



Ecole doctorale Sciences Fondamentales et Appliquées
Unité de recherche : Géoazur

Thèse de doctorat

Présentée en vue de l'obtention du
grade de docteur en Sciences de la Planète et de l'Univers
de
l'UNIVERSITÉ CÔTE D'AZUR

par

Stephen BELLER

Imagerie lithosphérique par inversion de formes d'ondes télésismiques

Application aux Alpes Occidentales

Dirigée par **Stéphane OPERTO** et **Guust NOLET**

Soutenue le 24 février 2017 devant le Jury composé de

Sébastien CHEVROT
Bertrand DELOUIS
Vadim MONTEILLER
Guust NOLET
Stéphane OPERTO
Anne PAUL
Luis RIVERA
Jean VIRIEUX

Directeur de Recherche CNRS
Professeur Université Côte d'Azur
Ingénieur de Recherche CNRS
Professeur Université Côte d'Azur
Directeur de Recherche CNRS
Directrice de Recherche CNRS
Professeur Université de Strasbourg
Professeur Université Grenoble-Alpes

Rapporteur
Examinateur
Invité
Co-directeur de Thèse
Directeur de Thèse
Examinateur
Examinateur
Rapporteur

Résumé

L'imagerie fine de la lithosphère est essentielle pour comprendre les mécanismes profonds qui gouvernent la dynamique terrestre et valider les modèles géodynamiques déduits des observations de surface. En sismologie, les méthodes de tomographie ou de migration de fonctions récepteur n'ont pas toujours la résolution suffisante pour proposer une interprétation géodynamique univoque. Les déploiements récents de réseaux denses de capteurs ainsi que l'augmentation des capacités de calcul permettent aujourd'hui l'inversion de formes d'ondes complètes pour l'imagerie de la lithosphère à partir de données téléseismiques. Cette technique permet d'obtenir des modèles quantitatifs du sous-sol avec une résolution spatiale proche de la longueur d'onde en exploitant la totalité de l'information contenue dans les sismogrammes.

Dans cette thèse, un algorithme massivement parallèle d'inversion de formes d'ondes téléseismiques est développé pour des milieux élastiques isotropes 3D. La modélisation sismique est effectuée efficacement avec une méthode hybride d'injection de champ d'onde. Une première modélisation est effectuée dans une Terre globale à symétrie sphérique avec le logiciel AxiSEM pour déterminer les champs d'ondes aux bords de la cible lithosphérique. Ces solutions sont ensuite propagées dans cette cible avec une méthode aux éléments finis spectraux. Le problème inverse est résolu avec un algorithme d'optimisation locale de type quasi-Newton (l-BFGS).

La sensibilité de la méthode à la configuration expérimentale (paramétrisation du milieu, modèle initial, géométrie et échantillonnage du dispositif de capteurs) est tout d'abord analysée avec un modèle synthétique réaliste des Alpes Occidentales. Les trois principales conclusions sont : (i) Des modèles précis et bien résolus sont obtenus en reconstruisant la densité et les paramètres de Lamé (ou les vitesses P et S) durant l'inversion et en recombinaison des impédances P et S a posteriori. (ii) Des modèles terrestres de référence tels que PREM fournissent des modèles initiaux acceptables. (iii) Des déploiements surfaciques de capteurs par opposition à des lignes denses améliorent significativement la résolution verticale et horizontale.

L'algorithme est finalement appliqué à neuf événements de la campagne CIFALPS dans les Alpes occidentales jusqu'à une fréquence de 0.2Hz. Les modèles de vitesses P et S et de densité révèlent les grandes structures lithosphériques de la chaîne alpine avec une résolution encore jamais atteinte, en particulier la géométrie des grands bassins sédimentaires, du corps d'Ivrée et des Moho européen et adriatique. Plus profondément, deux anomalies de vitesses lentes sous les Alpes occidentales sont imagées dans le manteau. La première à l'Ouest de la chaîne est interprétée comme la signature d'une remontée asthénosphérique, la seconde au niveau du corps d'Ivrée indique le détachement du panneau plongeant européen. L'étude valide ainsi l'hypothèse d'une subduction continentale de la croûte européenne et confirme un éventuel déchirement du panneau plongeant européen au cours de la collision continentale.

Abstract

Fine-scale imaging of lithospheric structures is crucial to understand the deep mechanisms that govern the Earth dynamics and to appraise geodynamical models inferred from surface observations. In seismology, traveltimes tomography or receiver functions migration have often an insufficient resolution to propose a unique geodynamical interpretation of lithospheric structures. However, the recent deployments of dense seismic arrays, together with the continuous growth in computing facilities, are amenable to full-waveform inversion (FWI) of teleseismic data for high-resolution lithospheric imaging. This technique can produce quantitative multi-parameter models of the subsurface with a resolution close to the wavelength by handling the full information carried out by very general wave types.

In this thesis, a massively-parallel teleseismic waveform inversion algorithm is developed for 3D elastic isotropic media. Efficient seismic modeling is performed with a wavefield injection hybrid approach. A first simulation is performed in a symmetric global Earth with the AxiSEM code to compute the wavefields on the borders of the lithospheric target. Then, these wavefields are propagated in the target with the spectral finite-element method. After linearization, the inverse problem is solved with a quasi-Newton (1-BFGS) local optimization algorithm.

The sensitivity of the teleseismic FWI to the experimental setup (subsurface parametrization, initial model, sampling and geometry of the station layout) is first assessed with a realistic synthetic model of the Western Alps. The main conclusions are : (i) Reliable models are obtained by updating the density and Lamé parameters (or the P and S wavespeeds) during inversion before recombining the P and S impedances for visualization and interpretation. These numerical results are consistent with the radiation pattern of the different investigated parametrizations. (ii) One-dimensional reference Earth models, such as PREM, are sufficiently-accurate to drive the inversion toward acceptable models. (iii) Coarse areal deployment of stations provide subsurface models of higher vertical and horizontal resolutions compared to dense linear deployments. The method is finally applied to nine events of the CIFALPS experiment carried out in the Western Alps, up to a frequency of 0.2Hz. Reliable models of P and S wavespeeds and density reveal with an unprecedented resolution the crustal and lithospheric structures of the Alpine Belt, in particular the geometry of the large sedimentary basins from the Po plain and Rhone valley, the Ivrea body, and the European and Adriatic Mohos. Deeper, two slow velocity anomalies beneath the Western Alps are imaged in the mantle. The first, to the west of the chain, is interpreted as the signature of an asthenospheric upwelling, the second near the location of the Ivrea body indicates the European slab break-off. The study validates the hypothesis of the European continental crust subduction and confirms the possible tearing of the European slab.

Remerciements

Avant toute chose, je tiens à remercier toutes les personnes qui ont participé ou contribué de près ou de loin à ce travail de thèse.

En premier lieu je tiens à remercier les membres du Jury. Je remercie Sébastien Chevrot et Jean Virieux qui ont tous deux accepté de rapporter mon travail, ainsi que Bertrand Delouis, Anne Paul et Luis Rivera qui ont accepté d'être examinateur. Merci à tous d'avoir considéré mon travail avec intérêt.

Mes remerciements s'adressent ensuite naturellement à mes directeurs de thèse Stéphane Operto et Guust Nolet. Je vous remercie de m'avoir accordé votre confiance pour mener à bien ce travail de recherche, d'avoir toujours été disponibles et de m'avoir guidé tout au long de ces trois années. J'ai grandement bénéficié de vos connaissances dans vos domaines respectifs, de votre implication dans mon travail, et de votre rigueur somme toute scientifique.

Ce travail n'aurait probablement pas été le même sans l'étroite participation des membres de l'équipe « ImOndes » et du groupe SEISCOPE. Je remercie tout particulièrement Vadim Monteiller qui a co-encadré mon travail de recherche depuis le début, ainsi que Laure Combe qui a toujours été là pour m'aider à debugger, me soutenir moralement et permis de paralléliser le code d'inversion. J'ai toujours eu plaisir à échanger, collaborer, discuter avec vous que ce soit pendant une pause café ou derrière un clavier d'ordinateur. De même je souhaite exprimer mes remerciements à Alain Miniussi, Alessandra Ribodetti, Jean-Xavier Dessa, Ebru Bozdağ, Tony Monfret, Romain Brossier et Ludovic Métivier avec lesquels j'ai pu discuter à maintes reprises.

Je tiens à remercier tous les membres du laboratoire Géoazur, les personnels administratifs, l'équipe informatique, les chercheurs ainsi que les doctorants. Je leur suis reconnaissant pour leur accueil, leur aide et leur soutien, ainsi que leur bonne humeur au quotidien.

Je remercie une seconde fois Anne Paul pour l'accès aux données de la campagne CIFALPS et pour les nombreuses discussions concernant l'interprétation des modèles tomographiques.

Ce travail ayant nécessité d'intenses calculs numériques, je tiens à remercier les structures qui m'ont permis d'accéder à des ressources de calcul hautes-performances. Je remercie donc l'administration et les équipes techniques du mesocentre de calcul SIGAMM de l'Observatoire de la Côte d'Azur ainsi que du centre national de calcul du CINES qui m'ont permis de mener à bien ce projet de recherche.

Enfin, je remercie toutes celles et ceux qui m'ont soutenu durant ces trois années en me témoignant leur amitié et leur affection.

A vous tous, Merci !

Table des matières

Remerciements	7
Introduction générale	17
I Modélisation de la propagation des ondes télésismiques	37
1 Généralités sur la propagation des ondes sismiques	39
1.1 Qu'est-ce qu'une onde ?	40
1.2 Les équations de l'élastodynamique	41
1.3 Ondes propagées dans un milieu élastique	44
1.3.1 Ondes de volume	44
1.3.2 Ondes de surface	47
1.4 Quelques solutions analytiques	48
1.4.1 Les fonctions de Green	48
1.4.2 Onde plane dans un demi-espace homogène	49
1.5 Ondes sismiques en milieux hétérogènes	51
1.5.1 La diffusion des ondes élastiques	52
1.5.2 Diffusion simple par l'approximation de Born	54
1.6 Caractéristiques du champ d'onde télésismique	58
1.6.1 Les sources télésismiques	58
1.6.2 Ondes générées à l'échelle globale, phases télésismiques	60
1.6.3 Multiples lithosphériques	61
1.6.4 Données synthétiques pour un modèle lithosphérique des Alpes Occidentales	64
2 Modélisation numérique	69
2.1 Résolution d'une équation aux dérivées partielles	70
2.1.1 Les différences-finies	71
2.1.2 Les éléments-finis	72
2.2 Méthode des éléments spectraux	73
2.2.1 Forme faible des équations de l'élastodynamique	74
2.2.2 Discrétisation : maillages et fonctions de forme	75
2.2.3 Représentation polynomiale et intégration sur un élément	76
2.2.4 Intégration numérique de la forme faible de l'équation d'onde	79

2.2.5	Intégration temporelle – Le schéma de Newmark	81
2.2.6	Implémentation	83
2.2.7	Exemples de modélisation et validations	86
2.3	Stratégie d’injection de champ d’onde	92
2.3.1	L’approche champ total / champ diffracté	93
2.3.2	Injection d’une onde plane	94
2.4	Injection d’un champ d’onde téléseismique avec AxiSEM	96
2.4.1	AxiSEM	97
2.4.2	Aspects pratiques du couplage	98
2.4.3	Exemple et validation du couplage SEM/AxiSEM	101
II	Inversion de formes d’ondes téléseismique	107
3	Le problème d’inversion des formes d’ondes	109
3.1	Notions fondamentales sur les problèmes inverses	110
3.1.1	Le problème inverse	110
3.1.2	Caractère mal posé des problèmes inverses	111
3.1.3	Résolution des problèmes inverses linéaire et non-linéaire	113
3.1.4	Méthodes de résolution locale	116
3.2	L’inversion de formes d’ondes, un problème aux moindres carrés linéarisé	122
3.2.1	La fonction coût	122
3.2.2	Linéarisation du problème	123
3.2.3	Les équations normales	124
3.2.4	Régularisation	127
3.2.5	Paramétrisation	128
3.3	Le gradient de la fonction coût	129
3.3.1	Construction par la méthode de l’état adjoint	129
3.3.2	Implémentation et validation	133
3.3.3	Anatomie d’un gradient téléseismique	136
3.4	Algorithme de la FWI	141
3.4.1	Implémentation	141
3.4.2	Exemple d’inversion synthétique	141
3.4.3	Estimation de la source	143
4	Analyse paramétrique de la FWI en configuration téléseismique	147
4.1	Introduction	149
4.2	Teleseismic full-waveform inversion	153
4.2.1	Theoretical background	153
4.2.2	Gradient of the misfit function	154
4.2.3	Resolution analysis	156
4.2.4	Radiation patterns for different subsurface parametrization	157
4.3	The Western Alps lithospheric model	162
4.3.1	Model description	162
4.3.2	Teleseismic data anatomy	163

4.4	Parametric analysis of teleseismic FWI	165
4.4.1	Experimental set-up	165
4.4.2	Which earth parametrisation for teleseismic FWI?	167
4.4.3	Which initial model for teleseismic FWI?	170
4.4.4	Sensitivity to acquisition design	176
4.5	Conclusions	184
III Application de la FWI aux données télésismiques		185
5	Les Alpes Occidentales	187
5.1	Présentation de la chaîne alpine	188
5.2	Histoire géodynamique des Alpes	192
5.3	Synthèse des connaissances géophysiques	194
5.3.1	Structures crustales	194
5.3.2	Structures mantelliques	199
5.4	Conclusions et problématiques actuelles	201
6	Application de l'inversion de formes d'ondes aux données CIFALPS	205
6.1	Introduction	207
6.2	Multi-parameter teleseismic waveform inversion	210
6.2.1	Teleseismic full wavefield modelling	210
6.2.2	Inverse problem	212
6.2.3	High-resolution multi-parameter imaging	213
6.3	Application to the Western Alps	215
6.3.1	The CIFALPS experiment	215
6.3.2	Full-waveform inversion workflow	216
6.3.3	FWI models	220
6.3.4	Resolution and parameter cross-talk appraisal	223
6.4	Discussion	228
6.4.1	Density parameter reconstruction	228
6.4.2	Geodynamic interpretation	229
6.5	Conclusions	232
Conclusions et perspectives		235
Bibliographie		241

Table des figures

1	Structure radiale de la Terre.	18
2	Schéma présentant la configuration téléseismique.	20
3	Rai versus noyau de sensibilité.	24
4	Méthode des fonctions récepteur.	27
5	Découplage d'échelles entre le modèle de référence et le modèle de réflectivité.	30
6	Principe de l'inversion de formes d'ondes complètes.	32
1.1	Période et longueur d'onde	41
1.2	Enregistrements du séisme de Tohoku-Oki.	45
1.3	Ondes propagées dans les milieux élastiques.	46
1.4	Coefficients de réflexion à la surface libre.	51
1.5	Régimes de diffusion.	53
1.6	Approximation de Born	56
1.7	Diagrammes de radiation d'un point diffusant (ρ, λ, μ).	57
1.8	Nomenclature des phases générées à l'échelle globale.	59
1.9	Hodochrones des phases générées à l'échelle globale.	61
1.10	Phases profondeur.	62
1.11	Phases lithosphériques.	63
1.12	Hodochrones lithosphériques pour un point diffractant et une interface plane.	64
1.13	Coefficients de réflexion et transmission pour un milieu à deux couches.	65
1.14	Exemple synthétique de données téléseismiques.	66
1.15	Effets de la topographie sur les données téléseismiques.	68
2.1	Discretisation du domaine de calcul par des hexaèdres.	76
2.2	Interpolateurs de Lagrange	77
2.3	Raffinement de maillage avec la brique « doubling »	84
2.4	Exemples de maillages et leurs décompositions en sous-domaines.	85
2.5	Validation de la modélisation pour une source explosive.	87
2.6	Validation de la modélisation pour une source de type force.	88
2.7	Validation du problème de Lamb.	90
2.8	Modélisation SEM pour une force ponctuelle à la surface libre.	91
2.9	Géométrie de la méthode champ total - champ diffracté	94
2.10	Injection d'une onde plane.	95
2.11	Validation de l'injection d'une onde plane.	96
2.12	Diagrammes de radiation monopolaires, dipolaires et quadripolaires.	97
2.13	Modélisation de champ d'onde par AxiSEM.	102

2.14	Injection d'un champ d'onde généré par AxiSEM.	103
2.15	Validation du couplage avec AxiSEM.	104
3.1	Schéma représentatif des problèmes directs et inverses.	111
3.2	Approche hiérarchique en fréquence.	114
3.3	Non-linéarité, minima locaux et approche hiérarchique en fréquence.	115
3.4	Recherche linéaire et conditions de Wolfe.	118
3.5	Construction du gradient par la méthode de l'état adjoint pour une source explosive	133
3.6	Validation du gradient.	135
3.7	Construction du gradient par la méthode de l'état adjoint pour une source télé-	
	sismique	137
3.8	Construction du gradient, sommation sur les sources et récepteurs.	138
3.9	Anatomie d'un gradient télé-sismique.	139
3.10	Anatomie des données synthétiques utilisée pour le calcul du gradient télé-sis-	
	mique.	140
3.11	Workflow de la FWI télé-sismique.	142
3.12	Reconstruction d'un inclusion cubique multiparamètres.	143
3.13	Reconstruction d'un inclusion cubique monoparamètre.	144
4.1	Sensitivity matrix and FWI imaging principle	155
4.2	Radiation patterns for elastic isotropic media.	158
4.3	Anatomy of a teleseismic gradient for one source-receiver pair and different	
	parametrisations.	159
4.4	Synthetic example design.	164
4.5	Synthetic data computed in the Alps lithospheric model.	166
4.6	Subsurface parametrisation assessment, velocities.	169
4.7	Subsurface parametrisation assessment, impedences.	171
4.8	Subsurface parametrisation assessment, elastic coefficients.	172
4.9	FWI model quality.	173
4.10	Starting models for FWI.	174
4.11	Assessment of initial models.	175
4.12	Effect of acquisition geometry on FWI reconstruction.	178
4.13	Resolution test for the areal acquisition.	182
4.14	Resolution test for the linear acquisition.	183
5.1	Localisation de la chaîne alpine au sein de la ceinture péri-méditerranéenne.	
	D'après De Graciansky et al. (2011).	189
5.2	Présentation de la chaîne Alpine et de ses grandes unités structurales. D'après	
	De Graciansky et al. (2011).	190
5.3	Présentation des grandes zones structurales des Alpes occidentales.	191
5.4	Évolution paléogéographique du domaine Téthysien (d'après Renard et al. (2015)).	192
5.5	Océan ligure et son devenir alpin le long de la transversale Pelvoux – Mont	
	Viso. Modifié d'après Renard et al. (2015).	193
5.6	Profondeur du Moho et anomalie de Bouguer dans les Alpes occidentales.	194
5.7	Profils sismiques ECORS-CROP des Alpes franco-italiennes.	195

5.8	Interprétation structurale des profils ECORS-CROP.	196
5.9	Tomographie de Diehl et al. (2009) (coupes horizontales).	197
5.10	Tomographie de Diehl et al. (2009) (coupe verticale).	198
5.11	Fonctions récepteur migrées de la campagne CIFALPS (Zhao et al., 2015).	199
5.12	Tomographies du manteau de Piromallo & Morelli (2003) et Bijwaard & Spakman (2000).	200
5.13	Tomographies du manteau de Lippitsch et al. (2003) et Zhao et al. (2016a).	202
5.14	Biréfringence des ondes SKS dans les Alpes Occidentales (Barruol et al., 2011).	204
6.1	Sketch of the teleseismic configuration.	211
6.2	Scattering patterns of elastic waves for the (ρ, V_p, V_s) parametrisation.	214
6.3	The CIFALPS experiment.	217
6.4	Records of event 22.	218
6.5	Data fit.	221
6.6	Vertical section of the Alps FWI model.	222
6.7	Depth slices of the Alps FWI model.	224
6.8	Resolution analysis, checkerboard test.	225
6.9	Resolution analysis and parameter cross-talk appraisal.	227
6.10	Resolution analysis for a 50 km-wide (ρ, V_p, V_s) inclusion located at the position of the negative V_s anomaly shown in Fig. 6.6c,e.	228
6.11	Alps FWI model without density update.	230
6.12	Gravity data against FWI density model.	231
6.13	Carte du réseau AlpArray.	240

Liste des tableaux

1.1	Coefficients indépendants de quelques groupes de symétrie.	43
2.1	Stabilités et propriétés associées à la méthode de Newmark en fonction des paramètres β et γ (Hughes, 2003).	82
4.1	Selected events for teleseismic waveform modelling.	165
4.2	Statistics of the FWI applications.	168
6.1	Selected event for teleseismic FWI.	218

Introduction générale

« *Au commencement il y eut Chaos. Puis Gaia. Puis Eros.* »
La théogonie, Hésiodes

Sismologie et imagerie des structures internes de la Terre

La géophysique interne est une science dont l'objet principal est de recueillir des connaissances détaillées sur la structure et la composition de l'intérieur de la Terre pour comprendre les mécanismes géodynamiques actuels et passés. Elle a pour vocation de sonder les profondeurs terrestres par des mesures indirectes et permet de comprendre et d'observer des objets qui ne seraient pas accessibles autrement¹. La géophysique interne regroupe de nombreuses disciplines : la gravimétrie, le géomagnétisme, la sismologie, les méthodes électriques ou électromagnétiques, qui se distinguent par les observables qu'elles considèrent et les phénomènes physiques qui les caractérisent. Parmi elles, la sismologie, qui concerne l'étude des séismes, les phénomènes associés et notamment la propagation des ondes sismiques, est probablement la plus riche en termes de résolution et de domaines d'applications.

La sismologie est une science jeune : ce n'est que vers la fin du XIX^e siècle que la sismologie observationnelle moderne voit le jour avec l'émergence des sismomètres et la première détection d'un télé-séisme, le séisme de Tokyo enregistré par von Rebeur-Paschwitz (1889) à Potsdam. La Terre n'est ni complètement opaque, ni parfaitement transparente aux ondes sismiques, elle est translucide (Virieux et al., 2012). Ainsi, les ondes sismiques qui se propagent en son sein portent, en plus des informations liées à la source, l'empreinte du milieu traversé qui réagit comme un filtre. En intégrant des techniques et des données provenant aussi bien des domaines de la physique, des mathématiques que de la géologie, l'étude de la propagation des ondes sismiques a permis, depuis quelques décennies, de dessiner une image précise des structures internes du globe, de la proche surface jusqu'au centre de la Terre.

Les prémices de notre compréhension de la structure interne de la Terre (Fig. 1) remontent aux travaux d'Oldham (1906)², qui construit un premier modèle sismologique de Terre après

1. À ce jour, le forage terrestre le plus profond est de 12 262 m, il s'agit du *Kola Superdeep Borehole* réalisé, entre les années 70 et 80, dans la péninsule de Kola en Russie. En mer, un forage de 7 740 m a été récemment entrepris dans la fosse du Japon ; il a battu le précédent record de 7 049 m dans la fosse des Mariannes.

2. L'hypothèse d'un noyau interne est déjà avancée par les travaux de Wiechert (1897), qui suppose que la Terre s'apparente à une météorite géante avec un noyau de fer-nickel.

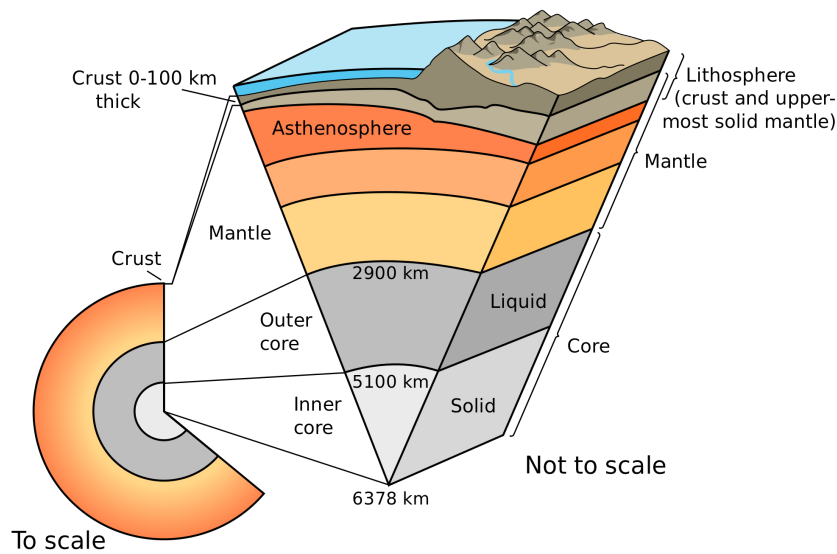


Figure 1 – Présentation de la structure radiale de la Terre. Source : USGS <https://pubs.usgs.gov/gip/dynamic/inside.html>.

avoir découvert l'existence du noyau par l'observation d'un retard sur l'arrivée des ondes S qu'il attribue à une discontinuité sismologique majeure à 3800 km de profondeur. Dans le même temps, [Mohorovicic \(1909\)](#) découvre la discontinuité entre la croûte et le manteau, qui porte aujourd'hui son nom, le Moho, en observant dans des données de sismique réfraction deux ondes P se propageant avec une vitesse différente : une première se propageant à 7.7 km/s (Pn) dans une couche profonde et une seconde à 5.5 km/s (Pg) dans une couche superficielle de 50 km d'épaisseur. [Gutenberg \(1914\)](#) étudie la « zone d'ombre » des ondes P et repositionne la discontinuité noyau-manteau d'Olhdam à 2900 km de profondeur. Plus tard, [Jeffreys \(1926\)](#), démontre que le noyau est fluide à partir de l'étude des marées terrestres et de l'absence de transmission des ondes S dans le noyau. Finalement, [Lehmann \(1936\)](#) découvre l'existence de la graine, au sein du noyau, par la détection d'ondes P' enregistrée par des stations situées dans la zone d'ombre des ondes P.

La connaissance précise de la structure radiale de la Terre arrive ensuite avec l'avènement de l'informatique et la détermination des modèles de références de Terre par l'inversion des modes normaux (PEM-C, PEM-O, PEM-A ([Dziewonski et al., 1975](#)) et PREM ([Dziewonski & Anderson, 1981](#))).

Les applications sismologiques ne se limitent pas aux questionnements fondamentaux abordés par la recherche académique. Les outils qu'elle procure, notamment pour la caractérisation des structures et propriétés du sous-sol, présentent des enjeux d'ordres aussi bien économiques, qu'environnementaux et sociétaux. L'imagerie sismique est notamment utilisée dans le domaine énergétique pour la recherche, le contrôle et l'exploitation d'énergies fossiles tels que les réservoirs d'hydrocarbures et les ressources minières ou le contrôle des sites géothermiques. Au niveau environnemental, elle intervient dans la surveillance des zones de stockage de CO₂ et des zones d'enfouissement de déchets radioactifs. Elle permet aussi, en génie civil et en ges-

tion des risques naturels, de caractériser l'instabilité des versants, de détecter des cavités ou de caractériser des failles. Enfin, la sismologie est un outil fondamental pour surveiller l'activité volcanique, caractériser l'aléa sismique, les sources sismiques et prévenir le risque tsunami ; sans oublier bien sûr la surveillance des essais nucléaires.

Ainsi, les applications de l'imagerie sismique couvrent un large spectre d'échelles d'observation, des premiers mètres du sous-sol à la globalité de la Terre. Ces observations se singularisent par les configurations expérimentales d'acquisition qui dépendent du type de sources utilisées, de la géométrie des réseaux de capteurs disponibles et de la distance de propagation entre les sources et les récepteurs considérés. En exploration sismique ou en sismologie à sources contrôlées (CSS), les sources et les capteurs sont généralement disposés à la surface³ de la cible que l'on souhaite imager. Les sources utilisées sont artificielles, elles sont généralement produites par des explosions (ou des vibrateurs) que l'on peut assez facilement caractériser et qui produisent des signaux relativement haute fréquence. L'acquisition est dans ces configurations entièrement contrôlée, les sources peuvent ainsi être répétées à des positions spatiales prédéfinies bien échantillonnées pour augmenter considérablement les données sismiques et les capteurs déplacés pour couvrir une zone géographique plus large. Ce type d'acquisition est idéal pour imager les structures de la proche surface au moyen des ondes réfléchies, les images produites par la migration des données de sismique réflexion sont alors hautement résolues et permettent d'accéder à la réflectivité du milieu. Pour imager des structures plus profondes et accéder aux grandes longueurs d'ondes de la cible étudiée, des dispositifs d'acquisition grands angles doivent être considérés. Ces dispositifs ayant une plus grande ouverture, ils permettent d'enregistrer des ondes réfractées et d'enrichir les données sismiques par des signaux plus basse fréquence propagés dans un régime de transmission. Aujourd'hui, le défi de l'imagerie dans ce type de configuration est de combler le déficit d'information basse fréquence pour imager les structures de la proche surface avec un spectre large.

Contrairement à l'exploration sismique, la sismologie considère principalement des sources naturelles non contrôlées : les tremblements de Terre. Selon les cibles considérées – la croûte, la lithosphère, le manteau supérieur ou la Terre globale – les géométries d'acquisition, les méthodes d'imagerie ainsi que la résolution des images du sous-sol varient. À l'échelle crustale, les sources correspondent à des séismes locaux de faibles magnitudes. Ils sont localisés à faibles profondeurs, inférieures à 30 km en général, et sont situés à proximité de réseaux permanents ou temporaires de stations généralement peu denses. L'information utilisée pour l'imagerie repose principalement sur la transmission et la mesure des temps de trajets des ondes propagées pour appliquer des méthodes de tomographie sismique. La résolution des images acquises est alors dépendante de la couverture des stations ainsi que de la positions des sources et elle diminue rapidement avec la profondeur. À l'échelle globale, les sources et les récepteurs sont réparties de manière éparse à la surface de la Terre. Les événements considérés sont de magnitudes élevées, supérieures à 5, pour pouvoir être détectés par les stations permanentes tout autour de la Terre. Les méthodes d'imagerie utilisées sont là encore principalement basées sur la tomographie sismique en transmission et recouvrent des modèles d'assez grandes longueurs d'ondes des hétérogénéités mantelliques. La résolution des images est alors très inégalement répartie

3. Ce n'est pas le cas en *time-lapse* où la cible est focalisée en profondeur (*target oriented imaging*), ni en VSP (*vertical seismic profiling*) où les sources et les capteurs sont disposés entre puits.

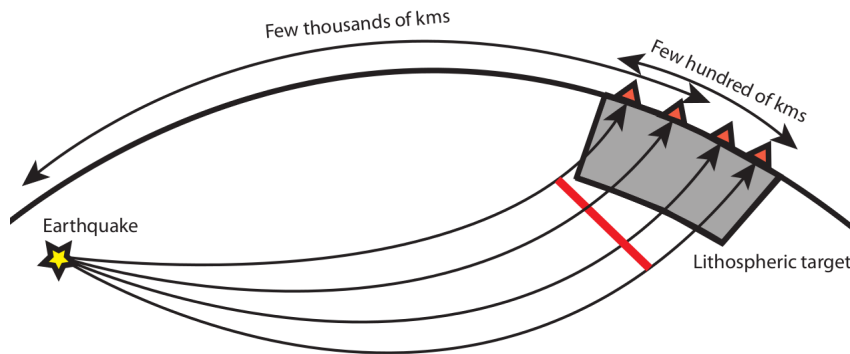


Figure 2 – Schéma présentant la géométrie de la configuration lithosphérique. Un téleséisme (étoile jaune) localisé à plusieurs milliers de kilomètres d’une cible lithosphérique (rectangle gris) génère un champ d’ondes qui se propage à l’intérieur de la Terre. Le champ d’ondes télé-sismiques en approche de la cible lithosphérique présente un front d’onde approximativement plan (trait rouge) qui éclaire la totalité de la cible lors de sa propagation. Les triangles oranges représentent les stations.

selon les zones géographiques. En effet, les séismes de fortes magnitudes sont localisés dans des zones géographiques très restreintes comme les frontières de plaque et les stations sont principalement déployées sur les continents. Il existe ainsi un déficit de couverture de stations dans les océans et notamment dans l’hémisphère sud. Le projet MERMAID porté par le laboratoire Géoazur vise à combler ce déficit par des réseaux sismiques de capteurs acoustiques flottants (Simons et al., 2006, 2009; Hello et al., 2011). Enfin, à l’échelle lithosphérique, l’acquisition est constituée de réseaux denses de capteurs (ex. USArray, AlpArray) disposés à l’aplomb de la cible étudiée. Les sources considérées sont des téleséismes localisés à plusieurs milliers de kilomètres de la zone d’étude comme présenté sur la figure 2. Le milieu est alors éclairé par des ondes approximativement planes qui entrent dans la cible par en dessous et l’éclairent dans sa totalité. Cette configuration est particulièrement adaptée pour l’étude de la croûte et du manteau supérieur, notamment au moyen de la tomographie de temps de trajet télé-sismique et de l’analyse des fonctions récepteur.

La lithosphère (Fig.1), d’une épaisseur d’environ 150 km⁴, constitue la partie rigide la plus externe de notre planète et qui bouge de manière cohérente avec le mouvement des plaques tectoniques (Artemieva, 2011). Elle est composée de la croûte et d’une partie du manteau supérieur (le manteau lithosphérique) et est caractérisée par un comportement élastique à l’échelle des temps géologiques. Deux types de lithosphère peuvent être distingués : la lithosphère océanique et la lithosphère continentale. Contrairement à la lithosphère océanique, qui est recyclée en environ 200 Ma, la lithosphère continentale est l’unique témoin permettant de retracer l’histoire de la Terre de part sa longévité. Certains cratons sont en effet âgés de plus de 3.5 milliards d’années. L’étude des structures et de l’évolution de la lithosphère est donc d’une importance majeure pour la compréhension des mécanismes qui gouvernent la tectonique des plaques, les processus géodynamiques et l’histoire géologique de la Terre : de la genèse des montagnes, au

4. Son épaisseur peut varier de 50 à 300 km.

recyclage de la lithosphère océanique jusqu'à l'étonnante stabilité des cratons archéens (Eaton et al., 2009; Fischer et al., 2010; Artemieva, 2011).

Méthodes d'imagerie lithosphérique

À l'heure actuelle, deux méthodes d'imagerie sont principalement utilisées en sismologie pour étudier la lithosphère. La première méthode, la tomographie sismique, est principalement fondée sur la théorie des rais et permet de déterminer des modèles de perturbation de vitesses en minimisant la différence entre les temps de propagation des ondes sismiques mesurés et ceux déterminés dans un modèle de Terre de référence. La seconde méthode, l'analyse des fonctions récepteur, permet par extraction du champ diffracté contenu dans les sismogrammes, d'accéder à la réflectivité du milieu situé sous une station. Ces deux méthodes possèdent des niveaux de détails différents. La tomographie, par l'exploitation de l'information en transmission, permet de reconstruire les grandes longueurs d'ondes des hétérogénéités du sous-sol. L'inversion ou la migration des fonctions récepteur permet au contraire d'accéder aux petites longueurs d'ondes du milieu lithosphérique. L'inversion de formes d'ondes, en expliquant le sismogramme complet peut théoriquement permettre de combler l'écart de résolution entre ces deux méthodes. Nous présentons dans les paragraphes suivants un état de l'art de la tomographie sismique, des méthodes liées aux fonctions récepteur et nous introduisons la technique d'inversion des formes d'ondes complètes.

La tomographie, temps de trajet – fréquences finies – adjoint

La tomographie de temps de trajet est une méthode d'imagerie qui permet, à partir de l'observation des temps de propagation des ondes sismiques, de contraindre les variations 3D des vitesses sismiques du sous-sol. Il s'agit d'un problème inverse d'estimation de paramètres, qui vise à déterminer les perturbations de lenteur du milieu pour expliquer les écarts de temps de trajets mesurés, en général, sur la première arrivée. Cette technique a été introduite par Aki & Lee (1976) pour la détermination des anomalies 3D de vitesses d'ondes P de la croûte dans la Bear Valley en Californie en utilisant les enregistrements des ondes sismiques émises par des séismes locaux⁵.

Le principe de la tomographie de temps de trajet est relativement simple, il est basé sur la théorie des rais et le principe de Fermat. Le problème direct qui permet de déterminer le temps de la première arrivée est résolu avec la théorie des rais⁶ pour un milieu discrétisé par blocs de vitesses homogènes. Au sein de ces blocs, les trajets des rais correspondent à des lignes droites. Le temps t de propagation d'une onde le long du rai $p(s)$ séparant une source et un récepteur est exprimé par l'intégrale (Tarantola & Nercessian, 1984) :

$$t(s) = \int_{p(s(l))} s(l) dl, \quad (1)$$

5. Cette référence présente une extension de la méthode de Geiger, qui permet de localiser des séismes locaux, à l'estimation des variations de vitesses. Ces deux paramètres sont ainsi estimés de manière simultanée.

6. Ou en résolvant l'équation iconale (Podvin & Lecomte, 1991; Noble et al., 2014).

où dl est un petit segment du rai et $s(l)$ est la lenteur (l'inverse de la vitesse) au sein d'un bloc. Dans un milieu de propagation perturbé, $s(l) + \delta s(l)$, cette équation devient

$$t(s + \delta s) = \int_{p(s(l)+\delta s(l))} s(l)dl + \int_{p(s(l)+\delta s(l))} \delta s(l)dl. \quad (2)$$

Selon le principe de Fermat, la trajectoire d'un rai est stationnaire et nous pouvons remplacer le domaine d'intégration le long des rais calculés dans le modèle perturbé par celui calculé le long des rais du modèle de référence

$$t(s + \delta s) = \int_{p(s(l))} s(l)dl + \int_{p(s(l))} \delta s(l)dl. \quad (3)$$

En soustrayant (3) de (1), nous déduisons l'expression de la différence de temps de trajet entre le milieu initial et le milieu perturbé en fonction des perturbations apportées au modèle de lenteur (sommées le long du rai) :

$$\delta t = t(s + \delta s) - t(s) = \int_{p(s(l))} \delta s(l)dl. \quad (4)$$

Sous forme discrète, l'équation (4) devient

$$\delta t = \sum_{i=1}^n \delta s_i \Delta l_i. \quad (5)$$

Pour un nombre n_d d'observations et un nombre n_m de paramètres, ce système linéaire peut se réécrire sous forme matricielle comme :

$$\Delta \mathbf{d} = \mathbf{G} \Delta \mathbf{m}, \quad (6)$$

où $\Delta \mathbf{d}$ est un vecteur de taille n_d contenant les temps de trajets différentiels observés. \mathbf{G} est la matrice des dérivées de Fréchet de taille $n_d \times n_m$ aussi appelée matrice de sensibilité qui contient la longueur des segments de chaque rai. \mathbf{m} est un vecteur de taille n_m contenant les paramètres de lenteur du milieu. Le problème inverse consiste à trouver des perturbations de lenteur qui minimisent l'erreur entre les temps de trajets observés et calculés :

$$\arg \min_{\mathbf{m}} \|\mathbf{d}_{obs} - \mathbf{d}_{cal}\|. \quad (7)$$

La solution au sens des moindres carrés (Menke, 2012) de l'équation (7) est obtenue en résolvant le système d'équations linéaires

$$\mathbf{G}^T \mathbf{G} \Delta \mathbf{m} = \mathbf{G}^T \Delta \mathbf{d}, \quad (8)$$

par exemple avec des algorithmes itératifs comme LSQR (Paige & Saunders, 1982; Nolet, 1985).

De part sa simplicité, la tomographie de temps de trajet est rapidement devenue populaire. Elle a été adaptée et améliorée aussi bien aux échelles locales, régionales que globales (Zelt & Barton, 1998; Spakman et al., 1993; Lippitsch et al., 2003). À l'échelle crustale, la tomographie

utilise des séismes locaux (LET pour *local earthquake tomography*) et implique, du fait de la proximité des sources et des stations, d'estimer conjointement la position des hypocentres et les paramètres du milieu. Les améliorations qui sont apportées à la LET sont principalement l'utilisation de méthodes de tracé de rai 3D (Lucio et al., 1995), la ré-expression du problème sous une forme non-linéaire et itérative (Rawlinson et al., 2010), et l'utilisation des doubles différences (Monteiller et al., 2005). Par ailleurs d'autres paramètres que les vitesses V_P peuvent être inversés, par exemple le rapport V_P/V_S (Vanorio et al., 2005).

Simultanément aux travaux d'Aki, la tomographie de temps de trajet est appliquée à l'échelle globale par Dziewonski et al. (1977) pour imager les hétérogénéités latérales (≈ 5000 km) de vitesse d'onde P dans le manteau. Ils inversent pour cela environ 700 000 temps de trajet en paramétrant le problème par une décomposition des anomalies de vitesse en harmoniques sphériques. Depuis lors, de nombreuses études ont été réitérées (Rawlinson et al., 2010; Liu & Gu, 2012), avec l'objectif d'augmenter la résolution des images multiparamètre du manteau en augmentant considérablement le nombre de données et en considérant d'autres arrivées sismologiques (PcP, PKP ou même des ondes S). La taille caractéristique des hétérogénéités imagées actuellement est de quelques centaines de kilomètres (Rawlinson et al., 2010).

Enfin, la tomographie de temps de trajet est directement appliquée aux échelles lithosphériques par Aki et al. (1977) en considérant des sources téléseismiques. Cette méthode est souvent nommée ACH dans la littérature, en référence aux noms de ses inventeurs. À partir des temps d'arrivées des ondes P téléseismiques du réseau norvégien Norsar, ils cartographient les anomalies latérales de vitesses dans cinq couches jusqu'à 126 km de profondeur elles-mêmes discrétisées par des blocs de 20 km. Comparablement à la LET, la méthode ACH a évolué et est appliquée en considérant des méthodes de tracé de rai 3D, des inversion non linéaires itératives, des corrections crustales et peut être adaptée pour estimer conjointement différents paramètres tels que les vitesses d'onde S, l'anisotropie ou la densité (Waldhauser et al., 2002; Lippitsch et al., 2003; Gao et al., 2004; Tiberi et al., 2003). Ces méthodes permettent d'imager des structures de l'ordre de 25 à 50 km.

La principale limitation de la tomographie de temps de trajet repose sur l'utilisation de la théorie des rais, c'est-à-dire, l'approximation asymptotique à fréquence infinie de l'équation d'onde (Červený, 2001). Ainsi, cette théorie n'est valide que si la taille caractéristique des hétérogénéités est grande par rapport à la longueur d'onde du signal propagé. Dans la réalité, les hétérogénéités ont des tailles variables et les ondes sismiques présentent un contenu fréquentiel fini. En effet, les ondes émises par un séisme ont un spectre fréquentiel dont la largeur est limitée, et elles sont atténuées au cours de leur propagation. La longueur d'onde peut ainsi varier de plusieurs centaines de mètres dans le domaine de l'exploration sismique à plusieurs centaines de kilomètres à l'échelle globale. Les temps de trajet mesurés ne sont pas uniquement sensibles au chemin emprunté par le rai, comme le suppose la théorie des rais, mais à toute hétérogénéité présente dans la première zone de Fresnel (Nolet, 2008). Cette dernière est un volume spatial, entourant le rai, au sein duquel l'énergie sismique interfère constructivement (Fig. 3b). Dans ce cas, les effets de diffraction contribuent de manière constructive au signal enregistré par une station. Autrement dit, un point diffractant se situe dans la première zone de Fresnel si les ondes diffractées par ce point sont enregistrées avec un décalage de phase inférieur ou égal à une demie période. Ainsi, des auteurs tels que Nolet (1987), Luo & Schuster (1991),

Woodward (1992), Marquering et al. (1998), Marquering et al. (1999) et Dahlen et al. (2000a) proposent de prendre en compte les effets de fréquences finies en tomographie.

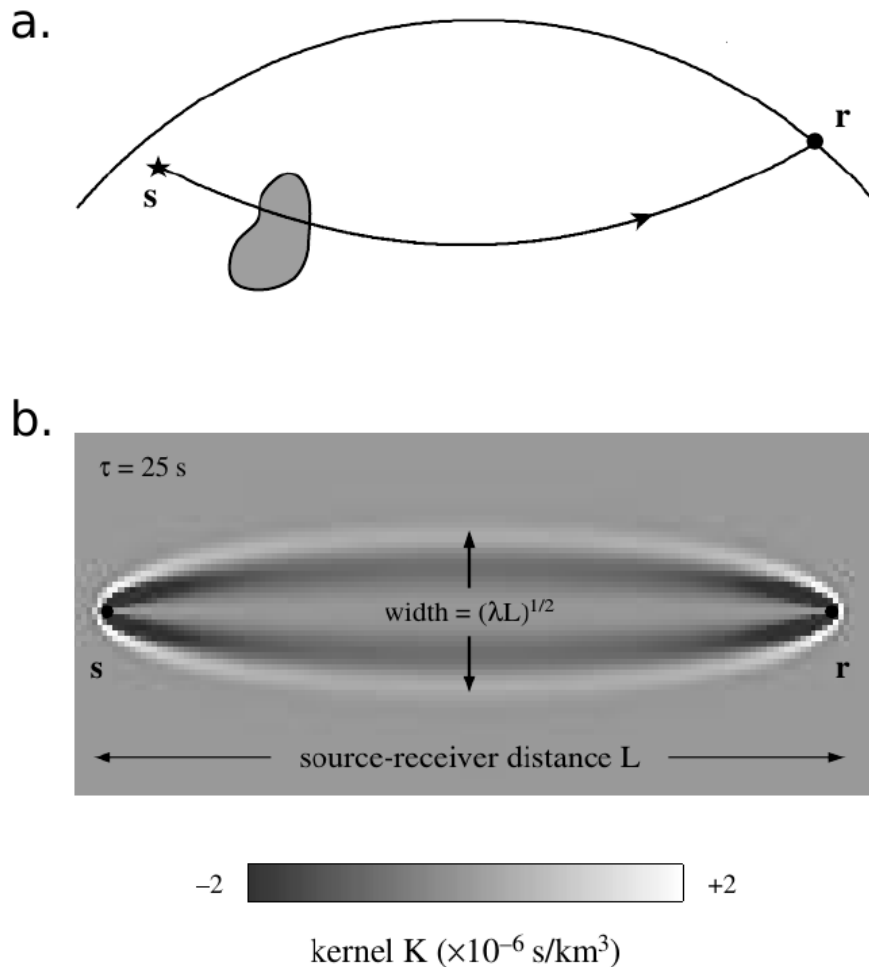


Figure 3 – Noyaux de sensibilité pour la tomographie des temps d’arrivées et la tomographie à fréquences finies. (a.) En tomographie de temps de trajet la sensibilité est distribuée le long du rai sismique. (b.) Noyau de sensibilité en tomographie à fréquences finies représentant la première zone de Fresnel. Il s’agit d’un noyau acoustique calculé pour une période caractéristique de 25 s. La forme du noyau est celle d’une « banane - doughnut ». La partie centrale du noyau qui correspond au trajet du rai ne présente aucune sensibilité. r et s représentent les positions de la source et du capteur. (Modifié d’après Hung et al. (2000)).

Plusieurs possibilités ont été proposées pour prendre en compte les phénomènes de diffraction se produisant dans la première zone de Fresnel. Husen & Kissling (2001) proposent une méthode pour modifier le noyau de sensibilité de la LET afin qu’elle rende compte de la zone de Fresnel. L’équation iconale est résolue deux fois, une depuis la source et l’autre depuis le récepteur, puis les temps de trajet obtenus sont sommés en chaque point du modèle. Le « fat ray » est finalement approximé en ne gardant que les points du milieu pour lesquels la différence

entre le temps calculé le long du rai et celui obtenu précédemment est inférieur à une demie période. Notons que, dans ce cas, la mesure reste un unique temps de trajet.

Une autre technique, la tomographie à fréquences finies (Marquering et al., 1998; Dahlen et al., 2000a), se base sur l'utilisation de la théorie des rais et détermine le noyau de sensibilité (en 3D) (Fig. 3) au moyen de l'approximation de Born (Hudson & Heritage, 1981). En tomographie à fréquences finies ces noyaux sont calculés pour des différences de temps de trajet mesurés par le maximum de l'inter-corrélation entre les formes d'ondes observées et celles calculées par l'approximation de Born⁷. L'équation (4) devient

$$\delta t = \int_V K(x) \frac{\delta c(x)}{c_0} d^3x, \quad (9)$$

où V est le volume d'intégration, $K(x)$ le noyau de sensibilité et $\delta c(x)/c_0$ est la perturbation de vitesse du milieu. Du fait de la linéarisation par l'approximation de Born, le problème inverse est linéaire et, tout comme la tomographie de temps de trajet, peut être résolu par des solveurs itératifs du type LSQR. Le gain en résolution apporté par cette méthode est montré par Montelli et al. (2004a) qui comparent la tomographie de temps de trajet à la tomographie fréquences finies. L'étude met en évidence l'importance des effets de diffractions et des effets de cicatrisation du front d'onde pour la résolution des images tomographiques du manteau. Elle leur permet de mettre en évidence une variété de panaches mantelliques difficilement identifiables auparavant, et de remarquer que la plupart d'entre-eux sont confinés dans la partie supérieure du manteau, accréditant l'hypothèse que la convection dans le manteau est partiellement séparée en deux régimes distincts (Montelli et al., 2004a,b).

Plus récemment, Tromp et al. (2005) ont réalisé un pas en avant vers l'inversion des formes d'ondes complètes en proposant de calculer le noyau de sensibilité ($K(x)$ dans l'équation (9)), pour le champ d'ondes complet, en résolvant numériquement l'équation d'onde. Le noyau de sensibilité peut alors être déterminé par construction explicite de la matrice de sensibilité (Chen et al., 2007). Une alternative, moins chère numériquement, consiste à évaluer le gradient de la fonction coût directement par la méthode de l'état adjoint (Tromp et al., 2005; Liu & Tromp, 2006; Fichtner et al., 2006a,b). Cette technique d'imagerie est généralement nommée « tomographie adjointe », du fait de l'utilisation de la méthode de l'état adjoint et de l'utilisation de temps de trajet mesurés par inter-corrélation sur une sélection de fenêtres temporelles contenant les principales arrivées (Maggi et al., 2009). La modélisation complète du champ d'ondes permet de rendre compte de tous les effets de propagation et rend le processus d'inversion non-linéaire. Le milieu est ainsi itérativement mis-à-jour jusqu'à la convergence du processus d'optimisation. L'application de la tomographie adjointe en Californie du Sud a permis de retrouver un modèle tomographique en ondes S de la croûte californienne très détaillé, qui révèle des structures très hétérogènes et qui permet d'envisager des études approfondies de l'aléa sismique ainsi que des interprétations géodynamiques poussées (Tape et al., 2009, 2010). Fichtner et al. (2009) et Bozdağ et al. (2011) réalisent un pas de plus vers l'inversion des formes d'ondes complètes en introduisant des fonctions coût non plus basées sur des inter-corrélations mais sur la phase et l'enveloppe des sismogrammes. Fichtner et al. (2010) appliquent la tomographie adjointe à l'échelle régionale et fournissent une image de l'anisotropie du manteau australasien.

7. La modélisation ne prend alors en compte que les effets de diffraction simple.

De même, la méthode est appliquée à l'échelle continentale par [Zhu et al. \(2012, 2015\)](#) pour imager les structures du manteau européen. Le coût important de la modélisation du champ d'ondes complet, requis pour l'application de la tomographie adjointe, rend la méthode difficile à appliquer à l'échelle globale. Néanmoins, grâce à l'augmentation des capacités de calcul, une première application de la tomographie adjointe a pu être réalisée à l'échelle globale par [Bozdağ et al. \(2016\)](#). Le modèle obtenu apporte des améliorations sur la résolution des zones de subductions connues et permet d'approcher la résolution des modèles obtenus aux échelles régionales.

Les fonctions récepteurs

Lorsqu'un champ d'ondes télésismique est propagée au sein d'un milieu lithosphérique, il interagit avec les structures et génère de multiples ondes secondaires diffractées, converties ou réfléchies. Un capteur situé en surface enregistre ainsi, en plus de l'arrivée principale, les arrivées de toutes ces phases lithosphériques secondaires qui caractérisent les structures du milieu traversé.

Pour une onde P incidente, les arrivées multiples les plus énergétiques correspondent aux conversions des ondes P en S (PS) sur les interfaces lithosphériques ainsi que les ondes PPS et PSS, réfléchies ou converties une première fois à la surface libre puis réfléchies une seconde fois sur l'interface lithosphérique (Fig. 4a). Les vitesses des ondes P étant plus rapides que celles des ondes S, ces ondes multiples arrivent avec un retard de plusieurs secondes par rapport au champ incident et sont enregistrées dans la coda de l'onde P (Fig. 4b). De manière similaire, lorsqu'une onde S incidente rencontre le milieu lithosphérique, elle génère des ondes converties S en P (SP). Dans ce cas, contrairement aux conversions PS, les conversions SP se propagent plus rapidement que l'onde S incidente, elles arrivent donc avant l'arrivée de l'onde S directe. En revanche, toutes les ondes diffractées une première fois à la surface libre arrivent en retard par rapport à l'onde S. Ainsi l'arrivée d'une onde S télésismique est précédée par ses multiples lithosphériques (SP) et suivie par ses multiples SSP et SPP.

Les temps d'arrivée des multiples lithosphériques sont dépendants de la profondeur des interfaces traversées, des vitesses du milieu et de l'angle d'incidence de l'onde. En isolant ces temps d'arrivée, il est possible de contraindre la profondeur des différents réflecteurs lithosphériques et de caractériser la réflectivité du milieu. Cependant, ces signaux sont généralement très faibles en comparaison de l'amplitude de l'onde P directe et ils peuvent être masqués par la signature de la source lorsque celle-ci est de durée assez longue. Le défi des méthodes en fonctions récepteur est d'analyser et d'extraire le signal des ondes converties pour obtenir des images précises des réflecteurs terrestres où se produisent ces conversions.

On ne s'intéresse à présent qu'aux ondes converties lors de la propagation d'une onde P incidente⁸. La calcul des fonctions récepteur, proposé par [Langston \(1979\)](#) et [Vinnik \(1977\)](#)⁹, permet d'isoler la contribution des conversions P en S des effets de la transmission en onde P

8. Pour les conversions S en P, voir [Farra & Vinnik \(2000\)](#) et [Yuan et al. \(2006\)](#).

9. Le premier pour des applications crustales et le second pour l'étude des interfaces dans la zone de transition du manteau.

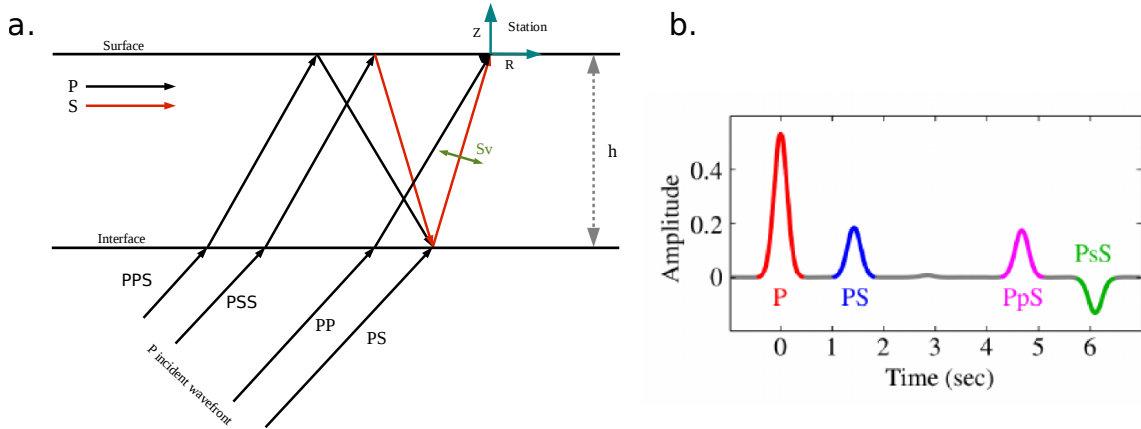


Figure 4 – Méthode des fonctions récepteur. (a.) Géométrie des chemins de rais téléseismiques de quelques conversions P en S sur une interface lithosphérique. Les flèches noires représentent des chemins en ondes P et les flèches rouges des chemins en ondes S. (b.) La fonction récepteur correspondante (Modifié d’après [Hetényi \(2007\)](#)).

et de la signature de la source. Le signal téléseismique enregistré $u_i(t)$ par une station peut se décomposer comme le produit de convolution entre $s(t)$ la signature temporelle de la source, $i(t)$ la réponse instrumentale du capteur et $g_i(t)$ la réponse impulsionnelle du milieu de propagation :

$$u_i(t) = s(t) * i(t) * g_i(t), \quad (10)$$

où i représente le signal enregistré dans les directions verticales, radiales et transverses (ZRT). Pour isoler la réponse impulsionnelle du milieu $g_i(t)$, [Langston \(1979\)](#) propose de déconvoluer les composantes horizontales (R et T) de la composante verticale (Z), généralement fenêtrée autour de la première arrivée, et qui est supposée être une bonne approximation du produit $s(t) * i(t)$. L’estimation de la fonction récepteur radiale est donnée en domaine fréquentiel par

$$RF_R(\omega) = \frac{U_Z(\omega) * U_R(\omega)}{U_Z(\omega) * U_Z(\omega)}, \quad (11)$$

où $*$ représente l’opérateur complexe conjugué. Alternativement, la fonction récepteur peut être estimée au moyen d’une déconvolution itérative en domaine temporel ([Ligorria & Ammon, 1999](#)) ou par des méthodes inter-correlation avec apodisations multiples en domaine fréquentiel ([Park & Levin, 2000](#)).

Les premières techniques d’inversion des fonctions récepteurs ont été présentées par [Owens et al. \(1984\)](#) et [Ammon et al. \(1990\)](#) en domaine temporel. Le problème inverse, non-linéaire ¹⁰, est formulé au sens des moindres carrés pour reconstruire un modèle 1D de vitesses d’ondes S. Le problème inverse est linéarisé autour d’un modèle initial et itérativement mis-à-jour pour prendre en compte les effets de non-linéarité. Le problème direct, la modélisation des fonctions récepteur, est quant à lui réalisé par une méthode de réflectivité qui permet de calculer des sismogrammes dans les milieux stratifiés ([Kennett, 1983](#)). Les sismogrammes différentiels, autre-

10. C’est à proprement parler un problème d’inversion de formes d’ondes.

ment dit les dérivés de Fréchet, sont alors calculées dans chaque couche en suivant le formalisme de Kennett (1983) (voir Randall (1989)). Ammon et al. (1990) étudient la non-unicité de l'inversion des fonctions récepteurs et montre qu'elles sont sensibles, non pas aux seules variations de vitesses d'ondes S, mais au produit de la vitesse des ondes S et de la profondeur. En outre, ces inversions sont sensibles aux contrastes d'impédances et non pas aux vitesses des couches elles-mêmes, elles ne permettent ainsi que de reconstruire les courtes longueurs d'ondes du milieu lithosphérique. Pour réduire la non-unicité et combler le manque de sensibilité aux grandes longueurs d'ondes des modèles de vitesses S, Julià et al. (2000) développent une méthode d'inversion conjointe des fonctions récepteurs et des courbes de dispersion des ondes de surface. Les développements récents concernant l'inversion des fonctions récepteurs proposent de modifier la fonction coût minimisée (Bodin et al., 2013), d'utiliser des méthodes d'optimisation globales et l'inférence bayésienne (Bodin et al., 2012, 2013) et d'inverser conjointement plusieurs jeux de données pour contraindre par exemple l'anisotropie (Bodin et al., 2016).

Une autre approche populaire pour le traitement des fonctions récepteur est l'opération de migration qui consiste à repositionner les arrivées temporelles sur la position des réflecteurs en profondeur. Pour ces opérations, il est nécessaire d'améliorer le rapport signal sur bruit des fonctions récepteur en les sommant en amont ou en aval du processus de migration (Owens et al., 1984; Zhu & Kanamori, 2000; Zhu, 2000). Zhu & Kanamori (2000) présentent une technique de sommation des fonctions récepteurs dans le domaine H-k pour estimer la profondeur d'une interface ainsi que le rapport V_P/V_S de la couche sous une station sismologique à partir des conversions PS. Cette méthode a l'avantage d'être peu sensible aux variations latérales de vitesses puisque les points de conversions sont proches de la station. Néanmoins, elle présente une forte ambiguïté entre l'estimation de la profondeur et le rapport de vitesses, qui peut toutefois être levée en prenant en compte les ondes converties à la surface libre (PPS, PSS). La profondeur et le rapport de vitesse sont déterminés par la valeur maximum de la sommation. Lorsque des réseaux denses de capteurs sont disponibles, il est possible de migrer les fonctions récepteur de chaque station pour obtenir une image de la continuité latérale des réflecteurs en profondeur. Zhu (2000) développe une technique de sommation des fonctions récepteur en profondeur, la sommation CCP (pour *common conversion point*). Premièrement, les chemins de rai sont déterminés au moyen d'un modèle de vitesse de référence, puis les amplitudes de chaque échantillon temporel des fonctions récepteur sont positionnées le long des rais, après correction des angles d'incidence. Ces positions correspondent aux points de conversions et sont déterminées à partir de la connaissance du temps différentiel de l'échantillon par rapport au temps d'arrivée de l'onde P directe, du modèle de vitesse et du rai considéré. L'image, dont l'amplitude représente les contrastes d'impédance, est finalement obtenue en sommant l'amplitude de tous les points de conversion. Cette méthode ne permet pas de prendre en compte les diffractions autres que les conversions PS. Elle ne permet pas non plus d'imager précisément des interfaces pentées, discontinues ou présentant de fortes topographies (Chen et al., 2005; Shang et al., 2012). Elle est néanmoins couramment utilisée de part sa simplicité de mise en œuvre. De récentes applications ont été présentées dans les Pyrénées et les Alpes occidentales par Chevrot et al. (2015) et Zhao et al. (2015)

Enfin, d'autres approches inspirées de l'exploration sismique ont été proposées pour migrer les fonctions récepteur en profondeur. Nous pouvons citer les méthodes basées sur la migration par renversement temporel (Fink, 1993) proposée pour les fonctions récepteurs par Shang et al.

(2012), celles basées sur la migration profondeur de Kirchhoff (Cheng et al., 2016) ou encore des méthodes basées sur la résolution de l'équation d'onde (Chen et al., 2005).

Migration du champ d'ondes diffractées

Plutôt que de migrer les fonctions récepteurs pour fournir des images de la réflectivité, Bostock & Rondenay (1999); Bostock et al. (2001); Shragge et al. (2001); Rondenay et al. (2001, 2005) ont directement adapté à des données télésismiques des méthodes de migration du champ d'ondes diffractées généralement utilisées en exploration sismique (Jin et al., 1992). En exploration sismique, ce champ d'ondes diffractées (c'est-à-dire l'ensemble des ondes qui ne seraient pas reproduites par un milieu de référence lisse dans l'hypothèse d'un régime de diffraction simple) correspond aux ondes réfléchies primaires renvoyées par les contrastes d'impédance du milieu. Ce champ d'onde diffracté peut être modélisé par linéarisation du problème direct en couplant l'approximation de Born à des fonctions de Green asymptotiques. L'utilisation de tracé de rai dynamique pour calculer différents attributs sismiques tels que les temps, les amplitudes géométriques et les angles est utile d'une part pour sélectionner des phases spécifiques lors de la migration (par exemples, des ondes converties PS ou des multiples de la surface libre) et pour développer l'expression analytique ou approchée d'opérateurs intervenant dans le formalisme du problème inverse (diagrammes de rayonnements, opérateur Hessien) (Lambaré et al., 1996; Thierry et al., 1996). Après linéarisation du problème direct, le problème inverse d'inversion des formes d'ondes du champ d'ondes diffracté est linéaire et peut être formulé comme la minimisation de l'écart entre les données diffractées enregistrées δd_{obs} et modélisées δd_{cal} pour mettre à jour un modèle de perturbation δm représentant les courtes longueurs d'onde du milieu (Jin et al., 1992) (Fig. 5)

$$\min_{\delta m}(\mathbf{x}) \sum_s \sum_r \int_t \| (\delta d_{cal}(s, t; r) - \delta d_{obs}(s, t; r)) \| dt \Delta s \Delta r, \quad (12)$$

où $\delta d_{cal} = \mathbf{B}_0 \delta m$ et \mathbf{B}_0 est l'opérateur linéaire de Born dépendant des fonctions de Green calculées dans le milieu de référence depuis la position des sources s et des capteurs r . L'expression du modèle de perturbation est alors donnée par

$$\delta m(\mathbf{x}) = - [\mathbf{B}_0^\dagger \mathbf{B}_0]^{-1} \mathbf{B}_0^\dagger \Delta \delta d. \quad (13)$$

où $\mathbf{B}_0^\dagger \mathbf{B}_0$ est l'opérateur Hessien, $\mathbf{B}_0^\dagger \Delta \delta d$ le gradient de la fonction coût et † indique l'opérateur adjoint.

Ce processus peut être itéré pour mettre à jour les perturbations par sommation et améliorer l'ajustement de l'amplitude du champ d'ondes diffracté tout en gardant le modèle de référence fixe en accord avec les hypothèses associées à la linéarisation du problème direct.

Ces approches ont été adaptées aux configurations lithosphériques par Bostock & Rondenay (1999); Bostock et al. (2001); Shragge et al. (2001); Rondenay et al. (2001, 2005). L'originalité de leur approche réside dans l'utilisation de la surface libre, dont la géométrie est connue a priori et de ce fait utilisable lors du tracé de rai pour modéliser les phénomènes de réflexion et

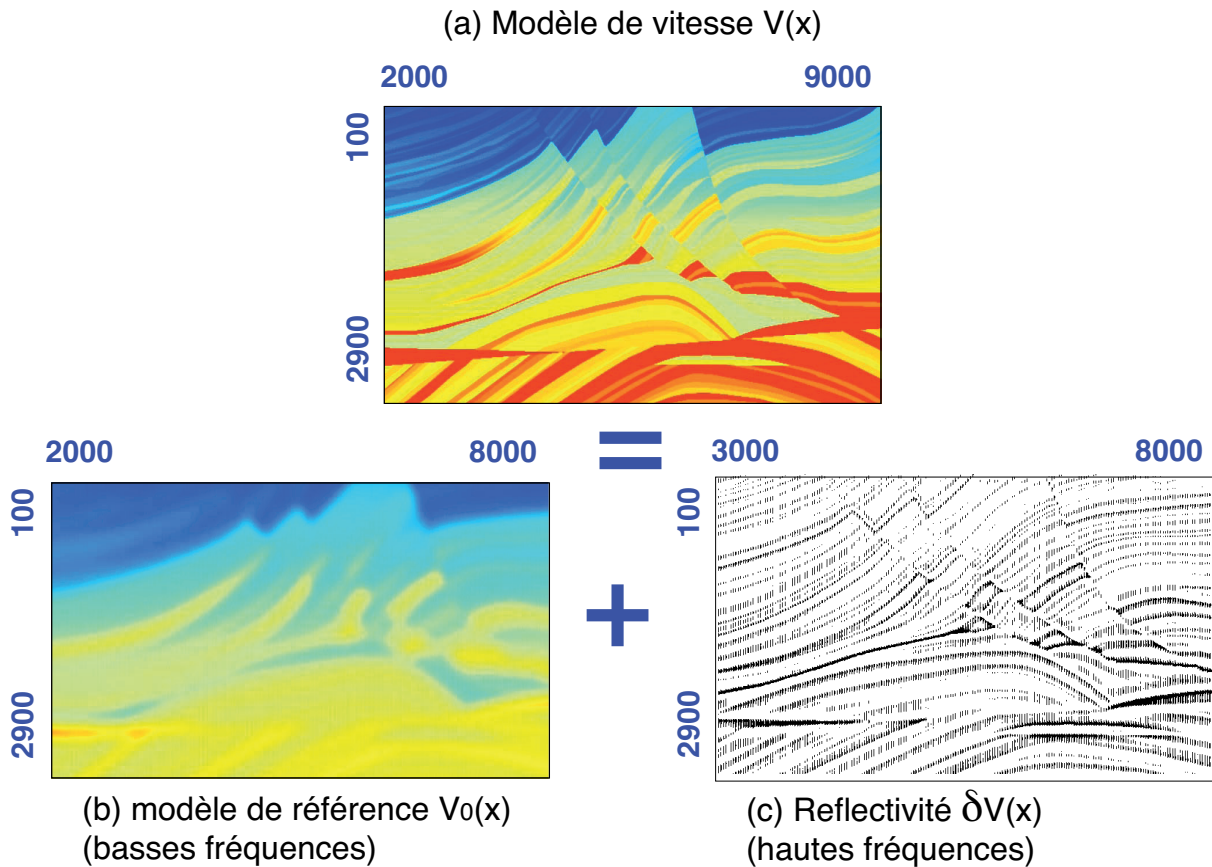


Figure 5 – Découplage d'échelles entre le modèle de référence et le modèle de réflectivité issu de la linéarisation du problème (approximation de *Born*) autour d'un modèle de référence (Chauris, 2000). Le modèle de vitesse recherché (a) peut être décomposé en un modèle de référence basse fréquence $V_0(x)$ et un modèle de réflectivité haute fréquence $\delta V(x)$. Le modèle de référence, qui peut être obtenu par exemple par tomographie des temps de première arrivée, explique la cinématique des ondes. La partie haute fréquence correspond à la réflectivité du milieu et peut être obtenu par une méthode de migration du champ diffracté qui consiste à refocaliser à la position des réflecteurs les arrivées réfléchies et diffractées. D'après Pageot (2012).

de conversion qui s’y produisent. Elle agit ainsi comme une source secondaire ré-émettant des ondes P et S dans le milieu.

Cette configuration permet de reproduire la géométrie d’une acquisition de sismique réflexion de surface et permet l’imagerie des réflecteurs lithosphériques par migration des multiples de la surface libre.

L’inversion des formes d’ondes complètes

Contrairement à la plupart des méthodes d’imagerie présentées précédemment, l’inversion de formes d’ondes complètes (FWI, pour *full-waveform inversion* (Lailly, 1983a; Tarantola, 1984b; Gauthier et al., 1986b; Mora, 1987; Woodward, 1992)) tente de reconstruire des modèles quantitatifs du sous-sol en exploitant toute l’information contenue dans les sismogrammes observés, c’est-à-dire, l’enregistrement complet (Virieux & Operto, 2009). Cette utilisation des formes d’ondes permet de tenir compte simultanément des informations liées aux temps d’arrivées, aux variations de phases ainsi qu’aux variations d’amplitudes des différentes ondelettes contenues dans le signal sismique qu’elles soient transmises, réfléchies ou converties. L’application de la FWI aux données télésismiques permet donc, par une méthode intégrée, de tirer partie de toutes ces informations et d’envisager de combler le hiatus de résolution entre les images obtenues par migrations/inversion des fonctions récepteur et celles obtenues par l’application de la tomographie sismique. Elle peut potentiellement permettre d’atteindre une résolution proche de la demie longueur d’onde propagée lorsque le dispositif d’acquisition permet d’enregistrer les réflexions verticales (ou à déport source-récepteur nul).

Le principe de fonctionnement de la FWI, schématiquement représenté sur la figure 6, est le suivant : étant donné un modèle initial du sous-sol on cherche à estimer un nouveau modèle de paramètres qui minimise l’écart, échantillon par échantillon, entre les sismogrammes enregistrés et ceux modélisés à partir du modèle de paramètre. L’application de la FWI requiert ainsi de résoudre le problème de propagation des ondes de manière complète pour rendre compte de tous les effets de propagation et notamment des effets de diffractions multiples. Néanmoins, la comparaison directe des formes d’ondes observées et calculées rend le problème inverse fortement non-linéaire. Ce dernier est de plus mal-posé du fait de la mise à jour d’un grand nombre de paramètres ainsi que la prise en compte de paramètres physiques différents. La grande taille du problème inverse ainsi que le coût de calcul de la méthode rendent inenvisageable l’utilisation de méthodes d’optimisation globales pour la résolution du problème inverse. Il est donc linéarisé autour d’un modèle initial (Lailly, 1983b; Tarantola, 1984b), et résolu de manière itérative par des méthodes d’optimisation locales. La méthode est donc très sensible au modèle initial utilisé qui doit être assez précis pour localiser le minimum global de la fonction coût que l’on cherche à minimiser. Précisons que contrairement aux approches de tomographie à fréquences finies ou de migration des champs diffractés pour lesquelles le problème direct est linéarisé et seuls les effets de diffractions simples sont pris en compte, la FWI par la mise à jour progressive du milieu, peut théoriquement expliquer les effets de diffraction multiple.

Historiquement, la FWI a été développée par la communauté de l’exploration sismique par Lailly (1983b) et Tarantola (1984b) qui cherchaient à reformuler le problème de migration des données de sismique réflexion sous la forme d’un problème inverse linéarisé. Le problème fut

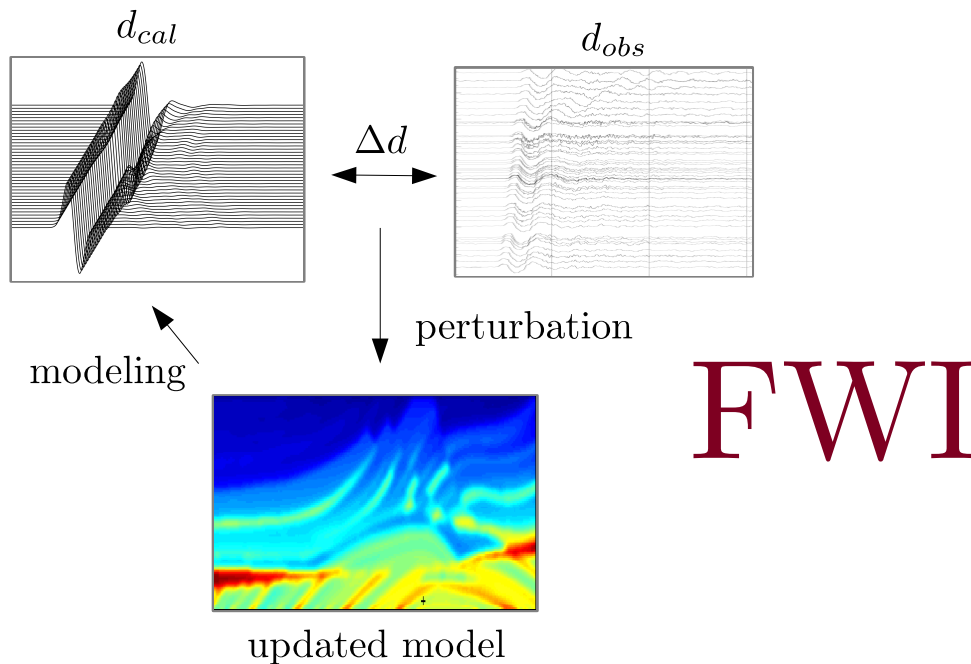


Figure 6 – Schéma présentant le concept de l’inversion de formes d’onde complètes (FWI). La comparaison échantillon par échantillon des données modélisées d_{cal} dans un modèle de Terre initial et des données observées d_{obs} crée un résidu Δd que l’inversion cherche à minimiser. Le processus d’optimisation détermine les perturbations à apporter au modèle initial qui est progressivement mis-à-jour, de manière itérative, jusqu’à atteindre un certain critère de convergence sur les résidus.

d’abord formulé dans le domaine temporel pour des milieux acoustiques (Tarantola, 1984b; Gauthier et al., 1986a) puis a été étendue aux milieux élastiques par Mora (1987). Gauthier et al. (1986b) ont montré la faisabilité de la méthode au moyen de tests synthétiques 2D ainsi que l’importance de l’information en transmission sur le signal par rapport à celle en réflexion. Cependant, le coût numérique de la méthode, sa forte non-linéarité et son caractère mal posé rendent difficile l’application de la FWI à cette époque. Par la suite, Pratt (1990a) et Pratt & Worthington (1990) reformulent la FWI dans le domaine fréquentiel. L’efficacité de l’approche fréquentielle en 2D (Sirgue & Pratt, 2004) a ensuite permis de mettre en œuvre plusieurs applications, aussi bien terrestres (Ravaut et al., 2004) que marines (Dessa et al., 2004; Prioux et al., 2013c,b; Operto et al., 2015), sur des données réelles en considérant des milieux visco-acoustiques ou visco-élastiques (Brossier et al., 2009b). tandis que Operto et al. (2015) ont démontré la faisabilité de l’approche fréquentielle en 3D pour des dispositifs d’acquisition de type OBC (Ocean Bottom Cable) pour des milieux visco-acoustiques. Durant la dernière décennie, les approches temporelles se sont néanmoins réimposées dans le domaine de l’exploration sismique pour effectuer l’imagerie en 3D en raison de leur versatilité pour implémenter des algorithmes massivement parallèles de modélisation sismique et prendre en compte différents types d’acquisition (streamer, OBC, OBN, terrestre) et de physique des ondes (e.g., Warner et al. (2013)).

Ces dernières années, les premières tentatives d’application de la FWI aux données télé-

sismiques ont été entreprises pour améliorer la résolution des modèles lithosphériques. Dans ce type d'étude, les données inversées correspondent aux formes d'ondes de l'onde P et de ses multiples diffractés par le milieu lithosphérique situés dans les premières dizaines de secondes de coda. Les premières études ont été réalisées en 2D ou 2.5D (Roecker et al., 2010; Pageot et al., 2013; Baker & Roecker, 2014) et implémentées en domaine fréquentiel (Pratt & Worthington, 1990). Le choix d'une approche fréquentielle a été motivée car les nombreuses modélisations, requises pour la FWI, peuvent être très efficacement réalisées par des techniques d'élimination de Gauss lorsqu'un nombre conséquent de sources est considéré et que peu de fréquences discrètes sont prises en compte (Pratt, 1999). Les particularités du milieu lithosphérique et des sources télésismiques, qui sont localisées en dehors de la cible, permettent de mettre à jour les paramètres uniquement dans le domaine spatial de la cible lithosphérique. La source est alors implémentée au moyen de techniques hybrides permettant d'injecter des champs d'ondes calculés analytiquement (Roecker et al., 2010; Pageot et al., 2013), par exemple des ondes planes. Pageot et al. (2013) réalisent une analyse paramétrique de la FWI télésismique en domaine fréquentiel pour des milieux 2.5 D. Ils concluent néanmoins que l'approche fréquentielle n'est pas adaptée pour l'application de la FWI car elle nécessite un échantillonnage très fin des fréquences pour compenser le déficit d'éclairage angulaire des cibles lithosphériques. Ce déficit est directement relié à l'utilisation d'ondes quasiment planes, et à un petit nombre de sources considérées. Ils proposent alors de reformuler la FWI en domaine temporel dans le cadre des applications télésismiques.

L'application de la FWI pour des milieux 3D en domaine temporel est ensuite investiguée par Tong et al. (2014b). Dans cette étude, la source télésismique est implémentée au moyen d'une technique d'injection de champs d'ondes proposée par Monteiller et al. (2013) qui couple une méthode de modélisation F-K (Zhu & Rivera, 2002) avec une méthode de modélisation aux éléments spectraux (Komatitsch & Tromp, 1999b). Une implémentation de la FWI télésismique est aussi présentée par Monteiller et al. (2015) qui considèrent l'injection d'un champ d'ondes télésismique calculé dans la Terre globale par le solveur DSM (Direct Solution Method) (Takeuchi et al., 2000). Ils démontrent l'efficacité de la méthode en reconstruisant un modèle synthétique représentatif de la lithosphère pyrénéenne au moyen de quatre sources télésismiques. La faisabilité et l'efficacité de la FWI à l'échelle lithosphérique est ensuite confirmée par Wang et al. (2016), en inversant les données de cinq événements télésismiques enregistrés par l'expérience PYROPE, une ligne ultra-dense de capteurs larges bandes déployée dans les Pyrénées orientales. Les modèles de vitesse d'ondes P et S reconstruits révèlent la subduction de la croûte ibérique sous la plaque européenne ainsi que la présence d'un morceau de manteau européen serpentinisé situé à faible profondeur sous le bassin de Mauléon.

Objectifs de cette thèse

Les interprétations géodynamiques requièrent des images de plus en plus fines des structures lithosphériques pour permettre de contraindre, infirmer ou confirmer des hypothèses sur l'évolution et la dynamique des objets géologiques. Cette perpétuelle quête de résolution a motivé, ces dernières années, les déploiements de réseaux toujours plus denses de capteurs tels que USarray ou AlpArray. En parallèle, l'augmentation constante des capacités de calcul rend aujourd'hui

possible l'application de l'inversion de formes d'ondes élastiques, en domaine temporel et pour des milieux tridimensionnels. Ensemble ces deux avancées permettent d'envisager l'application de la FWI aux données télésismiques et ainsi d'accroître sensiblement la résolution des images lithosphériques.

Cette thèse s'inscrit dans ce contexte et vise à développer un algorithme d'inversion de formes d'ondes, adapté à la configuration télésismique, pour caractériser les structures élastiques tridimensionnelles de la lithosphère et, en particulier, celles des Alpes Occidentales.

Pour cela, nous considérons trois grands axes de travail.

1. Développer une méthode de simulation efficace de la propagation 3D des ondes télésismiques dans les cibles lithosphériques pouvant générer des formes d'ondes avec un contenu fréquentiel proche du hertz.
2. Formuler le problème d'inversion de formes d'ondes multiparamètre. Développer l'algorithme et définir les stratégies de mise en application de la méthode. Étudier les potentiels et les limitations de la méthode au vu des caractéristiques de la configuration lithosphérique.
3. Confronter la méthode à un jeu de données réel, les données télésismiques recueillies par la campagne CIFALPS, pour vérifier le gain apporté par la méthode développée et compléter les connaissances actuelles sur la lithosphère des Alpes Occidentales.

Contours du manuscrit

Suivant les trois axes définis dans la section précédente, nous organisons ce manuscrit en trois parties distinctes qui traitent respectivement du problème direct, du problème inverse et de l'application aux données réelles.

Dans la première partie, nous présentons quelques aspects théoriques de la propagation des ondes élastiques ainsi que les méthodes numériques permettant de les résoudre pour des milieux arbitrairement hétérogènes. Cette partie est divisée en deux chapitres. Le premier chapitre est dédié à la présentation du phénomène de propagation d'ondes et à la description des champs d'ondes télésismiques. Les équations de l'élastodynamique y sont brièvement présentées, ainsi que quelques solutions analytiques qui nous permettront de valider les outils de modélisation numérique. Nous passons en revue les différents régimes de diffusion et présentons le cadre théorique de l'approximation de Born, utile à la compréhension de certains mécanismes de l'inversion de formes d'ondes. Enfin, nous terminons par une description du champ d'onde télésismique. Au cours du second chapitre, nous abordons les méthodes de modélisation numérique envisageables pour la résolution du problème direct. Nous présentons en détail la méthode de modélisation aux éléments-finis spectraux que nous utiliserons lors du processus d'inversion. Dans un second temps, nous présentons une technique d'injection de champs d'ondes, qui permet de modéliser, à faibles coûts, la propagation d'un champ d'onde télésismique dans une cible lithosphérique. Nous considérons pour cela le couplage des solutions fournies par AxiSEM (Nissen-Meyer et al., 2014), une méthode aux éléments spectraux 2D qui permet de simuler la propagation 3D des ondes à l'échelle globale, avec notre logiciel de modélisation local.

La seconde partie, quant à elle, traite exclusivement du problème inverse d'inversion de formes d'ondes, et de son application dans le cas spécifique des applications lithosphériques. Elle est constituée de deux chapitres. Le premier présente la théorie et les méthodes de résolution des problèmes inverses et précise le caractère non-linéaire et mal-posé de l'inversion de formes d'ondes. Nous posons ensuite l'inversion de formes d'ondes comme un problème aux moindres carrés linéarisé visant à minimiser une fonction coût qui prend en compte la différence, échantillon par échantillon, entre les formes d'ondes observées et modélisées. La résolution du problème est alors effectuée au moyen d'un algorithme local d'optimisation de type quasi-Newton (l-BFGS) qui nécessite de déterminer le gradient de la fonction coût que nous exprimons par la méthode de l'état adjoint. Une interprétation physique des structures du gradient est ensuite proposée. Enfin, nous présentons divers aspects, plus techniques, liés à l'implémentation de l'algorithme de la FWI. Le second chapitre présente, sous forme d'une publication scientifique, une étude paramétrique de la FWI en configuration téléseismique effectuée à l'aide d'un modèle lithosphérique réaliste représentatif des structures des Alpes Occidentales. Dans cette publication, nous analysons les effets de la paramétrisation, du choix du modèle initial et de l'acquisition sur les résultats de l'inversion des formes d'ondes téléseismiques. Nous montrons que des modèles de référence de la terre globale, tels que PREM, sont suffisamment précis pour garantir la convergence de l'inversion vers un modèle lithosphérique pertinent. Nous montrons également qu'il est important de distinguer les paramétrisations utilisées lors du processus d'optimisation de celles utilisées pour la visualisation des résultats de l'inversion. Notre étude indique que la paramétrisation faisant intervenir les paramètres élastiques est la plus naturelle pour recouvrir avec un spectre large les petites et grandes longueurs d'ondes des modèles lithosphériques. Nous montrons également que la recombinaison *a posteriori* de l'impédance des ondes P et S est celle qui fournit les images les mieux résolues, et est en conséquence la plus utile pour l'interprétation géodynamique. Finalement, nous insistons sur l'importance du déploiement de réseaux bi-dimensionnels de stations qui permettent de bénéficier d'une résolution verticale et latérale plus importante que celles obtenues pour des réseaux linéiques, même ultra-denses.

Enfin, la troisième partie de cette thèse porte sur l'application de la méthode aux données réelles de la campagne *CIFALPS* déployée dans les Alpes Occidentales. Un court chapitre introductif présente l'état des connaissances géodynamiques actuelles de la chaîne Alpine, les résultats géophysiques majeurs obtenus ces dernières années ainsi que les questionnements subsistants sur les structures lithosphériques de l'arc alpin et les processus géodynamiques qui leur sont associés. Le second chapitre, sous forme d'un article scientifique, présente les résultats de l'application de la FWI aux données *CIFALPS*. Nous montrons que la FWI permet d'accéder à des modèles multi-paramètres fiables de densité et de vitesses d'ondes élastiques. Nous mettons en évidence les indices clés de la subduction et de la collision continentale : les géométries particulières du corps d'Ivrée ainsi que le pendage, sous les Alpes internes, du Moho jusqu'à des profondeurs de l'ordre de 75 km qui témoignent de l'enfouissement à grande profondeur de la lithosphère européenne sous la lithosphère adriatique. La présence d'une anomalie de vitesse lente sur le modèle de vitesse S – le mieux résolu – indiquerait la position du détachement du panneau plongeant européen soutenant ainsi l'hypothèse de son déchirement au court de la collision continentale.

Je conclus ce manuscrit par une synthèse des principaux résultats et par les perspectives que

ce travail a permis de dresser.

Première partie

Modélisation de la propagation des ondes télésismiques

Chapitre 1

Généralités sur la propagation des ondes sismiques

Sommaire

1.1	Qu'est-ce qu'une onde ?	40
1.2	Les équations de l'élastodynamique	41
1.3	Ondes propagées dans un milieu élastique	44
1.3.1	Ondes de volume	44
1.3.2	Ondes de surface	47
1.4	Quelques solutions analytiques	48
1.4.1	Les fonctions de Green	48
1.4.2	Onde plane dans un demi-espace homogène	49
1.5	Ondes sismiques en milieux hétérogènes	51
1.5.1	La diffusion des ondes élastiques	52
1.5.2	Diffusion simple par l'approximation de Born	54
1.6	Caractéristiques du champ d'onde téléseismique	58
1.6.1	Les sources téléseismiques	58
1.6.2	Ondes générées à l'échelle globale, phases téléseismiques	60
1.6.3	Multiples lithosphériques	61
1.6.4	Données synthétiques pour un modèle lithosphérique des Alpes Occidentales	64

Introduction

Dans ce chapitre, nous abordons les notions de physiques permettant de décrire mathématiquement le phénomène de propagation des ondes élastiques à l'intérieur de la Terre. Ces développements nous permettront d'analyser et de décrire le comportement du champ d'onde sismologique depuis sa nucléation à l'hypocentre d'un séisme jusqu'à son enregistrement par des sismomètres. Ces connaissances sont importantes dans le cadre de l'inversion de formes d'ondes puisque ce sont sur ces notions mathématiques que reposent les méthodes de modélisation numériques de champs d'ondes élastiques que nous présenterons dans le chapitre suivant. La description physique du phénomène de propagation des ondes nous permettra de comprendre les interactions entre les ondes et les différentes structures de la Terre et d'analyser l'empreinte de ces interactions dans les enregistrements téléseismiques. En outre, nous présenterons quelques solutions analytiques correspondant à des exemples canoniques du problème de propagation des ondes qui nous permettront de valider les outils de modélisation numérique.

1.1 Qu'est-ce qu'une onde ?

Une onde¹ est définie comme la propagation d'une perturbation provoquant une variation réversible des propriétés physiques du milieu traversé. Ainsi, une onde permet de propager de l'énergie au sein d'un milieu sans transporter de matière. On en distingue deux types :

- les ondes électromagnétiques (ex. la lumière) qui propagent un champ électromagnétique sans support matériel connu,
- les ondes mécaniques (ex. les ondes sonores) qui propagent un champ de déformation - déplacement mécanique dans un milieu matériel.

Ces ondes, qui sont des phénomènes oscillatoires périodiques², sont caractérisées par des grandeurs qui dépendent du milieu traversé, de l'espace et du temps (figure 1.1). On définit particulièrement la célérité (1.1) qui correspond à la vitesse de propagation de l'onde en introduisant les concepts de longueur d'onde λ et de fréquence f qui correspond à l'inverse de la période T .

$$c = \lambda f = \frac{\lambda}{T}. \quad (1.1)$$

Conjointement à ces grandeurs, nous pouvons définir la pulsation ou fréquence angulaire ainsi que le nombre d'onde angulaire,

$$\omega = \frac{2\pi}{T} \quad \text{et} \quad k = \frac{2\pi}{\lambda}, \quad (1.2)$$

qui représentent respectivement le temps et la distance nécessaires pour effectuer un cycle complet.

1. Du latin *unda* qui signifie « eau en mouvement »

2. Une phénomène transitoire pouvant se décomposer en une somme de fonctions périodiques.

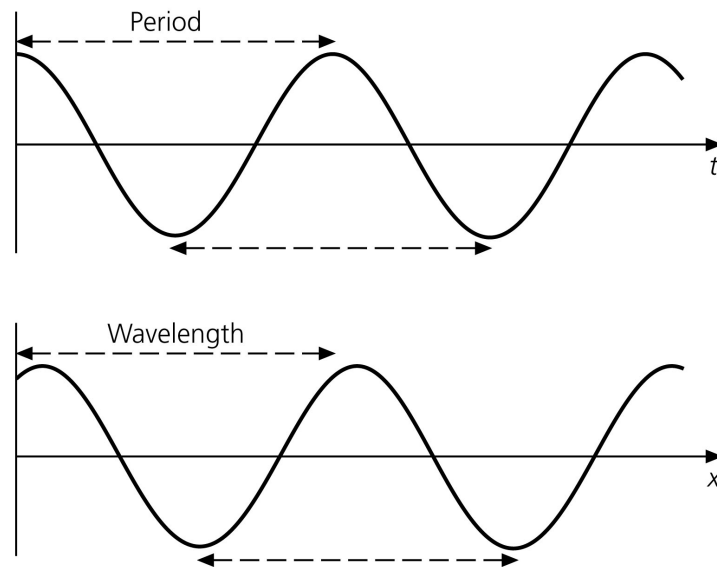


Figure 1.1 – Différence entre période et longueur d’onde (d’après Stein & Wysession (2009)).

Les ondes sismiques, en particulier, appartiennent à la famille des ondes mécaniques élastiques. Elles sont générées sous l’action d’un séisme ou d’une source artificielle, et sont propagées dans toutes les directions³ à l’intérieur de la Terre. L’enregistrement d’un sismomètre contient en conséquence une multitude d’informations liées à la source sismique – le séisme – et au milieu de propagation, tous deux de natures très complexes. Pour appliquer la méthode d’inversion de formes d’onde, il est nécessaire de simuler le signal sismique enregistré par les stations sismologiques. Autrement dit, il nous faut un moyen de modéliser le phénomène de propagation d’onde : son excitation et ses interactions avec le milieu terrestre le long de son parcours.

1.2 Les équations de l’élastodynamique

Le phénomène dit de propagation des ondes élastiques est gouverné par les lois physiques régissant l’élasticité linéaire (Duvaut, 1990), dérivées de la mécanique des milieux continus, appelées équations de l’élastodynamique. La mécanique des milieux continus repose sur l’hypothèse que le comportement macroscopique d’un milieu matériel continu – une poutre, une voiture ou dans notre cas la Terre – est régi à l’échelle microscopique par une succession discontinue de particules élémentaires infiniment proches⁴. Ces subdivisions de la matière sont choisies de sorte que leurs tailles caractéristiques soient assez réduites pour que des variations en leur sein puissent être négligées.

À partir de cette définition, nous pouvons reformuler la seconde loi de Newton dans le cadre

3. Avec toutefois une certaine directivité décrite par un diagramme de rayonnement.

4. Cette notion est fondamentale puisqu’elle implique que les quantités physiques considérées (densité, forces, déplacements) peuvent être représentées par des fonctions continues de l’espace et du temps, et différentiables. Elle nous permet ainsi d’avoir recours aux outils mathématiques de l’analyse fonctionnelle

d'un milieu continu et aboutir à l'équation de *conservation de la quantité de mouvement*. Pour un élément matériel situé à une position \mathbf{x} et soumis à une force volumique externe $\mathbf{f}(\mathbf{x}, t)$, elle s'exprime par la relation :

$$\rho(\mathbf{x})\partial_t^2\mathbf{u}(\mathbf{x}, t) = \nabla \cdot \boldsymbol{\sigma}(\mathbf{x}, t) + \mathbf{f}(\mathbf{x}, t), \quad (1.3)$$

où $\rho(\mathbf{x})$ correspond à la densité du milieu et $\mathbf{u}(\mathbf{x}, t)$ au déplacement particulaire. Contrairement à la seconde loi de Newton, cette formulation fait apparaître un terme supplémentaire $\nabla \cdot \boldsymbol{\sigma}(\mathbf{x}, t)$ qui correspond à la divergence du tenseur de contrainte. Cette dernière décrit l'action des forces de surface (ou de contact) entre un élément de volume et ses voisins.

Les effets conjugués des forces de volume et de surface induisent des déplacements infinitésimaux de l'élément qui se déforme en conséquence. Cette déformation s'exprime naturellement par le biais du tenseur des déformations $\boldsymbol{\epsilon}(\mathbf{x}, t)$, qui, dans le cas de petites déformations correspond à :

$$\boldsymbol{\epsilon}(\mathbf{x}, t) = \frac{1}{2} \left[\nabla \mathbf{u}(\mathbf{x}, t) + (\nabla \mathbf{u}(\mathbf{x}, t))^T \right]. \quad (1.4)$$

Un milieu est dit élastique s'il existe un état de référence sans contraintes et si le tenseur des contraintes ne dépend que du tenseur des déformations calculés à partir de l'état de référence (Duvaut, 1990). De ce fait, des petites variations de déformation induisent des petites variations de contraintes qui tendent à ramener le milieu à son état de référence. La loi qui relie les déformations aux contraintes est appelée loi de comportement. La plus simple, celle de l'élasticité linéaire⁵, est la loi de Hooke⁶. qui relie linéairement le tenseur des déformations au tenseur des contraintes :

$$\boldsymbol{\sigma}(\mathbf{x}, t) = \mathbf{c}(\mathbf{x}) : \boldsymbol{\epsilon}(\mathbf{x}, t), \quad (1.5)$$

où « : » représente le produit doublement contracté sur les deux indices externes tel que $\sigma_{ij} = c_{ijkl}\epsilon_{kl}$ et \mathbf{c} est le tenseur de raideur qui décrit le comportement du matériau étudié. Ce dernier est un tenseur d'ordre 4 et peut, de ce fait, posséder jusqu'à 81 coefficients indépendants. Néanmoins, la symétrie des tenseurs de contraintes et de petites déformations induit des propriétés de symétrie sur le tenseur de raideur, telles que

$$c_{ijkl} = c_{jikl} = c_{ijlk}. \quad (1.6)$$

Le nombre de coefficients élastiques indépendants passe ainsi de 81 à 36. De plus, si on considère que le tenseur de contrainte dérive d'une énergie potentielle élastique :

$$U = \frac{1}{2} c_{ijkl} \epsilon^{ij} \epsilon^{kl}, \quad (1.7)$$

nous obtenons une dernière relation de symétrie :

$$c_{ijkl} = c_{klij}, \quad (1.8)$$

qui réduit encore le nombre de paramètres indépendants à 21, le cas d'un milieu triclinique

5. Il s'agit en fait d'une linéarisation autour de l'état de référence.

6. ut tensio, sic vis

Paramètres	Symétrie	Type de minéraux
1	Fluide – Acoustique	eau
2	Isotrope	verre volcanique
3	Cubique	grenat
5	Hexagonal	glace
9	Orthorhombique	olivine
13	Monoclinique	hornblende
21	Triclinique	plagioclase

Table 1.1 – Tableau représentant le nombre de paramètres indépendants correspondant à quelques symétries cristallines (Babuska & Cara, 1991). Notons le cas particulier du milieu acoustique ne possédant qu’un seul coefficient.

c’est-à-dire un milieu totalement anisotrope. Au vu de ces symétries, il est possible et courant de représenter le tenseur de raideur sous forme d’une matrice ($c_{mn} = c_{6 \times 6}$) symétrique permettant de visualiser ces 21 coefficients. Nous introduisons pour cela la notation de Voigt (Voigt, 1889) telle que pour chaque couple d’indices $ij \rightarrow m$ et $kl \rightarrow n$,

$$11 \rightarrow 1; 22 \rightarrow 2; 33 \rightarrow 3; 32 \rightarrow 4; 31 \rightarrow 5; 21 \rightarrow 6.$$

Différents types d’anisotropie peuvent-être étudiés en considérant des matériaux présentant de plus hauts degrés de symétries (Babuska & Cara, 1991). Le tableau 1.1 présente les nombres de paramètres élastiques indépendants pour quelques systèmes de symétrie dont les cas extrêmes sont les milieux tricliniques et fluides. À l’échelle lithosphérique, les types les plus rencontrés sont : le milieu isotrope auquel nous nous restreindrons dans cette thèse et qui possède deux coefficients élastiques, le milieu hexagonal (isotrope transverse) à 5 coefficients qui possède un axe de symétrie et est souvent utilisé pour décrire l’anisotropie de milieu à couches et enfin, le milieu orthorhombique à 9 coefficients utile pour les études mantelliques dont le minéral dominant (l’olivine) est à symétrie orthorhombique.

