S}_d is a diagonal weighting matrix applied to the misfit vector to scale the relative contributions of each of its components. The

2D elastic FWI: VALHALL model

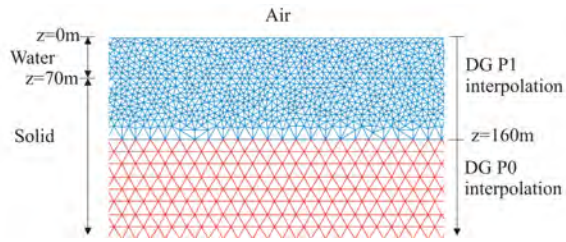


Figure 1: Close-up of the triangular mesh centred on the sea bottom, used for elastic wave propagation modeling in the Valhall model. P_1 and P_0 cells are displayed respectively in blue and red.

gradient $\mathcal{G}^{(k)}$ of the objective function is efficiently computed by the adjoint-state technique (Plessix, 2006), which requires solving only two forward problems per shot.

A second-order Taylor expansion of the objective function allows to find the Newton update $\delta \mathbf{m}$ which minimizes the objective function:

$$\mathbf{B}^{(k)} \delta \mathbf{m} = -\mathcal{G}^{(k)}, \quad (3)$$

where $\mathbf{B}^{(k)}$ is the Hessian matrix. Because of the huge computational time required to build $\mathbf{B}^{(k)}$, Newton method are generally avoided. Steepest Descent or Conjugate Gradient methods scaled by the diagonal terms of an approximate Hessian are generally used in FWI (Pratt et al., 1998; Operto et al., 2006; Shin et al., 2001). In this study, we use the Quasi-Newton L-BFGS method (Nocedal, 1980). L-BFGS provides an approximate inverse Hessian at a marginal computational extra cost, that is useful to improve the convergence rate of the iterative imaging process (Brossier et al., 2009).

Data preconditioning

Non-linearity of FWI can be efficiently mitigated by selecting a subset of specific arrivals (e.g., Sears et al., 2008). Data preconditioning can be applied in the frequency domain by means of complex frequencies ($\omega + i\gamma$), which is equivalent to damp seismograms in time (Brenders and Pratt, 2007). The Fourier transform of a signal $f(t)$ damped in time by $\exp^{-\gamma(t-t_0)}$ is given by:

$$F(\omega + i\gamma) \exp^{\gamma t_0} = \int_{-\infty}^{+\infty} f(t) \exp^{-\gamma(t-t_0)} \exp^{-i\omega t} dt, \quad (4)$$

where the damping, controlled by the γ parameter, can be applied from an arbitrarily arrival time t_0 , which usually corresponds to the first-arrival time. Time-damping applied from the first-arrival time can be viewed as a heuristic way to select aperture angles in the data. This aperture selection can be exploited to implement a second-level of hierarchy in the multi-scale inversion, in addition to that provided by the frequency selection: during each successive frequency inversion, we hierarchically performed several inversion iterations using decreasing imaginary terms γ to progressively inject in the inversion more resolving arrivals associated with later-arriving phases (i.e., shorter-aperture phases). This strategy revealed useful for inversion of land data (Brossier et al., 2009) and will be used in this study.

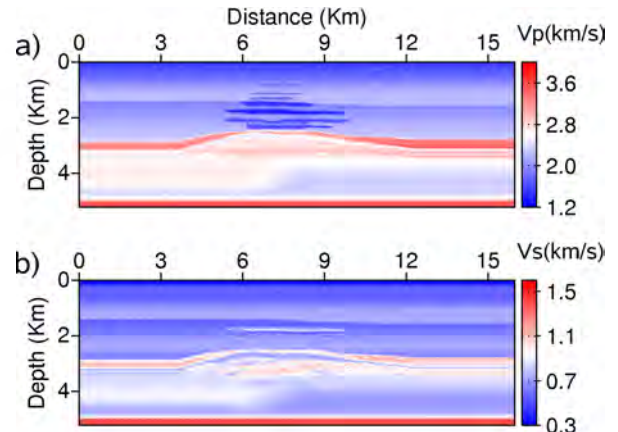


Figure 2: Synthetic Valhall models. a) V_P , b) V_S .

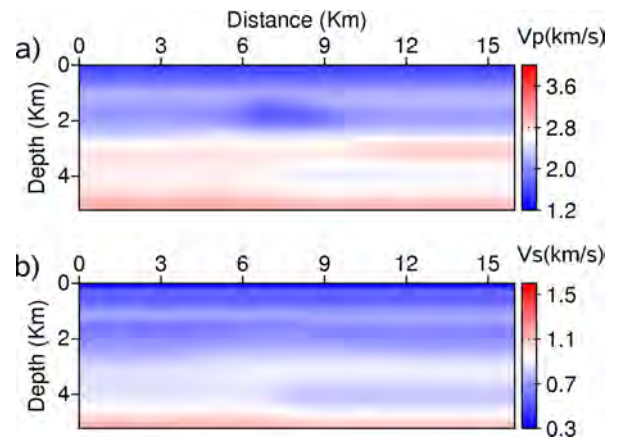


Figure 3: Initial V_P (a) and V_S (b) models for FWI.

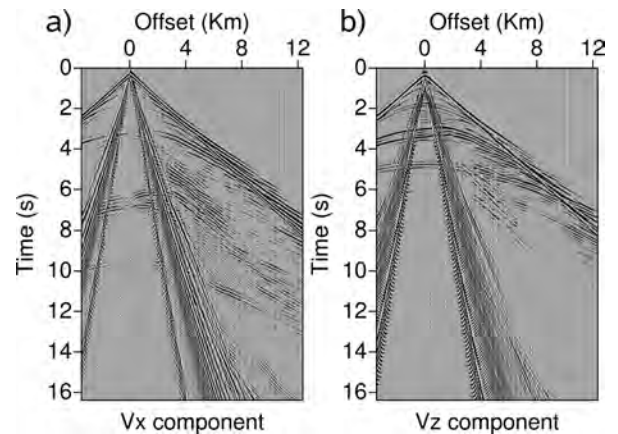


Figure 4: Example of time-domain shot gather for a source at $x = 3$ km in the synthetic Valhall elastic models for a) horizontal component and b) vertical one of geophones.