Quelle que soit la physique du milieu considéré, nous obtenons, en injectant les équations (1.4) et (1.5) dans l’équation (1.3), le système dit des équations de l’élastodynamique :

$$\rho(\mathbf{x}) \partial_t^2 \mathbf{u}(\mathbf{x}, t) - \nabla \cdot \left[\mathbf{c}(\mathbf{x}) : \frac{1}{2} \left(\nabla \mathbf{u}(\mathbf{x}, t) + (\nabla \mathbf{u}(\mathbf{x}, t))^T \right) \right] = \mathbf{f}(\mathbf{x}, t), \quad (1.9)$$

qui traduit le comportement du milieu, les variations temporelles et spatiales des déplacements infinitésimaux en tous points, sous l’action d’une force extérieure. Ce système rend ainsi compte de tous les effets relatifs à la propagation des ondes élastiques.

La Terre, en première approximation, est souvent considérée comme un milieu isotrope. C’est-à-dire, que les propriétés physiques des matériaux qui la composent sont invariantes suivant les directions de l’espace. Pour un tel milieu, le tenseur de raideur se réduit à :

$$c_{ijkl} = \lambda \delta_{ij} \delta_{kl} + \mu (\delta_{ik} \delta_{jl} + \delta_{il} \delta_{jk}), \quad (1.10)$$

avec δ_{ij} le symbole de Kronecker qui vaut 1 si $i = j$, 0 sinon. Nous voyons que le tenseur de raideur ne contient plus que deux coefficients élastiques indépendants, appelés paramètres de Lamé λ et μ , que nous pouvons représenter matriciellement par :

$$c_{ijkl} = c_{mn} = \begin{pmatrix} \lambda + 2\mu & \lambda & \lambda & 0 & 0 & 0 \\ \lambda & \lambda + 2\mu & \lambda & 0 & 0 & 0 \\ \lambda & \lambda & \lambda + 2\mu & 0 & 0 & 0 \\ 0 & 0 & 0 & \mu & 0 & 0 \\ 0 & 0 & 0 & 0 & \mu & 0 \\ 0 & 0 & 0 & 0 & 0 & \mu \end{pmatrix}. \quad (1.11)$$

Dans le strict cadre de l'élasticité linéaire isotrope, nous pouvons réécrire les équations de l'élastodynamique telles que

$$\boxed{\rho \partial_t^2 \mathbf{u} - (\lambda + 2\mu) \nabla \nabla \cdot \mathbf{u} + \mu \nabla \times \nabla \times \mathbf{u} = \mathbf{f}}. \quad (1.12)$$

Les termes $\nabla \cdot \mathbf{u}$ et $\nabla \times \mathbf{u}$ décrivent respectivement le transport d'une onde sous forme de variation de volume et le transport d'une onde sous forme d'une variation de rotation.

1.3 Ondes propagées dans un milieu élastique

La figure 1.2 présente l'enregistrement d'un événement télésismique, le séisme de Tohoku-Oki⁷, enregistré par un sismomètre large-bande (ECH) dans les Vosges. L'analyse du sismogramme, le mouvement du sol enregistré par le capteur, met en évidence l'émergence de multiples arrivées liées à la propagation de différents types d'onde élastique. On en distingue deux grandes catégories : les *ondes de volume* qui se propagent partout à l'intérieur de la Terre, et les *ondes de surface*, d'amplitudes plus importantes, dont la propagation est restreinte au voisinage de la surface.

1.3.1 Ondes de volume

Au sein d'un volume homogène isotrope et infini, l'élasticité linéaire prédit l'existence de deux ondes de volume distinctes dont les vitesses de propagation et les polarisations sont différentes. Les caractéristiques principales de ces ondes sont facilement déterminées en considérant le théorème de Helmholtz (Aki & Richards, 2002; Lay & Wallace, 1995; Stein & Wysession, 2009) qui permet de reformuler le champ vectoriel de déplacement \mathbf{u} comme la somme du gradient d'un potentiel scalaire Φ et du rotationnel d'un potentiel vecteur Ψ :

$$\mathbf{u} = \nabla \Phi + \nabla \times \Psi \quad (1.13)$$

Ces potentiels indiquent que le déplacement peut-être décomposé en une composante *longitudinale* (irrotationnelle $\nabla \times \Phi = 0$) et une composante transverse (solénoïdale $\nabla \cdot \Psi = 0$). En

7. Funeste responsable de la catastrophe nucléaire de la centrale de Fukushima

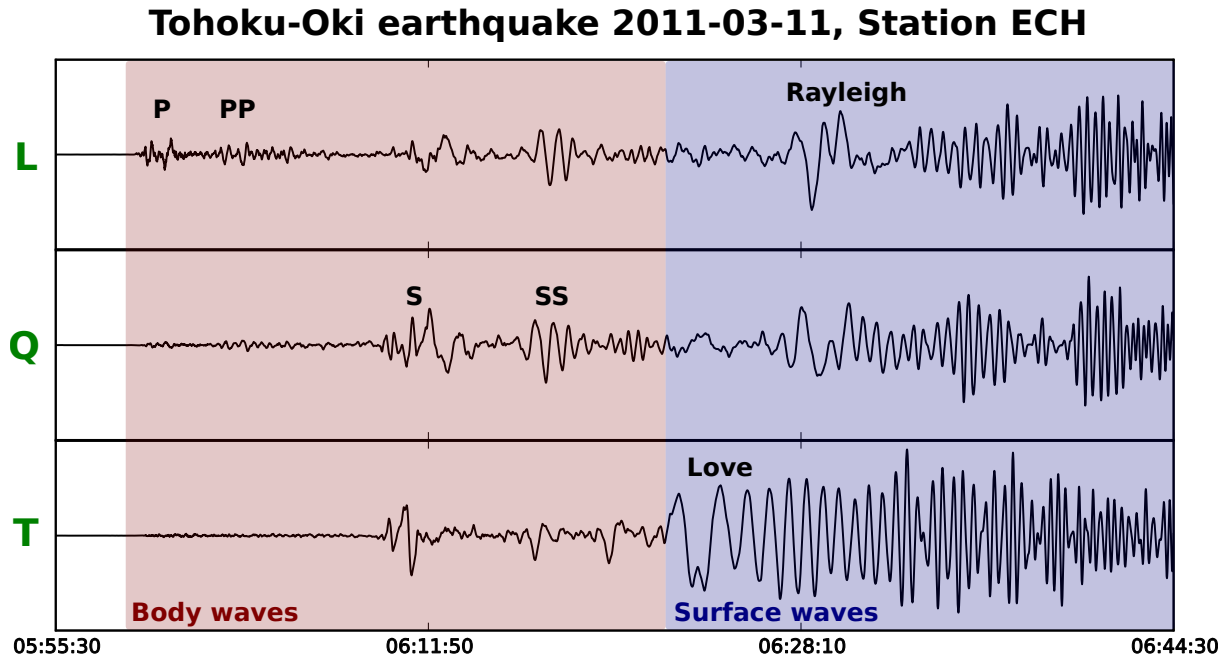


Figure 1.2 – Sismogrammes télésismiques du séisme de Tohoku-Oki ($M_w = 9.1$) enregistrés par la station ECH située à une distance épacentrale de 85° . Les zones rouge et bleue représentent les domaines d’arrivées des ondes de volume et de surface.

injectant cette expression dans l’équation d’onde (1.12), on obtient après quelques arrangements algébriques :

$$\nabla \left[\rho \partial_t^2 \Phi - (\lambda + 2\mu) \nabla^2 \Phi \right] = -\nabla \times \left[\rho \partial_t^2 \Psi - \mu \nabla^2 \Psi \right]. \quad (1.14)$$

Cette dernière équation possède une solution, si et seulement si, les deux termes entre crochets s’annulent de sorte que :

$$\nabla^2 \Phi - \frac{1}{\alpha^2} \partial_t^2 \Phi = 0, \quad \text{et} \quad \nabla^2 \Psi - \frac{1}{\beta^2} \partial_t^2 \Psi = 0. \quad (1.15)$$

Nous retrouvons deux équations de d’Alembert, la forme la plus simple de l’équation d’onde, pour les potentiels Φ et Ψ dont les vitesses de propagation sont respectivement α et β .

Ondes P

L’étude du potentiel scalaire Φ montre qu’un premier type d’onde de volume se propage de manière longitudinale avec une vitesse de propagation :

$$\alpha = v_p = \sqrt{\frac{\lambda + 2\mu}{\rho}}. \quad (1.16)$$

La polarisation longitudinale des ces ondes implique que le mouvement des particules est colinéaire à la direction de propagation, il s’ensuit des variations locales de volume qui propagent l’onde par une succession de pressions et de dilatations (Fig. 1.3). Il s’agit des *ondes P*, pour

Main types of seismic waves

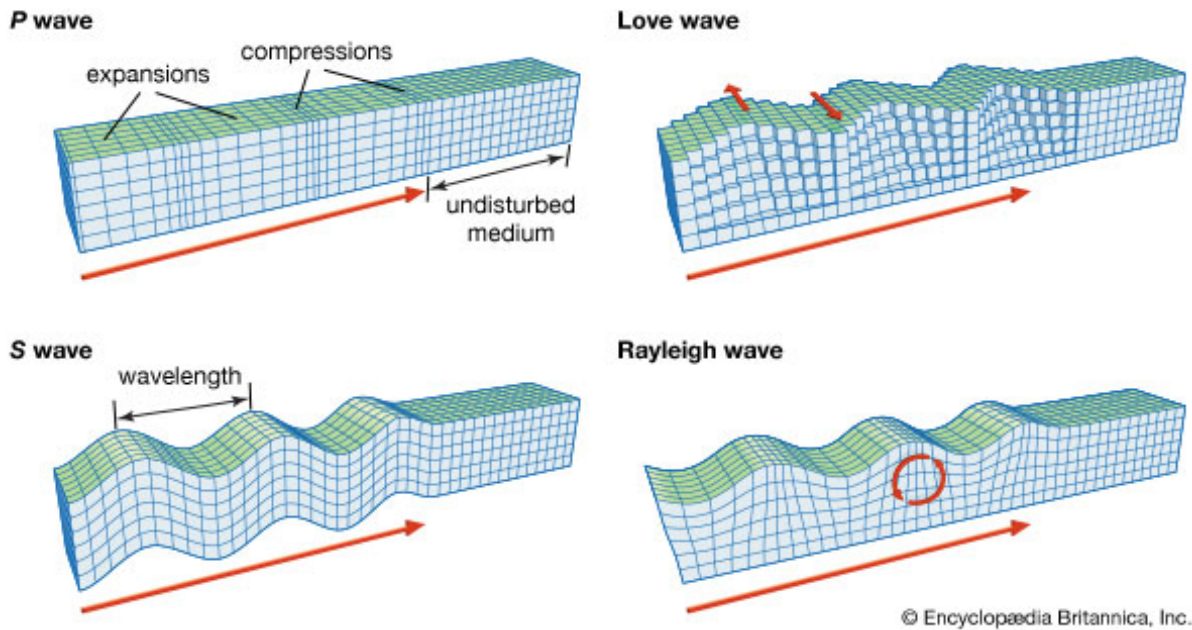


Figure 1.3 – Présentation des différentes ondes propagées au sein d'un milieu élastique avec leurs polarisations. Source : *Encyclopedia Britannica*.

« *pressure waves* » (les ondes de pression), aussi nommées ondes premières car ce sont les plus rapides et constituent, en général, les premières arrivées observées sur un sismogramme. Ces ondes dépendent à la fois du module d'incompressibilité λ et du module de cisaillement μ .

Ondes S

L'étude du potentiel vecteur Ψ montre qu'une autre onde de volume, de polarisation linéaire transverse, est propagée dans le milieu élastique à une vitesse de :

$$\beta = v_s = \sqrt{\frac{\mu}{\rho}}. \quad (1.17)$$

Il s'agit des *ondes S*, pour « *shear waves* » (les ondes de cisaillement), aussi appelées ondes secondaires car elles arrivent généralement après les ondes P du fait d'une vitesse moindre, communément d'un facteur inversement proportionnel à $\sqrt{3}$. Contrairement aux ondes P, elles ne génèrent pas de variations locales de volume mais cisailent le milieu traversé. En effet, le déplacement particulière associé aux ondes S est transversale, c'est-à-dire, que le mouvement se fait dans le plan orthogonal à la direction de propagation (Fig. 1.3). On distingue deux types d'onde S : les ondes Sv qui ont une polarisation dans le plan vertical contenant la source et le récepteur – on parle aussi de polarisation radiale – et les ondes Sh dont la polarisation est transverse c'est-à-dire perpendiculaire aux directions longitudinales et radiales. Les ondes S ne dépendent que du module de cisaillement μ et ne peuvent être propagées dans les milieux

fluides.

1.3.2 Ondes de surface

Lorsqu'un solide élastique homogène est borné par une surface⁸, les ondes de volume entrent en interaction avec cette interface et produisent une série d'ondes évanescentes⁹ P et S. L'interférence constructive de ces ondes évanescentes génèrent les ondes dites de surface. Ces dernières sont guidées par l'interface et leurs amplitudes décroît rapidement à mesure que l'on s'en éloigne. Elles sont, en outre, très énergétiques puisque leurs propagations essentiellement 2D induisent une décroissance géométrique inversement proportionnelle avec la distance (les ondes de volumes, quant à elles, décroissent proportionnellement avec l'inverse du carré de la distance). On en distingue deux types : les *ondes de Rayleigh* et les *ondes de Love*.

Ondes de Rayleigh

Les ondes de Rayleigh (Rayleigh, 1887) sont générées dans tout milieu possédant une surface libre, par interférences constructives d'ondes évanescentes P et Sv. Elles sont facilement reconnaissables de part leurs polarisations caractéristiques : elliptiques et rétrogrades dans le plan de propagation (Fig. 1.2 les composantes L et Q, et Fig. 1.3). Il n'existe pas de formule analytique simple pour déterminer la vitesse de propagation des ondes de Rayleigh V_r , néanmoins, une bonne approximation de cette vitesse est obtenue par la relation :

$$V_r = V_s \left(\frac{20 - \sqrt{256\gamma^4 - 336\gamma^2 + 130}}{16\gamma^2 + 9} \right). \quad (1.18)$$

où $\gamma = V_s/V_p$. Une autre approximation, plus courante, est de considérer un rapport $V_p/V_s \approx \sqrt{3}$ pour lequel la vitesse approximative de l'onde de Rayleigh vaut $V_r = 0.92V_s$. Notons que l'onde de Rayleigh, dans un demi-espace strictement homogène possédant une surface plane, est non dispersive. Dans tout autre cas, si la surface est courbe et/ou des hétérogénéités sont présentes dans le milieu, l'onde devient dispersive (dans ces cas, on parle aussi d'ondes de Lamb).

Ondes de Love

Les ondes de Love (Love, 1911), contrairement aux ondes de Rayleigh, n'apparaissent que lorsque le demi-espace possède, au niveau de la surface, une couche aux propriétés élastiques différentes du reste du milieu. Ces ondes sont créées exclusivement par l'interférence constructive d'ondes évanescentes Sh et sont, de plus, toujours dispersives. Les ondes de Love ont une polarisation linéaire horizontale (Fig. 1.3) et sont ainsi visibles sur la composante transversale

8. On se restreint à la notion de surface, mais la théorie est valable pour n'importe quel type d'interface (fluide/solide par exemple qui crée les ondes de Scholte/Stoneley (Stoneley, 1926)).

9. Des ondes dont l'amplitude décroît exponentiellement en fonction de leur distance à l'interface et qui font partie des solutions dites de champ proche.

d'un sismogramme (Fig. 1.2, composante T). Elles sont, par ailleurs, systématiquement plus rapides que les ondes de Rayleigh.

1.4 Quelques solutions analytiques

Nous présentons, à présent, quelques solutions canoniques au problème de propagation des ondes. La première est l'expression de la fonction de Green, pour un milieu homogène infini soumis à une force ponctuelle, qui nous permettra de valider les outils de modélisation numérique et de présenter le concept de fonction de Green. La seconde est l'expression analytique de la propagation d'une onde plane, dans un demi-espace homogène, que nous utiliserons à des fins de validation et comme solution de champ incident lors de l'injection de champ d'onde dans des demi-espace homogènes.

1.4.1 Les fonctions de Green

Une *fonction de Green* est une solution élémentaire d'une équation différentielle ou d'une équation aux dérivées partielles à coefficients constants. Ce type de fonction est obtenu en cherchant les solutions de l'équation différentielle pour une source impulsionnelle correspondant à une distribution $\delta = \delta(\mathbf{x}_s, \tau)$ de Dirac déclenchée à la position \mathbf{x}_s et au temps τ . La théorie formelle du problème de Green fait ainsi appel à la théorie des distributions. Chercher une fonction de Green du problème de propagation des ondes élastique (1.9) se réduit à résoudre le système suivant :

$$\rho \partial_t^2 \mathbf{G} - \nabla \cdot [\mathbf{c} : ((\nabla \mathbf{G})^T + \nabla \mathbf{G})] = \delta, \quad (1.19)$$

où $\mathbf{G} = \mathbf{G}(\mathbf{x}, t | \mathbf{x}_s, \tau)$ est la solution que nous recherchons. Pour assurer l'unicité de la solution, il est nécessaire d'ajouter des conditions aux limites (aux frontières et initiales) à ce problème.

Considérons un milieu isotrope, infini et homogène soumis à une excitation de type force ponctuelle. Pour un tel milieu, la fonction de Green G_{ij} donne le déplacement particulière dans la direction i occasionné par une force ponctuelle dans la direction j , par le biais de :

$$\boxed{G_{ij}(\mathbf{x}, t | \mathbf{x}_s, \tau) = \frac{\gamma_i \gamma_j}{4\pi \rho \alpha^2} \frac{1}{R} \delta \left(t - \tau - \frac{R}{\alpha} \right) + \frac{\delta_{ij} - \gamma_i \gamma_j}{4\pi \rho \beta^2} \frac{1}{R} \delta \left(t - \tau - \frac{R}{\beta} \right) + \frac{3\gamma_i \gamma_j - \delta_{ij}}{4\pi \rho} \frac{1}{R^3} \int_{R/\alpha}^{R/\beta} \tau' \delta(t - \tau - \tau') d\tau',} \quad (1.20)$$

où γ_i et γ_j sont les cosinus directeurs de la composante i étudiée, soumis à la composante j de la force ; ils sont donnés par :

$$\gamma_i = \frac{x_i}{|\mathbf{r}|} = \frac{\partial x_i}{\partial \mathbf{r}}. \quad (1.21)$$

L'équation (1.20) est très importante puisqu'elle permet de déterminer le champ d'onde généré par une source localisée en \mathbf{x}_s , en tout point \mathbf{x} du milieu et pour tout temps t . De plus, quelle que

soit la signature temporelle de la source s , le déplacement peut-être déterminé par une simple convolution $\mathbf{u} = \mathbf{G} * s$ en temps en vertu de la linéarité de l'équation d'onde.

Détaillons les différents termes de notre fonction de Green. Les deux premier représentent la propagation, en champ lointain, des ondes P et S caractérisées par leurs vitesses α et β et pour lesquelles l'amplitude décroît avec l'inverse de la distance. Notons au passage les expressions des diagrammes de radiation à la source des ondes P et S ($\gamma_i \gamma_j$ et $\delta_{ij} - \gamma_i \gamma_j$) qui décrivent l'amplitude du champ émis en fonction de la direction de propagation. Le dernier terme, dont l'expression est plus compliquée décrit une « rampe » entre l'arrivée de l'onde P et l'arrivée de l'onde S. Son amplitude décroît très rapidement, de manière inversement proportionnelle au cube de la distance. Il caractérise le champ proche.

Dans le cas d'un tenseur de moment complet, l'expression de la fonction de Green devient plus compliquée et nous invitons le lecteur à se référer à l'ouvrage d'Aki & Richards (2002). De même, plusieurs fonctions de Green analytiques correspondant à des conditions aux limites différentes sont disponibles dans l'ouvrage de Kausel (2006), et plus particulièrement la solution au problème de Lamb qui est donnée par Kausel (2012).

1.4.2 Onde plane dans un demi-espace homogène

Une onde plane propagative se décrit comme le déplacement dans un espace tri-dimensionnel d'un front d'onde plan de phase constante. La phase d'une onde plane se propageant à la vitesse v s'écrit :

$$t - \frac{\mathbf{n} \cdot (\mathbf{x} - \mathbf{x}_0)}{v}, \quad (1.22)$$

où \mathbf{n} est un vecteur unitaire, orthogonal au front d'onde, indiquant la direction de propagation. Considérons une onde plane harmonique, solution des équations de l'élastodynamique (1.9), qui s'écrit sous la forme :

$$\mathbf{u}(\mathbf{x}, t) = \mathbf{p} \cdot A \exp \left\{ i\omega \left(t - \frac{\mathbf{n} \cdot (\mathbf{x} - \mathbf{x}_0)}{v} \right) \right\} \quad (1.23)$$

et exprime le vecteur déplacement¹⁰ en fonction de ω la pulsation, A un facteur d'amplitude, \mathbf{n} la direction de propagation et de \mathbf{p} le vecteur polarisation unitaire de l'onde considérée.

On se limite à l'étude d'une onde P se propageant dans un demi-espace homogène dans la direction $\mathbf{n}_p = (n_{px}, n_{py}, n_{pz})^T$. Sa polarisation étant longitudinale, nous déduisons que $\mathbf{p}_p = \mathbf{n}_p$. L'expression de l'onde plane P incidente devient :

$$\mathbf{u}^p(\mathbf{x}, t) = \mathbf{p}_p \cdot \exp \left\{ i\omega \left(t - \frac{\mathbf{n}_p \cdot (\mathbf{x} - \mathbf{x}_0)}{v_p} \right) \right\} \quad (1.24)$$

pour laquelle nous imposons que l'amplitude A est unitaire. En présence d'une surface libre,

10. L'expression (1.23) contient une exponentielle a valeur complexe, or, seule sa partie réelle possède une signification physique. Utiliser l'exponentielle permet néanmoins de simplifier les développements, à la fin, le déplacement physique est obtenu en prenant la partie réelle de l'équation (1.23).

cette onde P incidente va générer une onde P réfléchiée et une onde Sv convertie dont les directions de propagation \mathbf{n}_{pp} et \mathbf{n}_{ps} sont données par la loi de Snell-Descartes. Cette loi énonce que les paramètres de rai, l'inverse de la vitesse apparente c_h , est conservé pour toutes les ondes réfléchiées et transmises à une interface :

$$p = \frac{1}{c_h} = \frac{\sin i_1}{\alpha_1} = \frac{\sin i_2}{\alpha_2} = \frac{\sin j_1}{\beta_1} = \frac{\sin j_2}{\beta_2} = \text{cste.} \quad (1.25)$$

où i et j correspondent aux angles d'incidence des ondes P et S respectivement, et les indices $_1$ et $_2$ aux milieux bornant l'interface. On en déduit les directions de propagation des ondes diffractées :

$$\begin{aligned} \mathbf{n}_{pp} &= (n_{px}, n_{py}, -n_{pz})^T \\ \mathbf{n}_{ps} &= (n_{px}, n_{py}, -n_{pz})^T (\beta/\alpha) \end{aligned} \quad (1.26)$$

De même que pour l'onde P incidente, l'onde PP réfléchiée est longitudinale et nous pouvons écrire que $\mathbf{p}_{pp} = \mathbf{n}_{pp}$. En revanche, nous avons vu que l'onde Sv est polarisée de manière orthogonale à la direction de propagation dans le plan vertical. Il s'ensuit que $\mathbf{p}_{ps} = (\mathbf{n}_{ps} \times \mathbf{n}_{ps_z}) \times \mathbf{n}_{ps}$, que nous pouvons réécrire :

$$\mathbf{p}_{ps} = n_{ps_z} (-n_{ps_z} n_{ps_x}, -n_{ps_z} n_{ps_y}, n_{ps_y}^2 + n_{ps_x}^2)^T \quad (1.27)$$

À la surface, la condition de surface libre permet de déterminer les coefficients de réflexion-conversion, pour les ondes P et Sv descendantes qui contraignent l'amplitude des déplacements \mathbf{u}^{pp} et \mathbf{u}^{ps} tel que :

$$\begin{aligned} \mathbf{u}^{pp}(\mathbf{x}, t) &= \mathbf{p}_{pp} \cdot R^{PP} \exp \left\{ i\omega \left(t - \frac{\mathbf{n}_{ps} \cdot (\mathbf{x} - \mathbf{x}_o)}{v_p} \right) \right\} \\ \mathbf{u}^{ps}(\mathbf{x}, t) &= \mathbf{p}_{ps} \cdot R^{PS} \exp \left\{ i\omega \left(t - \frac{\mathbf{n}_{pp} \cdot (\mathbf{x} - \mathbf{x}_o)}{v_s} \right) \right\} \end{aligned} \quad (1.28)$$

Le champ d'onde total est au final exprimé comme la somme des champs incidents et convertis de sorte que :

$$\mathbf{u} = \mathbf{u}^p + \mathbf{u}^{pp} + \mathbf{u}^{ps}. \quad (1.29)$$

Les coefficients de réflexion et de transmission sont donnés dans plusieurs ouvrages dont [Lay & Wallace \(1995\)](#) et [Aki & Richards \(2002\)](#) dans le cas d'une surface libre, et dans le cas d'interface solide-solide et fluide-solide. Nous rappelons uniquement l'expression des coefficients de réflexion à la surface pour une onde P montante :

$$R_{PPP} = \frac{-(\frac{1}{\beta^2} - 2p^2)^2 + 4p^2 \frac{\cos i}{\alpha} \frac{\cos j}{\beta}}{(\frac{1}{\beta^2} - 2p^2)^2 + 4p^2 \frac{\cos i}{\alpha} \frac{\cos j}{\beta}} \quad \text{et} \quad R_{PPS} = \frac{4 \frac{\alpha}{\beta} p \frac{\cos i}{\alpha} (\frac{1}{\beta^2} - 2p^2)^2}{(\frac{1}{\beta^2} - 2p^2)^2 + 4p^2 \frac{\cos i}{\alpha} \frac{\cos j}{\beta}} \quad (1.30)$$

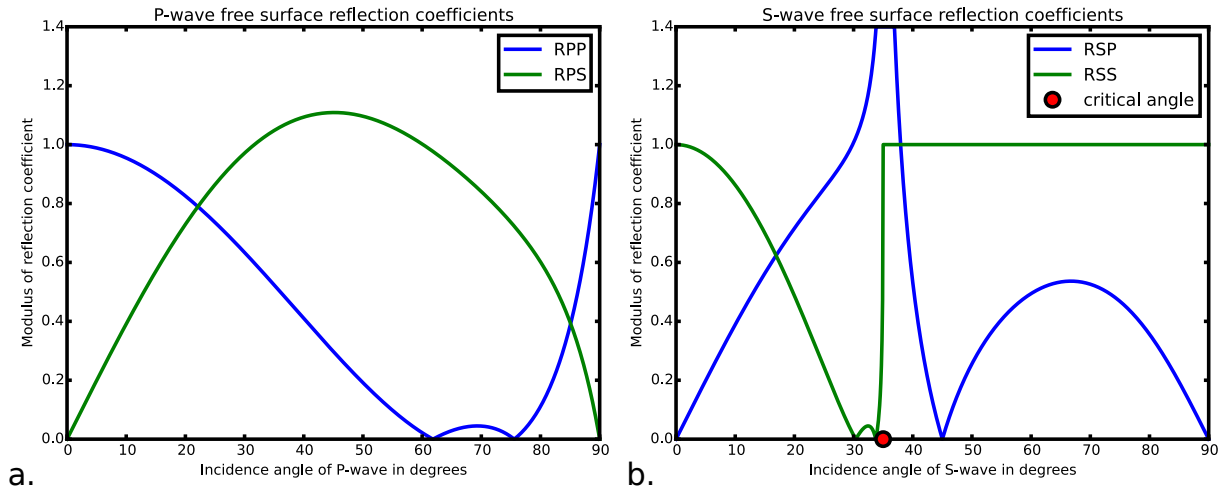


Figure 1.4 – Coefficients de réflexion et conversions à la surface libre. (a.) Coefficients pour une onde P incidente et (b.) pour une onde S incidente. Le milieu est paramétré par $\rho = 3 \text{ g.cm}^{-3}$, $\alpha = 7 \text{ km.s}^{-1}$ et $\beta = 4 \text{ km.s}^{-1}$. La pastille rouge indique l'angle critique à partir duquel l'onde devient évanescente.

ainsi que pour une onde S montante :

$$R_{SS} = \frac{\left(\frac{1}{\beta^2} - 2p^2\right)^2 - 4p^2 \frac{\cos i}{\alpha} \frac{\cos j}{\beta}}{\left(\frac{1}{\beta^2} - 2p^2\right)^2 + 4p^2 \frac{\cos i}{\alpha} \frac{\cos j}{\beta}} \quad \text{et} \quad R_{SP} = \frac{4 \frac{\beta}{\alpha} p \frac{\cos j}{\beta} \left(\frac{1}{\beta^2} - 2p^2\right)^2}{\left(\frac{1}{\beta^2} - 2p^2\right)^2 + 4p^2 \frac{\cos i}{\alpha} \frac{\cos j}{\beta}} \quad (1.31)$$

pour lesquelles, p est le paramètre du rai et i et j les angles d'incidence des ondes P et S.

La figure 1.4 montre la variation des coefficients de réflexion en fonction de l'angle incidence de l'onde P montante. Nous remarquons que pour la gamme d'incidence rencontrée en configuration télésismique, de 10 à 60 degrés, les ondes PP et PS réfléchies par la surface libre ont une amplitude importante. Pour de faibles incidences, inférieures à 20°, la réflexion PP domine tandis que pour des incidences plus importantes, au-delà de 20°, l'amplitude des ondes PS est dominante. Le champ secondaire provenant de la surface libre possède ainsi une amplitude comparable au champ incident. Ces phases peuvent ainsi être utilisées dans une optique d'imagerie.

Précisons que dans le cas de la propagation d'une onde plane dans un demi-espace homogène, aucune onde de surface n'est générée.

1.5 Ondes sismiques en milieux hétérogènes

Jusqu'à présent, nous avons considéré la Terre comme un espace élastique homogène ou radialement hétérogène pour lesquels s'appliquent l'optique géométrique et la théorie des rais. Ces approximations sont suffisantes pour décrire précisément nombre de phénomènes liés à la propagation des ondes dans la Terre globale et des milieux stratifiés. Cependant, l'intérieur du Globe est hétérogène puisque des facteurs tels que la tectonique, la pétrophysique et la

géothermie, entre autres, induisent des structures géophysiques complexes pouvant contenir des hétérogénéités de tailles variables : de l'ordre de quelques centimètres à plusieurs dizaines de kilomètres. Cette complexité affecte le champ d'onde à divers degrés qui sont notamment dépendant de la fréquence du signal considéré et de la taille des hétérogénéités. Selon les cas considérés, la propagation ne peut plus être décrite par des modèles physiques simples¹¹ et il est nécessaire de considérer des modèles physiques plus élaborés, notamment, en prenant en compte les effets de *diffusion* que nous présentons dans cette section.

1.5.1 La diffusion des ondes élastiques

Le phénomène de diffusion des ondes – *wave scattering* en anglais – intervient lorsqu'au cours de sa propagation, une onde par l'interaction avec une ou plusieurs hétérogénéités présentes dans le milieu, dévie de sa trajectoire initiale. On parle de diffusion simple lorsque l'onde interagit une seule fois, et de diffusion multiple lorsque plusieurs hétérogénéités sont impliquées. Au cours de leur propagation dans un milieu hétérogène, les ondes sismiques subissent des variations de formes d'ondes, d'amplitude, de phase pouvant dépendre de la fréquence. Ces variations peuvent aussi bien être isotropes qu'anisotropes.

La littérature est assez confuse quant à l'utilisation des termes diffusion et diffraction qui sont utilisés, abusivement, de manière interchangeable. La diffraction des ondes est à proprement parler un cas limite de la diffusion des ondes. Ce phénomène peut-être entièrement décrit par le principe de Huygens-Fresnel : chaque point diffusant se comporte comme une source secondaire et la diffraction résulte de l'interférence constructive des ondes diffusées par l'ensemble de ces sources secondaires. Il en découle que l'amplitude du champ d'onde est conservée. Dans la suite de ce manuscrit, le terme de diffraction sera utilisé aussi bien pour la diffraction que la diffusion.

Introduisant un ouvrage intitulé « *Scattering and Attenuations of Seismic Waves* » et portant sur la diffusion et l'atténuation des ondes sismiques, Wu & Aki (1988) ont tenté de classer les différents régimes de diffusion des ondes sismiques en fonction de paramètres caractérisant :

- la taille des hétérogénéités a
- l'amplitude relative δm de ces hétérogénéités
- la longueur d'onde du signal propagé λ

En introduisant le paramètre adimensionnel ka qui correspond au nombre d'onde normalisé par la taille des hétérogénéités, ils définissent quatre grands régimes de diffusion :

Quasi-homogène Apparaît lorsque $ka \ll 0.01$, les hétérogénéités sont trop petites pour être perçues par les ondes propagées et le milieu peut-être considéré comme homogène (on parle aussi de milieu effectif). Dans ce régime les effets de la diffusion sont négligeables.

Diffusion de Rayleigh Apparaît lorsque $0.01 < ka < 0.1$. En considérant de faibles fluctuations des paramètres élastiques, ce régime de diffusion peut-être modélisé par l'approximation de Born (en supposant un régime de diffusion simple). La quantité d'énergie diffusée en 3D est proportionnelle à k^4 impliquant une atténuation apparente des hautes fréquences.

11. Comme les méthodes asymptotiques fondées sur la théorie des rais.

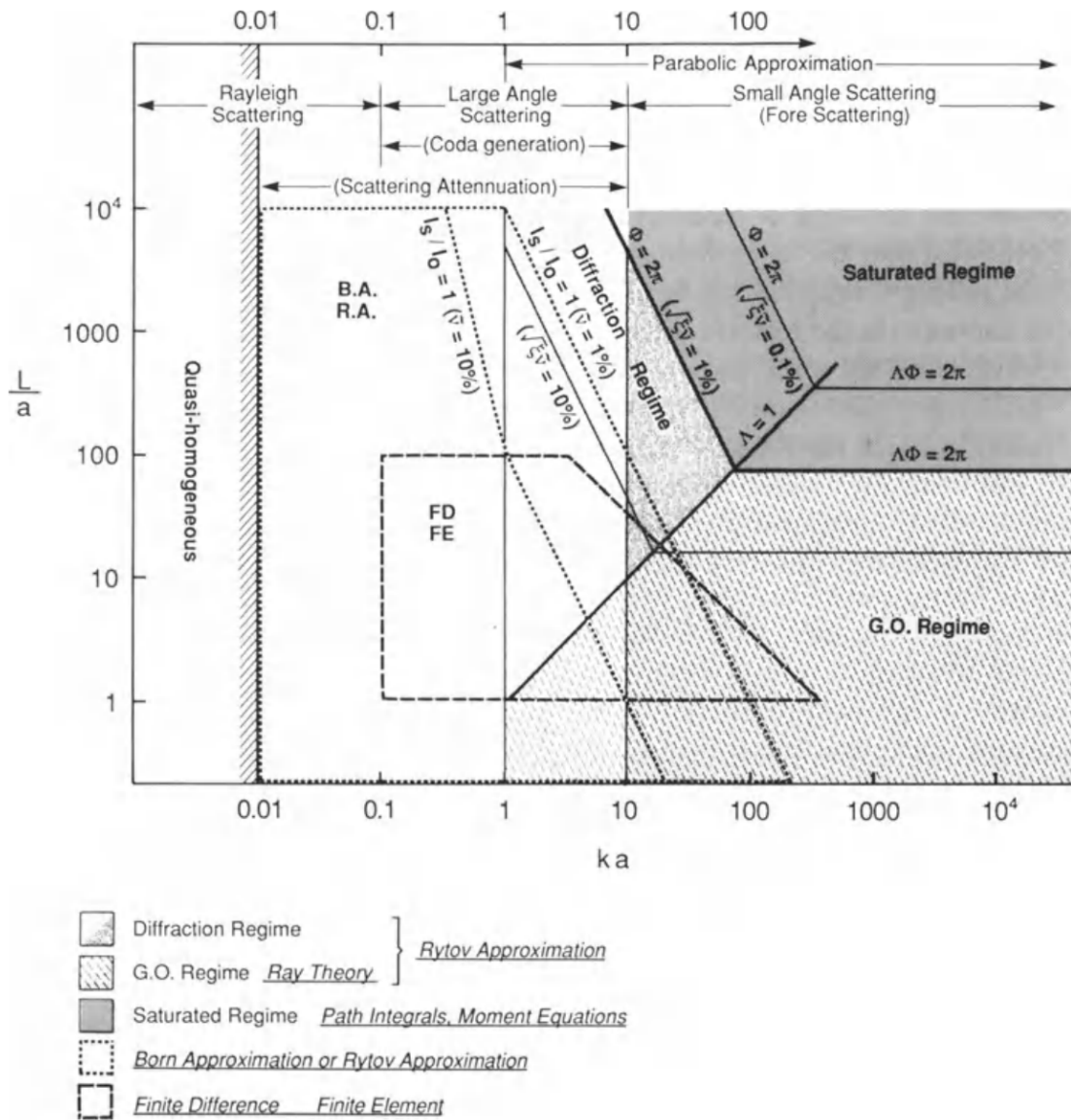


Figure 1.5 – Régimes de diffusion et domaines de validité de différentes méthodes d'approximation. a : échelle caractéristique des hétérogénéités, L : distance de propagation de l'onde et k : le nombre d'onde (d'après Wu & Aki (1988))

Diffusion de Mie Aussi nommé diffusion grand angle ou résonnante. Ce régime de diffusion intervient lorsque $0.1 < ka \approx 1 < 10$, c'est-à-dire que la taille des hétérogénéités est de l'ordre de la longueur d'onde. Les ondes diffusées sont alors déviées avec des angles très importants par rapport à l'onde incidente et une coda en réflexion peut être observées dans le champs d'onde.

Diffusion spéculaire Aussi nommée diffusion petit angle, elle apparaît lorsque la longueur d'onde est très faible au vue de la taille caractéristique des hétérogénéités ($ka \gg 10$). Il s'agit du régime de l'optique géométrique où l'énergie des ondes est diffusée principalement dans la direction de propagation. Dans ce régime, la diffusion arrière (*back-scattering*) est faible et les effets de focalisation, de diffraction et d'interférence deviennent importants. Dans ce régime de diffusion, l'approximation de Rytov peut être utilisée.

Notons que la diffusion de Rayleigh intervient lorsque la longueur d'onde du champs incident est très grande par rapport à la taille caractéristique des hétérogénéités. Ce mode de diffusion n'induit aucune variation d'énergie et l'onde diffusée conserve la longueur d'onde du champ incident. Pour des particules plus grandes, on entre dans le régime de diffusion de Mie, dont la théorie physique est exacte quelle que soit la taille des particules sphériques considérées¹².

La figure 1.5 présente les domaines de validités de différentes méthodes d'approximation de la propagation des ondes sismiques en fonction de ka et d'une grandeur sans-dimension L/a qui caractérise la distance de propagation. Pour des applications télésismiques, ces nombres d'ondes et distances de propagation sont extrêmement variées. Les champs d'ondes télésismiques couvrent en conséquence une large gamme de phénomène de diffusion. Bien que les méthodes de perturbation fondées sur une linéarisation de l'équation d'onde avec l'approximation de Born ou de Rytov (Woodward, 1992)¹³ permettent de caractériser une part importante de ces régimes, seules les méthodes numériques résolvant l'équation d'onde (différences-finies, éléments-finis) sont à même de rendre compte de toute cette complexité.

1.5.2 Diffusion simple par l'approximation de Born

Nous présentons dans cette section la théorie de Born (Hudson & Heritage, 1981) qui décrit le comportement du champ l'onde lorsque celui-ci entre en interaction simple avec une hétérogénéité ponctuelle des paramètres élastiques. Formellement, cette approximation est valide pour la modélisation d'un champ d'onde dans le cadre restreint du régime de diffusion de Rayleigh. Cependant, la théorie de Born permet l'analyse simple du champ d'onde diffusé et aussi de tirer des conclusions sur l'empreinte des différents paramètres élastiques dans les données sismiques (Wu & Aki, 1985; Tarantola, 1986). En outre, l'approximation de Born pour la modélisation des champs d'onde est intimement liée aux problèmes inverses de migration des champs diffusés ainsi que de la FWI (comme nous le verrons dans la seconde partie de cette thèse). Dans ce qui suit, nous considérons cette approximation pour l'étude de la diffusion des ondes de volume.

12. La diffusion de Rayleigh peut-être vue comme un cas limite de la diffusion de Mie.

13. L'approximation de Born perturbe l'amplitude et la phase du champ d'onde et permet de prendre en compte le back-scattering, Rytov perturbe uniquement la phase et est essentiellement utilisée pour étudier le forward scattering.

Pour les ondes de surface, nous invitons le lecteur intéressé à se référer aux études de [Snieder \(1986\)](#) et de [Snieder & Nolet \(1987\)](#), [Calvet et al. \(2006\)](#) et [Sieminski et al. \(2009\)](#) pour l'anisotropie et à [Ribodetti et al. \(2000\)](#) pour l'atténuation.

L'approche de Born consiste à linéariser, au premier ordre, la réponse du champ d'onde à une petite perturbation des paramètres du milieu de propagation. Le modèle est ainsi séparé en un modèle de référence \mathbf{m}_o ¹⁴ et une perturbation $\delta\mathbf{m}$ ponctuelle choisie de sorte que la fluctuation fractionnelle $\delta\mathbf{m}/\mathbf{m}_o \ll 1$. Nous décomposons donc la densité et le tenseur de raideur tels que

$$\rho = \rho_o + \delta\rho \quad \text{et} \quad \mathbf{c} = \mathbf{c}_o + \delta\mathbf{c}. \quad (1.32)$$

De façon similaire, nous décomposons le champ d'onde total \mathbf{u} en fonction du champ incident \mathbf{u}_o , solution du problème de propagation d'onde dans le modèle de référence, et du champ diffusé par l'hétérogénéité $\delta\mathbf{u}$ de sorte que

$$\mathbf{u} = \mathbf{u}_o + \delta\mathbf{u}. \quad (1.33)$$

On cherche à déterminer les équation qui caractérise la propagation du champs diffusé. Nous savons que le champ total satisfait l'équation d'onde que nous réécrivons sous sa forme fréquentielle qui est l'équation d'Helmholtz :

$$-\rho\omega^2\mathbf{u} - \nabla \cdot [\mathbf{c} : ((\nabla\mathbf{u})^T + \nabla\mathbf{u})] = \mathbf{f} \quad (1.34)$$

Les fonctions de Green \mathbf{G} correspondantes sont solutions de

$$-\rho\omega^2\mathbf{G} - \nabla \cdot [\mathbf{c} : ((\nabla\mathbf{G})^T + \nabla\mathbf{G})] = \delta, \quad (1.35)$$

que nous avons présentées en section 1.4.1. En injectant les approximations (1.32) et (1.33) dans l'équation (1.34), et en négligeant les termes d'ordre supérieurs à un, nous obtenons

$$\boxed{-\rho_o\omega^2\delta\mathbf{u} - \nabla \cdot [\mathbf{c}_o : (\nabla\delta\mathbf{u} + (\nabla\delta\mathbf{u})^T)] = \delta\rho\omega^2\mathbf{u}_o + \nabla \cdot [\delta\mathbf{c} : (\nabla\mathbf{u}_o + (\nabla\mathbf{u}_o)^T)]}. \quad (1.36)$$

Cette équation, qui a la même forme que l'équation (1.34), décrit la propagation dans le milieu de référence \mathbf{m}_o du champ diffusé $\delta\mathbf{u}$ généré par un nouveau terme source

$$\boxed{\delta\mathbf{f} = \delta\rho\omega^2\mathbf{u}_o + \nabla \cdot [\delta\mathbf{c} : (\nabla\mathbf{u}_o + (\nabla\mathbf{u}_o)^T)]}. \quad (1.37)$$

Cette source secondaire est similaire à un point source qui correspond à l'extraction du champ incident \mathbf{u}_o à la position du diffuseur pondérée par les perturbations locales $\delta\rho$ et $\delta\mathbf{c}$ des paramètres du milieu. Connaissant les fonction de Green \mathbf{G}' pour un point source localisé à la position du diffuseur et enregistré au récepteur, nous pouvons exprimer le champ diffusé sous forme d'une convolution :

$$\delta\mathbf{u} = \mathbf{G}' * \delta\mathbf{f} = \delta\rho\omega^2\mathbf{G}'\mathbf{u}_o + \mathbf{G}'\nabla \cdot [\delta\mathbf{c} : (\nabla\mathbf{u}_o + (\nabla\mathbf{u}_o)^T)]. \quad (1.38)$$

14. Généralement lisse ou homogène de sorte que les fonction de Green dans ce modèle soient faciles à calculer.

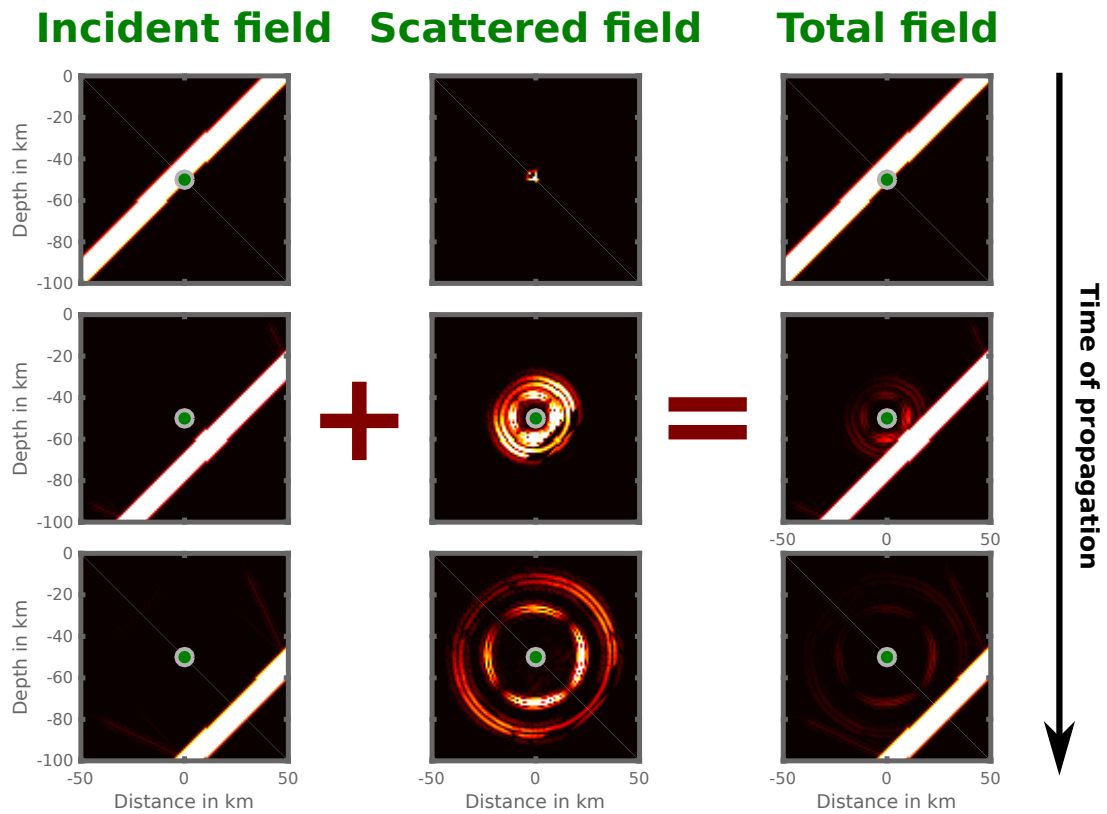


Figure 1.6 – Illustration numérique de l’approximation de Born. La première colonne présente la propagation d’une onde plane P incidente dans un milieu homogène. La deuxième, montre la propagation du champ diffusé par une source secondaire localisée au point diffusant (rond vert). La dernière colonne représente le champ total, soit la somme du champ incident et du champ diffusé. La perturbation est portée sur le modèle V_s .

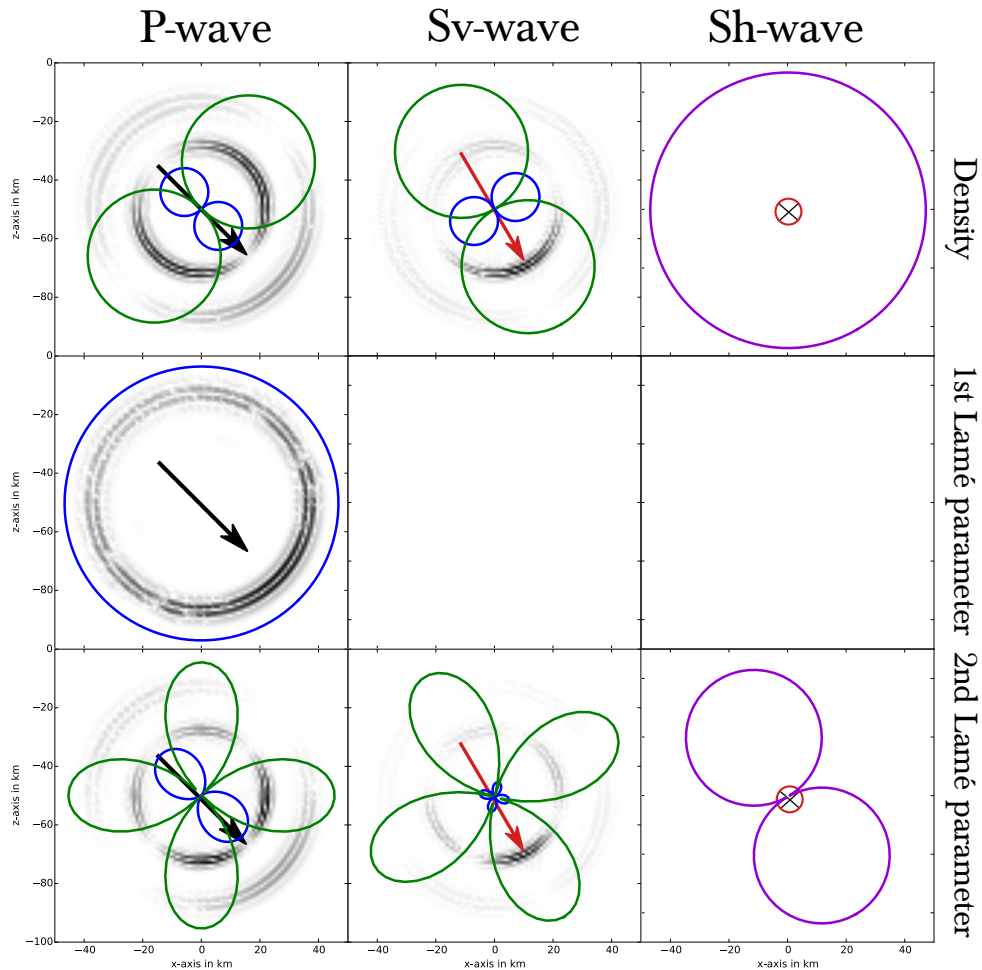


Figure 1.7 – Diagrammes de radiation pour une perturbation élastique isotrope dans la paramétrisation (ρ, λ, μ) . Les flèches noires et rouges indiquent les directions de propagation respectives des ondes incidentes P et Sv. La flèche rouge et noir indique la direction de propagation de l'onde incidente Sh. Les diagrammes bleus représentent l'amplitude des ondes P diffusées tandis que les diagrammes verts indiquent celles des ondes S. Le violet correspond à l'amplitude de l'onde Sh diffusée. Un exemple numérique de la diffusion est donnée en arrière plan de chaque interaction

La figure 1.6 illustre la diffusion du champ d'onde par une petite hétérogénéité de vitesse S selon l'approche de Born. Sous l'action de l'onde plane P incidente, le point diffusant émet deux champs d'onde sphériques sous forme d'onde P et d'onde S. Le champ d'onde total résulte de la sommation des champs incidents et diffusés. On remarque que l'amplitude des ondes rayonnées par l'hétérogénéité varie en fonction de la direction de propagation. Cette directivité peut se décrire au moyen des diagrammes de radiation¹⁵. Ces derniers sont déterminés analytiquement, dans le cadre de l'élasticité linéaire isotrope, en explicitant les fonctions de Green \mathbf{G}' par leur expression analytique et en choisissant une solution en onde plane P, Sv ou Sh pour le champ incident \mathbf{u}_o dans l'expression :

$$\delta \mathbf{u} = \delta \rho \omega^2 \mathbf{G}' \mathbf{u}_o - \delta \lambda (\nabla \cdot \mathbf{G}') (\nabla \cdot \mathbf{u}_o) - \delta \mu (\nabla \mathbf{G}' : [\nabla \mathbf{u}_o + (\nabla \mathbf{u}_o)^T]). \quad (1.39)$$

Ces diagrammes sont présentés sur la figure 1.7 pour des hétérogénéités de densité et de paramètres de Lamé λ et μ . Nous voyons que, selon le paramètre considéré, l'énergie sismique est diffusée préférentiellement dans une direction ou une autre qui dépend non seulement du paramètre mais aussi du mode de diffusion (PP, PS, SP ou SS). En conséquence, l'empreinte dans les données d'une perturbation d'un paramètre particulier varie selon les positions respectives des sources, des hétérogénéités et des récepteurs. Dans un processus d'imagerie cela signifie que certaines perturbations du sous-sol ne pourront pas être imagées du fait d'une trop faible sensibilité des données. Précisons que le diagramme de rayonnement d'un paramètre peut varier en fonction de la classe de paramètres considérée (vitesses, impédances, paramètres de Lamé, etc.) qui constitue la paramétrisation du problème. Nous discutons de cette dépendance dans le chapitre 4.

1.6 Caractéristiques du champ d'onde téléseismique

Nous avons vu que la configuration téléseismique se distingue des configurations rencontrées en exploration sismique, en sismologie à sources contrôlées, et de celles rencontrées en tomographie locale et régionale par une acquisition source-récepteur particulière qui permet d'imager des structures profondes. Les sources impliquées sont des téléseismes nucléés, sporadiquement, en dehors de la zone d'étude à des distances de plusieurs milliers de kilomètres dont les mécanismes au foyer ainsi que les signatures temporelles sont mal contraints. Cette particularité fait que le champ d'onde téléseismique éclaire la cible lithosphérique par en-dessous et que ce dernier est relativement complexe du fait des multiples interactions entre le champ d'onde primaire et la structure globale de la Terre, occasionnées au cours de la propagation.

1.6.1 Les sources téléseismiques

Les champs incidents éclairant la cible lithosphériques présentent une grande variabilité en terme de contenu fréquentiel, de rapport signal sur bruit, d'angle d'incidence et d'azimut. Le facteur dominant est la distance épacentrale (source-station) considérée, typiquement de 30° à

15. Il s'agit de l'analogie des diagrammes de radiation caractérisant la source sismique.

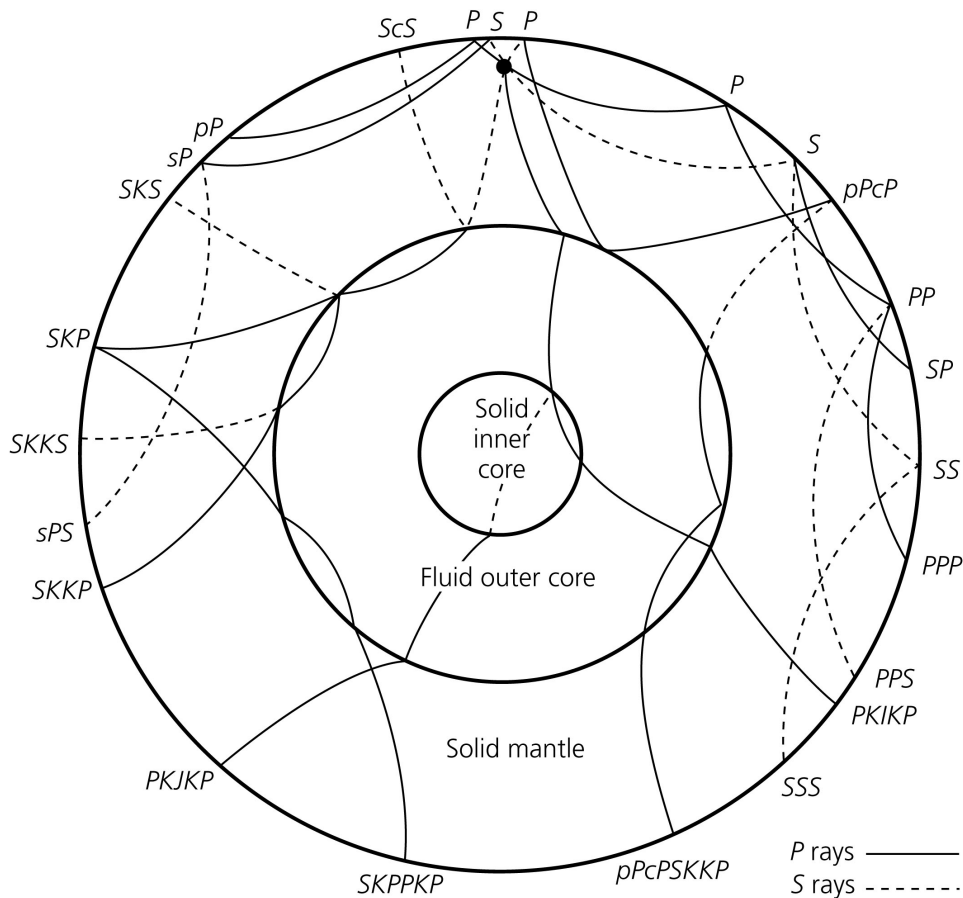


Figure 1.8 – Nomenclature des phases générées à l'échelle globale avec le chemin des rais correspondants (d'après Stein & Wysession (2009)).

100°, qui contraint l'angle d'incidence du champ d'onde téléseismique. Pour l'onde P directe, l'angle d'incidence varie ainsi de 60° à 25°. Un autre facteur limitant est la décroissance géométrique du champ d'onde, plus la distance de propagation est grande et plus l'onde P incidente est atténuée. De ce fait seuls des événements de forte magnitude, en général supérieure à 6, peuvent être pris en compte à grandes distances épicentrales pour obtenir un rapport signal sur bruit suffisant. Cependant, plus la magnitude augmente moins la source téléseismique peut être représentée par un point source. Pour des magnitudes comprises entre 7 et 9 les failles peuvent faire plusieurs centaines de km de long. La durée de rupture est ainsi plus longue et les périodes dominantes peuvent être supérieures à plusieurs dizaines de secondes. Pour des événements relativement proches, en revanche les magnitudes considérées peuvent être plus faibles et l'approximation en point source reste valable. Les périodes impliquées sont en conséquence plus faibles.

1.6.2 Ondes générées à l'échelle globale, phases téléseismiques

À l'échelle globale les ondes P et S générées par un séisme interagissent avec la structure radiale de la Terre. La Terre peut en effet, en première approximation, être assimilée à une superposition de couches concentriques dont les vitesses P et S sont soit constantes (croûte), soit des gradients de vitesse (manteau, noyau interne) ou dans le cas spécifique du noyau externe nulle pour les ondes S (fluide). En conséquence, les ondes P et S se convertissent et se réfléchissent de façon multiple à ces différentes interfaces et génèrent les phases dites téléseismiques. La figure 1.8 présente les chemins et dénominations de quelques-unes de ces phases. La nomenclature utilisée est la suivante :

P – S Ondes P et S suivant un chemin direct de la source au récepteur.

p – s Réflexion des ondes P et S à la surface du globe (proche du séisme, incidence $> 90^\circ$).

c Réflexion d'une onde P ou S à la limite noyau-manteau.

K Une onde P se propageant dans la noyau externe (fluide).

I – J Des ondes P et S se propageant dans la partie interne du noyau.

i Réflexion sur l'interface noyau externe-noyau interne.

De ce fait, une onde PP est une onde qui s'est réfléchiée une fois à la surface de la Terre avant d'être enregistrée et une onde PcP est une onde qui s'est réfléchiée une fois à la limite noyau-manteau. De même, une onde PKiKS est une onde P propagée dans le manteau, transmise dans le noyau externe, réfléchiée sur le noyau interne, et enfin convertie en onde S dans le manteau.

Ces différentes phases renforcent l'illumination de la cible lithosphérique fournie par l'onde incidente P. Les temps d'arrivées et les chemins empruntés sont en effets distincts de celui de l'onde P directe et varient de manière différente. Pour des distance épacentrales de 30° à 90° , une onde PcP arrivera aux stations avec des angles d'incidences compris entre 14° et 25° . Le temps d'arrivée différentiel de la PcP par rapport à la P varie, lui aussi, en fonction de la distance épacentrale (Fig 1.9). Ces temps sont approximativement de 3 minutes à 30° , 80 secondes à 50° , 20 secondes à 70° et sont identiques à la P pour des distances supérieures à 90° . L'utilisation de la phase PcP peut ainsi poser des problèmes en FWI. À des distances épacentrales faibles et intermédiaires le chemin qu'elle empreinte est totalement différent de celui de la P directe et dépend de la topographie de la CMB (limite noyau-manteau). Intégrer la PcP dans l'inversion présente donc le risque de générer des résidus dans le problème inverse dus à la modélisation imprécise de phénomènes de propagation d'ondes se produisant à l'extérieur de la zone d'étude.

Similairement, les phases profondeur (ou multiples à la sources) pP et sP peuvent aussi compromettre l'estimation de la source et polluer les données téléseismiques. Ces phases sont créées lorsque le séisme se déclenche en profondeur (Fig 1.10). Des ondes P et S rayonnées vers le haut se réfléchissent et se convertissent à la surface libre pour générer les ondes pP et sP. Ces phases, puisque la source est généralement peu profonde, suivent approximation le même chemin de rai que la P directe. Elles ont donc la même incidence au niveau de la cible lithosphérique. Leurs arrivées différentielles dépendent fortement de la profondeur du séisme : 5 et 8 s pour un séisme à 20 km de profondeur, 25 et 35 s pour un séisme à 100 km de profondeur, plus d'une minute pour un séisme à 600 km de profondeur. Dès lors, toute erreur sur l'estimation

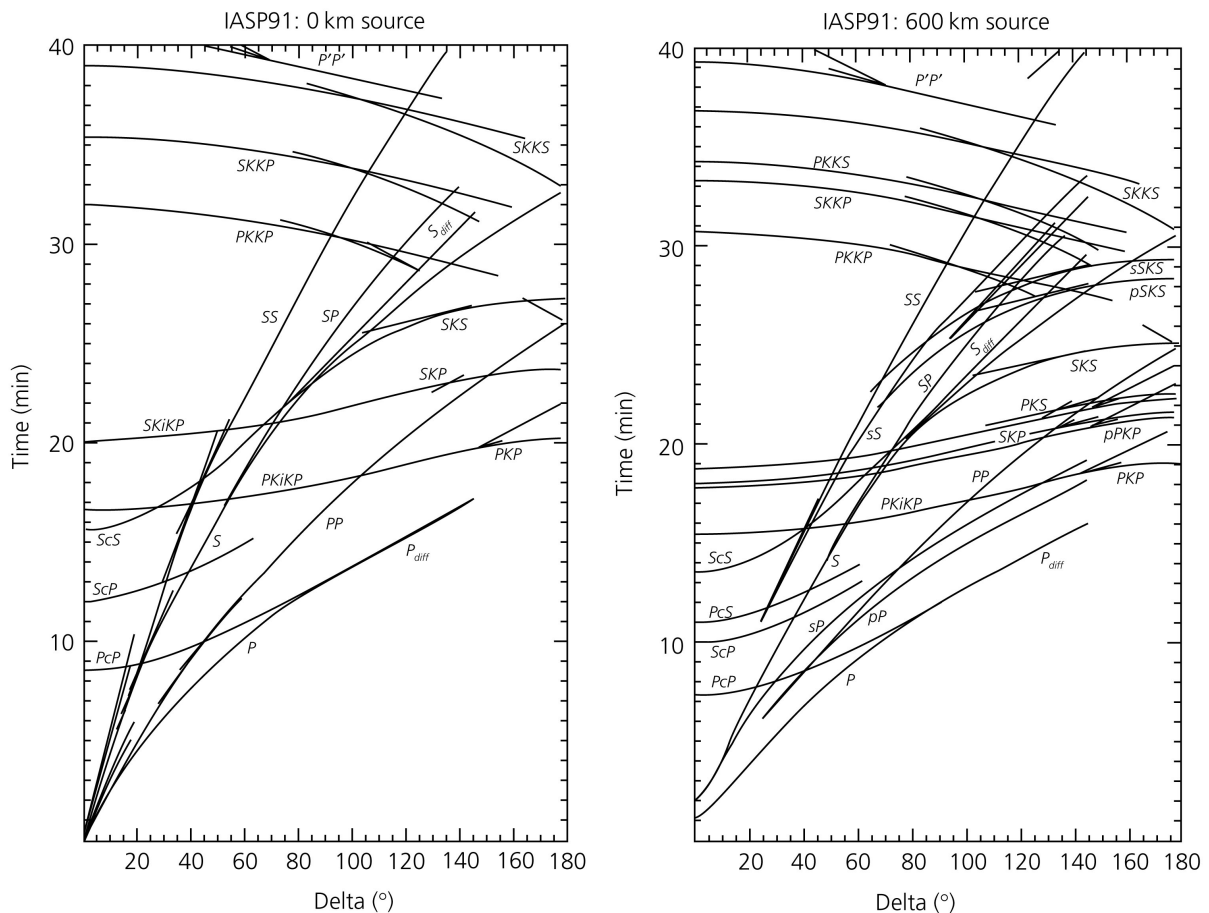


Figure 1.9 – Hodochrones des phases générées à l’échelle globale. A gauche pour une source en surface et à droite pour une source à 600 km de profondeur. D’après Stein & Wyession (2009).

de la profondeur¹⁶ de l’hypocentre peut se répercuter sur l’estimation de la signature de la source lors de la FWI.

1.6.3 Multiples lithosphériques

Au contact des hétérogénéités lithosphériques, les ondes incidentes créées à l’échelle globale génèrent de multiples arrivées P et S par transmission, réflexion et conversion lors de leurs interactions avec les hétérogénéités lithosphériques. La géométrie de ces interactions pour un milieu lithosphérique à une couche – spécifiquement la croûte et le manteau – est illustrée sur la figure 1.11. Ces interactions sont de deux types :

- les conversions en transmission au travers de l’interface (les ondes montantes PS et SP)
- les conversions à l’interface des ondes descendantes (provenant de la surface libre).

16. Fréquemment mal contrainte, du fait d’un fort couplage entre l’estimation du temps origine et de la profondeur

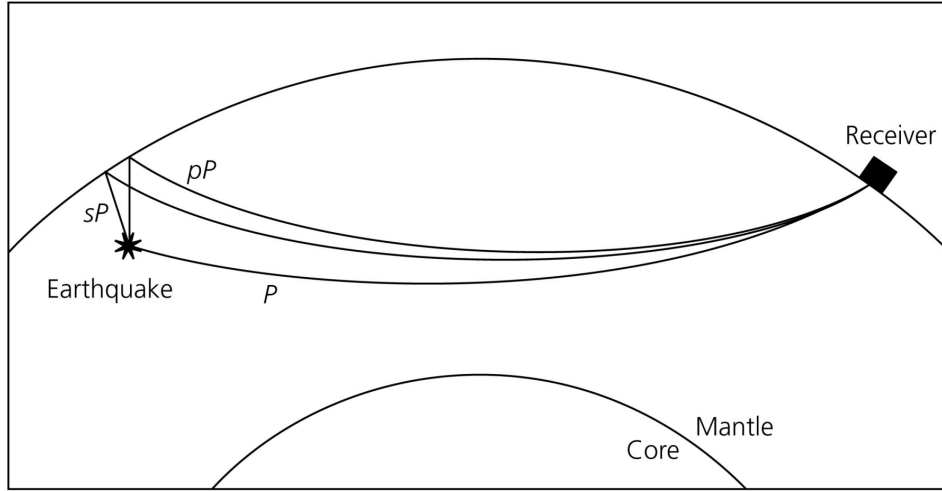


Figure 1.10 – Multiples pP et sP générés par la surface libre du côté source. Les temps d’arrivées de ces phases dépend de la profondeur de la source. Leurs paramètres de rai sont généralement proches de celui de la P directe. D’après [Stein & Wysession \(2009\)](#).

Ces ondes générées à proximité de la surface et du réseau de capteur arrivent ainsi dans les quelques secondes suivant l’onde P directe¹⁷. Les temps d’arrivée différentiels entre l’onde directe et les différentes conversions/réflexions sur l’interface peuvent être calculés géométriquement en appliquant la loi de Snell-Descartes. Ils sont obtenus à partir des relations suivantes ([Gurrola et al., 1994](#); [Zhu & Kanamori, 2000](#)) :

$$\Delta t_{pc'}^{\text{inter}} = h \left(\sqrt{\frac{1}{V_{c'}^2} - p_i^2} - \sqrt{\frac{1}{V_p^2} - p_i^2} \right) \quad (1.40)$$

dans le cas d’une transmission ; et dans le cas de multiples créés par la surface libre :

$$\Delta t_{pcc'}^{\text{inter}} = h \left(\sqrt{\frac{1}{V_{c'}^2} - p_i^2} + \sqrt{\frac{1}{V_c^2} - p_i^2} \right) \quad (1.41)$$

où h est la profondeur de l’interface, $p = \sin(i)/V$ le paramètre du rai, et où les indices c et c' indiquent les types d’ondes respectivement descendantes et montantes.

Les arrivées des multiples lithosphériques correspondent à des droites parallèles à l’arrivée de l’onde directe (Fig. 1.12, les lignes en tiretés). Comme pour la surface libre, nous pouvons déterminer les coefficients de réflexion et de transmission des ondes montantes et descendantes à l’interface (Fig. 1.13) et déterminer l’amplitude des multiples lithosphériques. Nous déduisons de la figure 1.13a que l’amplitude d’une conversion PS en transmission n’excède pas 20 % de l’amplitude de l’onde P incidente. De même les conversions et réflexions à l’interfaces dues aux ondes descendantes sont de l’ordre de 20 % pour les gammes d’incidences rencontrées dans les applications télésismiques. Puisque les ondes provenant de la surface libre ont des amplitudes

17. Le calcul des fonctions récepteurs vise notamment à extraire cette information des sismogrammes.

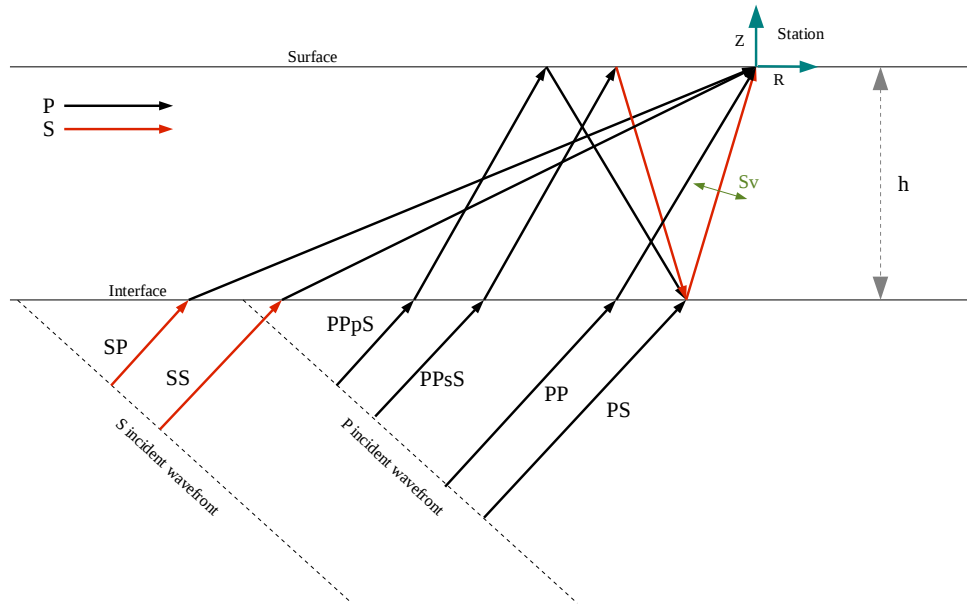


Figure 1.11 – Schéma illustrant la géométrie des conversions lithosphériques pour une interface plane située à une profondeur h . Les flèches noires et rouges indiquent des chemins en ondes P et Sv respectivement.

comparable à celle de l'onde incidente, nous concluons que l'amplitude des multiples de la surface libre réfléchies sur les réflecteurs lithosphériques est de l'ordre de 20 % de l'amplitude du champ d'onde incidente. Aussi, nous estimons que les multiples autres que ceux générés par la surface libre présentent une amplitude trop faible pour être discernés au sein des signaux télésismiques.

Si l'on considère un point diffusant, nous pouvons à nouveau déduire les temps d'arrivée différentiels des phases diffusées en appliquant les règles de l'optique géométrique. Elles s'expriment en fonction de la profondeur du point diffusant h , du paramètre du rai incident p_i et du paramètre du rai liant le point diffusant et la station $p' \neq p_i$ ¹⁸. Pour les ondes strictement montantes, nous obtenons :

$$\Delta t_{pc'}^{\text{diff}} = h \left(\frac{1 - p_i V_c^2 p_c'}{V_c' \sqrt{1 - p_c'^2 V_c'^2}} - \sqrt{\frac{1}{V_p^2} - p_i^2} \right), \quad (1.42)$$

et dans le cas de multiples créés par la surface libre :

$$\Delta t_{pcc'}^{\text{diff}} = h \left(\frac{1 - p_i V_c^2 p_c'}{V_c' \sqrt{1 - p_c'^2 V_c'^2}} + \sqrt{\frac{1}{V_c^2} - p_i^2} \right). \quad (1.43)$$

Les multiples lithosphériques créés par un point diffusant suivent des trajectoires hyperboliques dans le référentiel distance-temps (Fig. 1.12, les lignes pleines). L'amplitude du champ diffusé

18. L'égalité des deux paramètres correspond à loi de Snell-Descartes qui n'est valide que pour une interface, on peut vérifier que dans ce cas les équation (1.40) et (1.41) sont retrouvées.

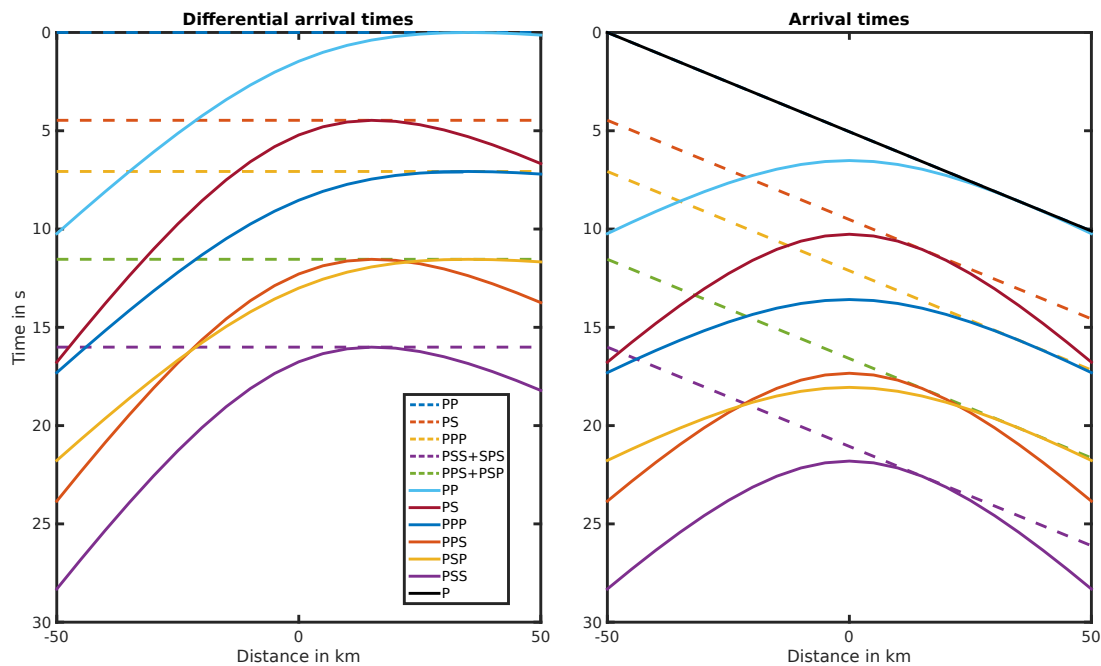


Figure 1.12 – Hodochrones lithosphériques pour un point diffractant (lignes continues) et une interface plane (lignes tiretées). Le point diffractant et l’interface sont situés à une profondeur $h = 35$ km dans un milieu de vitesse $V_P = 7$ km.s⁻¹ et $V_S = 4$ km.s⁻¹, l’onde P directe arrive avec une incidence de 45°. À gauche, les temps différentiels, à droite les hodochrones.

peut être calculée au moyen des diagrammes de radiation présentés dans la section 1.5.2.

Notons que dans les cas présentés jusqu’ici nous n’avons considéré que les conversions P en S pour lesquels tous les multiples lithosphériques sont enregistrés après l’arrivée de l’onde P directe. Si plusieurs interfaces ou points diffusants sont présents les multiples générés peuvent se superposer et rendre leur identification difficile. Au contraire, si l’on considère des conversions S en P, toutes les conversions SP arriveront avant l’onde S directe et tous les multiples liés à la surface libre après l’arrivée de l’onde S. Les conversions SP peuvent ainsi être détectées sans être polluées par les multiples émanant d’autres interfaces. Les temps différentiels pour les ondes S sont facilement obtenus en remplaçant V_P par V_S dans les équations (1.40 – 1.43).

1.6.4 Données synthétiques pour un modèle lithosphérique des Alpes Occidentales

Afin d’illustrer la complexité ainsi que la richesse des champs d’onde télésismiques nous présentons deux exemples synthétiques d’enregistrements télésismiques (Fig. 1.14) pour deux sources de périodes dominantes relativement élevées de deux secondes. L’acquisition est une ligne Ouest-Est dont les capteurs sont espacés de 5 km. La cible lithosphérique est représentative de la structure des Alpes Occidentales et présente deux bassins sédimentaires à faibles vitesses (V_S), une écaïlle de manteau indentée dans la croûte à faible profondeur (Ivr) et deux Moho distincts : plan à l’Est et plongeant jusqu’à une profondeur de 100 km à l’Ouest.

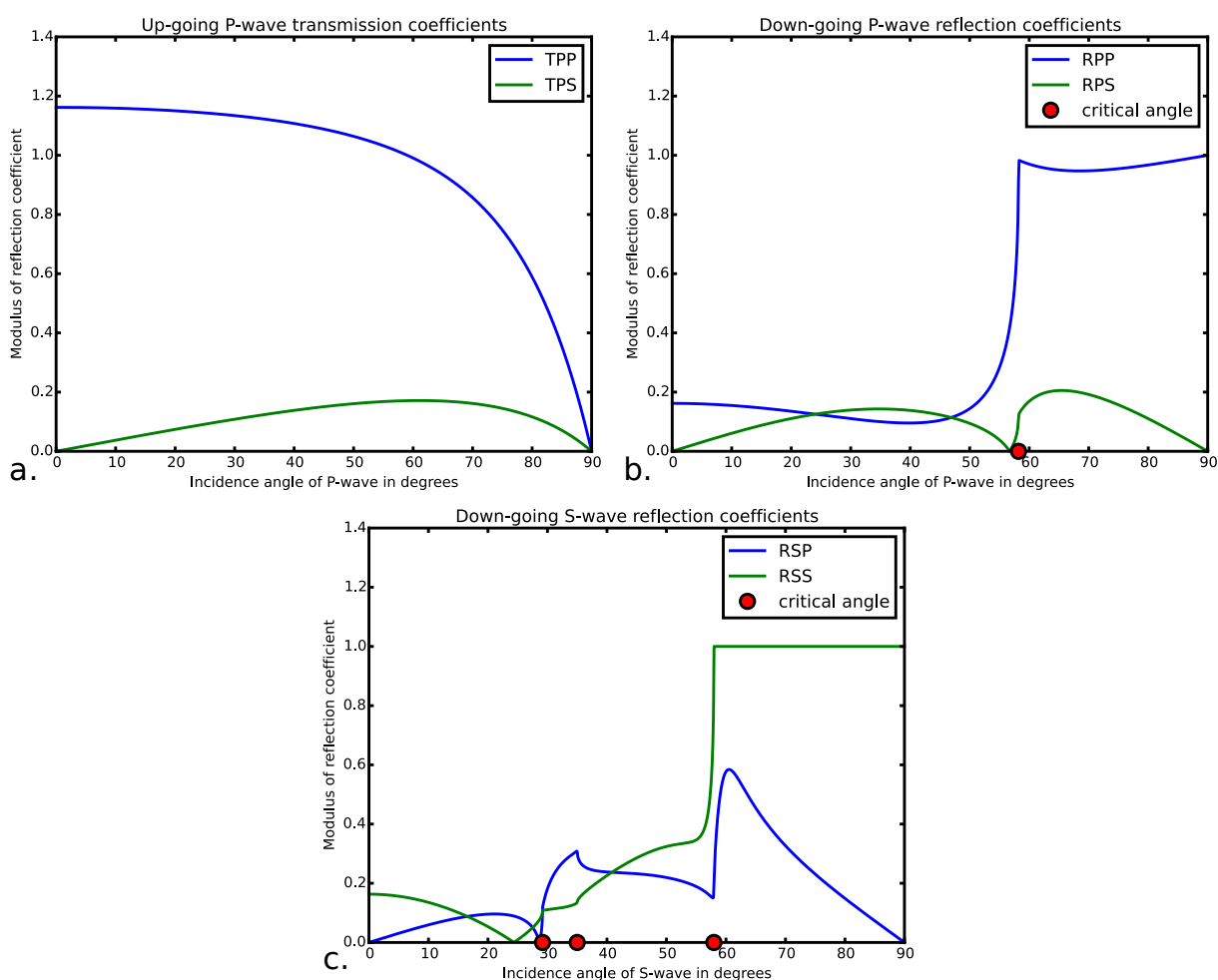


Figure 1.13 – Coefficients de réflexion et transmission pour un milieu à deux couches. (a.) Coefficients de transmission d’une onde P montante. (b.) et (c.) coefficients de réflexion pour une onde P et une onde S descendantes respectivement. Les couches supérieure et inférieure sont caractérisées par des densités de 2.8 et 3.3 $\text{g}\cdot\text{cm}^{-3}$, des vitesses P de 6.8 et 8 km/s et des vitesses S de 3.9 et 4.6 km/s. Les pastilles rouges indiquent les différents angles critiques.

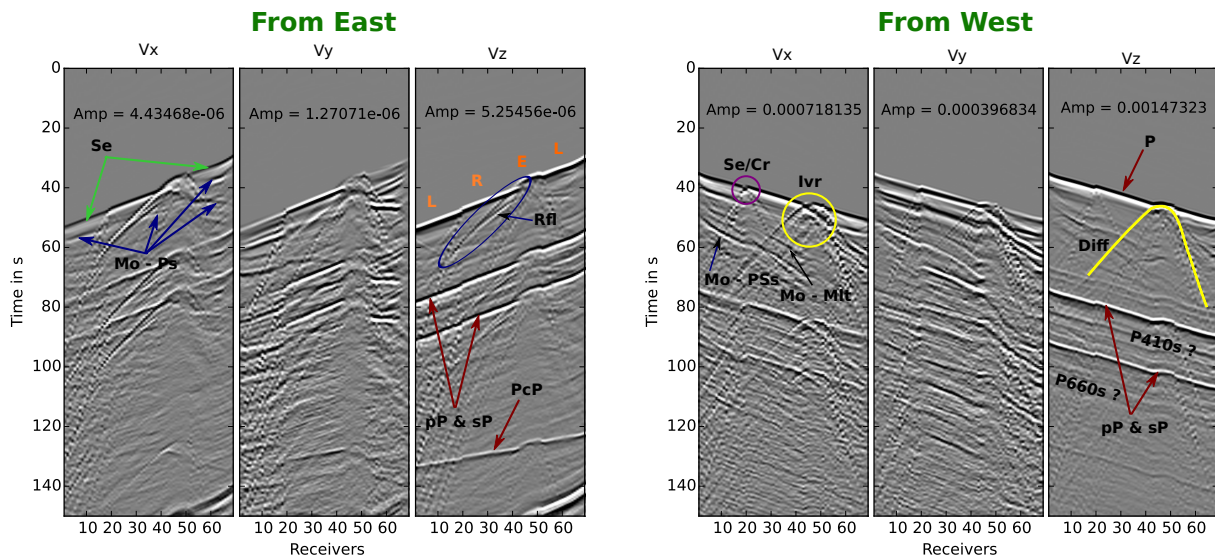


Figure 1.14 – Exemple de données télésismiques synthétiques calculées pour un modèle lithosphérique représentatif des Alpes. A gauche, sismogrammes pour un événement situé à 48° et éclairant le réseau de capteurs par l'Est. A droite, sismogrammes pour un événement situé à 83.6° et éclairant les capteurs par l'Ouest. Les données sont ici calculées pour une période dominante de 2s.

Le premier événement, éclaire le réseau par l'Ouest (avec un back-azimut de $75,4^\circ$) et correspond à un événement télésismique localisé à une distance épacentrale de 48° et une profondeur de 160 km. Le second éclaire le réseau par l'Ouest (avec un back-azimut de $266,4^\circ$) et correspond à un événement télésismique localisé à une distance épacentrale de $83,6^\circ$ et une profondeur de 110 km.

La profondeur intermédiaire des sources permet facilement d'identifier les phases profondeur pP et sP. Ces ondes sont, en effet, très énergétiques, de polarité opposée à l'onde P et leurs pentes sont parallèles à celle de l'onde P. Leurs temps d'arrivée différentiels théoriques par rapport à la P sont d'environ 24 et 37 secondes pour le premier événement, et de 42 et 55 secondes pour le second. Les sismogrammes générés par l'événement provenant de l'Est contiennent une arrivée énergétique plus tardive (environ 38 secondes après la P), de même polarité que l'onde P et dont la pente est plus horizontale. Il s'agit de l'onde PcP, visible du fait de la proximité du séisme. En revanche, cette phase n'est pas discernable sur l'événement provenant de l'Ouest car pour de grandes distances épacentrales la PcP et la P directe suivent approximativement le même chemin et arrivent en même temps.

Les effets des structures lithosphériques sont aussi très marquées sur les enregistrements. Premièrement, les arrivés globales subissent des déphasages et variations d'amplitudes créés par les variations de vitesses et de densité en proche surface. Ainsi, on remarque des arrivés tardives d'amplitude moindre (L) au niveau des bassins sédimentaires, précoces et d'amplitudes élevées (E) au niveau de l'écaïlle mantellique par rapport aux arrivés attendues (R). Les discontinuités entre ces structures sont soulignées par des hyperboles de diffraction (Diff, Iv, Se/Cr) très pentées. En outre, l'empreinte des variations géométriques du Moho sont visibles

sur les ondes converties Ps (Mo-Ps) et PSS (Mo-PSS). Bien que d'autres multiples du Moho sont discernables (Mo-Mlt), les conversions les plus énergétiques sont les Ps et PSS. Enfin, il est également possible de discerner la réflexion des ondes diffractées par l'écaïlle mantellique sur le Moho (Rfl) qui apparaît sous la forme d'une arrivée à double branchement correspondant à une hyperbole de réflexion.

Les enregistrements de l'événement provenant de l'Ouest montrent également les multiples lithosphériques reliés aux discontinuités de la zone de transition du manteau (à 410 et 660 km). Ces interfaces situées en dehors de la cible lithosphériques génèrent de faibles conversions Ps qui arrivent généralement entre 42 et 47 secondes après l'onde P pour la discontinuité à 410 km et entre 65 et 75 s pour celle à 660 km de profondeur. La topographie de ces interfaces peut varier selon les zones d'études et il est ainsi risqué de les considérer pour l'inversion de formes d'ondes si le modèle mis à jour ne les inclut pas.

Conclusion

Ce chapitre nous a permis de présenter les lois régissant le phénomène de propagation des ondes sismiques et de décrire l'allure générale des données télésismiques. Nous avons vu que les champs d'onde télésismiques résultent de multiples interactions entre le champ d'onde généré par la source du séisme et la complexité de la structure de la Terre, tant à l'échelle globale que locale. Ces interactions de nature plus ou moins complexes peuvent être modélisées au moyen de différentes approximations (milieux stratifiés, approximation de Born pour la diffusion). Néanmoins, pour modéliser de manière complète les champs d'ondes télésismiques, et ce pour des milieux arbitrairement hétérogènes, il est nécessaire d'employer des méthodes de modélisation rendant compte de tous les effets de propagation. Nous présentons dans le chapitre suivant les méthodes numériques de résolution des équations de l'élastodynamique permettant ce réaliser type de simulations.

Avant de conclure ce chapitre, précisons que d'autres facteurs peuvent impacter les champs d'onde. Tout d'abord, les effets d'atténuation des ondes qui apparaissent du fait que les milieux ne sont pas élastiques. Cette atténuation est séparée en deux catégories, l'atténuation due à la diffusion des ondes qui se caractérise par la partition de l'énergie sismique entre les différentes ondes converties ; et l'atténuation intrinsèque due à la dissipation de l'énergie sismique en chaleur du fait de l'anélasticité du milieu. Cette dernière impacte fortement les hautes fréquences des signaux sismiques et est décrite par des facteurs de qualité qui caractérisent la perte d'énergie sur un cycle de propagation. Elle est donc importante à l'échelle de propagation globale (de nombreux cycles sont parcourus) et aux échelles très locales (sismique de proche surface où les données sont très hautes-fréquences). À l'échelle lithosphérique les fréquences considérées ainsi que le nombre de cycles propagés font que les effets d'atténuation intrinsèque peuvent être négligés. Enfin, un autre facteur important à haute fréquence est l'irrégularité de la topographie (Bleibinhaus & Rondenay, 2009; Bouchon et al., 1996). Cette dernière génère de fortes diffractions et donne naissance à des ondes de surfaces (Monteiller et al., 2013), du fait de la conversion des ondes P en ondes de Rayleigh, fortement visibles sur les enregistrements télésismiques (Fig. 1.15).

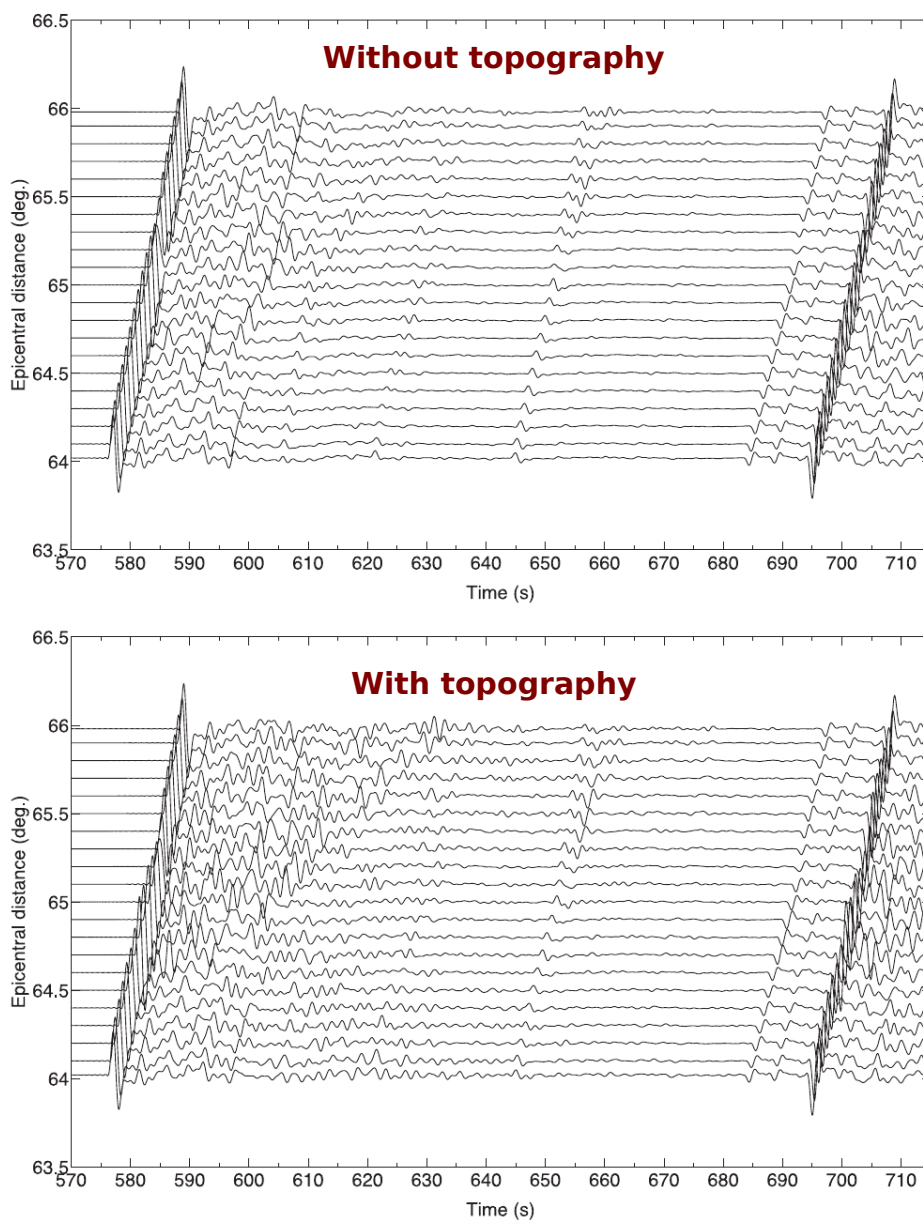


Figure 1.15 – Effets d’une topographie irrégulière sur les données télésismiques, adapté de [Monteiller et al. \(2013\)](#). En haut, des données télésismiques synthétiques sans topographie. En bas, même données avec prise en compte d’une topographie complexe.

Chapitre 2

Modélisation numérique

Sommaire

2.1	Résolution d'une équation aux dérivées partielles	70
2.1.1	Les différences-finies	71
2.1.2	Les éléments-finis	72
2.2	Méthode des éléments spectraux	73
2.2.1	Forme faible des équations de l'élastodynamique	74
2.2.2	Discretisation : maillages et fonctions de forme	75
2.2.3	Représentation polynomiale et intégration sur un élément	76
2.2.4	Intégration numérique de la forme faible de l'équation d'onde	79
2.2.5	Intégration temporelle – Le schéma de Newmark	81
2.2.6	Implémentation	83
2.2.7	Exemples de modélisation et validations	86
2.3	Stratégie d'injection de champ d'onde	92
2.3.1	L'approche champ total / champ diffracté	93
2.3.2	Injection d'une onde plane	94
2.4	Injection d'un champ d'onde téléseismique avec AxiSEM	96
2.4.1	AxiSEM	97
2.4.2	Aspects pratiques du couplage	98
2.4.3	Exemple et validation du couplage SEM/AxiSEM	101

Introduction

L'inversion de formes d'ondes en trois dimensions nécessite de modéliser la propagation des champs d'ondes dans des milieux arbitrairement hétérogènes pour lesquels il n'existe pas de solutions analytiques ou pseudo-analytiques connues. Nous devons ainsi, afin de résoudre le problème aux dérivées partielles, avoir recours à des méthodes de résolution numérique capable d'approcher la solution du problème physique considéré. De nombreuses méthodes permettent de faire cela, chacune possédant ses avantages et ses inconvénients. Dans ce chapitre, nous présentons succinctement les grandes catégories de techniques numériques qui permettent de résoudre le problème de propagation des ondes élastiques, et en analysons les limitations. Nous nous focalisons, en particulier, sur la méthode des éléments finis spectraux et détaillons son implémentation numérique pour des simulations élastiques 3D en domaine temporel. Nous présentons également une approche d'injection de champs d'ondes permettant de propager dans le milieu d'étude lithosphérique un champ d'ondes téléseismiques, préalablement calculé dans une Terre globale à symétrie sphérique avec le logiciel AxiSEM (Nissen-Meyer et al., 2014). Les méthodes présentées sont finalement validées dans des configurations expérimentales canoniques par comparaison entre les ondes modélisées numériquement et celles obtenues grâce aux solutions analytiques correspondantes.

2.1 Résolution d'une équation aux dérivées partielles

De multiples applications, allant de l'évaluation de sismogrammes en sismologie globale à l'inversion de formes d'onde, requièrent la modélisation de la propagation des ondes sismiques aussi précises que possible. S'il est possible de trouver des solutions analytiques des équations de l'élastodynamique (1.9) pour certaines configurations simples du milieu de propagation, ce fut notamment le cas au chapitre précédant pour des espaces ou demi-espaces élastiques homogènes (Aki & Richards, 2002; Kausel, 2006) ou encore des solutions semi-analytiques telles que les méthodes de réflectivité (Kennett & Kerry, 1979) et F-K (Bouchon & Aki, 1977; Coutant, 1989; Zhu & Rivera, 2002) pour des milieux tabulaires, de telles solutions n'existent pas pour des milieux arbitrairement hétérogènes. Pour pallier ce problème, des méthodes numériques ont été développées afin de modéliser certaines caractéristiques des ondes sismiques propagées dans ce type de configurations. Notamment, pour des applications tomographiques par inversion des temps de la première arrivée, des approches asymptotiques par tracé de rayons ont été développés dans le cadre de l'approximation haute fréquence de la théorie des rais (Červený et al., 1977; Červený, 2001). La résolution de l'équation iconale permet de décrire uniquement le comportement des temps de trajet, elle peut toutefois être complétée par la résolution de l'équation de transport fournissant l'amplitude de l'arrivée. Des évolutions ultérieures (Dahlen et al., 2000b), ont permis d'enrichir ces méthodes par la prise en compte des effets de fréquences-finies. Néanmoins, ces méthodes ne restent que des approximations qui ne peuvent être appliquées que pour des milieux peu perturbés et lisses. Dans le cadre de la FWI, tous les effets de propagation doivent être pris en compte et ce, pour des milieux arbitrairement complexes. Il est dès lors nécessaire de se tourner vers des méthodes numériques qui permettent de résoudre les équations aux dérivées partielles (EDP) qui régissent le phénomène de propa-

gation d'onde. Plusieurs catégories de méthodes sont disponibles, notamment : les méthodes aux différences-finies (Moczo et al., 2006), les méthodes aux éléments-finis (Zienkiewicz et al., 2005), les méthodes aux volumes-finis (LeVeque, 2002) ou les méthodes de Boltzmann sur réseaux (O'Brien et al., 2012; O'Brien, 2014). Nous présentons rapidement les caractéristiques des deux premières techniques qui sont, à l'heure actuelle, les plus répandues en géophysique.

2.1.1 Les différences-finies

La méthode aux différences-finies (DF) (Alterman & Karal, 1968; Kelly et al., 1976; Virieux, 1986; Levander, 1988; Graves, 1996; Moczo et al., 2006; Taflove & Hagness, 2005) développée en géophysique depuis la fin des années 60, est probablement la méthode de résolution des EDP la plus utilisée et la plus répandue dans le domaine du fait de sa rapidité d'exécution et de sa simplicité de mise en œuvre. La méthode DF repose sur la formulation forte des EDP et propose de résoudre le problème en discrétisant le milieu physique ainsi que ses solutions sur un ensemble fini de points répartis en général sur une grille cartésienne régulière. Les dérivées partielles sont alors approximées en exprimant la solution par des développements de Taylor tronqués à des ordres qui dépendent de la précision numérique recherchée :

$$f(x + \Delta x) = f(x) + \Delta x f'(x) + \frac{\Delta x^2}{2} f''(x) + \dots + \frac{\Delta x^n}{n!} f^n(x) + \mathcal{O}(\Delta x^{n+1}) \quad (2.1)$$

ou

$$f(x - \Delta x) = f(x) - \Delta x f'(x) + \frac{\Delta x^2}{2} f''(x) + \dots + (-1)^n \frac{\Delta x^n}{n!} f^n(x) + \mathcal{O}(\Delta x^{n+1}). \quad (2.2)$$

Un exemple d'approximation d'une dérivée spatiale est la DF centrée, précise au second ordre, obtenue en combinant les deux développements de Taylor précédents. Pour un champ d'onde u au point i , elle est donnée en fonction des champs aux points voisins $i + 1$ et $i - 1$ par

$$\partial_x u_i \approx \frac{u_{i+1} - u_{i-1}}{2h} \quad (2.3)$$

où h est le pas de grille. L'équation (2.3) est un exemple de schéma numérique à une dimension. De nombreux schémas numériques peuvent être envisagés, en 3D, citons particulièrement les schémas *full-grid* (Kelly et al., 1976), *staggered-grid* (Virieux, 1986), et *partly-staggered grid* (Saenger et al., 2000). Ces schémas ont chacun leurs qualités et leurs défauts en terme de dispersion, d'anisotropie et de stabilité numérique. Ils sont ainsi plus ou moins bien adaptés aux configurations des problèmes étudiés.

Bien que simples en théorie, les méthodes aux différences-finies souffrent de nombreux inconvénients intrinsèquement liés à l'utilisation de grilles régulières. Le pas d'échantillonnage spatial de ce type de grille est constant et est contraint par la vitesse minimale présente dans le milieu de propagation. Les relations de dispersion numérique imposent que la plus petite longueur d'onde propagée soit représentée par un nombre minimal de points (ex. 5 points pour le schéma de Levander (1988)). Même si la zone de plus faible vitesse est très restreinte comparativement à la taille du domaine étudié, la grille se doit de respecter cette condition et le

nombre de nœuds du maillage peut devenir très conséquent, limitant ainsi les domaines d'application de la méthode. Un autre inconvénient majeur des DF provient de la gestion explicite des conditions aux frontières et en particulier de la surface libre. Pour des interfaces planes, cette condition peut-être implémentée par des techniques d'antisymétrisation de contrainte « *image methods* » (Levander, 1988; Ohminato & Chouet, 1997; Zhang & Chen, 2006) sur le bord supérieur de la grille considérée. Pour des topographies complexes, en revanche, les techniques consistent soit à simuler la présence de vide dans le milieu « *vacuum formulation* » (Zeng et al., 2012), la topographie est dès lors représentée sous forme de marches d'escalier ce qui nécessite d'augmenter considérablement le nombre de points par longueur d'onde (Bohlen & Saenger, 2006) pour obtenir une précision satisfaisante ; soit à déformer le maillage pour que son bord supérieur puisse épouser les formes de la topographie (Tarrass et al., 2011) nécessitant là encore de raffiner fortement le maillage ; soit en considérant des interfaces immergées (Lombard et al., 2008; Gao et al., 2015) qui adaptent localement les coefficients du schéma numérique en accord avec la topographie. D'autre part, la majorité des schémas numériques tels que ceux de Virieux (1986) ou Levander (1988) ne sont pas compatibles géométriquement avec certaines formes d'anisotropie. Enfin, un dernier handicap est lié au positionnement de la source et des récepteur au sein de la grille. Lorsque ces positions ne sont pas colocalisées à un nœud de la grille, il est nécessaire de faire appel à des techniques d'interpolation (Hicks, 2002) pour obtenir un résultat précis.

2.1.2 Les éléments-finis

Lorsque des topographies complexes doivent être prises en compte, les méthodes DF deviennent difficilement applicables. La méthode aux éléments-finis (EF) (Zienkiewicz & Taylor, 1967; Marfurt, 1984) procure alors une alternative intéressante. Les EF, basés sur la forme variationnelle des EDP et une discrétisation irrégulière, permettent de restituer fidèlement des géométries complexes par l'usage d'un maillage adapté. Pour ce faire, le domaine de calcul est discrétisé en un nombre fini d'éléments de forme géométriques particulières – des tétraèdres ou des hexaèdres dans le cas de milieux 3D – qui définissent le maillage sur lequel est évaluée la solution. Chacun de ces éléments est lui-même discrétisé en un nombre fini de nœuds portant les propriétés physiques du milieu et auxquels sont attachés des fonctions de base, généralement polynômiales et interpolantes, permettant d'approximer les solutions recherchées. Les EF présentent de nombreuses variantes qui dépendent notamment du choix des fonctions de bases et des points de quadrature considérés. Elles peuvent être classées en deux sous-catégories : les EF continus pour lesquels les éléments voisins communiquent la solution par le biais de nœuds colocalisés sur leur frontière commune, et les EF discontinus pour lesquels ces nœuds sont propres à chaque élément (non forcément colocalisés) et où les solutions sont échangées par le biais de flux numériques aux frontières.

La méthode aux éléments spectraux (SEM), initialement développée en sciences de la Terre pour des milieux acoustique par Priolo et al. (1994) et Komatitsch (1997) pour les milieux élastiques, est certainement la technique d'élément-finis continus la plus utilisée actuellement en sismologie (Komatitsch & Vilotte, 1998; Komatitsch & Tromp, 1999b). Contrairement aux EF classiques les SEM utilisent un choix judicieux de fonctions de base (les polynômes de La-

grange) et de quadrature numérique (Gauss-Lobatto-Legendre) permettant d'obtenir une convergence spectrale de la solution avec l'élévation du degré polynômial. Ce choix particulier a pour effet de rendre diagonale par construction la matrice de masse du problème. La popularité de cette méthode résulte de sa précision et de son efficacité : un seul élément¹ par longueur d'onde est requis pour des polynômes d'ordre supérieurs à 4 ce qui permet de considérer des maillages assez grossiers. Le frein essentiel de cette méthode est l'utilisation de quadrangles et d'hexaèdres pour la discrétisation spatiale du problème. Lorsque les milieux sont fortement hétérogènes, il peut devenir difficile de générer un maillage respectant les interfaces. En particulier, dans le cas d'interfaces fluide-solide, de bathymétrie complexes ou de bord de mer généralement difficile à mailler et pour lesquelles les conditions de continuité doivent être respectées.

Une autre variante des EF activement développée ces dernières années est la méthode des EF de type Galerkin-Discontinus (GD) (Käser & Dumbser, 2006; Hesthaven & Warburton, 2007; Etienne et al., 2010; Wilcox et al., 2010; Ye et al., 2016). Contrairement aux SEM, les GD peuvent opérer sur des éléments tétraédriques qui simplifient la génération de maillages complexes. A l'inverse des EF continus, les GD représentent la solution de manière discontinue. Les éléments qui composent le maillage communiquent alors par le biais de flux numériques, sans forcément partager de nœuds communs. Cette particularité offre une grande flexibilité dans le choix des ordres d'interpolation entre des éléments voisins qui peuvent ainsi être localement adaptés, on parle de « *p-adaptivity* ». Elle rend aussi possible l'utilisation de maillages non-conformes. De même la caractéristique discontinue des GD facilite la considération des conditions de continuité pour des interfaces fluide-solide par le biais des flux numériques (Ye et al., 2016). Notons tout de même qu'à maillage et ordres d'interpolation équivalents, les GD considèrent un nombre plus important de degrés de liberté. En outre la méthode est nécessairement plus coûteuse que les EF continus puisqu'elle implique le calcul supplémentaire des flux numérique entre éléments voisins.

Dans la suite de ce travail, nous présentons en détail la méthode des éléments-finis spectraux qui sera utilisée pour modéliser le champ d'onde dans la cible lithosphérique et qui est aussi utilisée pour le calcul du champs d'onde téléseismique dans la Terre globale. Les raisons qui nous poussent à considérer les SEM sont principalement leur simplicité d'implémentation par rapport aux GD, leur rapidité d'exécution dans les cas élastiques et leur simplicité de gestion de la surface libre. En outre, la méthode utilisée pour générer le champ d'onde global (AxiSEM) est aussi une méthode aux éléments spectraux. En utilisant des SEM de même ordre polynômial, nous minimisons les erreurs qui pourraient être associées à l'utilisation de techniques de modélisation différentes lors du processus d'injection.

2.2 Méthode des éléments spectraux

Nous présentons dans cette section la méthode des éléments spectraux pour la modélisation de la propagation des ondes dans des milieux élastiques 3D. Les éléments théoriques développés sont en grande partie repris des développements de Komatitsch & Tromp (1999b) auquel nous invitons le lecteur à se référer pour de plus amples informations. Notamment, l'extension de la

1. 5 degrés de liberté à l'ordre 4.

méthode aux milieux anélastiques et à la Terre globale est présentée par Komatitsch & Tromp (2002a) et Komatitsch & Tromp (2002b)

2.2.1 Forme faible des équations de l'élastodynamique

On rappelle la forme forte des équations de l'élastodynamique (1.9), réécrite de sorte à rendre implicites les dépendances en temps et en espace

$$\rho \partial_t^2 \mathbf{u} = \nabla \cdot \boldsymbol{\sigma} + \mathbf{f}. \quad (2.4)$$

Contrairement aux différences-finies qui tentent de résoudre directement l'équation (2.4) en explicitant les conditions aux limites associées, les SEM cherchent à discrétiser et résoudre la forme variationnelle – aussi nommée forme faible ou intégrale – des équations de l'élastodynamique.

On considère le domaine de calcul Ω , borné par une frontière $\Gamma = \partial\Omega$, sur lequel nous cherchons à approximer la solution \mathbf{u} . La forme faible du problème est obtenue en multipliant les termes de l'équation (2.4) par une fonction test \mathbf{w} , quelconque et indépendante du temps, puis en intégrant sur le domaine d'évaluation Ω :

$$\int_{\Omega} \rho \mathbf{w} \cdot \partial_t^2 \mathbf{u} \, d^3\mathbf{x} = \int_{\Omega} \mathbf{w} \cdot (\nabla \cdot \boldsymbol{\sigma}) \, d^3\mathbf{x} + \int_{\Omega} \mathbf{w} \cdot \mathbf{f} \, d^3\mathbf{x}. \quad (2.5)$$

Les seules conditions que doivent vérifier les fonctions tests sont d'être, ainsi que leurs dérivés, de carré intégrable sur le domaine de définition Ω . En général, cela implique que ces fonctions soient continues et se « comportent bien » au sens où l'on peut les dériver suffisamment et qu'elles respectent les conditions aux limites du problème.

Dans l'équation (2.5), le premier terme du membre du droite contient la divergence du tenseur de contrainte $\nabla \cdot \boldsymbol{\sigma}$ qui fait intervenir des dérivées partielles secondes en espace du champ d'onde \mathbf{u} . En intégrant par partie sur le volume Ω et en appliquant le théorème de Green-Ostrogradsky (théorème de flux-divergence), nous trouvons :

$$\int_{\Omega} \rho \mathbf{w} \cdot \partial_t^2 \mathbf{u} \, d^3\mathbf{x} = \int_{\Gamma} \mathbf{w} \cdot (\boldsymbol{\sigma} \cdot \hat{\mathbf{n}}) \, d^2\mathbf{x} - \int_{\Omega} \nabla \mathbf{w} : \boldsymbol{\sigma} \, d^3\mathbf{x} + \int_{\Omega} \mathbf{w} \cdot \mathbf{f} \, d^3\mathbf{x}. \quad (2.6)$$

Cette expression est équivalente à (2.5), sauf qu'à présent, l'intégrale volumique impliquant le tenseur des contraintes a été remplacée par une intégrale surfacique faisant intervenir les tractions sur les bords du domaine $\mathbf{T} = \boldsymbol{\sigma} \cdot \hat{\mathbf{n}}$ et une intégrale volumique faisant intervenir le gradient de la fonction test. Cette manipulation permet ainsi de diminuer l'ordre de dérivation du champs d'onde \mathbf{u} au prix d'une élévation de celui des fonctions tests \mathbf{w} que nous pouvons déduire analytiquement.

L'analyse de l'équation (2.6) montre que la forme faible des équations de l'élastodynamique rend implicite la prise en compte des conditions aux frontières par le biais de l'intégrale de surface. En particulier, la condition de surface libre – ou condition de Neumann – implique que

la traction aux bords du domaine Γ , s'annule c'est-à-dire que

$$\boldsymbol{\sigma} \cdot \hat{\mathbf{n}} = 0. \quad (2.7)$$

Il en résulte que l'intégrale surfacique est nulle et la condition de surface libre est satisfaite de manière naturelle sans avoir à imposer de traitements particuliers. Cette simplicité de gestion de la surface libre, particulièrement lorsque des topographies complexes sont rencontrées, est un des avantages clés des méthodes aux éléments-spectraux et permet de modéliser précisément les ondes de surfaces.

Une autre condition aux frontières très importante dans le cadre de nos applications télésismiques est la condition de bords absorbants. Cette dernière permet de simuler la radiation du champ d'ondes en dehors de la cible télésismique, c'est-à-dire aux bords du domaine de calcul. La plus populaire des conditions aux frontières absorbantes est une approximation introduite par Clayton & Engquist (1977) et Stacey (1988) qui est basée sur les approximations paraxiales des équations de l'élastodynamique. L'implémentation que nous utilisons est une version modifiée de Clayton & Engquist (1977) présentée par Komatitsch et al. (1998) pour laquelle les tractions sur les bords du domaine Γ sont exprimées par :

$$\boldsymbol{\sigma} \cdot \hat{\mathbf{n}} = \rho [V_p (\hat{\mathbf{n}} \cdot \partial_t \mathbf{u}) \hat{\mathbf{n}} + V_s (\mathbf{I} - \hat{\mathbf{n}}\hat{\mathbf{n}}) \cdot \partial_t \mathbf{u}]. \quad (2.8)$$

Cette approximation est exacte pour des ondes se propageant orthogonalement aux surfaces absorbantes mais est de moins en moins précise au fur et à mesure que l'angle d'incidence de l'onde augmente. Elle donne alors lieu à des réflexions parasites émanant des bords du domaine. Pour des applications où l'approximation (2.8) n'est pas satisfaisante, des conditions absorbantes plus efficaces peuvent être envisagées, notamment, les « *Perfectly Matched Layers* » (Bérenger, 1994; Komatitsch & Martin, 2007; Zhinan et al., 2014) et les « *SMART layers* » (Tago et al., 2014). Néanmoins, l'efficacité de ces méthodes se fait au détriment du coût de calcul et elles sont malheureusement plus difficiles à mettre en œuvre.

2.2.2 Discrétisation : maillages et fonctions de forme

Les méthodes aux éléments-finis discrétisent le domaine d'évaluation Ω du problème (2.6) en un nombre fini d'éléments Ω_e où $e = 1, \dots, N_e$, de sorte que le domaine d'évaluation global soit représenté par l'union des N_e éléments sans recouvrement entre ceux-ci $\Omega = \cup_e^{N_e} \Omega_e$. De manière similaires, les surfaces qui nécessitent un traitement particulier, comme une condition absorbante, sont subdivisées de sorte que : $\Gamma = \cup_b^{N_b} \Gamma_b$ (Fig. 2.1a). Plusieurs géométries d'éléments peuvent être envisagées pour construire le maillage². Néanmoins la méthode des éléments spectraux privilégie l'utilisation d'hexaèdres (et de quadrangles pour la surface des éléments) (Fig. 2.1b) qui permettent d'exprimer les fonctions de base 3D sous forme de produit tensoriel des fonctions de base 1D (Komatitsch et al., 2005). Ces éléments sont déformés pour épouser les contours du domaine physique considéré.

2. Des tétraèdres peuvent aussi être utilisés. Dans ce cas, la convergence spectrale est obtenue en utilisant les points de Fekete (Pozrikidis, 2005) pour l'intégration, dont nous ne disposons pas de solutions analytiques

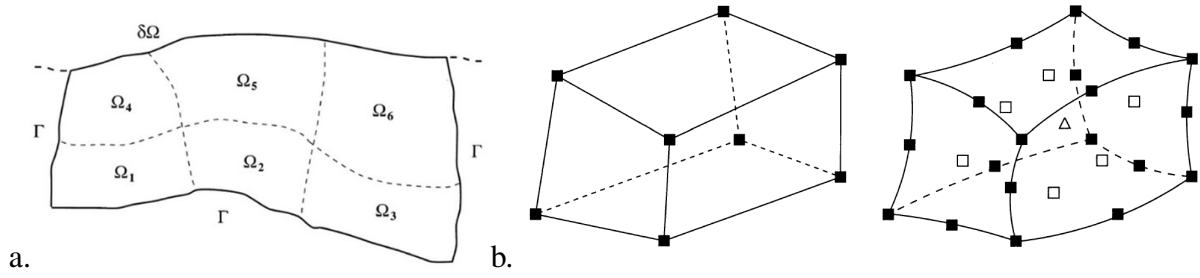


Figure 2.1 – Discretisation du domaine par des hexaèdres. (a.) Discretisation du domaine. (b.) Géométrie des hexaèdres à 8 et 27 points de contrôle. D’après Komatitsch & Tromp (1999a).

Les éléments hexaédriques étant isomorphes à un cube, il est possible de définir une matrice Jacobienne unique permettant d’effectuer la transformation d’un élément du maillage vers un élément de référence qui correspond à un cube. L’élément de référence est défini par ses coordonnées naturelles $\boldsymbol{\xi} = (\xi, \eta, \gamma)$ dont les valeurs sont comprises entre -1 et 1 . Chaque élément Ω_e est défini à partir de ses points de contrôle ou points d’ancrage $\mathbf{x}_a = \mathbf{x}(\boldsymbol{\xi}_a)$ et d’un jeu de fonctions de forme $N_a(\boldsymbol{\xi})$ au nombre de 8 pour un hexaèdre droit et au nombre de 27 pour un hexaèdre courbé (Fig. 2.1b). La passage d’un point physique \mathbf{x} de l’élément Ω_e au cube de référence s’exprime comme :

$$\mathbf{x}(\boldsymbol{\xi}) = \sum_{a=1}^{n_a} N_a(\boldsymbol{\xi}) \mathbf{x}_a. \quad (2.9)$$

où n_a est le nombre de points de contrôle. Les fonctions de forme sont calculées par des triples produits de polynômes de Lagrange de degrés 1 ou 2 (Eq. 2.12).

La description d’un élément de volume dV de l’élément Ω_e est donnée en fonction du Jacobien de la transformation \mathbf{J}_e de sorte que :

$$dV = dx dy dz = \mathbf{J}_e d\xi d\eta d\gamma, \quad (2.10)$$

et où,

$$\mathbf{J}_e = \left| \frac{\partial(x, y, z)}{\partial(\xi, \eta, \gamma)} \right|, \quad (2.11)$$

qui est lui-même obtenu en calculant les dérivées partielles de la transformation (2.9).

Un raisonnement similaire est appliqué pour obtenir le passage ainsi que le Jacobien de la transformation de la surface des éléments frontières Γ_b vers le carré de référence. Dans cette thèse, tous les éléments sont construits sur des hexaèdres à 27 points d’ancrage et des quadrangles à 9 points.

2.2.3 Représentation polynomiale et intégration sur un élément

De manière similaire à la représentation de la forme des éléments, les méthodes aux éléments-finis utilisent des bases polynômiales pour représenter la solution au sein d’un élément. Les

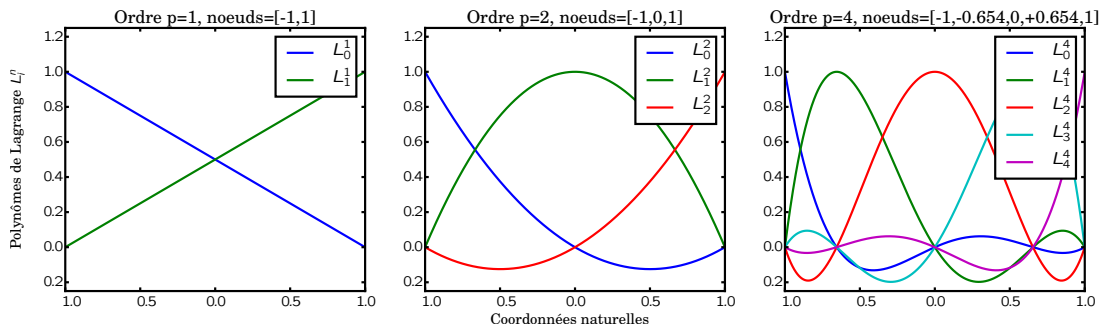


Figure 2.2 – Fonctions d’interpolation de Lagrange pour différents ordres polynomiaux ($p=1$, $p=2$, $p=4$). Ces fonctions sont définies à partir des points de colocations de la quadrature de Gauss-Lobatto-Legendre. Remarquons que les fonctions d’interpolation de Lagrange valent 0 ou 1 lorsqu’elles sont évaluées aux points de colocation.

éléments spectraux, en particulier, approximent ces solutions au moyen de polynômes de Lagrange de degrés élevés. Les polynômes d’interpolation de Lagrange i de degré n ainsi que leurs dérivées sont données par les relations suivantes :

$$L_i^n(x) = \prod_{k=0, k \neq i}^n \frac{x - x_k}{x_i - x_k}, \quad (2.12)$$

et

$$L_i^n(x)' = L_i^n(x) \sum_{k=0, k \neq i}^n \frac{1}{x - x_k}. \quad (2.13)$$

où x_k sont les points de colocation de la fonction à interpoler. Notons que les polynômes de Lagrange évalués aux points de colocation valent 0 ou 1 : $L_i^n(x_k) = \delta_{ik}$.

Plusieurs choix existent pour définir la position des points de colocation. Une approche naïve consiste à prendre des points équidistants. Cependant, pour des degrés polynomiaux élevés, cette approximation devient oscillante aux bords de l’intervalle d’interpolation, le phénomène de Runge. Une meilleure représentation polynomiale est obtenue en choisissant des points de colocation optimaux qui minimisent l’erreur d’approximation.

Dans la méthode aux éléments spectraux, le choix des points de colocation est intimement lié au choix de la quadrature utilisée pour évaluer l’intégration numérique sur un élément. Les quadratures de Gauss estiment l’intégrale de la fonction par

$$\int_{-1}^1 f(\xi) d\xi \approx \sum_{i=0}^N f(\xi_i) \omega_i \quad (2.14)$$

où ω_i est le poids de la quadrature aux points de colocation ξ irrégulièrement répartis sur l’intervalle $[-1, +1]$. Dans le cas particulier de la quadrature de Gauss, ces points sont strictement à l’intérieur du domaine $] -1, +1[$ et correspondent à des racines de polynômes caractéristiques de degré $N + 1$ (Tchebychev ou Legendre). Si la fonction à intégrer est un polynôme de degré $\leq 2N + 1$, cette quadrature donne un résultat exact. Pour des problèmes impliquant des conditions aux frontières, il est avantageux d’imposer des points de colocation aux extrémités

de l'intervalle $[-1, +1]$. Le type de quadrature est alors appelé quadrature de Gauss-Lobatto pour laquelle les points de colocation correspondent aux racines de

$$(1 - x^2)p'_N(x) = 0, \quad (2.15)$$

où p_N est un polynôme de Jacobi (Tchebychev, Legendre ou Legendre associés). Le résultat de ce type de quadrature est exact pour des polynômes de degré $\leq 2N - 1$. Dans le cas particulier de la quadrature de type Gauss-Lobatto-Legendre (GLL), les points de colocation sont déterminés à partir des dérivées des polynômes de Lagrange (2.12) et peuvent être déterminés numériquement (Canuto et al., 2006). Les poids de quadrature correspondant sont donnés par

$$\omega_i = \frac{2}{N(N+1)(L_N(\xi))^2}. \quad (2.16)$$

La figure 2.2 présente les $N + 1$ polynômes d'interpolation de Lagrange de degré 1, 2 et 4 pour des points de colocation de la quadrature de type Gauss-Lobatto-Legendre (GLL). Ce choix particulier de quadrature, en conjonction de l'utilisation des interpolateurs de Lagrange, permet non seulement de minimiser le phénomène de Runge (Davis, 1975) mais aussi, comme nous le verrons par la suite, de rendre la matrice de masse diagonale.

À trois dimensions, la fonction f peut donc être représentée sur l'élément Ω_e au moyen de la relation :

$$f(\mathbf{x}(\boldsymbol{\xi})) \approx \sum_{i,j,k=0}^n f^{ijk} L_i(\xi) L_j(\eta) L_k(\gamma) \quad \text{où} \quad f^{ijk} = f(\mathbf{x}(\xi_i, \eta_j, \gamma_k)) \quad (2.17)$$

De même, le gradient d'une fonction ∇f est représenté au moyen des dérivés des polynômes d'interpolation de Lagrange :

$$\begin{aligned} \nabla f(\mathbf{x}(\boldsymbol{\xi})) &= \sum_{\alpha}^3 \hat{\mathbf{x}}_{\alpha} \partial_{\alpha} f(\mathbf{x}(\boldsymbol{\xi})) \\ &\approx \sum_{\alpha}^3 \hat{\mathbf{x}}_{\alpha} \sum_{i,j,k=0}^n f^{ijk} [L'_i(\xi) L_j(\eta) L_k(\gamma) \partial_{\alpha} \xi + \\ &\quad L_i(\xi) L'_j(\eta) L_k(\gamma) \partial_{\alpha} \eta + L_i(\xi) L_j(\eta) L'_k(\gamma) \partial_{\alpha} \gamma]. \end{aligned} \quad (2.18)$$

Dans le cas particulier de l'évaluation du gradient d'une fonction, nous voyons qu'il est nécessaire d'accéder aux coefficients de l'inverse du jacobien $\frac{\partial \boldsymbol{\xi}}{\partial \mathbf{x}}$. Si le gradient est évalué aux points de colocation de type GLL, l'expression (2.18) se réduit à

$$\nabla f(\mathbf{x}(\boldsymbol{\xi})) \approx \sum_{\alpha}^3 \hat{\mathbf{x}}_{\alpha} \left[\sum_{i=0}^n f^{ij'k'} L'_i(\xi_{i'}) \partial_{\alpha} \xi + \sum_{j=0}^n f^{i'jk'} L'_j(\eta_{j'}) \partial_{\alpha} \eta + \sum_{k=0}^n f^{i'j'k} L'_k(\gamma_{k'}) \partial_{\alpha} \gamma \right]. \quad (2.19)$$

L'intégration numérique sur un élément de volume s'exprime au moyen de

$$\int_{\Omega_e} f(\mathbf{x}) d\mathbf{x}^3 = \int_{-1}^1 \int_{-1}^1 \int_{-1}^1 f(\mathbf{x}(\boldsymbol{\xi})) \mathbf{J}_e(\boldsymbol{\xi}) d\boldsymbol{\xi}^3 \approx \sum_{i,j,k=0}^n \omega_i \omega_j \omega_k f^{ijk} J_e^{ijk}. \quad (2.20)$$

qui fait intervenir le jacobien de la transformation.

2.2.4 Intégration numérique de la forme faible de l'équation d'onde

L'intégration numérique des termes présents dans la formulation faible des équations de l'élastodynamique (2.6) est réalisée en utilisant la discrétisation et les règles de quadrature vues précédemment. A cet effet, nous représentons les champs d'onde au moyen de l'expression (2.17) et choisissons les fonctions tests comme étant des polynômes de Lagrange.

Matrice de masse

La terme de gauche de l'équation (2.6) correspond à la matrice de masse. Son expression est obtenue en réalisant une triple intégration numérique sur l'élément de référence :

$$\int_{\Omega_e} \rho(\mathbf{x}) \mathbf{w}(\mathbf{x}) \cdot \partial_t^2 \mathbf{u}(\mathbf{x}) d^3\mathbf{x} = \int_{-1}^1 \int_{-1}^1 \int_{-1}^1 \rho(\mathbf{x}(\boldsymbol{\xi})) \mathbf{w}(\mathbf{x}(\boldsymbol{\xi})) \cdot \partial_t^2 \mathbf{u}(\mathbf{x}(\boldsymbol{\xi})) \mathbf{J}_e d^3\boldsymbol{\xi} = \sum_{l,m,n}^N \partial_t^2 u_{lmn} \sum_{i,j,k}^N \rho_{ijk} \omega_{ijk} \mathcal{J}_{ijk} \delta_{li} \delta_{mj} \delta_{nk} = \mathbf{M}\ddot{\mathbf{u}}. \quad (2.21)$$

où \mathbf{M} est la matrice de masse.

La présence des symboles de Kronecker dans l'expression (2.21) indique que, par construction, la matrice de masse est diagonale du fait de la colocation des points de quadrature et d'interpolations : ils correspondent tous deux aux points GLL. Le calcul des termes diagonaux s'effectue en sommant les produits entre la densité ρ , le Jacobien \mathbf{J}_e et les poids de la quadrature ω_{ijk} évalués à tous les degrés de liberté $(\xi_i, \eta_j, \gamma_k)$ de l'élément considéré.

Précisons que la diagonalité de la matrice de masse provient du fait que la quadrature de Gauss-Lobatto-Legendre est uniquement évaluée aux points nodaux. Cette astuce s'apparente aux techniques de condensation de masse (« *mass lumping* ») couramment utilisées dans les méthodes aux éléments finis (Zienkiewicz et al., 2005). Néanmoins, cela implique une légère dégradation de la précision puisque nous évaluons des polynômes de degré $2N$ et que l'intégration n'est exacte que pour des polynômes de degré inférieur ou égal à $2N - 1$. Cette im- précision peut toutefois être diminuée en considérant des degrés polynomiaux élevés. En outre, pour des schémas explicites d'intégration temporelle l'efficacité calculatoire fournie par cette

astuce compense largement la perte de précision. En effet, avoir une matrice de masse diagonale permet d'économiser énormément d'espace mémoire et rend simple, rapide et peu coûteuse son inversion.

Matrice de raideur

Pour réaliser l'intégration volumique de la matrice de raideur, le deuxième terme du membre de droite de l'équation (2.6), nous devons en premier lieu déterminer l'expression de $\nabla \mathbf{w}$ et de la contrainte $\boldsymbol{\sigma}$. Pour ce faire, les gradients des fonctions tests et des déplacements sont exprimés au moyen de l'équation (2.19). La contrainte est quant à elle obtenue en appliquant la loi de Hooke (1.5).

L'expression de l'intégrande $\nabla \mathbf{w} : \boldsymbol{\sigma}$ fait intervenir le double produit contracté « : » et s'écrit (pour un degré de liberté donné) :

$$\begin{aligned} \nabla \mathbf{w} : \boldsymbol{\sigma} &= \sum_{i,j=1}^3 \sigma_{ij} \partial_j w_i \approx \sum_{i,k=1}^3 \left(\sum_{j=1}^3 \sigma_{ij} \partial_j \boldsymbol{\xi}_k \right) \frac{\partial w_i}{\partial \boldsymbol{\xi}_k} \\ &\approx \sum_{i,k=1}^3 B_{ik} \frac{\partial w_i}{\partial \boldsymbol{\xi}_k}, \end{aligned} \quad (2.22)$$

où,

$$B_{ik} = \sum_{j=1}^3 \sigma_{ij} \partial_j \boldsymbol{\xi}_k. \quad (2.23)$$

Nous pouvons à présent intégrer numériquement en utilisant l'interpolation lagrangienne et la quadrature GLL :

$$\begin{aligned} \int_{\Omega_e} \nabla \mathbf{w} : \boldsymbol{\sigma} \, d\mathbf{x}^3 &\approx \iiint_{-1-1-1}^{1-1-1} \left(\sum_{i,k=1}^3 B_{ik} \frac{\partial w_i}{\partial \boldsymbol{\xi}_k} \right) \mathbf{J}_e \, d\boldsymbol{\xi}^3 \\ &\approx \sum_{i=1}^3 w_i^{\alpha\beta\gamma} [\omega_\beta \omega_\gamma \sum_{\alpha'=0}^N \omega_{\alpha'} \mathcal{J}^{\alpha'\beta\gamma} B_{i1}^{\alpha'\beta\gamma} L'_\alpha(\xi_{\alpha'}) \\ &\quad + \omega_\alpha \omega_\gamma \sum_{\beta'=0}^N \omega_{\beta'} \mathcal{J}^{\alpha\beta'\gamma} B_{i2}^{\alpha\beta'\gamma} L'_\beta(\eta_{\beta'}) \\ &\quad + \omega_\alpha \omega_\beta \sum_{\gamma'=0}^N \omega_{\gamma'} \mathcal{J}^{\alpha\beta\gamma'} B_{i3}^{\alpha\beta\gamma'} L'_\gamma(\gamma_{\gamma'})] \\ &= \mathbf{K} \mathbf{u}, \end{aligned} \quad (2.24)$$

où \mathbf{K} est la matrice de raideur. Contrairement à la matrice de masse, la matrice de raideur est pleine.

Force ponctuelle

L'intégration du terme source pour une force ponctuelle $\mathbf{f}(\mathbf{x}_s)$ sur l'élément Ω_s qui la contient est donnée directement par :

$$\int_{\Omega_s} \mathbf{w} \cdot \mathbf{f} \, dx^3 \approx \sum_{\alpha, \beta, \gamma=0}^N L_\alpha(\boldsymbol{\xi}(\mathbf{x}_s)) L_\beta(\boldsymbol{\xi}(\mathbf{x}_s)) L_\gamma(\boldsymbol{\xi}(\mathbf{x}_s)) f_i = \mathbf{f}, \quad (2.25)$$

où \mathbf{f} est le vecteur force. Dans le cas d'un tenseur de moment complet ou d'un champ de force, les expressions sont légèrement plus compliquées et sont données par (Komatitsch & Tromp, 1999a).

Conditions aux bords

Pour évaluer les conditions aux bords seule une intégrale de surface doit être calculée. Elle est donnée par :

$$\int_{\Gamma} \mathbf{w} \cdot (\hat{\mathbf{n}} \cdot \boldsymbol{\sigma}) \approx \sum_{\alpha, \beta=0}^N \omega_\alpha \omega_\beta \mathcal{J}^{\alpha\beta} t_i^{\alpha\beta} = \mathbf{C}\dot{\mathbf{u}}, \quad (2.26)$$

où $t_k^{\alpha\beta}$ est la traction sur le bord Γ qui vaut zéro dans le cas d'une surface libre ou la valeur donnée par la condition absorbante (2.8), et \mathbf{C} est la matrice d'amortissement.

2.2.5 Intégration temporelle – Le schéma de Newmark

Le schéma de Newmark (Hughes, 2003) est une méthode de résolution numérique d'équations différentielles aux différences-finies permettant de résoudre de manière itérative, on parle aussi d'algorithme d'évolution en temps, le problème

$$\mathbf{M}\ddot{\mathbf{u}}(t) + \mathbf{C}\dot{\mathbf{u}}(t) + \mathbf{K}\mathbf{u}(t) = \mathbf{f}(t), \quad (2.27)$$

où, \mathbf{M} est la matrice de masse, \mathbf{K} la matrice de raideur, \mathbf{C} la matrice d'amortissement et \mathbf{f} le terme source (force). Le principe de la méthode de Newmark est de discrétiser, au moyen d'un développement limité, les déplacements \mathbf{u} et vitesses particulières $\dot{\mathbf{u}}$ à l'instant $t + \Delta t$ à partir de ces mêmes grandeurs à l'instant t :

$$\begin{cases} \mathbf{u}(t + \Delta t) &= \mathbf{u}(t) + \Delta t \dot{\mathbf{u}}(t) + \frac{\Delta t^2}{2} \ddot{\mathbf{u}}(t) + \frac{\Delta t^3}{6} \dddot{\mathbf{u}}(t) + \dots \\ \dot{\mathbf{u}}(t + \Delta t) &= \dot{\mathbf{u}}(t) + \Delta t \ddot{\mathbf{u}}(t) + \frac{\Delta t^2}{2} \dddot{\mathbf{u}}(t) + \dots, \end{cases} \quad (2.28)$$

qui, après troncature et l'introduction de deux paramètres β et γ , deviennent :

$$\begin{cases} \mathbf{u}(t + \Delta t) &= \mathbf{u}(t) + \Delta t \dot{\mathbf{u}}(t) + \frac{\Delta t^2}{2} \ddot{\mathbf{u}}(t) + \beta \Delta t^3 \ddot{\mathbf{u}}(t) \\ \dot{\mathbf{u}}(t + \Delta t) &= \dot{\mathbf{u}}(t) + \Delta t \ddot{\mathbf{u}}(t) + \gamma \Delta t^2 \ddot{\mathbf{u}}(t). \end{cases} \quad (2.29)$$

En supposant que l'accélération est linéaire pour un pas de temps,

$$\ddot{\mathbf{u}}(t) = \frac{\dot{\mathbf{u}}(t + \Delta t) - \dot{\mathbf{u}}(t)}{\Delta t}, \quad (2.30)$$

Il vient :

$$\begin{cases} \mathbf{u}(t + \Delta t) &= \mathbf{u}(t) + \Delta t \dot{\mathbf{u}}(t) + \left(\frac{1}{2} - \beta\right) \Delta t^2 \ddot{\mathbf{u}}(t) + \beta \Delta t^2 \ddot{\mathbf{u}}(t + \Delta t) \\ \dot{\mathbf{u}}(t + \Delta t) &= \dot{\mathbf{u}}(t) + (1 - \gamma) \Delta t \ddot{\mathbf{u}}(t) + \gamma \Delta t \ddot{\mathbf{u}}(t + \Delta t) \end{cases} \quad (2.31)$$

Les constantes β et γ sont deux paramètres qui contrôlent la stabilité numérique du schéma ainsi que le taux d'amortissement numérique introduit dans le système par la méthode de Newmark. Leurs caractéristiques sont présentées dans les tableaux 2.1. Pour $\gamma = 0.5$ il n'y a pas d'amortissement tandis que pour $\gamma > 0.5$ le schéma devient dissipatif. Précisons, que le schéma d'intégration est d'ordre 2 si et seulement si $\gamma = 0.5$. Dans le cas où $\gamma = 0.5$ et $\beta = 0$ le schéma devient explicite et conditionnellement stable et est équivalent à un schéma aux différences-finies centrées. Ce type d'algorithme mène à une résolution itérative en temps :

$$\boxed{\begin{cases} \mathbf{u}(t + \Delta t) &= \mathbf{u}(t) + \Delta t \dot{\mathbf{u}}(t) + \frac{\Delta t^2}{2} \ddot{\mathbf{u}}(t) \\ \dot{\mathbf{u}}(t + \Delta t) &= \dot{\mathbf{u}}(t) + \frac{\Delta t}{2} \ddot{\mathbf{u}}(t) + \frac{\Delta t}{2} \ddot{\mathbf{u}}(t + \Delta t) \end{cases}} \quad (2.32)$$

La résolution numérique est effectuée sous forme d'un schéma prédictor-correcteur. Les déplacements et vitesses particulières sont ainsi prédites au temps $t + \Delta t$ en négligeant le terme d'accélération au temps $t + \Delta t$. Cette dernière est estimée par résolution de l'équation (2.27) et est utilisée pour corriger la prédiction des vitesses particulières. Le schéma se décompose ainsi en trois étapes distinctes :

Domaine		Stabilité	
$\gamma < 0.5$		instable	
$0.5 \leq \gamma$ et $2\beta \leq \gamma$		conditionnellement stable	
$0.5 \leq \gamma$ et $\gamma \leq 2\beta$		inconditionnellement stable	

Méthode	γ	β	Propriété
Différences centrées	0.5	0	explicite et conditionnellement stable
Fox Goodwin	0.5	1/12	conditionnellement stable
Accélération linéaire	0.5	1/6	conditionnellement stable
Accélération moyenne	0.5	1/4	inconditionnellement stable

Table 2.1 – Stabilités et propriétés associées à la méthode de Newmark en fonction des paramètres β et γ (Hughes, 2003).

Prediction

$$\begin{aligned} \mathbf{u}_{n+1} &= \mathbf{u}_n + \Delta t \mathbf{v}_n + (\Delta t/2) \mathbf{a}_n \\ \tilde{\mathbf{v}}_{n+1} &= \mathbf{v}_n + (\Delta t/2) \mathbf{a}_n + (\Delta t/2) \mathbf{a}_{n+1} \\ \mathbf{a}_{n+1} &= 0 \end{aligned} \quad (2.33)$$

Calcul de l'accélération

$$\mathbf{a}_{n+1} = \mathbf{M}^{-1}[\mathbf{f} - \mathbf{C}\mathbf{v}_{n+1} - \mathbf{K}\mathbf{u}_{n+1}] \quad (2.34)$$

Correction

$$\mathbf{v}_{n+1} = \tilde{\mathbf{v}}_{n+1} + (\Delta t/2) \mathbf{a}_{n+1} \quad (2.35)$$

La stabilité du schéma de Newmark est assurée si la condition de Courant-Friedrich-Levy (CFL) est respectée. Elle est donnée par la relation

$$\Delta t < \frac{\min(\Delta x)}{\max(V)} \quad (2.36)$$

où Δx est la plus petite distance entre deux points GLL, et V la vitesse maximale dans le milieu. Pour des polynômes de degré 4 et des éléments peu déformés, une règle approximative est de prendre $\Delta t < 0.17D/V_p^{\max}$ où D est la dimension caractéristique du plus petit élément présent dans le maillage. Les analyses de dispersion (Seriani & Oliveira, 2008; De Basabe & Sen, 2007) indiquent que cette taille doit être au minimum de l'ordre de la plus petite longueur d'onde propagée dans le milieu.

2.2.6 Implémentation

Nous avons implémenté un code de modélisation aux éléments spectraux pour des milieux strictement élastiques pouvant être arbitrairement hétérogènes et arbitrairement anisotropes. Le code est adapté aux plateformes de calcul hautes-performances par un double niveau de parallélisme MPI : une décomposition sur les sources et une décomposition du maillage en sous-domaines. La premier niveau de parallélisme permet de modéliser simultanément la propagation de champs d'ondes générés par différentes sources au lieu de les traiter séquentiellement³. Le second niveau permet quant à lui de répartir la charge de calcul liée au maillage global sur différents processeurs en le partitionnant en plusieurs sous-domaines. De cette manière, chaque processeur réalise les calculs sur les éléments contenus dans son sous-domaine particulier. Cette décomposition implique la communication de la solution entre les sous-domaines à chaque itération temporelle.

Bien qu'il soit possible et assez simple d'implémenter des kernels acoustiques, le code ne peut dans son état actuel modéliser la propagation d'ondes acoustiques. Pour des milieux présentant des interfaces fluides-solides, il serait nécessaire de mailler séparément les domaines

3. Ce parallélisme ne génère que de faibles communications entre les différentes tâches, pour un grand nombre de sources il est ainsi à privilégier.

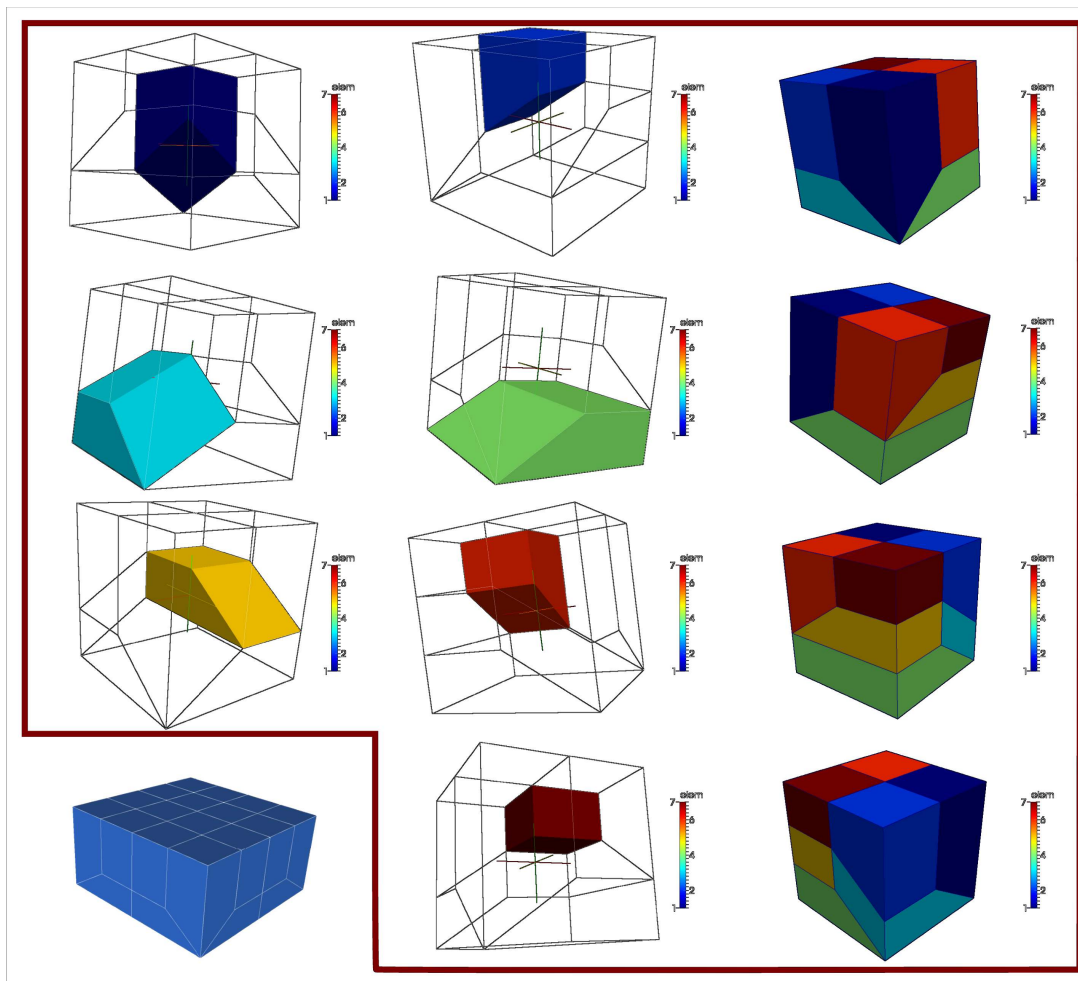


Figure 2.3 – Stratégie pour le raffinement du maillage du type « doubling ». Le brique *doubling* est composée de sept éléments et permet la transition entre des éléments de taille h à des éléments de taille $2h$. Présentation des sept éléments constituant la brique élémentaire. L'assemblage de quatre briques *doubling* est conforme si celles-ci sont tournées.

fluides et solides et d'imposer les conditions de continuité des champs d'onde aux abords de l'interface. Cela ne constitue pas un facteur limitant pour des applications exclusivement terrestres comme celles abordées dans ce travail.

Une étape essentielle pour appliquer les SEM consiste à définir le maillage dans lequel la propagation des ondes est modélisée. Notre implémentation discrétise le milieu lithosphérique par des hexaèdres à 27 points de contrôle pour lesquels les fonctions de formes sont calculées au moyen de produits de polynômes de Lagrange de degré 2 (voir section 2.2.2). Les éléments peuvent ainsi être déformés pour respecter la géométrie de surfaces courbes comme la sphéricité de la Terre ou des topographies lisses.

Pour mailler le domaine lithosphérique, nous procédons en quatre étapes :

- Nous commençons par définir un maillage structuré et régulier d'hexaèdres couvrant l'étendue du modèle lithosphérique.

- Si besoin, le maillage est raffiné à partir d’une profondeur donnée par l’utilisateur au moyen d’une super-brique « *doubling* » (Fig 2.3). Cette brique est elle-même constituée de sept sous-éléments qui permettent de réduire de moitié la taille des éléments tout en respectant la conformité du maillage.
- Si besoin, le maillage est déformé pour épouser la courbure de la Terre ou la forme d’une topographie complexe donnée par l’utilisateur⁴. Afin d’éviter que les éléments ne soient trop déformés en surface, cette dernière est parcimonieusement propagée sur toute la profondeur du maillage en suivant l’approche de Tarrass et al. (2011).
- Dans le cadre du couplage avec AxiSEM, les coordonnées locales et géographiques des bords du maillage sont déterminées.

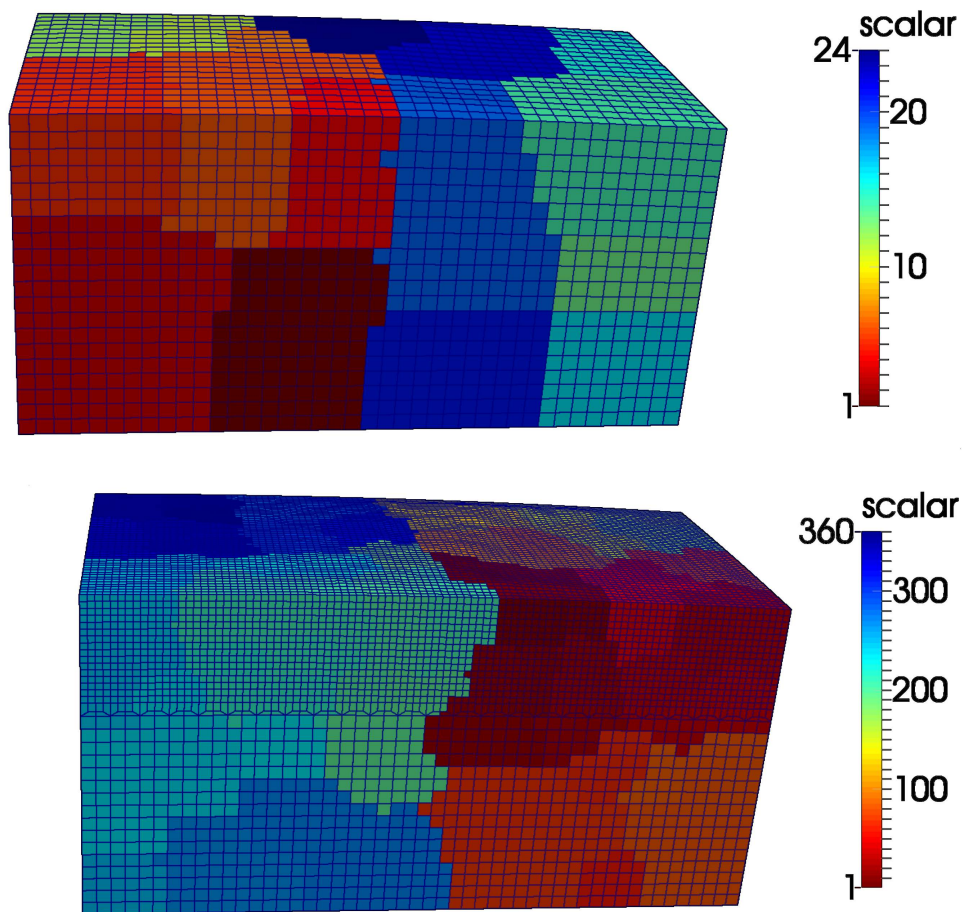


Figure 2.4 – Exemples de maillages hexaédrique et décomposition de domaine. En haut, maillage régulier de 16 000 éléments déformés pour épouser la courbure de la Terre. En bas, maillage de 108 750 éléments raffinés (d’un facteur 1/2) vers la surface et épousant une topographie complexe. L’échelle de couleur indique l’identification des sous-domaines après décomposition de domaine pour le calcul parallèle.

4. La topographie est préalablement filtrée pour respecter le théorème d’échantillonnage de Nyquist-Shannon.

Le maillage est ensuite décomposé en domaines par une librairie développée par L. Combe au sein du consortium SEISCOPE. Cette librairie de décomposition de domaine est elle-même basée sur la librairie de partitionnement de graphes METIS (Karypis & Kumar, 1999).

La figure 2.4 présente deux exemples de maillage et leur décomposition respective en 24 et 360 domaines. Le premier est conçu de manière régulière, avec des éléments de 10 km de côté, puis est déformé de sorte à rendre compte de la rotondité de la surface terrestre. Ce maillage présente peu d'éléments (16 000) et ne requiert pas d'être localement adapté en profondeur. Le second, en revanche, est maillé en raffinant la maille près de la surface par un facteur deux : la taille des éléments passe de 8 km à 4 km à 70 km de profondeur. Il est également déformé pour rendre compte d'une topographie complexe (dans ce cas précis, la topographie des Alpes). Le maillage résultant contient 108 750 éléments. S'il avait été maillé de façon purement régulière le nombre d'éléments serait de 250 000, augmentant très significativement le coût de calcul de la modélisation.

2.2.7 Exemples de modélisation et validations

Afin de valider l'implémentation de la méthode des éléments spectraux pour la modélisation des champs d'ondes 3D, nous considérons trois tests numériques canoniques dont nous connaissons les solutions analytiques et comparons ces solutions aux résultats fournis par la simulation numérique.

Pour chacun de ces tests, le milieu considéré est un espace ou demi-espace élastique isotrope homogène de densité $\rho = 3000 \text{ kg.m}^{-3}$, de vitesse P $V_P = 7000 \text{ m.s}^{-1}$ et de vitesse S $V_S = 4000 \text{ m.s}^{-1}$. Le domaine de calcul a pour dimensions $140\text{km} \times 140\text{km} \times 70\text{km}$. Il est discrétisé avec des éléments de 1400 m de côté, produisant un maillage de $100 \times 100 \times 50 = 500\,000$ éléments. La durée totale de la simulation est de 25s correspondant à 2500 pas de temps pour un échantillonnage à 0.01s.

Espace infini homogène

Le premier test de validation consiste à comparer la réponse d'un milieu homogène infini à une excitation de type explosion. L'explosion est simulée en considérant un tenseur de moment isotrope dont seuls les coefficients diagonaux sont différents de zéro et égaux entre eux. Pour une telle configuration, seule une onde P est propagée dans le volume. La figure 2.5 présente les composantes radiales et verticales des sismogrammes numériques (rouge) et analytiques (noir) enregistrés par des récepteurs situés à 1 km sous le toit (bord absorbant) du modèle. Sur la figure 2.5a, la source est localisée au centre du modèle et n'est pas concourante à la position d'un point de colocation et sa signature temporelle est un Ricker de fréquence dominante égale à 1 Hz. Les sismogrammes modélisés par SEM prédisent raisonnablement bien l'arrivée ainsi que l'amplitude de l'onde P mais sont également pollués par une onde S parasite. Il a été montré par Schubert (2003) que la génération de cette onde S est imputable à la représentation de la source au sein d'un élément. Cette dernière, lorsqu'elle ne coïncide pas avec un GLL peut introduire une légère anisotropie numérique. Cette conclusion est validée par la figure 2.5b pour laquelle nous avons déplacé la position de la source de sorte à ce qu'elle soit placée sur le nœud

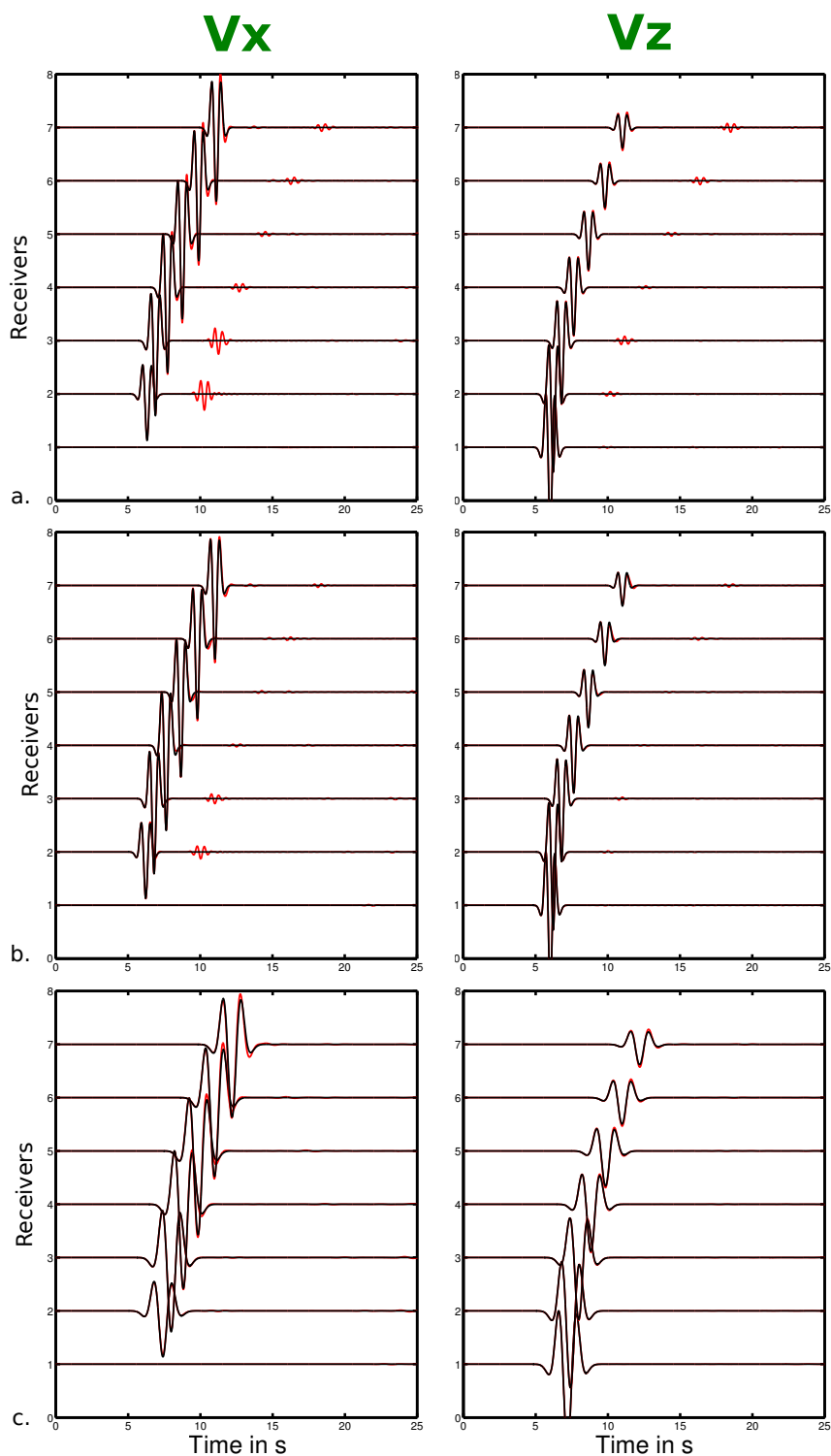


Figure 2.5 – Validation de la modélisation pour une source explosive. (a.) Résultats obtenus pour une source non-localisée sur un GLL dont la fréquence dominante est de 1 Hz. (b.) Similaire à (a.) pour une source localisé sur le GLL central de l’élément. (c.) Similaire à (b.) pour une fréquence dominante de 0.5 Hz. Les sismogrammes modélisés par SEM sont représentés en rouge tandis que les sismogrammes analytiques sont présentés en noir.

central d'un élément. Dans ce cas, l'onde S parasite est fortement atténuée mais néanmoins toujours visible sur les composantes radiales. Cette anisotropie peut-être encore diminuée en augmentant le nombre de points par longueur d'onde, c'est-à-dire en maillant plus finement, ou bien en diminuant la fréquence dominante de la source temporelle (Fig. 2.5c). Dans ce cas, la source est toujours un Ricker mais de fréquence dominante égale à 0.5 Hz et l'onde S disparaît complètement.

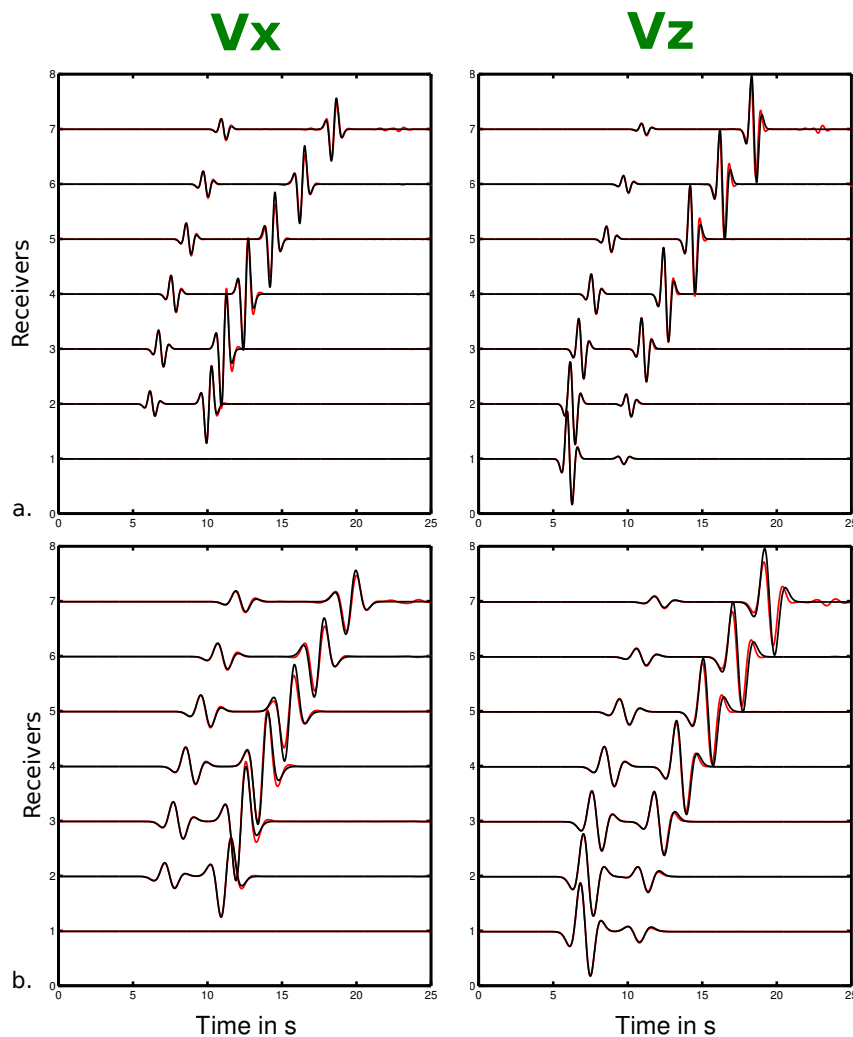


Figure 2.6 – Validation de la modélisation pour une source de type force. (a.) Résultats obtenus pour une source dont la fréquence dominante est de 1 Hz. (b.) Similaire à (a.) pour une fréquence dominante de 0.5 Hz. Les sismogrammes modélisés par SEM sont représentés en rouge tandis que les sismogrammes analytiques sont présentés en noir.

Le second test de validation est similaire au premier excepté que la source considérée est à présent une force ponctuelle verticale au lieu d'une explosion. Dans cette configuration, des ondes P et des ondes S sont toutes deux propagées dans le milieu élastique. L'ajustement entre les données modélisées et les données analytiques est montré sur la figure 2.6 pour deux Ricker de fréquences dominantes égales respectivement à 1 et 0.5 Hz. À nouveau, l'amplitude de l'onde P est fidèlement restituée tandis que l'onde S subit des variations d'amplitudes similaires à

celles observées pour la source ponctuelle explosive. Il s'agit toujours d'un effet d'anisotropie numérique engendrée par la localisation de la source au sein de l'élément. On remarque aussi une réflexion parasite sur les sismogrammes du récepteur numéro 7. Cette dernière provient de la réflexion du champ sur les bords absorbants du modèle.

Au vu de ces résultats numériques, nous concluons que notre implémentation des termes sources de type force ponctuelle et de type tenseur de moment sont validés dans le cas d'un milieu homogène infini.

Demi-espace homogène

Le test de validation suivant est celui du problème de Lamb (Lamb, 1904; Kausel, 2006, 2012). Le problème est de déterminer la réponse d'un demi-espace homogène pour une force ponctuelle verticale appliquée à la surface du demi-espace. Cette situation génère des ondes de volumes P et S ainsi qu'une onde de surface, l'onde de Rayleigh, qui dans ce cas précis est non dispersive. Ce test de validation est très important dans notre étude puisque cette configuration sera fréquemment rencontrée lors du processus d'inversion qui nécessite, comme nous le verrons, des modélisations pour des forces ponctuelles localisées aux récepteurs c'est-à-dire situés en surface.

Nous considérons donc un demi-espace homogène dont les propriétés physiques sont les mêmes que pour les deux cas tests précédents et appliquons une force verticale au centre de la surface du modèle. À nouveau, nous considérons deux fonctions source avec une fréquence dominante respective de 0.5 et 1 Hz. La comparaison des sismogrammes modélisés par SEM et des solutions analytiques correspondantes est présentée sur la figure 2.7.

Les résultats montrent que les ondes P, les ondes S et les ondes de Rayleigh sont fidèlement restituées, quelle que soit la fréquence considérée. Pour une fréquence dominante de 1 Hz, une légère oscillation apparaît juste après l'arrivée de l'onde de Rayleigh. Elle est néanmoins complètement atténuée lorsque la fréquence dominante de la source est diminuée, indiquant qu'il s'agit probablement d'un effet de dispersion lié à la discrétisation du problème. Comme pour la force ponctuelle, nous observons une forte réflexion parasite émanant des bords du maillage au niveau du récepteur numéro 7. Nous concluons que notre implémentation d'une force ponctuelle en présence d'une surface libre est correcte et est capable de rendre précisément compte des ondes de volume et des ondes de surface.

La figure 2.8 présente quelques instantanés de l'amplitude du champ d'onde généré par une force ponctuelle en surface. Nous voyons distinctement les fronts d'ondes sphériques des phases P, d'amplitude différentes, se séparer des ondes S du fait de leurs vitesses de propagation différentes. Il en résulte l'apparition d'un front d'onde plan caractérisant l'onde conique (ou onde réfractée). En surface, nous observons la propagation de l'onde de Rayleigh, de très forte amplitude, restreinte aux abords de la surface libre. À 15s, cette dernière commence à se « détacher » de l'onde S du fait de sa plus faible vitesse ($\approx 0.92V_S$).

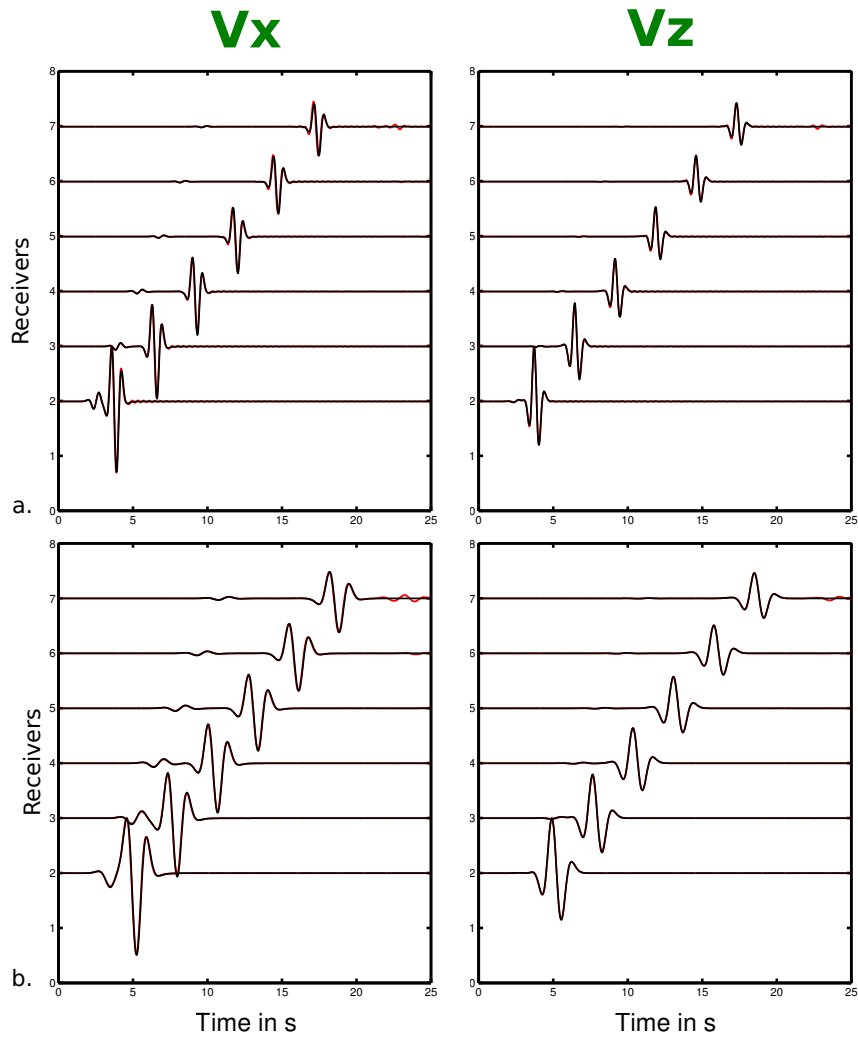


Figure 2.7 – Validation du problème de Lamb. La source correspond à une force verticale excitée à la surface. Les récepteurs sont eux aussi localisés sur la surface libre. (a.) Résultats obtenus pour une source dont la fréquence dominante est de 1 Hz. (b.) Similaire à (a.) pour une fréquence dominante de 0.5 Hz. Les sismogrammes modélisés par SEM sont représentés en rouge tandis que les sismogrammes analytiques sont présentés en noir.

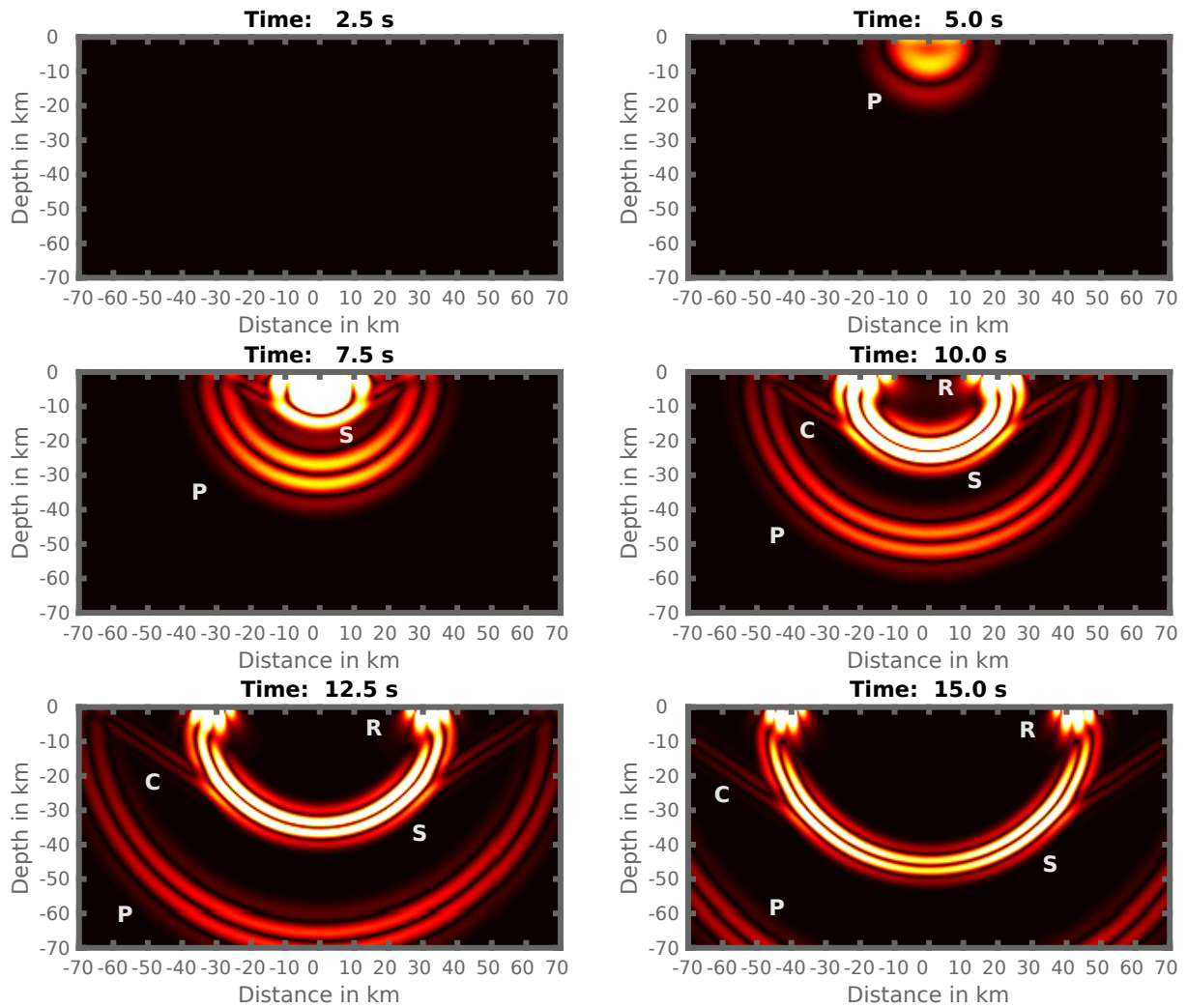


Figure 2.8 – Modélisation SEM pour une force ponctuelle à la surface libre. L'échelle de couleur représente l'amplitude des vitesses particulières. L'identification des différentes phases est superposée au champ d'onde. P : onde P, S : onde S, R : onde de Rayleigh et C : onde conique.

2.3 Stratégie d'injection de champ d'onde

Bien qu'il soit à l'heure actuelle possible de modéliser la propagation 3D des ondes sismiques dans la Terre globale (Komatitsch et al., 2002) à des périodes minimales proches de la seconde, il est encore inenvisageable de recourir à ce type de modélisation pour l'application de la FWI télésismique au vu du grand nombre de modélisations requises et des fréquences considérées. Notons toutefois que l'évolution actuelle des moyens de calculs a permis à Bozdağ et al. (2016) d'appliquer la tomographie adjointe – une variante de la FWI⁵ – à l'échelle globale pour des périodes minimales allant jusqu'à 17 secondes pour les ondes de volume et 45 secondes pour les ondes de surface. Il est en revanche possible, dans le cadre des applications lithosphériques, de recourir à des méthodes hybrides de modélisation qui permettent, au prix de quelques approximations, de calculer des champs d'ondes à relativement hautes fréquences et ce de manière répétée et à moindres coûts, rendant de fait envisageable l'application de la FWI. Ces méthodes hybrides consistent à injecter dans la cible lithosphérique, une solution de champ d'onde télésismique préalablement calculée dans la Terre globale. Ainsi, une seule modélisation globale est nécessaire quelle que soit le nombre de modélisations requis pour la FWI.

Les techniques d'injection de champs d'ondes ont été proposées en sismologie dès les années 60 par Alterman & Karal (1968) pour l'implémentation d'un point source dans les méthodes aux différences-finies. Elles reposent principalement sur la linéarité (superposition et continuité) du champ d'onde élastique qui peut ainsi être décomposé comme la somme d'un champ d'onde de référence et d'un champ d'onde diffracté par un sous-domaine contenant des hétérogénéités (Taflove & Hagness, 2000), d'où leur dénomination de technique « *Total-Field – Scattered-Field* » (champ total – champ diffracté). La popularité de ces techniques est grandissante aussi bien dans le domaine de l'exploration sismique (Robertsson & Chapman, 2000; Borisov, 2014) qu'en sismologie globale (Bielak et al., 2003; Roecker et al., 2010; Pageot et al., 2013; Monteiller et al., 2013; Tong et al., 2014a; Masson et al., 2014) pour simuler l'entrée d'un champ d'onde dans une cible particulière ou au contraire pour simuler la sortie du champ d'onde générée depuis la cible en question. Une alternative intéressante est le couplage complet (Capdeville et al., 2003) de champ d'onde (injection et extraction du champ d'onde dans une cible) particulièrement pour les études reliées à la CMB. Actuellement, ce type de problème est un champ actif de recherche et est investigué au moyen du théorème de représentation et des intégrales de Kirchhoff-Helmholtz (Masson & Romanowicz, 2016).

Dans nos applications, seule l'injection d'un champ d'onde dans la cible lithosphérique est nécessaire car les stations sont situés à la surface de cette cible. L'approche hybride ne requiert ainsi qu'un couplage partiel et nous considérons l'approche champ total – champ diffracté que nous présentons dans la section suivante.

5. Elle diffère de la FWI par la considération de paquets de phases particulières (Maggi et al., 2009). Son coût numérique reste toutefois identique à la FWI classique puisque le champ d'onde est aussi modélisé de manière complète.

2.3.1 L'approche champ total / champ diffracté

L'approche champ total – champ diffracté consiste à supposer, en vertu de la linéarité de l'équation d'onde, que le champ d'onde total \mathbf{u} dans un milieu hétérogène peut être décomposé comme la somme d'un champ d'onde incident \mathbf{u}^i calculé dans un milieu de référence et d'un champ d'onde diffracté \mathbf{u}^s par les hétérogénéités du milieu perturbé :

$$\mathbf{u} = \mathbf{u}^i + \mathbf{u}^s. \quad (2.37)$$

On peut par ailleurs décomposer le domaine de calcul en deux milieux distincts : le milieu interne Ω^I ne propageant que le champ total \mathbf{u} , et le milieu externe Ω^E qui lui ne propage que le champ diffracté \mathbf{u}^s (Fig. 2.9). Les deux domaines sont séparés par une frontière commune Γ sur laquelle le champ incident du milieu de référence \mathbf{u}^i est connu. Nous verrons dans la suite que nous exploitons la linéarité de l'équation d'onde par rapport au terme de source pour implémenter la méthode d'injection à la frontière entre les deux milieux. Cela présuppose néanmoins que ces deux milieux aient des propriétés identiques le long de la surface d'injection Γ . Il est ainsi important pour l'efficacité de la méthode que le modèle de référence et le modèle hétérogène soient identiques sur la surface d'injection, d'où l'addition d'une zone de transition lisse entre la frontière et le modèle lithosphérique (Fig. 2.9, la zone marron). Le champ diffracté sortant de la cible lithosphérique peut ensuite être atténué par l'ajout d'un domaine absorbant à l'extérieur de la zone champ diffracté.

La configuration du problème présenté s'applique indifféremment à n'importe quelle méthode numérique, différences-finies ou éléments-finis. Dans le cas particulier des SEM, les deux domaines de calcul Ω^I et Ω^E peuvent être considérés séparément dans la formulation faible des équation de l'élastodynamique (2.6). Pour le domaine interne, nous avons

$$\int_{\Omega^I} \rho \mathbf{w} \cdot \partial_t^2 \mathbf{u} \, d^3\mathbf{x} = \int_{\Gamma} \mathbf{w} \cdot \mathbf{T} \, d^2\mathbf{x} - \int_{\Omega^I} \nabla \mathbf{w} : \boldsymbol{\sigma} \, d^3\mathbf{x}, \quad (2.38)$$

et pour le domaine externe :

$$\int_{\Omega^E} \rho \mathbf{w} \cdot \partial_t^2 \mathbf{u}^s \, d^3\mathbf{x} = \int_{\Gamma} \mathbf{w} \cdot \mathbf{T}^s \, d^2\mathbf{x} - \int_{\Omega^E} \nabla \mathbf{w} : \boldsymbol{\sigma}^s \, d^3\mathbf{x}. \quad (2.39)$$

En vertu de la linéarité de l'équation d'onde nous réécrivons les tractions et contraintes comme $\mathbf{T} = \mathbf{T}^i + \mathbf{T}^s$ et $\boldsymbol{\sigma} = \boldsymbol{\sigma}^i + \boldsymbol{\sigma}^s$. Le champ total peut ainsi être calculé à partir du champ diffracté⁶ en imposant la traction incidente dans l'intégrale de surface de l'équation (2.38). Ceci fait, le champ diffracté peut être déterminé dans Ω^E en remplaçant \mathbf{T}^s par $\mathbf{T} - \mathbf{T}^i$ dans l'intégrale de surface de l'équation (2.39) (Zhinan et al., 2014).

Dans le cas particulier de la condition absorbante (2.8) le problème peut être simplifié en absorbant directement le champ diffracté sur les bords du domaine interne (Monteiller et al., 2013; Tong et al., 2014a) de sorte que :

$$\mathbf{T} - \mathbf{T}^i = \rho \left[V_p \left(\hat{\mathbf{n}} \cdot \partial_t (\mathbf{u} - \mathbf{u}^i) \right) \hat{\mathbf{n}} + V_s (\mathbf{I} - \hat{\mathbf{n}}\hat{\mathbf{n}}) \cdot \partial_t (\mathbf{u} - \mathbf{u}^i) \right]. \quad (2.40)$$

6. Nul à l'initialisation de la modélisation.

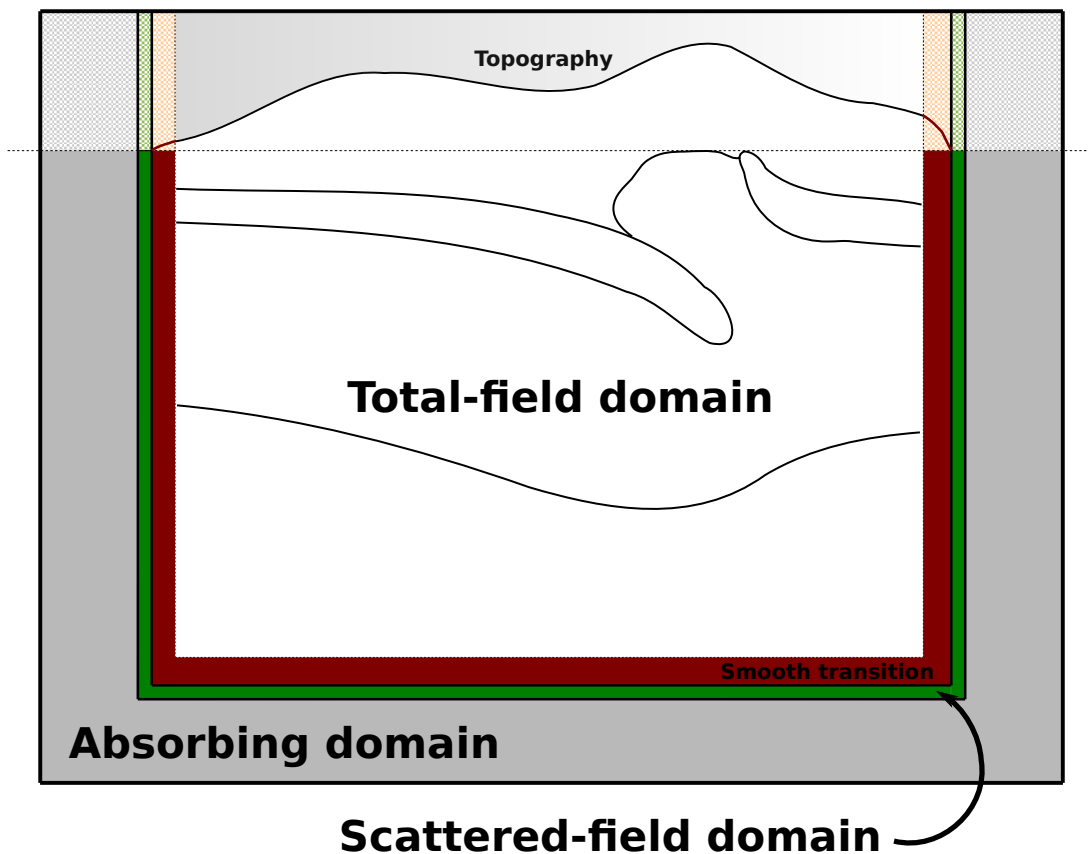


Figure 2.9 – Géométrie de la méthode d’injection champ total - champ diffracté. Le champ total est propagé dans le milieu interne (zone blanche) tandis que le champ diffracté est lui propagé dans la zone champ diffracté (en vert). Le domaine de calcul peut être complété par une couche absorbante (en gris) afin de forcer la propagation du champ diffracté vers l’extérieur de la cible, simulant ainsi un milieu infini. La zone marron assure une transition lisse entre le milieu perturbé et le milieu de référence.

Ce traitement aux frontières implémente naturellement l’injection du champ d’onde incident dans les SEM et nécessite simplement la connaissance des tractions et vitesses particulières incidentes sur le bord Γ . Notons que dans le cas où le milieu interne est peu perturbé par rapport au milieu de référence, le champ diffracté est assez faible et est efficacement absorbé par la condition de Stacey (Clayton & Engquist, 1980; Stacey, 1988).

2.3.2 Injection d’une onde plane

Nous présentons à présent un exemple d’injection d’une onde plane par la méthode champ total – champ diffracté. La solution analytique de l’onde plane dans un demi-espace homogène est calculée à partir des relations données dans la section 1.4.2. La cible considérée est un demi-espace homogène de vitesse P et S respectives de 7 km/s et de 4 km/s et de densité égale à 3 g.cm^{-3} . Le champ incident est une onde plane, de source temporelle gaussienne, de 10 km de longueur d’onde entrant dans la cible avec un angle d’incidence de 45° . Ce dernier est injecté

sur tous les bords de la cible à l'exception du bord supérieur où une condition de surface libre est implémentée.

La figure 2.10 présente quelques instantanés de la propagation de l'onde plane dans la cible. Nous observons la propagation de l'onde P incidente ainsi que la génération des multiples PP et PS à la surface libre. En outre, nous voyons que les processus d'injection et d'absorption n'engendrent aucune réflexion parasite significative attestant non seulement de l'efficacité de l'injection mais aussi de la précision de la modélisation des interactions à la surface libre par les SEM. Ces conclusions sont validées par la comparaison directe des sismogrammes analytiques en surface et ceux modélisés par les SEM (Fig. 2.11) qui se superposent parfaitement. En effet, dans cette configuration le modèle perturbé est identique au milieu de référence et donc $\mathbf{u} = \mathbf{u}^i$

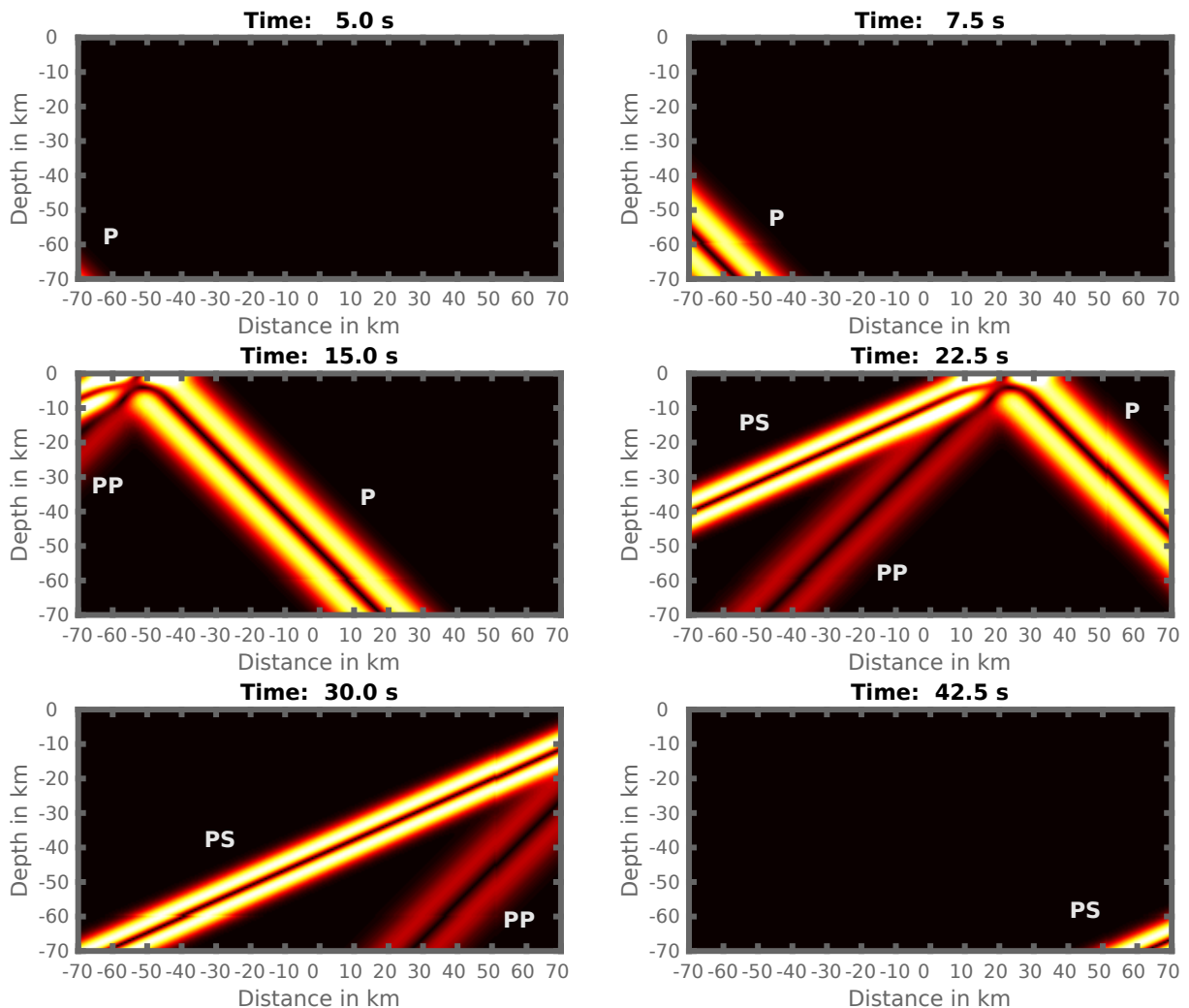


Figure 2.10 – Injection d'une onde plane analytique au sein de la cible lithosphérique. L'échelle de couleur représente l'amplitude de la vitesse particulière. L'identification des différentes phases est superposée au champ d'onde. Le champ d'ondes arrive par la gauche.

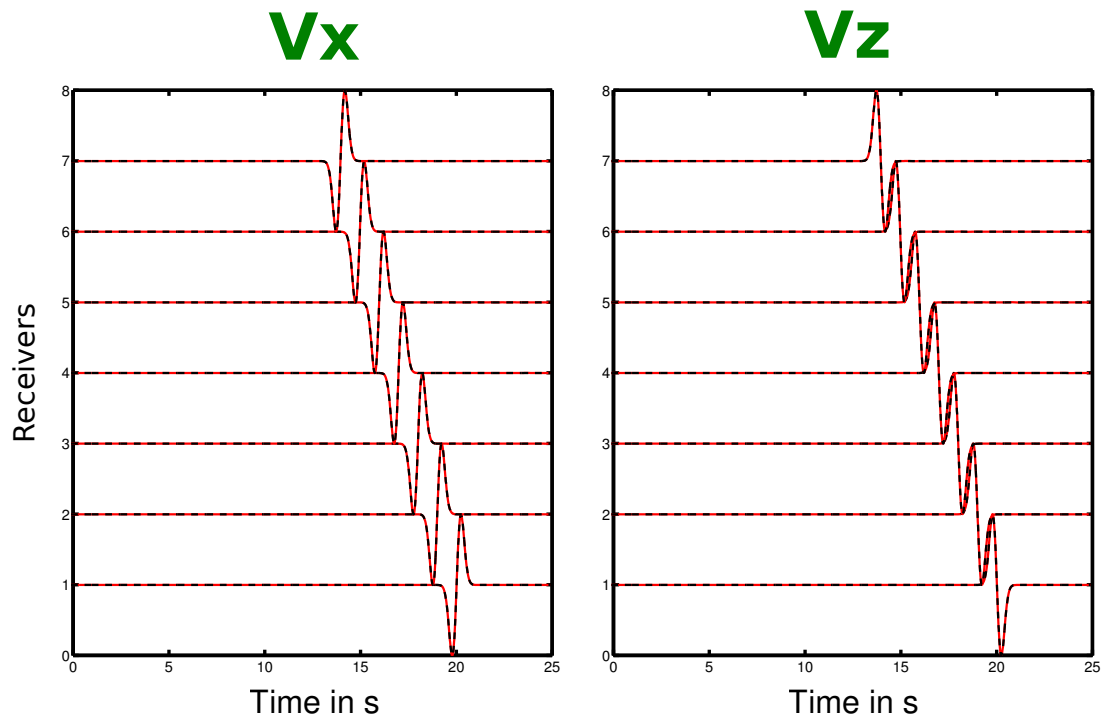


Figure 2.11 – Validation de l’injection d’une onde plane dans un demi-espace homogène. La longueur d’onde de l’onde P incidente est de 10 km, son angle d’incidence est de 45° , les récepteurs sont placés à la surface. Les sismogrammes modélisés par SEM sont représentés en rouge tandis que les sismogrammes analytiques sont présentés en tirets noirs.

2.4 Injection d’un champ d’onde télésismique avec AxiSEM

Nous avons vu que la méthode d’injection champ total – champ diffracté est efficace pour l’injection d’une onde plane dans une cible homogène. Cependant, dans les applications télésismiques le milieu de référence considéré est la Terre globale dont les structures sont essentiellement radiales et les champs d’ondes incidents plus complexes qu’une simple onde plane comme nous l’avons vu dans la section 1.6. Il est ainsi important d’injecter un champ d’onde incident qui rend compte au mieux des effets de propagation dans la Terre globale.

L’option qui s’offre à nous est alors d’injecter une solution de champ d’ondes calculée par une méthode de modélisation à l’échelle globale avec un coût numérique raisonnable. Plusieurs possibilités sont envisageables, citons notamment pour des modélisations hautes fréquences les méthodes F-K (Zhu & Rivera, 2002; Tong et al., 2014a,b), la méthode DSM (Takeuchi et al., 2000; Takeuchi & Geller, 2000; Monteiller et al., 2013, 2015) ou encore SpecFEM_Globe (Komatitsch & Tromp, 2002a,b; Capdeville et al., 2003).⁷

Pour cette thèse, nous avons fait le choix d’injecter les solutions de champ d’onde donnée par le code de modélisation AxiSEM que nous présentons dans la section suivante. La raison essentielle de ce choix est qu’actuellement seul AxiSEM peut à des coûts abordables modéliser précisément la propagation des ondes télésismiques à des fréquences supérieures au hertz.

7. Cette liste n’est pas exhaustive, nous pourrions ajouter GEMINI, Yspec, les modes normaux...

2.4.1 AxiSEM

AxiSEM (Nissen-Meyer et al., 2014) est un logiciel libre de modélisation aux éléments spectraux qui permet de modéliser la propagation de champs d'ondes élastiques 3D, à l'échelle globale, lesquels sont émis par des sources télésismiques réalistes. L'originalité ainsi que l'efficacité de cette méthode résident dans l'approximation axisymétrique de la structure terrestre, des sources sismologiques et du champ d'onde propagé. Sous ces hypothèses, il est possible de décomposer analytiquement (Nissen-Meyer et al., 2007a,b) la réponse d'un tenseur de moment complet M_{ij} comme une superposition linéaire de trois types de sources multipolaires elles-mêmes factorisées en fonction de leurs dépendances azimutales (Fig. 2.12). Néanmoins, pour que cette décomposition soit valide, il est impératif que la source soit localisée sur l'axe de symétrie c'est-à-dire au pôle.