VALHALL ELASTIC MODEL AND DATA

The figure 2 shows the V_P and V_S synthetic models provided by BP. The main targets are the gas cloud in the large sediments layer, and the trapped oil underneath the cap rock composed of chalk in a deeper part of the model. Gas clouds are easily identified by the low P-wave velocities, whereas their signature is

2D elastic FWI: VALHALL model

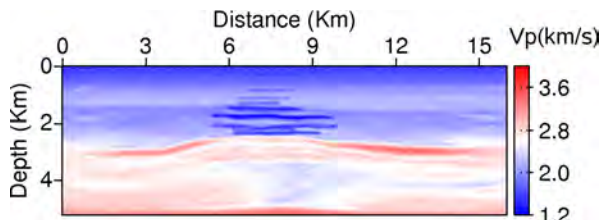


Figure 5: V_P model inferred from acoustic FWI of elastic data.

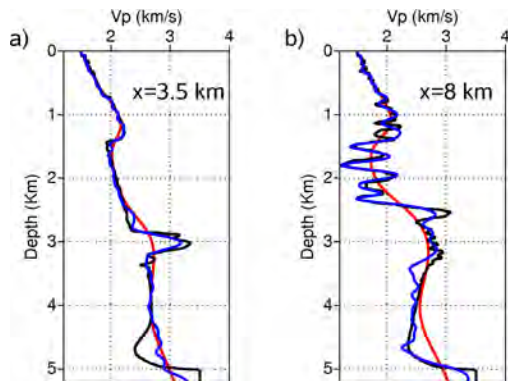


Figure 6: Vertical logs at distances a) 3.5 km and b) 8 km extracted from the FWI model of Figure 5. True, starting and FWI models are displayed in black, red and blue, respectively.

much weaker in the V_S model. In this study, density is considered constant and assumed to be known during the inversion. Starting models (Figure 3) are inferred from the true ones by Gaussian smoothing, the vertical correlation of which linearly increases from 25 m at the water-solid interface to 1000 m at 5200 m depth. The horizontal correlation length is 500 m. The acquisition geometry mimics an OBC acquisition. 315 pressure sources are located 6 m below the water surface with an horizontal spacing of 50 m. 315 3-C sensors spaced by 50 m, each composed of one hydrophone, and one vertical and inline horizontal geophones, are located at a depth of 70 m on the sea bottom.

A time-domain shot-gather is shown in figure 4 for the V_X and the V_Z components. Low V_S on the sea bottom and the relatively smooth V_S model led to small-amplitude converted P-to-S waves on both the vertical and horizontal components, and made the reconstruction of the V_S model ill-posed.

In this study, the elastic observed data have been computed with the frequency-domain P_0 - P_1 DG method for the frequencies and damping factors used in inversion. The maximum frequency involved in the inversion was 6 Hz, and the triangular mesh was designed according to the minimum S wavelength at this frequency (Figure 1). For all the tests, source signature is estimated by linear inversion during each frequency inversion iteration (Pratt, 1999).

ACOUSTIC INVERSION OF ELASTIC DATA

We should consider first acoustic FWI of elastic data. As observed in the figure 4, P arrivals dominate the seismic wavefield (especially, on the vertical geophone and hydrophone), leading to a favorable framework for acoustic FWI (Barnes and

Charara, 2008).

Acoustic FWI was performed with the FWT2D code (Sourbier et al., 2009) based on a finite-difference forward problem engine to solve the Helmholtz equation. Five frequencies were inverted successively (2, 3, 4, 5 and 6 Hz) using 3 damping factors ($\gamma = 2, 0.33, 0.1$) applied in cascade during each frequency inversion. The inversion was applied to the hydrophone component. The initial V_P model for acoustic FWI is shown in figure 3a. The steepest Descent method was used for optimization, preconditioned by the diagonal terms of the approximate Hessian (Operto et al., 2006). Fifteen iterations were performed per damping factor leading to 45 iterations per frequency inversion. The final FWI V_P model does not exhibit significant artifacts (Figure 5). Comparison between V_P logs, extracted from the true and final FWI models, shows a reliable reconstruction of the P-wave velocities, in spite of the low frequencies involved in the inversion (Figure 6).

These results suggest that acoustic FWI can be relevant in some marine case studies, where elastic effects have a relatively-weak signature in the data. The resulting V_P model may be used as a starting model to proceed with elastic FWI in later stage.

ELASTIC FULL-WAVEFORM INVERSION

Inversion set-up

We now applied elastic FWI to the 3-C OBC data. Sears et al. (2008) have illustrated that elastic FWI is ill-posed when the S-wave velocity models exhibit relatively small contrasts, that prevent recording of significant P-to-S converted waves. In order to mitigate the ill-posedness of FWI, they designed a multi-step strategy, that was adapted to the frequency-domain FWI as described below:

1- In a first step, V_P is reconstructed from the hydrophone data. The forward problem is performed with the elastic DG method but the V_S model is left unchanged during FWI. The aim of this first stage is to improve the V_P model to significantly decrease the P-wave residuals. During this first step, a coarse mesh adapted to the V_P wavelength was designed for computational efficiency. In this case, S-wave modeling was affected by numerical dispersion, that, however, did not significantly impact the V_P -model reconstruction. Note that the final V_P model of this first stage is close to that of acoustic inversion.

2- In a second step, the V_P and V_S models are reconstructed simultaneously from the horizontal and vertical components of geophones. A gain, given by the source-receiver offset to the power of 2, was applied to the data through the matrix S_d . The aim of this weighting was to increase the weight of the intermediate-to-long offset data, at which converted P-to-S arrivals are recorded.

Other tests performed without this two-step strategy, where both V_P and V_S parameters are reconstructed simultaneously from the beginning of the inversion process, failed and get stuck in a local minima.

FWI was performed for the same data as for the acoustic test: five frequencies were inverted successively (2, 3, 4, 5 and 6 Hz) with 3 damping factors applied in cascade for each frequency inversion ($\gamma = 2, 0.33, 0.1$). Starting models are shown in Figure 3. L-BFGS algorithm was used for optimization. We selected the diagonal of the Pseudo-Hessian (Shin et al., 2001) as

2D elastic FWI: VALHALL model

an initial guess for the L-BFGS iterations. Ten iterations were performed per damping factor, leading to 30 iterations per frequency inversion.