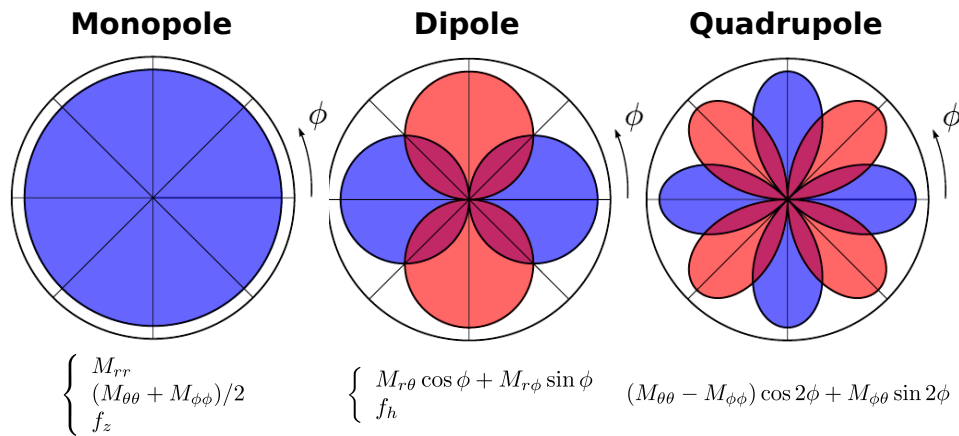


Figure 2.12 – Diagrammes de radiation pour les sources monopolaires, dipolaires et quadripolaires utilisées par AxiSEM pour reconstruire un champs d'ondes 3D à partir de champs d'ondes 2D. Pour un tenseur de moment complet, quatre sources sont nécessaires : deux sources monopole, une source dipôle et 1 source quadripôle. Adapté de Nissen-Meyer et al. (2014)

Les champs d'onde 3D créés par chacune de ces sources multipolaires $\mathbf{u}_m(\mathbf{x})$ s'écrivent ainsi comme :

$$\mathbf{u}_m(\mathbf{x}) = \begin{pmatrix} u_s(\tilde{\mathbf{x}}) \cos m\Phi \\ u_\Phi(\tilde{\mathbf{x}}) \sin m\Phi \\ u_z(\tilde{\mathbf{x}}) \cos m\Phi \end{pmatrix} \quad (2.41)$$

où $m = 0, 1, 2$ sont respectivement les types de radiation monopolaire, dipolaire et quadripolaire, $\mathbf{x} = (s, \Phi, z)$ sont les coordonnées cylindriques du milieu 3D et $\tilde{\mathbf{x}} = (s, z)$ les coordonnées 2D après « élimination » de la dimension azimutale Φ . L'équation (2.41) indique que le champ d'onde 3D axisymétrique associé à chaque composante multipolaire peut être reconstruit analytiquement à partir de modélisations 2D, effectuées sur des demi-disques, fournissant une diminution des temps calcul d'environ un ordre de grandeur comparativement à une simulation tridimensionnelle conventionnelle. Dans le cas d'un tenseur de moment complet, la reconstruc-

tion du champ d'onde total, nécessite quatre simulations indépendantes et est donnée par :

$$\mathbf{u}^p(\mathbf{x}, t) = \sum_{m=1}^4 \tilde{M}_m^p(\Phi, t) \tilde{\mathbf{G}}_m^p(\tilde{\mathbf{x}}, t), \quad (2.42)$$

où $\tilde{\mathbf{G}}_m^p$ est la fonction de Green 2D de chaque simulation m pour le point source p et pour laquelle \tilde{M}_m est donné par

$$\begin{aligned} \tilde{M}_1(\Phi) &= M_{rr} \\ \tilde{M}_2(\Phi) &= (M_{\theta\theta} + M_{\Phi\Phi})/2 \\ \tilde{M}_3(\Phi) &= M_{r\theta} \cos \Phi + M_{r\Phi} \sin \Phi \\ \tilde{M}_4(\Phi) &= (M_{\theta\theta} - M_{\Phi\Phi}) \cos 2\Phi + M_{\Phi\theta} \sin 2\Phi. \end{aligned} \quad (2.43)$$

Ainsi, contrairement aux codes de modélisation purement 3D, l'utilisation d'AxiSEM requiert de cruciales étapes de pré- et post-traitements :

Pré-traitements Une rotation des récepteurs et de la source est effectuée pour placer cette dernière au pôle. Du fait de la symétrie axiale, la dépendance en Φ est « éliminée » et l'on passe d'une géométrie 3D à une géométrie 2D.

Résolution numérique Quatre simulations 2D sont effectuées, une pour chaque terme de la décomposition de la source en multipôles.

Post-traitements Les champs d'ondes 3D sont construits par sommation des champs d'ondes 2D générés par chaque source après une pondération par le diagramme de radiation multipolaire. Une rotation est ensuite effectuée pour replacer les récepteurs et la source à leurs positions géographiques. Une autre rotation des composantes du champ d'onde est effectuée pour passer dans un repère choisi et les solutions sont convoluées avec la signature temporelle de la source.

2.4.2 Aspects pratiques du couplage

Tout comme AxiSEM nécessite des étapes de pré- et post-traitement pour générer des sismogrammes synthétiques, le couplage des méthodes SEM et AxiSEM nécessite des étapes intermédiaires de traitement. Ces étapes sont nécessaires pour notamment :

- indiquer à AxiSEM les positions géographiques exactes des points de la frontière du maillage SEM pour lesquels nous souhaitons obtenir les champs d'ondes 2D calculés par les quatre simulations AxiSEM,
- reconstruire la solution 3D à partir de ces quatre solutions 2D,
- effectuer la rotation pour replacer la source télesismique à sa position géographique,
- effectuer une seconde rotation pour exprimer le champ d'onde dans le maillage SEM,
- fenêtrer et interpoler le champ d'ondes respectivement à la durée et au pas d'intégration temporelle de la simulation SEM.

Ces étapes sont présentées ci-après.

Localisation du maillage

La localisation du maillage consiste à déterminer les coordonnées géographiques des bords du maillage SEM. Pour ce faire, nous définissons la position du centre du maillage à l'équateur sur le méridien de Greenwich et maillons de manière quasi-cartésienne avec les coordonnées locales (x, y, z) exprimées en kilomètres (voir la section 2.2.6). Nous transformons ensuite ces coordonnées locales en coordonnées géocentriques (X, Y, Z) (elles aussi exprimées en kilomètres) à partir de la relation

$$\begin{cases} X = z + R_T \\ Y = x \\ Z = y, \end{cases} \quad (2.44)$$

où R_T est le rayon terrestre. Nous passons des coordonnées géocentriques aux coordonnées géographiques (r, θ, Φ) au moyen des relations :

$$\begin{cases} r = \sqrt{X^2 + Y^2 + Z^2} \\ \lambda = \arcsin \frac{Z}{R} \\ \Phi = \arctan \frac{Y}{X}, \end{cases} \quad (2.45)$$

avec r le rayon, λ la latitude et Φ la longitude. Ces coordonnées sont ensuite fournies à AxiSEM qui les utilise pour fournir les quatre solutions 2D en ces points.

Post-traitement des champs d'onde incidents

Dès que les solutions sont obtenues, ces dernières sont recombinaées pour fournir les solutions des champs d'ondes 3D au moyen des relations (2.41–2.43). La rotation pour replacer la source du pôle à sa localisation géographique est donnée par la matrice

$$\mathcal{R}_s := \begin{pmatrix} X' \\ Y' \\ Z' \end{pmatrix} = \begin{pmatrix} \sin \lambda_s \cos \Phi_s & -\sin \Phi_s & \cos \lambda_s \cos \Phi_s \\ \sin \lambda_s \sin \Phi_s & \cos \Phi_s & \cos \lambda_s \sin \Phi_s \\ -\cos \lambda_s & 0 & \sin \lambda_s \end{pmatrix} \begin{pmatrix} X \\ Y \\ Z \end{pmatrix} \quad (2.46)$$

où λ_s et Φ_s sont la latitude et la longitude de la source considérée. La seconde rotation qui consiste à passer des coordonnées géocentriques aux coordonnées locales du maillage est donnée par :

$$\begin{pmatrix} X' \\ Y' \\ Z' \end{pmatrix} = \mathcal{R}_m^T(\lambda_m, \Phi_m, \alpha_m) \begin{pmatrix} X \\ Y \\ Z \end{pmatrix} \quad (2.47)$$

où la matrice de rotation \mathcal{R}_m ⁸ est donnée comme

$$\mathcal{R}_m = \begin{pmatrix} \cos \lambda \cos \Phi & -\sin \Phi \cos \alpha - \sin \lambda \cos \Phi \sin \alpha & \sin \Phi \sin \alpha - \sin \lambda \cos \Phi \cos \alpha \\ \cos \lambda \sin \Phi & \cos \Phi \cos \alpha - \sin \lambda \sin \Phi \sin \alpha & -\cos \Phi \sin \alpha - \sin \lambda \sin \Phi \cos \alpha \\ \sin \lambda & \cos \lambda \sin \alpha & \cos \lambda \cos \alpha \end{pmatrix} \quad (2.48)$$

où λ et Φ sont respectivement les latitude et longitude du centre du maillage et α l'obliquité du maillage (défini en sens direct depuis l'Est). Les rotations successives pour les champs de vitesses sont effectuées au moyen de :

$$\mathbf{u}' = \mathcal{R}\mathbf{u} \quad (2.49)$$

et, dans le cas du tenseur de contrainte, par

$$\mathbf{A}' = \mathcal{R}^T \mathbf{A} \mathcal{R}. \quad (2.50)$$

Interpolation – Problèmes de stockage

Malgré le gain substantiel en mémoire fourni par l'approche *Total-Field/Scattered-Field*, le stockage du champ incident à tous les noeuds situés aux frontières du modèle lithosphérique et à chaque pas de temps de la simulation reste assez prohibitif. Dans les configurations les moins favorables, ce stockage peut nécessiter jusqu'à 10 To d'espace disque par source. Néanmoins, en considérant que le pas de temps nécessaire à la stabilité d'une simulation AxiSEM est bien inférieur à celui requis par la condition de Nyquist, une réduction mémoire d'un à deux ordres de grandeur peut être obtenue en sous-échantillonnant les solutions d'AxiSEM aux frontières du modèle à la fréquence de Nyquist (Unser, 2000). Le champ incident est ainsi sous-échantillonné par AxiSEM pour être stocké à la fréquence de Nyquist. Ces solutions sont ensuite ré-interpolé en amont de la simulation SEM avec une fonction sinc) :

$$f(x) = \sum_{k \in \mathbb{Z}} f(kT) \text{sinc} \left(\frac{x}{T} - k \right) \quad (2.51)$$

où x est le pas de temps auquel on cherche à interpoler le signal, $f(x)$ le signal interpolé, T la période du signal à interpoler et k un indice de sommation. Notons que ce type d'interpolation reste assez coûteuse puisqu'elle utilise tous les échantillons des solutions sous-échantillonnées pour la reconstitution des solutions à chaque pas interpolé. Des alternatives moins coûteuse pourraient être considérées en utilisant des splines ou des fonctions sinc à largeur de bande limitées (Unser, 2000).

Pré-traitement des données

Tout comme les champs d'ondes d'AxiSEM ont besoin d'être tournés dans le repère local du maillage SEM, les données télésismiques observées, généralement enregistrées selon les

8. Attention 2.47 est défini comme la matrice de passage permettant d'aller du repère local au repère global.

directions géographiques (Z, N, E), doivent être tournées pour être dans le même repère que les champs d'ondes modélisés localement. Une première rotation est effectuée pour placer les récepteurs de leur position géographique vers la position définie par l'équateur et le méridien de Greenwich au moyen de :

$$\mathcal{R}_c := \begin{pmatrix} X \\ Y \\ Z \end{pmatrix} = \begin{pmatrix} \cos \lambda_r \cos \Phi_r & -\sin \lambda_r \cos \Phi_r & -\sin \Phi_r \\ \cos \lambda_r \sin \Phi_r & -\sin \lambda_r \sin \Phi_r & \cos \Phi_r \\ \sin \lambda_r & \cos \lambda_r & 0 \end{pmatrix} \begin{pmatrix} Z \\ N \\ E \end{pmatrix}, \quad (2.52)$$

où λ_r et Φ_r sont les latitudes et longitudes des récepteurs. Une dernière rotation est effectuée pour tourner les sismogrammes dans le repère du maillage local :

$$\begin{pmatrix} X' \\ Y' \\ Z' \end{pmatrix} = \mathcal{R}_m(\lambda_m, \Phi_m, \alpha_m) \begin{pmatrix} X \\ Y \\ Z \end{pmatrix}. \quad (2.53)$$

Puis les composantes locales sont obtenues par :

$$\begin{cases} x = Y' \\ y = Z' \\ z = R_T - X'. \end{cases} \quad (2.54)$$

2.4.3 Exemple et validation du couplage SEM/AxiSEM

Afin de valider l'approche hybride de modélisation par injection des solutions fournies par AxiSEM nous considérons un cas synthétique de modélisation dans la Terre globale et comparons les sismogrammes modélisés dans la cible lithosphérique avec SEM à ceux directement calculés dans la Terre globale par AxiSEM. Nous considérons pour cela une source téléseismique monopolaire $m_{rr} = 10^{24}$ N.m, localisée à 600 km de profondeur au point de latitude égale à 0° et de longitude égale de 6° . La source utilisée est une gaussienne de période dominante égale à 10 s. La cible lithosphérique fait une taille de 1000x1000x500 km et localisée à une latitude de 0° et de longitude 0° . Nous disposons sept récepteurs à l'équateur, régulièrement espacés par pas de 1° entre les longitudes -3° E et 3° E. Nous considérons une Terre à 3 couches homogènes : un manteau solide, un noyau externe fluide et un noyau interne solide. A part le manteau dont les vitesses et densité sont de $\rho = 3300$ kg.m $^{-3}$, $V_p = 8500$ m.s $^{-1}$ et $V_s = 4900$ m.s $^{-1}$, le reste du milieu est représenté par le modèle PREM.

Les figures 2.13 et 2.14 montrent respectivement des instantanées du champ d'onde 2D modélisé par AxiSEM et leurs équivalents propagés dans la cible lithosphérique. Les champs d'ondes présentent à l'échelle globale la propagation d'une onde P et d'une onde S émise à l'extérieur de la cible lithosphérique. Ces champ primaires interagissent avec la surface libre et donnent naissance à des réflexion PP et SS ainsi que des conversions PS et SP qui sont propagées au sein de la cible lithosphérique. Dans la cible lithosphérique, les champs d'ondes entrent et ressortent sans occasionner de réflexions parasites validant ainsi les différentes étapes du couplage.

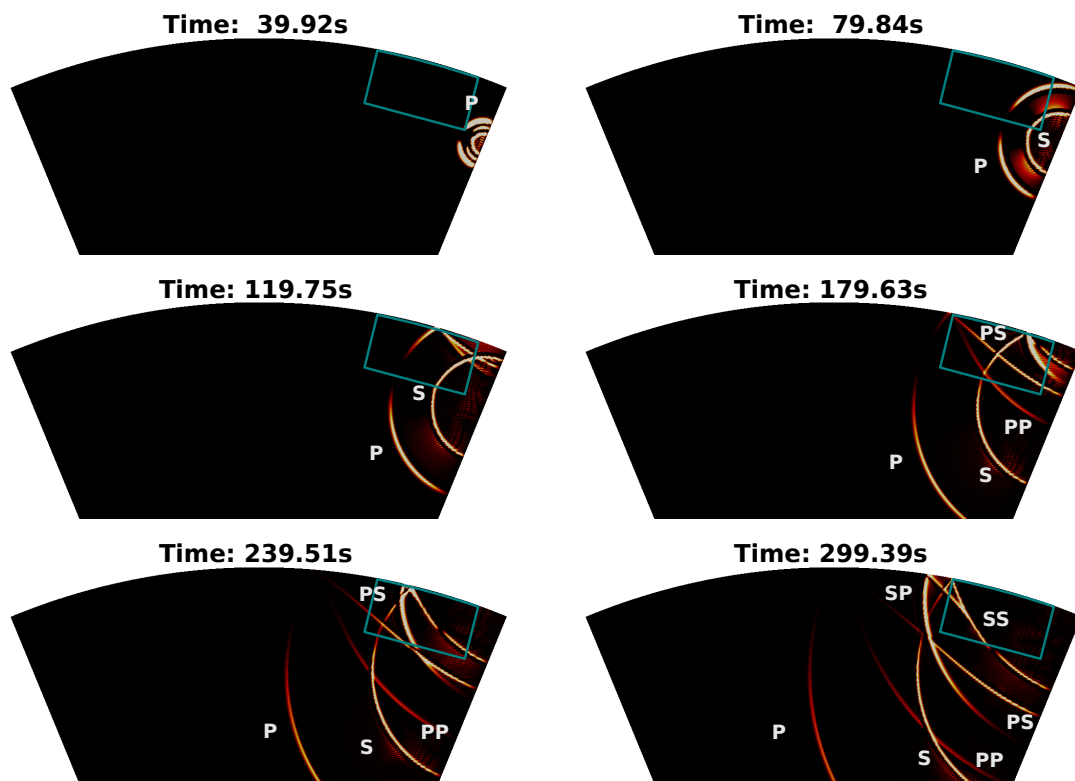


Figure 2.13 – Champ d’onde 2D modélisé par AxiSEM. L’échelle de couleur représente l’amplitude du déplacement particulière. L’identification des différentes phases est superposée au champs d’ondes. Le rectangle bleu est une représentation approximative de la cible lithosphérique.

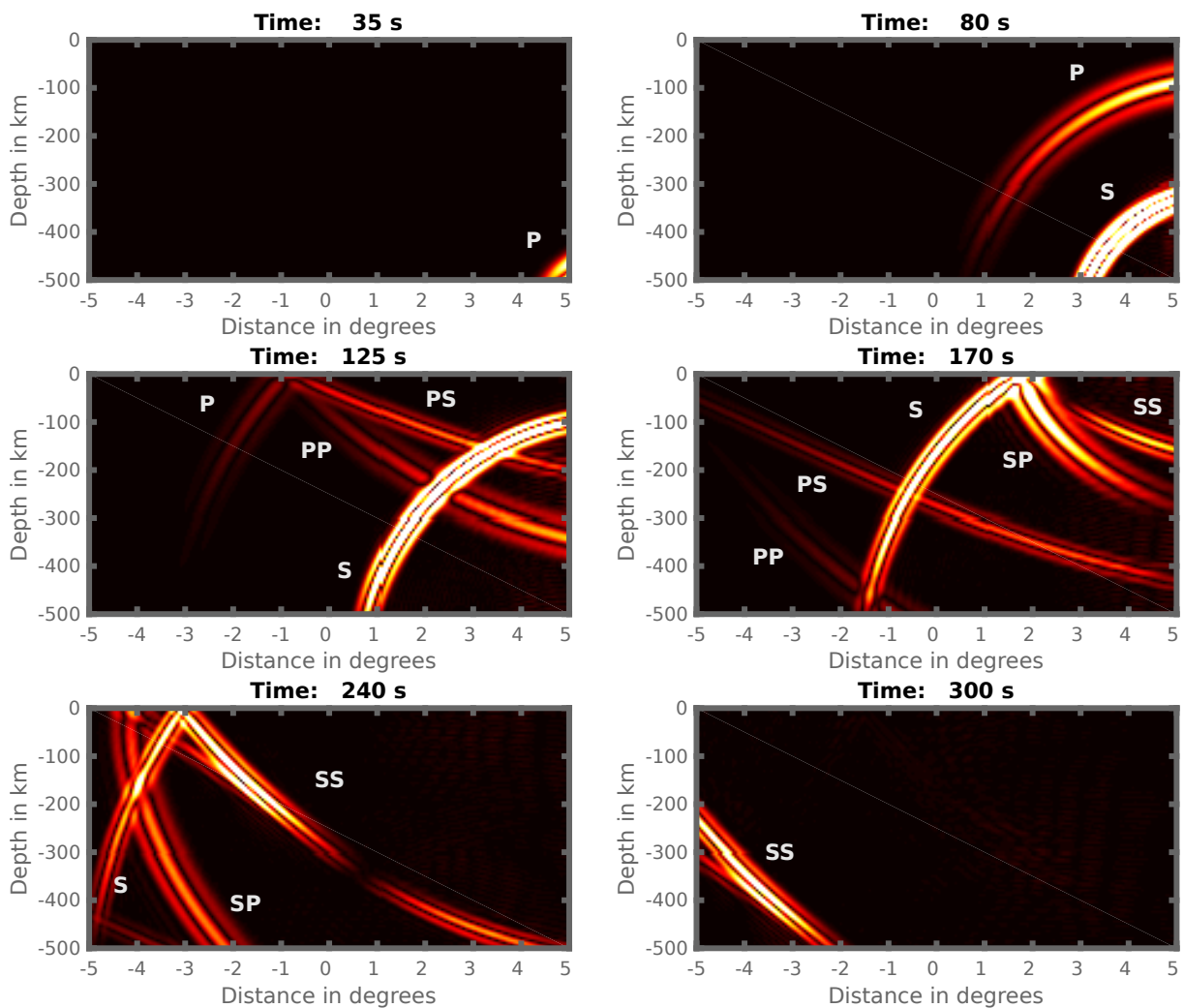


Figure 2.14 – Injection d’un champ d’onde généré par AxiSEM au sein de la cible lithosphérique. L’échelle de couleur représente l’amplitude de l’accélération particulière. L’identification des différentes phases est superposée au champ d’onde. Le champ d’onde arrive par la droite.

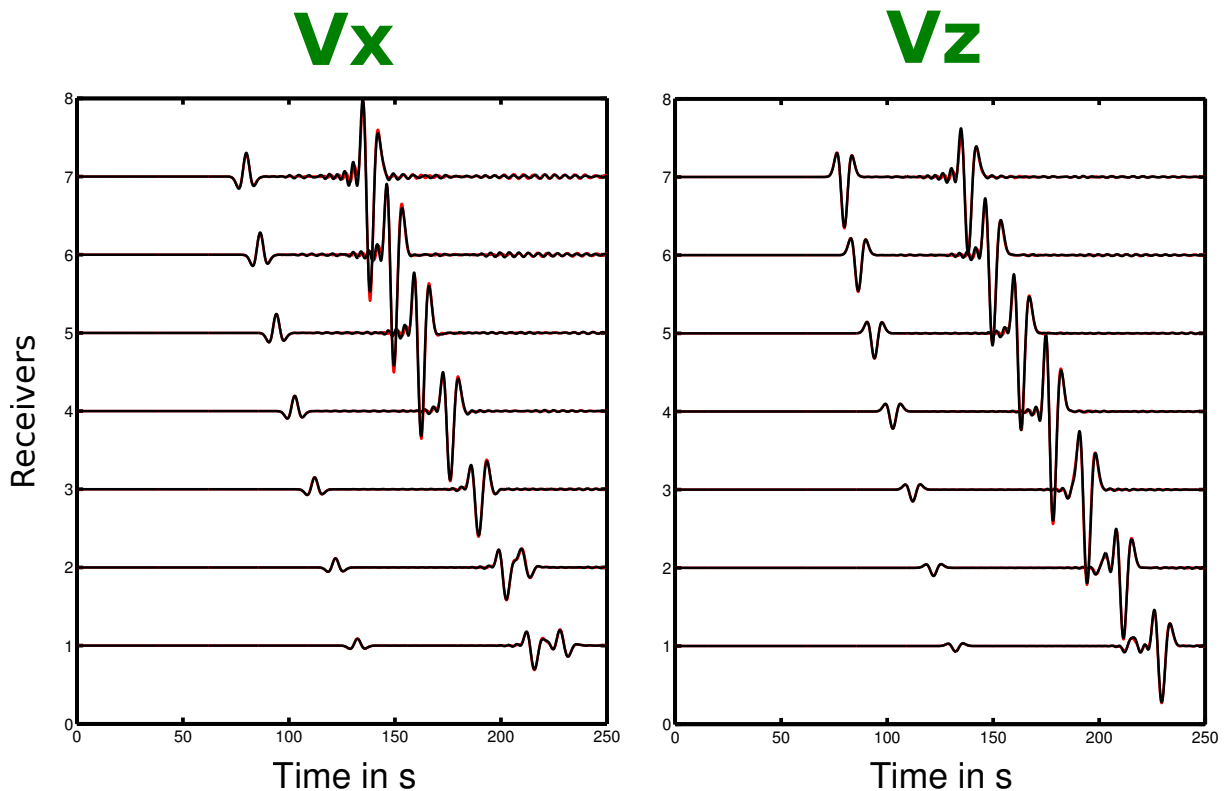


Figure 2.15 – Validation du couplage SEM-AxiSEM. La source possède une période dominante de 10 s et est localisée à une latitude de 0° et une longitude de 6° . La durée de la simulation est de 250s. Les récepteurs sont espacés de 1° le long de l'équateur à la surface. Les sismogrammes modélisés par SEM sont représentés en rouge tandis que les sismogrammes calculés par AxiSEM sont présentés en noir.

La figure 2.15 présente la comparaison des composantes radiales et verticales des sismogrammes modélisés par AxiSEM (rouges) et par SEM (noir). Globalement, nous observons que les formes d'onde sont en bon accord. La phase ainsi que l'amplitude de l'onde P sont très bien restituées. En revanche, bien que la forme d'onde de l'onde S soit en phase, son amplitude modélisée par SEM est légèrement différente de celle calculée par AxiSEM (particulièrement pour les récepteurs proche de la source).

Nous observons une légère oscillation sur les sismogrammes modélisés par SEM, principalement sur la composante horizontale, qui peut être imputée au processus de sous-échantillonnage temporel des champs d'ondes AxiSEM et de leur ré-échantillonnage au pas de temps de la simulation SEM par l'interpolation sinus-cardinal. Cette oscillation est plus importante pour les récepteurs proches de la source et s'atténue grandement à mesure que l'on s'en éloigne.

Conclusions

Dans ce chapitre, nous avons vu que différentes méthodes de modélisation sont envisageables pour résoudre le problème de propagation des ondes sismiques en milieu hétérogène. Le coût de calcul important de la modélisation du champ d'onde complet pour des milieux élastiques bornés par une surface libre nous a orienté vers l'implémentation de la méthode de modélisation aux éléments spectraux qui présente le meilleur rapport précision / efficacité. Cette implémentation a été validée au moyen de solutions analytiques pour trois exemples canoniques de propagation d'onde.

Nous avons vu que l'application de la FWI aux données télésismiques nécessite l'utilisation de méthodes hybrides d'injection de champ d'onde pour n'effectuer qu'une seule modélisation à l'échelle globale par événement. Nous avons adapté aux configurations télésismiques la méthode d'injection champ total - champ diffracté et validé son implémentation dans le cas d'une onde plane analytique. Pour prendre en compte les effets de propagation dans la Terre globale, nous avons de plus considéré le couplage d'AxiSEM, un code de modélisation à l'échelle globale basé sur des SEM 2D axisymétriques, lui-même validé par un test synthétique.

Nous disposons désormais des ingrédients nécessaires pour effectuer les modélisations requises lors de l'application de la FWI aux données télésismiques. Nous présentons dans la partie suivante les bases théoriques de la FWI ainsi qu'une analyse paramétrique de cette méthode en configuration télésismique.

Deuxième partie

Inversion de formes d'ondes télésismique

Chapitre 3

Le problème d'inversion des formes d'ondes

Sommaire

3.1	Notions fondamentales sur les problèmes inverses	110
3.1.1	Le problème inverse	110
3.1.2	Caractère mal posé des problèmes inverses	111
3.1.3	Résolution des problèmes inverses linéaire et non-linéaire	113
3.1.4	Méthodes de résolution locale	116
3.2	L'inversion de formes d'ondes, un problème aux moindres carrés linéarisé . .	122
3.2.1	La fonction coût	122
3.2.2	Linéarisation du problème	123
3.2.3	Les équations normales	124
3.2.4	Régularisation	127
3.2.5	Paramétrisation	128
3.3	Le gradient de la fonction coût	129
3.3.1	Construction par la méthode de l'état adjoint	129
3.3.2	Implémentation et validation	133
3.3.3	Anatomie d'un gradient télésismique	136
3.4	Algorithme de la FWI	141
3.4.1	Implémentation	141
3.4.2	Exemple d'inversion synthétique	141
3.4.3	Estimation de la source	143

Introduction

Dans ce chapitre, nous abordons le problème inverse d'inversion de formes d'ondes et son implémentation dans la configuration spécifique des acquisitions télésismiques. Avant cela, nous introduisons la théorie des problèmes inverses, et détaillons quelques méthodes locales de résolution des problèmes d'optimisation. Dans un second temps, nous formulons la FWI comme un problème inverse aux moindres carrés linéarisé que nous résolvons par une méthode d'optimisation locale. Nous détaillons le calcul du gradient de la fonctionnelle, que nous cherchons à minimiser, par la méthode de l'état adjoint et détaillons la structure de ce gradient dans des configurations télésismiques. Enfin, nous présentons l'implémentation et la validation du calcul du gradient ainsi que celui de l'algorithme d'inversion. Pour de plus amples informations sur les problèmes inverses, nous invitons le lecteur à regarder les ouvrages de [Menke \(2012\)](#), [Tarantola \(2005\)](#) et [Kern \(2002\)](#).

3.1 Notions fondamentales sur les problèmes inverses

3.1.1 Le problème inverse

Les problèmes inverses apparaissent lorsque l'on souhaite déterminer certaines propriétés (paramètres ou état) d'un système physique à partir d'une collection d'observables en lien direct ou indirect avec le système considéré. De telles situations sont naturellement rencontrées dans les sciences physiques, où l'acquisition intensive de données permet d'inférer des paramètres physiques caractéristiques de l'objet considéré.

Le problème inverse peut être défini simplement comme : *déterminer les causes, connaissant les effets*. Il s'oppose, en ce sens, aux problèmes dits directs qui visent à *déduire les effets, connaissant les causes* ([Kern, 2002](#)). Dans ces deux cas de figure, nous supposons qu'il existe un lien entre les causes (les paramètres) et les effets (les données). Ce lien ([Fig. 3.1](#)) est fourni par des modèles physiques \mathbf{G} , qui relient les données observées \mathbf{d}_{obs} aux paramètres \mathbf{m} du modèle physique considéré par le biais du problème direct :

$$\mathbf{d}_{obs} = \mathbf{G}(\mathbf{m}). \quad (3.1)$$

Dans le cas de la propagation des ondes élastiques, \mathbf{G} représente les équations de l'élastodynamique. Résoudre l'équation (3.1) permet ainsi de prédire les données observées \mathbf{d}_{obs} partir de la connaissance des propriétés élastiques du milieu \mathbf{m} . Cette résolution est un problème déterministe qui permet d'accéder à l'état du champ d'onde $\mathbf{u}(\mathbf{m})$ dans la totalité du domaine considéré. À contrario, le problème inverse consiste à résoudre l'équation (3.1) pour déterminer les paramètres élastiques \mathbf{m} à partir des données observées \mathbf{d}_{obs} . Pour ce faire, nous émettons l'hypothèse que la théorie physique de la propagation des ondes est connue et admettons que les champs d'ondes obéissent aux équations de l'élastodynamique.

En général, les champs d'ondes sismiques ne sont pas connus pour l'entièreté du système étudié. En sismologie par exemple, la mesure n'est effectuée qu'en un nombre restreint de sta-

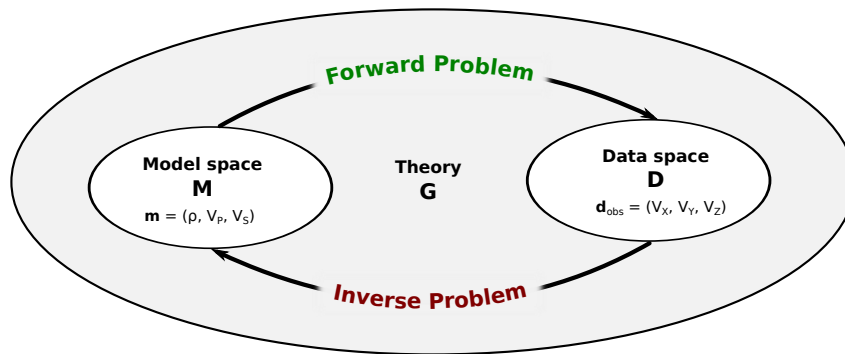


Figure 3.1 – Schéma représentatif des problèmes directs et inverses. Le problème direct consiste à prédire les données observées d_{obs} à partir des paramètres d'un modèle m au moyen de la théorie physique G . Le problème inverse consiste à déterminer un modèle de paramètres m à partir des données d sachant la théorie physique G .

tions elles-mêmes localisées en surface. Il est donc nécessaire d'introduire l'équation d'observation, implicite dans la relation (3.1), qui restreint l'état du champ d'onde $u(m)$ aux données observées aux récepteurs

$$d_{obs} = G(m) = Ru(m), \quad (3.2)$$

où R est un opérateur de restriction du champ d'onde qui extrait la solution aux capteurs¹.

Comme nous allons le voir, la résolution d'un problème inverse est intrinsèquement plus compliquée que la résolution d'un problème direct. Tout d'abord, alors que la plupart des problèmes directs sont bien posés, un problème inverse est en général mal posé et l'existence ou l'unicité de sa solution n'est pas assurée. En outre, même si le champ d'onde est linéaire – en vertu de la linéarité de l'équation d'onde – la relation G entre le modèle et les données est généralement non-linéaire. Ainsi, les problèmes inverses dépendent fortement de ces propriétés : bien posé – mal posé, linéaire – non-linéaire. Nous abordons à présent ces caractéristiques pour l'inversion de formes d'ondes, puis présentons les différentes classes de méthodes permettant de résoudre les problèmes inverses.

3.1.2 Caractère mal posé des problèmes inverses

Selon Hadamard (1902), pour qu'un problème mathématique soit *bien-posé*, il doit satisfaire les propriétés suivantes :

1. *Existence* : pour toutes les données, une solution au problème existe,
2. *Unicité* : pour toutes les données, cette solution est unique,
3. *Continuité* : la solution dépend de manière continue aux données d'entrée.

Ainsi, pour Hadamard seul un problème bien posé peut correctement modéliser un système physique. La majorité des problèmes directs sont ainsi bien posés, et il est possible de trouver des algorithmes stables de résolution (Kern, 2002). Au contraire, un problème est dit *mal-posé*

1. Un delta de Dirac par exemple.

s'il ne vérifie pas tout ou parti de ces propriétés, et il est dès lors difficile de résoudre le problème considéré. Les problèmes inverses, comme nous allons le voir, sont en général mal-posés. Il est donc nécessaire de les reformuler ou de les transformer pour d'une part, assurer l'existence de la solution, et d'autre part assurer ou du moins discuter l'unicité de la solution. Ci-après, nous analysons l'origine et les effets de ces trois propriétés dans la cadre de l'inversion de formes d'onde.

La non-existence d'une solution stricte² au problème inverse provient principalement des hypothèses qui sont posées sur la théorie physique et les données observées. En FWI, nous cherchons à déterminer les propriétés du milieu de propagation en supposant certaine la physique du milieu de propagation³, par exemple, un milieu purement élastique et isotrope. Ce choix est en général biaisé puisque nous ne connaissons pas la « vraie » physique qui caractérise le milieu de propagation. Ce dernier pouvant parfaitement être anisotrope et/ou atténuant. Les données observées dépendent ainsi de paramètres qui ne sont pas pris en compte par le modèle physique utilisé et elles ne peuvent être expliquées de manière exacte. Notons par ailleurs que, même si la physique considérée est exacte, les données sont entachées par la présence de bruit. Il est ainsi impossible de trouver une solution stricte au problème posé.

La non-existence n'est pourtant pas si problématique, nous pouvons toujours relaxer la notion de solution en cherchant une solution approchée au problème. Soit en cherchant la solution optimale : celle qui explique *au mieux* les données, (Nocedal & Wright, 2006a; Sen & Stoffa, 1995). Soit en cherchant la solution *la plus probable*, au moyen de l'inférence bayésienne (Tarantola, 1987; Sambridge & Mosegaard, 2002). Le problème est alors ramené à un problème d'optimisation visant à minimiser une fonction coût qui caractérise l'erreur, au sens d'une certaine norme, entre les données observées et les données calculées

$$\arg \min_{\mathbf{m}} \|\mathbf{d}_{obs} - \mathbf{R}\mathbf{u}(\mathbf{m})\|. \quad (3.3)$$

Précisons qu'il peut paraître intéressant de considérer une théorie physique plus complète pour améliorer le caractère mal-posé du problème. Ce choix n'est pourtant pas évident. Tout d'abord, cela peut engendrer un surcoût de calcul pour la résolution des problèmes directs. Enfin, elle implique de considérer un plus grand nombre de paramètres, et accentue de fait la non-unicité du problème.

Les problèmes linéaires sur-déterminés sont un bon exemple de problèmes pour lesquels il n'existe pas de solution stricte.

La non-unicité est une propriété plus délicate à analyser que la précédente. En effet, puisque le problème possède différentes solutions, il est nécessaire de définir des critères pour isoler l'une d'entre elles. La non-unicité provient principalement du caractère *sous-déterminé* des problèmes inverses, c'est-à-dire, du fait que les données accessibles – par les conditions expérimentales – ne sont pas en nombre suffisant pour contraindre tous les paramètres du modèle de manière unique. De ce fait il est possible de trouver une infinité de jeux de paramètres capables de produire les mêmes observations et donc d'expliquer les données de manière équivalente.

2. Qui explique parfaitement les données.

3. Notons que les méthodes probabilistes permettent de prendre en compte l'incertitude liée à la théorie physique du problème.

La non-unicité peut témoigner de différents effets. Elle peut tout d'abord traduire un manque de sensibilité des données aux paramètres situés dans des zones non ou mal contraintes par les données du fait des dispositifs expérimentaux. Un exemple typique est la tomographie de temps de trajet où la sensibilité des données dépend essentiellement de la couverture spatiale du milieu par les rai reliant les sources aux capteurs. Ainsi, lorsqu'un paramètre n'est traversé par aucun rai, aucune contrainte ne peut être apportée à son estimation. Autrement dit, quelque soit la valeur donnée à ce paramètre les données modélisées restent les mêmes. Par ailleurs, la non-unicité peut aussi traduire des couplages entre différents paramètres du modèle. Par exemple, si deux paramètres sont traversés par un seul et unique rai, le temps de trajet mesuré peut être expliqué de manière équivalente par l'un ou l'autre des paramètres ou par une combinaison linéaire des deux paramètres. Les effets conjoints du couplage des paramètres et de la sensibilité des données à ces paramètres sont critiques en FWI, particulièrement lorsqu'on considère des problèmes impliquant plusieurs classes de paramètres comme l'estimation conjointe de la densité, des vitesse de propagation d'ondes P et d'ondes S.

Il existe tout de même quelques moyens pour réduire la non-unicité des problème inverses. Une première option est d'utiliser de l'information *a priori* additionnelle (Tarantola & Valette, 1982) pour améliorer le conditionnement du problème. Une seconde option est d'imposer la régularité de la solution en régularisant le problème (Tikhonov, 1963; Tikhonov & Arsenin, 1977). Il est important de noter que, quelle que soit la méthode adoptée, l'information introduite biaise l'estimation du modèle obtenue. Il n'est pas évident d'estimer le biais qui a été introduit, néanmoins des analyses de résolution peuvent aider à apprécier la qualité de la solution (Scales & Snieder, 2000).

Le manque de continuité est essentiellement dû à la présence de bruit dans les données. Même si l'on peut théoriquement trouver une solution pour des données exactes, on ne peut pas résoudre le problème pour des données inexacts car la solution ne peut qu'approcher les données observées (même si le niveau de bruit est relativement faible). Cette difficulté est dramatique pour la résolution numérique des problèmes inverses puisqu'elle peut conduire à des algorithmes instables. Dans ce cas, d'infimes variations sur les données observées peuvent engendrer d'importantes variations des paramètres estimés.

3.1.3 Résolution des problèmes inverses linéaire et non-linéaire

Mathématiquement, un problème inverse est dit linéaire lorsque l'opérateur G est linéaire, c'est-à-dire, lorsque les données dépendent linéairement des paramètres du modèle. L'estimation de la signature temporelle d'une source sismique à partir de la connaissance des fonctions de Green est un exemple typique de problème inverse linéaire. Les problèmes linéaires sont assez bien compris, il est en effet possible d'obtenir des résultats précis et des algorithmes stables en ayant recours aux outils de l'algèbre linéaire et de l'analyse fonctionnelle, en particulier, la décomposition en valeur singulières de l'opérateur considéré (Kern, 2002; Menke, 2012).

À contrario, le problème inverse est dit non-linéaire lorsque les données sont non linéairement dépendantes des paramètres du modèle. Le plus souvent, ce sont des problèmes d'estimation de paramètres d'équations différentielles ou aux dérivées partielles. La FWI est un exemple

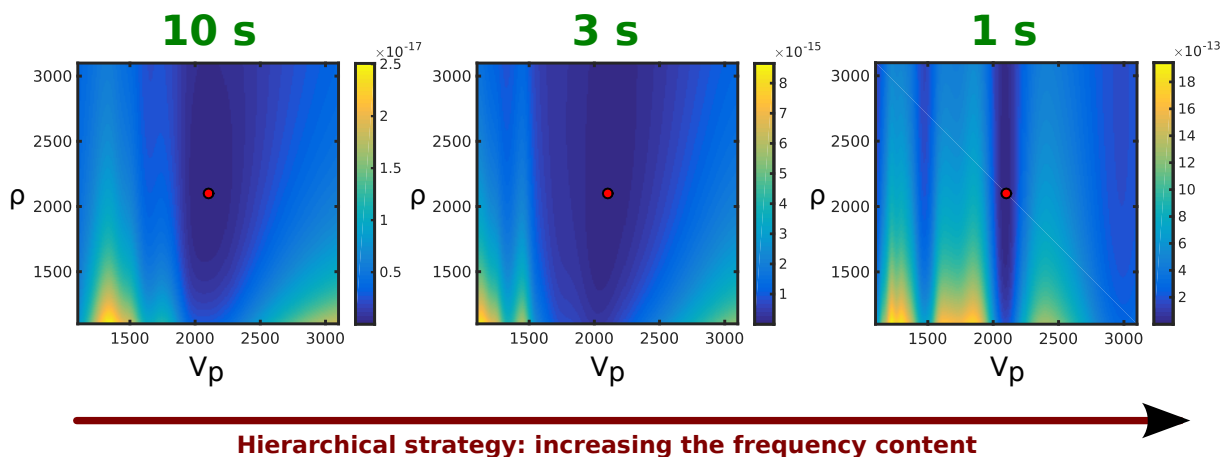


Figure 3.2 – Approche hiérarchique en fréquence. De gauche à droite, la fonction coût d'un problème FWI à une dimension pour des données filtrées à 10, 3 et 1 s de période minimale. Le modèle recherché (matérialisé par la pastille rouge) est un milieu homogène paramétré par une vitesse de 2100 m.s^{-1} et une densité de 2100 kg.m^{-3} . L'augmentation du contenu fréquentiel des données augmente la non-linéarité de la fonction coût qui montre de plus en plus de minima locaux dont les bassins d'attraction sont de plus en plus étroits. Notons au passage la forme particulière des minima locaux qui traduisent une sensibilité différente de la fonction coût selon le type de paramètre.

typique de problème non-linéaire puisque les sismogrammes enregistrés ne dépendent pas linéairement des paramètres élastiques du milieu. Cette non-linéarité s'ajoute à la non-unicité du problème et se manifeste par la présence de multiples minima locaux dans la fonction coût du problème d'optimisation. Des exemples de fonctions coût présentant plusieurs minima sont donnés sur les figures 3.2 et 3.3.

Les problèmes non-linéaires sont ainsi plus difficiles à résoudre que les problèmes linéaires. Ils sont généralement résolus au moyen d'algorithmes d'optimisation globaux qui permettent une exploration complète ou partielle de l'espace des modèles (Sen & Stoffa, 1995). L'exploration est réalisée en effectuant une recherche directe sur grille ou en employant des méthodes stochastiques du type Monte-Carlo (Mosegaard & Tarantola, 1995) qui tirent des modèles de paramètres de façon aléatoire. Ces méthodes requièrent un nombre important de modélisation pour évaluer la fonction coût. Ce surcoût peut être légèrement amorti en optimisant le tirage pour accélérer la convergence en utilisant des algorithmes de types recuit simulé (Metropolis & Ulam, 1949; Kirkpatrick et al., 1983), génétiques (Holland, 1975; Sen & Stoffa, 1991) ou de voisinage (Sambridge, 1999a,b). Bien que ces approches permettent d'obtenir le modèle optimal (minimum global), elle ne permettent pas de conclure sur l'unicité de la solution ni sur la confiance de la solution. Les approches bayésienne (Tarantola & Valette, 1982; Sambridge & Mosegaard, 2002; Tarantola, 2005) ont ainsi été développées pour décrire l'ensemble de l'espace des solutions. Dans ce type d'approche, on ne cherche plus le modèle optimal mais la densité de probabilité de la solution dont la confiance est caractérisée au moyen d'outils statistiques.

Bien que très attrayantes, les méthodes globales restent actuellement trop onéreuses pour

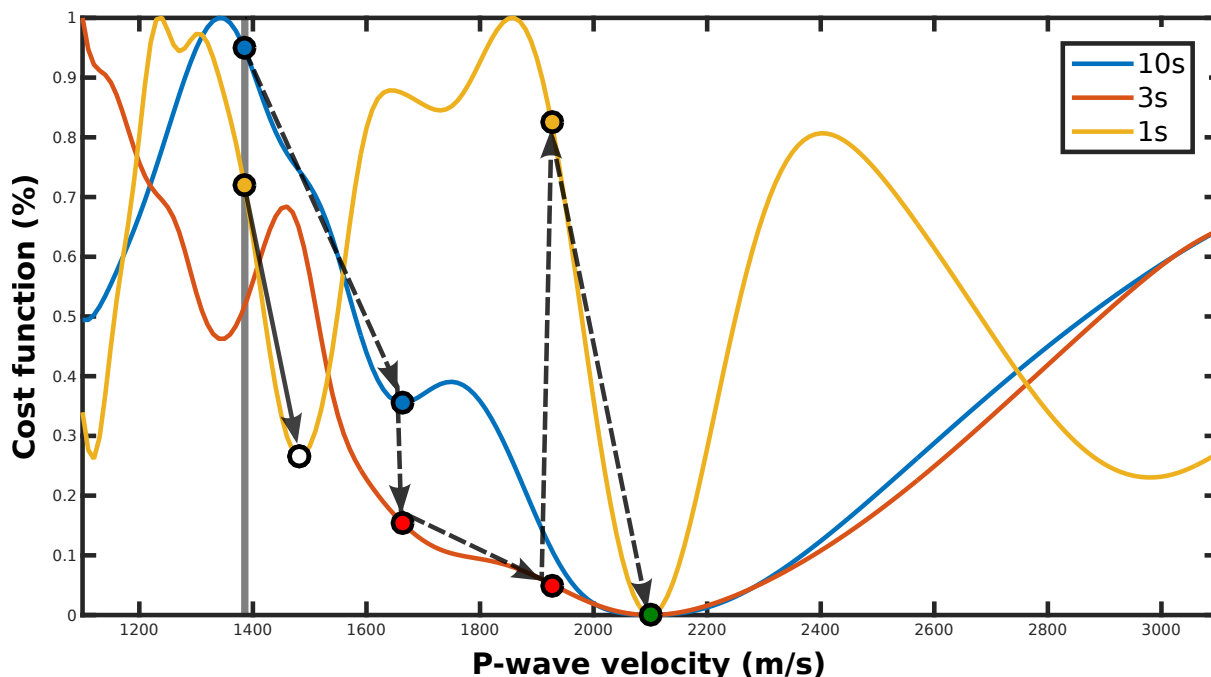


Figure 3.3 – Non-linéarité, minima locaux, et approche hiérarchique en fréquence. Les courbes présentent les fonction coût de la figure 3.2, dans trois bandes de fréquences et pour un paramètres de densité fixé. À 1 s, la fonction coût présente de multiples minima locaux et l’optimisation converge vers un minimum local (pastille blanche). À plus basses fréquences, la fonction coût présente une forme plus convexe. L’optimisation successive pour des données de contenu fréquentiel croissant permet à l’algorithme d’optimisation d’approcher le bassin d’attraction du minimum global et finalement de converger vers le modèle optimal.

être sérieusement envisagées dans des problèmes impliquant un nombre important de paramètres, comme c’est le cas en FWI. Il est néanmoins possible, dans le cas de problème faiblement non-linéaires, de se tourner vers des méthodes d’optimisation locales en linéarisant le problème inverse (Lailly, 1984; Tarantola, 1984b). On suppose pour cela que la fonction coût est localement quadratique aux alentours d’une estimation initiale des paramètres du milieu. L’optimisation est alors effectuée en utilisant des algorithmes de descente basés sur le calcul du gradient de la fonction coût. La non-linéarité est prise en compte, en mettant à jour les paramètres du milieu jusqu’à atteindre un critère donné de convergence. De manière équivalente aux méthodes globales, il est possible d’utiliser l’inférence bayésienne pour analyser, localement, la confiance que l’on peut apporter au modèle estimé (Gouveia & Scales, 1997, 1998)⁴.

Une analogie existe entre la résolution de la FWI sous forme d’un problème linéarisé et les approches d’inversions de formes d’ondes par migration "ray+Born" du champ d’ondes diffractées (Miller et al., 1987; Beydoun & Tarantola, 1988; Beylkin & Burridge, 1990; Bostock & Rondenay, 1999). Pour ces dernières, le problème direct est linéarisé au moyen de l’approximation de Born (voir section 1.5.2) si bien qu’un problème inverse linéaire est résolu. Le modèle initial est ainsi fixe et seules des interactions de premier ordre sont prises en compte lors de la

4. Ces approches ne sont valides que si l’algorithme local a convergé dans le bassin d’attraction global

résolution du problème direct. Au contraire, le problème direct est résolu de manière exact en FWI, i.e. sans linéarisation ; il rend ainsi compte de tous les effets de propagation. Néanmoins, la linéarisation du problème inverse implique que les perturbations apportées au modèle sont elles aussi de premier ordre. Cependant, le processus itératif permet de mettre à jour le modèle à l'issue de chaque itération et de prendre progressivement en compte des effets de diffractions plus complexes lors de la modélisation sismique.

Un effet de la linéarisation est la grande sensibilité de l'optimisation à l'estimé initial du modèle. S'il n'est pas assez précis, le point de départ de l'optimisation se situe dans le bassin d'attraction d'un minimum local et l'inversion ne converge pas vers le modèle optimal (Fig. 3.3). Une approche pour contourner partiellement ce problème, proposée par [Bunks et al. \(1995\)](#), est de réaliser l'inversion de manière hiérarchique en augmentant progressivement le contenu fréquentiel des données observées. L'intérêt est double : d'une part elle permet d'atténuer la non-linéarité du problème en réduisant les effets de « *cycle-skipping*⁵ » ([Virieux & Operto, 2009](#)), favorisant ainsi la convergence vers le minimum global en augmentant la convexité de la fonction coût. D'autre part, elle définit naturellement une approche multi-échelles de l'imagerie en injectant progressivement des détails de plus en plus fins (Fig 3.2 et 3.3).

En amont de la résolution d'un problème inverse, il est essentiel de définir les variables représentant les données ainsi que les paramètres du modèle, c'est-à-dire, la *paramétrisation* du problème. Dans bien des cas le choix de cette paramétrisation est arbitraire et il n'y a pas de raisons fortes pour en choisir une plutôt qu'une autre. Toutefois, ce choix devient très important lorsqu'on s'intéresse aux problèmes inverses non-linéaires dont les solutions dépendent des paramétrisations considérées. Ainsi, selon la paramétrisation choisie, une solution différente peut être obtenue⁶.

3.1.4 Méthodes de résolution locale

Partant d'une solution approchée initiale \mathbf{m}_0 du problème d'optimisation, les méthodes itératives locales génèrent une séquence de solutions $\mathbf{m}_k = (\mathbf{m}_0, \mathbf{m}_1, \dots, \mathbf{m}_k)$ incrémentées à chaque itération par la relation :

$$\begin{aligned} \mathbf{m}_{k+1} &= \mathbf{m}_k + \delta\mathbf{m}_k \\ &= \mathbf{m}_k + \alpha_k \mathbf{p}_k, \end{aligned} \tag{3.4}$$

où $\delta\mathbf{m}_k$ correspond aux perturbations à apporter aux paramètres pour améliorer la solution estimée. Ces perturbations qui dépendent de deux quantités, \mathbf{p}_k la *direction de descente* et α_k le *pas de descente* calculés à chaque itération, sont déterminées de manière à assurer une décroissance et un taux de convergence suffisants de la fonction coût.

5. Une différence de phase, entre les données observées et calculées, supérieure à une demi-période. Ce problème est lié à la précision cinématique du modèle et, est responsable de l'apparition de minima locaux.

6. Ce n'est pas le cas des problèmes linéaires pour lesquels la solution obtenue est invariante de toute reparamétrisation linéaire des données et des paramètres du modèle.

Directions de descente

Une direction de descente est par définition tout vecteur \mathbf{p}_k qui permet une réduction de la fonctionnelle \mathcal{C} entre les itération k et $k + 1$, c'est-à-dire lorsque :

$$\mathcal{C}(\mathbf{m}_{k+1}) < \mathcal{C}(\mathbf{m}_k). \quad (3.5)$$

En développant au premier ordre l'expression de la fonctionnelle \mathcal{C} , il vient :

$$\mathcal{C}(\mathbf{m}_k + \alpha_k \mathbf{p}_k) \approx \mathcal{C}(\mathbf{m}_k) + \alpha_k \mathbf{p}_k^T \nabla \mathcal{C}(\mathbf{m}_k) + \mathcal{O}(\alpha_k^2 \mathbf{p}_k^2) \quad (3.6)$$

que nous pouvons injecter dans l'inéquation (3.5). Nous déduisons donc que si \mathbf{p}_k est une direction de descente, alors $\alpha_k \mathbf{p}_k^T \nabla \mathcal{C}(\mathbf{m}_k) < 0$. Or, puisque $\alpha_k > 0$, il vient :

$$\boxed{\mathbf{p}_k^T \nabla \mathcal{C}(\mathbf{m}_k) < 0.} \quad (3.7)$$

Ainsi, \mathbf{p}_k est une direction de descente si et seulement si le produit scalaire entre le gradient de la fonctionnelle et la direction de descente est négatif.

Le pas de descente, recherche linéaire

La recherche d'un pas de descente est nécessaire pour forcer la convergence des algorithmes d'optimisation. En effet, une trop faible valeur de α aura tendance à ralentir la convergence du processus, tandis qu'une valeur trop élevée peut empêcher la fonction coût de décroître et, dans des cas extrêmes, mener à des algorithmes oscillants⁷.

La recherche de α constitue, en soit, un problème d'optimisation qu'il nous faut résoudre à chaque itération de l'algorithme de résolution locale. Deux classes de méthodes permettent de définir ce pas, les méthodes de régions de confiance et les méthodes de recherche linéaire (Nocedal & Wright, 2006a). Nous ne nous intéressons qu'à ce dernier type de méthodes qui sont plus faciles à mettre en œuvre que les méthodes à régions de confiance. Les algorithmes de recherche linéaire visent à déterminer le pas de descente optimal le long de la direction de descente. Une approche naturelle est de procéder à une recherche linéaire *exacte*, c'est-à-dire, à résoudre le problème d'optimisation pour trouver le pas de descente optimal (point rouge sur la figure 3.4). Cette stratégie est cependant difficile à appliquer dans le cadre de problèmes fortement non-linéaires puisqu'elle implique d'évaluer de nombreuses fois la fonction coût.

Alternativement des méthodes de recherche linéaire *inexactes* peuvent être considérées. Ces dernières tentent, à partir d'une valeur de α_k initiale, de trouver un pas vérifiant certains critères de décroissance de la fonction coût. Des critères populaires sont les conditions de Wolfe (Nocedal & Wright, 2006a) que nous détaillons ci-après.

Supposons qu'une direction de descente \mathbf{p}_k soit connue, le problème d'optimisation à résoudre est

$$\arg \min_{\alpha > 0} \mathcal{C}(\mathbf{m}_k + \alpha \mathbf{p}_k). \quad (3.8)$$

7. Il est ainsi vivement déconseillé de recourir à des pas de descente fixes.

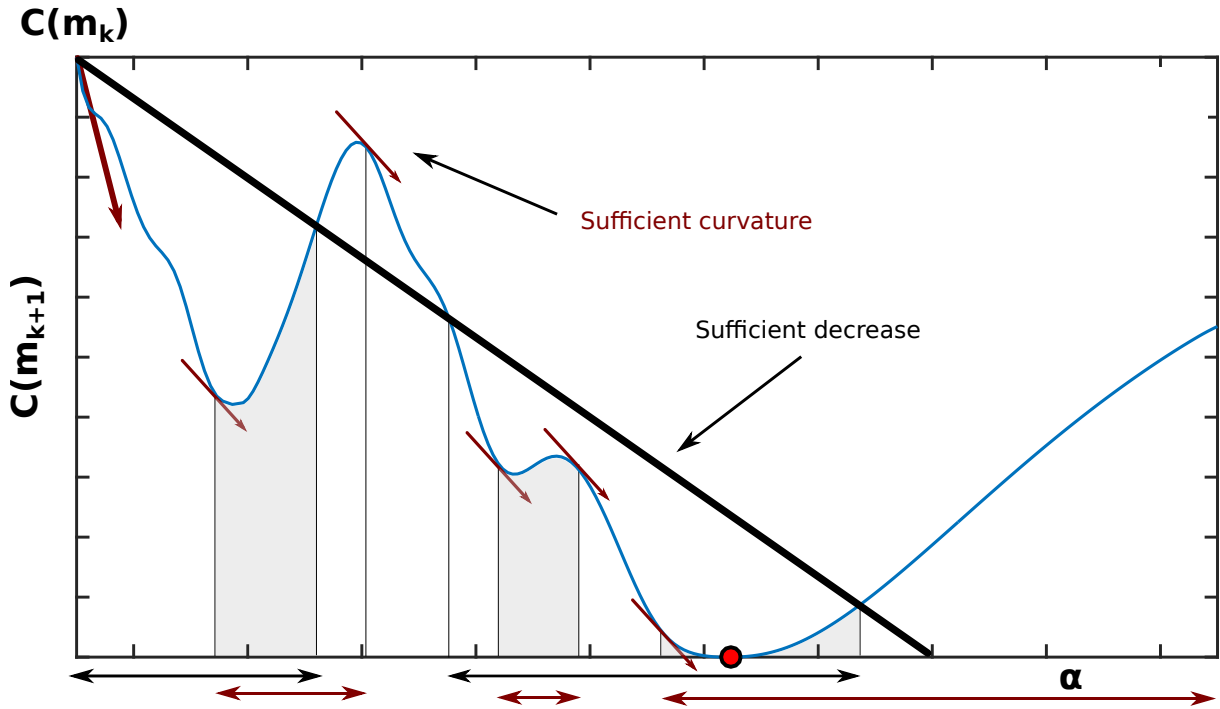


Figure 3.4 – Méthode de recherche linéaire, les conditions de Wolfe. La courbe bleue présente la fonction coût à l'itération $k + 1$ (dans la direction de descente) en fonction de α le pas de descente. L'optimum de α est matérialisé par une pastille rouge, il correspond à la valeur du pas si une recherche linéaire exacte avait été utilisée. La ligne noire correspond à la première condition de Wolfe (décroissance suffisante), tout α donnant une valeur de \mathcal{C} inférieure à cette courbe satisfait la condition et est acceptable (flèches noires). Les tangentes rouges représentent les pentes de la fonction coût satisfaisant la seconde condition de Wolfe (courbure suffisante). Les valeurs de α satisfaisant cette condition sont représentées par les flèches rouges bi-directionnelles. L'intersection de ces intervalles (zones grisées) correspond aux α satisfaisant les deux conditions à la fois. Dès qu'un α est dans une zone grisée, un pas de descente est trouvé et la recherche linéaire s'arrête.

Une première condition essentielle pour minimiser le problème est que la fonction coût décroît d'une itération à l'autre, c'est-à-dire que $\mathcal{C}(\mathbf{m}_{k+1}) < \mathcal{C}(\mathbf{m}_k)$. Pour accélérer la convergence, nous pouvons être plus restrictif et imposer que la fonction coût décroît au minimum d'une quantité donnée :

$$\mathcal{C}(\mathbf{m}_{k+1}) < \mathcal{C}(\mathbf{m}_k) + C, \quad (3.9)$$

où C est une constante négative. En développant au second ordre $\mathcal{C}(\mathbf{m}_{k+1})$, il vient :

$$\mathcal{C}(\mathbf{m}_{k+1}) \approx \mathcal{C}(\mathbf{m}_k) + \alpha_k \mathbf{p}_k^T \nabla \mathcal{C}(\mathbf{m}_k) + \mathcal{O}(\alpha_k^2). \quad (3.10)$$

En injectant cette équation dans (3.9) et en remarquant qu'une direction de descente n'est obtenue que si $\mathbf{p}_k^T \nabla \mathcal{C}(\mathbf{m}_k) < 0$ nous pouvons écrire que

$$\mathcal{C}(\mathbf{m}_{k+1}) \leq \mathcal{C}(\mathbf{m}_k) + c_1 \alpha_k \mathbf{p}_k^T \nabla \mathcal{C}(\mathbf{m}_k). \quad (3.11)$$

qui correspond à la première condition de Wolfe aussi nommée règle d'Armijo. Cette relation impose un taux de décroissance en fonction du paramètre $c_1 \in [0, 1]$. Plus c_1 est proche de 1, et plus nous imposons un taux de décroissance important et plus le pas devient difficile à trouver. Cette condition est matérialisée sur la figure 3.4 par un trait noir ; toutes valeurs de α pour lesquelles la fonction coût est inférieure à cette ligne est acceptable.

En général, la première condition de Wolfe n'est pas utilisée toute seule puisqu'elle nous expose au risque d'accepter des pas de descente trop petits entraînant de fait une convergence lente. Pour éviter cela, nous introduisons la seconde condition de Wolfe qui est une condition de courbure suffisante :

$$\mathbf{p}_k^T \nabla \mathcal{C}(\mathbf{m}_{k+1}) \geq c_2 \mathbf{p}_k^T \nabla \mathcal{C}(\mathbf{m}_k), \quad (3.12)$$

elle-même contrôlée par le paramètre $c_2 \in [c_1, 1]$. Elle impose que la pente de la fonction coût à l'itération $k + 1$ soit supérieure à celle de l'itération k fois c_2 . On s'assure ainsi que la recherche linéaire ne s'arrête pas pour des pas trop petits qui ralentiraient la convergence, c'est-à-dire en des points où la pente autorise localement une meilleur réduction de la fonction coût. Au contraire, cette condition force la recherche linéaire à trouver un α proche d'un extremum local (Fig. 3.4, les flèches rouges).

En pratique (Monteiller et al., 2015), on donne un estimé initial de α , si ce pas satisfait les deux conditions de Wolfe (3.11) et (3.12) alors le pas est gardé. Dans le cas ou aucune de ces condition n'est satisfaite, un nouveau α est testé :

- S'il ne respecte pas la première condition (3.11), alors α est trop grand et on teste un α plus petit.
- Sinon, s'il ne respecte pas la seconde condition (3.12), alors α est trop petit et on teste un pas plus grand.

Enfin, s'il n'est plus possible de trouver un α qui satisfasse ces relations, la convergence est atteinte et l'algorithme est stoppé. Évidemment, le comportement de cet algorithme dépend grandement du choix des paramètres c_1 et c_2 . Nocedal & Wright (2006a) préconisent de prendre des valeurs de $c_1 = 10^{-4}$ et $c_2 = 0.9$, dans le cas de l'utilisation d'algorithmes de Newton ou de quasi-Newton.

Méthode du gradient

Un choix naturel de direction de descente est de prendre la direction de la pente locale de la fonction coût évaluée à l'itération k . Il s'agit de la méthode de plus grande pente (« *steepest descent* », en anglais). Cette direction est obtenue en prenant la direction opposée du gradient de la fonction coût qui, lui, définit la direction de plus grande montée. La direction de descente est donc donnée par :

$$\mathbf{p}_k = -\nabla \mathcal{C}_k, \quad (3.13)$$

qui satisfait nécessairement la définition donnée en (3.7). Cette méthode, plutôt simple puisqu'elle ne nécessite que l'évaluation du gradient de la fonction n'est toutefois pas optimale pour la convergence du problème d'optimisation. En effet, même pour des problèmes d'optimisation linéaires, le taux de convergence de la solution est linéaire (Luenberger, 1973). En

outre, l'efficacité de la méthode du gradient est très dépendante de la mise à l'échelle des paramètres lorsque plusieurs classes de paramètres de nature différente sont reconstruites et il est généralement nécessaire de préconditionner le gradient.

Méthode de Newton

Une meilleure direction de descente peut être obtenue en effectuant un développement de Taylor au second ordre de la fonction coût :

$$\mathcal{C}(\mathbf{m}_k + \alpha_k \mathbf{p}_k) \approx \mathcal{C}(\mathbf{m}_k) + \alpha_k \mathbf{p}_k^T \nabla \mathcal{C}(\mathbf{m}_k) + \frac{1}{2} \alpha_k^2 \mathbf{p}_k^T \nabla^2 \mathcal{C}(\mathbf{m}_k) \mathbf{p}_k + \mathcal{O}(\alpha_k^3 \mathbf{p}_k^3). \quad (3.14)$$

En dérivant cette équation par rapport à \mathbf{p}_k , il vient :

$$\alpha_k \nabla^2 \mathcal{C}(\mathbf{m}_k) \mathbf{p}_k = -\nabla \mathcal{C}(\mathbf{m}_k), \quad (3.15)$$

qui correspond à l'équation de Newton et que nous pouvons réécrire, si l'inverse de $\nabla^2 \mathcal{C}(\mathbf{m}_k)$ existe, comme :

$$\alpha_k \mathbf{p}_k = -(\nabla^2 \mathcal{C}(\mathbf{m}_k))^{-1} \nabla \mathcal{C}(\mathbf{m}_k) = \delta \mathbf{m}_k. \quad (3.16)$$

Cette dernière relation indique que la direction de descente peut être obtenue en multipliant les dérivées premières de la fonction coût (le gradient) par l'inverse de la matrice contenant les dérivées secondes de la fonction coût (le Hessien). Elle indique aussi que, dans le cas où la fonction coût est quadratique et donc que le problème est linéaire, le minimum de la fonction est obtenu en une itération. Qui plus est, la direction de Newton donne directement l'expression de la perturbation à apporter au modèle $\delta \mathbf{m}_k$. C'est-à-dire qu'il n'est pas nécessaire de déterminer le pas de descente par recherche linéaire.

Tant que la fonction coût peut être localement approximée par une fonction convexe et tant que le Hessien est défini positif la méthode de Newton procure une direction de descente fiable qui est de surcroît bien mise à l'échelle par la prise en compte du Hessien. Son utilisation est ainsi très avantageuse pour les problèmes multi-paramètres. Un autre avantage réside dans son taux de convergence quadratique, le minimum est ainsi trouvé plus rapidement qu'en considérant la méthode du gradient.

Néanmoins, l'inverse du Hessien n'est généralement pas accessible. Tout d'abord, parce que cette matrice peut être très mal conditionnée et donc difficile voire impossible à inverser. Mais aussi car le Hessien, ainsi que son inverse, peuvent être excessivement coûteux à calculer et à stocker. Il existe cependant diverses alternatives à la méthode de Newton qui permettent d'approcher implicitement (et économiquement) les effets du Hessien : les méthodes de quasi-Newton dont les taux de convergence sont super-linéaires (Kelley, 1999). Nous présentons dans la section suivante une méthode de quasi-Newton permettant d'estimer itérativement les effets de l'inverse du Hessien, la méthode l-BFGS (Nocedal, 1980). D'autres méthodes de second ordre comme les méthodes de Gauss-Newton (Pratt et al., 1998) ou de Newton tronqué (Fichtner & Trampert, 2011a; Métivier et al., 2013) peuvent être envisagées pour approximer les effets du Hessien.

Méthodes de quasi-Newton : BFGS et l-BFGS

La méthode BFGS⁸ est une méthode de quasi-Newton qui cherche à approximer l'inverse du Hessien \mathbf{H}_k^{-1} et à raffiner cette approximation par un processus itératif. La direction de descente s'exprime de manière similaire à la direction de Newton (3.16) par :

$$\mathbf{p}_k = -\mathbf{H}_k^{-1} \nabla \mathcal{C}(\mathbf{m}_k). \quad (3.17)$$

La mise-à-jour de l'approximation de l'inverse du Hessien à l'itération $k + 1$ est donnée par (Nocedal & Wright, 2006a) :

$$\mathbf{H}_{k+1}^{-1} = \left(\mathbf{I} - \frac{\mathbf{s}_k \mathbf{y}_k^T}{\mathbf{y}_k^T \mathbf{s}_k} \right) \mathbf{H}_k^{-1} \left(\mathbf{I} - \frac{\mathbf{s}_k \mathbf{y}_k^T}{\mathbf{y}_k^T \mathbf{s}_k} \right) + \frac{\mathbf{s}_k \mathbf{s}_k^T}{\mathbf{y}_k^T \mathbf{s}_k}, \quad (3.18)$$

qui ne dépend que de l'estimation de l'inverse du Hessien à l'itération précédente et de différences finies sur les modèles \mathbf{s} et gradients \mathbf{y} des itérations courante $k + 1$ et précédente k telles que $\mathbf{s}_k = \mathbf{m}_{k+1} - \mathbf{m}_k$ et $\mathbf{y}_k = \nabla \mathcal{C}(\mathbf{m}_{k+1}) - \nabla \mathcal{C}(\mathbf{m}_k)$. Bien que l'équation (3.18) permette de construire itérativement une approximation du Hessien, il est toujours nécessaire de stocker et de manipuler la matrice \mathbf{H}_{k+1}^{-1} qui est très dense pour des problèmes de tailles conséquentes (sa taille est $N_p \times N_p$).

Néanmoins, l'équation (3.17) indique que la direction est obtenue par un produit matrice-vecteur entre l'approximation de l'inverse du Hessien et le gradient. Il n'est ainsi pas nécessaire de stocker la matrice de manière explicite, mais seulement d'être capable de calculer l'effet de cette approximation sur le vecteur gradient. Ceci est possible au moyen de la méthode l-BFGS (pour *limited-memory BFGS*) (Nocedal, 1980; Nocedal & Wright, 2006a), qui est une modification de la méthode BFGS. Son principe est d'utiliser la relation (3.18) de façon récursive de sorte à exprimer \mathbf{H}_{k+1}^{-1} en fonction de tous les vecteurs $\mathbf{s}_{1\dots k}$ et $\mathbf{y}_{1\dots k}$. De ce fait, seuls $2 \times (k + 1)$ vecteurs de taille N_p doivent être stockés en mémoire. Ce nombre peut encore être réduit en se défaussant des modèles et gradients les plus anciens, ne gardant de fait qu'un nombre très limité de vecteurs en mémoire.

Notons que les algorithmes BFGS ou l-BFGS nécessitent un estimé initial de l'inverse du Hessien \mathbf{H}_0^{-1} . Lorsqu'on ne dispose pas de cette estimation, un bon choix (Nocedal & Wright, 2006a) est de prendre $\mathbf{H}_0^{-1} = \gamma \mathbf{I}$ où γ est donné par :

$$\gamma = \frac{\mathbf{s}_0^T \mathbf{y}_0}{\mathbf{y}_0^T \mathbf{y}_0}, \quad (3.19)$$

qui permet notamment de calibrer la direction de descente \mathbf{p}_k de sorte que $\alpha = 1$ soit en général accepté. Dans le reste de ce manuscrit, nous utilisons exclusivement l'algorithme d'optimisation l-BFGS pour l'application de la FWI.

8. Nommée d'après ses inventeurs : Broyden, Fletcher, Goldfarb et Shanno.

3.2 L'inversion de formes d'ondes, un problème aux moindres carrés linéarisé

L'inversion de formes d'ondes (Tarantola, 1984a, 1987; Gauthier et al., 1986a) est un problème inverse non-linéaire et mal-posé dont l'objet est de reconstruire un modèle quantitatif du sous-sol paramétré par un jeu de propriétés physiques qui gouvernent la propagation des ondes sismiques. Ces inconnues de notre problème sont trop nombreuses pour envisager des méthodes d'optimisations globales qui requièrent de surcroît un nombre conséquent de coûteuses modélisations numériques. Nous formulons donc la FWI comme un problème d'optimisation locale aux moindres carrés (Tarantola, 1984b).

3.2.1 La fonction coût

La FWI est fondée sur la minimisation de l'écart, échantillon par échantillon, entre les formes d'ondes observées \mathbf{d}_{obs} et les formes d'ondes calculées \mathbf{d}_{cal} à partir d'un modèle du sous-sol \mathbf{m} et de la théorie physique \mathbf{G} de la propagation des ondes. Les données calculées sont ainsi recueillies par extraction du champ d'onde modélisé de sorte que $\mathbf{d}_{cal} = \mathbf{R}\mathbf{u}(\mathbf{x}, t)$ où \mathbf{R} est un opérateur d'extraction qui dans sa forme la plus simple correspond à un delta de Dirac positionné au récepteur.

L'écart sur les données à minimiser est défini par le vecteur résidu $\Delta\mathbf{d}$ qui correspond à la différence entre les données observées et les données modélisées tel que

$$\Delta\mathbf{d} = \mathbf{d}_{cal}(\mathbf{m}) - \mathbf{d}_{obs}. \quad (3.20)$$

Dans le cadre d'une application sismologique, les données correspondent aux sismogrammes enregistrés, soient des séries temporelles de N_t échantillons, par un réseau de N_r capteurs à N_c composantes pour un certain nombre de séismes N_s . La taille totale du vecteur résidu sur les données N_d sera en conséquence égale à $N_s \times N_r \times N_c \times N_t$. Une application télésismique impliquant une dizaine d'événements enregistrés par une centaine de stations large-bandes induit ainsi un nombre de données de plusieurs millions d'observations.

Afin de minimiser chacun des termes du vecteur résidu, nous introduisons un scalaire représentant l'ajustement sur les données appelée fonction coût ou fonctionnelle d'erreur. Cette fonction coût est définie par le type de résidu considéré – dans notre cas la forme d'onde mais nous pourrions tout aussi bien considérer l'enveloppe, la phase instantanée ou même des temps de trajet (Bozdağ et al., 2011) – ainsi que par le choix d'une norme $\|\cdot\|$ du vecteur $\Delta\mathbf{d}$ caractérisant la distance entre les données observées et modélisées. Le choix de la norme n'est pas évident pour la FWI. Dans le cas où le bruit sur les données vérifie une loi normale, c'est à dire que la statistique du bruit est gaussienne, un choix pertinent est la norme L_2 ou norme euclidienne carrée (Tarantola, 1987). Lorsque le bruit ne suit pas une statistique gaussienne et présente un nombre important de valeurs aberrantes (*outliers*), la norme L_2 aura tendance à donner plus de poids aux forts résidus puisque ces derniers sont élevés au carré. Il peut-être dans ce cas plus judicieux de considérer des normes plus robustes et moins sensibles à ces anomalies telles que la norme L_1 qui minimise la valeur absolue du résidu ou de faire appel à des normes

hybrides comme la norme d'Huber se comportant comme une norme L_2 pour de faibles résidus et comme une L_1 pour des résidus importants (Tarantola, 1987; Crase et al., 1990; Brossier et al., 2009a).

Dans ce qui suit, nous considérons que les données observées ont été sélectionnées après élimination des données trop bruitées et nous supposons que l'hypothèse d'une statistique gaussienne est vérifiée, ce choix est fait de manière assez classique en FWI (Pratt & Worthington, 1990; Pratt, 1990b). Nous nous focalisons donc sur la norme L_2 qui réduit la FWI à un problème aux moindres carrés et définissons donc la fonction coût que nous cherchons à minimiser par :

$$\mathcal{C}(\mathbf{m}) = \frac{1}{2} \|\Delta \mathbf{d}\|^2 = \frac{1}{2} \Delta \mathbf{d}^T \Delta \mathbf{d} \quad (3.21)$$

où T représente l'opérateur transposé.

3.2.2 Linéarisation du problème

Nous avons vu dans la section 3.1.3 que la minimisation de la fonction coût (3.21) peut se faire aussi bien par des algorithmes d'optimisation locaux que globaux. Cependant, au vu du très grand nombre de paramètres définissant le modèle \mathbf{m} et du coût numérique de la résolution d'un problème direct, il est impossible ne serait-ce que d'envisager l'utilisation de méthodes globales pour des application tri-dimensionnelles en temps. Nous reformulons donc la FWI sous-forme d'un problème itératif linéarisé afin de pouvoir appliquer des méthodes d'optimisation locale. Pour ce faire, nous considérons que le minimum de la fonction coût $\mathcal{C}(\mathbf{m}^k)$ à l'itération k est recherché aux alentours d'un modèle \mathbf{m}^{k-1} obtenu lors de l'itération précédente⁹. Si les perturbations $\Delta \mathbf{m}^k$ à apporter au modèle \mathbf{m}^{k-1} sont relativement faibles, nous pouvons écrire qu'au premier ordre :

$$\mathbf{m}^k = \mathbf{m}^{k-1} + \Delta \mathbf{m}^k. \quad (3.22)$$

Considérant (3.22), nous pouvons réaliser un développement Taylor de la fonctionnelle (3.21) limité à l'ordre deux¹⁰, qui exprime la valeur de $\mathcal{C}(\mathbf{m}^k)$ en fonction de $\mathcal{C}(\mathbf{m}^{k-1})$:

$$\begin{aligned} \mathcal{C}(\mathbf{m}^k) = \mathcal{C}(\mathbf{m}^{k-1} + \Delta \mathbf{m}^k) &= \mathcal{C}(\mathbf{m}^{k-1}) + \sum_{i=1}^{N_p} \frac{\partial \mathcal{C}(\mathbf{m}^{k-1})}{\partial m_i} \Delta m_i^k \\ &+ \frac{1}{2} \sum_{i=1}^{N_p} \sum_{j=1}^{N_p} \frac{\partial^2 \mathcal{C}(\mathbf{m}^{k-1})}{\partial m_i \partial m_j} \Delta m_i^k \Delta m_j^k + \mathcal{O}(\mathbf{m}^3). \end{aligned} \quad (3.23)$$

La différentiation de cette équation par rapport à un paramètre m_l du modèle nous donne :

$$\frac{\partial \mathcal{C}(\mathbf{m}^k)}{\partial m_l} = \frac{\partial \mathcal{C}(\mathbf{m}^{k-1})}{\partial m_l} + \sum_{i=1}^{N_p} \frac{\partial^2 \mathcal{C}(\mathbf{m}^{k-1})}{\partial m_i \partial m_l} \Delta m_i^k. \quad (3.24)$$

9. À l'initialisation du processus, il est évidemment nécessaire de définir un modèle de départ \mathbf{m}_0 , nous discutons le choix de ce modèle dans le chapitre 4.

10. Afin d'appliquer des méthodes d'optimisation de type Newton ou Quasi-Newton vues précédemment.

Or, un point critique – minimum, maximum ou point selle – de la fonctionnelle $\mathcal{C}(\mathbf{m}^k)$ est atteint dès que sa dérivée première (3.24) par rapport à \mathbf{m} est nulle ¹¹. Annulant le membre de gauche de l'équation (3.24), il vient :

$$\Delta \mathbf{m}^k = - \left[\underbrace{\frac{\partial^2 \mathcal{C}(\mathbf{m}^{k-1})}{\partial \mathbf{m}^2}}_{\text{Hessian}} \right]^{-1} \underbrace{\frac{\partial \mathcal{C}(\mathbf{m}^{k-1})}{\partial \mathbf{m}}}_{\text{Gradient}}. \quad (3.25)$$

Cette relation correspond à l'équation de Newton qui exprime le vecteur de perturbation des paramètres du modèle et permet d'obtenir un minimum local de la fonction coût lorsque celle-ci est localement convexe. Le terme $\Delta \mathbf{m}^k$ est déterminé par le vecteur gradient ¹² et l'inverse de la matrice hessienne de la fonction coût qui définissent la direction de plus grande pente ainsi que la courbure au voisinage du modèle \mathbf{m}_{k-1} .

Il est intéressant de noter que l'équation (3.25) donne le minimum de la fonctionnelle en une itération si la fonction coût est quadratique, par exemple, lorsqu'on considère un problème inverse linéaire. En revanche, pour un problème non-linéaire il est nécessaire de linéariser la mise à jour des perturbations et de procéder à une résolution itérative.

3.2.3 Les équations normales

L'équation de Newton (3.25) permet d'exprimer le vecteur des perturbations à apporter au modèle à chaque itération en fonction du gradient et de l'inverse du hessien de la fonction coût. En explicitant les dépendances de la fonction coût aux données observées et calculées nous pouvons réécrire le gradient et le hessien en fonction de la matrice des dérivées de Fréchet ou matrice de sensibilité $\mathbf{J} = \frac{\partial \mathbf{d}_{cal}(\mathbf{m})}{\partial \mathbf{m}}$ et retrouver les équations normales du problème aux moindres carrés.

Le gradient

Le gradient $\nabla_{\mathbf{m}} \mathcal{C}$ est défini comme la dérivée première de la fonction coût par rapport aux paramètres du modèle \mathbf{m} . Après différentiation, il vient :

$$\nabla_{\mathbf{m}} \mathcal{C} = \left(\frac{\partial \mathbf{d}_{cal}(\mathbf{m})}{\partial \mathbf{m}} \right)^T \Delta \mathbf{d}(\mathbf{m}) = \mathbf{J}^T \Delta \mathbf{d}. \quad (3.26)$$

Le gradient de la fonction coût est donc un vecteur (de taille n_p) correspondant au produit de la transposé de la matrice de sensibilité \mathbf{J} et du résidu sur les données $\Delta \mathbf{d}$, il correspond à une intercorrélacion à décalage nul.

Pour obtenir le gradient, il est ainsi nécessaire de déterminer la matrice de sensibilité \mathbf{J} . Dans la plupart des cas, il n'est pas possible de la déterminer analytiquement puisque nous

11. Première condition d'optimalité.

12. Aussi dénommée matrice jacobienne

ne disposons pas des solutions analytiques du champ d'ondes, et il doit donc être déterminé numériquement. Une approche naïve consiste à la déterminer par différences-finies, ce qui nécessiterait de réaliser un nombre de modélisations numériques équivalent au produit du nombre de données et de paramètres.

Une seconde option, présentée par Pratt et al. (1998), consiste à réaliser qu'une colonne de \mathbf{J} est physiquement équivalente au champ d'ondes diffracté par une hétérogénéité m_i , le champ d'onde aux dérivées partielles, échantillonné par les récepteurs¹³. Ce champ d'onde satisfait l'équation d'onde pour une source virtuelle, localisée à la position du point diffractant m_i , et caractérisée par la diagramme de radiation du paramètre m_i (Forgues, 1996). La matrice \mathbf{J} peut ainsi être construite en effectuant autant de modélisations que de paramètres décrivant le modèle.

Ce coût peut encore être réduit en exploitant la réciprocité spatiale des fonctions de Green (Shin et al., 2001; Chen et al., 2007). Les dérivées de Fréchet sont alors obtenues en effectuant autant de modélisations que de couples source-récepteur.

Alternativement, au lieu de construire la matrice de sensibilité \mathbf{J} de manière explicite, nous pouvons directement déterminer le gradient de la fonction coût au moyen de deux simulations par source en utilisant le formalisme de l'état adjoint (Lions, 1968; Tromp et al., 2005; Fichtner et al., 2006a,b; Plessix, 2006a). Le calcul du gradient est alors indépendant du nombre de paramètre et du nombre de récepteur. Nous détaillons ces développements dans la section 3.3.

Le Hessien

Suivant le même raisonnement, il est possible de déduire l'expression du hessien en dérivant le gradient de la fonction coût par rapport à \mathbf{m} , nous obtenons ainsi :

$$\nabla_{\mathbf{m}}^2 \mathcal{C} = \nabla_{\mathbf{m}} \nabla_{\mathbf{m}} \mathcal{C} = \mathbf{J}^T \mathbf{J} + \left(\frac{\partial \mathbf{J}}{\partial \mathbf{m}} \right)^T (\Delta \mathbf{d} \dots \Delta \mathbf{d}). \quad (3.27)$$

Le Hessien est une matrice $n_p \times n_p$ qui est formée par la somme de deux termes :

1. Un terme linéaire $\mathbf{J}^T \mathbf{J}$, souvent appelé Hessien approximé. Il correspond à l'intercorrélation à décalage nul entre les champs aux dérivées partielles contenus dans les matrices \mathbf{J} créés par deux paramètres du modèle m_i et m_j . Le Hessien approximé est dominé par ses termes diagonaux qui correspondent à la somme des autocorrélations des champs aux dérivées partielles (la somme des amplitudes au carré) pour chaque paramètre m_i . Appliqués au gradient, ils permettent de corriger les effets liés à la décroissance géométrique des champs diffractés. Les termes non-diagonaux sont calculés par sommation des intercorrélations des champs aux dérivées partielles entre deux paramètres m_i et m_j qui peuvent être de classes différentes et/ou de positions différentes. Ils renseignent sur la corrélation entre différents paramètres et prennent en compte les effets de largeur de bande limitée de la source et du dispositif d'acquisition. Son application au gradient peut être vue comme

13. La matrice \mathbf{J} représente l'opérateur de Born (Lailly, 1983b; Pratt, 1990a; Pratt & Sams, 1996), c'est-à-dire, les hyperboles de diffraction pour chaque paramètre \mathbf{m} du milieu.

un opérateur de déconvolution (Pratt et al., 1998) re-focalisant les perturbations déterminées par le gradient. Il permet en outre, dans le cas de problèmes multi-paramètres, de dimensionner les amplitudes respectives des gradient par rapport aux différentes classes de paramètre. Tout comme le gradient de la fonction coût, ce terme du Hessien ne prend en compte que les effets de diffraction simple.

2. Un terme non-linéaire $\left(\frac{\partial \mathbf{J}}{\partial \mathbf{m}}\right)^T (\Delta \mathbf{d} \dots \Delta \mathbf{d})$ qui correspond au produit des dérivées partielles secondes du champ d'onde et des résidus sur les données. Il caractérise les effets de diffraction double (Pratt et al., 1998) : il corrige l'empreinte des champs doublement diffractés interprétés comme des diffractions simples par le gradient. Ce terme est généralement considéré comme négligeable si le problème est quasi-linéaire ou que les résidus sur les données sont faibles (Tarantola, 1987).

Le calcul du premier terme du Hessien est assez coûteux en domaine temporel puisqu'il nécessite de construire la matrice des dérivées de Fréchet. Contrairement au gradient, elle ne peut être obtenue à faible coût par des méthodes adjointes mais peut être calculée en convoluant les fonctions de Green pour chaque couple source-récepteurs (Sheen et al., 2006; Chen et al., 2007). La seule prise en compte de ce terme mène aux algorithmes d'optimisation locale de type Gauss-Newton. Le terme de second ordre du Hessien est d'autant plus cher à déterminer qu'il nécessite d'évaluer les dérivées de la matrice de sensibilité par rapport aux paramètres du modèle. Des techniques de double adjoint (Fichtner & Trampert, 2011a; Métivier et al., 2013) permettent néanmoins d'évaluer l'effet d'une ligne de l'inverse du Hessien complet sur un vecteur (typiquement le gradient). Ces techniques permettent d'envisager des algorithmes d'optimisation du type Newton tronqué ainsi que des analyses locales de résolution (Fichtner & Trampert, 2011b).

Les équations normales

En injectant les expressions du gradient (3.26) et du hessien (3.27) dans l'équation de Newton (3.25), nous obtenons les équations normales :

$$\Delta \mathbf{m} = \left[\mathbf{J}^T \mathbf{J} + \left(\frac{\partial \mathbf{J}}{\partial \mathbf{m}}\right)^T (\Delta \mathbf{d} \dots \Delta \mathbf{d}) \right]^{-1} \mathbf{J}^T \Delta \mathbf{d}. \quad (3.28)$$

L'analyse dimensionnelle du gradient indique que ce dernier n'est pas homogène à la dimension d'une perturbation sur le modèle $\frac{[D] \times [D]}{[M]}$. Le gradient n'est donc pas suffisant pour obtenir des perturbations réalistes des paramètres du modèle. Néanmoins, les équations normales, par la prise en compte de l'inverse du Hessien dont la dimension est $\frac{[D] \times [D]}{[M] \times [M]}$, permettent de dimensionner les perturbations données par le gradient puisque $\frac{[D] \times [D]}{[M]} \times \left(\frac{[D] \times [D]}{[M] \times [M]}\right)^{-1} = [M]$ est homogène aux paramètres du modèle.

3.2.4 Régularisation

Le caractère mal-posé de la FWI implique que le problème de minimisation (3.21) ne possède pas de solution unique. En d'autres termes, les données peuvent être ajustées de la même manière par des modèles de paramètres différents. Pour mieux contraindre l'inversion, nous pouvons imposer des contraintes sur la solution en introduisant d'un terme de régularisation. Elle permet de stabiliser et de réduire le nombre de solutions en ajoutant de l'information *a priori* sur la solution recherchée.

La méthode de régularisation la plus fréquemment utilisée est celle de Tikhonov (Tikhonov & Arsenin, 1977). Elle consiste à ajouter à la fonction coût un terme de pénalité $\mathbf{R}(\mathbf{m})$ sur le modèle visant à pénaliser la « rugosité » du modèle de paramètre, c'est-à-dire, à forcer une reconstruction lisse du modèle de paramètres. La fonction coût régularisée devient

$$\mathcal{C}^R(\mathbf{m}) = \mathcal{C}(\mathbf{m}) + \frac{\lambda}{2} \|\mathbf{R}(\mathbf{m})\|_2^2, \quad (3.29)$$

où λ est le paramètre de régularisation. Le nouveau problème (3.29) consiste ainsi à déterminer un modèle fournissant le meilleur compromis possible entre l'ajustement des données et garantissant une solution aussi « simple » que possible. Si le paramètre de régularisation λ est trop faible, le problème (3.29) est très proche du problème original (3.21) et des instabilités peuvent être générées sous forme de reconstructions irréalistes de paramètres. Si λ est au contraire trop grand, le système (3.29) est figé et les données ne peuvent pas être ajustées.

Trouver une valeur optimale du paramètre λ est un problème compliqué qui n'a pas de solution évidente. Il existe toutefois quelques heuristiques dont la plus populaire repose sur l'utilisation de « *L-curves* » (Hansen, 1992). Cette méthode consiste à analyser le graphe $\mathcal{C}(\mathbf{m})$ en fonction de $\mathbf{R}(\mathbf{m})$ pour différentes valeurs du paramètre λ . L'idée de la méthode est que si λ est sous-optimal, la norme de l'ajustement ne change que très peu tandis que la norme du modèle est grande (puisque le problème est sous-régularisé), la pente de la L-curve est alors quasi-horizontale. Dans ce cas, on n'améliore pas l'ajustement des données et l'on complique inutilement la solution. Au contraire si λ est sur-optimal, la norme du modèle ne varie que très peu et la norme de l'ajustement est grande, la pente est alors quasi-verticale. Cette fois-ci, le modèle contient trop d'information *a priori* pour permettre d'améliorer l'ajustement. Ainsi, en échelle $\log - \log$, la courbe obtenue a généralement l'aspect d'un « L » et le paramètre λ optimal est choisi à proximité du coude formé par la courbe.

Différents types de pénalités \mathbf{R} peuvent être envisagées. Elles dépendent de l'*a priori* que l'on veut ajouter au problème. Sous sa forme standard, cet opérateur est l'identité \mathbf{I} , on amortit les paramètres ayant des valeurs élevées. Pour la FWI, ce choix n'est pas intéressant en pratique, et l'on préfère plutôt pénaliser les discontinuités présentes dans le modèle $\mathbf{R} = \nabla$ ou la rugosité du modèle $\mathbf{R} = \nabla^2$ pour privilégier des modèles lisses. L'ajout d'un terme de régularisation sur la fonction implique de modifier l'expression du gradient en conséquence sous forme de termes additifs qui sont respectivement $\nabla^2 \mathbf{m}$ et $\nabla^2 \nabla^2 \mathbf{m}$ (Castellanos, 2014).

La recherche du paramètre de régularisation est généralement une tâche fastidieuse, délicate et coûteuse. Une alternative pragmatique consiste à lisser le gradient de la fonction coût au moyen d'un filtre gaussien (Sirgue, 2003). Les longueurs de corrélation sont ainsi facilement

paramétrables et la régularité de la solution peut-être contrôlée de manière simple. Ce type de régularisation s'apparente au préconditionnement de la direction de descente par reparamétrisation (Guitten et al., 2012).

3.2.5 Paramétrisation

Pour des problèmes inverses non-linéaires, la solution du problème d'optimisation dépend de la paramétrisation choisie pour la mise-à-jour des paramètres du modèle. Les données observées ont en effet une sensibilité différente aux classes de paramètres considérées lesquels présentent alors des couplages différents selon les paramètres pris en compte (Tarantola, 1986; Köhn et al., 2012; Prieux et al., 2013a). Puisque le comportement de l'inversion dépend de la paramétrisation¹⁴, il peut être intéressant de modifier la paramétrisation en fonction de la géométrie d'acquisition ou même d'envisager des inversions successives pour différentes classes de paramètres (Tarantola, 1986).

Le passage d'une paramétrisation à une autre peut être réalisé simplement au moyen de la dérivation chaînée (Mora, 1987; Köhn et al., 2012) :

$$\nabla_{m'_i} \mathcal{C}(\mathbf{m}') = \sum_{j=1}^{N_{par}} \nabla_{m_j} \mathcal{C}(\mathbf{m}) \frac{dm_j}{dm'_i}. \quad (3.30)$$

Les gradients exprimés dans une paramétrisation \mathbf{m}' s'expriment ainsi par combinaisons linéaires des gradients calculés dans une paramétrisation de référence \mathbf{m} . Dans le cas d'un milieu élastique isotrope, la paramétrisation de référence la plus naturelle est celle qui fait directement intervenir les coefficients élastiques présents dans les équations de l'élastodynamique, c'est-à-dire, (ρ, λ, μ) . Nous détaillons le calcul de ces gradients dans la section suivante.

En sismologie, les paramétrisations les plus utilisées pour l'interprétation des images tomographiques sont celles reliées aux vitesses des ondes élastiques (ρ, V_P, V_S) et aux impédances (ρ, I_P, I_S) . Les relations entre ces différentes paramétrisations sont données par

$$\lambda = \rho(V_P^2 - 2V_S^2) = \frac{I_P^2 - 2I_S^2}{\rho} \quad \text{et} \quad \mu = \rho V_S^2 = \frac{I_S^2}{\rho}. \quad (3.31)$$

En appliquant la dérivation chaînée, nous trouvons pour la paramétrisation (ρ, V_P, V_S)

$$\begin{aligned} \nabla_{\rho^v} \mathcal{C}(\mathbf{m}') &= \nabla_{\rho} \mathcal{C}(\mathbf{m}) + (V_P^2 - 2V_S^2) \nabla_{\lambda} \mathcal{C}(\mathbf{m}) + V_S^2 \nabla_{\mu} \mathcal{C}(\mathbf{m}) \\ \nabla_{V_P} \mathcal{C}(\mathbf{m}') &= 2\rho V_P \nabla_{\lambda} \mathcal{C}(\mathbf{m}) \\ \nabla_{V_S} \mathcal{C}(\mathbf{m}') &= -4\rho V_S \nabla_{\lambda} \mathcal{C}(\mathbf{m}) + 2\rho V_S \nabla_{\mu} \mathcal{C}(\mathbf{m}) \end{aligned} \quad (3.32)$$

14. Nous discutons des effets liés au choix de la paramétrisation dans le chapitre 4

et,

$$\begin{aligned}\nabla_{\rho^i}\mathcal{C}(\mathbf{m}') &= \nabla_{\rho}\mathcal{C}(\mathbf{m}) + (2V_s^2 - V_p^2)\nabla_{\lambda}\mathcal{C}(\mathbf{m}) - V_s^2\nabla_{\mu}\mathcal{C}(\mathbf{m}) \\ \nabla_{I_p}\mathcal{C}(\mathbf{m}') &= 2V_p\nabla_{\lambda}\mathcal{C}(\mathbf{m}) \\ \nabla_{I_s}\mathcal{C}(\mathbf{m}') &= -4V_s\nabla_{\lambda}\mathcal{C}(\mathbf{m}) + 2V_s\nabla_{\mu}\mathcal{C}(\mathbf{m})\end{aligned}\quad (3.33)$$

pour la paramétrisation (ρ, I_p, I_s) .

De manière similaire, nous pouvons envisager la prise en compte d'une sous-paramétrisation pour préconditionner le processus d'inversion. Par exemple, nous pouvons considérer des paramètres adimensionnés $\mathbf{m}' = \mathbf{m}/\mathbf{m}_o$ pour équilibrer les poids respectifs des différents gradients auquel cas,

$$\nabla_{\mathbf{m}'}\mathcal{C}(\mathbf{m}') = \mathbf{m}_o\nabla_{\mathbf{m}}\mathcal{C}(\mathbf{m}). \quad (3.34)$$

Ou bien le logarithme des paramètres $\mathbf{m}' = \ln \mathbf{m}$ pour lequel,

$$\nabla_{\mathbf{m}'}\mathcal{C}(\mathbf{m}') = \mathbf{m}\nabla_{\mathbf{m}}\mathcal{C}(\mathbf{m}), \quad (3.35)$$

qui permet de préconditionner les gradients en fonction de l'amplitude locale des paramètres et donc de la profondeur.

3.3 Le gradient de la fonction coût

3.3.1 Construction par la méthode de l'état adjoint

Nous développons maintenant l'expression du gradient de la fonction coût à partir du formalisme de l'état adjoint. La méthode de l'état adjoint, initialement développé dans le cadre de la théorie du contrôle optimal (Lions, 1968), est une technique permettant d'évaluer le gradient de la fonction coût sans former la matrice de sensibilité du problème. En particulier, le calcul du gradient devient indépendant du nombre de paramètre et ne nécessite plus que deux simulations par source. Elle a été introduite dans la théorie du problème inverse par Chavent (1974) et appliquée en géophysique par de nombreux auteurs (Lailly, 1983a; Tarantola, 1984b; Tromp et al., 2005; Liu & Tromp, 2006). Des présentations approfondis de l'application de cette méthode en géophysique et en sismologie sont données par Plessix (2006b) et Fichtner et al. (2006a,b).

Avant de commencer le calcul du gradient par la méthode de l'état adjoint, nous introduisons quelques définitions mathématiques. Premièrement, nous introduisons la dérivée directionnelle (ou dérivée de Gâteaux) :

$$\nabla_{\mathbf{m}}\mathcal{C}(\mathbf{m}) \cdot \delta\mathbf{m} = \mathcal{C}(\mathbf{m} + \delta\mathbf{m}) - \mathcal{C}(\mathbf{m}), \quad (3.36)$$

et, afin de simplifier les notations, nous exprimons le produit scalaire sous forme de bra-ket

$$\langle \mathbf{u} | \mathbf{v} \rangle_{[0,T] \times \Omega} = \int_0^T \int_{\Omega} \mathbf{u}^*(\mathbf{x}, t, \mathbf{m}) \mathbf{v}(\mathbf{x}, t, \mathbf{m}) dt d\Omega \quad (3.37)$$

où $*$ est la transposé complexe.

Nous cherchons à minimiser un problème sous contrainte et commençons par définir le Lagrangien \mathcal{L} qui est une fonctionnelle augmentée de la fonction coût (3.21) pour laquelle les contraintes correspondent aux équations d'état et d'observation, c'est-à-dire dans le cas présent, les équations de l'élastodynamique ainsi que la restriction du champ d'onde aux récepteurs. Pour une source et un récepteur \mathbf{d}_{cal_r} , le Lagrangien s'exprime comme :

$$\begin{aligned} \mathcal{L}(\mathbf{m}, \mathbf{u}, \mathbf{d}_{cal_r}, \mathbf{u}^a, \mathbf{d}_{cal_r}^a) &= \langle 1 | \mathcal{C}(\mathbf{m}) \rangle \\ &\quad - \langle \mathbf{u}^a | \rho \partial_t^2 \mathbf{u} - \nabla \cdot [\mathbf{c} : \nabla \mathbf{u}] - \mathbf{f} \rangle \\ &\quad - \langle \mathbf{d}_{cal_r}^a | \mathbf{d}_{cal_r} - R_r \mathbf{u} \rangle, \end{aligned} \quad (3.38)$$

où, \mathbf{u}^a et $\mathbf{d}_{cal_r}^a$ sont des multiplicateurs de Lagrange correspondant aux champs et données ajointes respectivement. Un point selle de la fonctionnelle est trouvé lorsque les dérivées du Lagrangien \mathcal{L} par rapport aux variables d'état \mathbf{u} et \mathbf{d}_{cal_r} , aux variables adjointes \mathbf{u}^a et $\mathbf{d}_{cal_r}^a$ et aux paramètres du modèle \mathbf{m} sont annulées.

Équations d'état

En annulant les dérivées directionnelles du Lagrangien par rapport aux variables adjointes, nous retrouvons directement les expressions de l'équation d'état :

$$\nabla_{\mathbf{u}^a} \mathcal{L} \cdot \delta \mathbf{u}^a = - \langle \delta \mathbf{u}^a | \rho \partial_t^2 \mathbf{u} - \nabla \cdot [\mathbf{c} : \nabla \mathbf{u}] - \mathbf{f} \rangle = 0, \quad (3.39)$$

et de l'équation d'observation :

$$\nabla_{\mathbf{d}_{cal_r}^a} \mathcal{L} \cdot \delta \mathbf{d}_{cal_r}^a = - \langle \delta \mathbf{d}_{cal_r}^a | \mathbf{d}_{cal_r} - R_r \mathbf{u} \rangle = 0, \quad (3.40)$$

exprimées sur le domaine $[0, T] \times \partial\Omega$.

$$\boxed{\begin{cases} \int_0^T \int_{\Omega} \rho \partial_t^2 \mathbf{u}(\mathbf{x}, t) - \nabla \cdot [\mathbf{c} : \nabla \mathbf{u}(\mathbf{x}, t)] - \mathbf{f}(\mathbf{x}_s, t) \, d\mathbf{x}^3 \, dt = 0, \\ \mathbf{u}(\mathbf{x}, 0) = 0, \\ \partial_t \mathbf{u}(\mathbf{x}, 0) = 0, \\ (\mathbf{c} : \nabla \mathbf{u}(\partial\Omega, t)) \cdot \hat{\mathbf{n}} = 0. \end{cases}} \quad (3.41)$$

Équations adjointes

Au contraire, l'annulation des dérivées directionnelles par rapport aux variables d'observation et d'état fournit l'équation satisfaite par le champ adjoint. Nous obtenons :

$$\nabla_{\mathbf{d}_{cal_r}} \mathcal{L} \cdot \delta \mathbf{d}_{cal_r} = \langle \delta \mathbf{d}_{cal_r} | \Delta \mathbf{d} \rangle - \langle \delta \mathbf{d}_{cal_r} | \mathbf{d}_{cal_r}^a \rangle = 0, \quad (3.42)$$

et

$$\nabla_{\mathbf{u}} \mathcal{L} \cdot \delta \mathbf{u} = - \left\langle \mathbf{u}^a \left| \rho \partial_t^2 \delta \mathbf{u} - \nabla \cdot [\mathbf{c} : \nabla \delta \mathbf{u}] \right. \right\rangle + \langle \mathbf{u}^a | R_r \delta \mathbf{u} \rangle = 0. \quad (3.43)$$

En notant que :

- en vertu de la symétrie du tenseur de raideur : $\mathbf{c} = \mathbf{c}^T$,
- l'opérateur adjoint de la divergence est le gradient : $(\nabla \cdot \mathbf{u})^T = \nabla \mathbf{u}$,
- une intégration par partie en temps est donnée par : $\langle \mathbf{u} | \partial_t \mathbf{v} \rangle = \mathbf{u} \mathbf{v} |_{[0, T]} - \langle \partial_t \mathbf{u} | \mathbf{v} \rangle$,
- une intégration par partie en espace est donnée par : $\langle \mathbf{u} | \nabla \mathbf{v} \rangle = \mathbf{u} \mathbf{v} |_{\partial \Omega} - \langle \nabla \cdot \mathbf{u} | \mathbf{v} \rangle$,

nous pouvons appliquer une double intégration temporelle par partie ainsi qu'une double intégration spatiale par partie sur l'équation (3.43), ce faisant nous aboutissons à :

$$\begin{aligned} \nabla_{\mathbf{u}} \mathcal{L} \cdot \delta \mathbf{u} = & - \left\langle \rho \partial_t^2 \mathbf{u}^a - \nabla \cdot [\mathbf{c} : \nabla \mathbf{u}^a] \right| \delta \mathbf{u} \rangle + \left\langle R_r^T \mathbf{d}_{cal_r}^a \right| \delta \mathbf{u} \rangle \\ & - (\mathbf{u}^a)(\partial_t \delta \mathbf{u})|_0^T + (\mathbf{u}^a)(\delta \mathbf{u})|_0^T \\ & - (\mathbf{u}^a)(\mathbf{c} : \nabla \delta \mathbf{u})|_{\partial \Omega} + (\mathbf{c} : \nabla \mathbf{u}^a)(\delta \mathbf{u})|_{\partial \Omega} = 0. \end{aligned} \quad (3.44)$$

Les conditions aux limites sur le champ d'onde \mathbf{u} imposent par linéarité de l'équation d'onde :

1. une condition initiale, $\delta \mathbf{u}(\mathbf{x}, t) = 0$ et $\partial_t \delta \mathbf{u}(\mathbf{x}, t) = 0$ sur les perturbations des champs de déplacement et de vitesse,
2. une condition de surface libre sur la frontière du domaine, $(\mathbf{c} : \nabla \delta \mathbf{u}(\partial \Omega, t)) \cdot \hat{\mathbf{n}} = 0$.

De même, nous imposons des conditions aux limites pour la variable adjointe en définissant :

1. une condition finale pour les champs de déplacement et de vitesse adjointes, $\mathbf{u}^a(\mathbf{x}, T) = 0$ et $\partial_t \mathbf{u}^a(\mathbf{x}, T) = 0$,
2. une condition de surface libre sur la frontière du domaine, $(\mathbf{c} : \nabla \mathbf{u}^a(\partial \Omega, t)) \cdot \hat{\mathbf{n}} = 0$.

Nous obtenons finalement,

$$\begin{aligned} \nabla_{\mathbf{u}} \mathcal{L} \cdot \delta \mathbf{u} = & - \left\langle \rho \partial_t^2 \mathbf{u}^a - \nabla \cdot [\mathbf{c} : \nabla \mathbf{u}^a] \right| \delta \mathbf{u} \rangle + \left\langle R_r^T \Delta \mathbf{d} \right| \delta \mathbf{u} \rangle \\ & - \left\langle \rho \partial_t^2 \mathbf{u}^a - \nabla \cdot [\mathbf{c} : \nabla \mathbf{u}^a] - R_r^T \Delta \mathbf{d} \right| \delta \mathbf{u} \rangle = 0 \end{aligned} \quad (3.45)$$

qui correspond à l'équation du problème adjoint. Nous remarquons que cette équation a la même forme que l'équation d'état. Il s'agit, en effet, de l'équation d'onde satisfaite par le champ adjoint dont la source est, à présent, le résidu sur les données extraites aux positions des capteurs : $\mathbf{f}^a = R_r \Delta \mathbf{d}$. Cette équation est elle aussi définie sur le domaine $[0, T] \times \partial \Omega$, mais satisfait une condition finale (en $t = T$). Le champ adjoint est alors obtenu en résolvant le système :

$$\boxed{\begin{cases} \int_0^T \int_{\Omega} \rho \partial_t^2 \mathbf{u}^a(\mathbf{x}, t) - \nabla \cdot [\mathbf{c} : \nabla \mathbf{u}^a(\mathbf{x}, t)] - \mathbf{f}^a(\mathbf{x}_r, t) \, d\Omega dt = 0, \\ \mathbf{u}^a(\mathbf{x}, T) = 0, \\ \partial_t \mathbf{u}^a(\mathbf{x}, T) = 0, \\ (\mathbf{c} : \nabla \mathbf{u}^a(\partial \Omega, t)) \cdot \hat{\mathbf{n}} = 0. \end{cases}} \quad (3.46)$$

Pour réaliser numériquement cette intégration, il est nécessaire de transformer le problème de sorte qu'il satisfasse une condition initiale. Pour cela, nous appliquons un changement de variable tel que $t' = T - t$, et introduisons le champ d'onde des résidus rétropropagés $\boldsymbol{\lambda}(\mathbf{x}, t') = \mathbf{u}^a(\mathbf{x}, T - t')$. Il correspond au champ adjoint « renversé en temps », c'est à dire, propagé en sens contraire du temps :

$$\left\{ \begin{array}{l} \int_0^T \int_{\Omega} \rho \partial_t'^2 \boldsymbol{\lambda}(\mathbf{x}, t') - \nabla \cdot [\mathbf{c} : \nabla \boldsymbol{\lambda}(\mathbf{x}, t')] - \mathbf{f}^a(\mathbf{x}_r, T - t') \, d\Omega dt' = 0, \\ \boldsymbol{\lambda}(\mathbf{x}, 0) = 0, \\ \partial_t \boldsymbol{\lambda}(\mathbf{x}, 0) = 0, \\ (\mathbf{c} : \nabla \boldsymbol{\lambda}(\partial\Omega, t')) \cdot \hat{\mathbf{n}} = 0. \end{array} \right. \quad (3.47)$$

Le problème est maintenant exactement similaire au problème permettant de déterminer le champ d'état (3.41) dont le terme source correspond aux résidus renversés en temps.

Gradient

Finalement, l'expression du gradient est obtenu en annulant les dérivées du Lagrangien par rapport aux paramètres du modèle

$$\nabla_{\mathbf{m}} \mathcal{L} \cdot \delta \mathbf{m} = \nabla_{\mathbf{m}} \mathcal{C} \cdot \delta \mathbf{m} = - \langle \mathbf{u}^a | \delta \rho \partial_t'^2 \mathbf{u} - \nabla \cdot [\delta \mathbf{c} : \nabla \mathbf{u}] - \mathbf{f} \rangle = 0. \quad (3.48)$$

Les gradients respectifs à la densité et aux paramètres élastiques s'expriment :

$$\nabla_{\rho} \mathcal{L} \cdot \delta \rho = \nabla_{\rho} \mathcal{C} \cdot \delta \rho - \langle \mathbf{u}^a | \delta \rho \partial_t'^2 \mathbf{u} \rangle = 0, \quad (3.49)$$

et

$$\begin{aligned} \nabla_{\mathbf{c}} \mathcal{L} \cdot \delta \mathbf{c} &= \nabla_{\mathbf{c}} \mathcal{C} \cdot \delta \mathbf{c} + \langle \mathbf{u}^a | \nabla \cdot [\delta \mathbf{c} : \nabla \mathbf{u}] \rangle \\ &= \nabla_{\mathbf{c}} \mathcal{C} \cdot \delta \mathbf{c} - \langle \nabla \mathbf{u}^a | : \delta \mathbf{c} : \nabla \mathbf{u} \rangle = 0. \end{aligned} \quad (3.50)$$

Le gradient de la fonction coût peut alors s'exprimer en fonction des champs incident et adjoint

$$\nabla_{\mathbf{m}} \mathcal{C} = \begin{cases} \nabla_{\rho} \mathcal{C} = \int_0^T \int_{\Omega} \mathbf{u}^a(\mathbf{x}, t) \cdot \partial_t'^2 \mathbf{u}(\mathbf{x}, t) \, d\mathbf{x}^3 dt, \\ \nabla_{\mathbf{c}} \mathcal{C} = \int_0^T \int_{\Omega} \nabla \mathbf{u}^a(\mathbf{x}, t) :: \nabla \mathbf{u}(\mathbf{x}, t) \, d\mathbf{x}^3 dt, \end{cases} \quad (3.51)$$

ou en fonction du champ incident et du champs adjoint rétropropagé :

$$\nabla_{\mathbf{m}} \mathcal{C} = \begin{cases} \nabla_{\rho} \mathcal{C} = \int_0^T \int_{\Omega} \boldsymbol{\lambda}(\mathbf{x}, T - t) \cdot \partial_t'^2 \mathbf{u}(\mathbf{x}, t) \, d\mathbf{x}^3 dt, \\ \nabla_{\mathbf{c}} \mathcal{C} = \int_0^T \int_{\Omega} \nabla \boldsymbol{\lambda}(\mathbf{x}, T - t) :: \nabla \mathbf{u}(\mathbf{x}, t) \, d\mathbf{x}^3 dt. \end{cases} \quad (3.52)$$

Les équation (3.51) et (3.52) indiquent que le gradient de la fonction coût est construit par corrélation à décalage nul entre les champs incident et adjoint. Plus spécifiquement, le gradient par rapport à la densité est obtenu en corrélant les accélérations particulières du champ incident

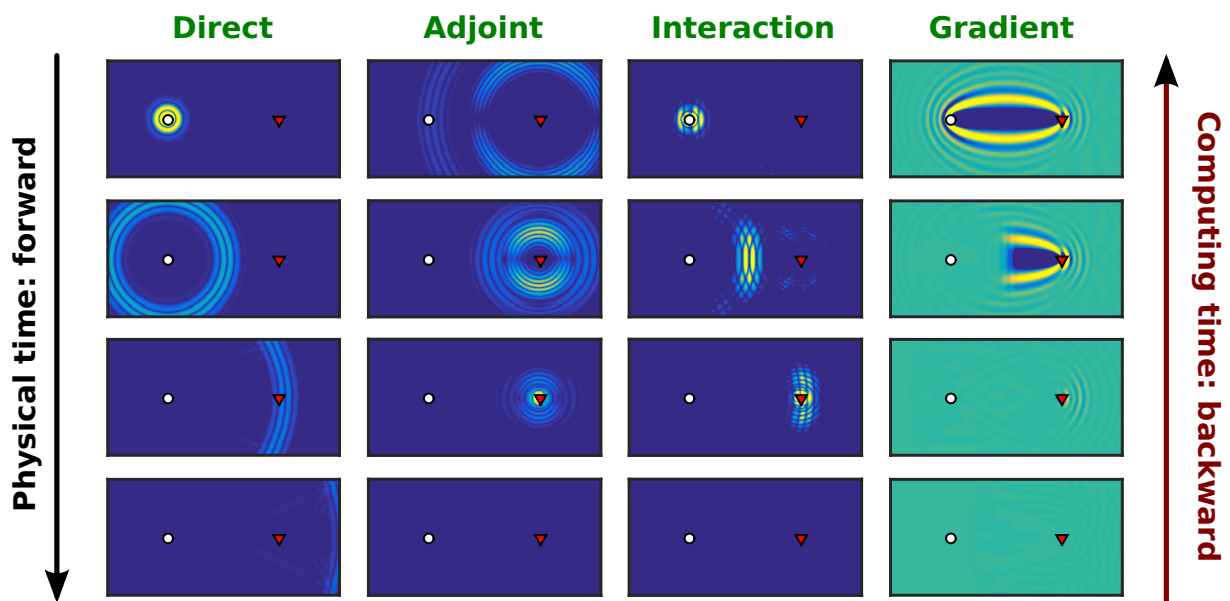


Figure 3.5 – Construction du gradient par la méthode de l'état adjoint pour une source explosive. De gauche à droite sont présentés le champ incident, le champ adjoint, l'interaction entre les deux champs, et le gradient (λ) construit par sommation au cours du temps des interactions. La source et le récepteur sont respectivement matérialisés par le rond blanc et le triangle rouge.

et les déplacements particuliers du champ adjoint, tandis que le gradient par rapport aux paramètres élastiques est construit en corrélant les champs de déformations incident et adjoints.

La figure 3.5 présente la construction du gradient par la méthode de l'état adjoint pour un couple source-récepteur dans un milieu homogène infini. Ce milieu est perturbé par une inclusion sphérique localisée à égale distance de la source et du récepteur. Les première et deuxième colonnes présentent les propagations respectives des champ incident (depuis la source) et adjoint (depuis le récepteur) dans le milieu homogène. La corrélation à décalage nul est effectuée en sommant la contribution des interactions entre les deux champs à chaque pas de temps (les troisième et quatrième colonne). La figure du gradient résultante présente des franges d'interférences qui correspondent à des isochrones elliptiques. Ces ellipses indiquent la localisation des points diffractants qui permettent de générer le résidu observé dans les données. Ces franges sont dominées par l'ellipse centrale qui correspond à la première zone de Fresnel (transmission), les franges externes impliquent des phases générées par des réflexions ou conversions.

3.3.2 Implémentation et validation

Nous avons que le calcul du gradient de la fonction coût par la méthode de l'état adjoint nécessite d'effectuer deux simulations différentes : une pour les champs d'ondes d'état satisfaisant une condition initiale et une seconde pour les champs d'ondes adjoints satisfaisant une condition finale. Le gradient est lui-même obtenu en sommant, à chaque temps t , l'interaction entre ces deux champs. Numériquement, ceci peut être réalisé de quatre façons différentes.

1. Réaliser les deux modélisations et les sauver en mémoire pour tous les temps de la simulation. Cette méthode ne nécessite que deux simulations mais est en revanche très coûteuse en mémoire.
2. Réaliser la simulation pour le champ adjoint et relancer, à chacun de ses pas de temps, la simulation du champs d'état pour obtenir le champ d'onde au même temps. Cette fois-ci aucun stockage mémoire n'est requis mais le nombre de modélisations devient prohibitif.
3. Calculer le gradient par rétropropagation au moyen de trois simulations. Une première pour obtenir l'instantané du champs d'état au pas de temps final ainsi que son historique sur les bords du domaine de calcul. Puis deux modélisations pour rétro-propager les champs adjoint et d'états simultanément (Liu & Tromp, 2006). Cette méthode ne nécessite que trois modélisations mais ne permet pas de considérer des milieux atténuants pour lesquels la rétropropagation du champ incident est instable.
4. Utiliser une stratégie de check-pointing (Symes, 2007). Il s'agit d'un compromis entre les méthodes 1 et 2 pour laquelle quelques instantanés du champs d'état sont sauvés en mémoire. Le champs d'état est alors propagé par temps croissants depuis un instantané jusqu'au temps du champ adjoint. La méthode est ainsi viable quelle que soit la physique du milieu considéré (atténuant ou non). Le nombre de modélisations requises dépend alors du nombre d'instantanés considérés. Pour un nombre suffisant, cette méthode est aussi coûteuse que la méthode 3.

Dans ce travail, nous ne considérons pas les effets d'atténuation, nous choisissons donc d'implémenter la méthode numéro 3 qui présente le meilleur compromis nombre de simulation / mémoire et qui est plus simple à implémenter.

Pour ce faire, nous propageons le champs d'état dans le sens croissant du temps et nous sauvons sur disque l'histoire du champ d'onde sur les bords de la cible lithosphérique. Une fois la simulation terminée, nous calculons les résidus ainsi que le fonction coût. L'instantané final étant toujours en mémoire, nous itérons dans le sens décroissant du temps. Pour chaque pas de temps, nous rétro-propageons le champ d'état et propageons le champ adjoint. Nous calculons l'interaction entre ces deux champs et nous sommions ces contributions sur tous les pas de temps.

Le gradient étant l'ingrédient essentiel du processus de minimisation de la FWI, nous validons à présent son implémentation¹⁵. Pour ce faire, nous comparons dans un cas simple le gradient calculé par la méthode de l'état adjoint au gradient estimé par différences finies de la fonction coût :

$$\nabla_{\mathbf{m}_i} \mathcal{C} \approx \frac{\mathcal{C}(\mathbf{m} + \Delta \mathbf{m}_i) - \mathcal{C}(\mathbf{m})}{\Delta \mathbf{m}_i} \quad (3.53)$$

où $\Delta \mathbf{m}_i$ correspond à la perturbation d'un unique paramètre à une position donnée du modèle. Le gradient est ainsi obtenu en n'effectuant que des modélisations : une dans le modèle initial permettant d'obtenir $\mathcal{C}(\mathbf{m})$ et n_p simulations (une pour chaque paramètre du milieu). Cette méthode de calcul est donc très coûteuse en temps de calcul. En outre, l'implémentation est légèrement problématique puisque la précision de l'approximation aux différences finies dépend de la perturbation $\Delta \mathbf{m}$. Plus elle est faible et plus l'approximation est sensée être précise,

15. Il est aussi possible de vérifier la validité de la modélisation du champ adjoint par le test du produit scalaire (Kern, 2002).

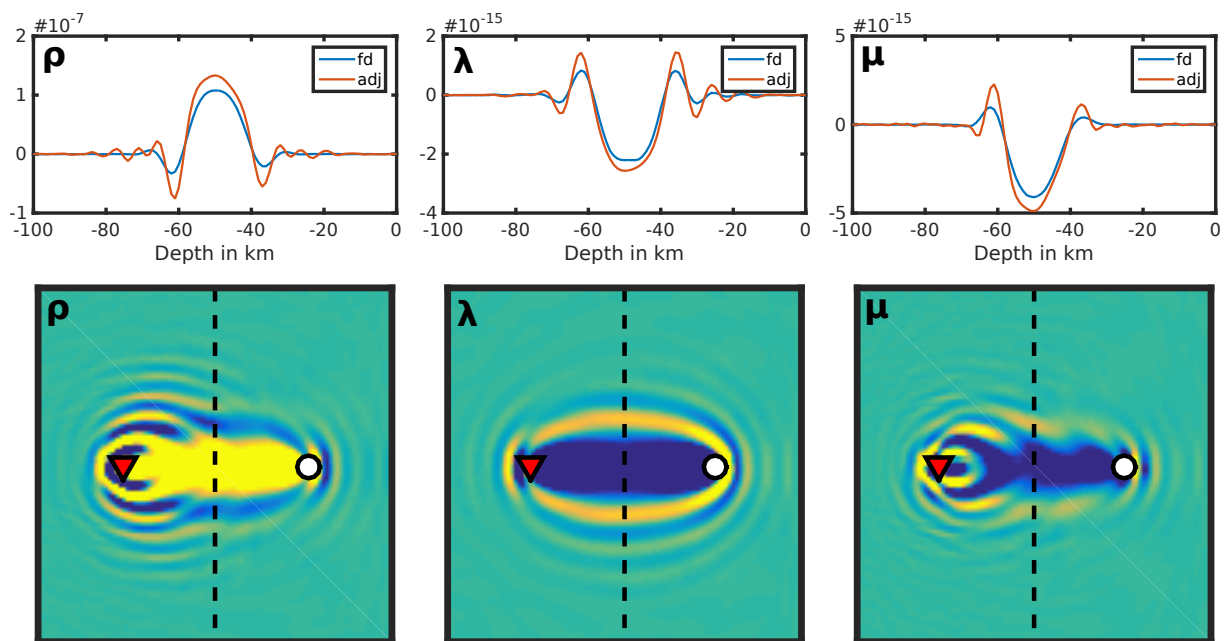


Figure 3.6 – Validation du gradient de référence – paramétrisation (ρ, λ, μ) – construit par l'état adjoint (courbes rouge). À titre de comparaison, le gradient approximé par différences-finies le long d'un profil vertical (courbes bleue). En haut, la comparaison des deux gradients le long du profil pour les trois paramètres. En bas, profil vertical des gradients respectifs obtenus par la méthode de l'état adjoint. La source et le récepteur sont respectivement matérialisés par un cercle blanc et un triangle rouge.

néanmoins, il faut qu'elle soit assez importante pour générer une variation significative de la fonction coût, c'est-à-dire supérieure à la précision numérique machine pour éviter les erreurs d'arrondis.

Nous considérons un milieu infini homogène de densité $\rho = 3 \text{ g/cm}^3$, de vitesses $V_P = 7 \text{ km/s}$ et $V_S = 3 \text{ km/s}$. Nous générons des données synthétiques en perturbant le milieu avec une petite inclusion sphérique au centre du modèle dont les perturbations sont égales à 10 % des paramètres du modèle homogène. La source est une explosion ponctuelle. Les gradients obtenus par la méthode de l'état adjoint et par différences finies sont présentés sur la figure 3.6. Comme l'estimation du gradient par différences finies est assez coûteuse en temps calcul, seuls les gradients le long d'un profil 1D sont présentés (ce dernier est matérialisé par le tireté noir sur la figure 3.6). Nous voyons que la forme générale, la polarité ainsi que l'amplitude des gradients par rapport à ρ , V_P et V_S sont en bon accord. Nous imputons les légères différences en amplitude aux erreurs d'arrondis numériques et concluons que l'implémentation du gradient est validée.

3.3.3 Anatomie d'un gradient télésismique

Les spécificités de la configuration télésismique font que les champs d'ondes incidents entrant dans la cible lithosphérique correspondent approximativement à des champs d'ondes plans. L'interaction d'une onde P plane avec la surface libre génère, comme nous l'avons vu, des ondes réfléchies et converties PP et PS. Au vu de cette géométrie, il est important d'analyser l'anatomie générale ainsi que les contributions particulières des ondes montantes et descendantes pour la construction du gradient télésismique. Ce dernier correspondant aux perturbations de premier ordre que la FWI apporte au modèle initial, la connaissance de ses constituants permet de faire le lien entre la résolution escomptée pour les différents paramètres mis à jour.

À ce titre, nous considérons un demi-espace homogène de densité $\rho = 3 \text{ g/cm}^3$ et de vitesses sismiques $V_P = 7 \text{ km/s}$ et $V_S = 4 \text{ km/s}$. Nous générons des données synthétiques en insérant une inclusion sphérique au centre du modèle dont les perturbations sont égales à +10 % et en considérant une onde plane d'incidence égale à 45 degrés.

La figure 3.7 présente, de manière similaire à la figure 3.5, la construction du gradient par la méthode de l'état adjoint dans une configuration télésismique. La présence de la surface libre ainsi que l'interaction des multiples PP et PS avec l'inclusion, génère un motif de franges d'interférences plus complexe que celui rencontré pour une configuration en point source. Cette complexité est typique de la configuration lithosphérique et présente :

- des franges d'interférence sub-verticales dominantes sous forme de « demi-banane » (comme celles rencontrées en tomographie à fréquences finies (Dahlen et al., 2000a; Dahlen & Baig, 2002)) qui sont dues à la transmission de l'onde P directe,
- des franges d'interférences sub-horizontales semblables aux ellipses de réflexion rencontrées en migration (Claerbout, 1985), qui proviennent des champ secondaires PP et PS générés par la surface libre.

Les perturbations à apporter au modèle initial pour converger vers le modèle optimal sont obtenues en sommant la contribution des gradients pour différents couples source-récepteur.

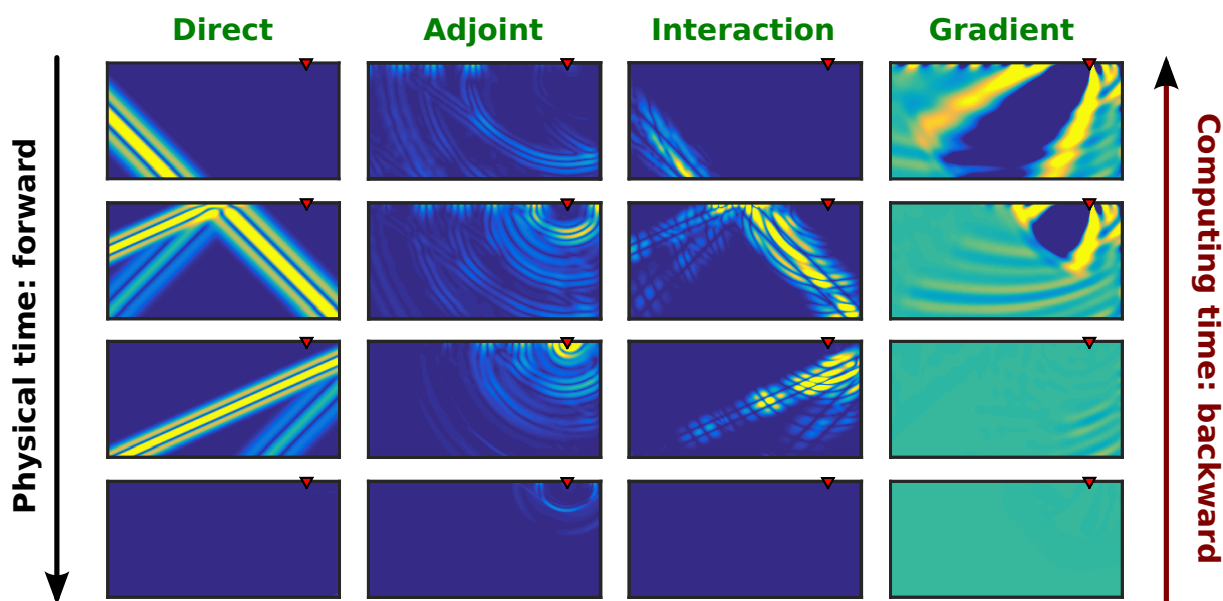


Figure 3.7 – Construction du gradient par la méthode de l'état adjoint pour une source sismique. De gauche à droite sont présentés le champ direct, le champ adjoint, l'interaction entre les deux champs, et le gradient construit par sommation au cours du temps des interactions. Le récepteur est matérialisé un triangle rouge. Le gradient présenté correspond au paramètre λ .

Cette sommation est présentée sur la figure 3.8. La première ligne présente les gradients dans la paramétrisation (ρ, V_P, V_S) pour un couple source-récepteur (la configuration est la même que pour la figure 3.7). Le calcul du gradient pour la même source, mais un nombre plus important de récepteurs (ici, une nappe 2D constituées de 484 récepteurs) est présenté sur la seconde ligne. Il permet de focaliser le gradient de la fonction coût sur la position de l'hétérogénéité par interférences constructives et destructives des contributions de chaque récepteur. La sommation des gradients calculés pour plusieurs sources, présentée sur la dernière ligne, améliore la focalisation des perturbations au niveau de l'hétérogénéité et contribue à réduire l'effet des franges parasites (les sourires de migration visibles sur les premières et deuxièmes lignes).

Puisque le champ incident est une onde plane analytique, nous pouvons séparer¹⁶ la contribution de l'onde montante (P directe) et des ondes descendantes (PP et SP) et analyser les apports respectifs de chacune des phases sismiques au gradient de la fonction coût (Fig. 3.9). Pour l'onde P directe, nous retrouvons la forme caractéristique de « banane » générée par la transmission P-P (l'ellipse de migration correspondante est indiquée par une courbe bleue). Les gradients par rapport à la densité et la vitesse des ondes S présentent une forme caractéristique correspondant à l'ellipse de migration de la conversion P-S en transmission (courbe verte). Celle-ci n'est pas visible sur le gradient de V_P puisqu'une hétérogénéité de vitesse P ne peut induire que des diffractions en mode P-P (Wu & Aki, 1985; Tarantola, 1986). Lorsqu'on ne considère que les ondes descendantes (troisième et quatrième lignes de la figure 3.9), Les franges d'interférences sont toutes approximativement horizontales et correspondent aux

16. Cette séparation est effectuée en enlevant la condition de surface libre et en injectant séparément les solutions analytiques des ondes montantes et descendantes.

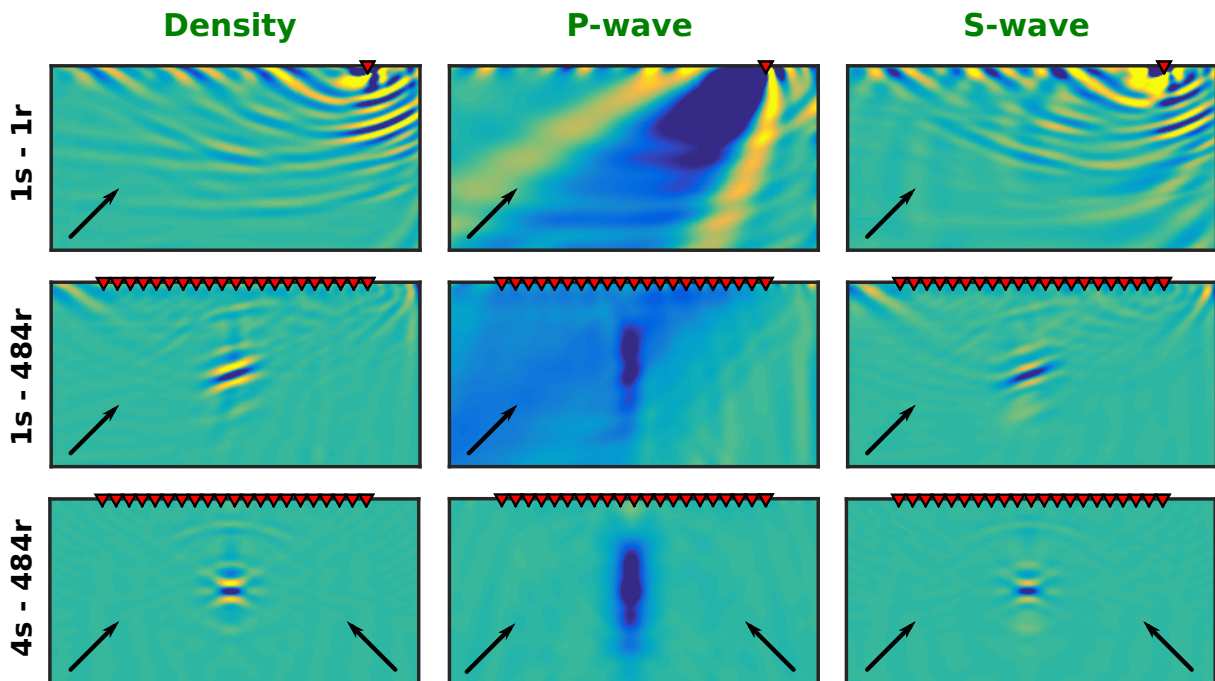


Figure 3.8 – Construction du gradient par la méthode de l'état adjoint. De gauche à droites sont présentés les gradients dans la paramétrisation (ρ, V_p, V_s) . De haut en bas : gradient obtenu pour un couple source-récepteur, gradient obtenu pour une source et une nappe de récepteurs en surface et enfin, gradient obtenus pour 4 sources et une nappe de récepteurs. La sommation sur les sources et les récepteurs focalise les gradients à la position de l'hétérogénéité (par interférence constructives et destructives). Ce résultat donne la perturbation à apporter au milieu pour décroître la fonction coût et améliorer le modèle du sous-sol.

ellipses de migration en réflexion (PPP, PPS, PSP, PSS). On observe l'apparition de franges multiples ne correspondant pas aux ellipses théoriques. Ces fantômes sont en fait générés par la mauvaise interprétation du gradient : les phases générées en transmission (PP et PS) sont alors interprétées comme des diffraction en réflexion (PPP, PPS, PSP, PSS) et réciproquement. La contributions de plusieurs sources et plusieurs récepteurs permet d'atténuer par interférences destructives ces artefacts.

Notons que les gradients par rapport à la densité et à la vitesse des ondes S sont dominés par les phases PS et PSS. Ce sont, comme nous l'avons vu dans la section 1.4.2, les phases les plus énergétiques. La correspondance entre les différentes phases télésismiques et les données synthétiques utilisées pour le calcul du gradient est présentée sur la figure 3.10 Leur importance est donc prépondérante dans le gradient télésismique. Au contraire, le gradient par rapport à la vitesse des ondes P est dominé par la transmission PP qui est hautement plus énergétique que les multiples PPP.

Il est aussi intéressant d'analyser l'anatomie du gradient en fonction de la paramétrisation élastique choisie (ρ, λ, μ) , (ρ, V_P, V_S) ou (ρ, I_P, I_S) . Cette analyse n'est pas détaillée ici mais sera présentée dans le chapitre 4.

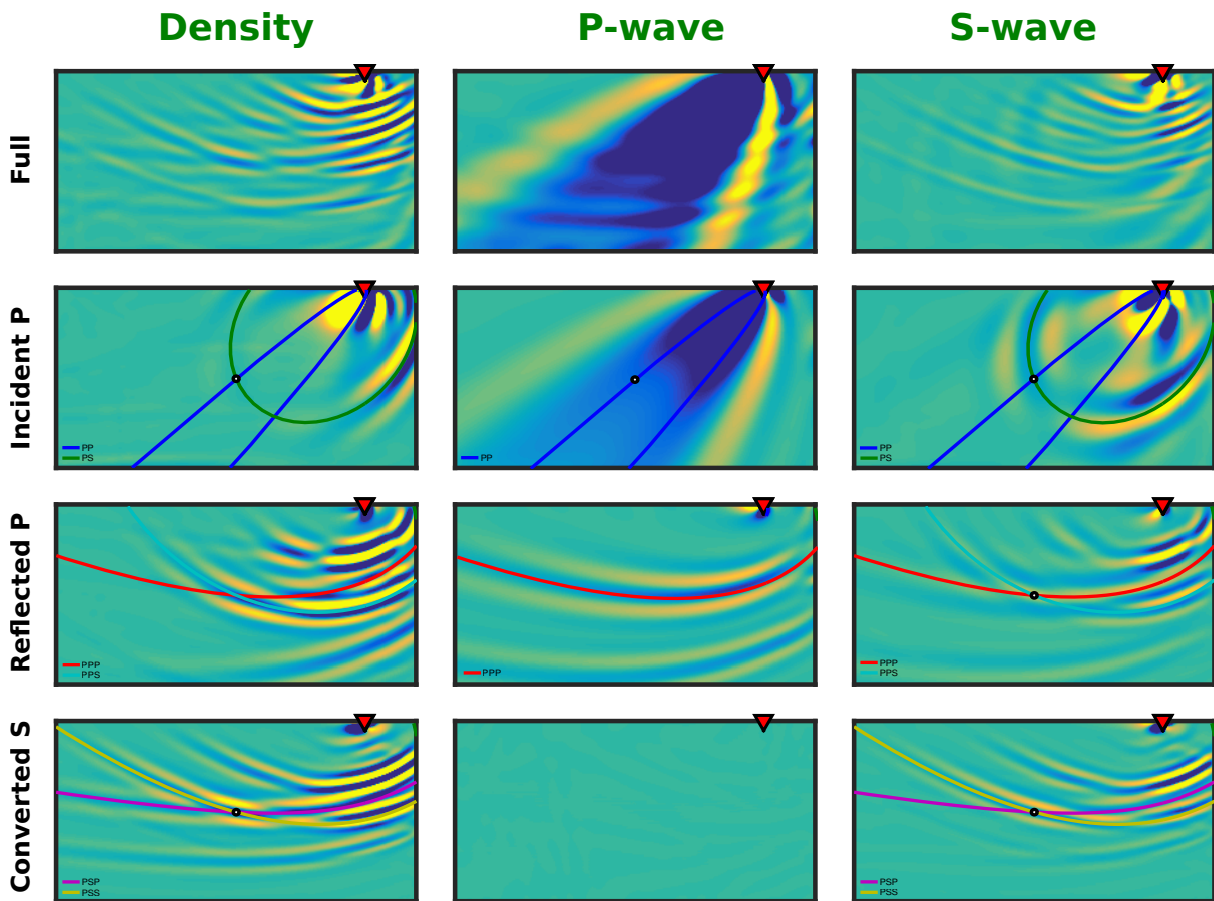


Figure 3.9 – Anatomie d'un gradient télésismique pour un seul couple source-récepteur. De gauche à droite sont présentés les gradients par rapport à ρ , V_p et V_s . De haut en bas : le gradient total, la partie uniquement dues à l'onde P montante (transmission), la partie due à l'onde P réfléchi (descendante) et la partie due à l'onde S convertie (descendante). L'hétérogénéité du milieu est représentée par un cercle blanc et le récepteur par un triangle rouge. Sur les gradients sont superposés les ellipses de migrations correspondant aux arrivées théoriques des différentes diffractions générées par l'hétérogénéité.

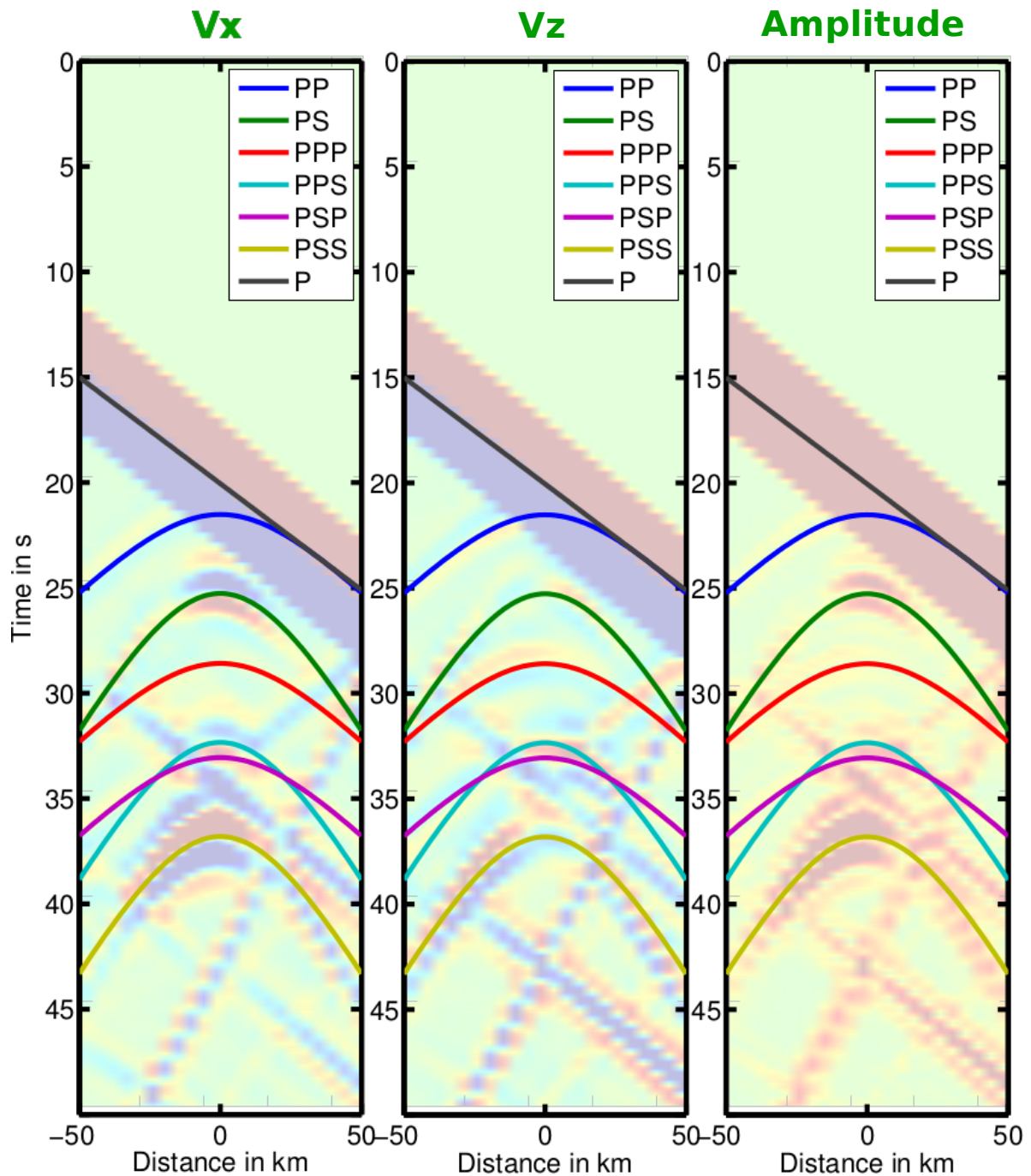


Figure 3.10 – Anatomie des données synthétiques utilisées pour le calcul du gradient téléseismique (propagation d'une onde plane dans un milieu présentant une inclusion sphérique). On remarque que les diffractions les plus énergétiques sont celles impliquant des phases en ondes S, notamment les phases PS (vert) et PSS (jaune).

3.4 Algorithme de la FWI

3.4.1 Implémentation

Notre programme d'inversion de formes d'ondes est implémenté sous la forme d'un problème de minimisation local. Ce problème est résolu au moyen d'une librairie d'optimisation développée par L. Métivier au sein du consortium SEISCOPE (Métivier & Brossier, 2016a). Cette dernière contient plusieurs algorithmes de calcul de direction de descente : la méthode de plus grande pente, la méthode de gradient conjugué, l-BFGS ainsi que celle de Newton tronqué. En outre, cette librairie implémente directement la recherche linéaire d'un pas de descente satisfaisant les conditions de Wolfe. Les seuls pré-requis de cette librairie sont les valeurs de la fonction coût ainsi que les gradients.

Le processus de l'inversion de formes d'ondes est présenté sur la figure 3.11. En amont de la FWI, les étapes de maillage et de calculs des champs d'ondes téléseismiques avec AxiSEM doivent être réalisées de la manière présentée dans la première partie de cette thèse. Au besoin, la signature de la source téléseismique doit être estimée et convoluée aux champs d'ondes incidents (nous discutons de cette estimation dans la section 3.4.3). Le processus d'optimisation est ensuite effectué pour des données filtrées dans une certaine bande de fréquence. Pour chaque itération, une première modélisation est effectuée pour obtenir les données modélisées. Puis le gradient de la fonction est calculé moyennant deux simulations par source comme expliqué dans la section 3.3. Dès que le critère de convergence est atteint, le processus de FWI est stoppé et le modèle sauvegardé. Si l'on souhaite effectuer une inversion hiérarchique en fréquence (Bunks et al., 1995), le processus peut être relancé en fournissant le modèle de l'inversion précédente et en considérant une bande de fréquence plus élevée.

Dans sa forme actuel, le programme FWI est capable de considérer plusieurs types de fonctions coûts et de normes. Les normes implémentées sont les normes L2 et L1. Les fonctions coûts disponibles sont quant à elles :

- la différence de formes d'ondes (comme présentée précédemment),
- les différences d'enveloppes et de phases instantanées (Bozdağ et al., 2011),
- la fonction coût indépendante de la source (Choi & Alkhalifah, 2011).

Concernant la régularisation du problème, nous avons implémenté les termes de Tikhonov pénalisant soit le gradient soit le laplacien du modèle. Il est aussi possible de simplement lisser le gradient de la fonction coût au moyen d'un filtre gaussien.

3.4.2 Exemple d'inversion synthétique

Afin de vérifier le comportement de notre algorithme d'inversion de formes d'ondes, nous réalisons quatre tests synthétiques. Les données synthétiques sont obtenues en propageant huit ondes planes dans un milieu élastique homogène de densité $\rho = 3 \text{ g/cm}^3$, de vitesses $V_P = 7 \text{ km/s}$ et $V_S = 3 \text{ km/s}$, possédant une inclusion cubique de 20 km de côté localisée à 30 km de profondeur. Dans le premier test, l'inclusion est portée par tous les paramètres avec $\rho = 3.3 \text{ g/cm}^3$, $V_P = 7.3 \text{ km/s}$ et $V_S = 3.3 \text{ km/s}$. Dans les trois tests suivants, les données sont générées à

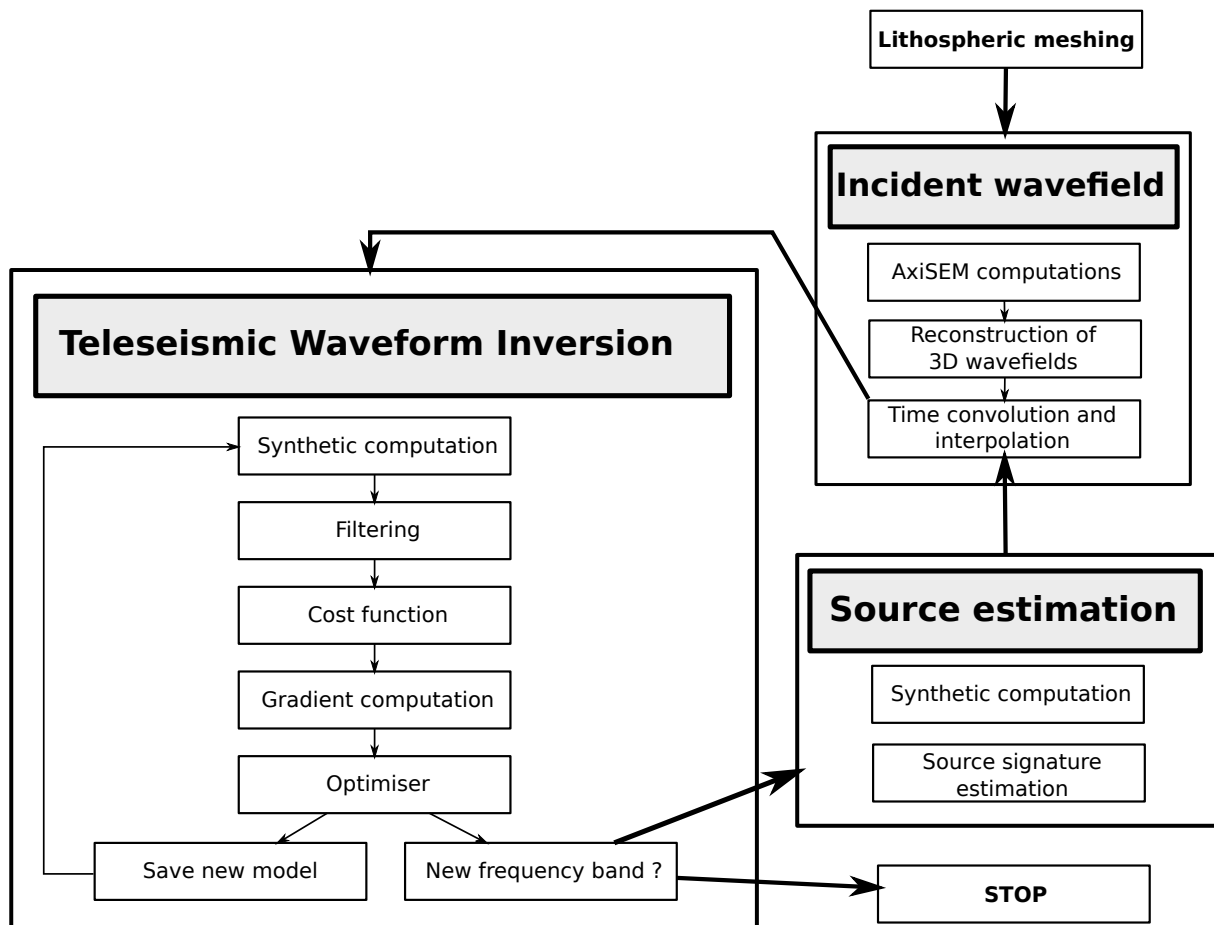


Figure 3.11 – Workflow de l'inversion de formes d'onde téléseismique.

l'aide d'une inclusion mono-paramètre portant sur la densité ou une des vitesses de propagation. Les données sont enregistrées par une nappe régulière de 361 récepteurs à trois composantes espacés de 5 km. Les huit ondes planes illuminent les quatre bords du milieu lithosphérique avec deux angles d'incidences de 20 et 45 degrés. Leur période dominante est de 4.2 secondes.

Toutes les inversions sont réalisées en partant d'un milieu homogène de densité $\rho = 3 \text{ g/cm}^3$, de vitesses $V_P = 7 \text{ km/s}$ et $V_S = 3 \text{ km/s}$. L'algorithme d'optimisation local utilisé est l-BFGS avec 10 gradients mémoires pour pouvoir prendre en compte les effets de l'inverse du Hessien. Les inversions sont réalisées de manière hiérarchique en filtrant les données dans trois bandes de fréquence dont les périodes minimales sont successivement de 20, 10 et 5s. Aucune régularisation et aucun préconditionnement n'est utilisé. Durant chaque cas test, les trois paramètres élastiques sont mis à jour. Dans les trois derniers tests (inclusion mono-paramètre), toute image fantôme de l'inclusion dans la reconstruction d'un paramètre n'ayant pas généré de résidus reflétera un couplage entre ce paramètre et celui ayant généré les résidus.

La figure 3.12 présente les résultats de l'inversion de formes d'onde pour une inclusion multi-paramètres. À l'exception du modèle V_P qui est un peu plus bruité que ses homologues V_S et ρ , les modèles de vitesse et de densité reproduisent parfaitement bien la forme de l'inclusion cubique. En outre, les profils 1D indiquent qu'en plus de la restitution de la forme de l'anomalie,

la FWI est capable de restituer fidèlement l'amplitude des perturbations de vitesses et de densité. Néanmoins, nous ne pouvons conclure quant à l'impact de la reconstruction d'un paramètre par rapport à l'autre.

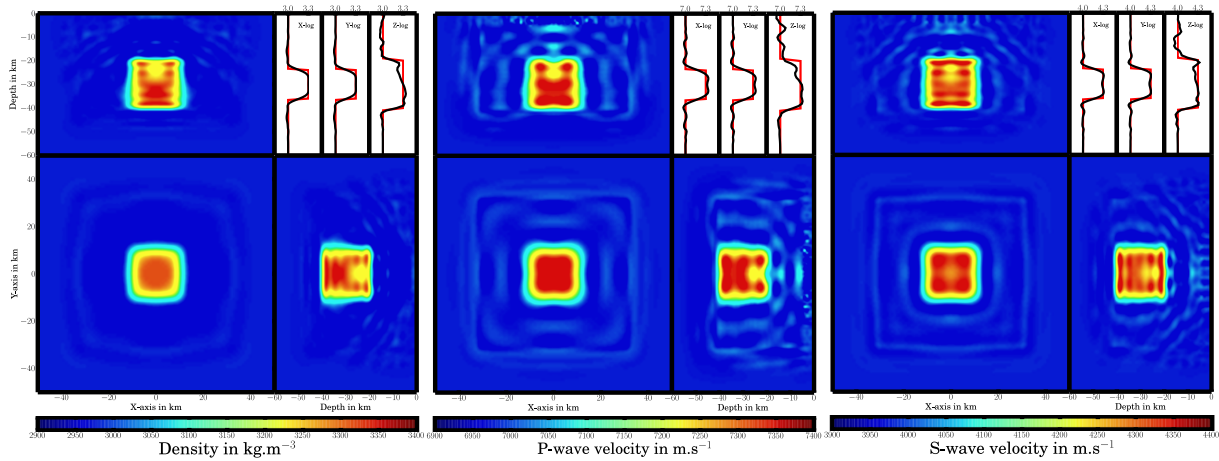


Figure 3.12 – Reconstruction multiparamètres (ρ, V_P, V_S) dans le cas d'une inclusion cubique multiparamètres (ρ, V_P, V_S). De gauche à droite, les modèles FWI ρ, V_P et V_S .

Pour estimer la capacité de la FWI à reconstruire les paramètres indépendamment et analyser les possibles couplages entre les paramètres estimés, nous présentons sur la figure 3.13 les résultats de l'inversion pour les trois inclusions monoparamètres. Ces résultats indiquent que peu de couplages sont présents et que l'inversion est théoriquement capable de discerner une perturbation de densité, d'une perturbation de vitesse V_P et/ou V_S .

3.4.3 Estimation de la source

Jusqu'à présent nous avons considéré le problème d'inversion de formes d'ondes pour la mise-à-jour du milieu lithosphérique en supposant les paramètres de la source téléseismique connus. Néanmoins, lorsqu'on s'intéresse à des applications sur données réelles les paramètres de la source sont inconnus ou mal-contraints. Il est dans ce cas nécessaire d'estimer la position, le mécanisme au foyer et surtout la signature temporelle de la source. Bien qu'il soit possible de reformuler le problème inverse pour estimer les paramètres de la source de manière conjointe aux structures (Kim et al., 2011), cela augmente considérablement la non-linéarité du problème. En pratique, la source est généralement estimée en amont de la mise à jour des structures, et peut être au besoin raffinée dans le processus de FWI (Bozdağ et al., 2016).

Une difficulté supplémentaire apparaît dans les configurations téléseismiques du fait que les sources sont localisées en dehors de la cible lithosphérique. Ré-estimer la position ou le mécanisme au foyer des événements téléseismiques impliquerait de relancer de coûteuses simulations à l'échelle globale avec AxiSEM. Il est donc inenvisageable de ré-estimer ces paramètres, et nous nous contentons des bulletins données de manière routinière par les centre internationaux tels que l'ISC, le NEIC et le globalCMT (Ekström et al., 2012). L'estimation de la source tem-

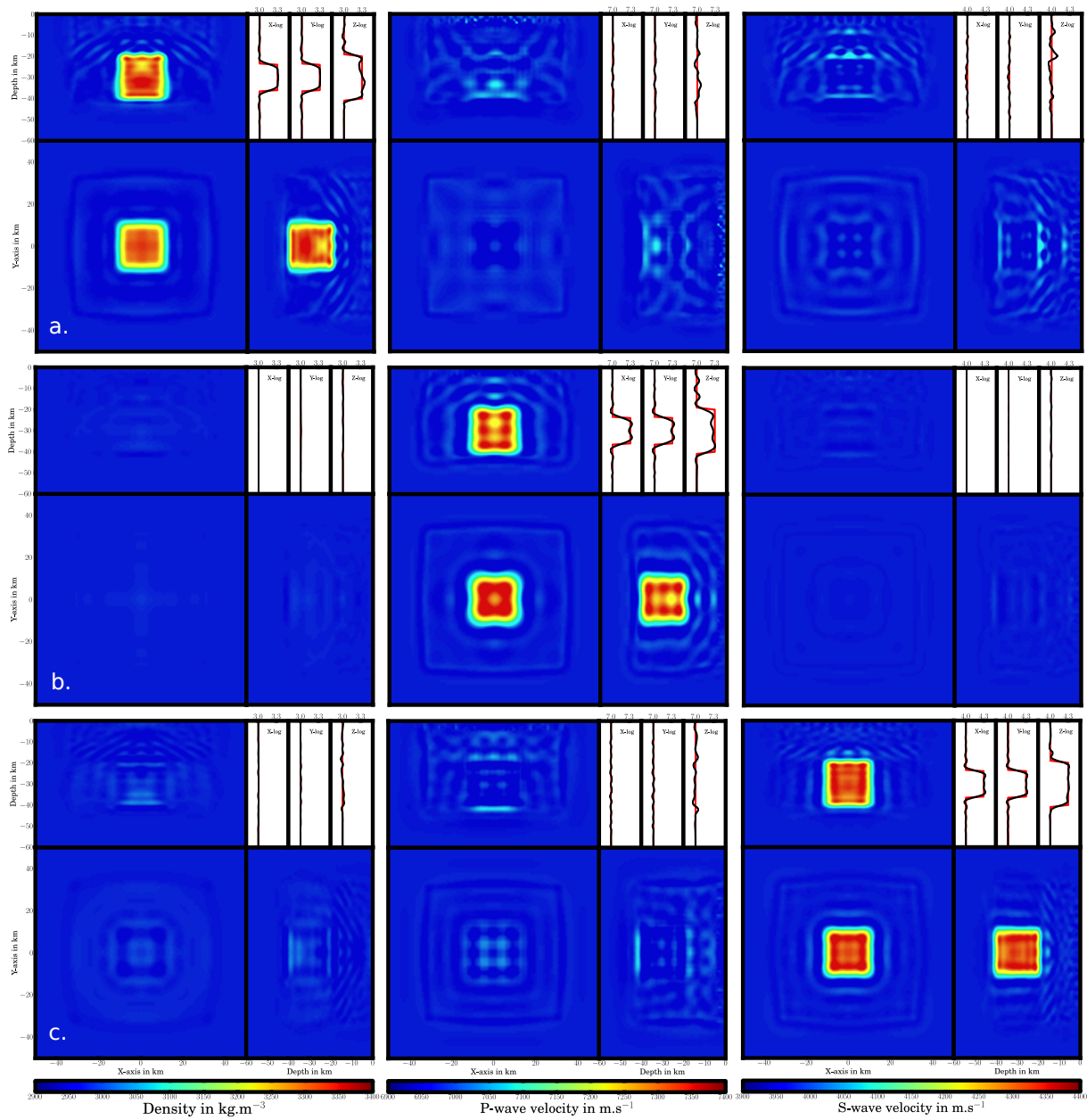


Figure 3.13 – Reconstruction multiparamètres (ρ, V_P, V_S) dans le cas d'une inclusion cubique monoparamètre. (a.) Les perturbations sont portées par le modèle de ρ , (b.) V_P et (c.) V_S . De gauche à droite, les modèles FWI ρ, V_P et V_S .

porielle¹⁷ est un problème linéaire que nous réalisons simplement par une déconvolution en domaine fréquentiel (Pratt, 1999) :

$$s(\omega) = \frac{\mathbf{d}_{cal}^*(\omega)\mathbf{d}_{obs}(\omega)}{\mathbf{d}_{cal}^*(\omega)\mathbf{d}_{cal}(\omega)}, \quad (3.54)$$

où * dénote l'opérateur complexe conjugué.

Les sources sismiques estimées doivent ensuite être convoluées avec les solutions AxiSEM en chaque point des bords de la cible lithosphérique. Cette opération peut être assez coûteuse selon le maillage considéré. Alternativement, il est possible de convoluer la source estimée avec les données calculées dans la cible lithosphérique. La fonction coût devient alors

$$\mathcal{C}(\mathbf{m}) = \frac{1}{2} \|s * \mathbf{d}_{cal} - \mathbf{d}_{obs}\|_2^2, \quad (3.55)$$

où s est la signature temporelle de la source, et pour laquelle la source adjointe devient

$$s^a = s * (\mathbf{d}_{cal} - \mathbf{d}_{obs}), \quad (3.56)$$

où * et \star sont respectivement le produit de convolution et l'intercorrelation.

D'autres alternatives intéressantes, plus aveugles, consistent à modifier la fonction coût pour la rendre insensible à la signature temporelle de la source, soit en convoluant les données observées et calculées avec une trace de référence calculée et observée (Choi & Alkhalifah, 2011; Zhang et al., 2016; Beller et al., 2015) ou bien en considérant une fonction coût basée sur des doubles-différences (Yuan et al., 2016).

Conclusions

Ce chapitre nous a permis de présenter et décrire les problèmes inverses de manière générale, leurs principales caractéristiques ainsi que les méthodes numériques permettant de les résoudre. Considérant le grand nombre de paramètres impliqués dans la FWI et sa non-linéarité, nous avons formulé le problème inverse sous la forme d'un problème de minimisation aux moindres carrés lui-même résolu par des méthodes d'optimisation locales. L'élément clé de ce type de méthode est le gradient de la fonction coût que nous exprimons par la méthode de l'état adjoint et dont nous avons validé le calcul en le comparant avec son homologue obtenu par différences finies. L'étude de la structure du gradient nous a aussi permis de mettre en relief les analogies qui existent entre la FWI et le processus de migration. Elle permet aussi de comprendre les mécanismes du processus d'imagerie et de déterminer la contribution des données à l'estimation des paramètres du modèle. Nous avons ensuite présenté l'implémentation de l'algorithme d'inversion de formes d'ondes et donné un exemple de reconstruction pour un modèle synthétique simple. Finalement, nous avons discuté du problème d'estimation de la source télé-sismique.

17. Il existe aussi des catalogues de fonctions source pour les grands séismes ($M_w \geq 7.5$) comme le catalogue SCARDEC (Vallée et al., 2011; Vallée & Douet, 2016).

Dans le prochain chapitre, nous analysons plus en détails le comportement de l'inversion de formes d'ondes en configurations télésismiques à l'aide d'un modèle réaliste de lithosphère représentatif de la structure des Alpes Occidentales. Nous étudions notamment le comportement de la FWI en fonction des modèles initiaux utilisées, du choix de la paramétrisation et de la géométrie expérimentale d'acquisition. Nous discuterons aussi de leurs impacts en terme de résolution.

Chapitre 4

Analyse paramétrique de la FWI en configuration téléseismique

Sommaire

4.1	Introduction	149
4.2	Teleseismic full-waveform inversion	153
4.2.1	Theoretical background	153
4.2.2	Gradient of the misfit function	154
4.2.3	Resolution analysis	156
4.2.4	Radiation patterns for different subsurface parametrization	157
4.3	The Western Alps lithospheric model	162
4.3.1	Model description	162
4.3.2	Teleseismic data anatomy	163
4.4	Parametric analysis of teleseismic FWI	165
4.4.1	Experimental set-up	165
4.4.2	Which earth parametrisation for teleseismic FWI ?	167
4.4.3	Which initial model for teleseismic FWI ?	170
4.4.4	Sensitivity to acquisition design	176
4.5	Conclusions	184

Ce chapitre est une première version d'une publication que nous comptons soumettre au journal *Geophysical Journal International*. Ce travail présente une analyse paramétrique de l'application de la méthode d'inversions de formes d'ondes aux données télésismique à partir d'un modèle lithosphérique synthétique, représentatif de la collision continentale des Alpes occidentales. Cette étude vise à analyser les effets spécifiques de trois facteurs pouvant impacter le comportement et la résolution spatiale de l'inversion de formes d'ondes. Ces facteurs expérimentaux portent sur le choix des paramétrisations élastiques isotropes utilisées d'une part durant le processus d'optimisation et d'autre part lors de la visualisation/interprétation des modèles, les modèles initiaux du processus d'optimisation et la géométrie et l'échantillonnage du dispositif de capteurs. Cette analyse indique que les modèles FWI les mieux résolus sont obtenus en utilisant la paramétrisation (ρ, λ, μ) durant l'optimisation et en recombinaison les impédances P et S lors de la phase de visualisation et d'interprétation. Nous confrontons ces conclusions héritées de l'expérience numérique avec l'analyse des diagrammes de rayonnement des différentes paramétrisations testées : (ρ, λ, μ) , (ρ, V_p, V_s) et (ρ, I_p, I_s) . Elle indique aussi que les modèles 1D de références tels que le modèle PREM sont suffisamment précis pour l'application de la FWI à l'échelle lithosphérique en considérant le contenu basse fréquence des données télésismiques. Enfin, elle montre sans ambiguïté que les géométries d'acquisition 2D sont à privilégier au détriment de géométries 1D, même ultra-denses, puisqu'elles augmentent considérablement les résolutions horizontales et verticales en prenant plus naturellement en compte des effets de diffraction hors plan.

**On the sensitivity of teleseismic full waveform inversion to earth
parametrisation, initial model and acquisition design**

S. Beller, V. Monteiller, L. Combe, S. Operto and G. Nolet
to be submitted to *Geophysical Journal International*.

Abstract

Full waveform inversion is not yet a mature imaging technology for lithospheric imaging from teleseismic data. Therefore, its promise and pitfalls need to be assessed more accurately according to the specifications of teleseismic experiments. Three important issues are related to the choice of the lithospheric parametrisation for optimisation and visualization, the initial model and the acquisition design, in particular in terms of receiver spread and sampling. These three issues are investigated with a realistic synthetic example that has been inspired by the CI-FALPS experiment in the Western Alps. Isotropic elastic full waveform inversion is implemented with an adjoint-state formalism and aims to update three parameter classes by minimization of a classical least-squares difference-based misfit function. Three different subsurface parametrisations, combining density (ρ) with P and S wave speeds (V_p and V_s), P and S impedances (I_p and I_s) or elastic moduli (λ and μ) are first discussed based on their radiation patterns before their assessment by FWI. We conclude that the (ρ, λ, μ) parametrisation provides the FWI models that best correlate with the true ones after recombining a posteriori the (ρ, λ, μ) optimization parameters into I_p and I_s . One-dimensional reference global models as PREM provide sufficiently-accurate initial models for FWI owing the low frequency content of teleseismic data

after smoothing, that is necessary to remove the imprint of the layering. Two kinds of station deployments are assessed : coarse areal geometry versus dense linear one. We unambiguously conclude that a coarse areal geometry should be favoured as it dramatically increases the penetration in depth of the imaging as well as the horizontal resolution. This results because the areal geometry significantly increases local wavenumber coverage, through a broader sampling of the scattering and dip angles, compared to a linear deployment.

Keywords : Teleseismic FWI – Parametrisation – Initial model – Acquisition

4.1 Introduction

Building high-resolution and quantitative images of the lithosphere from body waves is a key challenge in earthquake seismology. The methodological challenge results from the heterogeneous nature of the crust and the complex interaction of the wavefields with both these heterogeneities and the free surface. From the geodynamical viewpoint, building images of the lithosphere with a sufficient resolution is crucial to correlate tectonic deformation in the crust with deeper mantle processes arising in the asthenosphere. Lithospheric images can be built either from regional earthquakes (earthquakes nucleated in the lithospheric target) or from distant earthquakes, located several thousands of kilometres away from the target and commonly referred to as teleseisms. One key advantage of the teleseismic configuration is to potentially provide a large catalogue of events that are suitable for lithospheric imaging by either travel-time tomography, receiver function analysis or waveform inversion techniques. The teleseismic wave field can fairly be approximated by an incident planar wave field incoming from outside of the lithospheric target with different incidence angles and back-azimuths. This planar configuration is conducive to a rather uniform illumination of the lithospheric target but might be a limiting resolution factor due to the limited angular illumination provided by plane-wave sources as opposed to point sources.

Teleseismic datasets are conventionally processed by either traveltime tomography or receiver function analysis. Traveltime tomography, when limited to the first-arrival P-wave, provides subsurface model with a resolution of the order of the first Fresnel zone width (Williamson, 1991). In the teleseismic setting where the incident planar P-wave wavefields impinge the base of the lithospheric target with incidence angles ranging between 20 to 60 degrees, the tomographic models tend to exhibit significant vertical smearing of lithospheric structures along banana-shape sensitivity kernel (Marquering et al., 1999), hence attesting to the limited resolution of these approaches. This vertical resolution issue has prompted some authors to enrich teleseismic datasets with either regional (Bavalia et al., 2016) or global (Weidle & Widiyantoro, 2005) datasets to perform P-wave traveltime tomography. Conversely, migration of receiver functions (Vinnik, 1977; Langston, 1979) can provide sharper images of lithospheric discontinuities (Zhu & Kanamori, 2000; Ryberg & Weber, 2000). However, stacking of P-wave receiver functions are inherently polluted by secondary arrivals, that limit their use to first order PS converted waves generated by strong impedance contrast at major lithospheric discontinuities (Moho, 410 km and, 660 km boundaries).

The current densification of permanent and temporary arrays in conjunction with the conti-

nuous growth of computing facilities permit the consideration of waveform inversion methods that are expected to provide quantitative images of the subsurface with a resolution close to the wavelength. Among teleseismic waveform inversion techniques, least-squares migration of scattered teleseismic body waves has been first proposed by [Bostock et al. \(2001\)](#), [Shragge et al. \(2001\)](#), [Rondenay et al. \(2001\)](#) and [Rondenay et al. \(2005\)](#). In their approach, the forward problem is linearised with the single-scattering Born approximation, Green functions are computed by dynamic ray tracing and the migration is recast as a linear waveform inversion accordingly, where the misfit between the recorded and modelled scattered wavefields is minimised in a least-squares sense. [Bostock et al. \(2001\)](#) have emphasized the key role of the free surface as a secondary source of down-going P and S waves. Compared to migration of receiver function, that is often focused on the forward-scattered converted P-S waves, migration of the reflected waves that are back-scattered by the lithospheric reflectors after a first reflection from the free surface produces sharp images of the lithospheric structures.

A natural extension of the former linear waveform inversion is non-linear full-waveform inversion (FWI), a technique also borrowed from exploration geophysics ([Tarantola, 1984a](#); [Mora, 1987](#); [Pratt, 1999](#); [Virieux & Operto, 2009](#)). The main difference with the above-mentioned linear waveform inversion is that the full wave field residuals are minimized in FWI as opposed to the single-scattered wave field residuals in scattering migration. This implies that the background subsurface model, in which seismic modelling is performed, is updated at each FWI iteration, while the background model is kept the same over iterations in scattering migration. A key advantage of the non linear inversion is that, as the resolution of the subsurface models in which seismic modelling is performed improves over iterations, the resulting discontinuities can help to more fully account for converted waves and multi-scattering during full-waveform modelling and inversion. Conversely, a potential advantage of the linear inversion is that the explicit separation between the background wave field and the scattered wave field during a preprocessing stage and the use of ray tracing to select specific phases can be helpful to drive the inversion toward the most resolving information in the data (back-scattered P-S converted waves). Alternatively to FWI, [Shang et al. \(2012\)](#) proposed to image lithospheric reflectors by reverse time migration, while the background velocity models are built by reflection tomography from the free surface multiples ([Burdick et al., 2014](#)). This purely reflection approach discards the contribution of the incident wavefields and uncouple the update of the long wavelengths of the lithosphere by migration-based velocity analysis from the shorter ones by migration through an explicit scale separation as conventionally performed in multichannel seismic reflection processing.

In its conventional form, FWI seeks to minimize in a least-squares sense the sample-to-sample misfit between the recorded and modelled seismograms. This implies that both traveltimes and amplitudes of all of the arrivals are involved in the misfit function. One question which arises in relation to the use of the amplitudes is whether secondary parameters such as density and attenuation can be reliably updated during FWI. While multi-parameter reconstruction for density and attenuation is an active field of research in controlled-source seismology ([Hicks & Pratt, 2001](#); [Malinowski et al., 2011](#); [Kamei & Pratt, 2013](#); [Prioux et al., 2013a](#); [Kurzman et al., 2013](#); [Groos et al., 2014](#)), the path taken by the earthquake seismology community has been rather to modify the sensitivity kernels of the FWI to remove amplitudes effects and recast the waveform inversion as a finite-frequency traveltimes inversion of the first arrival ([Marquering](#)

et al., 1999) or selected wave packets (Maggi et al., 2009). This approach is generally referred to as adjoint tomography by the earthquake seismological community when seismic modelling is performed with full-wave numerical techniques such as the spectral element method (Komatitsch & Tromp, 1999a). Here, by adjoint is meant a mathematical technique that allows for the efficient computation of the gradient of a functional without forming the sensitivity matrix (see Tromp et al. (2005); Fichtner et al. (2006a,b); Plessix (2006a) for a review). An alternative to the adjoint approach is the scattering-integral approach which relies on the explicit computation of the sensitivity matrix (Chen et al., 2007). Several applications of adjoint tomography at regional, continental and global scales have been presented by Tape et al. (2010), Fichtner et al. (2009), Fichtner et al. (2010), Zhu et al. (2012), Fichtner et al. (2013), Zhu et al. (2015) and Bozdağ et al. (2016).

During these last years, there have been a few attempts to apply FWI on teleseismic data for high-resolution imaging of lithospheric targets. The first attempts were performed in 2D and 2.5D (Roecker et al., 2010; Baker & Roecker, 2014; Pageot et al., 2013) using a frequency-domain implementation (Pratt, 1999). A frequency domain implementation was used because 2D seismic modelling can be performed efficiently with Gauss elimination techniques when a large number of sources are processed for a few discrete frequencies (Pratt, 1999). Pageot et al. (2013) presented a parametric analysis that highlighted the contribution of the back-scattered waves from the free surface to improve the resolution of the imaging as originally pointed out by Bostock et al. (2001). They also show that a fine sampling of the frequencies was required during the inversion to balance the narrow and coarse scattering angle illumination provided by a few plane wave sources. This prompts them to conclude that a time-domain implementation of FWI was more suitable than the frequency-domain counterpart according to the specificity of the teleseismic experiments.

Three-dimensional time-domain FWI of teleseismic data has been first investigated by Tong et al. (2014a), who assessed the ability of FWI to reconstruct reflector geometry and density, V_p and V_s models of simple volumetric structures such as cubic inclusion or slab through a careful analysis of sensitivity kernels of specific phases. Monteiller et al. (2015) illustrate the resolution power of FWI relative to traveltimes tomography to image a simple crustal model with a sharp Moho even when only four teleseismic events are used for inversion. The reliability of FWI for teleseismic application was further demonstrated with an application to real data collected in the framework of the Pyrope experiment across the Pyrénées range (Wang et al., 2016). FWI of five events recorded by 29 stations provided tomographic P and S wave velocity (V_p and V_s) models of the Pyrenean lithospheric structure with an unprecedented resolution. Moreover, Wang et al. (2016) showed the consistency between the FWI V_p and V_s models and migrated images inferred from receiver functions. In parallel with this, Beller et al. (2017) applied FWI on nine teleseismic events collected during the CIFALPS experiment (Zhao et al., 2015) in the Western Alps. They use a FWI approach similar to Monteiller et al. (2015) except that the density (ρ) was updated in addition to V_p and V_s . The FWI models show convincing images of the continental subduction and the Ivrea body as well as a low V_s zone in the lower lithosphere supporting the hypothesis of a slab detachment. The inversion was pushed up to a maximum frequency of 0.2Hz. The ρ model shows structures that are consistent with those shown in the V_s model in the upper lithosphere. However, the maximum depth at which ρ can be imaged is smaller than for V_s suggesting that ρ has been mostly reconstructed from free-

surface multiples. Another striking result is the good resolution of the imaging in the cross-line direction, perpendicular to the main receiver line of the CIFALPS experiment, although a quite sparse distribution of stations in the cross-line direction. In particular, [Beller et al. \(2017\)](#) successfully recovered the arc shape of the Ivrea body, whose ρ model allows to match the Bouguer anomaly.

Although these recent success, there is still a need to appraise the promises and pitfalls of 3D teleseismic FWI. First, the plane-wave configuration raises the issue of the resolution with which lithospheric structures can be imaged from a few teleseismic events. Although [Bostock et al. \(2001\)](#) have emphasized the key role of the free surface multiples, it is not yet clear to which extent second-order back-scattered waves can be exploited by FWI when the full wave field is processed in one go, i.e., without explicit separation between the incident and scattered wavefields. A second issue is related to the choice of the best subsurface parametrization (the parameter set that fully describe the subsurface) and the optimization parameters (the subset of parameters that are updated during the inversion as opposed to the passive parameters that are kept fixed) for teleseismic FWI. This issue is far to be neutral as long as one seeks to update secondary parameters such as density, attenuation or anisotropic parameters (see [Operto et al. \(2013\)](#) for a tutorial). The choice of the subsurface parametrization and the optimization parameters is mainly driven by the need to mitigate parameter cross-talk, while preserving a sufficient resolution in the imaging. The most common tools to define a suitable parametrization for FWI relies on the analysis of the so-called radiation patterns ([Tarantola, 1986](#); [Forgues & Lambaré, 1997](#); [Gholami et al., 2013](#); [Alkhalifah & Plessix, 2014](#)), the principal component analysis of the sensitivity matrix ([Sieminski et al., 2009](#)) or the eigenvector analysis of the Hessian operator ([Plessix & Cao, 2011](#)).

A third issue is related to the role of the initial model in teleseismic FWI. It is well acknowledged that this initial model should allow to predict traveltimes with an error that does not exceed half the period to prevent cycle skipping artefacts when a conventional least-squares difference-based misfit function is used (e.g. [Virieux & Operto, 2009](#)). This condition is challenging to fulfil as long as a large number of wavelengths are propagated. This arises when seismic data lack low frequencies as in controlled-source seismology and the acquisition geometry involve long propagation distances. The low frequency content of teleseismic data should make FWI reasonably immune to cycle skipping. However, it is unclear if other factors than the kinematic accuracy of the initial model, for example the smoothness versus blocky character of the velocity model or site effects near the free surface, can drive the FWI toward a local minimum. A fourth obvious factor is related to the footprint of the station spread and sampling in teleseismic FWI. Clarifying this issue is crucial to design reliable surveys amenable to high resolution imaging methods such as FWI. Using too coarse stations network can generate aliasing artefacts of different nature ([Shragge et al., 2001](#); [Rondenay et al., 2005](#)) as well as wraparound of geological structures ([Mulder & Plessix, 2004](#); [Pageot et al., 2010](#)). This has prompted the design of several teleseismic acquisitions in the form of dense profile of stations spaced 5 km apart at the expense of coarser areal deployments ([Rondenay et al. \(2001, Cascadia experiment\)](#), [Iglesias et al. \(2010, MASE experiment\)](#), [Zhao et al. \(2015, CIFALPS experiment\)](#)). While such a dense linear set-up might be suitable to sample teleseismic events arriving with an azimuth aligned with the receiver line, it is not be suitable to record waves that are scattered off the receiver line or process events reaching the array with a significant obliquity.

The aim of this study is to gain new insights on the four above-mentioned issues through the application of 3D time-domain FWI on a realistic synthetic case study inspired from the CIFALPS experiment.

In the sequel of this study, we first review the main theoretical ingredients of the FWI. Then, we present the synthetic lithospheric model of the western Alps and describe the anatomy of the data computed in this model. We interpret the dominant arrivals recorded by the three geophone components and identify part of the signal that underwent aliasing. This is followed by a parametric analysis of the FWI. We first assess three different subsurface parametrizations : (ρ, λ, μ) , (ρ, v_p, v_s) , (ρ, I_p, I_s) , where λ and μ are the Lamé parameters and I_p and I_s the P and S impedances. Our numerical examples suggest that the (ρ, λ, μ) parametrization provides the best FWI results after recombining a posteriori the FWI ρ , λ and μ models into I_p and I_s . Second, we assess the sensitivity of the FWI to the initial models. We show that ideally a smooth tomographic model is the most suitable one for FWI, although 1D reference models as PREM are enough to achieve acceptable FWI results. Third, we assess the sensitivity of the teleseismic FWI to two different kinds of acquisition geometry : coarse areal configuration versus dense linear acquisition. Our results strongly support that areal deployment should be firstly designed at the expense of a dense linear deployment.

4.2 Teleseismic full-waveform inversion

This section reviews some basic principles of FWI that will be useful to interpret the results of the numerical experiments carried out in this study. For sake of compactness, we review these principles with a matrix formalism suitable for the frequency-domain formulation of FWI, although our teleseismic FWI is fully implemented in the time domain.

4.2.1 Theoretical background

Let's assume that a modelling engine such as those based upon hybrid methods is available to compute synthetic teleseismic seismograms (Monteiller et al., 2013; Tong et al., 2014b; Beller et al., 2017). The elastodynamic equations can be written in matrix form as

$$\mathbf{A}(\mathbf{m})\mathbf{u} = \mathbf{s}, \quad (4.1)$$

where \mathbf{s} denotes the seismic source term and \mathbf{A} the wave modelling operator. The synthetic waveforms \mathbf{u} are computed from subsurface properties gathered in the vector \mathbf{m} . The most natural parametrization of the subsurface involve the density ρ and elastic moduli c_{ijkl} since this parametrization directly results from the equation of motion and the Hooke's constitutive law.

FWI is a data fitting procedure that aims to find a model \mathbf{m} that minimises a certain measure \mathcal{C} of the data misfit, in its most conventional form, the least-squares norm of the sample-to-

sample difference between the observed and the modelled data

$$\mathbf{m} = \arg \min_{\mathbf{m}} \mathcal{C}(\mathbf{m}) \quad (4.2)$$

where $\mathcal{C}(\mathbf{m}) = \frac{1}{2} \|\Delta \mathbf{d}\|_2^2$ and $\Delta \mathbf{d} = \mathbf{R}\mathbf{u}(\mathbf{m}) - \mathbf{d}$ is the difference between the recorded data \mathbf{d} and the modelled wave field sampled at the receiver positions by the operator \mathbf{R} . Because of the computational burden of the seismic modelling and the huge dimension of the model space, FWI is recast as an iterative local optimisation problem that seeks the minimum of the misfit function in the vicinity of a starting model \mathbf{m}_k at iteration k . Zeroing the gradient of the misfit function after a second-order Taylor expansion of the misfit function (4.2) around the starting model provides the Newton descent direction. Multiplying this descent direction with a step length γ_k , that can be found by a line search procedure (Nocedal & Wright, 2006a), provides the model perturbation $\Delta \mathbf{m}_{k+1}$ at iteration k

$$\Delta \mathbf{m}_{k+1} = -\gamma_k \mathbf{H}^{-1}(\mathbf{m}_k) \nabla_{\mathbf{m}} \mathcal{C}(\mathbf{m}_k), \quad (4.3)$$

where \mathbf{H} and $\nabla_{\mathbf{m}} \mathcal{C}$ are the Hessian (the second derivative of the misfit function) and the gradient of the misfit function with respect to \mathbf{m} , respectively.

4.2.2 Gradient of the misfit function

The gradient of the misfit function is given by

$$\nabla_{\mathbf{m}} \mathcal{C} = \left(\frac{\partial \mathbf{R} \mathbf{u}(\mathbf{m})}{\partial \mathbf{m}} \right)^T \Delta \mathbf{d} = \mathbf{J}^T \Delta \mathbf{d}, \quad (4.4)$$

where \mathbf{J} denotes the Fréchet derivatives matrix or sensitivity matrix (Pratt et al., 1998). Differentiation of equation (4.1) with respect to a model parameter m gives

$$\mathbf{A} \frac{\partial \mathbf{u}}{\partial m} = -\frac{\partial \mathbf{A}}{\partial m} \mathbf{u}. \quad (4.5)$$

Analogy between equations (4.1) and (4.5) shows that the partial derivative wave field $\frac{\partial \mathbf{u}}{\partial m}$ is solution of the elastodynamic equations for a secondary virtual source $-\frac{\partial \mathbf{A}}{\partial m} \mathbf{u}$. The sparse operator $\frac{\partial \mathbf{A}}{\partial m}$ defines the spatial and temporal supports of the virtual source at the position of m in space and at the arrival time of \mathbf{u} at m in time, respectively as well as its radiation pattern. Therefore, the partial derivative wave field can be interpreted as the wave field scattered by a point diffractor located at m . The traveltime curve followed by this partial derivative wave field sampled at the receiver positions (this sampling gives one column of the sensitivity matrix) describes a diffraction hyperbola (Fig. 4.1). The gradient of the misfit function is simply formed by the zero-lag correlation between the sampled partial derivative wave field and the data residuals or in other words by the summation along the diffraction hyperbola of the product between the partial derivative wave field and the data residuals. The aim of this correlation is to pick among all of the residuals those which result from a potential missing heterogeneity located at m .

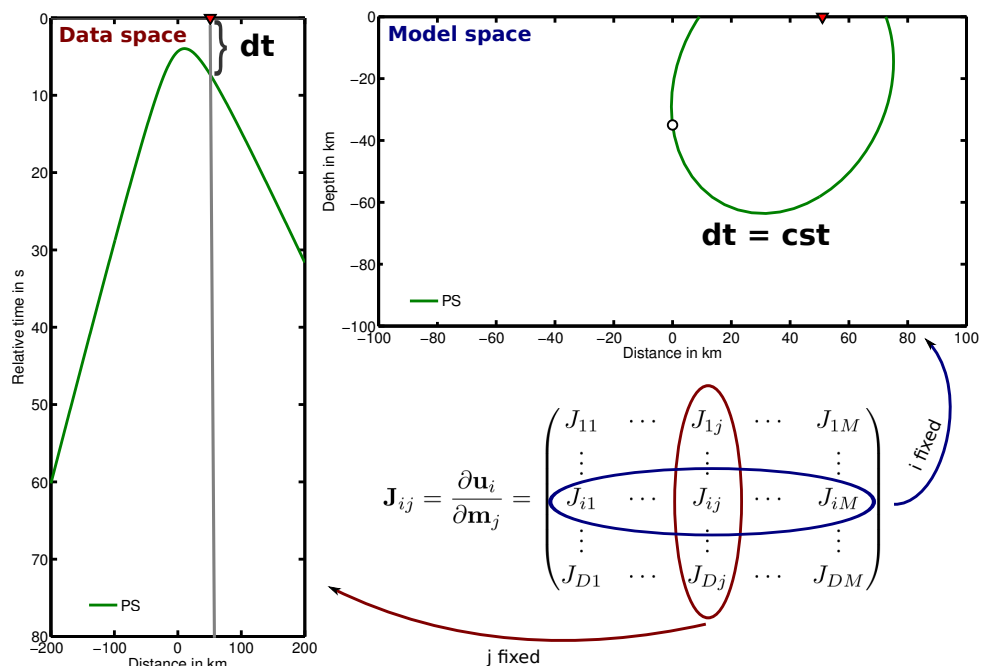


Figure 4.1 – Sensitivity matrix and FWI imaging principle. The column j of the sensitivity matrix describes the synthetic wave field that would be scattered by a diffractor located at m_j and sampled at the stations positions. In this figure, we show the case of a P-S diffraction. The support of this scattered wave field in the time-distance domain follows a diffraction hyperbola. The gradient of the misfit function is formed by a weighted stack of the data residuals along these trajectories. Conversely, the line i of the sensitivity matrix describes the positions in space along which a sample of the data residuals ∂u_i is projected. These positions describe an isochronal ellipse as in scattering migration. Summation over all possible ellipses focuses by constructive interference the image of the scatterer with a resolution controlled by the source bandwidth and the angular illumination provided by the acquisition geometry.

Although equations 4.5 and 4.6 draw some clear connection between FWI and diffraction tomography, the gradient of the misfit function can be more efficiently computed with the adjoint-state method at the cost of two forward simulations per source (Plessix (2006a) for a review). Following Pratt et al. (1998), injecting the expression of the partial derivative wave field (4.5) into (4.6) gives

$$\nabla_m \mathcal{C} = -\mathbf{u}^T \left(\frac{\partial \mathbf{A}}{\partial m} \right)^T (\mathbf{A}^{-1})^T \Delta \mathbf{d}^* = - \left(\frac{\partial \mathbf{A}}{\partial m} \mathbf{u} \right)^T \lambda^*, \quad (4.6)$$

where the adjoint wave field λ satisfies

$$\mathbf{A}^T \lambda^* = \mathbf{R}^T \Delta \mathbf{d}^*. \quad (4.7)$$

This expression shows that the gradient is obtained by a zero-lag correlation of the incident wave field \mathbf{u} weighted by the scattering matrix $-\frac{\partial \mathbf{A}}{\partial m}$ with the adjoint wave field λ , that is back-propagated in the subsurface using the residuals as a composite source term. Considering only one sample of the data residuals associated with a source s , a receiver r and a traveltime T , its energy will be smeared in the subsurface along positions x that satisfy the imaging condition $T_{sx} + T_{rx} = T$. Here, T_{sx} and T_{rx} denote the traveltimes of a wave connecting the source and the receiver to the scatterer x , respectively. These positions described an isochronal surface, which is an ellipse in a 2D homogeneous background medium (Fig. 4.1).

All of these scattering concepts and imaging condition are typically met in migration techniques (such as ray+Kirchhoff or ray+Born migration) (McMechan & Fuis, 1987; Miller et al., 1987; Lumley et al., 1994; Rondenay et al., 2005) and highlight the connections that can be drawn between FWI and generalized diffraction tomography (Devaney, 1984; Wu & Toksöz, 1987) as reviewed by Pratt et al. (1998).

For a general elastic medium parametrized by the density ρ and the coefficients c_{ijkl} of the stiffness tensor, the gradient is simply given by :

$$\begin{aligned} \nabla_\rho \mathcal{C}(\mathbf{x}) &= \int_0^T \mathbf{u}^\dagger(\mathbf{x}, t) \cdot \partial_t^2 \mathbf{u}(\mathbf{x}, t) dt, \\ \nabla_{c_{ijkl}} \mathcal{C}(\mathbf{x}) &= \int_0^T \boldsymbol{\epsilon}^\dagger(\mathbf{x}, t) : \frac{\partial \mathbf{c}}{\partial c_{ijkl}} : \boldsymbol{\epsilon}(\mathbf{x}, t) dt, \end{aligned} \quad (4.8)$$

where \mathbf{c} denotes the stiffness tensor, \mathbf{u} the displacement wave field and $\boldsymbol{\epsilon}$ and $\boldsymbol{\epsilon}^\dagger$, the strain tensor and the adjoint strain tensor, respectively (Tromp et al., 2005). Other parametrisations may be considered and deduced from the former using the chain-rule of derivatives (Köhn et al., 2012), three of them are inspected in section 4.2.4.

4.2.3 Resolution analysis

We review here the main experimental factors that control the spatial resolution of the FWI. In the framework of diffraction tomography and FWI respectively, Wu & Toksöz (1987) and Sirgue & Pratt (2004) showed that the gradient of the misfit function with respect to one parameter has the form of a truncated inverse Fourier series in which the arguments of the basis functions are the wavenumber components locally injected in the targeted model at the position

of the parameter. The truncation of this Fourier series, which results from the limited bandwidth of the source and the limited spread of the acquisition, limits the resolving power of FWI.

In the high-frequency approximation (or local plane wave approximation), the local wavenumber vectors \mathbf{k} injected at a given position in a 2D subsurface model (Fig. 4.2a) can be inferred from the ray paths connecting the sources and the receivers to the scatterer leading to the following expression

$$\mathbf{k} = \mathbf{k}_s + \mathbf{k}_r = (k_x, k_z) = k(\cos \phi, \sin \phi) \quad (4.9)$$

where

$$k = \frac{2\omega}{c} \sin \left(\frac{\theta}{2} \right) \quad (4.10)$$

and c is the local wave speed, ω the angular frequency, ϕ and θ are the dip and scattering angles, respectively (Miller et al., 1987; Wu & Toksöz, 1987; Lambaré et al., 2003).

The expression of \mathbf{k} shows that the local wavelength and the angles ϕ and θ are the three parameters that control the spatial resolution of the FWI. In teleseismic applications when a few up-going wavefields can lead to a coarse and narrow sampling of θ , the use of a broad frequency band is useful to guarantee a proper sampling of the wavenumbers in the spatial directions spanned by the dip angle ϕ . Moreover, free-surface multiples may enrich the range of dip and short scattering angles sampled by the acquisition, hence increasing the vertical resolution of the imaging.

Beyond the resolution limitations imposed by the limited frequency bandwidth of the source and the limited spread of the acquisition, the radiation pattern of the virtual sources further controls the amplitude versus θ variations of the partial derivative wavefields and hence potentially narrow the real range of wavenumbers mapped in the FWI gradient. In the next section, we analyse the radiation patterns of the virtual sources for the three subsurface parametrizations that are used in this study for a theoretical assessment of their impact on the FWI resolution and the trade-off between parameters.

4.2.4 Radiation patterns for different subsurface parametrization

Inferences on the sensitivity of the wave field to different parameter classes can be drawn from the radiation patterns of elastic waves scattered by small elastic heterogeneities (Wu & Aki, 1985; Tarantola, 1986). Such patterns give clear insights on the influence of a single parameter onto the data for different scattering regime (forward versus backward) (Fig. 4.2a). Moreover, the overlap of the radiation patterns of different parameters with θ indicates potential trade-off between these parameters. On the one hand, one may want to mitigate these trade-offs by choosing a parametrisation that minimises these overlaps. On the other hand, this strategy might narrow the influence of the parameters with θ , hence degrading the resolution with which parameters are reconstructed accordingly. Many different parametrisations can be considered for FWI. However, in the case of isotropic elasticity, the most common ones combine ρ with either wave speeds (V_p, V_s), impedances (I_p, I_s) or elastic moduli (λ, μ). The radiation patterns of these three parametrisations are shown in Fig. 4.2(b-d) for a 2D homogeneous medium.