Results

Reconstructed V_P and V_S models and vertical logs at distances 3.5 and 8 km are displayed in Figures 7 and 8, respectively. The V_P model shows an accurate reconstruction of the main geological features, such as the anticlinal structure and the gas cloud in the sediments. The V_P model inferred from elastic inversion (Figure 7a) is close to that inferred from acoustic FWI (Figure 5), except in the deep part of the model (below 2.5 km), where elastic FWI provided improved reconstructions (compare Figures 6 and 8). The V_S model suffers from a deficit of short wavenumbers, especially in the deeper part where the starting model is less accurate. The very low velocities in the model endows the V_S imaging with a high resolution power, even at the low frequency of 2 Hz, that requires a very accurate starting model, a key issue in the prospect of real data processing. Amplitudes in the V_S model are not reconstructed as well as the V_P ones, and more iterations may be necessary to retrieve the correct amplitudes. The reconstructed models give, however, a useful information on the shallowest part of the target even for the V_S velocity model: these models would allow the identification and characterization of gas clouds, thanks to elastic properties variations.

CONCLUSION

This study presents an application of frequency-domain Full-Waveform Inversion of multi-components OBC data to the Valhall synthetic case study. The FWI algorithm is implemented through a two-loop scheme, over frequencies and time damping, to efficiently mitigate the non-linearities of the inversion. The quasi-Newton L-BFGS method is implemented for optimization.

The geological structure of the Valhall model makes the P arrivals dominant in the seismic wavefield, contrary to the S ones. On the one hand, P-wave velocity model can be successfully reconstructed by acoustic FWI, hence showing the relevance of the acoustic approximation for some marine case studies. On the other hand, elastic-parameters imaging is a difficult task, and complex multi-step strategies involving hierarchic inversions of data components and hierarchic reconstructions of different parameter classes must be designed to converge towards acceptable velocity models, if low frequencies and sufficiently accurate starting models are not available. Future works will focus on the starting model building by Laplace-domain inversion (Shin and Cha, 2008) and PP-PS stereotomography (Alerini et al., 2002), before considering applications to real data.

ACKNOWLEDGMENTS

This study has been funded by the SEISCOPE consortium <http://seiscope.unice.fr>, sponsored by BP, CGG-VERITAS, EXXON-MOBIL, SHELL and TOTAL, and by ANR under project ANR-05-NT05-2-42427. Access to the high-performance computing facilities of MESOCENTRE SIGAMM (OCA) and IDRIS (CNRS) computer centers provided the required computer resources.

Many thanks go to J. Koppal and L. Sirgue from BP for providing the elastic synthetic models of Valhall.

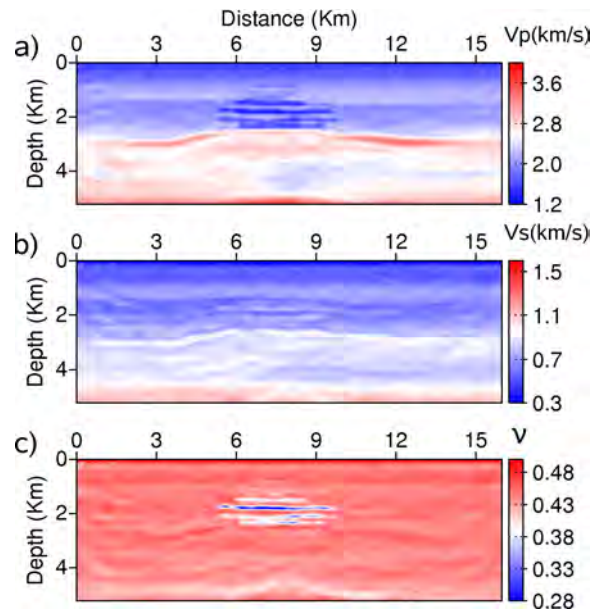


Figure 7: Reconstructed (a) V_P and (b) V_S models inferred from elastic FWI. (c) associated Poisson's ratio.

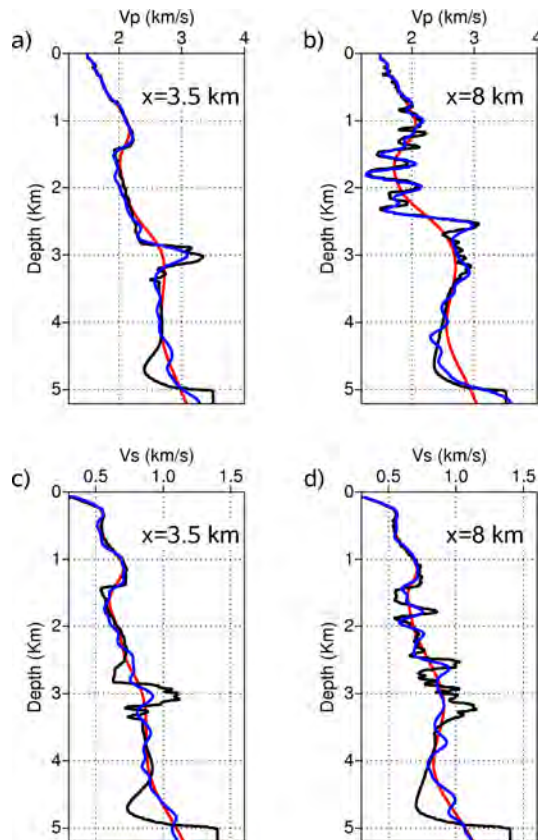


Figure 8: Vertical logs at distances a) 3.5 km and b) 8 km extracted from the FWI model of Figure 7. True, starting and FWI models are displayed in black, red and blue, respectively.

EDITED REFERENCES

Note: This reference list is a copy-edited version of the reference list submitted by the author. Reference lists for the 2009 SEG Technical Program Expanded Abstracts have been copy edited so that references provided with the online metadata for each paper will achieve a high degree of linking to cited sources that appear on the Web.