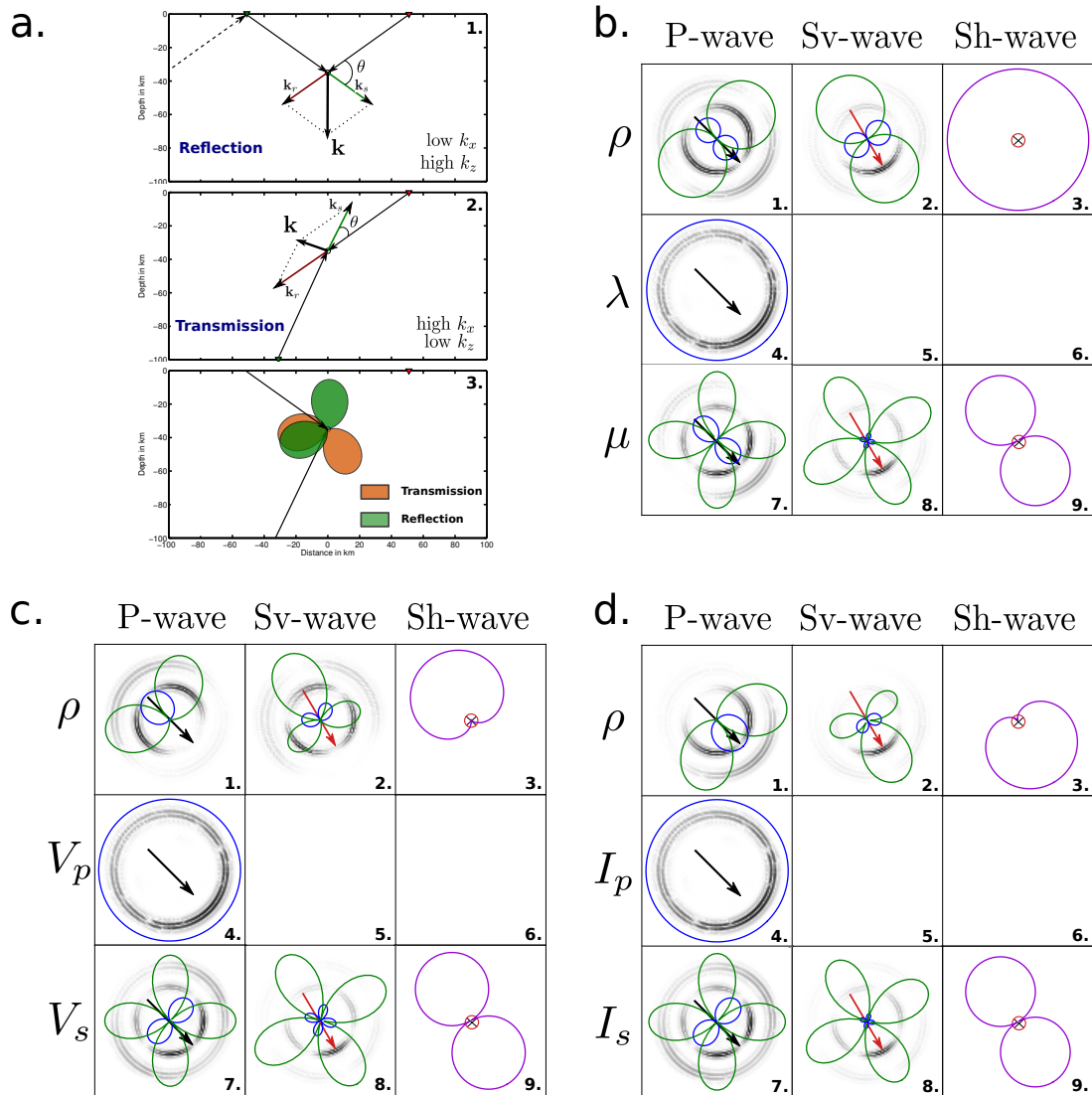


Figure 4.2 – Radiation patterns for elastic isotropic media. (a.) Schematic illustration of the relationship between the wavenumber components mapped in the subsurface at a given position with the angular illumination provided by the acquisition and the reflectors of the background subsurface model (here, the free surface). In a teleseismic setting, the forward scattering generated during the up-going propagation of the incident wave field toward the surface will contribute to preferentially map the low-to-intermediate horizontal wavenumbers, while the backward scattering generated after the reflection of the incident wave field from the free surface will contribute to preferentially map the high-to-intermediate vertical wavenumbers. (b,c,d) Radiation patterns of elastic scatterers for the b) (ρ , λ , μ), c) (ρ , V_p , V_s) and d) (ρ , I_p , I_s) parametrisations . From left to right, each column presents the interaction of a P (1,4,7), Sv (2,5,8) and Sh-wave (3,6,9) incident plane wave with, from top to bottom, a ρ (1-3), λ (4-6) (resp. V_p and I_p) and μ (7-9) (resp. V_s and I_s) scatterer. All the analytical radiation patterns (Wu & Aki, 1985) are superimposed with the scattered wavefields that are computed numerically. Arrows indicate the direction of the incident plane wave (black for P-wave, red for Sv wave and red-black for Sh wave). Radiation-pattern curves are plotted in blue, green and purple for the P, Sv and Sh scattered modes, respectively. See text for the interpretation.

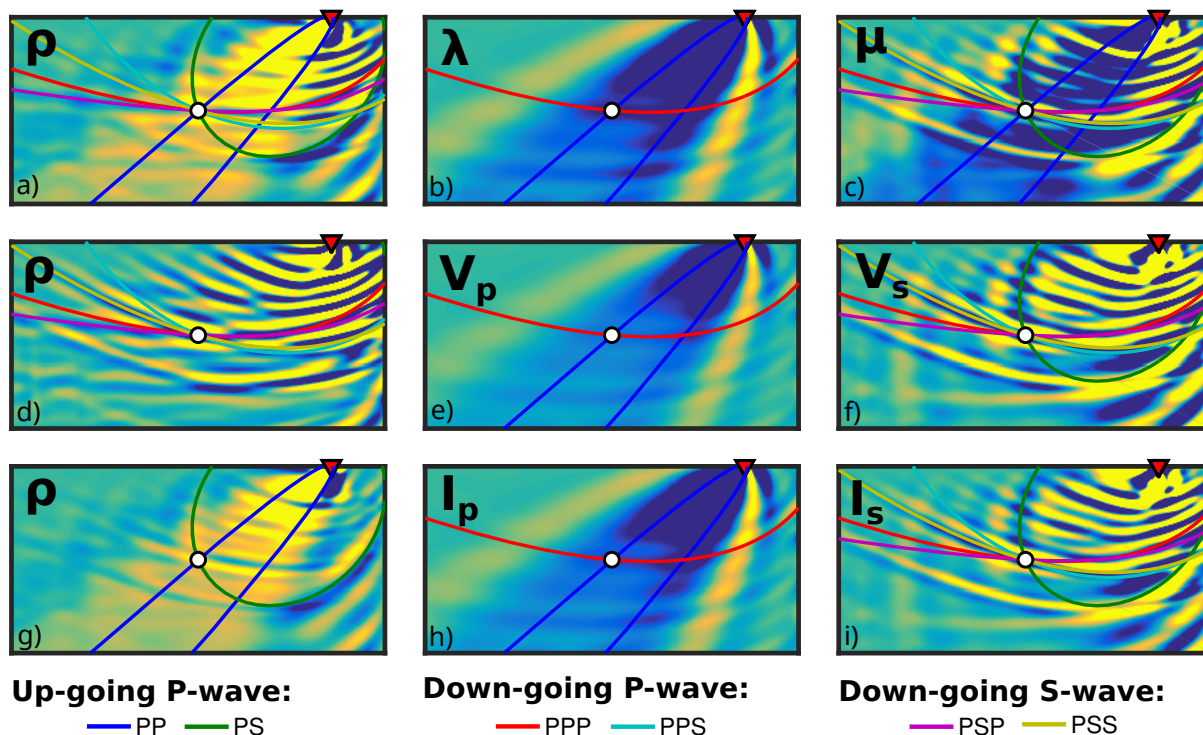


Figure 4.3 – Anatomy of a teleseismic gradient for one source-receiver pair and different parametrisations. FWI gradients computed for (a-c) (ρ, λ, μ) , (d-f) (ρ, V_p, V_s) and (g-i) (ρ, I_p, I_s) parametrisations. The source is a 45° incident plane-wave travelling through an homogeneous half-space with a small heterogeneity located on the centre of the figure (white circle). The gradient is computed for a 3-component sensor located at the position of the red triangle. The isochronal curves of the scattered-waves generated through the interaction of the up-going P-wave and down-going P and S-wave with the small heterogeneity are superimposed on the gradients. See text for interpretation.

The presentation of the radiation patterns is complemented with multi-parameter FWI gradients computed for an incident P wave with an incidence angle of 45° and one station (Fig. 4.3). The background medium is a homogeneous half space to which is added a small circular inclusion at 50km depth to generate the data residuals.

A first noticeable feature, that is shared by the three considered parametrisations, is that λ (resp. V_p and I_p) generates only P-P scattering with an isotropic radiation pattern (Fig. 4.2(b-d),(4-6)). Therefore, the radiation pattern of λ (resp. V_p and I_p) has no impact on the resolution with which this parameter is reconstructed : the local resolution is only controlled by the local P wavelength and the range of P-P scattering angles spanned by the acquisition. The FWI gradient with respect to λ (resp. V_p and I_p) highlights the dominant contribution of the forward-scattered P-P mode delineated by the half-ellipse isochronal curve with one focus at the station (Fig. 4.3b,e,h, blue curve). This forward-scattering regime is associated with wide θ and will contribute to update the long wavelengths of λ (resp. V_p and I_p) (Fig. 4.2a). Superimposed on the forward-scattered P-P sensitivity kernel, we also see, with much smaller amplitudes, the migration isochrone associated with the doubly P-P back-scattered wave from the free-surface and the inclusion (Fig. 4.3b,e,h, red curve). This back-scattering regime is associated with smaller θ and hence will contribute to update shorter wavelengths of λ (resp. V_p and I_p) (Fig. 4.2a). Since the forward-scattering regime is dominant, the FWI will tend to update the long wavelengths first before updating the shorter ones from the back-scattered waves once the residuals of the incident wave will have been reduced. This hierarchy between forward and backward scattering naturally honours a multi-scale approach which is useful to mitigate the nonlinearity of the FWI.

In contrast, μ (resp. V_s and I_s) gives rise to P-P, P-S, S-S and S-P scattering, with stronger diffracted S waves relative to the P counterparts (Fig. 4.2(b-d),(7-9)). The P-S, S-S and S-P radiation patterns are identical for the μ , V_s and I_s and the union of these three radiation patterns covers a broad range of θ , that is conducive to a broadband reconstruction of these parameters if the imprint of all of these scattering modes is significant in the data. However, reminding that only incident P waves are considered in this study, the two lobes of the P-Sv radiation pattern pointing forward in the direction of the arrow with an angle of 45° in Fig. 4.2(a-c)-(7, orange curves) show that the long wavelengths of μ (resp. V_s and I_s) cannot be updated from the P-Sv mode since there is no radiation at very wide θ . This is supported by the gradient with respect to μ (resp. V_s and I_s) in Fig. 4.3c,f,i, green curve, which shows the more focused spatial domain spanned by the P-Sv forward scattering mode relative to the P-P one (Fig. 4.3b,e,h, blue curve). Note also that the amplitudes of the P-Sv forward scattering mode are small compared to those of the doubly back-scattered modes in Fig. 4.3c,f,i due to the low values of the P-S transmission coefficient relative to those of the P-Sv reflection coefficient at the free surface (Pageot et al., 2013, Their Fig. 3).

Interestingly, the only difference between the radiation patterns of μ , V_s and I_s is that μ gives rise to two lobes of P-P scattering at wide and small θ (Fig. 4.2(b),(7)), while V_s and I_s give rise to P-P scattering at intermediate θ (Fig. 4.2(c-d),(7)). This difference manifests in Fig. 4.3c, blue curve, by the clear imprint of the P-P forward-scattering mode in the μ gradient within a half-ellipse sensitivity kernel, unlike in the V_s and I_s gradients. Owing that the long wavelengths of μ (resp. V_s and I_s) cannot be updated from the P-Sv scattering mode as above-mentioned,

the ability to update the long wavelengths of μ from the P-P mode might be an argument in favour to the (ρ, λ, μ) parametrisation in particular if the inversion is started from a crude initial μ model.

Potential trade-off between λ (resp. V_p and I_p) and μ (resp. V_s and I_s) would result from the overlap between the P-P radiation patterns of μ (resp. V_s and I_s) and λ (resp. V_p and I_p). These trade-offs might be more significant during the update of the long wavelengths of λ and μ because the P-P forward-scattering sensitivity kernel of these two parameters are very similar (Fig. 4.3(b-c)). This comment highlights the fact that it is often challenging to update multiple parameter classes with a good resolution, while minimizing the trade-offs between them.

The radiation patterns of ρ show significant differences for the three parametrisations, although ρ generates scattering for the P-P, P-S, S-S and S-P scattering modes in the three cases (Fig. 4.2(b-d),(1)).

For (ρ, λ, μ) , the radiation patterns are symmetric with respect to the incident plane wave-front. In other words, forward and backward scattering are equally generated by ρ with this parametrisation. Moreover, combining the contribution of the P and S scattered waves provides a sensitivity over the full range of scattering angles. Conversely, ρ scatters waves in a preferential direction with the two other parametrisations : backward for (ρ, V_p, V_s) (Fig. 4.2(c),(1)) and forward for (ρ, I_p, I_s) (Fig. 4.2(d),(1)). This implies that (ρ, I_p, I_s) is suitable to update the long to intermediate wavelengths of ρ , while (ρ, V_p, V_s) is more suitable to update the short to intermediate wavelengths. We might conclude that (ρ, λ, μ) might be the most suitable one to update ρ because the first-order P-P and P-S forward-scattering modes can yet generate models of good resolution, which can be further improved by the contribution of the doubly Sv-P and Sv-Sv back-scattered waves. Note also that the Sh-Sh radiation pattern is isotropic and hence can contribute to significantly increase the resolution of ρ .

This theoretical analysis is supported by the gradients with respect to ρ (Fig. 4.3a,d,g). For (ρ, λ, μ) , we show well-balanced contributions of the forward- and backward scattering regimes, while the broad P-P elliptical sensitivity kernel of the forward-scattering (delineated by the blue curve) is dominant in the (ρ, I_p, I_s) parametrisation. Conversely, the more spatially-focused migration-like isochronal surfaces of the backward scattering are dominant for the (ρ, V_p, V_s) parametrisation.

Significant trade-offs are expected between λ (resp. V_p, I_p) and ρ and between μ (resp. V_s, I_s) and ρ according to their respective radiation patterns. However, it is worth remembering that λ (resp. V_p, I_p) and μ (resp. V_s, I_s) control both the kinematic and the dynamic attributes of seismic waves, while density mostly influences amplitudes. Heuristically, we may think that the long-wavelength reconstructions of λ (resp. V_p, I_p) and μ (resp. V_s, I_s) will be primarily tied to the need to fit traveltimes and that amplitudes will play a second-order role in their update, hence limiting the imprint of these parameter trade-offs. Controlling the trade-off between the update of the small to intermediate wavelengths of ρ and λ (resp. V_p, I_p) or ρ and μ (resp. V_s, I_s) from reflections at short to intermediate scattering angles is probably more challenging.

A last comment concerns the presence of ghost isochronal surfaces in the gradients that have been focused off the position of the inclusion (Fig. 4.3). These isochrones result from the misinterpretation by the gradient of doubly-scattered waves from the free surface and the

inclusion as single scattered waves from the inclusion and vice versa. These artefacts can be partially cancelled out during the gradient computation by the destructive interferences of the contributions of several receivers and events. Moreover, the Gauss-Newton Hessian can further contribute to remove these artefacts by a deconvolution-like processing since the zero-lag correlation between the partial derivative wavefields will mimic the same undesired correlations than the zero-lag correlation between the partial derivative wavefields and the data residuals. Note that the artefacts related to doubly-scattered arrivals generated by the discontinuities of the background model should not be confused with those generated by double scattering between missing heterogeneities in the background model. These doubly-scattered arrivals are present in the recorded data but are not modelled during the FWI gradient calculation, which relies on the single-scattering approximation. In this case, the artefacts generated by the recorded doubly-scattered arrivals can be removed from the FWI gradient by the second-order (non linear) term of the Hessian.

4.3 The Western Alps lithospheric model

4.3.1 Model description

Beller et al. (2017) applied FWI on nine teleseismic events recorded by the CIFALPS temporary array to image the lithospheric structure of the South-Western Alps (Fig. 4.4a). In the light of this study and for the purpose of the current study, we design a lithospheric model that is representative of the retrieved structure of the Alpine lithosphere.

The synthetic model (Fig. 4.4b) represents the subduction of the European continental crust below the Adriatic plate in the Western Alps. This model is $400 \text{ km} \times 200 \text{ km}$ long and 200 km deep. In its shallowest part, the model exhibits slow velocities / low density sedimentary basins, that are 10km thick to mimic the presence of the SE-France and Po plain sedimentary basins (Fig. 4.4a, red). The upper European crust is indented by a piece of Adriatic lithospheric mantle (Fig. 4.4b, light blue) referred to as the Ivrea body mantle wedge. From the West to the East, the model displays two distinct lower crusts (Fig. 4.4a, green), the first dips toward the east up to a maximum depth of 100 km and characterises the subduction of the European continental crust that under-thrusts the Ivrea body, whilst the other depicts the more conventional lower crust of the Adriatic plate. Below, we designed a lithospheric mantle that gently varies from 100 km to 125 km depth beneath the Ivrea body. The asthenospheric mantle located below also contains a piece of detached slab at the bottom end of the model (Fig. 4.4b).

Overall velocities of the synthetic model are based on the IASP91 model (Kennett & Engdahl, 1991) and densities are taken from the PEM-C model (Dziewonski et al., 1975). Seismic modelling in the lithospheric target is performed with a grid injection technique described in Beller et al. (2017) : we first perform a global-earth simulation in the PREM model (Dziewonski & Anderson, 1981) with the AxiSEM software (Nissen-Meyer et al., 2014) and store the 3D teleseismic wave field at the boundaries of the lithospheric target. Second, the teleseismic wave field is propagated in the lithospheric target with the spectral element method (Komatsich et al., 1998) through a grid injection technique (Monteiller et al., 2013). To avoid numerical

artefacts during the grid injection, we force a smooth transition between our synthetic model and the outer PREM model with a cosine taper in a 20 km thick layer along the boundaries of the lithospheric target.

4.3.2 Teleseismic data anatomy

Since FWI is a data-fitting procedure, it is relevant to interpret the potential information content of the teleseismic wave field and correlate "with one's own eyes" the geological features of the synthetic model with their seismic response. For that purpose, we show in Fig. 4.5 synthetic seismograms recorded by a dense profile of stations sampled 5km apart for an event, the characteristics of which are given in Table 4.1.

The relatively shallow depth (12 km) of this event together with the dominant period of the source signature (≈ 5 s) prevents the identification of source-related multiples such as pP and sP phases, which are hidden within the coda of the direct P-wave arrival (their theoretical arrival times from the onset of the direct P arrival are 3.95 and 5.6 seconds, respectively). In contrast, the intermediate epicentral distance of the event (63°) allows the PcP phase to be well-separated from the incident P-wave within the displayed time window. This phase has a theoretical arrival time of 37.6 s, according to PREM and the TauP toolkit (Crotwell et al., 1999), and is characterized by a sub-planar move-out (Fig. 4.5c).

At first glance, owing to the low frequency content of the data, it is challenging to identify reflections from major lithospheric boundaries in the recorded data, which are dominated by the incident P wave field and energetic diffraction from the Ivrea body. However, the footprint of several lithospheric structures, such as sedimentary basins and Ivrea body, is clearly visible in the onset time of the direct P arrival. In particular, the slow / buoyant sedimentary basins delay the traveltimes of the direct P arrival and increase its amplitudes (Fig. 4.5, L), while their pinch-out generate energetic diffractions with sharp dips (Fig. 4.5, SD). In contrast, the fast Ivrea body advances the direct P arrival onset and decreases its amplitudes (Fig. 4.5, E). Moreover, it behaves as a strong secondary point source near the surface, that scatters waves downward in the lithosphere. With respect to this comment, we interpret a late arrival with an hyperbolic moveout as the reflection from the Moho generated from this diffracting source (Fig. 4.5, M). The signature of the upper-mantle discontinuities corresponding to the top and bottom of the transition zone at 410 and 660 km depth are also visible at around 40s and 70s after the first arrival. These discontinuities are not part of the lithospheric target and hence will not be imaged by FWI. However, their signature in the data will not impact upon the imaging since they will be injected in the modelled synthetic seismograms by the grid injection technique.

The computation of receiver functions by frequency-domain deconvolution (Langston, 1979) further helps us to analyse the contribution of late reflections from crustal heterogeneities by removing the imprint of the incident wave field (figure 4.5(d-e)). The radial receiver function nicely shows the geometry of the Moho of the European and Adriatic plate, with a pronounced dip for the former. On the transverse component, receiver functions are dominated by three diffraction hyperbola created by the pinch-out of the sedimentary basins and the Ivrea body. Note that, due to their significant dips, these arrivals are spatially aliased. Later arrivals might indicate the location of the lithosphere-asthenosphere boundary (LAB). However, as is frequently

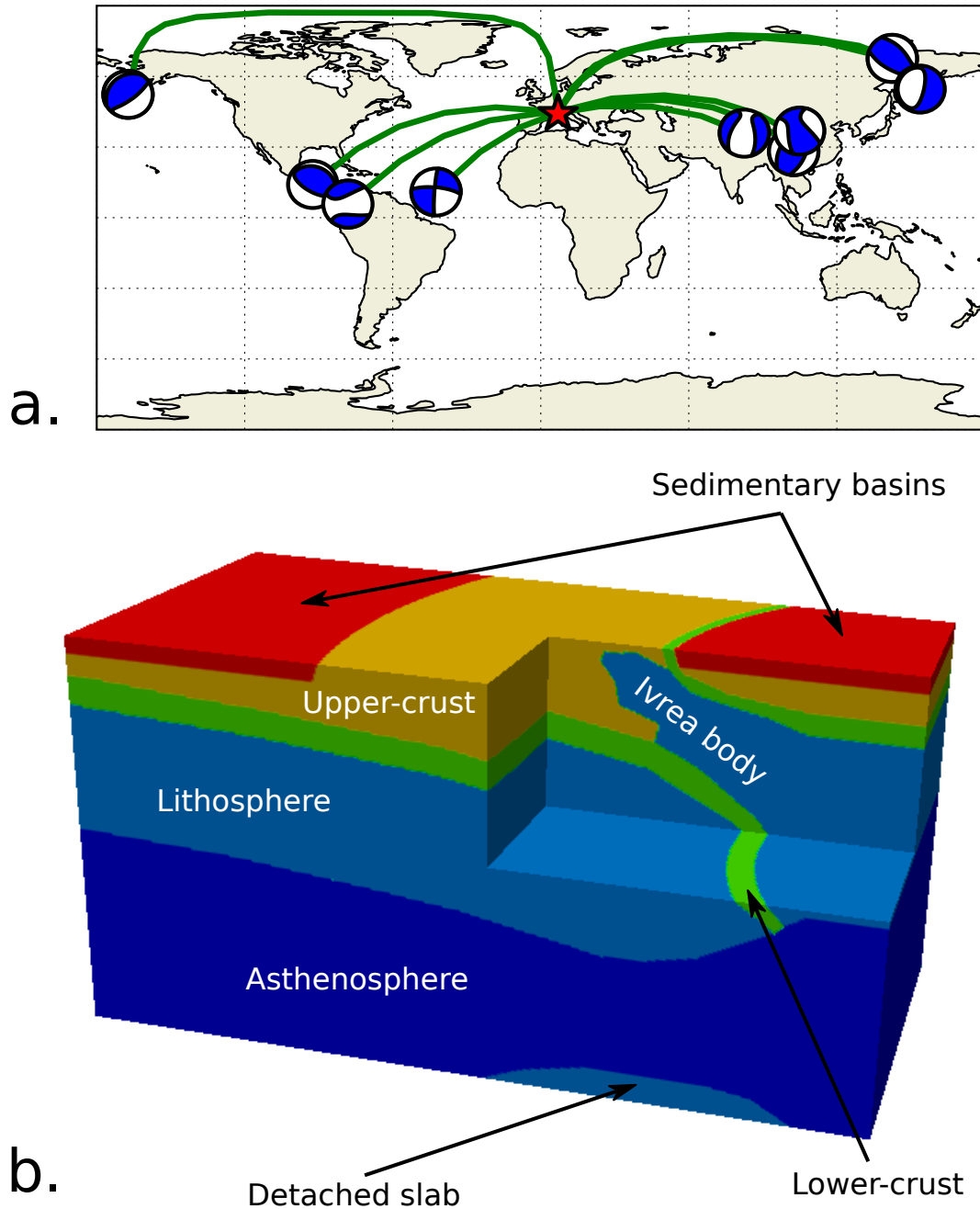


Figure 4.4 – Synthetic example design. (a.) The nine teleseismic events selected from the CI-FALPS dataset for FWI. (b.) The 3D Alps lithospheric model representative of a continental collision.


Id	Event	Lat (°)	Dist (°)	Hdur	Depth	Mw
		Lon (°)	Baz (°)	Tshift	(km)	CMT
22	2013-02-14	67.65	62.98	5.30	12	6.7
		13 :13 :53.1	142.512	17.37	6.26	

Table 4.1 – Selected event for teleseismic waveform modelling indicating origin time, location and moment tensor solution from Global CMT project (Dziewonski et al., 1981; Ekström et al., 2012). Half duration (Hdur) and time shift (Tshift) are given in seconds.

the case in P-wave receiver function analysis, the signature of the LAB is hampered by the Moho free-surface multiples, which makes its identification challenging.

4.4 Parametric analysis of teleseismic FWI

4.4.1 Experimental set-up

For all the experiments involved in the parametric study, we perform FWI for the same nine events as those used for the CIFALPS experiment (Fig. 4.4a). We refer the reader to Beller et al. (2017, Their table 1) for the precise location of the nine events used for FWI. We performed a multi-scale FWI (Bunks et al., 1995) by successive inversion of datasets of increasingly high-frequency content. We have generated three datasets by low-pass filtering the raw data with cut-off periods of 20, 10 and 5s, respectively.

Unlike the real CIFALPS data application, the temporal source signature is assumed to be known and hence is not estimated during FWI. We do not apply any time-dependent preconditioning (namely, time windowing or amplitude balancing) on the data, but we normalise each observed and computed common-event gathers by the maximal amplitude of the vertical component of the observed data to equally balance their contribution in the gradient of the misfit function. We use the l-BFGS optimisation algorithm (Nocedal, 1980), a quasi-Newton approach that recursively estimates the action of the inverse Hessian on the gradient, without other additional preconditioning on the gradient. The optimization parameters are the logarithm of the physical parameters to further balance the parameters in depth. A mild regularization is applied by smoothing the gradient with a Gaussian function with a correlation length of 1km in all three Cartesian directions.

In the following numerical examples, each experimental configuration deviates from the one we consider as the reference one in only one respect : subsurface parametrisation, starting model or acquisition geometry. In the reference case, FWI is applied using :

- a (ρ, V_p, V_s) parametrisation,
- a starting model that consists in a smoothed version of the true model (this model will be referred to as the Tomographic model),
- a 2-D regular grid of stations with a receiver-sampling interval of 5 km in the two horizontal Cartesian directions,

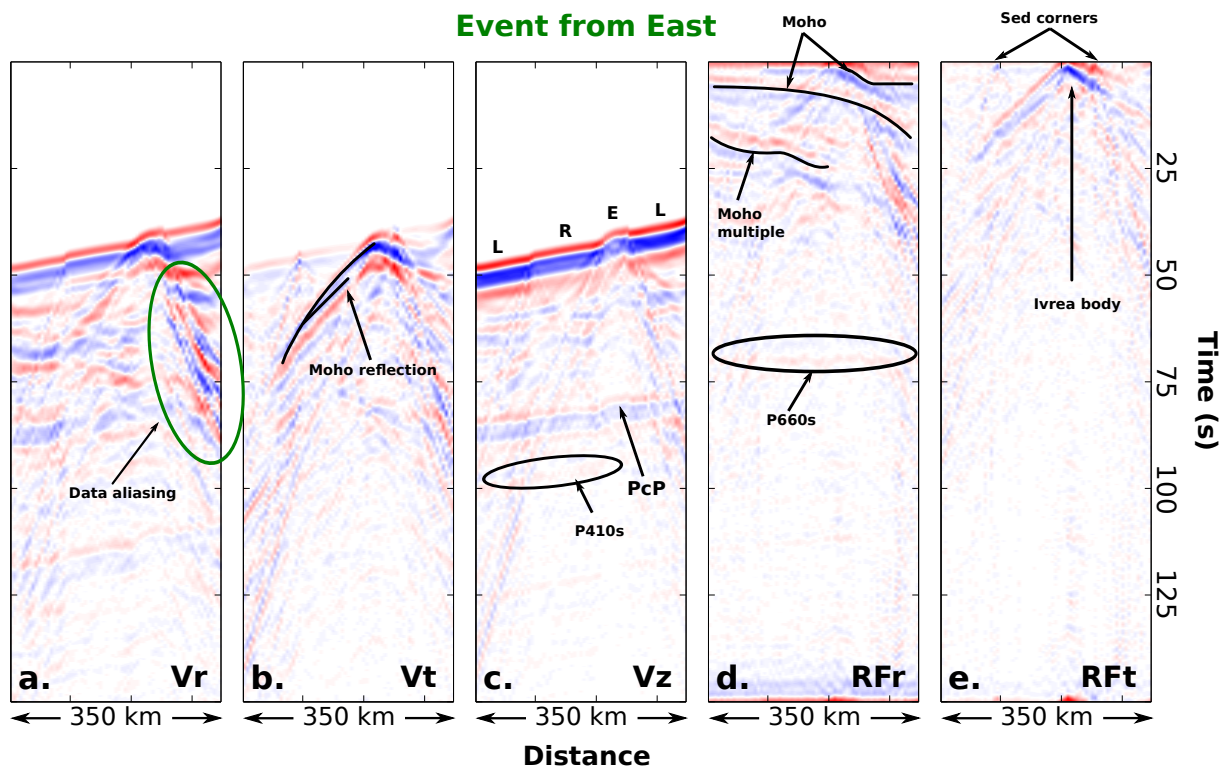


Figure 4.5 – Synthetic data computed in the Alps lithospheric model. From left to right : radial, transverse and vertical components of particle velocities followed by radial and transverse receiver functions corresponding to the event 22 in Beller et al. (2017). E = early arrivals, L = late arrivals, R = reference arrivals. Black ellipses indicate P-to-S wave conversions at upper-mantle reflectors, while the green ellipse highlights data aliasing of steep diffractions generated by the Ivrea body.

as shown in the second line of table 4.2.

For a quantitative insight on the relevance of the FWI models, we computed a correlation-based misfit between a vertical section of the FWI models and the Alps lithospheric model, this vertical section being aligned with the CIFALPS profile (Beller et al., 2017, Their Fig. 3). Even though such a norm may not be the most suitable one (in the sense that it cannot discriminate the footprint of resolution and amplitude errors), it complements in a more quantitative way the qualitative visual assessment of the subsurface models. The statistics of the results of the parametric analysis are outlined in table 4.2.

4.4.2 Which earth parametrisation for teleseismic FWI ?

We first provide some insights on the most suitable subsurface parametrisation for teleseismic FWI. We perform FWI for three different elastic parametrisations, (ρ, λ, μ) , (ρ, V_p, V_s) and (ρ, I_p, I_s) , and we jointly update the three parameter classes involved in the chosen parametrisation. For each of the three inversions, we show the parameters that have been updated by the inversion, referred to as optimization parameters, as well as the parameters of the two other parametrisations, these parameters being built by the non linear recombination of the optimization parameters. These parameters will be referred to as visualization parameters.

The (ρ, V_p, V_s) , (ρ, I_p, I_s) and (ρ, λ, μ) models are shown in Figures 4.6, 4.7 and 4.8, respectively. For each of these figures, we show from top to bottom the FWI results obtained with the (ρ, λ, μ) , (ρ, V_p, V_s) and (ρ, I_p, I_s) parametrisations, respectively. This implies that, for each figure, one row shows the optimization parameters, while the two other rows show the visualization parameters. The presentation of the subsurface models are complemented by the Fig. 4.9, whose histograms show the level of correlation between the reconstructed models generated by each FWI and the true subsurface models.

For a visual assessment of the subsurface models, we focus on the reconstruction of the sedimentary basins, the Ivrea body and the subducting slab. According to these 3 criteria, we conclude that the best subsurface models for the kinematic parameters, namely, V_p , λ and I_p and V_s , μ and I_s , are the I_p and I_s (visualization) models inferred from the (ρ, λ, μ) optimization models (Fig. 4.7(b-c)). This qualitative assessment is supported by the metric outlined in Fig. 4.9, that confirms that these I_s and I_s models show the best correlation with the true ones.

The ρ model inferred from the (ρ, λ, μ) parametrisation seems slightly better resolved than the one inferred from the (ρ, V_p, V_s) parametrisation, at the expense of the noise level. In contrast, the ρ model inferred from the (ρ, I_p, I_s) parametrization is much smoother, that is consistent with the former analysis of the radiation pattern. This visual assessment is supported by Fig. 4.9 in the sense that the ρ model inferred from the (ρ, V_p, V_s) parametrisation shows the best correlation with the true model, while the one inferred from the (ρ, λ, μ) parametrisation shows the poorest one, which might result from the noise impacting this model. However, this noise did not prevent a reliable recombination of the I_p and I_s models from the (ρ, λ, μ) optimization parameters.

One may wonder why the (ρ, λ, μ) parametrisation provided the best subsurface models after recombination of I_p and I_s . If we limit our comparative analysis to the optimisation parameters, μ is the one which shows the best correlation with the true model among the kinematic

m_0	Setup param	acqui	Misfit function (Ss)			Misfit reduction (%)			Model misfit		Iterations			
			f_0	f_{end}	%	20s	10s	5s	m_0 - m_{true}	m - m_{true}	%	20s	10s	5s
Tom	(ρ, λ, μ)	N05	7.57×10^3	4.00×10^2	5.28	1.88	7.87	12.86	3.89×10^7	2.73×10^7	70.33	44	47	45
Tom	(ρ, V_p, V_s)	N05	7.57×10^3	8.95×10^2	11.82	1.51	29.38	23.31	3.89×10^7	2.51×10^7	64.76	42	14	36
Tom	(ρ, I_p, I_s)	N05	7.57×10^3	1.17×10^3	15.46	3.32	8.22	26.29	3.89×10^7	3.22×10^7	82.98	44	42	19
Smooth	(ρ, V_p, V_s)	N05	1.64×10^4	1.07×10^3	6.41	1.72	19.4	18.84	4.59×10^7	2.82×10^7	61.37	34	41	31
PREM	(ρ, V_p, V_s)	N05	1.61×10^4	1.01×10^3	6.27	1.56	19.04	21.58	4.66×10^7	3.17×10^7	68	35	11	40
Gradient	(ρ, V_p, V_s)	N05	5.84×10^4	1.73×10^3	2.96	1.58	15.08	18.44	1.33×10^8	5.76×10^7	43.15	35	39	37
Tom	(ρ, V_p, V_s)	N10	2.04×10^3	2.39×10^2	11.72	1.7	18.12	16.04	3.89×10^7	2.90×10^7	74.7	36	44	29
Tom	(ρ, V_p, V_s)	N25	4.00×10^2	1.68×10^2	42	1.16	14.97	75.34	3.89×10^7	2.92×10^7	75.24	38	14	2
Tom	(ρ, V_p, V_s)	N50	9.13×10^1	3.79×10^1	41.51	2.16	12.14	48.97	3.89×10^7	4.33×10^7	111.49	35	13	4
Tom	(ρ, V_p, V_s)	L05	3.21×10^2	2.58×10^2	67.07	2.36	24.16	100	3.89×10^7	4.87×10^7	125.19	44	11	0
Tom	(ρ, V_p, V_s)	L10	1.64×10^2	1.10×10^2	80.37	2.35	7.8	93.22	3.89×10^7	4.80×10^7	123.39	44	42	2
Tom	(ρ, V_p, V_s)	CIF	6.43×10^2	1.37×10^2	21.31	7.78	11.34	20.79	3.89×10^7	5.11×10^7	131.36	30	35	44

Table 4.2 – Statistics of the FWI applications. **Setup** : experimental set-up in terms of initial model (m_0), optimisation parameters ($param$) and acquisition geometry ($acqui$). **Misfit function (Ss)** : initial/final data misfit function and data misfit function reduction in % for the last multi-scale step (cut-off period : 5s). **Misfit reduction** : data misfit reduction in % for each frequency band (20s, 10s and 5s). **Model misfit** : initial/final correlation-based model misfit and misfit reduction in %. **Iterations** : number of FWI iteration for each frequency band (20s, 10s, 5s).

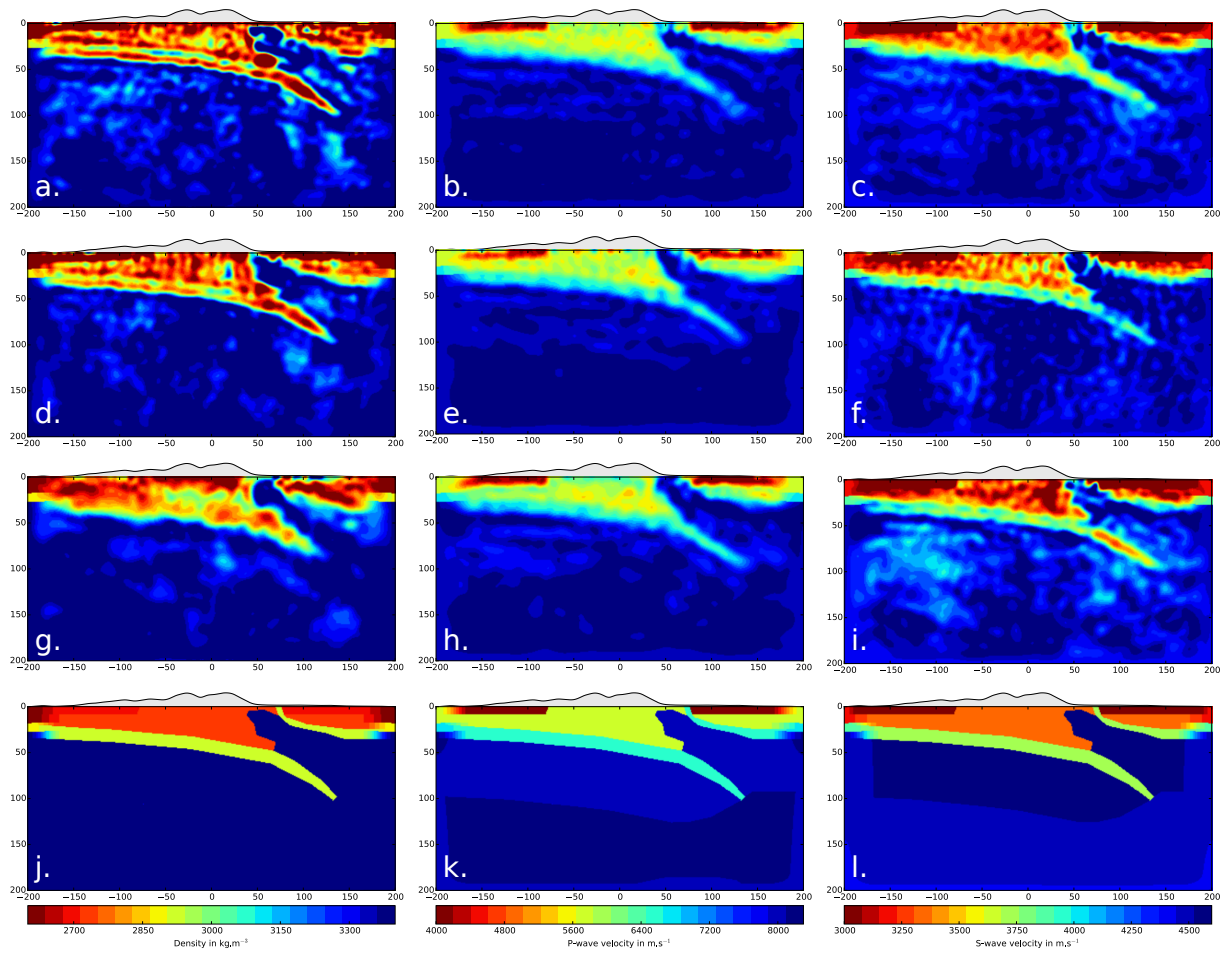


Figure 4.6 – Subsurface parametrization assessment. From left to right are shown the ρ , V_p and V_s models inferred from the FWI reconstruction of (a,b,c) (ρ, λ, μ) , (d,e,f) (ρ, V_p, V_s) and (g,h,i) ρ, I_p, I_s). The last row (j,k,l) shows the true (ρ, V_p, V_s) models for comparison. The optimization parameters are shown on the second row.

parameter more closely related to shear waves (V_s , I_s , μ), (Fig. 4.9). This might result from the key contribution of the forward-scattered P-P mode during the long-wavelength reconstruction of μ (Fig. 4.3c). A second reason may be related to the high-wavenumber content of the ρ model obtained with the (ρ, λ, μ) parametrisation (Fig. 4.3a). When I_p is inferred from ρ and μ by multiplication ($I_p = \sqrt{\rho \mu}$), it is likely that the high wavenumber content of ρ has complemented the small-to-intermediate wavenumber content of μ to focus a broadband I_s model. This improved resolution of the I_p and I_s models relative to the λ and μ models is quite visible in the reconstruction of the Ivrea body and subducting slab (compare Fig. 4.7(b-c) with 4.8(b-c)). A similar line of thought would explain why the sedimentary basins in the I_p and I_s models inferred from the optimization (ρ, λ, μ) parameters (Fig. 4.7(b-c)) are better focused than those of the optimization I_p and I_s models (Fig. 4.7(h-i)). The lower high-wavenumber content of the optimization μ model, in particular at great depths (Fig. 4.8c, subducting slab), relative to the visualization I_p and V_s models might result from the dominant imprint of the P-P forward scattering in the μ gradient relative to that of the back-scattering components. This might have dominantly driven the deep update of μ toward a low-to-intermediate wavenumber reconstruction at the expense of the high wavenumber components. The same comment could apply to the I_p reconstruction from λ and μ . Another issue concerns potential cross-talk between ρ and the two kinematic parameters. The recombination of ρ with a kinematic parameter to build the impedance might help to remove cross-talks as suggested by [Prioux et al. \(2013a\)](#).

4.4.3 Which initial model for teleseismic FWI ?

FWI is a non-linear and ill-posed inverse problem that is generally tackled with local optimisation techniques. In this framework, a requirement for the optimisation algorithm to converge towards a reliable subsurface model is that the initial model should be accurate enough to position the starting point within the attraction basin of the global minimum. Forwardly, this issue is closely related to the cycle skipping phenomenon that occurs in FWI whenever a seismic wiggle obtained from an inaccurate predictive model is shifted by more than half the period with respect to the corresponding observed wiggle ([Virieux & Operto, 2009](#)). If cycle-skipping occurs data, residuals will be back-projected onto a wrong isochrone surface hence conducting FWI towards a local minimum.

A conventional way of mitigating cycle skipping effects is to progressively increase the number of propagated wavelengths during seismic modelling, since the relative time error is inversely proportional to this number. A first data-driven strategy to achieve this goal relies on a frequency continuation approach where datasets of increasing high-frequency content are hierarchically processed ([Bunks et al., 1995](#)). This frequency continuation can be complemented by a time continuation that can be designed by gently increasing the selected time window over iterations. Both data continuation strategies contribute to implement a multi-scale imaging proceeding from the low wavenumber updates to the higher ones.

The frequency continuation approach requires that the recorded data have a sufficiently low frequency content, which is not always the case in exploration seismology although the design of broadband technology has become an active field of research. To overcome this issue, this community has focused its efforts on the upstream building of reliable initial model by tra-

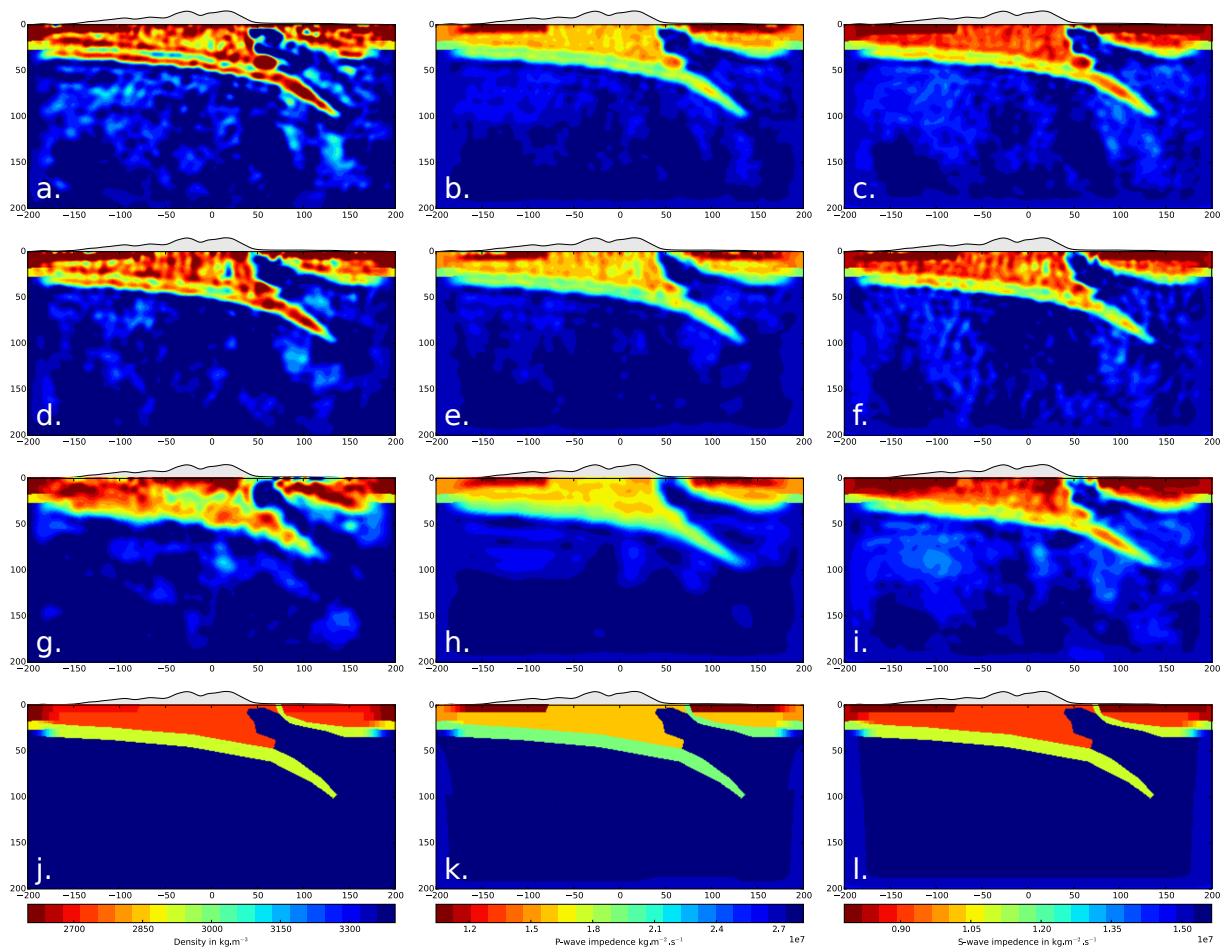


Figure 4.7 – Subsurface parametrization assessment. From left to right are shown the ρ , I_p and I_s models inferred from the FWI reconstruction of (a,b,c) (ρ, λ, μ) , (d,e,f) (ρ, V_p, V_s) and (g,h,i) ρ, I_p, I_s). The last row (j,k,l) shows the true (ρ, I_p, I_s) models for comparison. The optimization parameters are shown on the third row.

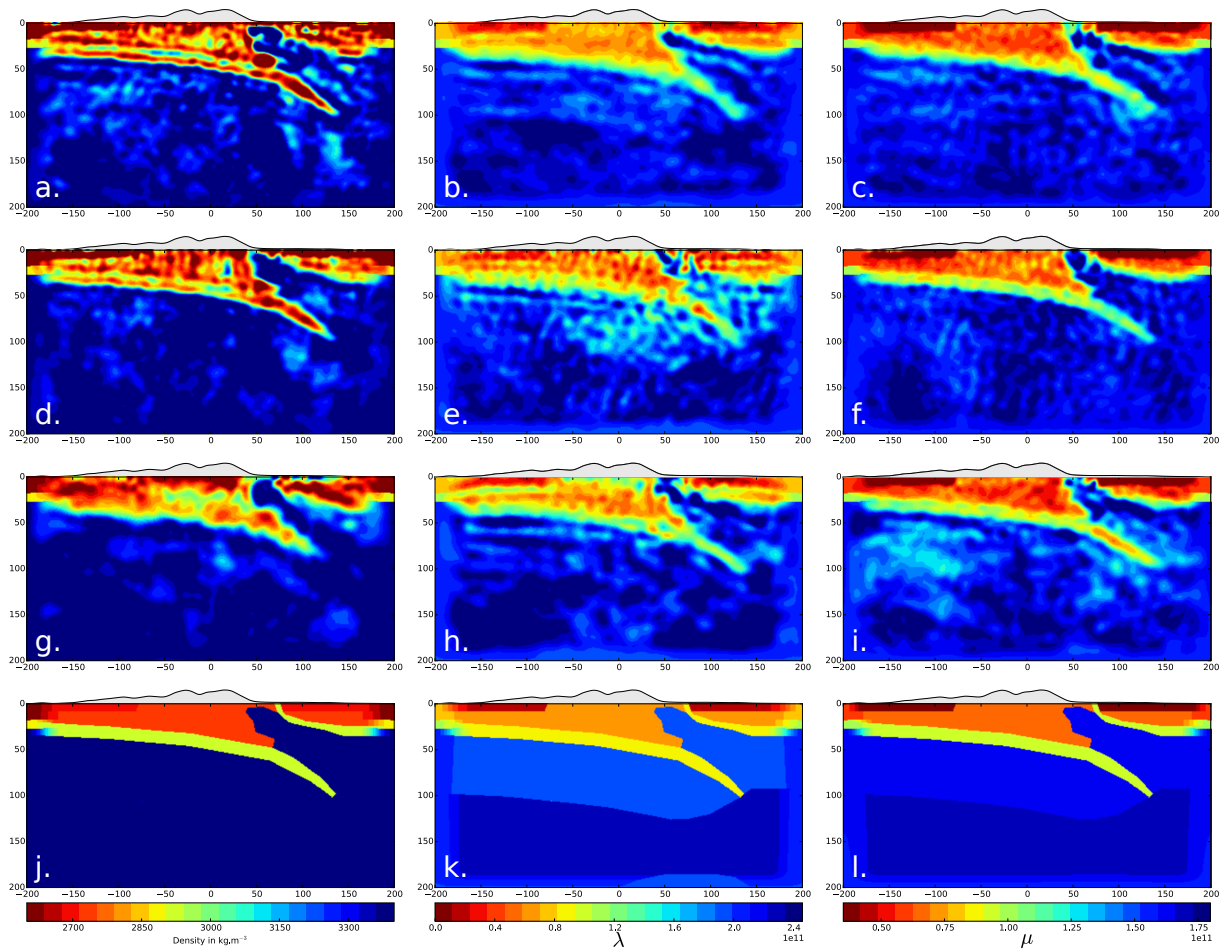


Figure 4.8 – Subsurface parametrisation assessment. From left to right are shown the ρ , λ and μ models inferred from the FWI reconstruction of (a,b,c) (ρ , λ , μ), (d,e,f) (ρ , V_p , V_s) and (g,h,i) (ρ , I_p , I_s). The last row (j,k,l) shows the true (ρ , I_p , I_s) models for comparison. The optimization parameters are shown on the first row.

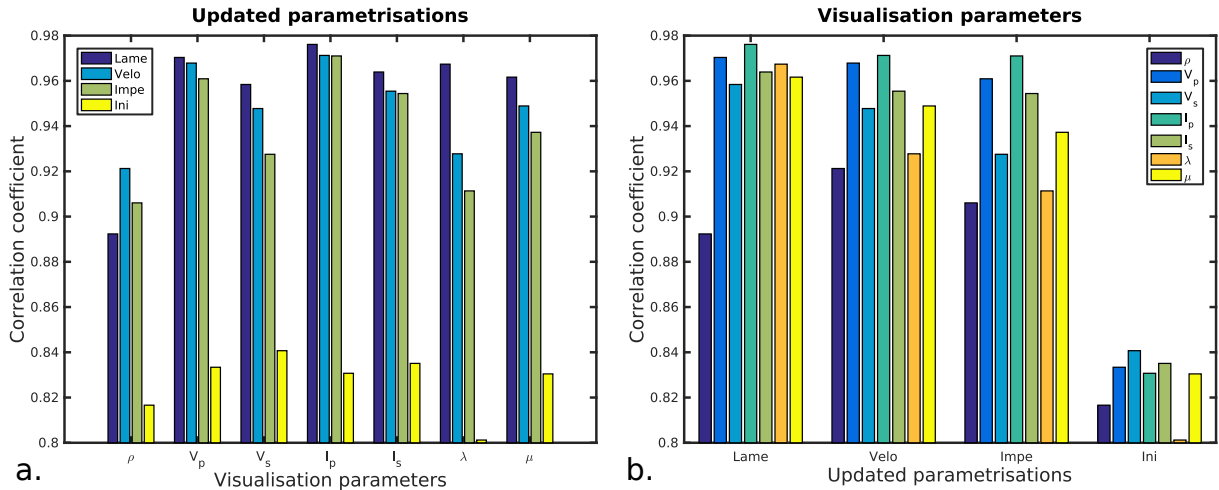


Figure 4.9 – FWI model quality. a) The colours label the subsurface parametrisation used for FWI or in other words the optimisation parameters of the FWI. *Lame*, *Velo*, *Impe* stand for (ρ, λ, μ) , (ρ, V_p, V_s) , (ρ, I_p, I_s) , respectively. *Ini* refers to the initial models for each parameter class. The horizontal axis labels the visualization parameter that has been built a posteriori by non linear recombination of the optimization parameters. The vertical axis labels the correlation-based misfit function between the visualization and the true models. b) Same as (a) with a transposed representation. The colours label the visualization parameter, while the horizontal axis label the optimization parameters.

veltime and slope tomography and migration-based velocity analysis, the design of elaborated misfit functions with better convexity properties or by extending the model space. In the case of teleseismic data, the cycle skipping issue should be prevented by the availability of long period (low frequency) data. Furthermore, reference 1-dimensional global earth model such as PREM, IASP91 or AK135 are generally accurate enough to predict the arrivals of global seismic phases with a confidence interval in between one and two seconds therefore preventing the cycle skipping issue for representative teleseismic periods (20s to 1s).

To verify the previous statement, we analyse the sensitivity of the FWI to four different starting models. A first initial model (Fig. 4.10d-f) is built by Gaussian smoothing of the true Alps lithospheric model with a correlation length of 50 km in the three spatial directions. Despite being smooth, this model yet captures some of the 3D complexity of the target and can be considered as the analogous of an initial model built by first-arrival traveltime tomography. In absence of such a knowledge about the lithospheric velocity and density structure, another option is to resort to available 1-dimensional reference earth models. We investigate here two possible initial models based on the PREM reference model : the layered original one and a smoothed version obtained by Gaussian filtering with a 10km correlation length (Fig. 4.10g-i) and (Fig. 4.10j-l). A last starting model is a simple vertical velocity gradient (Fig. 4.10m-o).

The FWI results for the different initial models are shown in figure 4.11. As expected, the best FWI model has been inferred from the initial tomography-like model (Fig. 4.11d-f and table 4.2). Using the smooth PREM model as an initial model leads to acceptable FWI results, that may be improved by a more aggressive regularization (Fig. 4.11g-i) and table 4.2). The

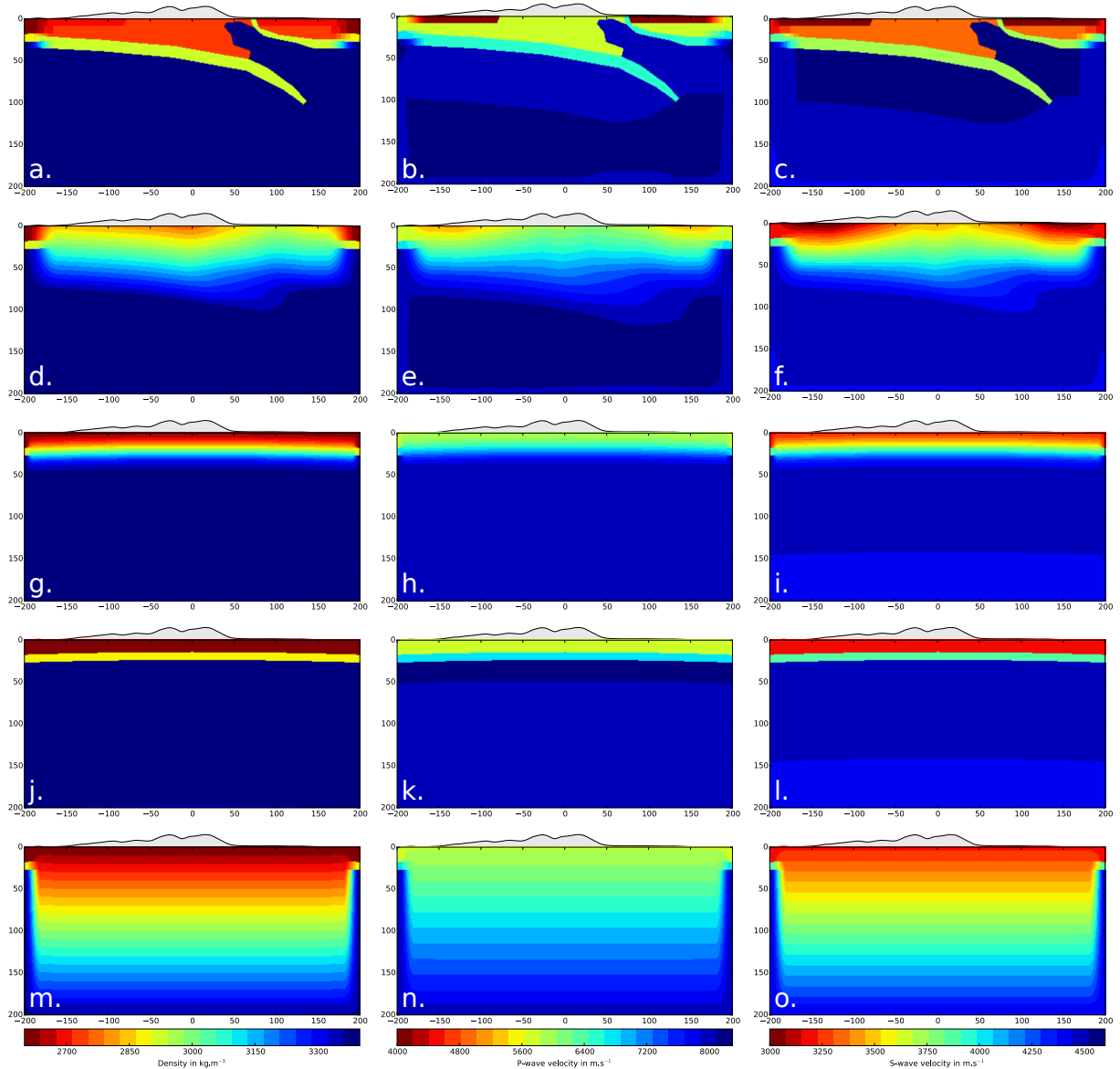


Figure 4.10 – Starting models for FWI. From left to right ρ , v_p and v_s models for : (a,b,c) the input model used to generate synthetic data, (d,e,f) tomographic model computed by smoothing of the input model, (g,h,i) smoothed version of PREM model, (j,k,l) PREM model and (m,n,o) vertical gradient.

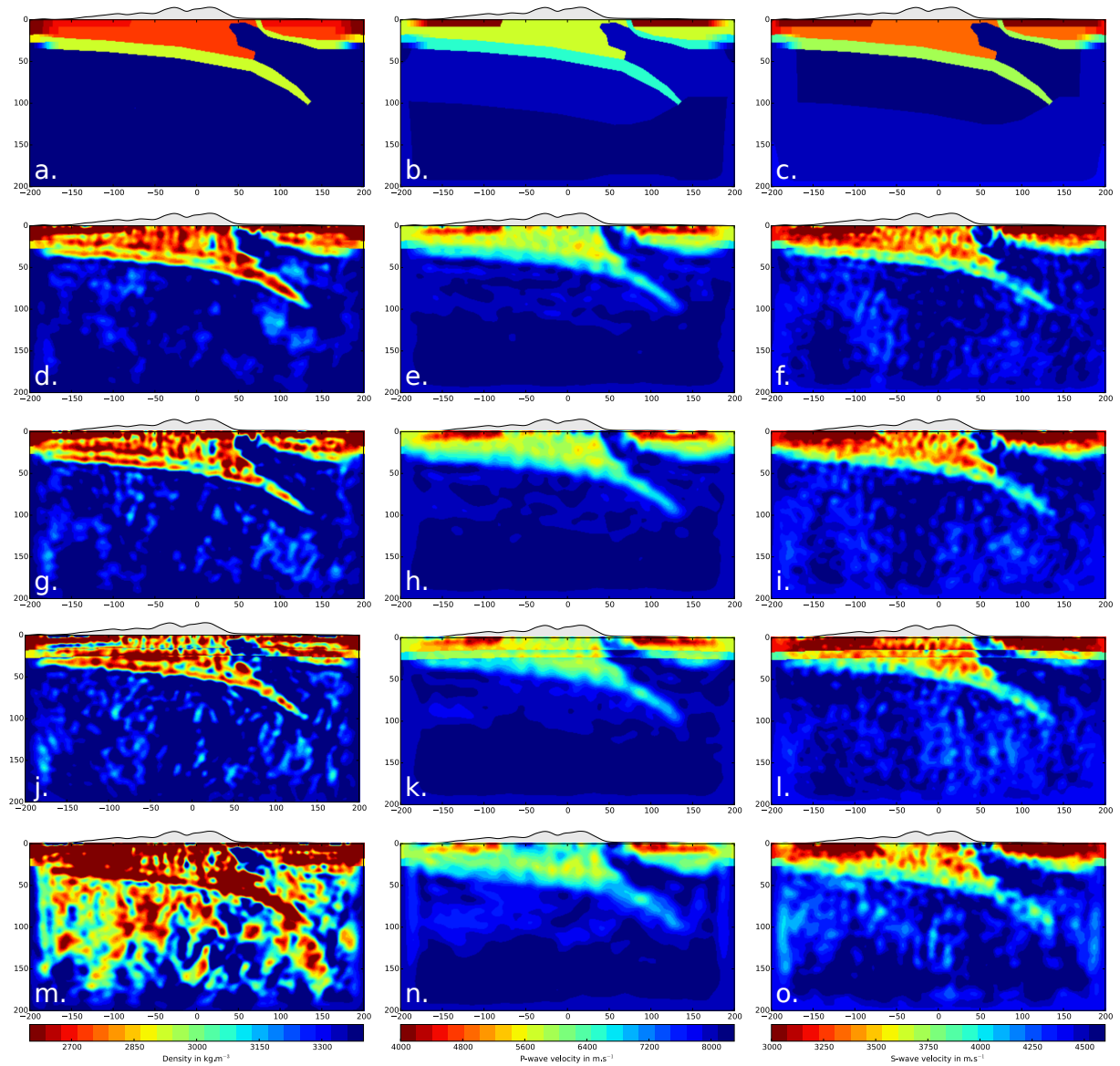


Figure 4.11 – Assessment of initial models. From left to right, ρ , V_p and V_s models. (a-c) True models. (d-o) FWI models inferred from (d-f) the smooth tomography-like model, (g-i) the smoothed version of the PREM model, (j-l) the original PREM model and (m-o) the vertical gradient model.

FWI V_p model is very similar to the one inferred from the tomography-like model, while the ρ and V_s models are slightly noisier. Overall, this result confirms that smoothed reference earth models are accurate enough for teleseismic FWI. When the original PREM model is used as initial model, the footprint of the sharp crustal layering of the initial model overlap the smoother structures updated by FWI (figures 4.11j-l). Indeed, this high-wavenumber footprint cannot be removed or updated by FWI because its spectral content is beyond the resolution limits of FWI. Moreover, this prior high-wavenumber content in the initial model, which can be mispositioned in depth, can drive the inversion toward a local minimum by generating, during seismic modelling, transmitted and reflected waves that are badly positioned in time.

Interestingly the vertical gradient model is accurate enough to produce acceptable velocity models (Fig. 4.11m-o and table 4.2). However, the ρ model shows saturated amplitudes suggesting a higher sensitivity of the inversion to this parameter. A possible reason is that, among all of the tested initial models, the vertical gradient model is the one which does not generate any wave partitioning except at the free surface, hence leading to significant amplitude residuals. In this case, the ρ update might have been driven by the need to remove these amplitude residuals during the early iterations. This higher sensitivity to ρ might also have favoured some leakage between the velocities and ρ updates. This leakage from the velocity to the ρ updates is suggested by the less contrasted reconstruction of the V_p parameter in Fig. 4.11n.

In summary, a good initial model for teleseismic FWI should be smooth enough to avoid involving wavenumber components in the initial models that cannot be easily updated by FWI or which are beyond its resolution limit. On the other hand, our numerical experiments suggest that the initial model should be accurate enough to predict well enough amplitude phenomena in particular if second-order parameter such as density is jointly updated with the wave speeds. If this condition is not satisfied, the update of the density might become exaggeratedly sensitive to these amplitude residuals and might prevent the update of the high wavenumber components of the wave speeds in the upper part of the lithospheric model.

4.4.4 Sensitivity to acquisition design

Linear versus areal acquisitions

High-resolution seismic imaging techniques such as receiver function migration or waveform tomography require dense array of seismic sensors to adequately sample the teleseismic wave field as well as the imaging kernel. On the one hand, seismic experiments have been designed for high-resolution receiver function analysis or scattering migration with dense profiles of stations spaced 5km to 10km apart. Cascadia, Hi-Climb, MASE, CIFALPS or Pyrop experiments are examples of such dense linear acquisitions. On the other hand, areal geometries designed with a coarse grid of stations (30km to 70km interval) have been carried out to perform tomography at a larger scale. USArray, Iberarray or the incoming AlpArray are examples of such experiments, where the receiver grid is either fixed or moving.

To the best of our knowledge, no synthetic experiment has been performed to assess which acquisition geometry is the most suitable one for teleseismic FWI at the lithospheric scale. To fill this gap, we performed FWI for seven station layouts, that have been designed based

upon the two above-mentioned acquisition trends. Four areal acquisitions consist of 2D receiver grids with a station-sampling interval of 5, 10, 25 and 50 km, leading to pools of 2001, 525, 105 and 21 stations, respectively (Fig. 4.12d-o). Two acquisitions consist of one line of stations spaced 5km and 10km apart, (Fig. 4.12p-u), leading to pools of 69 and 35 stations, respectively (Fig. 4.12p-u). Finally, for sake of comparison with the real data case study presented in [Beller et al. \(2017\)](#), we also assess the CIFALPS acquisition, that combines a dense profile of stations spaced 5km apart with a sparse irregular spread of permanent stations (Fig. 4.12v-x).

The FWI models for each experiment are shown in Fig. 4.12. As expected, the denser areal geometry gives the best resolved models for each of the three parameter classes. The coarsening from 5 to 10 km of the receiver grid does not significantly degrade the reconstruction of the synthetic model neither in terms of spatial resolution nor in terms of noise (Fig 4.12d-i). Furthermore, the 25 km acquisition still provides reliable reconstruction of the synthetic model even though its resolution is clearly more limited. Here, this poorer resolution results because the FWI stopped after two iterations during the processing of the last frequency band (see table 4.2). For a 50 km acquisition, despite a quite low resolution, the main structural heterogeneities of the synthetic model are still recognisable. However, a strong acquisition footprint in the form of energetic velocity anomalies at the station positions every 50 km is quite visible in the P-wave velocity model (Fig. 4.12n). Overall, we show a progressive degradation of the spatial resolution as the receiver sampling becomes coarser. This results because the range of scattering and dip angles sampled by the source-receiver layout becomes narrower and sparser as the station interval increases. The narrowing of the scattering angle illumination will limit the resolution of the reconstruction perpendicularly to the dip, while the coarsening of the scattering-angle illumination can generate wraparound of the structures if narrow-frequency bands are processed. The scattering-angle illumination as a function of the dip angle (noted ϕ in the current study, equation 4.9), is illustrated in [Rondenay et al. \(2005, Their Fig. 3\)](#) for a linear teleseismic acquisition and two different station samplings.

Compared to areal geometries, linear acquisitions drastically affect the ability of FWI in recovering well-resolved lithospheric models. Although six of the nine earthquakes have back-azimuths approximately aligned with the receiver line (Fig. 4.4a), the FWI models show a degraded resolution of the V_p and V_s models as well as incomplete reconstruction of density at great depths. This results because the linear deployment illuminates more incompletely (in terms of bandwidth and sampling) the scattering and the two dip angles than an areal geometry, even for the image points located in the vertical plane of the receiver line. As an illustrative example, any non-cylindrical structure cross-cutting the vertical plane of the profile such as the Ivrea body will generate out-of-plane scattering which cannot be recorded without an areal geometry. The limited number of off-line stations available in the real CIFALPS acquisition helps increasing the spatial resolution at the expense of the signal-to-noise ratio, which might have been degraded due to spatial aliasing (Fig. 4.12v-x). It is worth remembering that, for sake of consistency between the different experiments, we did not apply significant regularisation during FWI. When dealing with such sparse acquisition, more aggressive smoothing constraints along the spatial directions that are not sampled well by the acquisition geometry would improve the results as suggested by the results of [Beller et al. \(2017\)](#) on the CIFALPS real data.

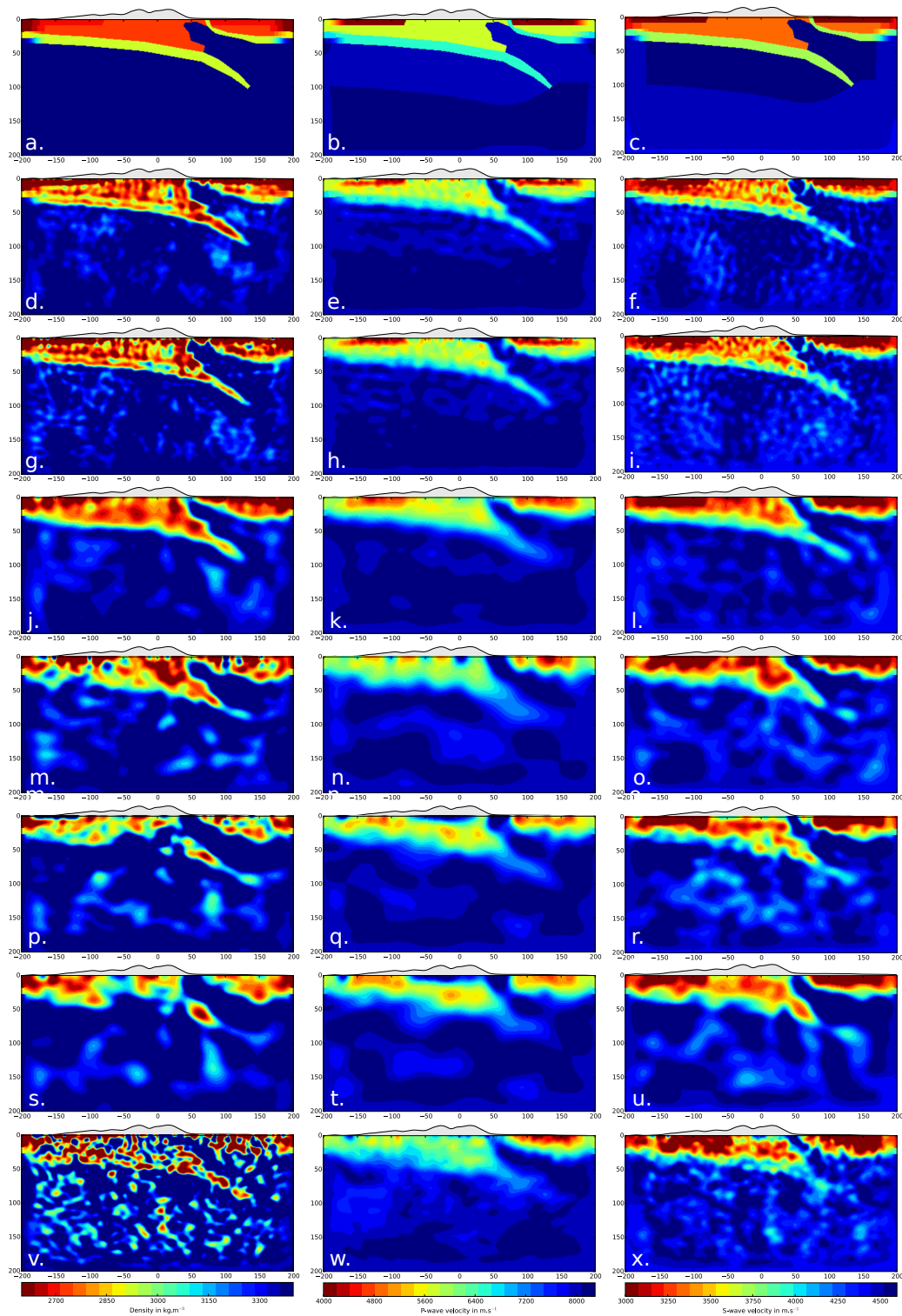


Figure 4.12 – Effect of acquisition geometry on FWI reconstruction. From left to right ρ , V_p and V_s models for : (a,b,c) the input model used to generate synthetic data ; (d,e,f), (g,h,i), (j,k,l) and (m,n,o) results using a 2D array of station with respective inter-station spacings of 5, 10, 25 and 50 km ; (p,q,r) and (s,t,u) results using a linear acquisition with respective spacings of 10 and 5 km ; and (v,w,x) the results considering the CIFALPS acquisition as used in (Beller et al., 2017).

Acquisition footprint and aliasing

The noise in the recovered FWI model might be due to the conjunction of two factors. Firstly, FWI models are polluted by acquisition footprint in the form of high-amplitude perturbations at the station positions, this acquisition footprint being particularly obvious for the 50km station grid. These localized perturbations result from the high amplitudes of the adjoint wavefields at the station positions and the arrival time of the incident planar wavefields at the station. Therefore, the product of these two wavefields generate some bright spots at the station positions if no efficient correction for these amplitude effects are applied. The Hessian, in particular its diagonal coefficients, is expected to correct for them. However, the poor approximation of the Hessian provided by l-BFGS during the first FWI iterations might have contributed to leave a significant imprint of this amplitude singularities in the FWI models. These high-amplitude perturbations can behave as strong diffractors that propagate noise in the FWI models over iterations.

Second, aliasing effects may pollute the recovered image at high-frequencies therefore adding noise in the reconstructed image. In the framework of Kirchhoff migration, Lumley et al. (1994) defined three kinds of aliasing : data, image and operator aliasing.

Data aliasing occurs when the recorded wave field is not properly sampled in one or more dimension according to the Nyquist theorem. Spatial aliasing can occur when the apparent wavelength of the incoming wave field at the surface is not sampled adequately by the receiver grid. Even though the 5s-period incident P wavefield satisfies the sampling criterion, sharp small-scale structures near the surface can behave as energetic diffractors that scatter energy along steep diffraction hyperbola in the time-distance domain. In this study, the edges of the Ivrea body and the pinch-out of the sedimentary basins generate such energetic diffractions that undergo aliasing even for a 5km station-sampling interval (Fig. 4.5, green ellipse). An obvious means to mitigate data aliasing is to limit the inversion to low-frequency data at the expense of the spatial resolution of the imaging. Another strategy would consist in re-sampling the acquisition by interpolation of seismic traces (Wilson & Guitton, 2007).

Image aliasing occurs when the mesh, that parametrizes the subsurface target, is too coarse to properly sample the output of the imaging. In FWI, the maximum resolution is half a wavelength and the subsurface should be meshed accordingly, that is with at least four grid points per wavelengths in a finite-difference framework. In our case, we design the hexahedral mesh with at least five degrees of freedom per wavelength according to the above-mentioned requirement.

Operator aliasing occurs when the integral operator that transforms the data in model updates is not properly sampled. This can happen even if the data are not aliased. In the framework of inverse scattering techniques, Rondenay et al. (2005, their Fig. 4) illustrated how the sampling of the stations impacts upon the sampling of the weighting function applied to the scattered data during their weighted stack along diffraction hyperbolas. This migration operator becomes increasingly smooth with depth, which means that operator aliasing is mostly generated at shallow depths. Undersampling this weighting function leads to artefacts in the form of high wavenumber "speckle" in the image. Considering that the FWI gradient and the migration kernel obey the same imaging principle, it is likely that this kinds of aliasing has affected our FWI results.

Other sources of noise can result from the poor fold of teleseismic acquisitions. The subsurface imaging in FWI relies on the Huygen's principle : any continuous structure in the subsurface is imaged by the interference of the isochrone surfaces generated by back-projection of data residual samples. These interferences are constructive at the locus of the missing subsurface structures and destructive off this locus. If the stations sampling is too coarse, the destructive interferences off the locus of the missing heterogeneities is not efficient enough to cancel out the undesired contributions of the isochrones.

Beyond the receiver sampling, the source distribution in teleseismic imaging can potentially generate aliasing noise and degrade spatial resolution. In a parametric study of 2.5D frequency domain teleseismic FWI, Pageot et al. (2013) have shown the footprint of the source distribution and sampling on the quality of the imaging. Narrowing the range of incidence angles sampled by the events contribute to narrow the wavenumber coverage accordingly and hence, degrade the spatial resolution of the imaging. However, the source sampling for a fixed range of incidence angles has little effects on the imaging quality provided that the temporal frequency dimension is finely sampled. This results because a planar wave field leads to a uniform spatial sampling of the subsurface at the expense of the angular illumination. More serious aliasing artefacts occurred when a 2D lithospheric target is illuminated from only one side, leading to significant gaps in the dip angle coverage. In 3D applications, this aliasing may be more prominent since two dip angles need to be sampled.

In the current study, it is not obvious to decipher which of the above-mentioned effects is responsible for the pollution of FWI reconstruction.

Resolution analysis

We assess now more precisely the footprint of the receiver sampling on the spatial resolution of the FWI. We performed two spike tests, the aim of which is to image a grid of small spherical inclusions spaced 20km apart. These inclusions are superimposed on the smooth tomographic lithospheric model, which is used as a starting model for FWI. The diameter of the anomalies is 20km and the amplitude of the perturbations are $\pm 100 \text{ m.s}^{-1}$ and $\pm 100 \text{ kg.m}^{-3}$ for wave speeds and ρ , respectively, with an alternating sign from one inclusion to the next. We perform this resolution analysis for two station deployments : the first consists of a grid of stations sampled 25km apart (Fig. 4.13), while the second one involves a line of stations spaced 5km apart (Fig. 4.14). The number of stations involved in the two acquisitions is roughly equivalent. In Figures 4.13 and 4.14, we extract from the FWI models a vertical section along the receiver line of the linear acquisition and several horizontal slices at increasing depths.

For the areal geometry, FWI succeeds in reconstructing the three parameters at shallow depths. The resolution and the amplitudes of the V_p updates degrade with depth much faster than those of V_s and ρ . This might result because all of the scattering modes contribute to the update of V_s and ρ unlike V_p hence leading to a higher fold during the V_s and ρ updates. Moreover, the P-S reflection coefficient at the free surface is higher than the P-P one (Pageot et al., 2013, Their Fig. 3), making the doubly P-S-S reflections from the lithospheric reflectors more energetic than the P-P-P counterparts. Increasing the fold is quite important when the imaging mostly rely on the contribution of second-order back-scattered waves which is likely the case

at these depths. Moreover, the signature of the forward scattered wave field diffracted from the deep inclusions might have been healed when arriving at the surface, hence contributing to degrade the V_p resolution at depth. We also show alternating positive and negative perturbations in the ρ vertical section at large depths. This distortion of the inclusion shape highlight limited bandwidth effects, which result because ρ is mainly updated from back-scattered waves according to its radiation pattern. The range of scattering angles sampled by the acquisition decreases with depth hence impacting the vertical resolution accordingly. More surprisingly, the horizontal resolution is not impacted significantly by the limited back-azimuthal coverage provided by the nine events. In other words, the reconstruction of the inclusions in the horizontal planes do not show any preferential direction except below 130km depth in the V_p reconstructions. This can be however explain by the fact that spherical inclusions behave as point diffractors that scatter waves in all the directions. These scattered waves are recorded by the areal geometry along all of the horizontal directions hence providing a significant coverage of the horizontal components of the wavenumber vectors.

For the line geometry, the reconstruction of the three parameters has been significantly degraded hence confirming the previous FWI results on the Alps model. In the vertical section, only the shallow V_p , V_s and ρ inclusions are reconstructed at 20km depth, while the inversion failed to reconstruct them at greater depths. Comparing these results with those obtained with the areal geometry highlights the key contribution of off-line stations to reconstruct the vertical wavenumber components of the inclusions at great depths. Again, this can be easily understood by the fact that the spherical inclusions behave as point diffractor that scatter waves in all spatial directions. Therefore, the wave field scattered by the in-line inclusions can be recorded by the full station layout leading to a broad scattering-angle and dip illumination and a high fold. Note however that it is not guaranteed that more cylindrical geological structures would emphasize the same add-value of the areal geometry relative to the linear one. A second possible factor that explains the rapidly decreasing sensitivity in depth of the imaging is that the surface waves that are generated by conversion of the incident P wave onto Rayleigh waves near the surface might have a significant contribution in both the vertical and horizontal resolution. The wavelength of the Rayleigh wave at 0.1Hz (30km) is consistent with the depth resolution shown in Fig. 4.14a.

Concerning horizontal resolution, the depth slices of the ρ and V_s inclusions show a good resolution at 10km depth, while only the inclusions along the receiver line are reconstructed in the V_p slices. In both case, we see a clear smearing along the cross-line (y) direction, which reflects a deficit of wavenumber coverage along this dimension. The failure of the V_p reconstruction off the receiver line might result from the small amplitudes of the P-P waves that are back-scattered towards the receiver line (back-scattered waves are those which dominantly sample the horizontal components of the wavenumber vectors), while V_s and ρ reconstructions might have benefited from the various off-line back-scattering modes generated by the inclusions and the free surface. As for the vertical resolution, a second factor which can explain the contrasted horizontal resolution in the shallow updates of V_p compared to V_s and ρ is related to the contribution of Rayleigh waves, which are more sensitive to V_s perturbations.

2D array - 25 km

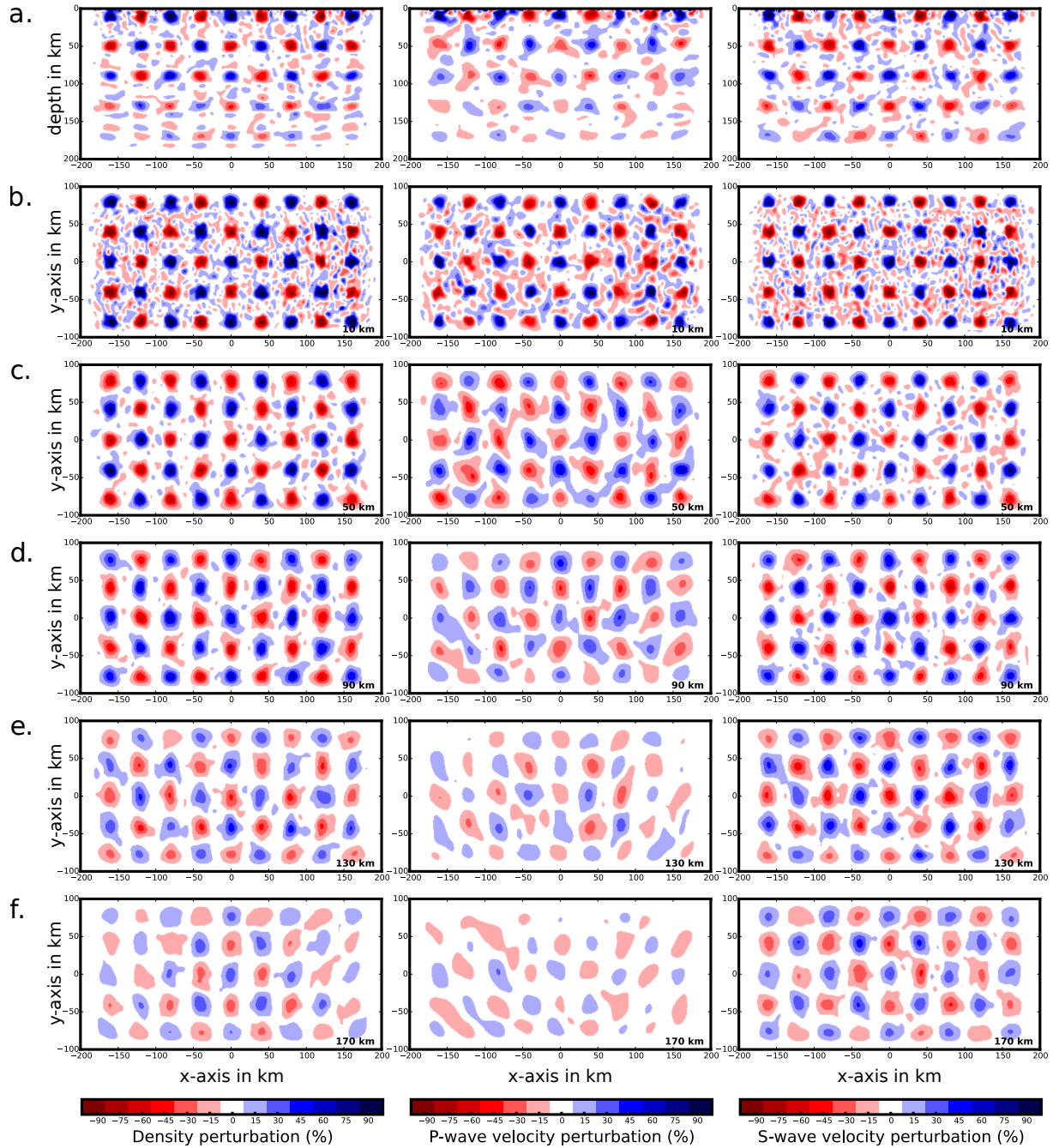


Figure 4.13 – Resolution test for the areal acquisition. a) Vertical section in the middle of the receiver grid. (b-f) Depth slices at 10km, 50km, 90km, 130km and 170km depth. From left to right, reconstructed ρ , V_p and V_s inclusions.

1D array - 5 km

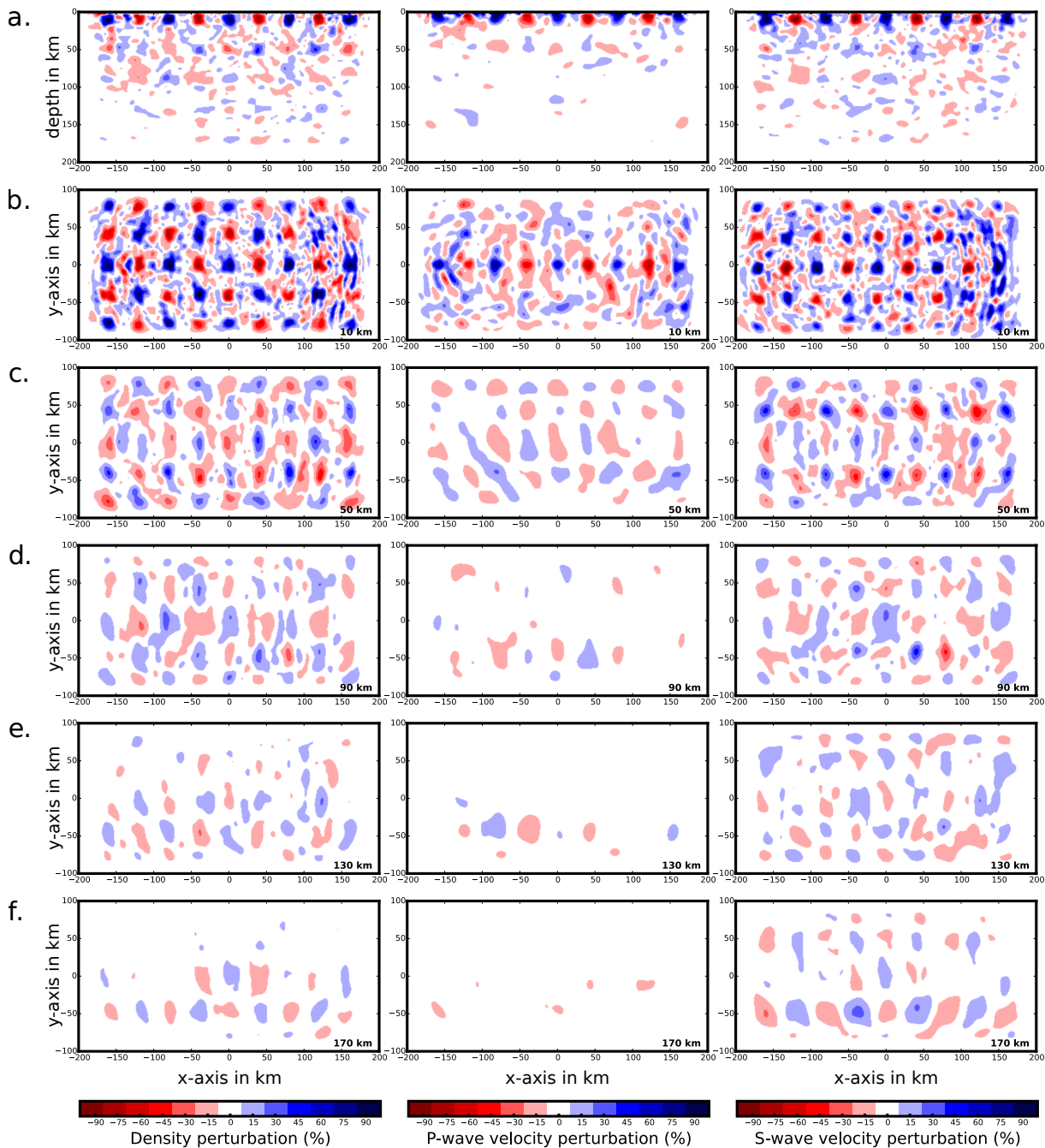


Figure 4.14 – Resolution test for the linear acquisition. a) Vertical section along the receiver line. (b-f) Depth slices at 10km, 50km, 90km, 130km and 170km depth. From left to right, reconstructed ρ , V_p and V_s inclusions.

4.5 Conclusions

We have analysed several theoretical and experimental factors that have a significant influence on teleseismic FWI for lithospheric imaging. We have designed a realistic lithospheric model of the Western Alps and used nine events collected during the CIFALPS experiment to assess teleseismic FWI with a realistic experimental set-up.

The first open question that has been addressed deals with the best subsurface parametrization for teleseismic FWI in the elastic isotropic approximation. We have concluded that (ρ, λ, μ) were the most suitable optimization parameters based on the analysis of the radiation patterns. Recombining a posteriori these three parameters into P and S impedances provide the lithospheric reconstruction that correlates best with the true model. During this recombination, density through the multiplication with the Lamé parameters, contribute to better focus some key structures such as shallow sedimentary basin, the Ivrea body and the subducting slab by injecting high wavenumber components. However, (ρ, V_s, V_s) or (ρ, I_p, I_s) optimization parameters also provide reliable and close results without recombination. The conclusions drawn from this synthetic experiment will need to be validated against real data applications where the presence of noise might hamper reconstruction of second-order parameters such as density.

A second important conclusion is that 1-D global reference models such as PREM provide reliable initial models for teleseismic FWI after smoothing, that is necessary to remove the imprint of the layering. The low frequency content of the teleseismic data makes the inversion almost immune to cycle skipping. However, too strong amplitude residuals generated by too simple initial models, such as vertical gradient models, can make the inversion to over-update the density parameter.

Designing reliable acquisition for teleseismic FWI is a key issue since the best trade-off before station sampling and layout spread should be found for a given pool of stations. We have tested two main acquisition trends corresponding to areal geometries designed with coarse regular grids of stations and dense linear acquisitions. Our results clearly indicate that areal geometry should be designed to improve both the penetration in depth of the imaging and the horizontal resolution. The main reason is that, compared to a linear acquisition, an areal geometry will tremendously increase the fold and the wavenumber coverage with which a diffractor point will be imaged in the subsurface

Acknowledgments

This study was partially funded by the SEISCOPE consortium (<http://seiscope2.osug.fr>), sponsored by CGG, CHEVRON, EXXON-MOBIL, JGI, SHELL, SINOPEC, STATOIL, TOTAL and WOODSIDE. This study was granted access to the HPC resources of SIGAMM infrastructure (<http://crimson.oca.eu>), hosted by Observatoire de la Côte d'Azur and which is supported by the Provence-Alpes Côte d'Azur region, and the HPC resources of CINES/IDRIS under the allocation 046091 made by GENCI.

Troisième partie

Application de la FWI aux données télésismiques

Chapitre 5

Les Alpes Occidentales

Sommaire

5.1	Présentation de la chaîne alpine	188
5.2	Histoire géodynamique des Alpes	192
5.3	Synthèse des connaissances géophysiques	194
5.3.1	Structures crustales	194
5.3.2	Structures mantelliques	199
5.4	Conclusions et problématiques actuelles	201

Introduction

La chaîne Alpine, particulièrement sa partie Occidentale, est historiquement un domaine privilégiée d'études géologiques. Depuis les premières études effectuées par les naturalistes des 18^e et 19^e siècles, l'orogène alpin n'a cessé d'être documentée ce qui a permis de développer un nombre important de concepts fondamentaux en géologie. La documentation et les données, tant géologiques que géophysiques, disponibles sont ainsi conséquentes et les synthétiser représente un exercice difficile. D'un point de vu géodynamique, les Alpes Occidentales sont un exemple typique de *collision continentale* résultant de la convergence des marges européenne et africaine suite à la fermeture d'un océan aujourd'hui disparu : la Téthys ligure.

Dans ce chapitre, nous présentons les connaissances géologiques et structurales essentielles pour la compréhension de l'histoire tectonique et géodynamique des Alpes occidentales ainsi que l'interprétation des images tomographiques. Dans un second temps, nous détaillons l'état actuel des connaissances géophysiques sur l'arc Alpin, notamment, leurs apports pour la compréhension des structures lithosphériques alpines et le comportement géodynamique de la chaîne. Enfin, nous discutons quelques problématiques actuelles concernant la dynamique et la structure des Alpes occidentales.

Ces informations sont particulièrement importantes pour interpréter les résultats de l'inversion des formes d'ondes des données télésismiques de l'expérience CIFALPS menée dans les Alpes occidentales et que nous présenterons dans le prochain chapitre.

5.1 Présentation de la chaîne alpine

Les Alpes font partie d'un système de montagnes, la ceinture alpine aussi appelée chaîne péri-téthysienne, qui constitue un ensemble d'orogènes plus ou moins continu s'étendant de Gibraltar dans sa partie la plus occidentale jusqu'à l'Indonésie dans sa partie orientale en passant par l'Himalaya. La ceinture alpine, formée durant le Mésozoïque et le Cénozoïque, présente ainsi une structuration Est-Ouest et est liée à la fermeture de l'océan Téthys et à la dislocation de la Pangée (Lemoine et al., 2000; De Graciansky et al., 2011). Les parties de la ceinture péri-téthysienne situées entre Gibraltar et la Turquie constituent les chaînes péri-méditerranéennes (Fig. 5.1). Ces dernières résultent de l'ouverture puis de la fermeture de bassins océaniques du système téthysien et de la convergence entre les continents africain et européen et principalement de la subduction du continent africain sous le continent européen (Stampfli et al., 2002; Renard et al., 2015).

Les Alpes, au sens strict, forment une chaîne montagneuse d'environ 1000 km de long (Fig. 5.2), allant de Nice jusqu'à Vienne, pour 100 à 400 km de large, recouvrant pas moins de huit pays européens (Allemagne, Autriche, France, Italie, Liechtenstein, Monaco, Slovénie et Suisse). Contrairement à la majorité des chaînes péri-méditerranéennes, les Alpes résultent de la subduction « inversée » de la plaque européenne sous la plaque adriatique, un promontoire de la plaque africaine aussi nommée plaque apulienne (Laubscher, 1991). Cette subduction entraîne la fermeture (suture océanique) des océans ligure, aussi nommé néo-téthys ou océan liguro-piemontais, et valaisan qui engendre la collision continentale entre les marges européenne et



Figure 5.1 – Localisation de la chaîne alpine au sein de la ceinture péri-méditerranéenne. D’après De Graciansky et al. (2011).

adriatique (Dèzes et al., 2004; Ziegler & Dèzes, 2006; Handy et al., 2010). Cette dernière est caractérisée par un édifice de nappes chevauchant l’avant pays européen.

L’architecture de la chaîne, dans son ensemble, est simple : la marge européenne (du côté externe) est recouverte par les unités d’origine océanique qui résultent d’une obduction. Ces unités sont elles-mêmes recouvertes par les unités de la marge opposée de l’Afrique (le promontoire adriatique) qui constituent les unités les plus internes. Ces formations se retrouvent à l’affleurement dans les grandes unités structurales de la chaîne alpine, présentées sur la figure 5.2. D’Est en Ouest, la chaîne peut être subdivisée en trois parties distinctes (De Graciansky et al., 2011) :

- les Alpes orientales, qui font partie de la marge continentale africaine (promontoire adriatique) principalement représenté par les domaines auto- et sud-alpin,
- les Alpes centrales, qui font partie de la marge continentale européenne,
- les Alpes occidentales, allant de Genève jusqu’à Gênes, qui font partie de la marge continentale européenne et dessinent un arc prononcé.

Contrairement aux Alpes orientales et centrales qui présentent une structuration rectiligne Nord-Sud, les Alpes occidentales se singularisent par une courbure et une structuration Ouest-Est.

Les Alpes occidentales sont classiquement subdivisées en deux zones structurales majeures, les zones internes et les zones externes, séparées par un grand chevauchement vers l’Ouest : le *chevauchement Pennique frontal* (Fig. 5.3) qui charrie les zones internes sur les zones externes.

Les zones externes (ou dauphinoises) Situées à l’Ouest du front Pennique, elles sont essentiellement composées d’unités non ou peu métamorphisées rattachées à la marge européenne. On y distingue les chaînes subalpines comportant des roches sédimentaires du Mésozoïque et du Cénozoïque (Chartreuse, Vercors) et les massifs cristallins externes

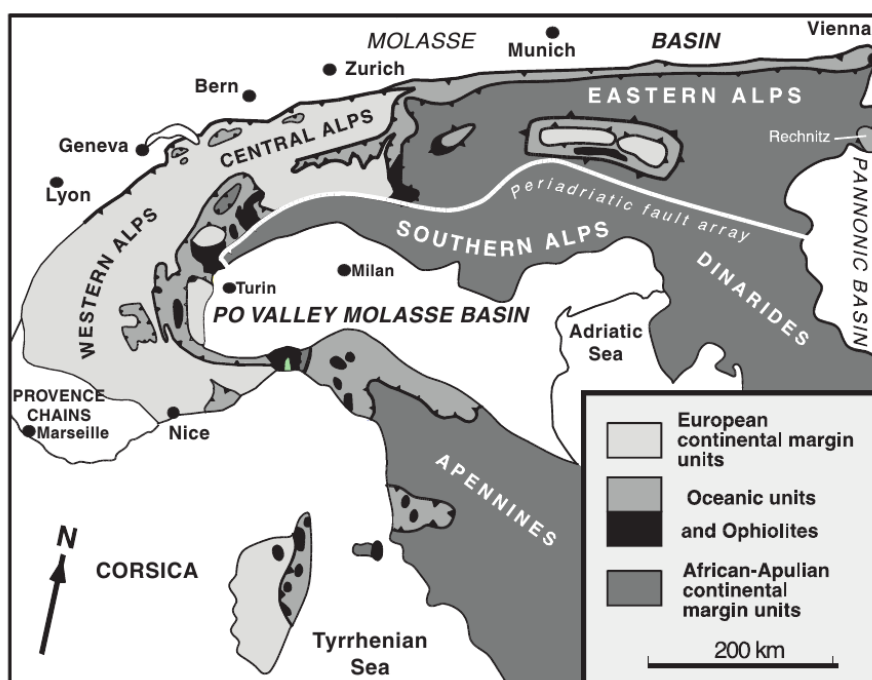


Figure 5.2 – Présentation de la chaîne Alpine et de ses grandes unités structurales. D’après De Graciansky et al. (2011).

(Argentera, Belledonne, Pelvoux, A. Rouges, Mt Blanc) constituées de roches Paléozoïques qui correspondent au sous-bassement de la marge européenne. Elles ont principalement enregistré les étapes de la collision continentale depuis l’Eocène.

Les zones internes (ou Pennique) Situées à l’Est du front Pennique et affectées de manière plus ou moins importante par le métamorphisme alpin. Elles contiennent les zones Briançonnaise (continentale), piémontaise (marge océanique) et liguro-piémontaise (d’origine océanique ex. Chenaillet et Mt Viso). Ces zones présentent des faciès métamorphiques HP-BT (schistes bleus à éclogite) caractéristique des zones de subduction. On y rattache aussi les massifs cristallins internes (Grand Paradis et Dora Maira). Ceux-ci témoignent de la subduction continentale de la marge européenne. En effet, la présence de coesite¹ dans le massif de Dora Maira indique un enfouissement de la marge continentale européenne à des profondeurs de l’ordre de 100 km (Chopin, 1984). La zone interne a tout d’abord enregistré la subduction continentale à l’Eocène moyen et supérieur, puis la collision continentale.

Cette organisation montre que les Alpes sont aussi structurées par un degré croissant de métamorphisme depuis les zones externes vers les zones internes. Cette évolution correspond à une augmentation des conditions de pression et température qui enregistre un gradient de haute pression - basse température (HP-BT) caractéristique des zones de subduction. Le degré croissant de métamorphisme, d’Ouest en Est, suggère que la polarité de la subduction est orientée vers l’Est.

1. Un polymorphe du quartz.

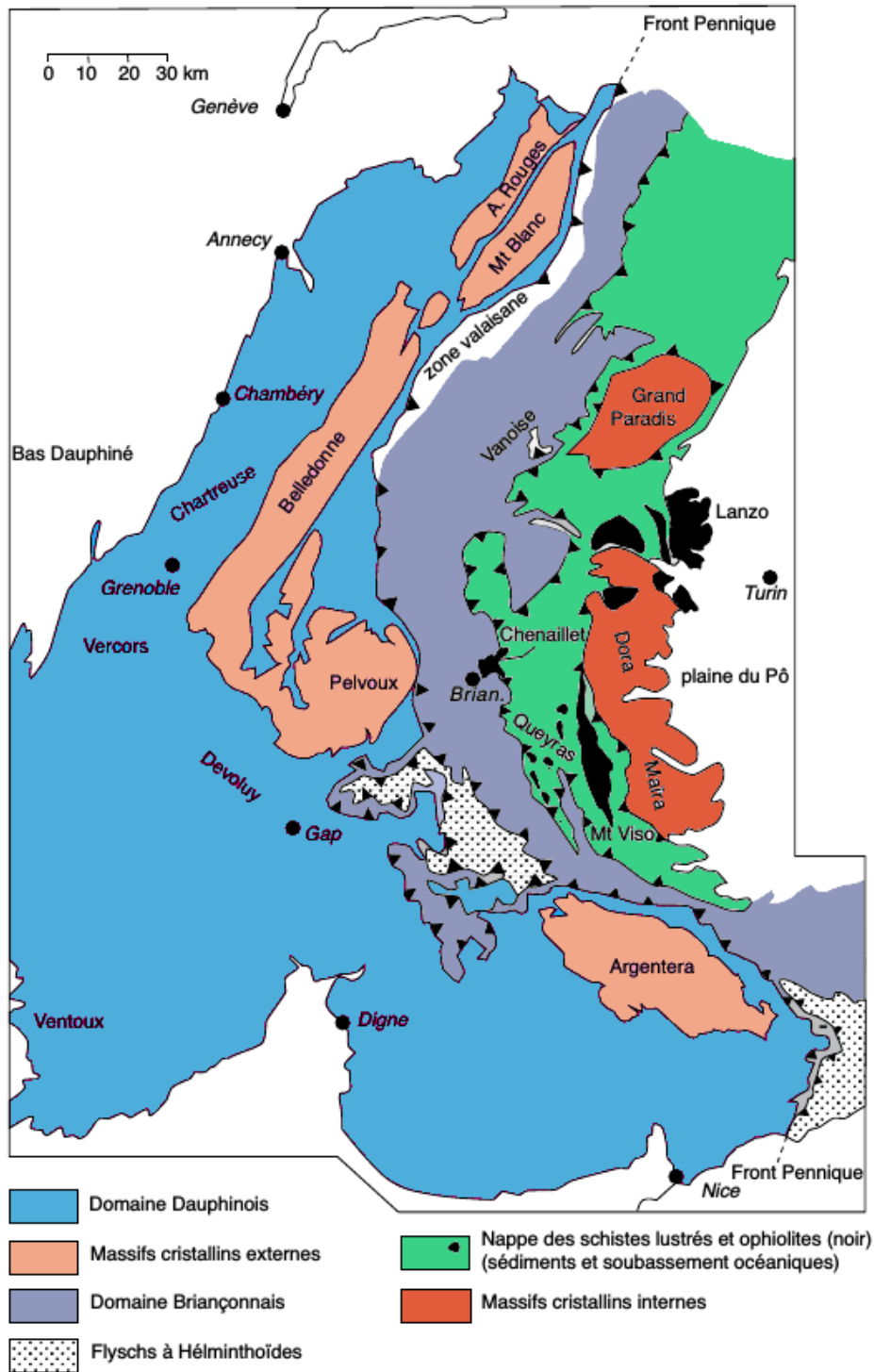


Figure 5.3 – Présentation des grandes zones structurales des Alpes occidentales. A l’Ouest du Front Pennique se situent les zones externes, peu métamorphisées, composées du domaine Dauphinois et des massifs cristallins externes. A l’Est du Front Pennique se situent les zones internes, affectées par le métamorphisme alpin et composées du domaine Briançonnais, Piémontais et Liguro-Piémontais. Modifié de Renard et al. (2015).

5.2 Histoire géodynamique des Alpes

La structure géologique des Alpes nous indique que son histoire géodynamique est intimement liée au devenir des océans du système téthysien. Les principales étapes de l'évolution de la Téthys alpine sont présentées sur la figure 5.4. Au début du Mésozoïque, au Trias précisément, la croûte continentale de la Pangée subit un phénomène d'extension, le *rifting*, qui aboutit au Jurassique à l'ouverture de l'océan atlantique central et de l'océan ligure qui en est une dépendance (De Graciansky et al., 2011; Pfiffner, 2014). Au milieu du Crétacé, l'ouverture de l'océan atlantique se poursuit vers le Nord. Dans un premier temps, cette ouverture perturbe l'expansion de l'océan ligure et cause son avortement. Dans un second temps, elle entraîne un début de convergence entre l'Afrique et l'Europe qui se traduit par la fermeture de l'océan ligure.

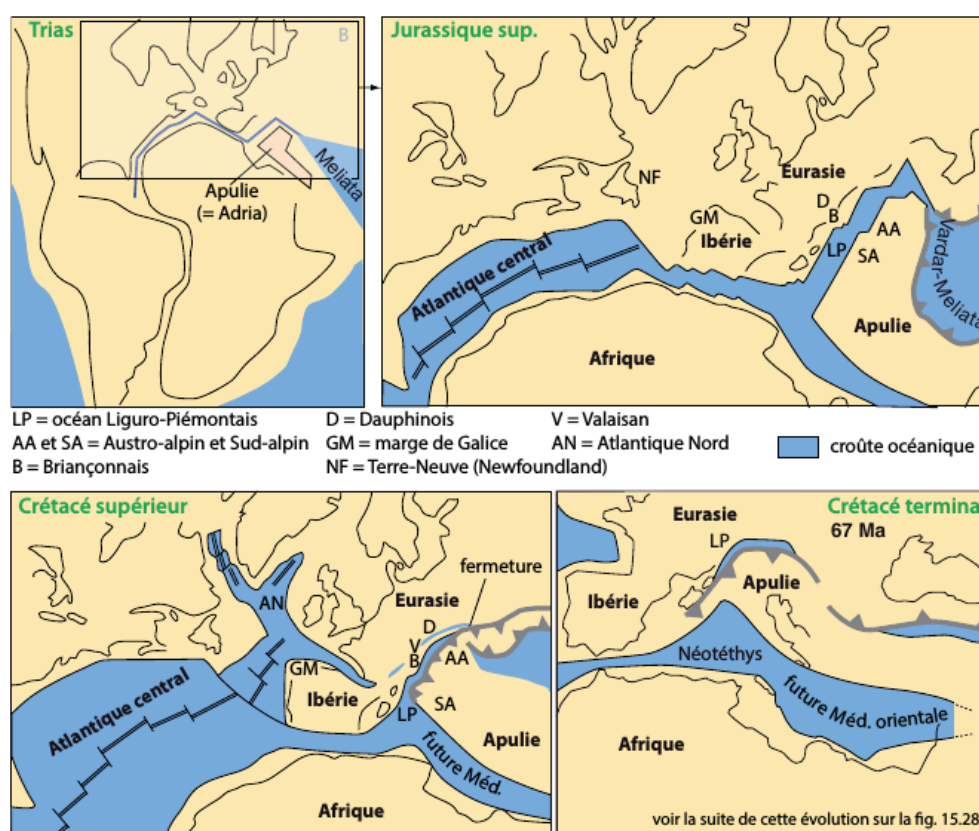


Figure 5.4 – Évolution paléogéographique du domaine Téthysien (d'après Renard et al. (2015)). L'ouverture de l'océan atlantique central au Trias induit l'ouverture de l'océan ligure au Jurassique supérieur et moyen. L'avancement de l'ouverture de l'océan atlantique vers le Nord induit l'avortement puis la fermeture de l'océan ligure au Crétacé.

La figure 5.5, présente schématiquement les stades majeurs de l'histoire géodynamique de la chaîne alpine. Nous pouvons résumer cette histoire par trois épisodes successifs² : la subduction

2. Détailler l'histoire géodynamique complète de la chaîne alpine sort du cadre de ce manuscrit. Et nous conseillons au lecteur de se reporter à la description récente de Handy et al. (2010)

océanique, la subduction continentale et la collision continentale (Renard et al., 2015).

La subduction océanique Elle prend place il y a environ 70 Ma avec l'enfouissement de la lithosphère océanique ligure sous la plaque adriatique. Les sédiments accumulés dans le prisme d'accrétion, ainsi que la croûte océanique, qui sont entraînés en subduction et métamorphisés, formeront les Schistes Lustrés, du Mont Viso et du Chenaillet (domaine liguro-piemontais).

La subduction continentale Aux alentours de 45 Ma, la subduction océanique devient continentale. Les domaines piémontais (marge océanique) et Briançonnais (marge continentale) entrent en subduction sous la plaque adriatique. Les formations cristallines de Dora Maira sont entraînées à très grande profondeur par la subduction continentale. Celle-ci s'arrête après quelques millions d'années et laisse place à la collision continentale.

La collision continentale Les croûtes continentales européenne et adriatique entrent en collision. Il s'ensuit une phase de surrection des zones internes avec exhumation des unités HP-BT dès l'Oligocène et une progression, à partir du Miocène, des chevauchements vers l'avant de la chaîne. La mise en place du front Pennique et les rétrochevauchements vers l'Est contribuent au raccourcissement horizontal de la chaîne et à son épaississement crustal.

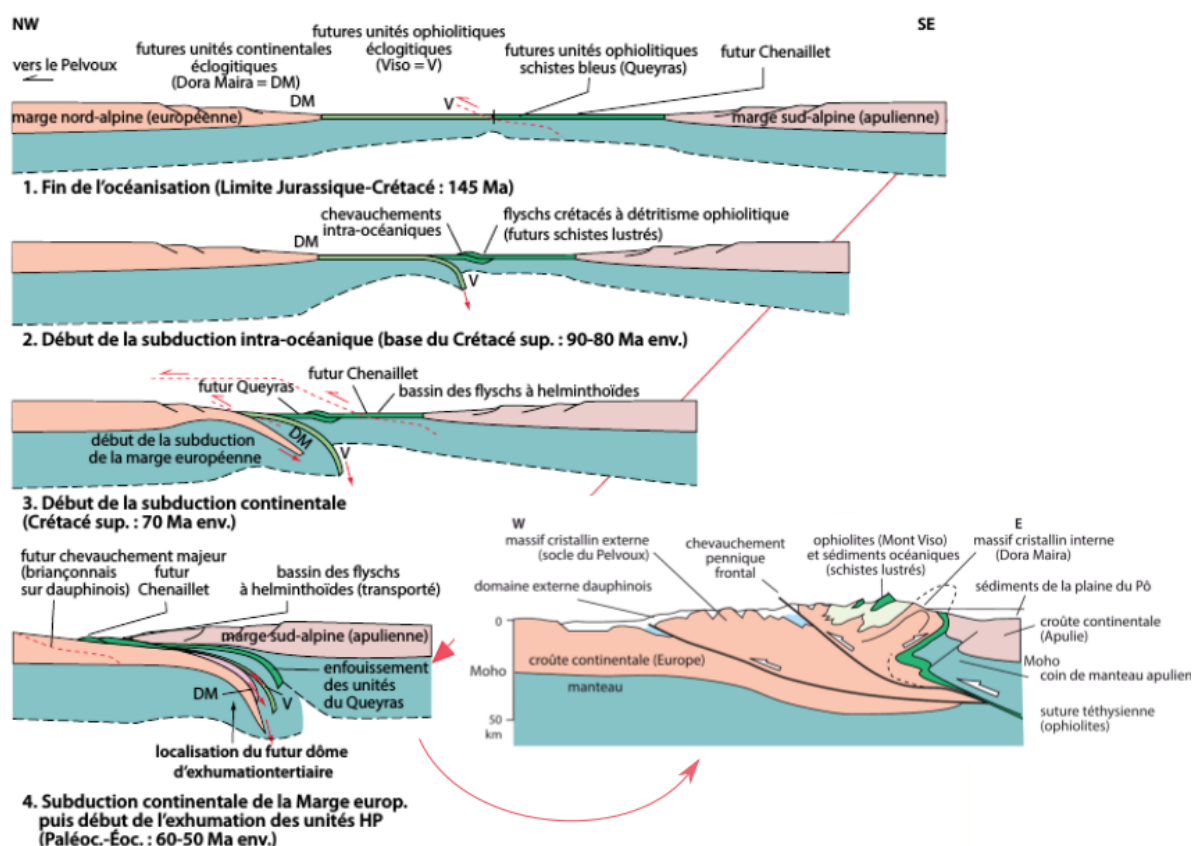


Figure 5.5 – Océan ligurien et son devenir alpin le long de la transversale Pelvoux – Mont Viso. Modifié d'après Renard et al. (2015).

5.3 Synthèse des connaissances géophysiques

Les méthodes géophysiques, en particulier la sismologie, permettent d'inférer des connaissances sur le comportement tectonique actuel des objets géologiques et d'obtenir des images indirectes de leurs structures profondes. Ces informations sont essentielles pour comprendre l'évolution tectonique de la lithosphère et apporter des contraintes supplémentaires pour replacer les unités géologiques dans un cadre géodynamique et établir des modèles de formation et d'évolution des chaînes de montagne. Dans les Alpes, la littérature est si fournie qu'il est impossible de tout résumer ici, aussi, nous n'effleurons que quelques études que nous jugeons essentielles pour la suite de ce manuscrit. Des synthèses géophysiques plus approfondies sont données par [Kissling \(1993\)](#) et [Kissling et al. \(2006\)](#).

5.3.1 Structures crustales

Les mesures gravimétriques ([Masson et al., 1999](#)) dans la chaîne alpine sont caractérisées par la présence de deux anomalies de Bouguer associées ([Fig. 5.6](#)), une négative et une positive, qui sont typiques des zones de collision. L'anomalie négative s'amplifie progressivement d'Ouest en Est : elle passe de -30 mgal dans la vallée du Rhône à -150 mgal dans les zones internes. Ce comportement traduit l'épaississement vers l'Est de la croûte continentale européenne dont l'épaisseur dépasse 50 km au milieu de la chaîne ([Spada et al., 2013](#)). Immédiatement à l'Est des zones internes, une anomalie positive atteignant 80 mgal correspond à la signature gravimétrique du corps d'Ivrée, interprété comme une écaïlle de manteau adriatique imbriquée à faible profondeur dans la croûte européenne ([Menard & Thouvenot, 1984](#)).

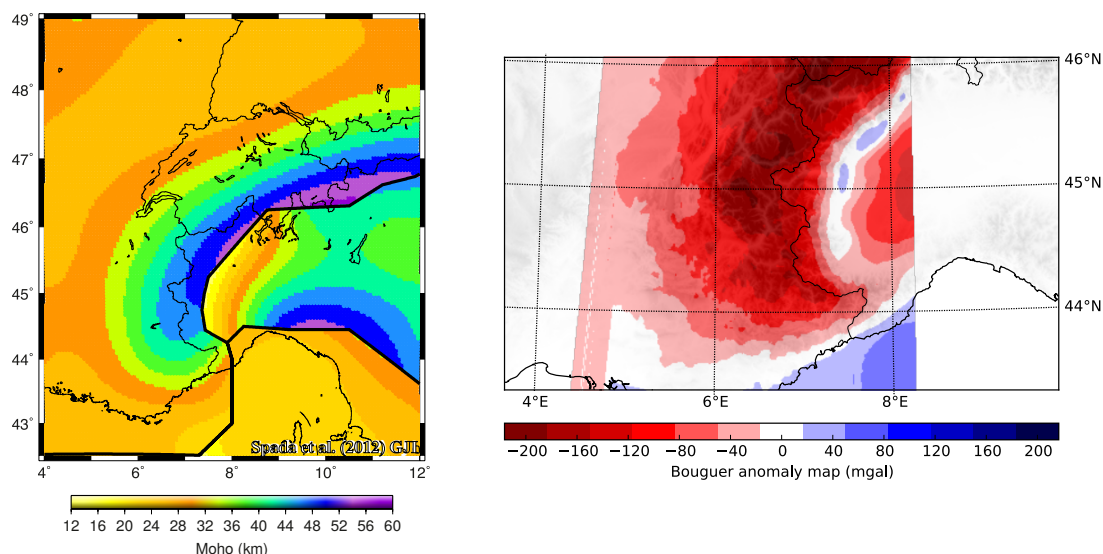


Figure 5.6 – A gauche, profondeurs du Moho fournies par [Spada et al. \(2013\)](#) à partir des données de sismique active et de l'analyse de fonctions récepteur. Au nord, le Moho européen, à l'Est le Moho adriatique et au sud, le Moho Thyrrénéen. A droite, carte d'anomalie de Bouguer obtenue par [Masson et al. \(1999\)](#).

Vers la fin des années 80, la mise en place du programme de sismique profonde ECORS-CROP (Bayer et al., 1987; Nicolas et al., 1990) a permis de réaliser un grand profil de sismique active, d'environ 200 km de long et orienté NO-SE, à travers la partie occidentale des Alpes au niveau du massif du Grand Paradis. Le profil, présenté sur la figure 5.7, précise la structure superficielle de la croûte alpine et a donné lieu à de nombreuses interprétations (Bayer et al., 1987; Nicolas et al., 1990; Roure, 1996; Schmid & Kissling, 2000; Lardeaux et al., 2006). Un exemple d'interprétation structurale, celle de Schmid & Kissling (2000), est présenté sur la figure 5.8.

Les principaux résultats du profil ECORS-CROP sont :

- Un Moho européen bien défini jusqu'aux Alpes internes. Sa profondeur varie graduellement de 35 km de profondeur dans les zones externes à 55 km de profondeur sous les zones internes tandis que le Moho adriatique est peu marqué. La base de la croûte supérieure européenne est elle-aussi bien représentée.
- En proche surface, le profil permet de clairement identifier les bassins sédimentaires, la géométrie du front Pennique, le prisme d'accrétion et les grands chevauchements du domaine dauphinois. À l'Est, il révèle également la présence de rétrochevauchements dans la croûte adriatique.
- Une marge européenne chevauchée par la plaque adriatique.
- Ils précisent la géométrie du corps d'Ivrée constitué, selon les auteurs, de plusieurs écailles de manteau et/ou de croûte inférieure.
- La présence sous la zone Briançonnaise d'une écaille de manteau adriatique (Bayer et al., 1987; Nicolas et al., 1990) ou de croûte inférieure européenne (Roure, 1996; Schmid & Kissling, 2000), confirmée par (Thouvenot et al., 2007).

Les divergences d'interprétation résultent principalement des incertitudes sur la géométrie de grandes structures profondes où les réflecteurs sont assez diffus et donc peu contraignant pour l'interprétation.

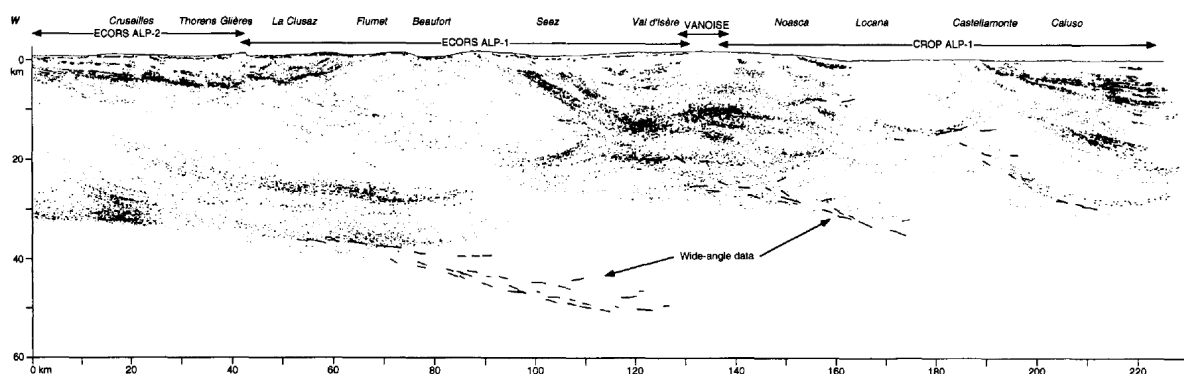


Figure 5.7 – Profils sismiques réflexion et réfraction ECORS-CROP des Alpes franco-italiennes. D'après Marchant & Stampfli (1997).

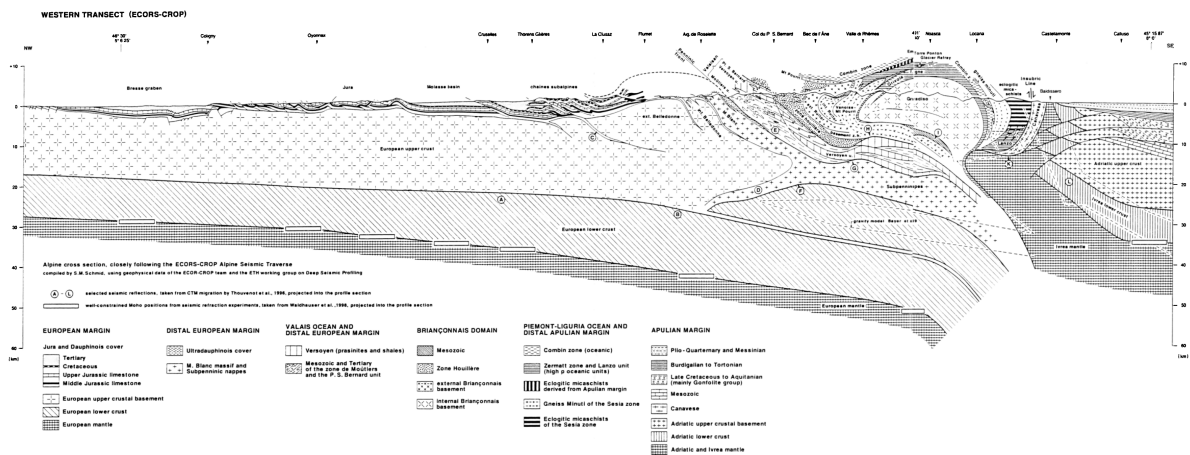


Figure 5.8 – Interprétation structurale des profils sismiques ECORS-CROP d’après Schmid & Kissling (2000).

Les résultats du programme ECORS-CROP ont été complétés par de nombreuses études de tomographie locale (Paul et al., 2001; Vernant et al., 2002; Béthoux et al., 2007; Diehl et al., 2009; Wagner et al., 2012; Potin et al., 2015). Nous nous concentrons sur les résultats des études de Paul et al. (2001) et Diehl et al. (2009).

Le modèle tomographique de Paul et al. (2001) présente une résolution limitée en profondeur due à la faible ouverture du réseau considéré. En conséquence, il ne contraint les vitesses d’ondes P de la partie occidentale des Alpes que dans la partie supérieure de la croûte. Le modèle tomographique indique que la croûte est dominée par de forts contrastes de vitesses. Les plus notables sont les anomalies de faibles vitesses dans les nappes externes de Digne et Castellane (visibles jusqu’à 10 km de profondeur) et des anomalies de vitesses très importantes (7.4 à 7.7 km/s) sous le massif de Dora Maira et la partie la plus occidentale de la plaine du Pô. Cette dernière est corrélée avec le corps d’Ivrée, qui est ainsi interprété comme une écaille mantellique adriatique, dont le toit se situe à 10 km de profondeur et peut être suivie continuellement jusqu’à 30 km de profondeur. Horizontalement, sa géométrie est allongée dans la direction Nord-Sud et fine dans la direction Est-Ouest (10 à 15 km de large).

Le modèle tomographique de (Diehl et al., 2009), qui fait usage d’un réseau de stations plus important et de plus grande ouverture, permet de contraindre les vitesses d’ondes P jusqu’à la croûte inférieure (Fig. 5.9). Ce modèle confirme les observations de Paul et al. (2001). En proche surface, il atteste l’augmentation de vitesse liée aux zones internes de l’orogène ainsi que leur diminution dans les bassins sédimentaires molassiques et de la plaine du Pô. La croûte supérieure est dominée par le corps d’Ivrée dont la signature en profondeur est visible jusqu’à 30 km de profondeur avant de progressivement disparaître aux profondeurs où ses vitesses se confondent avec celles du manteau. A 45 km de profondeur, des anomalies de faible vitesse à l’Ouest du corps d’Ivrée témoignent de l’épaississement crustal, visible jusqu’à 60km de profondeur, du fait de l’orogénèse alpine. La figure 5.10 présente une coupe verticale du modèle de vitesse P le long du profil ECORS. Elle indique l’origine adriatique du corps d’Ivrée qui poinçonne la croûte continentale européenne. Malgré une faible résolution verticale, elle confirme aussi l’origine crustale de l’écaille Briançonnaise. Néanmoins, elle ne permet pas de suivre la

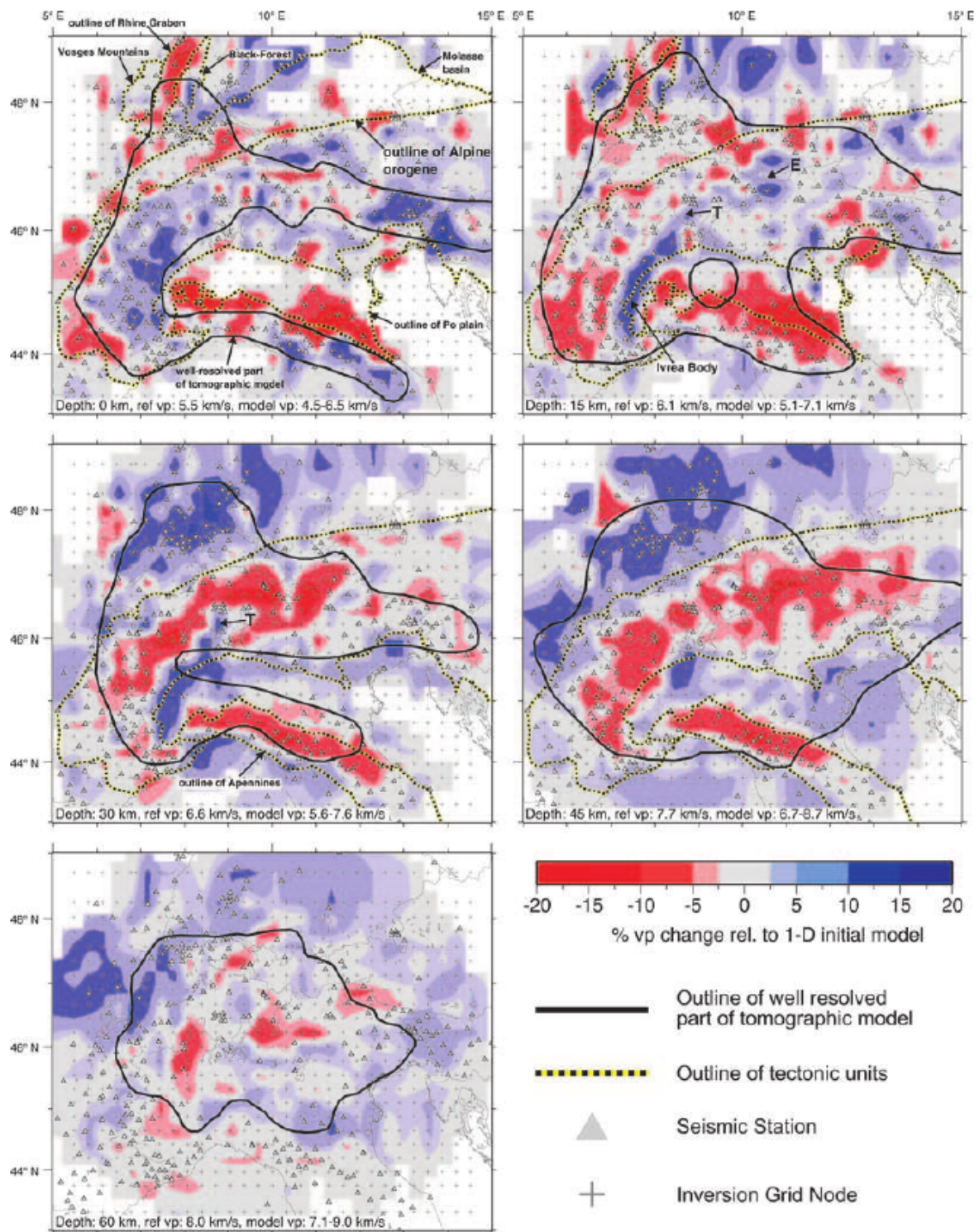


Figure 5.9 – Coupes horizontales du modèle tomographique crustal obtenu par Diehl et al. (2009), représenté sous forme de perturbations de vitesses d’onde P.

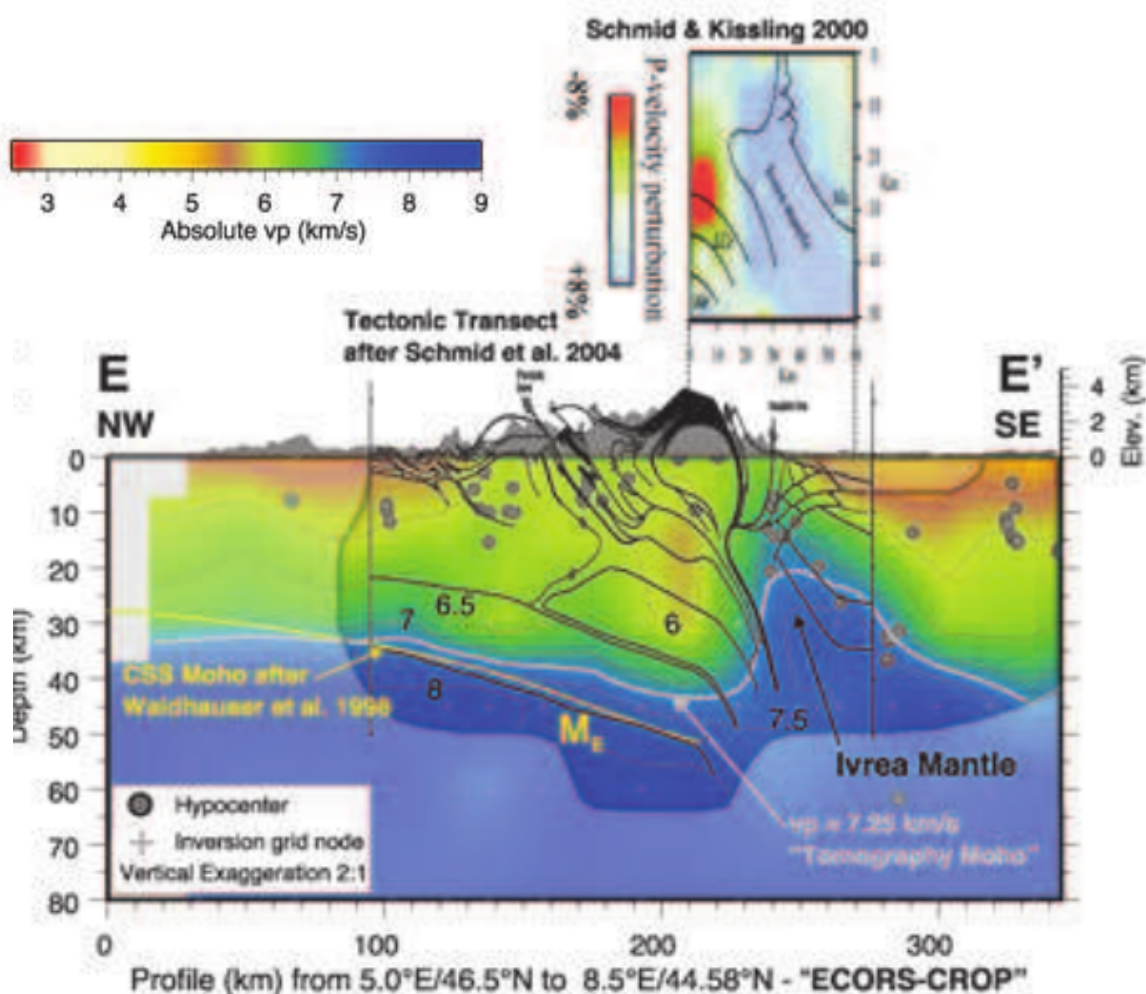


Figure 5.10 – Coupe verticale du modèle tomographique de vitesses d’ondes P de Diehl et al. (2009) le long du profil ECORS.

géométrie de la croûte inférieure européenne sous la plaque adriatique.

La question du Moho a toujours été un problème dans les Alpes occidentales (Laubscher, 1990; Marchant & Stampfli, 1997; Lombardi et al., 2008; Stehly et al., 2009; Spada et al., 2013). Bien que les travaux de Chopin (1984) attestent de l’enfouissement de la croûte continentale à des profondeurs de l’ordre de 90 à 100 km, aucune étude sismologique n’a permis de révéler une signature du Moho à de telles profondeurs. Ceci est notamment dû au fait que les tomographies locales ne permettent pas de contraindre des vitesses à de telles profondeurs et que les tomographies télésismiques n’ont pas une résolution suffisante pour imaginer une telle structure. Le déploiement récent du réseau ultra-dense CICALPS, a toutefois permis à Zhao et al. (2015) de fournir une première preuve sismologique de l’enfouissement à grandes profondeurs par le calcul et la migration de fonctions récepteur (Fig. 5.11). L’image obtenue montre un Moho européen plongeant de 35 km à l’Ouest jusqu’à 40 km de profondeur au niveau du front Pennique

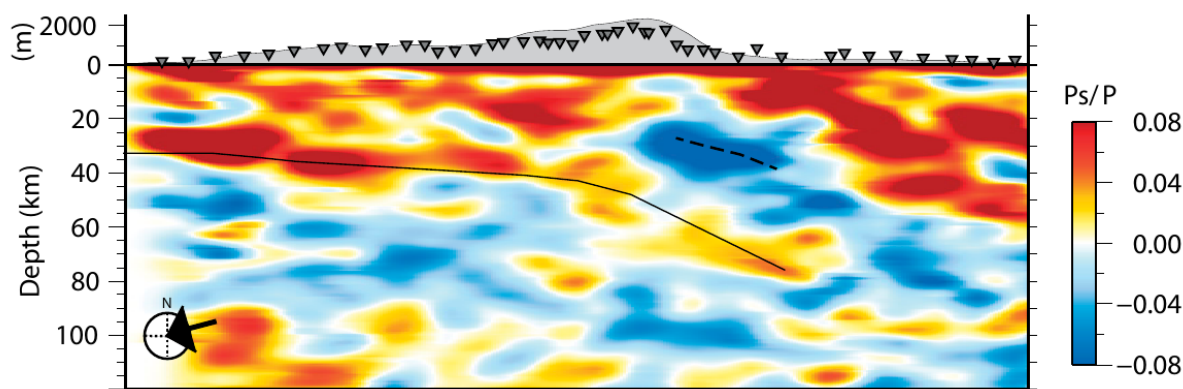


Figure 5.11 – Image obtenue par migration des fonctions récepteur P en S calculées à partir des données téléseismiques de la campagne CIFALPS par Zhao et al. (2015).

avec une pente inférieure à 5 degrés. A l'Est du front Pennique, le pendage augmente pour atteindre plus de 20 degrés, le Moho plonge alors de 40 à 75 km de profondeur sous le massif de Dora Maira et la partie occidentale de la plaine du Pô. Un réflecteur de polarité inverse, qui caractérise une augmentation de vitesse, délimite la frontière entre la croûte continentale subduite et le corps d'Ivrée. Plus à l'Est, les réflecteurs de polarité positive indiquent la partie supérieure du corps d'Ivrée et le Moho adriatique.

5.3.2 Structures mantelliques

L'évolution tectonique et la structure des Alpes ne peuvent être comprises ou reconstruites sans la connaissance précise des structures 3D du manteau lithosphérique et de la géométrie de la frontière entre la lithosphère et l'asthénosphère (Kissling et al., 2006; Fischer et al., 2010).

Les premières études tomographiques, par inversion de temps de trajet sans corrections crustales, menées par (Spakman et al., 1993; Kissling, 1993; Bijwaard & Spakman, 2000; Piromallo & Morelli, 2003) indiquent la présence d'anomalies de vitesses rapides continues jusqu'à 150 km de profondeur sous la ceinture alpine, interprétées comme de la lithosphère froide et indiquant le panneau plongeant européen qui subducte vers le Sud et le Sud-Est. Plus en profondeur, la zone de transition du manteau entre 410 et 660 km de profondeur est elle aussi dominée par du matériel rapide, lui aussi interprété comme de la lithosphère froide et constituant les reliques d'anciens panneaux plongeants subductés (Fig. 5.12).

Plus récemment, Lippitsch et al. (2003) ont réalisé une étude tomographique du manteau supérieur alpin par inversion de temps de trajets téléseismiques eux-mêmes corrigés des effets de propagation 3D dans la croûte. Leur modèle de perturbation de vitesse est présenté sur la figure 5.13-b,e. Ce modèle, mieux résolu que les précédents indique une géométrie différente des panneaux plongeants le long de la chaîne alpine. Dans les Alpes occidentales et centrales les panneaux plongeants sont orientés NO-SE tandis que dans les Alpes orientales son orientation est inversée (il plonge vers le Nord). De plus, alors que les panneaux plongeants sont relativement continus dans les Alpes orientales et centrales, le panneau plongeant dans les Alpes

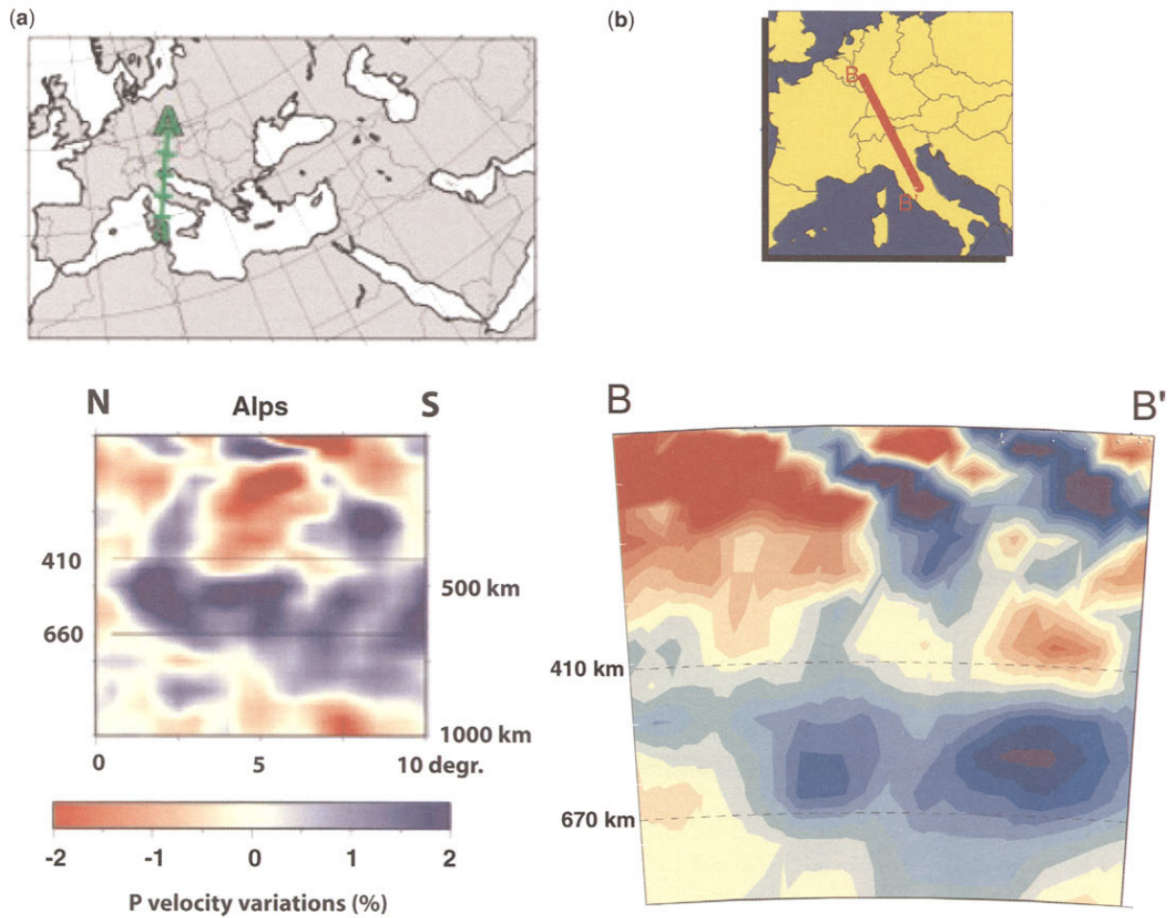


Figure 5.12 – Comparaison des modèles tomographiques du manteau supérieur obtenus par (a) Piromallo & Morelli (2003) et (b) Bijwaard & Spakman (2000). D’après (Kissling et al., 2006).

occidentales apparaît discontinu (Fig. 5.13-a, entre 150 et 300 km de profondeur). Cette discontinuité est interprétée comme le détachement du panneau plongeant océanique européen sous les Alpes occidentales, le « *slab break-off* ». Immédiatement à l’Ouest de cette discontinuité, le modèle présente une anomalie de vitesse faible elle-même interprétée comme une remontée asthénosphérique occasionnée par le détachement. L’asymétrie entre les panneaux plongeants des Alpes occidentales et centrales est elle-même interprétée comme un déchirement latéral progressif du panneau plongeant, le « *slab tear* ».

Le concept de détachement de panneaux plongeants n’est pas nouveau. Il a initialement été proposé par (von Blanckenburg & Davies, 1995) pour expliquer les épisodes magmatiques apparus durant le Cénozoïque. Il a ensuite été repris pour expliquer l’extension syn-convergence et tardi-orogénique (Sue et al., 1999; Wortel & Spakman, 2000) de la chaîne alpine et les effets de topographie dynamique (Fox et al., 2015; Nocquet et al., 2016). Il apparaît comme un processus majeur des chaînes de collision (Wortel & Spakman, 2000) et explique la forme arquée des Alpes occidentales ainsi que le déficit de matériel lithosphérique dans la partie supérieure du manteau (Piromallo & Faccenna, 2004). En effet, la convergence dans les Alpes implique un

raccourcissement horizontal d'au moins 500 km, en revanche, les images tomographiques ne décèlent les traces de la subduction que jusqu'à 200-300 km de profondeur. Piromallo & Faccenna (2004) proposent que les anomalies de vitesse rapides, localisées dans la zone de transition du manteau, correspondent à de la lithosphère subductée jusqu'au Paléogène, détachée entre 35 et 30 Ma, puis empilées à 660 km de profondeur.

Très récemment, Zhao et al. (2016a) ont réalisé une tomographie à fréquences finies du manteau supérieur dans les Alpes et les Apennins en inversant des temps de trajet corrigés de la propagation verticale dans la croûte. Ce modèle est présenté sur le figure 5.13b-d. Contrairement au modèle de Lippitsch et al. (2003), le modèle de Zhao et al. (2016a) présente une continuité latérale et verticale du panneau plongeant dans les Alpes occidentales et centrales. Du fait de la continuité du panneau plongeant, Zhao et al. (2016a) remettent en cause l'hypothèse du détachement ou déchirement du panneau plongeant. Ils retrouvent toutefois la forte anomalie négative de vitesse à l'Ouest du panneau plongeant attestant de la remontée asthénosphérique et attribuent eux-aussi l'extension observée dans les Alpes internes à des effets de topographie dynamique. L'origine de la remontée asthénosphérique à l'Ouest du panneau plongeant est attribuée à un effet de succion causé par le retrait du panneau plongeant apennique (*slab roll-back*).

Cette dernière hypothèse a déjà été proposée par Barruol et al. (2004, 2011) pour expliquer les mesures d'anisotropie du manteau supérieur obtenues par l'étude de la biréfringence des ondes SKS. Ils observent des vitesses rapides dans une direction parallèle à la chaîne alpine, témoins de la déformation et du flux mantellique (Fig. 5.14). Ils proposent trois hypothèses pour expliquer cette anisotropie :

- un effet « quille » du panneau plongeant entraîné par le mouvement de la plaque européenne,
- un flux asthénosphérique lié au déchirement progressif du panneau plongeant,
- un flux asthénosphérique généré par le retrait du panneau plongeant apennique.

L'étude note aussi que l'amplitude de l'anisotropie varie de part en part de la chaîne : elle est plus importante dans les parties externes que dans les parties internes. L'absence de mesures fiables au centre des Alpes internes est cohérent avec la remontée asthénosphérique visible dans les modèles tomographiques de Lippitsch et al. (2003) et Zhao et al. (2016a).

5.4 Conclusions et problématiques actuelles

Dans ce chapitre nous avons présenté les résultats structuraux et géodynamiques majeurs apportés par la géologie et la géophysique aux cours des dernières années. Il est admis aujourd'hui que les Alpes résultent de la collision continentale entre la marge européenne et le promontoire adriatique. À l'échelle crustale, cette collision se traduit par l'épaississement de la marge européenne avec un rétro-chevauchement ou dédoublement de la croûte continentale inférieure, elle-même poinçonnée par une écaille de manteau adriatique, le corps d'Ivrée (Kissling et al., 2006). Pour les structures mantelliennes, les études tomographiques concordent vers une subduction de la lithosphère européenne sous la lithosphère adriatique dans les Alpes occidentales ainsi que sur une remontée asthénosphérique à l'Ouest de la chaîne.

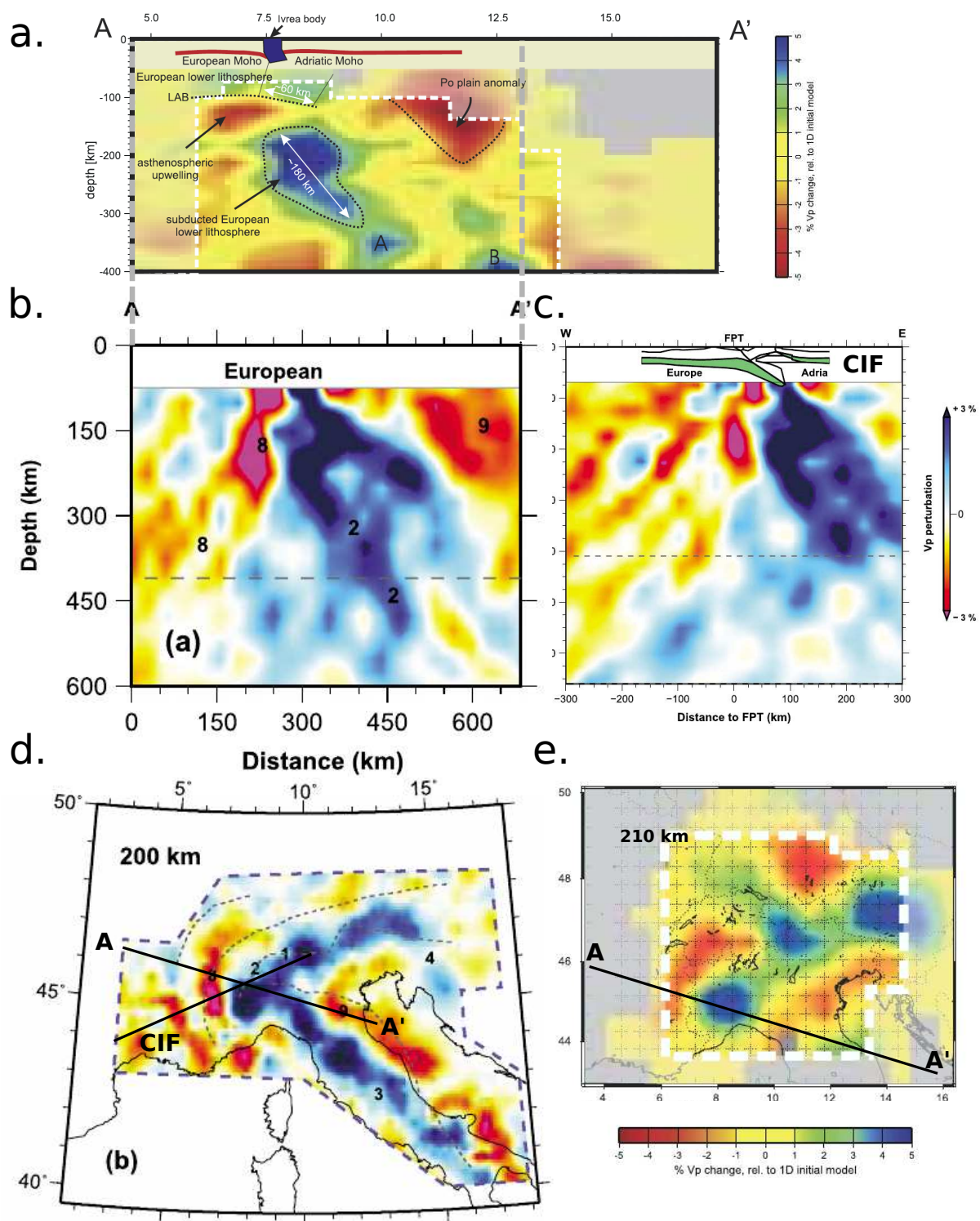


Figure 5.13 – Comparaison des modèles tomographiques du manteau supérieur obtenus par Lippitsch et al. (2003) (M1) et Zhao et al. (2016a) (M2). (a.) Coupe verticale Est-Ouest du modèle M1. (b.) Coupe verticale du modèle M2 le long du profil est-ouest représenté en (a). (c) Coupe verticale du modèle M2 le long du profil CIFALPS (CIF). (d) Coupe du modèle M2 à 200 km de profondeur. (e) Coupe du modèle M1 à 210 km de profondeur.

Néanmoins, des questionnements géodynamiques et structuraux demeurent. Il est notamment nécessaire de préciser la géométrie des principales unités structurales dans la croûte. En particulier, la géométrie du Moho européen pour caractériser plus précisément la subduction continentale à grandes profondeurs ainsi que celle du Moho adriatique généralement mal contrainte sous la plaine du Pô. Il est également nécessaire de préciser la nature de l'écaille Briançonnaise (Thouvenot et al., 2007) et de confirmer ou non un branchement possible du front Pennique sur une écaille du manteau. Il est également important de préciser la géométrie du corps d'Ivrée, notamment sa bordure occidentale. À l'échelle mantellique, des images tomographiques de haute résolution sont requises pour contraindre la géométrie 3D des panneaux plongeants européens (Alpes centrales et occidentales) et adriatiques (dinariques et apenniques). De telles images permettraient de déterminer les relations entre les différentes subductions (Vignaroli et al., 2008), et de déterminer si le panneau plongeant européen s'est détaché (Lippitsch et al., 2003) ou, au contraire, s'il est continu en profondeur (Zhao et al., 2016a).

Pour tenter de répondre à ces questionnements, nous présentons dans le prochain chapitre l'application de notre méthode d'inversion de formes d'ondes aux données télésismiques de la campagne CICALPS permettant de reconstruire un modèle multi-paramètres haute résolution de la croûte et du manteau supérieur.

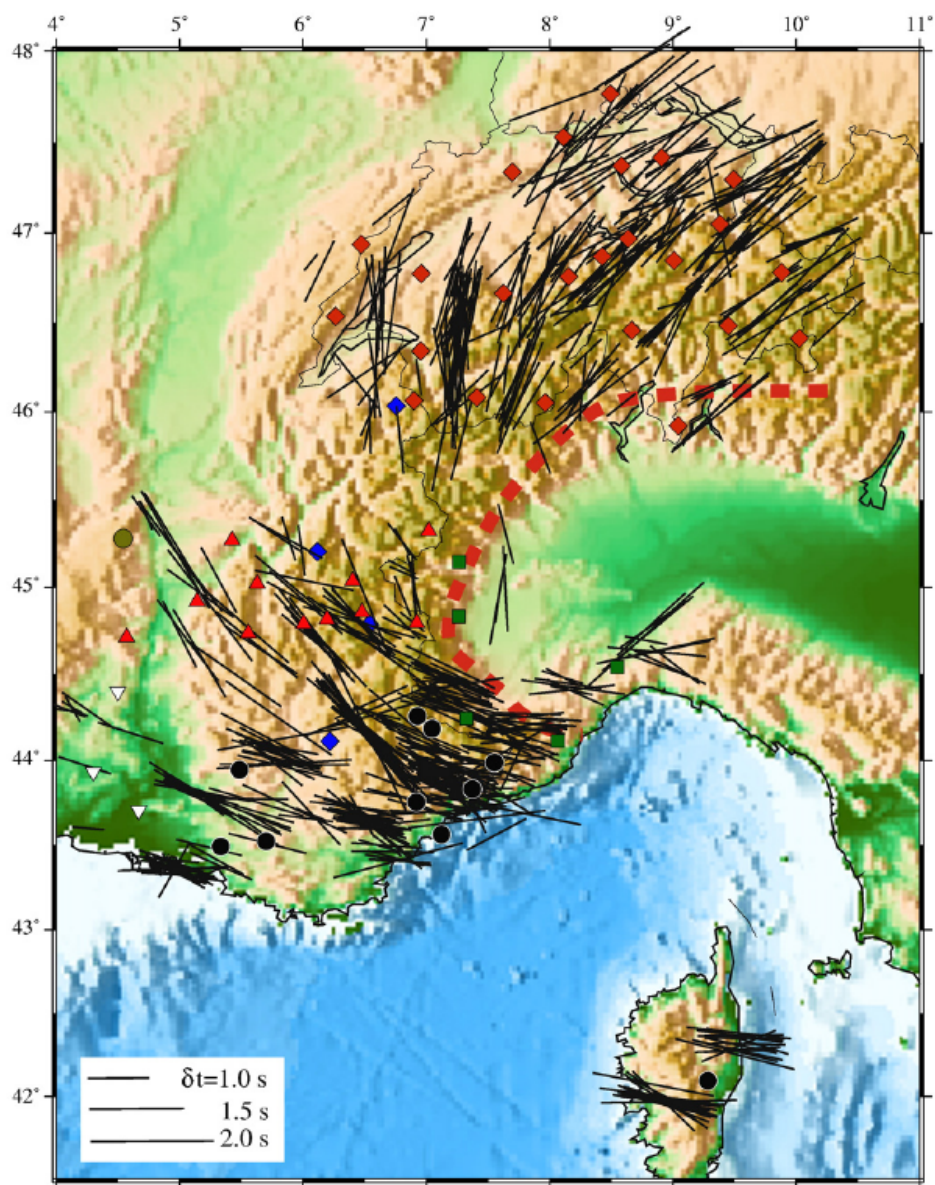


Figure 5.14 – Mesures de direction d’axe rapide de la biréfringence des ondes SKS dans les Alpes occidentales. D’après Barruol et al. (2011).

Chapitre 6

Application de l'inversion de formes d'ondes aux données CIFALPS

Sommaire

6.1	Introduction	207
6.2	Multi-parameter teleseismic waveform inversion	210
6.2.1	Teleseismic full waveform modelling	210
6.2.2	Inverse problem	212
6.2.3	High-resolution multi-parameter imaging	213
6.3	Application to the Western Alps	215
6.3.1	The CIFALPS experiment	215
6.3.2	Full-waveform inversion workflow	216
6.3.3	FWI models	220
6.3.4	Resolution and parameter cross-talk appraisal	223
6.4	Discussion	228
6.4.1	Density parameter reconstruction	228
6.4.2	Geodynamic interpretation	229
6.5	Conclusions	232

Ce chapitre est présenté sous forme d'une publication que nous avons soumis au journal *Geophysical Journal International*. Il présente l'application de la méthode d'inversion de formes d'ondes aux données télésismiques enregistrées lors de l'expérience *CIFALPS* dans les Alpes occidentales, pour l'imagerie multiparamètres (densité et vitesses sismiques P et S) des structures lithosphériques superficielles et profondes de la chaîne alpine. Dans l'article, nous décrivons succinctement la méthode d'inversion de formes d'ondes ainsi que les données utilisées. Nous détaillons les traitements apportés aux données télésismiques et le déroulement du processus d'inversion des formes d'ondes. Nous présentons le modèle FWI obtenu pour la lithosphère des Alpes occidentales et analysons sa résolution au moyen de tests de résolution. Finalement, nous discutons de l'apport de la FWI à l'imagerie lithosphérique en terme de résolution et de reconstruction multiparamètres, et proposons une interprétation des structures lithosphériques observées qui semble confirmer l'hypothèse d'un détachement du panneau plongeant européen sous les Alpes occidentales.

Lithospheric architecture of the South-Western Alps revealed by multi-parameter teleseismic full-waveform inversion

S. Beller, V. Monteiller, S. Operto, G. Nolet, A. Paul and L. Zhao
submitted to *Geophysical Journal International*.

Abstract

The Western Alps, although being intensively investigated, remains elusive when it comes to determine its lithospheric structure. New inferences on the lithospheric structure are crucial for the understanding of processes and mechanisms of orogeny needed to unravel the dynamic evolution of the Alps. This situation led to the deployment of the *CIFALPS* temporary experiment, conducted to address the lack of seismological data amenable to high-resolution seismic imaging of the crust and the upper-mantle. We perform a 3D isotropic full-waveform inversion of nine teleseismic events recorded by the *CIFALPS* experiment to infer three-dimensional models of both density and P and S wave velocities of the Alpine lithosphere. Here, by full waveform inversion is meant the inversion of the full seismograms including phase and amplitude effects within a time window following the first arrival up to a frequency of 0.2 Hz. We show that the application of the FWI at the lithospheric scale is able to generate images of the lithosphere with unprecedented resolution and can furnish a reliable density model of the upper-lithosphere. In the shallowest part of the crust, we retrieve the shape of the fast / dense Ivrea body anomaly and detect the imprints of the slow Po and SE France sedimentary basins. The geometry of the Ivrea body as revealed by our density model is consistent with the Bouguer anomaly. A concise Moho transition is followed from the external part (30 km depth) to the internal part of the Alps (70-80 km depth), giving clear evidence of a continental subduction event during the formation of the Alpine Belt. A low velocity zone in the lower lithosphere of the S wave velocity model supports the hypothesis of a slab detachment in the Western part of the Alps that is followed by asthenospheric up-welling. The application of FWI to teleseismic data helps to fill the gap of resolution between traditional imaging techniques, and enables integrated interpretations of both upper and lower lithospheric structures.

Keywords : Western Alps – Full-waveform inversion – Density – Teleseismic

6.1 Introduction

The Alpine Belt results from the collision between the European margin and the Adriatic margin, initiated during the Paleogene 35 Ma ago as a consequence of the Alpine Tethys closure (Stampfli et al., 2002; Handy et al., 2010). It is nowadays one of the most studied orogenic domains in the world, and therefore an ideal target to understand the deep processes and mechanisms of orogeny. Notably, the Western part of the Alps has preserved its full metamorphic history and has been the cradle of the continental subduction concept with the discovery by Chopin (1986) of HP-LT coesite-bearing metamorphic rocks in the Dora Maira massif. These rocks were the first evidence of the burial of the continental crust at depths greater than 90 km within the Alpine lithosphere (Marchant & Stampfli, 1997; Zhao et al., 2015).

However, at that time, no clear seismic evidence was available for such a mechanism in orogenies and the concept of continental subduction was highly debated. While high-resolution tomographic images of the deep lithosphere were lacking, key insights on the upper-lithospheric structure of the Western Alps were inferred from various geophysical studies. Notable features like the Ivrea body (early revealed by gravity anomalies) were interpreted as uplifted mantle materials within the upper-crust. The crustal velocities and the Moho discontinuity inferred from the prominent ECORS-CROP seismic reflection profiling in the late 1980s suggested a mantle wedge intruding the Alpine crust located further west from the Ivrea body (Bayer et al., 1987; Nicolas et al., 1990). A reinterpretation of available geophysical knowledge by Schmid & Kissling (2000) attributed this anomaly to a European lower crustal wedge. Nonetheless, seismic details on the deep structure of the south-western part remain sparse (Thouvenot et al., 2007) and as pointed out by Kissling (1993) : “without imaging techniques independent of assumptions about tectonic and dynamic processes, such as seismic methods, the various models of deep Alpine crustal structure could not be tested and improved”.

This has motivated intense seismic experiments and studies (Kissling et al., 2006). At crustal scale, local earthquake tomography (LET) (Paul et al., 2001; Béthoux et al., 2007), sometimes combined with gravity inversion (Vernant et al., 2002), succeed in revealing details of the shallowest structure of the Alpine crust. They notably underlined the extent of a high P-wave velocity anomaly (from 7.4 to 7.7 km/s) up to 10 km depth interpreted as the Ivrea body and pointed out its Adriatic mantle origin. Nonetheless, the limited aperture of the available seismic arrays restricted the resolution of these tomographic images to the upper-crust. More recent LET studies made use of wider seismic arrays (Diehl et al., 2009; Potin et al., 2015) with a larger number of data allowing the inference of crustal structures down to the lower crust. The tomographic model of Diehl et al. (2009) shows an upper to middle crust dominated by the high velocity anomaly associated with the Ivrea body that can be traced down to 45 km depth. In the upper-crust, this Ivrea body is surrounded by low velocity anomalies attributed to the Molasse and Po sedimentary basins that are also well-visible on ambient-noise tomography (Stehly et al., 2009). From 30 km depth, they observed reduced P-wave velocities associated with the thick crustal root of the Western Alps.

The upper-mantle structure of the Alps was derived by Lippitsch et al. (2003) from crustal corrected teleseismic P-wave traveltimes tomography. Their study revealed fast velocity anomalies interpreted as remnants of subducting slabs from 100 to 300 km depth in the vicinity of the

Alpine Belt. They notably observed a change in the subduction regime along the Alpine Belt. In the Western Alps, they denote a gap of fast velocity in conjunction with a slow velocity anomaly, interpreted as a testimony of slab detachment (Davies & von Blanckenburg, 1995; von Blanckenburg & Davies, 1995) and a related asthenospheric upwelling. However, in Central and Eastern Alps this fast anomaly can be traced from the crust down to 200 km depth with a reverse polarity in the Eastern Alps.

Despite the significant contribution of these tomographic studies, the need to improve the spatial resolution of the lithospheric models of the Alps and to more completely describe their seismic properties remains. To fill this gap, a dense profile of broadband stations was deployed during the CIFALPS temporary experiment to collect seismological data amenable to new high-resolution seismic imaging techniques (Zhao et al., 2013). A first exploitation of the CIFALPS data by common conversion point (CCP) migration of the teleseismic P-wave receiver functions has provided the first conclusive seismic image of a subducting continental crust in the Western Alps (Zhao et al., 2015).

The aim of this study is a second attempt to develop high-resolution images of the lithosphere of the western Alps from the CIFALPS data. We present an application of full waveform inversion (FWI) on the CIFALPS data to derive 3D high-resolution P and S wave velocity and density models of the lithosphere of the Western Alps and compare them with the migrated image of Zhao et al. (2015).

Beyond receiver function analysis, waveform inversion methods are expected to provide quantitative lithospheric models with a resolution close to the wavelength. Among teleseismic waveform inversion techniques, migration of scattered teleseismic body waves has been first proposed by Bostock et al. (2001), Shragge et al. (2001), Rondenay et al. (2001) and Rondenay et al. (2005). In their approach, the forward problem is linearized with the single-scattering Born approximation and the migration is recast as a ray-based linear waveform inversion accordingly, where the misfit between the recorded and modelled scattered wavefields is minimised in a least-squares sense. In this framework, Bostock et al. (2001) stress on the key contribution of the known free-surface reflector as a second-order source of down-going P and S waves. These down-going P and S waves are amenable to reflection from the lithospheric reflectors before being recorded at the surface by the stations. Migration of these back-scattered waves provide sharp quantitative images of lithospheric reflectors which are parametrised by P and S wave velocity perturbations. An application on the teleseismic data collected during the 1993 Cascadia experiment reveals the geometry of the continental Moho and the oceanic crust of the subduction Juan de Fuca plate (Rondenay et al., 2001).

A natural extension of the former linear waveform inversion is non-linear full-waveform inversion (FWI), a technique also borrowed from exploration geophysics (Tarantola, 1984a; Mora, 1987; Crase et al., 1990; Pratt, 1999). In a similar way to the linear waveform inversion, FWI relies on the principles of generalized diffraction tomography (Devaney, 1984; Wu & Toksöz, 1987) : the gradient of the misfit function is computed by the zero-lag correlation between the recorded and modelled scattered wavefields. The modelled scattered wavefield corresponds to the partial derivative of the wavefield with respect to the model parameters. In scattering migration, the recorded scattered wavefield is separated from the full wavefield during a preprocessing step once and for all, while, in FWI, the recorded scattered wavefield is built

on the fly at each iteration by the differences between the full and the modelled seismograms (namely, the data residuals). In FWI, the forward problem is performed without linearisation and a non-linear inversion is performed with a local optimization technique. This implies that the initial earth model at each iteration is updated with the final model of the former iteration. As the resolution of the subsurface models in which seismic modelling is performed improves over iterations, the resulting discontinuities can help to more fully account for converted waves and multi-scattering during the inversion process.

In its conventional form, FWI seeks to minimize in a least-squares sense the sample-to-sample misfit between the recorded and modelled seismograms. This implies that both travel-times and amplitudes of all of the arrivals are involved in the misfit function. Alternatively, the sensitivity kernels of the FWI can be modified in order to recast the waveform inversion in a finite-frequency traveltimes inversion of the first arrival (Marquering et al., 1999) or selected wave packets (Maggi et al., 2009). This approach, which accounts for amplitudes during seismic modelling for accurate extraction of time delays but removes their imprint in the misfit function, is generally referred to as adjoint tomography by the earthquake seismological community when seismic modelling is performed with full-wave numerical techniques such as the spectral element method (Komatitsch & Tromp, 1999a). Here, by adjoint is meant a mathematical technique that allows for the efficient computation of the gradient of a functional without forming the sensitivity matrix (see Tromp et al. (2005); Fichtner et al. (2006a,b); Plessix (2006a) for a review). An alternative to the adjoint approach is the scattering-integral approach which relies on the explicit computation of the sensitivity matrix (Chen et al., 2007). Several applications of adjoint tomography at regional, continental and global scales have been presented by Tape et al. (2010), Fichtner et al. (2009), Fichtner et al. (2010), Zhu et al. (2012), Fichtner et al. (2013) Zhu et al. (2015) and Bozdağ et al. (2016).

First attempts of teleseismic FWI were performed in 2D and 2.5D (Roecker et al., 2010; Baker & Roecker, 2014; Pageot et al., 2013) using a frequency-domain implementation (Pratt, 1999). A frequency domain implementation was used because 2D seismic modelling can be performed efficiently with Gauss elimination techniques when a large number of sources are processed for a few discrete frequencies (Pratt, 1999). Pageot et al. (2013) presented a parametric analysis that highlighted the contribution of the back-scattered waves from the free surface to improve the resolution of the imaging as originally pointed out by Bostock et al. (2001). They also show that a fine sampling of the frequencies was required during the inversion to balance the narrow and coarse scattering angle illumination provided by a few plane wave sources. This prompts them to conclude that a time-domain implementation of FWI was more suitable than the frequency-domain counterpart according to the specificity of the teleseismic experiments.

A first feasibility study of 3D time-domain FWI of teleseismic data has been performed by Monteiller et al. (2015). In their approach, full waveform modelling is performed for lithospheric targets with a hybrid approach : first, a simulation is performed in a radially symmetric global earth with the Direct Solution Method (DSM) method (Geller & Ohminato, 1994) and the wavefield solutions are stored on the edges of the lithospheric target. Second, a grid injection technique is used to propagate the wavefields in the lithospheric target with the spectral-element method from the storage of the DSM solutions at the edges of the computational mesh (Monteiller et al., 2013). Monteiller et al. (2015) illustrate the resolution power of FWI relative to

traveltime tomography to image a simple crustal model with a sharp Moho even when only four teleseismic events are used for inversion. The reliability of FWI for teleseismic application was further demonstrated with an application to real data collected in the framework of the Pyrope experiment across the Pyrenees range (Wang et al., 2016). FWI of five events recorded by 29 stations provided tomographic P and S wave velocity models of the Pyrenean lithospheric structure with an unprecedented resolution. Moreover, Wang et al. (2016) showed the consistency between the P and S wave velocity models developed by FWI and migrated images inferred from receiver functions.

In this study, we present an application of FWI on the CIFALPS dataset following an approach similar to that of Monteiller et al. (2015). A key difference is that we update the density parameter (ρ) in addition to the P and S wave velocity parameters (V_p and V_s). In the first section of the paper, we review some methodological aspects of teleseismic FWI. From the analysis of the radiation patterns of waves scattered by elastic heterogeneities, we also provide some insights on the resolution with which V_p , V_s and ρ can be imaged as well as potential leakage between parameters that might impact upon the multi-parameter reconstruction. Then, we show the FWI results and discuss their reliability based on a resolution analysis. We conclude with a discussion on the role of the density in teleseismic FWI and the geodynamical implications that can be drawn from our FWI results. Notably, we highlight the FWI potential in recovering a reliable density model of the upper-lithosphere that shows clear correlations with regional gravity anomalies. Upper-lithospheric velocity models well delineate known structures of the upper-crust, particularly the geometry of the Ivrea body that is clearly connected to the Adriatic mantle. Deeper, models display a clear crust-mantle boundary that dips down to 80-90 km beneath the internal Alps indicating a continental subduction event as a starting of the continental collision. At lower-lithospheric scale, our results give evidence of slab detachment and support the assumption of a slab tear in the Alpine Belt.

6.2 Multi-parameter teleseismic waveform inversion

6.2.1 Teleseismic full wavefield modelling

Teleseismic experiments deviate from traditional seismic applications such as local earthquake tomography or controlled-source seismology by rather different acquisition geometry. Classically, seismic sensors are located on top of the lithospheric target to be imaged and basically consist of either linear one-dimensional or sparse two-dimensional arrays. The special feature of teleseismic configuration arises from the use of passive and distant sources, located several thousands of kilometres away from the target, commonly referred to as teleseismic events (Fig. 6.1). In this source-receiver geometry, the teleseismic wavefield can fairly be approximated by an incident planar wavefield incoming from outside of the lithospheric target. Such plane-wave configuration leads to a rather uniform illumination of the lithospheric target (in the sense that planar waves do not suffer from geometrical spreading) at the expense of the incidence-angle coverage, which is much narrower and coarser than for point-source acquisition.

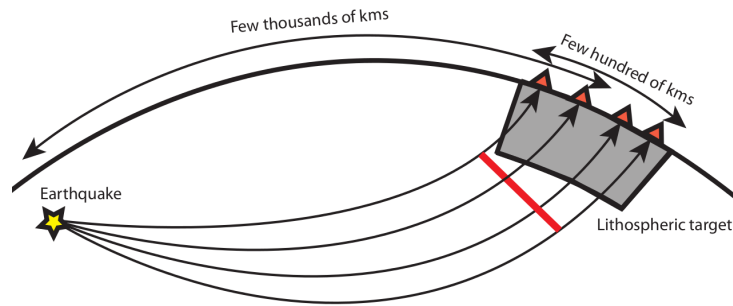


Figure 6.1 – Sketch of the teleseismic configuration.

From the computational perspective, the teleseismic setting raises the issue of the burden associated with the simulation of the teleseismic wavefields in a global earth within the 0.05-1 Hz frequency band. Several approaches based upon hybrid strategies have been proposed to reduce the computational effort of teleseismic wavefield simulation. The governing idea relies on the use of a background wavefield that can be efficiently computed in a simplified earth model and on the efficient computation of either the scattered or the full wavefield after model alteration.

Roecker et al. (2010) introduced the so-called scattered field method (Taflove & Hagness, 2005), where the source term of the wave equation is non zero where the background model differs from the model in question. Although this method is suitable in the frequency domain, its time-domain implementation would require the storage of the full 3D incident wavefields at each time step, that makes it intractable. A more affordable strategy we use in the present study is the so-called total-field / scattered-field grid injection method, which was first proposed to implement point sources in finite-difference simulations (Alterman & Karal, 1968; Taflove & Hagness, 2005; Monteiller et al., 2013; Masson et al., 2014; Tong et al., 2014a). As the scattered-field approach, it also relies on the pre-computation of an incident wavefield in a simplified earth model. However, the storage of the incident wavefield is only required along the outer boundaries of the lithospheric target. This wavefield is then injected within the lithospheric medium by applying a time-dependent boundary condition which accounts both for the incoming incident wavefield and the outgoing wavefield scattered by the lithospheric heterogeneities.

Our approach relies on the 3D spectral-element method (Komatitsch et al., 1998; Komatitsch & Tromp, 1999a), an efficient method for elastic wave propagation in presence of smoothly-varying topography, to simulate the teleseismic wave propagation within the lithospheric target. Unlike Monteiller et al. (2013) who use the DSM method to compute the incident wavefield in a radially symmetric global earth, we use the spectral-element *AxiSEM* software to compute the 3D incident wavefield (Nissen-Meyer et al., 2014). *AxiSEM* is a spectral element solver that allows one to infer 3D wavefields in a global axisymmetric earth for full-moment tensor sources from four 2D simulations (Nissen-Meyer et al., 2007b, 2008, 2007a). Its use greatly decreases the computational burden of simulations performed with *DSM* in a radially symmetric earth (Takeuchi et al., 2000) or with *SPECFEM3D_GLOBE* code in a fully 3D earth (Komatitsch

& Tromp, 2002a). Moreover, *AxiSEM* relies on the same discretization scheme than the one used to perform seismic modelling in the lithospheric target, hence mitigating the imprint of the numerical noise generated during the wavefield injection.

6.2.2 Inverse problem

Full-waveform inversion (FWI) is an iterative data fitting procedure that aims to update model parameters \mathbf{m} of the subsurface by minimising the waveform misfit between modelled and recorded seismograms (Virieux & Operto, 2009). Mathematically, it can be recast as a local optimisation problem that seeks a minimum of the least-squares misfit function :

$$\chi(\mathbf{m}) = \sum_{s=1}^{N_s} \sum_{r=1}^{N_r} \int_0^T \frac{1}{2} \|\mathbf{u}(\mathbf{x}_r, \mathbf{x}_s; t) - \mathbf{d}_{s,r}\|^2 dt, \quad (6.1)$$

where \mathbf{u} denotes the modelled wavefield, $\mathbf{d}_{s,r}$ the seismogram recorded by the source-receiver pair (s,r) , \mathbf{x}_s and \mathbf{x}_r the source and receiver positions, respectively, and t the time. Minimisation of the misfit function (6.1) around a starting model leads to the following expression of the model perturbation at iteration k :

$$\Delta \mathbf{m}_{k+1} = -\gamma_k \mathbf{H}^{-1}(\mathbf{m}_k) \nabla_{\mathbf{m}} \chi(\mathbf{m}_k), \quad (6.2)$$

where γ denotes the step length, \mathbf{H} the Hessian (the second derivative of the misfit function) and $\nabla_{\mathbf{m}} \chi$ the gradient of the misfit function with respect to \mathbf{m} . The gradient can be computed efficiently at the cost of two forward simulations per source with the adjoint state method (Chavent, 1974) which was popularised by Tarantola (1984a) and has become of common use in global and exploration seismology (Tromp et al., 2005; Plessix, 2006a; Fichtner et al., 2006a; Liu & Tromp, 2006). The first simulation computes the incident wavefield \mathbf{u} from the source while the second one computes the adjoint wavefield by back-propagating the residuals from the receiver positions. For an elastic medium parametrized by the density ρ and the coefficients c_{ijkl} of the stiffness tensor, the gradient is simply given by the zero-lag correlation of two wavefields :

$$\begin{aligned} \nabla_{\rho} \chi(\mathbf{x}) &= \int_0^T \mathbf{u}^{\dagger}(\mathbf{x}, t) \cdot \partial_t^2 \mathbf{u}(\mathbf{x}, t) dt, \\ \nabla_{c_{ijkl}} \chi(\mathbf{x}) &= \int_0^T \boldsymbol{\epsilon}^{\dagger}(\mathbf{x}, t) : \frac{\partial \mathbf{c}}{\partial c_{ijkl}} : \boldsymbol{\epsilon}(\mathbf{x}, t) dt, \end{aligned} \quad (6.3)$$

where \mathbf{c} denotes the stiffness tensor, \mathbf{u} the displacement wavefield and $\boldsymbol{\epsilon}$ and $\boldsymbol{\epsilon}^{\dagger}$, the strain tensor and the adjoint strain tensor, respectively (Tromp et al., 2005).

While the gradient (6.3) indicates the direction of greatest ascent and depends on first derivatives of the misfit function, the Hessian is based on second derivatives of the misfit function and accounts for its curvature. The Hessian is of particular interest for FWI, since it deconvolves the gradient from limited bandwidth effects and removes the imprint of wavefield amplitudes. In the framework of multi-parameter reconstruction, it contributes to mitigating the cross-talk between multiple parameter classes.

Even though the full Hessian cannot be computed easily (Fichtner & Trampert, 2011b), its action on the gradient can be approached considering limited-memory quasi-Newton (Nocedal,

1980; Nocedal & Wright, 2006b) or truncated Newton (Métivier et al., 2013, 2014) methods. In our study, the Hessian is iteratively approximated by the L-BFGS method, a matrix-free algorithm that estimates the action of the inverse Hessian onto a vector by applying finite-differences over previous gradients and model parameters. The optimization process including the line search is implemented in our code with the SEISCOPE toolbox (Métivier & Brossier, 2016b).

6.2.3 High-resolution multi-parameter imaging

Traditional seismic imaging methods such as traveltimes tomography or receiver function techniques are focused on determining perturbations of wave velocities from the kinematic attributes of the seismic wavefield (namely, traveltimes). On the other hand, the use of the waveforms, through the information carried out by traveltimes and amplitudes of forward- and backward- scattered waves, make FWI amenable to reconstruction of second-order parameters such as density (ρ) in addition to the P and S wave velocities (V_p and V_s). Inferences on the sensitivity of the wavefield to different parameter classes can be drawn from the radiation pattern of elastic waves scattered by small elastic heterogeneities (Wu & Aki, 1985; Tarantola, 1986). These radiation patterns represent the amplitude of P and S waves scattered by a point perturbation of one elastic parameter as a function of the scattering angle. These radiation patterns are shown in Fig. 6.2 for an isotropic elastic medium parametrised by V_p , V_s and ρ . Such patterns give clear insights on the influence of a single parameter onto the data depending on the scattering regime (forward versus backward). Moreover, the overlap of the radiation patterns of different parameters with scattering angles indicates potential trade-off between these parameters. On the one hand, one may want to mitigate these trade-offs by choosing a parametrisation that minimises these overlaps. On the other hand, this strategy might narrow the influence of the parameters with scattering angle, hence degrading the resolution with which parameters are reconstructed.

Limiting our analysis to the (ρ, V_p, V_s) parametrisation, a first noticeable feature is that a V_p scatterer generates only P-P scattering with an isotropic radiation pattern. Therefore, the radiation pattern of V_p has no impact on the resolution with which this parameter is reconstructed (Fig 6.2d-f) : the local resolution of V_p is controlled by the local P wavelength and the range of P-P scattering angles spanned by the acquisition geometry.

In contrast, a V_s perturbation gives rise to P-P, P-S, S-S and S-P scattering, with stronger diffracted S waves relative to the P counterparts (Fig. 6.2g-h). The union of these radiation patterns covers a broad range of scattering angles, that is conducive to a broadband reconstruction of the V_s parameter if the imprint of all of these scattering modes is significant in the data. Trade-off between V_p and V_s updates is expected to be limited since the V_s parameter generates P-P scattered waves only at intermediate angles with much smaller amplitudes than those generated by V_p . Since the radiation patterns of V_p and V_s should have little influence on the resolution with which they are reconstructed, the resolution of V_s is expected to be superior to that of V_p according to their respective wavelengths.

Similarly to a V_s perturbation, a ρ perturbation generates scattering for the P-P, P-S, S-S and S-P scattering modes. However, a key difference is that ρ mainly scatters waves backward

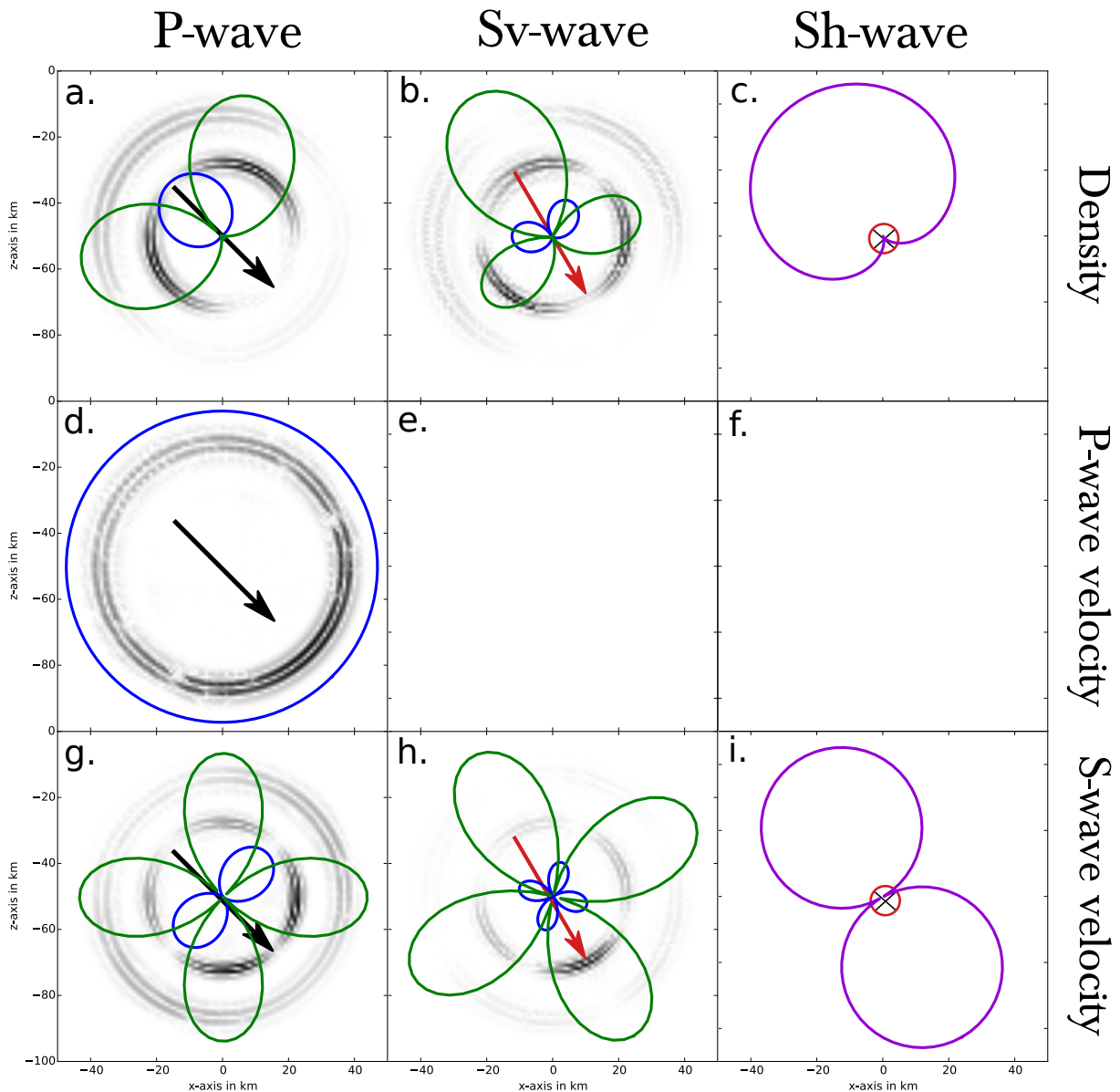


Figure 6.2 – Scattering patterns of elastic waves for the (ρ, V_p, V_s) parametrisation. From left to right, each column presents the interaction of a P (a,d,g), Sv (b,e,h) and Sh (c,f,i) incident plane wave with, from top to bottom, a ρ (a-c), V_p (d-f) and V_s (g-i) scattering point. For the seven different interactions, the analytical radiation patterns computed in the framework of the Born approximation (Wu & Aki, 1985; Tarantola, 1986) are superimposed on the scattered wavefields that have been computed numerically. Arrows indicate the direction of the incident plane wave (black for P wave, red for Sv wave and red-black for Sh wave). Radiation patterns are plotted in blue, green and purple for the P, Sv and Sh scattered modes, respectively.

(in the reflection regime). This is especially true for the P-P and S-P modes, while P-S and S-S modes show significant amplitudes at intermediate scattering angles. Overall, this deficit of sensitivity at large scattering angles can limit the ability of FWI to update the long wavelengths of ρ when the (ρ, V_p, V_s) parametrisation is used, which can be to some extent balanced by the low frequency content of teleseismic data. This comment also suggests that the second-order P and S wave sources generated by the free surface might have a leading role in the ρ reconstruction since they generate back-scattered waves from the lithospheric reflectors. The ability to image ρ by FWI is therefore prone to decrease rapidly with depth due to the smaller amplitudes of these second-order scattered waves relative to the primary counterparts.

Significant trade-offs are expected at intermediate to small scattering angles between V_p and ρ and between V_s and ρ according to their respective radiation patterns. However, it is worth remembering that V_p and V_s parameters control both the kinematic and the dynamic attributes of seismic waves, while density mostly influences amplitudes. Heuristically, we may think that the long-wavelength reconstructions of V_p and V_s will be primarily tied by the need to fit traveltimes and that amplitudes will play a second-order role in the parameter update, hence limiting the imprint of these parameter trade-offs.

In the sequel of this study, we perform FWI with the (ρ, V_p, V_s) parametrisation. Indeed, other combinations of parameters involving P and S impedance or stiffness coefficients can be viewed. A sensitivity analysis of teleseismic FWI to subsurface parametrisation will be the aim of a future study.

6.3 Application to the Western Alps

6.3.1 The CIFALPS experiment

The CIFALPS array (for China-Italy-France-ALPS seismic transect) (Zhao et al., 2015) was deployed from June 2012 to September 2013 between the Rhône valley in France and the Po plain in Italy (Fig. 6.3a). The network was composed of 56 three-components broadband stations from which 46 were forming a 350 km long transect across the Western Alps from SE France to NW Italy (Zhao et al., 2016b). The additional 10 stations were located off-profile 40 km to the North and South to fill gaps in French and Italian permanent networks (network code FR : RESIF (1995) ; network code GU : University of Genova (1967) ; network code IV : INGV Seismological Data Centre (1997)). On average, the inter-station spacing was less than 10 km for the transect with a 5 km densification in the internal Alps.

Teleseismic records gave access to a catalogue of 200 earthquakes with magnitude greater than 5.8. We deconvolved each record by its instrumental transfer function, performed a careful inspection of the dataset and selected 9 teleseismic events (Fig. 6.3b) that complied to the following criteria :

1. a range of epicentral distances between 30 to 90 degrees,
2. hypocentral depths lower than 30 km or higher than 200 km,
3. a well-discernible spatial coherency of arrivals on vertical and radial records (Fig. 6.4),

4. a high signal-to-noise ratio of individual traces.

The lower bound of the first criterion aims to mitigate the imprint in the selected time window of complex shallow multipathing arising outside the lithospheric target, while the upper bound ensures that the direct P wave is the first arrival. The second one controls the complexity of the incident wavefield with respect to source-side multiples : as the event depth increases, the time delay between direct P and pP or sP multiples increases and their travelling paths deviate. In this case, the pP and sP multiples can overprint the useful signals in the selected time window and make the estimation of the source time function and subsurface parameters more coupled and sensitive to the error on the hypocentral depth. For very deep events, however, this problem is not a concern since the arrival time of these multiples is far beyond the considered time-window (≈ 50 s).

Though most earthquakes from the selected catalogue are fairly aligned with the CIFALPS transect, the catalogue covers a broad range of magnitude, distance and spectral content. This is highlighted by the wide range of half-duration in the Global CMT solutions from 2.4 s to 35.7 s (table 6.1). This can be cumbersome for FWI, particularly at high frequency, in the sense that events will have contrasted spectral contents. Considering the nearly planar nature of teleseismic wavefields, it is crucial for imaging that events cover a wide range of epicentral distances to sample a wide range of incidence angles (Pageot et al., 2013). Typically, teleseismic wavefields are dominated by the incident P and source-side depth multiples (Fig. 6.4). These are followed by a mixture of lithospheric converted phases such as Ps conversion from the Moho discontinuity, the lithosphere-asthenosphere boundary or 410 and 660 mantle discontinuities. The key challenge of FWI is to extract the information carried out by these secondary signals, which are hidden in the P-wave coda, to enhance the resolution of lithospheric models.

6.3.2 Full-waveform inversion workflow

Before FWI, we computed and stored the wavefields computed with *AxiSEM* on the side and bottom boundaries of the lithospheric target. Then, we estimated a source signature for each individual earthquake by performing a least-squares frequency-domain deconvolution (Pratt, 1999) of the recorded vertical seismograms from the synthetic counterparts computed in the initial FWI model. As initial model we use a smoothed version of the *PREM* reference model (Dziewonski & Anderson, 1981). Note that the topography was not accounted for in our simulations. An example of the data fit obtained with the initial model is shown in Fig. 6.5c-e. Interestingly, the first arrival of the synthetic seismograms computed in the initial model are already in fair agreement with the recorded ones for vertical, radial and transverse components. A partial conclusion is that a layered one-dimensional reference Earth model and automatic moment tensor solutions refined by an estimation of the source time function provide the necessary ingredients to predict the low frequency content of the data without any risk of cycle skipping. Nonetheless, slight shifts and amplitude mismatches are visible in the early arrival and more significant waveform differences, whose lateral coherency can be followed from one stations to the next, are showing up for later arrivals testifying the complexity of receiver-side lithospheric structures (Fig. 6.5i-k).

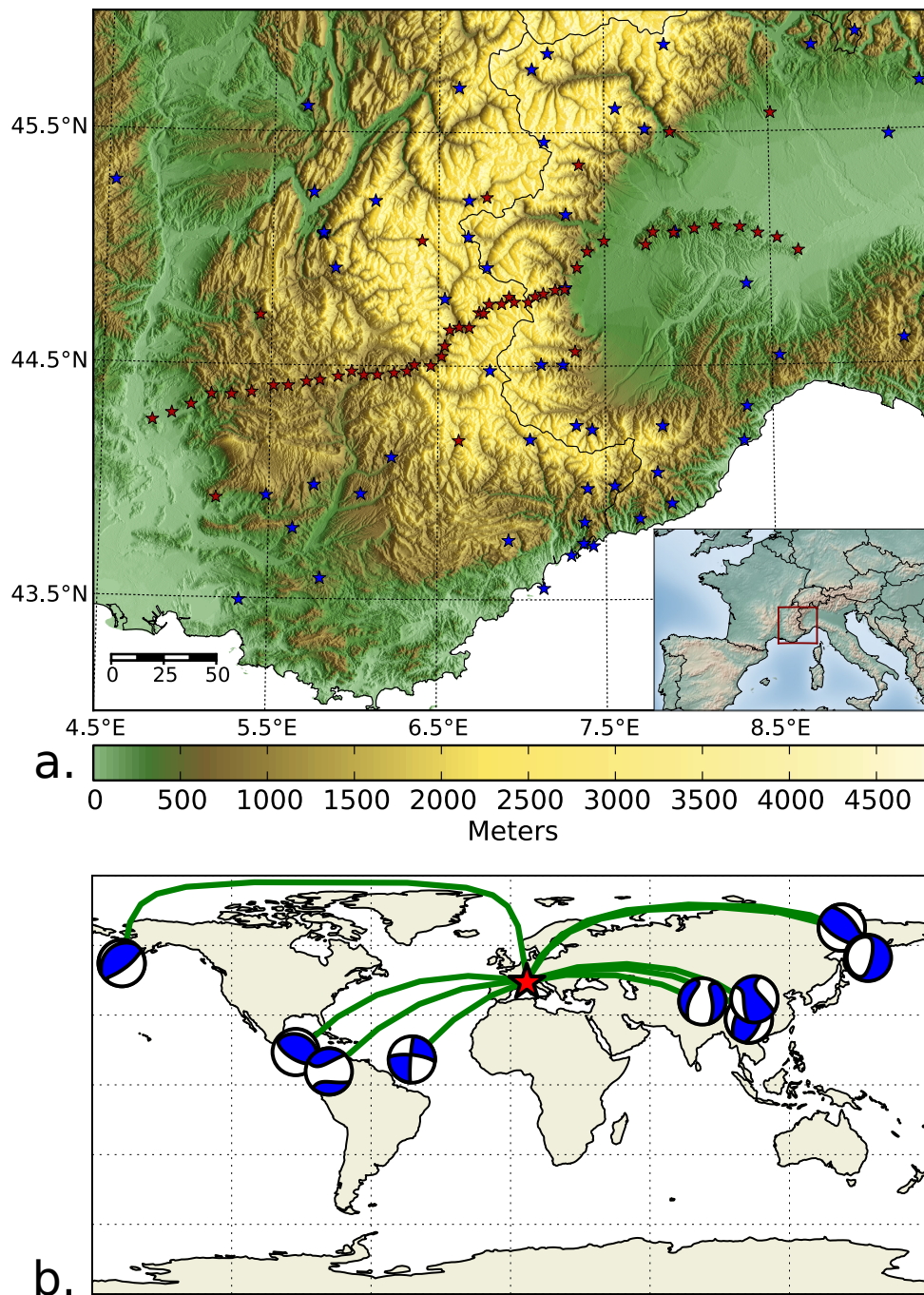


Figure 6.3 – The CIFALPS experiment. (a) Topography of the Western Alps and source-receiver layout of the CIFALPS experiment. Red stars stand for CIFALPS temporary stations whilst blue stars denote local permanent stations. (b) The selected 9 teleseismic events symbolised by their respective moment tensor solutions.