REFERENCES

- Alerini, M., S. Le Bégat, G. Lambaré, and R. Baina, 2002, 2D PP- and PS-stereotomography for a multicomponent dataset: 72nd Annual International Meeting, SEG, Expanded Abstracts, 838–841.
- Amestoy, P. R., A. Guermouche, J. Y. L'Excellent, and S. Pralet, 2006, Hybrid scheduling for the parallel solution of linear systems: *Parallel Computing*, **32**, 136–156.
- Barnes, C. and M. Charara, 2008, Full-waveform inversion results when using acoustic approximation instead of elastic medium: 78th Annual International Meeting, SEG, Expanded Abstracts, 1895–1899.
- Bleibinhaus, F., J. A. Hole, T. Ryberg, and G. S. Fuis, 2007, Structure of the California Coast Ranges and San Andreas Fault at SAFOD from seismic waveform inversion and reflection imaging: *Journal of Geophysical Research*, **112**, doi:10.1029/2006JB004611.
- Brenders, A. J. and R. G. Pratt, 2007, Full waveform tomography for lithospheric imaging: Results from a blind test in a realistic crustal model: *Geophysical Journal International*, **168**, 133–151.
- Brossier, R., S. Operto, and J. Virieux, 2009, 2D elastic frequency-domain full-waveform inversion for imaging complex onshore structures: 71th Conference & Exhibition, EAGE, Extended Abstracts, in press.
- Brossier, R., J. Virieux, and S. Operto, 2008, Parsimonious finite-volume frequency-domain method for 2-D P-SV-wave modelling: *Geophysical Journal International*, **175**, 541–559.
- Gelis, C., J. Virieux, and G. Grandjean, 2007, 2D elastic waveform inversion using Born and Rytov approximations in the frequency domain: *Geophysical Journal International*, **168**, 605–633.
- Nocedal, J., 1980, Updating Quasi-Newton matrices with limited storage: *Mathematics of Computation*, **35**, 773–782.
- Operto, S., J. Virieux, J. X. Dessa, and G. Pascal, 2006, Crustal imaging from multifold ocean bottom seismometers data by frequency-domain full-waveform tomography: Application to the eastern Nankai trough: *Journal of Geophysical Research*, **111**, doi:10.1029/2005JB003835.
- Plessix, R.-E., 2006, A review of the adjoint-state method for computing the gradient of a functional with geophysical applications: *Geophysical Journal International*, **167**, 495–503.
- Pratt, R. G., 1999, Seismic waveform inversion in the frequency domain, part I: Theory and verification in a physics scale model: *Geophysical Journal International*, **64**, 888–901.
- Pratt, R. G., C. Shin, and G. J. Hicks, 1998, Gauss-Newton and full Newton methods in frequency-space seismic waveform inversion: *Geophysical Journal International*, **133**, 341–362.
- Pratt, R. G., and M. H. Worthington, 1990, Inverse theory applied to multi-source cross-hole tomography. Part 1: Acoustic wave-equation method: *Geophysical Prospecting*, **38**, 287–310.
- Ravaut, C., S. Operto, L. Improta, J. Virieux, A. Herrero, and P. dell'Aversana, 2004, Multi-scale imaging of complex structures from multi-fold wide-aperture seismic data by frequency-domain full-wavefield inversions: application to a thrust belt: *Geophysical Journal International*, **159**, 1032–1056.
- Sears, T., S. Singh, and P. Barton, 2008, Elastic full waveform inversion of multi-component OBC seismic data: *Geophysical Prospecting*, **56**, 843–862.
- Shi, Y., W. Zhao, and H. Cao, 2007, Nonlinear process control of wave-equation inversion and its application in the detection of gas: *Geophysics*, **72**, no. 1, R9–R18.
- Shin, C. and Y. H. Cha, 2008, Waveform inversion in the Laplace domain: *Geophysical Journal International*, **173**, 922–931.
- Shin, C., S. Jang, and D. J. Min, 2001, Improved amplitude preservation for prestack depth migration by inverse scattering theory: *Geophysical Prospecting*, **49**, 592–606.
- Soubier, F., S. Operto, J. Virieux, P. Amestoy, and J.-Y. L'Excellent, 2009, FWT2D: A massively parallel program for frequency-domain full-waveform tomography of wide-aperture seismic data — Part 1: Algorithm: *Computers & Geosciences*, **35**, 487–495.



Robust elastic frequency-domain full-waveform inversion using the L_1 norm

R. Brossier,¹ S. Operto,² and J. Virieux³

Received 12 June 2009; accepted 21 September 2009; published 23 October 2009.

[1] Elastic frequency-domain full-waveform inversion is here applied to a realistic synthetic offshore model to study the behavior of the least-squares and least-absolute-value norms, in the presence of noisy multi-component seismic data. As expected, the inversion results show that the L_2 norm is highly sensitive to non-Gaussian errors in the data and gives rise to high amplitude artifacts in the reconstructed models. The L_1 norm shows more robust behavior, whatever the noise characteristics, and allows convergence towards acceptable models, for both the compressional-wave and shear-wave velocity models. This study highlights the sensitivity of the full-waveform inversion results for the chosen norm in the case of noisy data. The marginally used L_1 norm reveals a robust alternative to the L_2 norm for frequency-domain full-waveform inversion of weakly redundant data. **Citation:** Brossier, R., S. Operto, and J. Virieux (2009), Robust elastic frequency-domain full-waveform inversion using the L_1 norm, *Geophys. Res. Lett.*, 36, L20310, doi:10.1029/2009GL039458.

1. Introduction

[2] The physical properties of the Earth are often estimated through seismic-wave analysis. Wave arrival-time information is generally used to determine the tomography of the Earth. In the 1980s, the full-waveform inversion (FWI) method was introduced by *Tarantola* [1984] to exploit the full information contained in the seismograms, in terms of the full waveform, and to infer high-resolution models of the subsurface.

[3] Originally developed in the time domain [*Tarantola*, 1984, 1987], FWI has become tractable and popular since the pioneering work of *Pratt and Worthington* [1990] and *Pratt* [1990] in the frequency domain, and is now investigated more, particularly for active seismic surveys at various scales. Computationally efficient frequency-domain FWI was designed by limiting the inversion to a few discrete frequencies, taking advantage of the redundant wavenumber coverage provided by wide-aperture surveys [*Sirgue and Pratt*, 2004; *Brenders and Pratt*, 2007]. This wavenumber redundancy can be partly sacrificed by using a judicious subset of frequencies in the inversion, at the expense of the signal-to-noise ratio of the reconstructed

models. FWI potentially provides high-resolution quantitative images of physical parameters, but suffers from two main difficulties. The first is related to the computational cost of the forward problem; namely, the numerical resolution of the wave equation in heterogeneous media for multiple sources. The second difficulty is related to the ill-posedness and the non-linearity of the inverse problem, which is generally formulated as a least-squares local optimization [*Tarantola and Valette*, 1982], to manage the numerical cost of the forward problem. The ill-posedness and the non-linearity of the FWI arise mainly from the limited accuracy of the starting model in the framework of local optimization approaches, the incomplete illumination of the subsurface provided by conventional seismic surveys, the limited bandwidth sources and presence of noise. Several hierarchical multiscale strategies that proceed from low frequencies to higher frequencies [*Pratt and Worthington*, 1990; *Bunks et al.*, 1995; *Sirgue and Pratt*, 2004; *Brossier et al.*, 2009a], have been proposed to mitigate the non-linearity of the inverse problem.

[4] The noise footprint in seismic imaging is conventionally mitigated by stacking highly redundant multifold data. Mitigating the sensitivity of the inversion to noise is however a key issue in efficient frequency-domain FWI, where the data redundancy is partly sacrificed for computational efficiency. The least-squares objective function remains the most commonly used, although it suffers from poor robustness when dealing with large isolated and non-Gaussian errors. Other norms should therefore be considered. The L_1 norm was introduced in time-domain FWI by *Tarantola* [1987] and *Crase et al.* [1990], and has been shown to be weakly sensitive to noise. *Djikpéssé and Tarantola* [1999] used this norm successfully to invert field data from the Gulf of Mexico. Alternatives like the Huber criterion [*Huber*, 1973; *Guittou and Symes*, 2003; *Ha et al.*, 2009], the *sech* criterion [*Crase et al.*, 1990; *Monteiller et al.*, 2005] and the hybrid L_1/L_2 criterion [*Bube and Langan*, 1997] can also be considered. All these criteria behave as the L_2 norm for small residuals and as the L_1 norm for large ones. A threshold that needs to be provided defines where the transition between these two different behaviors takes place. These hybrid criteria are efficient for dealing with outliers in data. However, they assume Gaussian statistics as soon as the L_2 norm is used, leading to the same limitations as for the pure L_2 norm in the presence of ambient noise if the threshold is not chosen correctly.

[5] In this study, we assess two-dimensional (2D) elastic, frequency-domain FWI from noisy, multi-component seismic data of the oil & gas field of Valhall computed in synthetic velocity models. The results clearly demonstrate the lack of robustness of the classic least-squares norm, whereas the L_1 norm that has seen marginal use in the FWI

¹Géoazur, Observatoire de la Côte d'Azur, Université de Nice Sophia Antipolis, CNRS, Valbonne, France.

²Géoazur, Observatoire de la Côte d'Azur, Université de Nice Sophia Antipolis, CNRS, Villefranche-sur-Mer, France.

³LGIT, Université Joseph Fourier, Grenoble, France.

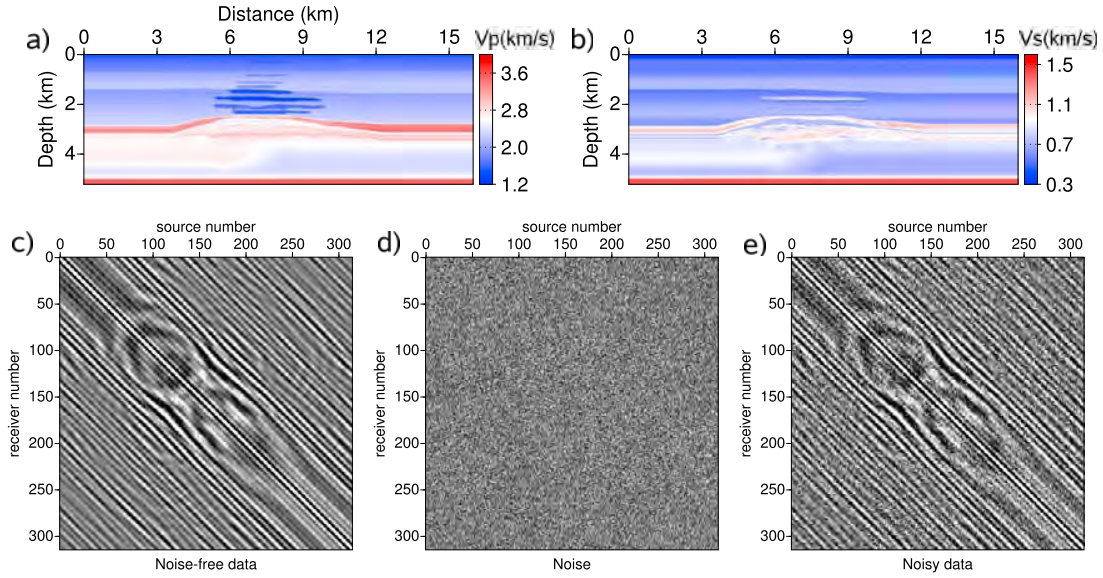


Figure 1. Synthetic true Valhall models for (a) V_P and (b) V_S . Real part of the hydrophone data for frequency $(4 + 10i)$ Hz. (c) Noise-free data, (d) added noise and (e) noisy data.

community, is shown to be very robust, even in the case of highly decimated data.

2. Theory and Algorithm

2.1. Least-Squares Norm

[6] The least-squares criterion provides the most usual framework for the development of frequency-domain FWI [Pratt and Worthington, 1990; Pratt, 1990]. The L_2 functional is usually written in the following form:

$$C_{L_2}^{(k)} = \frac{1}{2} \Delta \mathbf{d}' \mathbf{S}_d' \mathbf{S}_d^* \Delta \mathbf{d}^*, \quad (1)$$

where $\Delta \mathbf{d} = \mathbf{d}_{\text{obs}} - \mathbf{d}_{\text{calc}}^{(k)}$ is the data misfit vector, the difference between the observed data \mathbf{d}_{obs} and the modeled data $\mathbf{d}_{\text{calc}}^{(k)}$ computed in the model $\mathbf{m}^{(k)}$. k is the iteration number of the non-linear iterative inversion. Superscript * indicates the conjugate and \mathbf{S}_d is a diagonal weighting matrix applied to the misfit vector to scale the relative contributions of each of its components.

[7] Differentiation of $C_{L_2}^{(k)}$ with respect to the model parameters gives the following expression of the gradient:

$$\mathcal{G}_{L_2}^{(k)} = \mathcal{R}[\mathbf{J}' \mathbf{S}_d' \mathbf{S}_d^* \Delta \mathbf{d}^*], \quad (2)$$

where \mathbf{J} is the Fréchet derivative matrix. The gradient can be derived from the adjoint-state method that allows the computing of $\mathcal{G}_{L_2}^{(k)}$ without explicitly forming \mathbf{J} [Plessix, 2006], giving the expression:

$$\mathcal{G}_{m_i}^{(k)} = \mathcal{R} \left[\mathbf{v}' \left[\frac{\partial \mathbf{A}'}{\partial m_i} \right] \mathbf{A}^{-1} \mathbf{S}_d' \mathbf{S}_d^* \Delta \mathbf{d}^* \right], \quad (3)$$

where the incident wavefield \mathbf{v} is linearly related to the source \mathbf{s} through the forward problem operator \mathbf{A} : $\mathbf{A} \mathbf{v} = \mathbf{s}$.

[8] The gradient can be seen as a weighted zero-lag convolution between the incident wavefield \mathbf{v} emitted by

the seismic source and the residual wavefield back-propagated from the receiver positions $\mathbf{A}^{-1} \mathbf{S}_d' \mathbf{S}_d^* \Delta \mathbf{d}^*$.

2.2. Least-Absolute-Value Norm

[9] For complex arithmetic numbers, such as frequency-domain data, we can extend the L_1 norm developed by Tarantola [1987] and Crase *et al.* [1990], written as:

$$C_{L_1}^{(k)} = (\Delta \mathbf{d}' \mathbf{S}_d' \mathbf{S}_d^* \Delta \mathbf{d}^*)^{1/2}. \quad (4)$$

[10] The gradient $\mathcal{G}_{L_1}^{(k)}$ is given by:

$$\mathcal{G}_{L_1}^{(k)} = \mathcal{R} \left[\mathbf{J}' \mathbf{S}_d' \frac{\Delta \mathbf{d}^*}{|\Delta \mathbf{d}|} \right], \quad (5)$$

assuming that $|\Delta \mathbf{d}| > 0$, considering the machine precision used. For all of the tests that we performed, we never met any case where $|\Delta \mathbf{d}| = 0$.

[11] The L_1 and L_2 gradients have similar forms, as seen by equations (2) and (5), except that the residual term differs. For the L_1 norm, the data residuals are normalized according to their amplitudes, which gives clear insights why the L_1 norm is expected to be less sensitive to large residuals.

[12] In the case of real arithmetic numbers, the term $\frac{\Delta \mathbf{d}^*}{|\Delta \mathbf{d}|}$ of expression (5) corresponds to the function *sign* [Tarantola, 1987; Crase *et al.*, 1990].

2.3. Algorithm

[13] The 2D elastic, frequency-domain FWI algorithm used in this study is described by Brossier *et al.* [2009a]. The reader is thus referred to Brossier *et al.* [2009a] for a complete description of the algorithm. The algorithm embeds three main loops: the outer one is over frequency groups; namely, a set of frequencies simultaneously inverted. The second loop is over time-damping factors that control the amount of information over time that is preserved for inversion in the seismograms. Time damping is

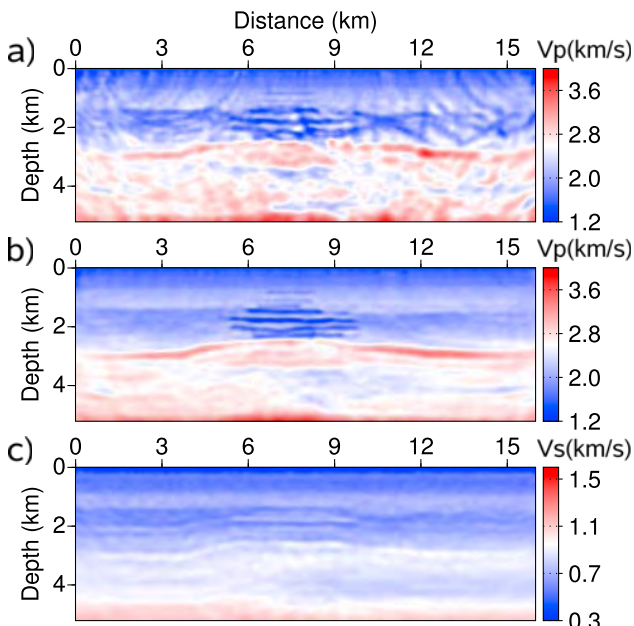


Figure 2. Reconstructed models for the test 1 with outliers, for (a) the V_P parameter with L_2 and the (b) V_P and (c) V_S parameters with L_1 norm.

applied in the frequency-domain modeling by using complex-valued frequencies which is equivalent to damp seismograms in time by $\exp^{-(t-t_0)\gamma}$ where t_0 is the first-arrival traveltimes. The third loop is over iterations of one frequency group inversion. The two outer nested loops define two hierarchical multiscale levels in the inversion that are helpful to mitigate the non-linearity of the inversion. The forward problem is performed with a discontinuous Galerkin method for solving the elastodynamic equations in the frequency domain [Brossier *et al.*, 2008]. The linear system that results is solved in parallel using the MUMPS LU solver [Amestoy *et al.*, 2006]. The optimization is solved with the quasi-Newton L-BFGS algorithm [Nocedal, 1980]. The diagonal of the pseudo-Hessian matrix [Shin *et al.*, 2001] is used as an initial guess for the L-BFGS algorithm for both the L_1 and the L_2 criteria.

3. Numerical Example: Offshore Valhall Model

[14] The numerical example is based on the synthetic Valhall model (Figure 1), which is representative of oil and gas fields in shallow water environments of the North Sea [Munns, 1985]. The main targets are the gas cloud in the large sediment layer, and the trapped oil underneath the cap rock, which is composed of chalk, in a deeper part of the model. Gas clouds are easily identified by the low P-wave velocities, whereas their signature is much weaker in the V_S model. The selected acquisition mimics a three-component ocean bottom cable survey [Kommedal *et al.*, 2004]. This environment shows a particularly ill-posed problem for S-wave velocity reconstruction, due to the relatively small shear-wave velocity contrast at the sea bed that prevents recording of significant P-to-S converted waves. A successful inversion requires a multi-step hierarchical strategy in the manner of Sears *et al.* [2008], as developed by Brossier

et al. [2009b] for noise-free data. We follow the same approach for noisy data:

[15] 1. In the first step, the V_P velocity is reconstructed from the hydrophone data. The forward problem is performed with the elastic discontinuous Galerkin method, but the V_S model is left unchanged during the FWI. The aim of this first stage is to improve the V_P model so as to significantly decrease the P-wave residuals. During this first step, a coarse mesh that is adapted to the V_P wavelength is designed for computational efficiency. In this case, S-wave modeling is affected by numerical dispersion that, however, does not significantly impact on the V_P -model reconstruction.

[16] 2. In the second step, the V_P and V_S models are reconstructed simultaneously from the horizontal and vertical components of the geophones. An amplification, with a gain given by the power of 2 of the source-receiver offset, is applied to the data through the matrix S_d . This weighting increases the weight of the intermediate-to-long-offset data at which the converted P-to-S arrivals are recorded.

[17] Five frequencies were inverted successively (2, 3, 4, 5 and 6 Hz) with 3 regularization damping factors applied in cascade for each frequency inversion ($\gamma = 2, 0.33, 0.1 \text{ s}^{-1}$). Starting models were built by smoothing the true models with a Gaussian filter, the vertical correlation length of which increased linearly from 25 m to 1000 m with depth, and the horizontal one was fixed at 500 m. Ten iterations were performed per damping factor, leading to 30 iterations per frequency inversion. Density is constant and assumed to be known in the inversion. The source is estimated during the inversion by linear inversion [Pratt, 1999].

[18] Two tests were performed. For both tests, uniform white noise was introduced into the observed data, computed using the forward-problem engine implemented in the inversion code (the so-called inverse crime). The signal-to-noise ratio was set to 10. Figures 1c–1e show the 4-Hz noise-free and noisy data in the source-receiver domain.

[19] During the first test, we introduced outliers into the data: large errors (noise had been multiplied by 20) were introduced randomly in one trace out of one hundred, to simulate a poorly preprocessed data set. The resulting noise is consequently no more uniform for this test. The V_P models obtained after the first inversion step with the L_2 and L_1 norms are shown in Figure 2. The L_1 norm provides an accurate V_P model, whereas the inversion rapidly converged towards a local minimum when the L_2 norm was used. For the L_2 norm, the inversion was stopped close to the first step because of the failure. However, the second step was performed for the L_1 norm and shows good quality of reconstruction for both V_P and V_S models.

[20] A second test was performed without any outliers, and considering only the ambient noise. The V_P and V_S models obtained with the L_2 and L_1 norms after the second step are shown in Figure 3 and reveal again the improved robustness of the L_1 norm.

4. Discussion

[21] The use of the L_2 and L_1 norms in elastic, frequency-domain FWI highlights the sensitivity of the optimization convergence to the noise.

[22] The first test with outliers illustrates the known behavior of the L_2 norm in the presence of high-amplitude

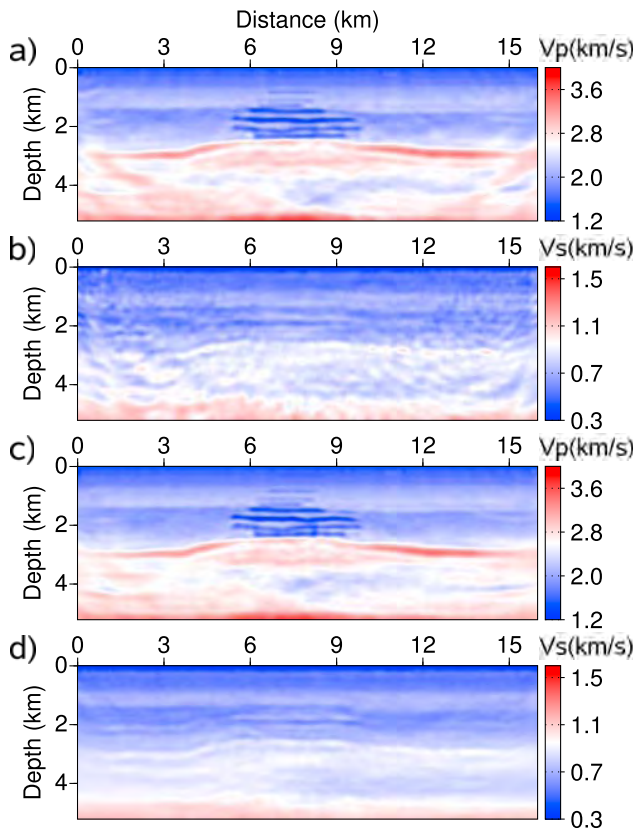


Figure 3. Reconstructed V_P and V_S models for the test 2 without outliers, for (a and b) the L_2 and (c and d) the L_1 norm.

isolated noise. The L_2 criterion intrinsically amplifies the weight of the high-amplitude residuals during inversion, hence causing divergence of the optimization if the residuals do not correspond to coherent signals. The L_1 norm shows stable behavior in this unfavorable test, because the isolated, high-amplitude outliers have negligible contributions in the final images. This confirms the robustness of this criterion applied to FWI, even when a limited amount of data is considered.

[23] The second test without outliers shows that the V_P reconstruction is robust for both norms, whereas only the L_1 norm allowed the successful reconstruction of the V_S model. In this shallow-water environment with low velocity contrasts at the sea bed, the V_P imaging is more linear than the V_S imaging for two main reasons. First, the larger P-wavelengths are less resolving than the S counterparts, and are therefore less sensitive to the inaccuracy of the starting model in the framework of a multiscale reconstruction [Brossier et al., 2009a]. Second, the P-waves dominate the seismic wavefield, whereas the P-to-S waves have a weaker footprint in the data. The limited signature of the S-waves in the data makes the inversion poorly conditioned for the S-wave parameter class, even with noise-free data. Brossier et al. [2009b] showed how the hierarchical two-step strategy allow the strengthening of the sensitivity of the inversion to the V_S parameter during the second step, and hence the successful reconstruction of the V_S model with the

L_2 norm in the case of noise-free data. However, adding noise to the data still contributes to the degradation of the sensitivity of FWI to the P-to-S arrivals. In this case, the two-step strategy implemented with the L_2 norm failed to reconstruct the V_S model, and the L_1 norm revealed that it was necessary to converge towards acceptable V_S models by mitigating the contribution of the amplitudes in the reconstruction.

5. Conclusion

[24] This study presents an application of elastic, frequency-domain FWI of controlled-source synthetic seismic data, to compare the behaviors of the classic L_2 and the robust L_1 functionals in the presence of noisy data. Our results show first the robust behavior of the L_1 norm in the presence of high-amplitude non-Gaussian noise, such as for outliers, whereas the L_2 criterion failed to produce acceptable models as expected. Second, the L_1 norm allows the successful exploitation of the P-to-S arrivals, which have a relatively weak footprint in the data, to image the V_S velocity model in the presence of white noise. In this case too, the L_2 norm failed to reconstruct the V_S model. The robustness of the L_1 norm with respect to noise is attributed to the fact that the footprints of the amplitudes are mitigated in the imaging. The L_1 norm should be an alternative to the classic L_2 norm, particularly when a limited amount of data are to be inverted. This is the case in efficient frequency-domain FWI, where only a limited number of discrete frequencies are inverted independently. The L_1 norm can also be useful to invert low-fold passive data, such as teleseismic data for lithospheric imaging.

[25] **Acknowledgments.** This study is funded by the SEISCOPE consortium (<http://seiscope.oca.eu/>), which is sponsored by BP, CGG-VERITAS, EXXON-MOBIL, SHELL and TOTAL, and by Agence Nationale de la Recherche (ANR) under project ANR-05-NT05-2-42427. The LU factorization of the impedance matrix was performed with MUMPS (<http://graal.ens-lyon.fr/MUMPS/index.html>). The mesh generation was performed with the help of TRIANGLE (<http://www.cs.cmu.edu/~quake/triangle.html>). Access to the high-performance computing facilities of the MESOCENTRE SIGAMM computer centre and IDRIS national (project 082280) centers provided the required computer resources, and we gratefully acknowledge both of these facilities and the support of their staff. Many thanks go to J. Kommendal and L. Sirgue from BP, for providing the elastic synthetic models of Valhall.

References

- Amestoy, P. R., A. Guermouche, J. Y. L'Excellent, and S. Pralet (2006), Hybrid scheduling for the parallel solution of linear systems., *Parallel Comput.*, 32, 136–156.
- Brenders, A. J., and R. G. Pratt (2007), Efficient waveform tomography for lithospheric imaging: Implications for realistic 2D acquisition geometries and low frequency data, *Geophys. J. Int.*, 168, 152–170.
- Brossier, R., J. Virieux, and S. Operto (2008), Parsimonious finite-volume frequency-domain method for 2-D P-SV-wave modelling, *Geophys. J. Int.*, 175, 541–559.
- Brossier, R., S. Operto, and J. Virieux (2009a), Seismic imaging of complex onshore structures by two-dimensional elastic frequency-domain full-waveform inversion, *Geophysics*, in press.
- Brossier, R., S. Operto, and J. Virieux (2009b), Two-dimensional seismic imaging of the valhall model from synthetic OBC data by frequency-domain elastic full-waveform inversion, *SEG Tech. Program Expanded Abstr.*, in press.
- Bube, K. P., and R. T. Langan (1997), Hybrid l_1/l_2 minimization with applications to tomography, *Geophysics*, 62, 1183–1195.
- Bunks, C., F. M. Saleck, S. Zaleski, and G. Chavent (1995), Multiscale seismic waveform inversion, *Geophysics*, 60, 1457–1473.

- Cruse, E., A. Pica, M. Noble, J. McDonald, and A. Tarantola (1990), Robust elastic non-linear waveform inversion: Application to real data, *Geophysics*, *55*, 527–538.
- Djickpéssé, H. A., and A. Tarantola (1999), Multiparameter l_1 norm waveform fitting: Interpretation of Gulf of Mexico reflection seismograms, *Geophysics*, *64*, 1023–1035.
- Guittou, A., and W. W. Symes (2003), Robust inversion of seismic data using the Huber norm, *Geophysics*, *68*, 1310–1319.
- Ha, T., W. Chung, and C. Shin (2009), Waveform inversion using a back-propagation algorithm and a Huber function norm, *Geophysics*, *74*, R15–R24.
- Huber, P. J. (1973), Robust regression: Asymptotics, conjectures, and Monte Carlo, *Ann. Stat.*, *1*(5), 799–821.
- Kommedal, J. H., O. I. Barkved, and D. J. Howe (2004), Initial experience operating a permanent 4C seabed array for reservoir monitoring at Valhall, *SEG Tech. Program Expanded Abstr.*, *23*, 2239–2242.
- Monteiller, V., J.-L. Got, J. Virieux, and P. Okubo (2005), An efficient algorithm for double-difference tomography and location in heterogeneous media, with an application to the Kilauea volcano, *J. Geophys. Res.*, *110*, B12306, doi:10.1029/2004JB003466.
- Munns, J. W. (1985), The Valhall field: A geological overview, *Mar. Pet. Geol.*, *2*, 23–43.
- Nocedal, J. (1980), Updating quasi-Newton matrices with limited storage, *Math. Comput.*, *35*(151), 773–782.
- Plessix, R.-E. (2006), A review of the adjoint-state method for computing the gradient of a functional with geophysical applications, *Geophys. J. Int.*, *167*, 495–503.
- Pratt, R. G. (1990), Inverse theory applied to multi-source cross-hole tomography. Part II: Elastic wave-equation method, *Geophys. Prospect.*, *38*, 311–330.
- Pratt, R. G. (1999), Seismic waveform inversion in the frequency domain, part I: Theory and verification in a physical scale model, *Geophys. J. Int.*, *64*, 888–901.
- Pratt, R. G., and M. H. Worthington (1990), Inverse theory applied to multi-source cross-hole tomography. Part I: Acoustic wave-equation method, *Geophys. Prospect.*, *38*, 287–310.
- Sears, T., S. Singh, and P. Barton (2008), Elastic full waveform inversion of multi-component OBC seismic data, *Geophys. Prospect.*, *56*, 843–862.
- Shin, C., S. Jang, and D. J. Min (2001), Improved amplitude preservation for prestack depth migration by inverse scattering theory, *Geophys. Prospect.*, *49*, 592–606.
- Sirgue, L., and R. G. Pratt (2004), Efficient waveform inversion and imaging: A strategy for selecting temporal frequencies, *Geophysics*, *69*, 231–248.
- Tarantola, A. (1984), Inversion of seismic reflection data in the acoustic approximation, *Geophysics*, *49*, 1259–1266.
- Tarantola, A. (1987), *Inverse Problem Theory: Methods for Data Fitting and Model Parameter Estimation*, Elsevier, New York.
- Tarantola, A., and B. Valette (1982), Generalized nonlinear inverse problems solved using the least square criterion, *Rev. Geophys. Space Phys.*, *20*, 219–232.

R. Brossier, Géoazur, Observatoire de la Côte d'Azur, Université de Nice Sophia Antipolis, CNRS, 250 rue Albert Einstein, F-06560 Valbonne, France. (brossier@geoazur.unice.fr)

S. Operto, Géoazur, Observatoire de la Côte d'Azur, Université de Nice Sophia Antipolis, CNRS, La Darse, B.P. 48, F-06235 Villefranche-sur-Mer CEDEX, France. (operto@geoazur.obs-vlfr.fr)

J. Virieux, LGIT, Université Joseph Fourier, B.P. 53, F-38041 Grenoble CEDEX 9, France. (jean.virieux@obs.ujf-grenoble.fr)