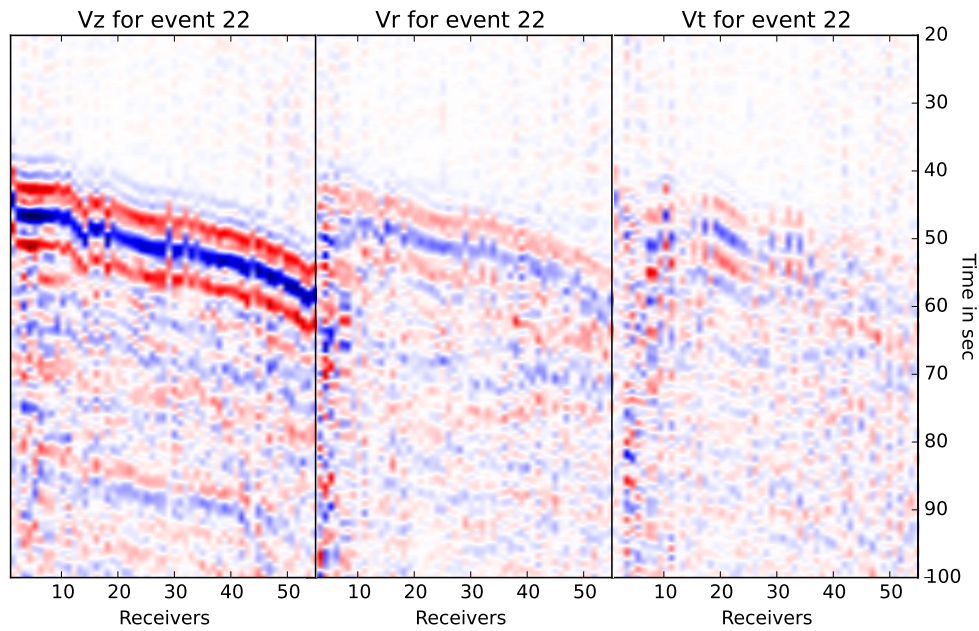


Figure 6.4 – Records of event 22 for the vertical, transverse and radial components. Note some coherent signal hidden within the P-wave coda related to receiver side lithospheric structure (reflections and conversions).

Id	Event	Lat (°)	Dist (°)	Hdur	Depth	Mw	CMT
		Lon (°)	Baz (°)	Tshift	(km)		
06	2012-08-10 18 :37 :43.1	52.39	82.99	3.00	13	6.2	
		-167.32	356.36	5.80			
08	2012-08-12 10 :47 :06.4	35.88	56.50	3.10	17	6.3	
		82.52	70.58	5.60			
16	2012-11-07 16 :35 :46.7	14.11	86.58	11.90	21	7.4	
		-92.43	286.53	9.60			
22	2013-02-14 13 :13 :53.1	67.65	62.98	5.30	12	6.7	
		142.512	17.37	6.26			
26	2013-04-20 00 :02 :47.5	30.22	73.54	4.90	22	6.8	
		103.12	63.61	5.57			
31	2013-05-24 05 :44 :49.0	54.61	77.02	35.70	611	8.4	
		153.77	18.89	19.35			
36	2013-06-24 22 :04 :13.5	10.91	53.91	4.20	18	6.6	
		-42.46	247.06	4.40			
39	2013-07-21 23 :45 :56.7	34.6	71.35	2.40	17	6.2	
		104.31	59.51	1.95			
41	2013-08-13 15 :43 :15.2	5.77	82.40	5.00	21	6.7	
		-78.19	270.70	5.91			

Table 6.1 – Selected events for teleseismic waveform inversions indicating origin time, location and moment tensor solution from Global CMT project (Dziewonski et al., 1981; Ekström et al., 2012). Half duration (Hdur) and time shift (Tshift) are given in seconds.

Afterwards, we perform a multi-scale FWI of the vertical and radial components by successive inversions of data sets of increasingly high-frequency content with cut-off periods of 20, 10 and 5 seconds. Although transverse component data are accessible and may carry information related to the 3D lithospheric structure geometry (Cassidy, 1992), they are usually influenced by anisotropic effects from the lithospheric medium (Frederiksen & Bostock, 2000) and we safely decided to omit them for FWI. In addition to the frequency continuation strategy, a second level of hierarchy can be implemented by progressively increasing the time window involved in the inversion (Monteiller et al., 2015). This time continuation strategy can help mitigating the non-linearity of the FWI by injecting progressively more complex information in the inversion such as back-scattered waves. However, according to the variability of our teleseismic catalogue, notably the heterogeneous spectral content underlined by rather long duration of source-time functions; and by considering the theoretical traveltimes of the lithospheric phases, we used a fixed 50 s-long time window starting 5 seconds before the P-wave onset time. We also make the inversion equally sensitive to each event (whose data amplitudes are mainly controlled by their epicentral distance and moment magnitude) by normalising common-event seismograms and their corresponding incident wavefields in each frequency band.

The reduction of the misfit functional over iterations is shown in Fig. 6.5b. FWI converges after 15, 20, and 24 iterations for cut-off periods of 20, 10 and 5 seconds, respectively, and achieves an overall 54 % misfit reduction in 59 iterations. It is worth noting that we penalised the misfit functional with a Tikhonov regularization term (Tikhonov & Arsenin, 1977) implemented by minimising the Laplacian of the model parameters. Without this regularization term, the data misfit can be decreased by more than 54 % with however the convergence toward an unreliable ρ model and without significant improvements of the V_p and V_s models. We interpret this behaviour as a data over-fitting issue when FWI starts minimising noise in the data. This noise generates significant waveform residuals due to inconsistent amplitudes that can be more easily accommodated by second-order parameters that are mostly sensitive to amplitudes such as density (Przebindowska et al., 2012). This statement is supported by a χ^2 analysis. Assuming an overall data noise standard deviation equal to 10% of the maximum recorded amplitude of each event and that all data are matched within this confidence interval, the value of the χ^2 criterion would correspond to a misfit reduction that is approximately 50%.

The data fit achieved by the the final FWI model for the event 22 is shown in Fig. 6.5f-h. Synthetic and recorded seismograms are now in remarkable agreement at least for the vertical and radial components. The data fitting procedure has not only removed subtle phase shifts between the recorded and modelled first arrivals but also more prominent residuals of strong secondary wiggles. A more quantitative assessment of the data fit is provided by the data residuals computed in the starting and final FWI models (Figs. 6.5e-h and 6.5i-n). Residuals decrease on the three components except for highly noisy stations (number 1 to 10) located in the Po plain sedimentary basin, which generates strong and local amplification or attenuation of the wavefields (Stehly et al., 2009). At these locations, the assumption of non-attenuating isotropic medium is likely to be unrealistic. Noticeably, the residuals of the transverse component are significantly decreased, although this component was not involved in the inversion (Fig. 6.5k,n). From the misfit reduction computed separately for each component and each event (Fig. 6.5a), we note that such a decrease is mainly shown for earthquakes (events 6, 22 and 31) whose transverse component is oriented in the azimuth of the CIFALPS transect. For event 6, the transverse

component is even better fitted than the radial one. We also note that, on average, the vertical component is better fitted than radial and transverse components.

6.3.3 FWI models

The vertical sections of the final FWI models along the CIFALPS transect are displayed in Fig. 6.6(a,c,e) for the V_p , V_s and ρ parameters. This showing is complemented by that of the model perturbations (computed with respect to the initial model) to highlight the resolution improvement achieved by FWI (Fig. 6.6(b,d,f)). The model perturbations roughly remove the imprint of the low-wavenumber components from the FWI model and, hence, mimic the intermediate wavenumber components that could be imaged by the migration of receiver functions. We superimpose on these vertical sections the structural interpretation proposed by Zhao et al. (2015) to facilitate the identification of the main structural units (Fig. 6.6, black lines). The ρ , V_p and V_s models show similar structures with however their own resolution. As expected, the V_p model is less resolved than the V_s one according to the P and S wavelengths. In contrast, the ρ and V_s models show a similar high-wavenumber content in the upper part of the lithospheric target, which is consistent with the analysis of their radiation patterns (Fig. 6.2). However, the FWI clearly fails to significantly update ρ beyond 70 km depth, unlike V_s . As suggested in section 2, this might result from the deficit of wavefield sensitivity to ρ perturbations in the forward-scattering regime. According to the teleseismic setting, this forward scattering is more suitable to update the deep part of the lithosphere. Since back-scattered waves should undergo two reflections from the free surface and the lithospheric reflectors to sample the deep lithosphere, it is unlikely that such second-order scattered waves can be recorded with a sufficient signal-to-noise ratio for FWI.

Of primary attention is the Moho discontinuity. Being a first order seismic discontinuity this boundary generally exhibits strong impedance gradients that are nicely depicted by receiver functions computations and migrations. Our models delineate clear but rather complex Moho geometries beneath the Alpine orogeny. On the western part, the European Moho dips from roughly 30 km in the Rhône valley to 50 km beneath the frontal penninic thrust (FPT, see figure 6.6). Further East, its dip increases and the Moho reaches a depth of around 65 km (on V_p and ρ models) at the transition between the internal Alps and the Po plain underlining a high density / fast velocity anomaly corresponding to mantle material within the crust (blue blob on V_p model). On the V_s model, the Moho can be followed at greater depths, down to approximately 80 km. At this depth, the Moho is interrupted by a slow velocity anomaly that spreads horizontally from 50 to roughly 100 km of distance. On the eastern side, the Moho transition is more difficult to follow and exhibits strong topographic variations. Nonetheless, absolute velocity variations (Fig. 6.6c-e) suggest a 40-50 km deep Moho boundary beneath the Po plain. In between, the area corresponding to the Ivrea body gravity anomaly is well delineated by a high density / fast velocity anomaly. In particular, the top-eastern boundary of this anomaly matches well the shape of the Ivrea body in the interpretative model of Zhao et al. (2015) (Fig. 6.6b). However, its top-western boundary highlights a very different geometry with a pronounced sub-vertical orientation (Fig. 6.6f). The horizontal extent of this feature is depicted at multiple depths in figure 6.7 by a characteristic North-South pattern.

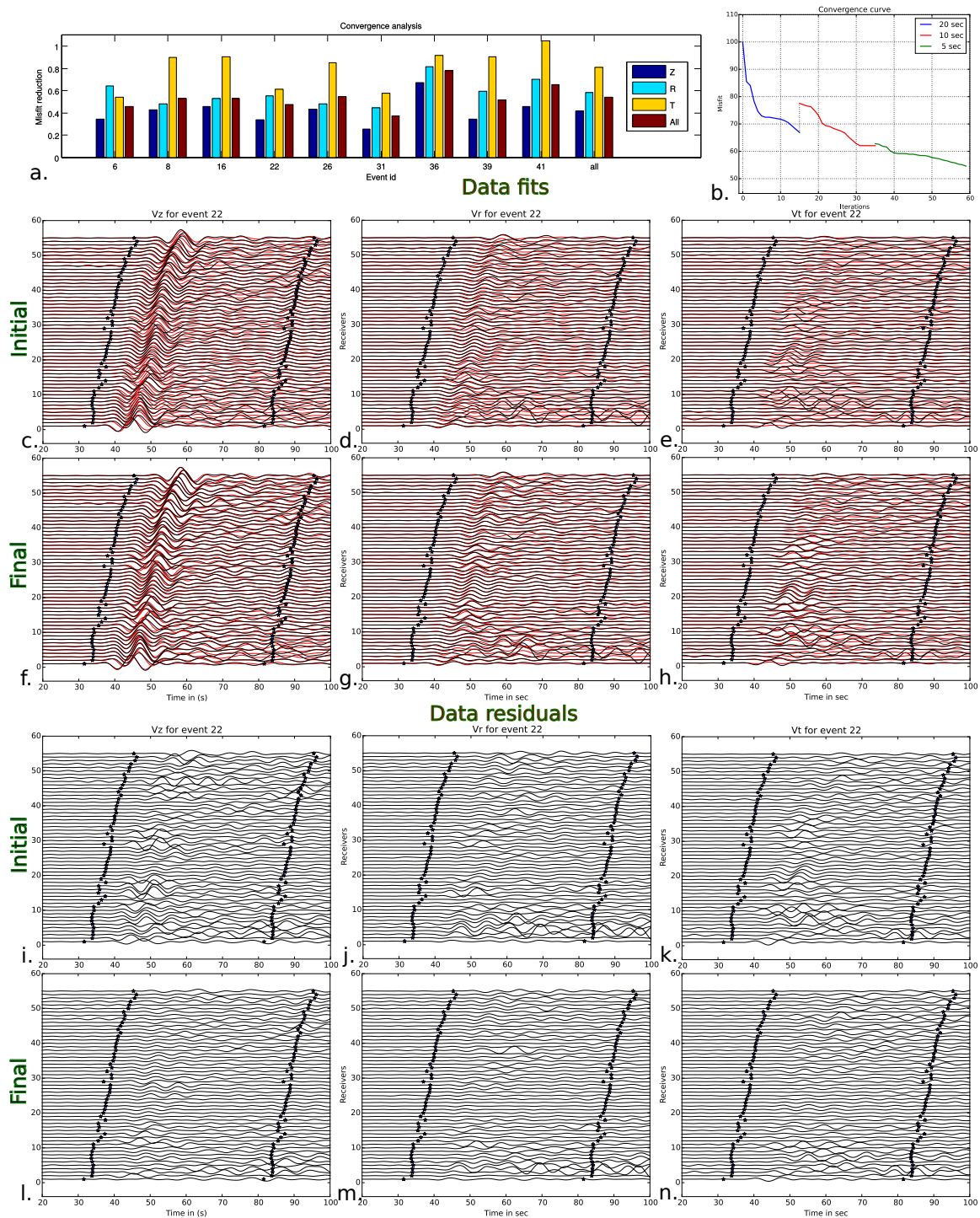


Figure 6.5 – Data fit. (a.) Misfit reduction for each component of each event (the right bar shows the misfit function for all components). (b.) Overall misfit reduction over iterations. Dash black lines indicate when FWI jumps to a new frequency band (see text for details). (c-n) Direct comparison between recorded (black) and modelled (red) seismograms for event 22 (Fig. 6.3c). (c-e) Modelled seismograms are computed in the initial model for vertical (c), radial (d) and transverse (e) components. (f-h) Same as (c-e) modelled seismograms computed in the final FWI model. (i-n) Data residuals computed in the initial (i-k) and final (l-n) FWI models, respectively. Blue stars define the time window used for FWI.

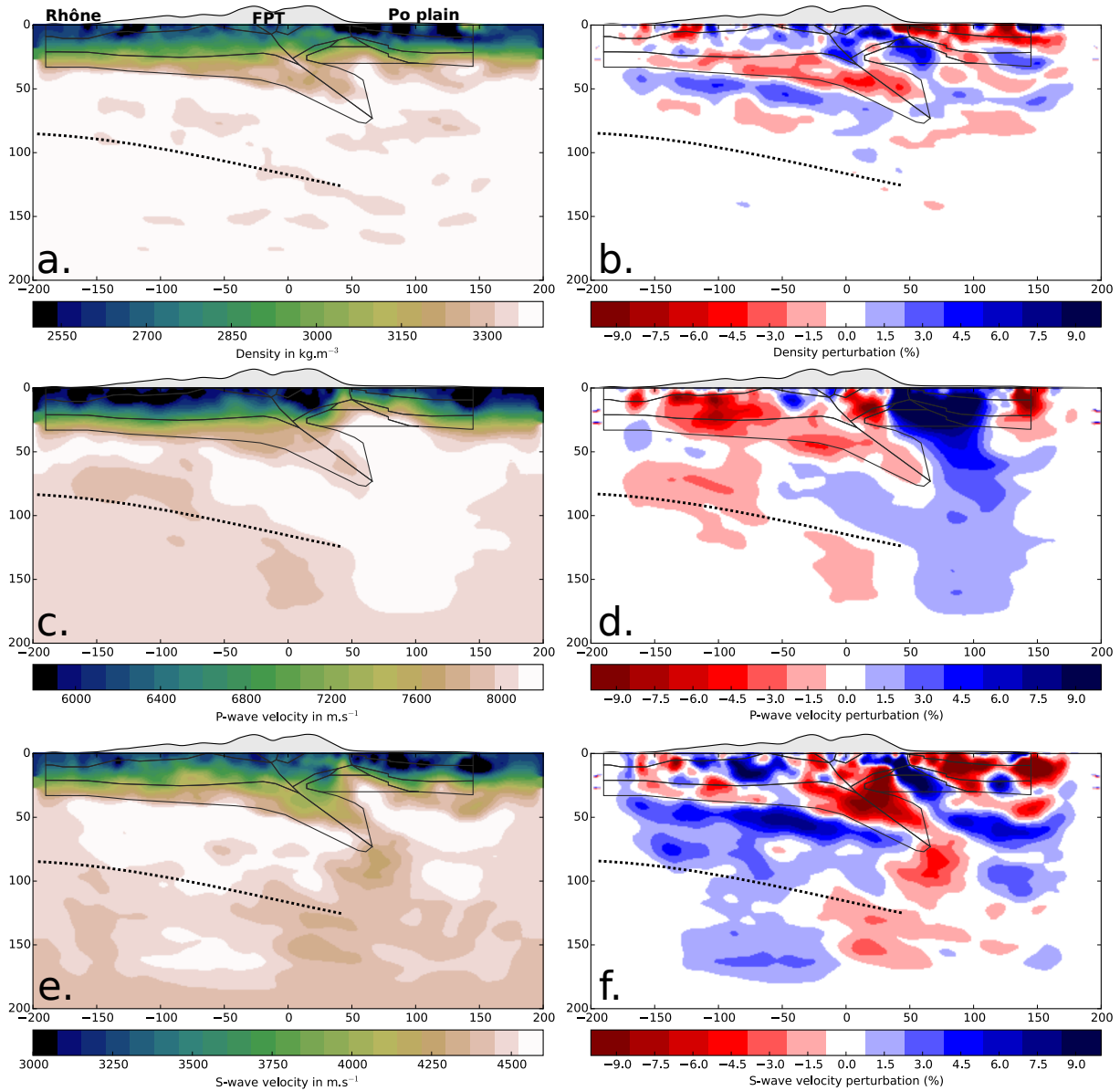


Figure 6.6 – Vertical sections of the FWI models along the CIFALPS transect (W-E). (a-b) ρ . (c-d) V_p . (e-f) V_s . The absolute values of the parameters and the perturbations are shown in (a,c,e) and (b,d,f), respectively. Black lines denote units as defined by Zhao et al. (2015). Perturbation models are computed with respect to the initial model. The topography is shown with a vertical exaggeration for illustration purpose. The black dashed line shows the location of the european LAB.

At shallower depth, from the surface to 20 km, the three models exhibit strong parameter perturbations. Caution must be taken examining these anomalies that are locally polluted by the acquisition footprint which takes the form of singularities at receiver locations. Nonetheless lateral structures can be followed that are consistent with the location and shape of sedimentary basins. Notably all parameters show strong negative perturbations – 2000 to 2300 kg.m⁻³, 5000 to 5500 m.s⁻¹ and 2600 to 2800 m.s⁻¹ for ρ , V_p and V_s , respectively – in the Po plain that correlates well with the basement of the sedimentary basins assumed to be 10 km thick. A similar behaviour is visible for the Rhône valley in the western part where slow P and S-wave (5500 m.s⁻¹ and 3000 m.s⁻¹) velocities are encountered.

At the upper-mantle scale, our models show small amplitude variations. Nonetheless, at depths ranging from 80 km to 120 km, we can follow the lateral continuity of a vertical transition corresponding to a decrease of the S-wave velocity. This transition may correspond to the lithosphere-asthenosphere boundary (LAB) (Kissling & Spakman, 1996; Eaton et al., 2009; Fischer et al., 2010). The depth of the LAB would thus vary from 80-90 km beneath the Rhône valley to 100-120 km at the transition between internal Alps and the Po plain. Interestingly, the structure of the upper-mantle is sensibly different on the V_p and V_s models beneath the Ivrea zone. While a wide-spread high velocity anomaly characterises the V_p model, a 50 km blob of low V_s is imaged just beneath the dipping European Moho. Laterally, this anomaly follows the arc of the Western Alps from 100 km to 140 km depth (Fig 6.7).

6.3.4 Resolution and parameter cross-talk appraisal

Before discussing the geodynamical implications that can be drawn from the FWI models, it is worth appraising the ability of the teleseismic waveform inversion to resolve small-scale parameter perturbations by mean of resolution tests. Even though these resolution tests, which generally rely on local analysis around the minimum of the misfit function, cannot reveal whether the inversion has approached the global minimum of the misfit function (Lévêque et al., 1993; Fichtner & Trampert, 2011b; Trampert et al., 2013; Rawlinson & Spakman, 2016), they still provide a cheap manner of assessing the resolution with which heterogeneities were reconstructed. As explained in section 2, FWI should be able to resolve anomalies whose sizes are of the order of the minimum propagated wavelength. At lithospheric scale, considering averaged V_p and V_s values of 7500 and 4000 m.s⁻¹, and considering a minimal period of 5s, we expect our FWI model to catch heterogeneities the sizes of which are of the order of 20 km. To verify this assumption and assess the spatial resolution of the FWI model, we designed a first resolution test by superimposing onto the final FWI model a coarse grid of 20 km-wide Gaussian (ρ, V_p, V_s) inclusions with positive and negative perturbations in alternates (± 100 m.s⁻¹ for V_p and V_s and ± 100 kg.m⁻³ for ρ). The spacing between the inclusions is sufficiently large and their amplitude sufficiently small to prevent a significant imprint of multi-scattering (non-linearity) during FWI.

We use the final FWI model as initial model of the resolution test such that this initial model reproduces all of the wave phenomena that occurred during the latter stage of the FWI, a necessary condition for a reliable resolution appraisal (Rawlinson & Spakman, 2016). Synthetic data computed in the FWI model with the inclusions are then inverted following the same strategy as

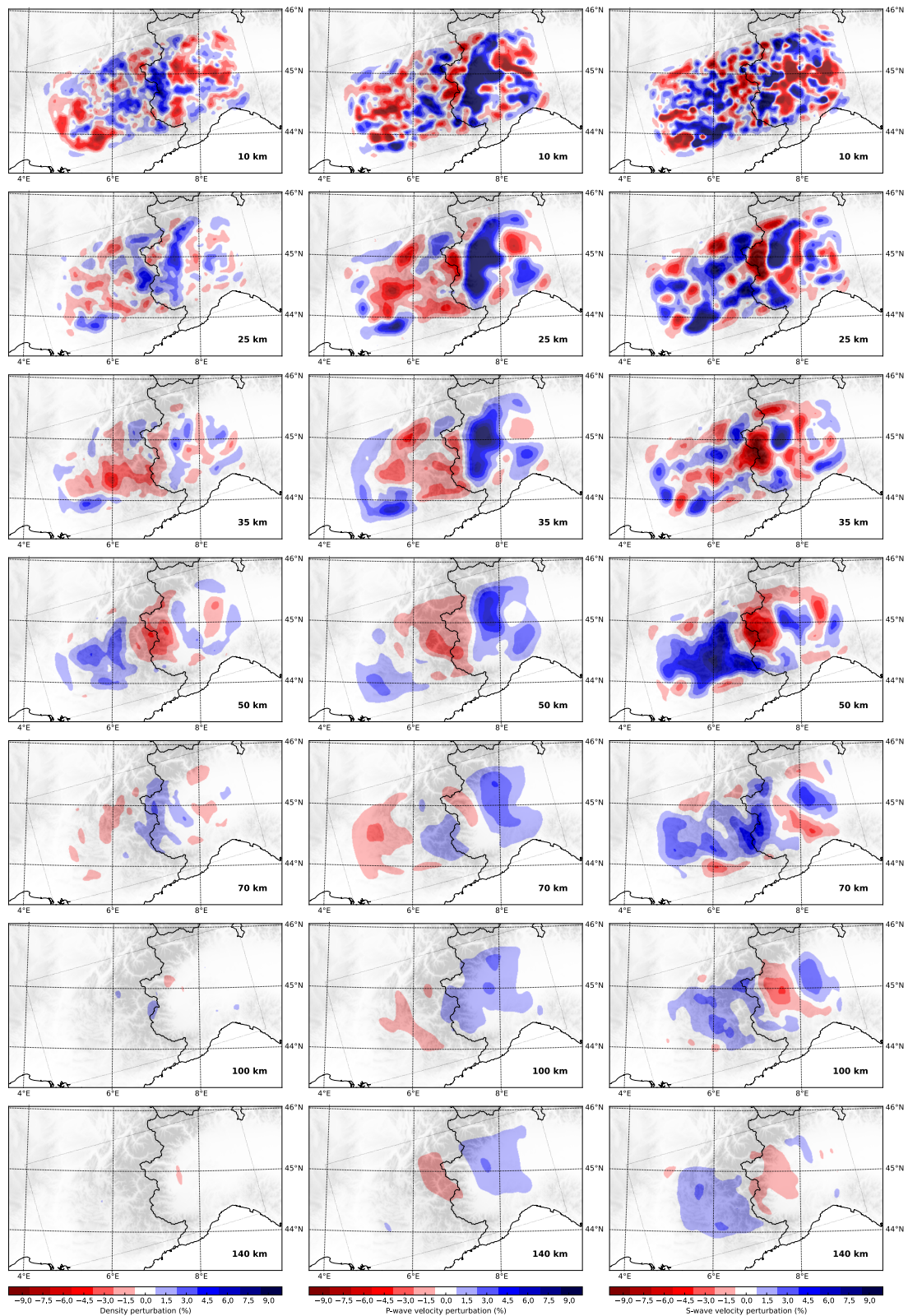


Figure 6.7 – Depth slices of the FWI perturbation models. From left to right, ρ , V_p and V_s parameters. From top to bottom, slice at 10, 25, 35, 50, 70, 100 and 140 km depth. Perturbations are computed with respect to the initial model.

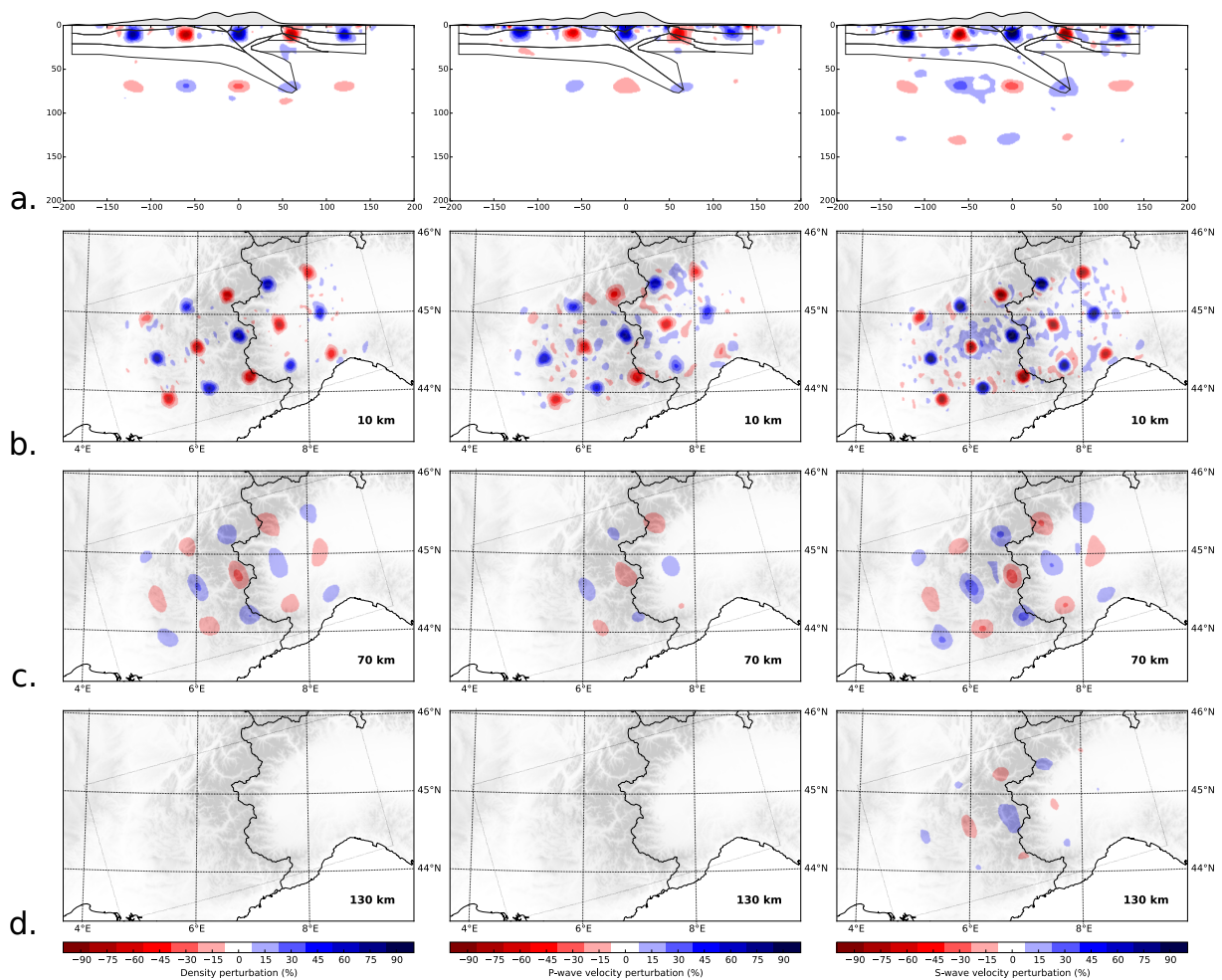


Figure 6.8 – Resolution analysis. Reconstruction of fast and slow (ρ, V_p, V_s) Gaussian inclusions by multi-parameter FWI (see text for details). From left to right, ρ, V_p and V_s FWI perturbation models. (a) Vertical section along the CIFALPS transect. (b-d) Depth slices at 10 km (b), 70 km (c) and 130 km (d) depth. The colour scale corresponds to the percentage of the true inclusion amplitude (100 kg.m^{-3} for V_p and V_s and 100 m.s^{-1} for ρ).

for the real data application with the final FWI model as initial model. The reconstructed inclusions are shown in Fig. 6.8. As expected from the teleseismic configuration, the ability of FWI to resolve parameter perturbations decreases with depth. While all of the three parameters are well-retrieved in amplitude and shape at 10 km depth, which is consistent with the theoretical resolution of FWI, FWI hardly recovers the true shape and amplitude of the anomalies at 70 km depth with horizontal and vertical smearing effects. This results from the combination of several factors. First, the increase of the wavelengths with depth decreases the resolution. Second, the inversions becomes increasingly ill-conditioned with depth due to incomplete correction for wave amplitude effects. This can make the imprint of deeply located scatterers in the wavefield too small to be exploited by the inversion. Third, the amount and strength of scattering might become more limited with depth as on the one hand the resolution of the FWI decreases and on the other hand the real heterogeneities of the deep lithosphere become milder and the free surface more distant. At 130 km depth, only the V_s parameter seems to be resolvable for 20 km-wide anomalies. This again might result on the one hand from the lack of penetration in depth of the ρ reconstruction and on the other hand to the higher wavelength of V_p relative to V_s .

It is also worth noting that, despite the one-dimensional nature of the receiver layout, our FWI model successfully recovers horizontal sections of the anomalies down to 70 km depth. Surprisingly, the depth slice at 10 km depth seems not affected by aliasing artefacts in spite of the poor sampling of the receivers in the direction perpendicular to the CIFALPS transect (Fig. 6.9b,c). To further assess how the resolution evolves with depth, we designed two additional resolution tests with only one 20 km-wide fast (ρ, V_p, V_s) inclusion located at 50 and 100 km depth, respectively (Fig. 6.9a,b). We use one inclusion per test to ensure that the reconstruction of this inclusion is not impacted upon by other close heterogeneities. The V_p FWI perturbation model shows significant vertical smearing supporting the fact that the vertical wavenumber components of this parameter are mainly reconstructed by forward scattering. In contrast, the ρ FWI perturbation model shows a horizontal smearing indicating a better vertical resolution relative to the horizontal one. This is consistent with the fact that the ρ reconstruction is mainly sensitive to back scattering generated by the double interaction of the wavefield with the free surface and the inclusion. This is well illustrated by the vertical alternates of positive and negative perturbations in the reconstruction at 100 km depth that highlights the deficit of low vertical wavenumbers in the ρ perturbation model.

In the framework of multi-parameter reconstruction, it is crucial to assess to which extent the FWI models are affected by parameter cross-talk. We perform three multi-parameter FWI by considering for each test a single-parameter inclusion involving V_p , V_s and ρ , respectively (Fig. 6.9c-e). Any ghost of the inclusion in the reconstruction of a parameter that has no imprint in the data residuals will be the expression of the parameter cross-talk. The reconstruction of the ρ and V_p inclusions do not generate significant leakage in the models of the two other parameters (Fig. 6.9c-d). In contrast, the reconstruction of the V_s inclusion generates some moderate leakage in the ρ and V_p perturbation models in the form of small negative perturbations. However, this leakage remains small, and prompts us to conclude that the (ρ, V_p, V_s) parametrisation is reliable for teleseismic FWI.

A final resolution test is performed to check the reliability of the negative anomaly located

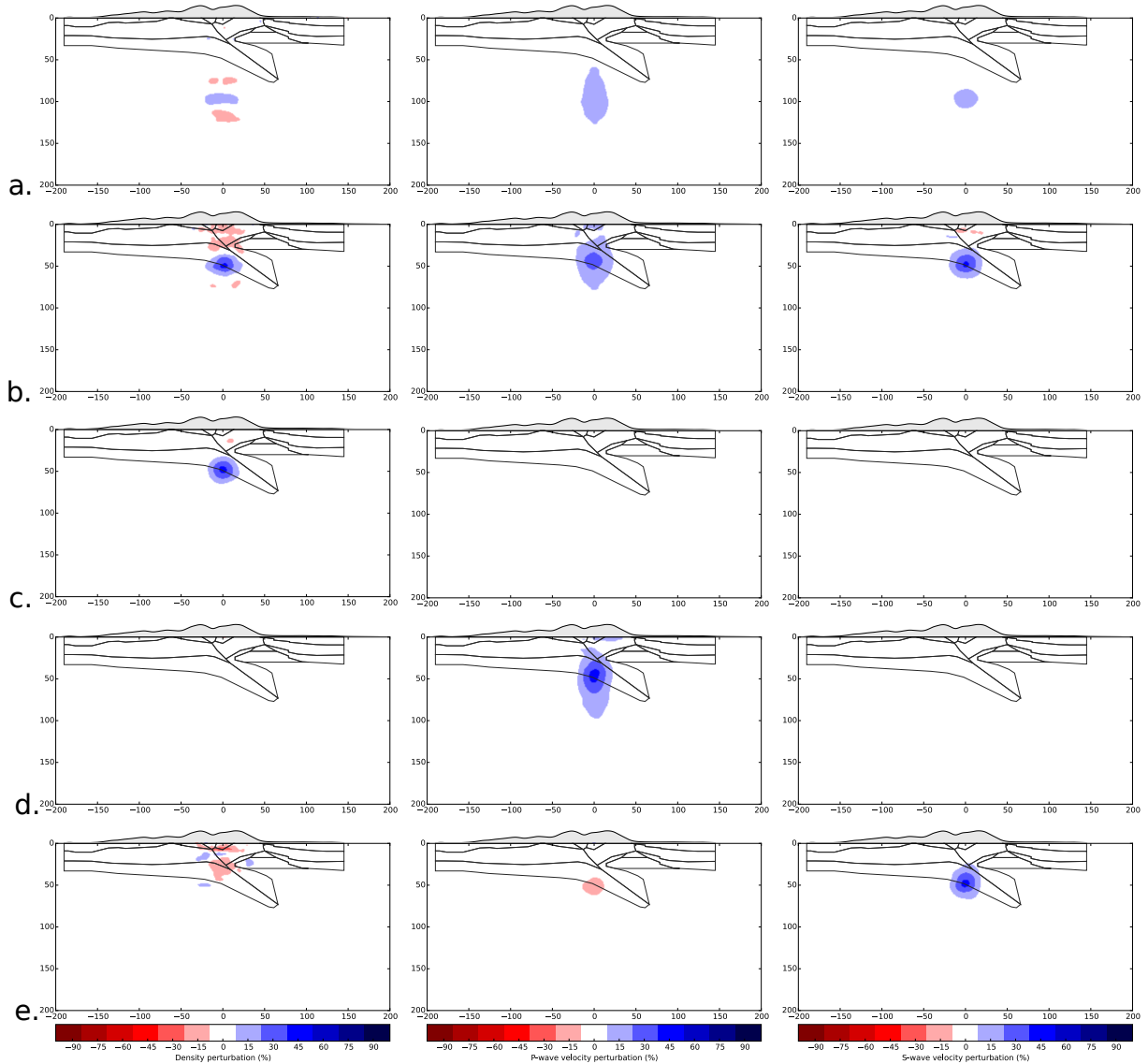


Figure 6.9 – Resolution analysis and parameter cross-talk appraisal. From left to right, vertical section along the CIFALPS transect of the ρ , V_p and V_s FWI perturbation models. All of the panels show results of multi-parameter FWI for ρ , V_p and V_s updates for either multi-parameter (a-b) or mono-parameter (c-e) inclusion models (see text for details). (a-b) Resolution analysis. Reconstruction of a fast 20 km-wide (ρ, V_p, V_s) inclusion at 100 km (a) and 50 km (b) depths (c-e) Parameter trade-off. (c) Reconstruction of a fast 20 km-wide ρ inclusion at 50 km depth. (d) Same as (c) for a V_p inclusion. (e) Same as (c) for a V_s inclusion. Ghost of the inclusion in V_p and V_s models of (c), in ρ and V_s models of (d) and in ρ and V_p models of (e) highlight parameter cross-talk. The colour scale corresponds to the percentage of the true inclusion amplitude (100 kg.m^{-3} for V_p and V_s and 100 m.s^{-1} for ρ).

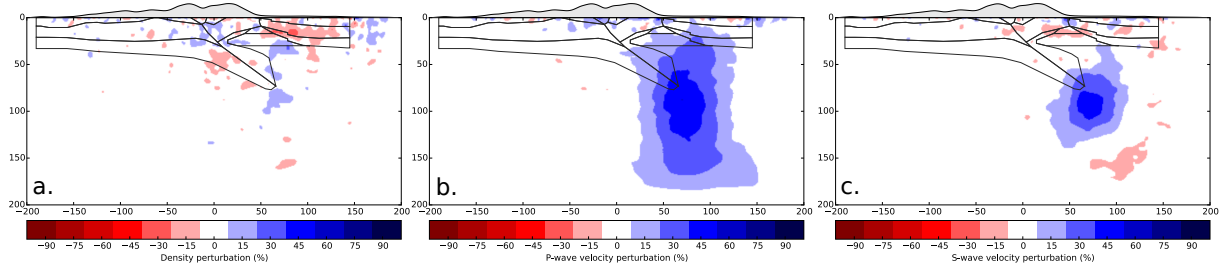


Figure 6.10 – Resolution analysis for a 50 km-wide (ρ, V_p, V_s) inclusion located at the position of the negative V_s anomaly shown in Fig. 6.6c,e. (a-c) FWI perturbation model for ρ (a), V_p (b) and V_s (c). The colour scale corresponds to the percentage of the true inclusion amplitude (100 kg.m^{-3} for V_p and V_s and 100 m.s^{-1} for ρ).

at 100 km beneath the Ivrea zone in the V_s FWI model (Fig. 6.6c,e). We add a 50 km-wide fast inclusion to the final V_p, V_s and ρ FWI models. The size of the inclusion was chosen according to that of the anomaly shown in the V_s FWI model of Fig. 6.6c,e. Results are presented in Fig. 6.10. Again, the shape of the anomaly is nicely recovered in the V_s FWI perturbation model, whereas the V_p reconstruction is affected by strong vertical smearing. For such large anomaly at rather great depth, FWI fails in recovering the ρ perturbation, as expected. A partial conclusion is that the S-wave anomaly can be trusted while the P-wave one must be interpreted with caution in view of its intrinsic low vertical resolution.

6.4 Discussion

6.4.1 Density parameter reconstruction

Through the analysis of scattering pattern in section 2.3 and the recovery of sensitivity models in section 3.4, we have seen that the joint use of amplitude and phase informations contained in the waveform data enables the resolution of the density. The reconstruction of the density parameter is however restricted to the shallow part of the lithospheric model since it might require the recording of a doubly back-scattered wavefield : reflected from the free surface and from a lithospheric heterogeneity. We also emphasized in section 3.2, the inherent sensitivity of the density to the noise content of the data. It is therefore relevant studying to which extent the update of V_p and V_s during FWI is sensitive to the way in which density is involved in the inversion, namely, as a passive parameter that is kept fixed over iteration as in Wang et al. (2016) or as an optimization parameter as in the current study. Figure 6.11 presents the results of the FWI of CIFALPS datasets obtained by updating only V_p and V_s and by keeping ρ fixed to the initial smoothed PREM model. A quick comparison of these results with those of figure 6.6 indicates that the update of the ρ parameter does not significantly impact the reconstruction of the V_p and V_s parameters, that are very close to those obtained without involving the ρ parameter in the inversion. We conclude that accounting for ρ in teleseismic FWI does not degrade nor improve the recovery of kinematic parameters. Beyond that, it may allow for (shallow) inference on the density model through the use of the amplitude information.

This statement is consistent with the conclusions of [Prioux et al. \(2013a\)](#), their figure 9) drawn from FWI applications on controlled-source seismic data. We thus believe it is worth updating the density to mitigate the effects of incorrect density information when no prior information on density is available ([Plonka et al., 2016](#)). This is supported by [Yuan et al. \(2015\)](#) who concluded that simultaneous update of ρ , V_p and V_s starting from an erroneous density model gives V_p and V_s models that are comparable to those we would obtain using the exact density information.

The relevance of the FWI ρ model of the Western Alps is further supported by the consistency between the measured and the computed regional Bouguer anomalies of the Alp, this latter being inferred from our FWI density model (Fig. 6.12). The Bouguer anomaly map of the Western Alps is dominated by negative anomalies that characterise the mass deficit generated by the thick crustal root of the Alps. These less dense crustal materials manifest in our density model as reddish negative anomalies as expected. This negative trend is flanked by a positive gravity anomaly associated with the North-South oriented Ivrea body, the shape of which (green lines on Fig. 6.12) is in good agreement with a high density anomaly at 25 km depth in our density model.

6.4.2 Geodynamic interpretation

The Western Alps result from the convergence of Adriatic and European margins over the past 40 Ma and points to several subduction events since 100 Ma. Despite numerous geophysical studies, an integrated interpretation of surface observations, crustal and upper-mantle structures remains difficult because of complex 3D structures and of the various sensitivities and resolutions of available imaging techniques that further complicate the understanding of the geodynamics evolution of the Alps. We believe that our FWI model provides a good compromise to fill the resolution gap between high-resolution images of the crust obtained from CSS, receiver functions and LET studies and lower resolution traveltimes tomography images of the upper-mantle.

For instance, our models of the Alpine lithosphere give a concise overview of the known Western Alps crustal structures. Notably, as seen in previous LET studies ([Paul et al., 2001](#); [Diehl et al., 2009](#)), the structure of the upper crust is dominated by the high-velocity / high-density Ivrea body. This feature can be continuously traced down from 10 km to 50 km depth and appears to be connected to the Adriatic mantle as mentioned by [Nicolas et al. \(1990\)](#) and [Schmid & Kissling \(2000\)](#) who interpreted it as a mantle wedge indenter. The relatively high V_p ($\approx 7400 \text{ m.s}^{-1}$) together with a high V_p/V_s ratios clearly indicate the upper-mantle origin of this body. The shape of the ρ anomaly nicely follows the limits of the upper-crust model of [Zhao et al. \(2015\)](#). In agreement with [Potin et al. \(2015\)](#), the horizontal extent of the Ivrea body seems to be separated into two independent blocks : one to the North near Torino and the other one to the South near Cuneo (see V_p and particularly V_s at 25-35 km depth in Fig. 6.7).

The question of the Moho geometry has long been an issue in the Western part of the Alps ([Laubscher, 1990](#); [Marchant & Stampfli, 1997](#); [Lombardi et al., 2008](#); [Spada et al., 2013](#)). Evidences of continental subduction such as HP-LT coesites found by [Chopin \(1984\)](#) in the Dora Maira massif indicates continental crust burial at a depth of the order of 100 km. Neither LET nor teleseismic tomography studies had a sufficient resolution to precisely image this first order

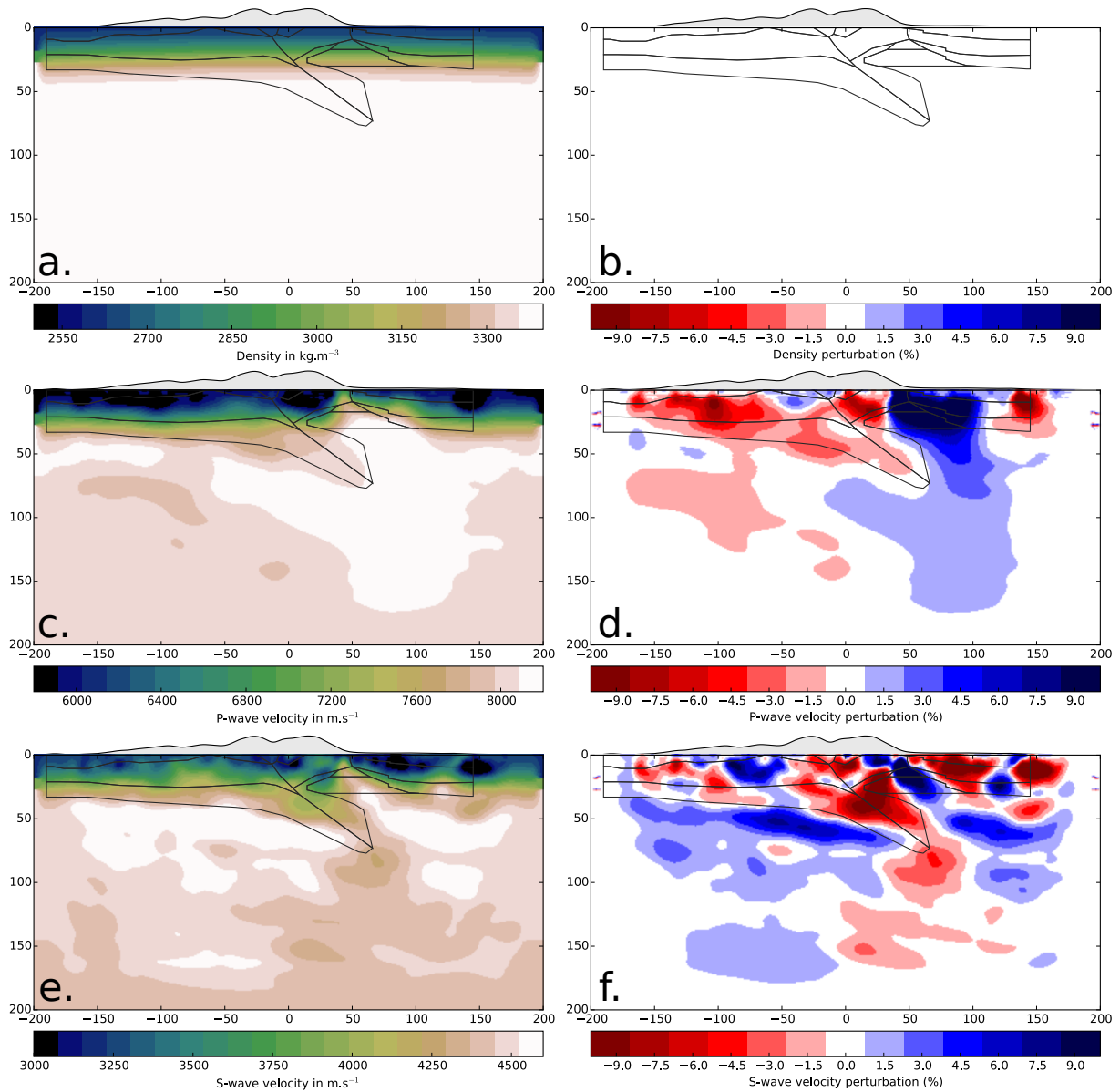


Figure 6.11 – Same as Fig. 6.6 except that the density model is kept fixed to the initial smoothed PREM model : only V_p and V_s are updated during FWI.

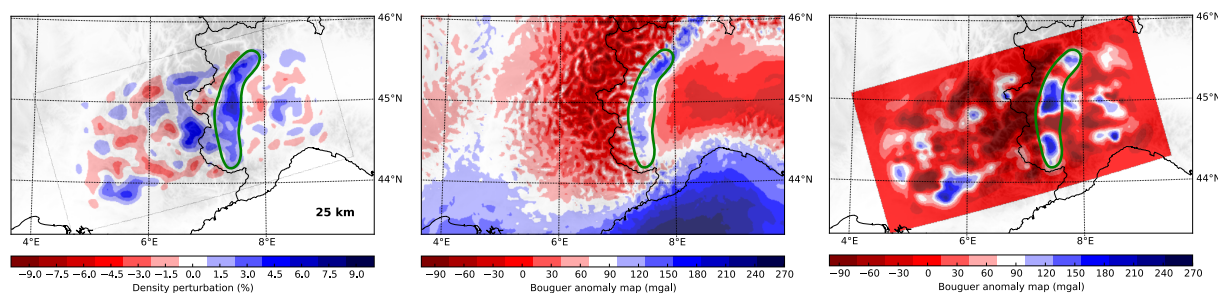


Figure 6.12 – Comparison of the density perturbation model at 25 km depth, the regional Bouguer Anomaly map from Balmino et al. (2012) and the Bouguer anomaly map computed from our density perturbation model. The green line highlights the horizontal geometry of the Ivrea body.

discontinuity of the crust that was finally imaged thanks to receiver function migration (Zhao et al., 2015) along the same CIFALPS transect. Our FWI model clearly delineates the Moho topography beneath the external and internal part of the Alps that is in good accordance with the results of Zhao et al. (2015). Our V_p and ρ models agree in following the European Moho discontinuity down to approximately 70 km depth while our (better resolved) V_s model tends to define a Moho boundary down to 80-90 km depth beneath the Ivrea body. Our models therefore support the assumption of subduction of the European continental lithosphere.

The subduction of continental lithosphere raises the question of the fate of continental and oceanic slabs. Early regional traveltimes tomography studies of the Alpine region (Spakman et al., 1993; Bijwaard & Spakman, 2000; Piromallo & Morelli, 2003) have indicated the presence of continuous fast velocity slabs underlying the Alpine Belt down to 150 km depth. Underneath, the mantle transition zone also appears to be dominated by high-velocity materials. The fast anomalies have been interpreted as cold material that are remnants of subducted slabs. Recently, Lippitsch et al. (2003) built an image of the Western Alps upper-mantle with an improved resolution by crustal-corrected traveltimes inversion. Their investigation revealed striking complexities in the velocity structure of the uppermost mantle. They observed a detached fast anomaly from 150 to 300 km depth beneath the Western part of the Alps that is bordered in the West by a slow velocity anomaly. The upper part of their model – from 80 to 120 km depth – displays a two-block (separated at the location of the Ivrea body) fast velocity anomaly that may indicate the extent of the European and Adriatic LAB. These observations agree with our V_p and V_s FWI models that also identify a variation of LAB from 80 to 120 km depth. We particularly recover the anomalously slow velocities in the Western part of the Alps lithosphere that indicates a zone of anomalously high temperature therefore positively buoyant mantle that has previously been interpreted as asthenospheric up-welling (Lippitsch et al., 2003). Such asthenospheric flow has been interpreted by von Blanckenburg & Davies (1995) as a consequence of the oceanic slab detachment that may explain extensional earthquakes in the Western Alps as suggested by Sue et al. (1999). Moreover, Nocquet et al. (2016) recently proposed that present-day uplift rates in the Alps (that are several times larger than horizontal velocities) may result from upward traction at the base of the Alpine lithosphere. They propose that this uplift may thus be due to the replacement of dense mantle material by more buoyant one, therefore sustaining the idea of asthenospheric up-welling.

The block separation of the lithosphere beneath the Ivrea body is not visible in our V_p model that displays a continuous positive anomaly down to 150 km depth, nonetheless our V_s model indicates a blob of slow velocity at the same location. We remind that the V_s velocity model is better resolved than the V_p model, whose resolution might be hampered by strong vertical smearing at these depths. The wide positive anomaly in the V_p model might result both from the downward smearing of the fast Ivrea body and from the upward smearing of the detached slab that would be located just beneath the depth limit of our FWI model according to the tomographic model of Lippitsch et al. (2003). The negative V_s anomaly can either result from thermal effects or the presence of fluid. The fact that the negative V_s anomaly does not correlate with a V_p negative anomaly strongly supports the presence of fluid at the base of the continental crust at the expense of the thermal effect. These fluids would have been released from the subducted hydrated oceanic lithosphere located beneath and possibly indicate the location of the slab detachment at 100 km depth beneath the Ivrea body.

In summary, our FWI image of the Alpine lithosphere suggests and confirms the hypothesis of the slab detachment (or slab break-off) in the Western part of the Alpine Belt as proposed by von Blanckenburg & Davies (1995). The lithospheric structure of the whole Alpine Belt as imaged by Lippitsch et al. (2003) further indicates that the detachment is restricted to the Western part since in the Central Alps the European slab can continuously be traced within the lower-lithosphere suggesting a lateral migration of the slab detachment : the slab tear hypothesis (Wortel & Spakman, 2000). Lippitsch et al. (2003) proposed that the slab break-off at the Ivrea location might be caused by the Ivrea body that could have acted as a buttress during the collision of European and Adriatic lithosphere, forcing the European slab to sink beneath the Adriatic lithosphere.

6.5 Conclusions

We presented the results of a multiparameter 3D elastic full-waveform inversion of the CIFALPS teleseismic dataset recorded in the Western Alps. We showed that FWI, by using the full information content of the seismograms, is able to generate multiparameter images of the lithosphere with unprecedented resolution. Particularly, we emphasised the ability of FWI to recover reliable density information on the upper lithosphere. Recovered images of the Western Alps lithosphere agree well with previous receiver function, LET and teleseismic tomography studies as well as regional gravity measurements. Our models nicely display sharp crustal structures of the Western-Alps, notably, the Ivrea body and the Moho boundary and support the idea of a continental subduction event during the Alpine orogeny. Additionally, the revealed structures of the lower lithosphere support the hypothesis of a slab detachment or slab tear in the Western Alps. We show that the application of FWI at lithospheric scale fills the resolution gap between highly-resolved crustal structures provided by receiver function analysis and classical tomographic images of the upper-mantle. Further applications may be considered for more complex elastic properties of the medium using the transverse component or even bigger time window to include first S-wave arrival and its multiples. Added constrains on density reconstructions may be provided if one consider a joint inversion with gravity data. The new datasets obtained

from current AlpArray or future CIFALPS-2 experiments will provide additional lithospheric information on the Alpine Belt and may confirm current interpretations.

Acknowledgments

We warmly thank L. Combe for invaluable help in the code development, especially for designing the domain decomposition library that allowed high-performance parallelism, and J. Virieux who encouraged the use of the CIFALPS dataset and motivated fruitful discussions. This study also benefits from discussions with E. Bozdağ, Y. Rolland, J.-X. Dessa and A. Sladen. The CIFALPS project is funded by the State Key Laboratory of Lithospheric Evolution, China, the National Natural Science Foundation of China (Grant 41625016) and by a grant from Labex OSUG@2020 (Investissements d'avenir – ANR10 LABX56, France). This study was partially funded by the SEISCOPE consortium (<http://seiscope2.osug.fr>), sponsored by BP, CGG, CHEVRON, EXXON-MOBIL, JGI, PETROBRAS, SAUDI ARAMCO, SCHLUMBERGER, SHELL, SINOPEC, STATOIL, TOTAL and WOODSIDE. This study was granted access to the HPC resources of SIGAMM infrastructure (<http://crimson.oca.eu>), hosted by Observatoire de la Côte d'Azur and which is supported by the Provence-Alpes Côte d'Azur region, and the HPC resources of CINES/IDRIS under the allocation 046091 made by GENCI.

Conclusions et perspectives

*« Il est souvent plus difficile de finir que de commencer... »
Mon nom est personne, (1973)*

L'objectif de cette thèse était de développer une méthode d'inversion de formes d'ondes adaptée à la configuration télesismique pour l'imagerie 3D haute résolution des structures élastiques de la lithosphère, et particulièrement celles des Alpes Occidentales.

La mise en œuvre de cette méthodologie a nécessité de lourds développements et validations numériques, notamment pour la modélisation des champs d'ondes télesismiques nécessaires au processus d'inversion dont le coût de calcul, intimement lié aux simulations 3D en temps, impose l'adaptation des programmes développées à des architectures massivement parallèles de calcul intensif. Parallèlement à ces développements, nous avons évalué la faisabilité, le potentiel et les limites de la méthode à l'aide de tests synthétiques et d'une étude paramétrique menée sur un modèle synthétique réaliste représentatif de l'état des connaissances actuelles de la lithosphère des Alpes Occidentales. Nous avons enfin appliqué notre méthode aux données télesismiques enregistrées par l'expérience CIFALPS, menée dans les Alpes Occidentales, et obtenu une image tomographique multiparamètre de la structure lithosphérique des Alpes. Les modèles de densité et de vitesses obtenus ont permis d'imager des structures connues de la lithosphère Alpine, notamment le corps d'Ivrée et les bassins sédimentaires de la vallée de Rhône et de la plaine de Pô. De surcroît, ces images lèvent toute équivoque quant à l'existence d'une subduction continentale lors de l'orogénèse des Alpes. En effet, ces images montrent une discontinuité croûte-manteau très nette dont le pendage augmente à mesure que l'on s'approche de la partie interne des Alpes. Plus en profondeur, les modèles présentent des anomalies de vitesse en accord avec l'hypothèse d'un détachement du panneau plongeant européen.

La présentation de ces travaux au sein de ce manuscrit était divisée en trois parties dont nous rappelons les grandes lignes dans la section suivante.

Conclusions générales

Modélisation de la propagation des ondes télesismiques

La première partie de cette thèse portait sur la modélisation tridimensionnelle des champs d'ondes télesismiques et se divisait en deux chapitres.

Dans le premier chapitre, nous avons présenté le cadre général des théories permettant de comprendre, décrire et simuler la propagation des champs d'ondes élastiques. Ces bases théoriques sont importantes car elles décrivent la propagation des ondes dans des milieux arbitrairement hétérogènes et mènent aux équations de l'élastodynamique que nous cherchons à résoudre numériquement. Par ailleurs, ces bases permettent aussi de décrypter le contenu des signaux enregistrés par les stations sismologiques : les champs d'ondes propagées dans les milieux élastiques, la part de complexité générée par la structure globale de la Terre ainsi que les manifestations propres aux structures lithosphériques. D'autre part, certains aspects développés dans ce chapitre ont permis de tisser d'importants liens entre la physique de la propagation des ondes et le principe d'imagerie inhérent au problème inverse et d'analyser les résultats de cette imagerie en termes de résolution spatiale.

Le second chapitre était dédié à la simulation numérique de la propagation des ondes élastiques. Nous avons présenté succinctement les méthodes de résolution d'équations aux dérivées partielles communément utilisées en sismologie et nous nous sommes focalisés sur la description de la méthode des éléments finis spectraux que nous avons utilisée pour simuler la propagation des ondes dans le milieu lithosphérique. Enfin, nous avons détaillé la méthode d'injection de champ d'onde basée sur le couplage entre un code aux éléments spectraux 2D, permettant le calcul efficace de champs d'ondes 3D dans une terre à symétrie sphérique (*AxiSEM*), et notre code aux éléments spectraux local qui permet de simuler la propagation du champ d'onde télé-sismique dans la cible lithosphérique.

L'inversion de formes d'ondes télésismiques

La seconde partie de ce manuscrit traitait exclusivement du problème inverse correspondant à l'inversion de formes d'ondes télésismiques.

Dans le quatrième chapitre nous avons rappelé la définition et les principales difficultés inhérentes à la résolution de problèmes inverses. Nous avons présenté les différentes approches possibles pour résoudre ces problèmes et détaillé précisément les méthodes d'optimisation locales. Par la suite, nous avons formulé l'inversion de formes d'ondes comme un problème inverse linéarisé résolu de manière itérative par un processus d'optimisation locale. Nous avons détaillé l'expression du gradient de la fonction coût et fourni une interprétation physique de la condition d'imagerie qui lui est associée, nous avons ensuite explicité comment effectuer efficacement son calcul numérique par la méthode de l'état adjoint. Nous avons finalement présenté différents aspects plus techniques liés à l'implémentation de l'algorithme de la FWI.

Le cinquième chapitre présentait, sous forme d'une publication scientifique, une étude paramétrique de la FWI en configuration télésismique. Cette étude a été effectuée à l'aide d'un modèle lithosphérique réaliste représentatif des structures des Alpes Occidentales. Nous avons analysé les effets conjoints de la paramétrisation, du choix du modèle initial et de l'acquisition sur les résultats de l'inversion de formes d'ondes télésismiques. Nous avons conclu que les données télésismique ont un contenu basse fréquence suffisant pour initier l'inversion de formes d'ondes à partir de modèles de référence de Terre sans risque de cycle-skipping ou de convergences vers un minimum local inacceptable. En outre, nous avons comparé les résultats de l'inversion lorsque celle-ci est effectuée pour différentes paramétrisations du milieu

lithosphérique et nous avons confronté les conclusions de ces expériences numériques avec une analyse de résolution théorique fondée sur l'analyse des diagrammes de rayonnement des différentes paramétrisations testées. Nous avons conclu que la paramétrisation faisant intervenir les paramètres élastiques et la densité était la plus favorable à une reconstruction large bande (en termes de contenu spectral de nombres d'ondes) des trois classes paramètres. Néanmoins, les modèles de meilleure qualité, c'est-à-dire fournissant le meilleur compromis entre la résolution atteinte et le rapport signal sur bruit, ont été obtenus par la recombinaison *a posteriori* des trois classes de paramètres sus-mentionnées en impédances P et S. Cette étude a ainsi permis d'illustrer la nécessité de distinguer les paramètres d'optimisation (ceux mis à jour lors de l'inversion) et les paramètres utilisés pour la visualisation ou l'interprétation générés par recombinaison non linéaire des paramètres d'optimisation. Enfin, étant donné les contraintes logistiques importantes imposées par les acquisitions sismiques, nous avons évalué deux grandes familles de dispositifs de capteurs pour l'imagerie télésismique : des acquisitions fondées sur des profils denses de stations versus des déploiements plus clairsemés en nappe. Considérant le nombre limité d'événements télésismiques généralement utilisé et l'éclairage angulaire (ou azimutal) incomplet qui peut en résulter, nous avons conclu que les réseaux bi-dimensionnels de stations sont à privilégier par rapports aux réseaux linéiques ultra-denses pour bénéficier d'une résolution verticale et horizontale acceptables. Au vu des fréquences contenues dans les données télésismiques et donc de la résolution escomptée, un espacement inter-station d'environ 25 km suffit pour caractériser finement les structures lithosphériques majeures.

Application aux Alpes Occidentales

La dernière partie de ce manuscrit portait sur l'application de notre algorithme d'inversion de formes d'ondes aux données télésismiques enregistrées par la campagne CIFALPS pour analyser la structure lithosphérique de la chaîne alpine dans les Alpes Occidentales.

Nous avons, au cours du cinquième chapitre, présenté les grandes lignes de l'histoire géodynamique des Alpes ainsi que ses grandes unités structurales qui résultent d'épisodes successifs de subduction océanique, subduction continentale et collision continentale. Une synthèse des connaissances géophysiques actuelles a permis de passer en revue quelques questionnements importants sur la structure et l'évolution de l'arc alpin. L'écaïlle briançonnaise est-elle d'origine crustale ou mantellique ? Jusqu'à quelle profondeur peut-on suivre les traces du Moho européen et de la subduction continentale ? Quelle est la géométrie de la partie occidentale du corps d'Ivrée ? Le panneau plongeant européen est-il continu ou détaché ?

Le dernier chapitre était une publication scientifique que nous avons soumis au « *Geophysical Journal International* » et qui présentait les résultats de l'application de l'inversion de formes d'ondes dans les Alpes Occidentales. Nos résultats ont permis de tirer deux conclusions majeures. Premièrement, ils ont montré que, malgré un nombre relativement limité de séismes auxquels elle est appliquée, la FWI peut fournir des modèles de vitesses et de densité de résolution suffisante pour permettre des interprétations intégrées des structures crustales et mantelliées. La fiabilité du modèle de densité (qui est un paramètre de second ordre) a été vérifiée en comparant sa signature gravimétrique à l'anomalie de Bouguer mesurée dans les Alpes occidentales. En outre, les tests de résolution synthétiques effectués ont montré que l'inversion

multi-paramètre n'était que faiblement impactée par des couplages inter-paramètres, confirmant ainsi les conclusions de l'analyse paramétrique. Deuxièmement, nos modèles FWI ont permis d'imager les principaux facteurs structuraux jouant un rôle majeur sur la géodynamique de l'orogène alpine. Les modèles imagés révèlent l'enfouissement de la lithosphère européenne à grandes profondeurs (entre 75 et 80 km) sous les zones internes de l'arc Alpin témoignant ainsi de la subduction continentale. Ils précisent aussi la géométrie verticale de la bordure occidentale du corps d'Ivrée lequel est clairement connecté au manteau adriatique. Nos modèles attestent également de l'origine crustale de l'écaïlle Briançonnaise. À l'échelle mantellique, ils confirment l'hypothèse d'une remontée asthénosphérique sous les Alpes externes, et indiquent un possible détachement du panneau plongeant européen à 100 km de profondeur sous le corps d'Ivrée. Ces résultats constituent des arguments en faveur de l'hypothèse d'un déchirement du panneau plongeant européen faisant suite à la collision continentale.

Perspectives

Amélioration de la méthode

Plusieurs axes de recherche à caractère méthodologique sont envisageables pour améliorer la qualité des images lithosphériques et mieux contraindre les modèles lithosphériques obtenus par inversion de formes d'ondes.

Tout d'abord, la méthode pourrait être étendue à des physiques de propagation d'ondes plus complexes prenant en compte des anisotropies particulières ou des effets d'atténuation. Ces informations apporteraient des contraintes supplémentaires pour l'interprétation géodynamique des modèles lithosphériques. Même si ces développements ne présentent pas de difficulté particulière en ce qui concerne l'implémentation du problème direct (la modélisation sismique), la prise en compte de ces phénomènes de propagation dans le processus d'inversion nécessite d'être étudiée minutieusement pour analyser la sensibilité respective des données et de la FWI à ces nouvelles classes de paramètres. De la même manière, le développement d'un problème inverse multi-physiques inversant les formes d'ondes télésismiques de manière conjointe aux données gravimétriques (anomalie de Bouguer, ondulation du géoïde) permettrait de mieux contraindre les modèles de densité estimés par la FWI.

D'autre part, la résolution et la qualité des modèles FWI pourraient être améliorées par la prise en compte de phases plus tardives comme les ondes S ou SKS. Celles-ci permettraient d'augmenter l'éclairage angulaire de la cible lithosphérique et d'enrichir les modes de diffractions utilisés par le processus d'imagerie. Les autres perspectives pour améliorer la résolution des modèles lithosphériques impliquent le déploiement de nappes de stations de plus en plus denses, malgré les contraintes logistiques auxquelles sont soumises ces acquisitions, tout en considérant un nombre plus important de sources pour renforcer l'éclairage angulaire du milieu lithosphérique. Enfin, l'apport d'inversions à plus hautes fréquences reste à évaluer en considérant le rapport signal sur bruit des données télésismiques, l'hétérogénéité des sources télésismiques et le surcoût calculatoire que ces inversions vont engendrer.

Précisons qu'il pourrait être envisageable de réduire les coûts numériques, qui croissent linéairement avec le nombre de sources considérées, en usant des méthodes d'encodage de source développées en exploration sismique (Castellanos-Lopez et al., 2013). Avec ces méthodes, une seule modélisation est effectuée pour un terme source encodé et ce quel que soit le nombre d'événements considéré. Néanmoins, la réduction du coût de calcul d'une itération se fait au détriment d'une dégradation du rapport signal sur bruit du gradient nécessitant un nombre plus important d'itérations. La mise en œuvre de ces techniques d'assemblage et d'encodage des sources est par ailleurs potentiellement problématique en configuration télé-sismique en raison de l'utilisation de la méthodes d'injection de champs d'ondes.

Les Alpes Occidentales

Notre application de l'inversion de formes d'ondes se limitait à l'imagerie de la croûte et de la partie supérieure du manteau (jusqu'à 200 km de profondeur) dans la partie occidentale de la chaîne alpine. Pour accroître notre compréhension de la chaîne il serait nécessaire de l'imager dans sa globalité, c'est-à-dire, de l'extrémité Sud des Alpes occidentales à l'extrémité Est des Alpes orientales et en profondeur jusqu'à la zone de transition du manteau. Cela permettrait de préciser la géométrie des différents panneaux plongeants présents dans les Alpes occidentales pour confirmer le déchirement du panneau plongeant européen et vérifier la polarité de panneau plongeant dinarique. En outre, cela permettrait d'analyser les interactions entre les panneaux plongeants apennique et européen.

Cette étude pourra être menée prochainement grâce au déploiement du réseau AlpArray (Fig. 6.13) qui couvre la totalité de la chaîne alpine et de la ligne ultra dense CIFALPS-2 qui sera prochainement déployée entre Genève et Gênes.

Au delà de la FWI en configuration télé-sismique

La méthode d'injection de champs d'ondes pourrait être employée et/ou modifiée pour de nombreuses applications en sismologie.

Dans sa forme actuelle, la modélisation hybride, pourrait permettre d'analyser les effets respectifs des structures 3D, de la topographie, de l'anisotropie et de l'atténuation dans les données télé-sismiques. Il est ainsi possible, au moyen de modèles synthétiques réalistes, de vérifier la pertinence des méthodes classiques d'estimation de l'anisotropie ou de l'atténuation.

D'autre part, nous pouvons envisager de modifier la méthode d'injection pour analyser les effets des structures 3D du côté de la source, notamment pour les grands séismes. La modélisation locale serait alors effectuée dans un modèle complexe contenant la source, les champs d'ondes seraient extraits sur les bords du domaine et renvoyés vers les récepteurs télé-sismiques au moyen de fonctions de Green calculées par AxiSEM.

Enfin, un couplage complet – injection et extraction – de champs d'ondes pourrait être envisagé pour l'étude de structures enfouies à très grandes profondeur. Ce genre de méthode est particulièrement intéressant pour l'étude de la frontière noyau-manteau (Masson et al., 2014).

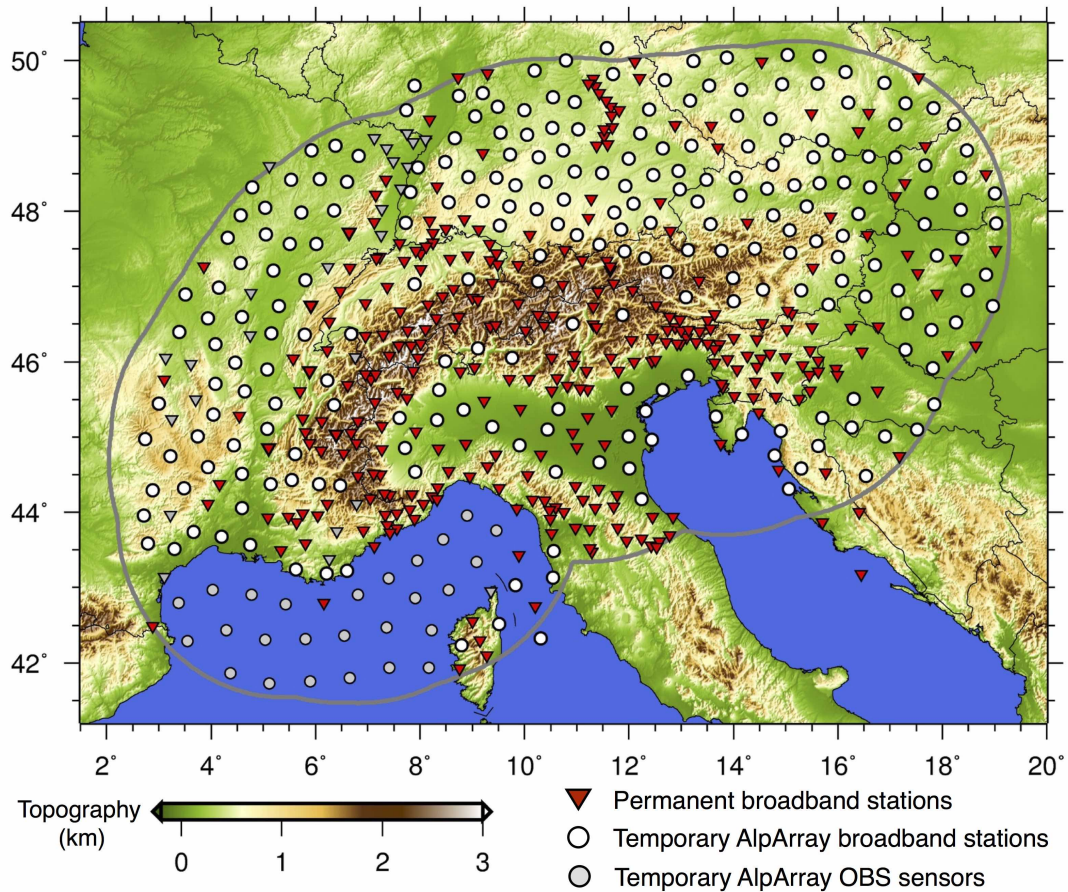


Figure 6.13 – Présentation du réseau temporaire AlpArray déployé dans les Alpes. Ce réseau est le fruit d’une initiative européenne pour améliorer nos connaissances géodynamiques sur l’orogénèse alpine.

Bibliographie

- Aki, K., Christoffersson, A., & Husebye, E. S. (1977). Determination of the three-dimensional seismic structure of the lithosphere. *Journal of Geophysical Research*, 82(2), 277–296.
- Aki, K. & Lee, W. (1976). Determination of three-dimensional velocity anomalies under a seismic array using first p-arrival times from local earthquakes. *Journal of Geophysical Research*, 81(4), 4381–4399.
- Aki, K. & Richards, P. G. (2002). *Quantitative seismology, theory and methods, second edition*. Sausalito, California : University Science Books.
- Alkhalifah, T. & Plessix, R. (2014). A recipe for practical full-waveform inversion in anisotropic media : An analytical parameter resolution study. *Geophysics*, 79(3), R91–R101.
- Alterman, Z. & Karal, F. C. (1968). Propagation of elastic waves in layered media by finite-difference methods. *Bulletin of the Seismological Society of America*, 58, 367–398.
- Ammon, C. J., Randall, G. E., & Zandt, G. (1990). On the nonuniqueness of receiver function inversions. *Journal of Geophysical Research*, 95(B10), 15,303 – 15,318.
- Artemieva, I. (2011). *The lithosphere : An interdisciplinary approach*. Cambridge University Press.
- Babuska, V. & Cara, M. (1991). *Seismic Anisotropy in the Earth*. Boston : Kluwer Academic Publishers.
- Baker, B. & Roecker, S. (2014). A full waveform tomography algorithm for teleseismic body and surface waves in 2.5 dimensions. *Geophysical Journal International*, 198, 1775–1794.
- Balmino, G., Vales, N., Bonvalot, S., & Briais, A. (2012). Spherical harmonic modelling to ultra-high degree of bouguer and isostatic anomalies. *Journal of Geodesy*, 86(7), 499–520.
- Barruol, G., Bonnin, M., Pedersen, H., Bokelmann, G. H., & Tiberi, C. (2011). Belt-parallel mantle flow beneath a halted continental collision : The western alps. *Earth and Planetary Science Letters*, 302(3), 429–438.
- Barruol, G., Deschamps, A., & Coutant, O. (2004). Mapping upper mantle anisotropy beneath se france by sks splitting indicates neogene asthenospheric flow induced by apenninic slab roll-back and deflected by the deep alpine roots. *Tectonophysics*, 394(1), 125–138.
- Bavalia, K., Motaghia, K., Soboutia, F., Ghodsa, A., Abbasib, M., Priestley, K., Mortezaejada, G., & Rezaei, M. (2016). Lithospheric structure beneath nw iran using regional and teleseismic travel-time tomography. *Physics of the Earth and Planetary Interiors*, 253, 97–107.

- Bayer, R., Cazes, M., Dal Piaz, G., Damotte, B., Elter, G., Gosso, G., Hirn, A., Lanza, R., Lombardo, B., Mugnier, J.-L., et al. (1987). Premiers résultats de la traversée des alpes occidentales par sismique réflexion verticale (programme ecors-crop). *Comptes rendus de l'Académie des sciences. Série 2, Mécanique, Physique, Chimie, Sciences de l'univers, Sciences de la Terre*, 305(18), 1461–1470.
- Beller, S., Monteiller, V., Operto, S., & Nolet, G. (2015). Multi-parameter high-resolution lithospheric imaging by source-independent full-waveform inversion of teleseismic data. In *2015 AGU Fall Meeting : Agu*.
- Beller, S., Monteiller, V., Operto, S., Nolet, G., & Paul, A. (2017). Lithospheric architecture of the south-western alps revealed by multi-parameter teleseismic full-waveform inversion. *Geophysical Journal International*, Submitted.
- Bérenger, J.-P. (1994). A perfectly matched layer for absorption of electromagnetic waves. *Journal of Computational Physics*, 114, 185–200.
- Béthoux, N., Sue, C., Paul, A., Virieux, J., Fréchet, J., Thouvenot, F., & Cattaneo, M. (2007). Local tomography and focal mechanisms in the south-western alps : comparison of methods and tectonic implications. *Tectonophysics*, 432(1), 1–19.
- Beydoun, W. B. & Tarantola, A. (1988). First Born and Rytov approximation : Modeling and inversion conditions in a canonical example. *Journal of the Acoustical Society of America*, 83, 1045–1055.
- Beylkin, G. & Burridge, R. (1990). Linearized inverse scattering problems in acoustics and elasticity. *Wave motion*, 12, 15–52.
- 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.
- Bijwaard, H. & Spakman, W. (2000). Non-linear global p-wave tomography by iterated linearized inversion. *Geophysical Journal International*, 141, 71–82.
- Bleibinhaus, F. & Rondenay, S. (2009). Effects of surface scattering in full-waveform inversion. *Geophysics*, 74(6), WCC69–WCC77.
- Bodin, T., Leiva, J., Romanowicz, B., Maupin, V., & Yuan, H. (2016). Imaging anisotropic layering with bayesian inversion of multiple data types. *Geophysical Journal International*, 206(1), 605–629.
- Bodin, T., Sambridge, M., Tkalčić, H., Arroucau, P., Gallagher, K., & Rawlinson, N. (2012). Transdimensional inversion of receiver functions and surface wave dispersion. *Journal of Geophysical Research : Solid Earth*, 117(B2).
- Bodin, T., Yuan, H., & Romanowicz, B. (2013). Inversion of receiver functions without deconvolution—application to the indian craton. *Geophysical Journal International*, (pp. ggt431).
- Bohlen, T. & Saenger, E. H. (2006). Accuracy of heterogeneous staggered-grid finite-difference modeling of Rayleigh waves. *Geophysics*, 71, 109–115.
- Borisov, D. (2014). *3D elastic full waveform inversion of time lapse seismic data using grid injection method*. PhD thesis, Institut de Physique du Globe de Paris.
- Bostock, M. & Rondenay, S. (1999). Migration of scattered teleseismic body waves. *Geophysical Journal International*, 137(3), 732–746.

- Bostock, M., 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. & Aki, K. (1977). Discrete wave-number representation of seismic-source wave fields. *Bulletin of the Seismological Society of America*, 67(2), 259–277.
- Bouchon, M., Schultz, C. A., & Toksöz, M. N. (1996). Effect of 3D topography on seismic motion. *Journal Geophysical Research*, 101, 5835–5846.
- Bozdağ, E., Peter, D., Lefebvre, M., Komatitsch, D., Tromp, J., Hill, J., Podhorszki, N., & Pugmire, D. (2016). Global adjoint tomography : first-generation model. *Geophysical Journal International*, 207(3), 1739–1766.
- Bozdağ, E., Trampert, J., & Tromp, J. (2011). Misfit functions for full waveform inversion based on instantaneous phase and envelope measurements. *Geophysical Journal International*, 185(2), 845–870.
- Brossier, R., Operto, S., & Virieux, J. (2009a). Robust frequency-domain full-waveform inversion using the l_1 norm. *Geophysical Research Letters*, 36, L20310.
- Brossier, R., Operto, S., & Virieux, J. (2009b). Seismic imaging of complex onshore structures by 2D elastic frequency-domain full-waveform inversion. *Geophysics*, 74(6), WCC105–WCC118.
- Bunks, C., Salek, F. M., Zaleski, S., & Chavent, G. (1995). Multiscale seismic waveform inversion. *Geophysics*, 60(5), 1457–1473.
- Burdick, S., de Hoop, M. V., Wang, S., & van der Hilst, R. D. (2014). Reverse-time migration-based reflection tomography using teleseismic free surface multiples. *Geophysical Journal International*, 16(2), 996–1017.
- Calvet, M., Chevrot, S., & Souriau, A. (2006). P-wave propagation in transversely isotropic media i. finite-frequency theory. *Physics of the Earth and Planetary Interiors*, 56, 12–20.
- Canuto, C., Hussaini, M. Y., Quarteroni, A., & Zang, T. A. (2006). *Spectral Methods Fundamentals in Single Domains*. Springer.
- Capdeville, Y., Chaljub, E., Vilotte, J., & Montagner, J. (2003). Coupling the spectral element method with a modal solution for elastic wave propagation in global earth models. *Geophysical Journal International*, 152, 34–67.
- Cassidy, J. (1992). Numerical experiments in broadband receiver function analysis. *Bulletin of the Seismological Society of America*, 82(3), 1453–1474.
- Castellanos, C. L. (2014). *Speed-up and regularization techniques for seismic full waveform inversion*. PhD thesis, Université de Nice-Sophia Antipolis.
- Castellanos-Lopez, C., Métivier, L., Operto, S., & Brossier, R. (2013). Fast full waveform inversion with source encoding and second order optimization methods. In *75th EAGE Conference & Exhibition incorporating SPE EUROPEC 2013* (pp. We 11 10).
- Červený, V. (2001). *Seismic Ray Theory*. Cambridge : Cambridge University Press.
- Červený, V., Molotkov, I. A., & Pšenčík, I. (1977). *Ray Method in Seismology*. Praha : Charles University Press.

- Chauris, H. (2000). *Analyse de vitesse par migration pour l'imagerie des structures complexes en sismique réflexion*. PhD thesis, Ecole des Mines de Paris.
- Chavent, G. (1974). Identification of parameter distributed systems. In R. Goodson & M. Polis (Eds.), *Identification of function parameters in partial differential equations* (pp. 31–48). American Society of Mechanical Engineers, New York.
- Chen, L., Wen, L., & Zheng, T. (2005). A wave equation migration method for receiver function imaging : 1. theory. *Journal of Geophysical Research : Solid Earth*, 110(B11).
- Chen, P., Jordan, T., & Zhao, L. (2007). Full three-dimensional tomography : a comparison between the scattering-integral and adjoint-wavefield methods. *Geophysical Journal International*, 170, 175–181.
- Cheng, C., Bodin, T., & Allen, R. M. (2016). Three-dimensional pre-stack depth migration of receiver functions with the fast-marching method : a Kirchhoff approach. *Geophysical Journal International*, 205, 819–829.
- Chevrot, S., Sylvander, M., Diaz, J., Ruiz, M., Paul, A., Group, P. W., et al. (2015). The pyrenean architecture as revealed by teleseismic p-to-s converted waves recorded along two dense transects. *Geophysical Journal International*, 200(2), 1096–1107.
- Choi, Y. & Alkhalifah, T. (2011). Source-independent time-domain waveform inversion using convolved wavefields : application to the encoded multisource waveform inversion. *Geophysics*, 76(5), R125–R134.
- Chopin, C. (1984). Coesite and pure pyrope in high-grade blueschists of the western alps : a first record and some consequences. *Contributions to Mineralogy and Petrology*, 86(2), 107–118.
- Chopin, C. (1986). Phase relationships of ellenbergerite, a new high-pressure mg-al-ti-silicate in pyrope-coesite-quartzite from the western alps. *Geological Society of America Memoirs*, 164, 31–41.
- Claerbout, J. (1985). *Imaging the Earth's interior*. Blackwell Scientific Publication.
- Clayton, R. & Engquist, B. (1977). Absorbing boundary conditions for acoustic and elastic wave equations. *Bulletin of the Seismological Society of America*, 67, 1529–1540.
- Clayton, R. W. & Engquist, B. (1980). Absorbing boundary conditions for wave equation migration. *Geophysics*, 45(5), 895–904.
- Coutant, O. (1989). *Program of Numerical Simulation AXITRA*. Technical report, Laboratoire de Géophysique Interne et Tectonophysique, Université J. Fourier, Grenoble I.
- Crase, E., Pica, A., Noble, M., McDonald, J., & Tarantola, A. (1990). Robust elastic non-linear waveform inversion : application to real data. *Geophysics*, 55, 527–538.
- Crotwell, H. P., Owens, J. T., & Ritsema, J. (1999). The TauP Toolkit : Flexible seismic travel-time and ray-path utilities. *Seismological Research Letters*, 70, 154–160.
- Dahlen, F. A. & Baig, A. M. (2002). Fréchet kernels for body waves amplitudes. *Geophysical Journal International*, 150, 440–466.
- Dahlen, F. A., Hung, S. H., & Nolet, G. (2000a). Fréchet kernels for finite-difference traveltimes - I. theory. *Geophysical Journal International*, 141, 157–174.
- Dahlen, F. A., Hung, S.-H., & Nolet, G. (2000b). Fréchet kernels for finite-frequency traveltimes - I. theory. *Geophysical Journal International*, 141, 157–174.

- Davies, J. H. & von Blanckenburg, F. (1995). Slab breakoff : a model of lithosphere detachment and its test in the magmatism and deformation of collisional orogens. *Earth and Planetary Science Letters*, 129(1), 85–102.
- Davis, P. J. (1975). *Interpolation and Approximation*. New York : Dover Publications.
- De Basabe, J. D. & Sen, M. K. (2007). Grid dispersion and stability criteria of some common finite-element methods for acoustic and elastic wave equations. *Geophysics*, 72(6), T81–T95.
- De Graciansky, P. C., Roberts, D. G., & Tricart, P. (2011). *The Western Alps, from rift to passive margin to orogenic belt : an integrated geoscience overview*, volume 14. Elsevier.
- Dessa, J. X., Operto, S., Kodaira, S., Nakanishi, A., Pascal, G., Virieux, J., & Kaneda, Y. (2004). Multiscale seismic imaging of the eastern Nankai trough by full waveform inversion. *Geophysical Research Letters*, 31(L18606), doi :10.1029/2004GL020453.
- Devaney, A. (1984). Geophysical diffraction tomography. *Geoscience and Remote Sensing, IEEE Transactions on*, GE-22(1), 3–13.
- Dèzes, P., Schmid, S., & Ziegler, P. (2004). Evolution of the European Cenozoic rift system : interaction of the alpine and Pyrenean orogens with their foreland lithosphere. *Tectonophysics*, 389(1), 1–33.
- Diehl, T., Husen, S., Kissling, E., & Deichmann, N. (2009). High-resolution 3-d p-wave model of the alpine crust. *Geophysical Journal International*, 179(2), 1133–1147.
- Duvaut, G. (1990). *Mécanique des milieux continus*. Masson.
- Dziewonski, A., Chou, T.-A., & Woodhouse, J. (1981). Determination of earthquake source parameters from waveform data for studies of global and regional seismicity. *Journal of Geophysical Research : Solid Earth*, 86(B4), 2825–2852.
- Dziewonski, A., Hales, A., & Lapwood, E. (1975). Parametrically simple earth models consistent with geophysical data. *Physics of the Earth and Planetary Interiors*, 10(1), 12–48.
- Dziewonski, A. M. & Anderson, D. L. (1981). Preliminary reference Earth model. *Physics of the Earth and Planetary Interiors*, 25, 297–356.
- Dziewonski, A. M., Hager, B. H., & O’Connell, R. J. (1977). Large-scale heterogeneities in the lower mantle. *Journal of Geophysical Research*, 82(2), 239–255.
- Eaton, D. W., Darbyshire, F., Evans, R. L., Grütter, H., Jones, A. G., & Yuan, X. (2009). The elusive lithosphere–asthenosphere boundary (lab) beneath cratons. *Lithos*, 109(1), 1–22.
- Ekström, G., Nettles, M., & Dziewoński, A. (2012). The global cmt project 2004–2010 : centroid-moment tensors for 13,017 earthquakes. *Physics of the Earth and Planetary Interiors*, 200, 1–9.
- Etienne, V., Chaljub, E., Virieux, J., & Glinsky, N. (2010). An hp-adaptive discontinuous Galerkin finite-element method for 3D elastic wave modelling. *Geophysical Journal International*, 183(2), 941–962.
- Farra, V. & Vinnik, L. (2000). Upper mantle stratification by p and s receiver functions. *Geophysical Journal International*, 141(3), 699–712.
- Fichtner, A., Bunge, H. P., & Igel, H. (2006a). The adjoint method in seismology : I - Theory. *Physics of the Earth and Planetary Interiors*, 157(1-2), 86–104.

- Fichtner, A., Bunge, H. P., & Igel, H. (2006b). The adjoint method in seismology : II - Applications. *Physics of the Earth and Planetary Interiors*, 157(1-2), 105–123.
- Fichtner, A., Kennett, B. L. N., Igel, H., & Bunge, H. P. (2009). Full seismic waveform tomography for upper-mantle structure in the Australasian region using adjoint methods. *Geophysical Journal International*, 179(3), 1703–1725.
- Fichtner, A., Kennett, B. L. N., Igel, H., & Bunge, H. P. (2010). Full waveform tomography for radially anisotropic structure : New insights into present and past states of the Australasian upper mantle. *Earth and Planetary Science Letters*, 290(3-4), 270–280.
- Fichtner, A. & Trampert, J. (2011a). Hessian kernels of seismic data functionals based upon adjoint techniques. *Geophysical Journal International*, 185(2), 775–798.
- Fichtner, A. & Trampert, J. (2011b). Resolution analysis in full waveform inversion. *Geophysical Journal International*, 187, 1604–1624.
- Fichtner, A., Trampert, J., Cupillard, P., Saygin, E., Taymaz, T., Capdeville, Y., & nor, A. V. (2013). Multiscale full waveform inversion. *Geophysical Journal International*, 194, 534–556.
- Fink, M. (1993). Time-reversal mirrors. *Journal of Physics D : Applied Physics*, 26(9), 1333–1350.
- Fischer, K. M., Ford, H. A., Abt, D. L., & Rychert, C. A. (2010). The lithosphere-asthenosphere boundary. *Annual Review of Earth and Planetary Sciences*, 38, 551–575.
- Forgues, E. (1996). *Inversion linearisée multi-paramètres via la théorie des rais*. PhD thesis, 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.
- Fox, M., Herman, F., Kissling, E., & Willett, S. D. (2015). Rapid exhumation in the western alps driven by slab detachment and glacial erosion. *Geology*, 43(5), 379–382.
- Frederiksen, A. & Bostock, M. (2000). Modelling teleseismic waves in dipping anisotropic structures. *Geophysical Journal International*, 141(2), 401–412.
- Gao, L., Brossier, R., Pajot, B., Tago, J., & Virieux, J. (2015). An immersed free surface boundary treatment for seismic wave simulation. *Geophysics*, 80(5), T193–T209.
- Gao, W., Grand, S. P., Baldrige, W. S., Wilson, D., West, M., Ni, J. F., & Aster, R. (2004). Upper mantle convection beneath the central rio grande rift imaged by p and s wave tomography. *Journal of Geophysical Research : Solid Earth*, 109(B3).
- Gauthier, O., Virieux, J., & Tarantola, A. (1986a). Two-dimensional nonlinear inversion of seismic waveform : numerical results. *Geophysics*, 51, 1387–1403.
- Gauthier, O., Virieux, J., & Tarantola, A. (1986b). Two-dimensional nonlinear inversion of seismic waveforms : numerical results. *Geophysics*, 51(7), 1387–1403.
- Geller, R. & Ohminato, T. (1994). Computation of synthetic seismograms and their partial derivatives for heterogeneous media with arbitrary natural boundary conditions using the direct solution method. *Geophysical Journal International*, 116, 421–446.
- Gholami, Y., Brossier, R., Operto, S., Ribodetti, A., & Virieux, J. (2013). Which parametrization is suitable for acoustic VTI full waveform inversion ? - Part 1 : sensitivity and trade-off analysis. *Geophysics*, 78(2), R81–R105.

- Gouveia, W. P. & Scales, J. A. (1997). Resolution of seismic waveform inversion : Bayes versus occam. *Inverse Problems*, 13, 323–349.
- Gouveia, W. P. & Scales, J. A. (1998). Bayesian seismic waveform inversion : parameter estimation and uncertainty analysis. *Journal of Geophysical Research*, 103(B2), 2759–2779.
- Graves, R. (1996). Simulating seismic wave propagation in 3D elastic media using staggered-grid finite differences. *Bulletin of the Seismological Society of America*, 86, 1091–1106.
- Groos, L., Schäfer, M., Forbriger, T., & Bohlen, T. (2014). The role of attenuation in 2D full-waveform inversion of shallow-seismic body and Rayleigh waves. *Geophysics*, 79(6), R247–R261.
- Guitton, A., Ayeni, G., & Diaz, E. (2012). Constrained full-waveform inversion by model reparameterization. *Geophysics*, 77(2), R117–R127.
- Gurrola, H., Minster, J., & Owens, T. (1994). The use of velocity spectrum for stacking receiver functions and imaging upper mantle discontinuities. *Geophysical Journal International*, 117(2), 427–440.
- Gutenberg, B. (1914). Über erdbebenwellen viia. beobachtungen an registrierungen von fernbeben in göttingen und folgerungen über die konstitution des erdkörpers. *Nachrichten von der Königlich Gesellschaft der Wissenschaften zu Göttinge, Mathematisch-Physikalische Klasse*, (pp. 125–176).
- Hadamard, J. (1902). Sur les problèmes aux dérivés partielles et leur signification physique. *Princeton University Bulletin*, (pp. 49–52).
- Handy, M. R., Schmid, S. M., Bousquet, R., Kissling, E., & Bernoulli, D. (2010). Reconciling plate-tectonic reconstructions of alpine tethys with the geological–geophysical record of spreading and subduction in the alps. *Earth-Science Reviews*, 102(3), 121–158.
- Hansen, P. C. (1992). Analysis of discrete ill-posed problems by means of the l-curve. *SIAM Review*, 34(4), 561–580.
- Hello, Y., Ogé, A., Sukhovich, A., & Nolet, G. (2011). Modern mermaids : New floats image the deep earth. *Eos, Transactions American Geophysical Union*, 92(40), 337–338.
- Hesthaven, J. S. & Warburton, T. (2007). *Nodal discontinuous Galerkin methods : algorithms, analysis, and applications*, volume 54. Springer.
- Hetényi, G. (2007). *Evolution of deformation of the Himalayan prism : from imaging to modeling*. PhD thesis, Université Paris Sud-Paris XI.
- Hicks, G. J. (2002). Arbitrary source and receiver positioning in finite-difference schemes using Kaiser windowed sinc functions. *Geophysics*, 67, 156–166.
- 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.
- Holland, J. H. (1975). Adaptation in natural and artificial system. In *Adaptation in Natural and Artificial System*. University Michigan Press.
- Hudson, J. A. & Heritage, J. R. (1981). *Use of the Born approximation in seismic scattering problems*. Cambridge University Press.
- Hughes, T. J. R. (2003). *The finite element method : linear static and dynamic finite element analysis*. Dover Publications Inc.

- Hung, S.-H., Dahlen, F. A., & Nolet, G. (2000). Fréchet kernels for finite-frequency traveltimes - II. examples. *Geophysical Journal International*, 141, 175–203.
- Husen, S. & Kissling, E. (2001). Local earthquake tomography between rays and waves : fat ray tomography. *Physics of the earth and Planetary Interiors*, 123(2), 127–147.
- Iglesias, A., Clayton, R. W., Perez-Campos, X., Singh, S. K., Pacheco, J. F., Garcia, D., & Valdes-Gonzalez, C. (2010). S wave velocity structure below central Mexico using high-resolution surface wave tomography. *Journal of Geophysical Research*, 115.
- INGV Seismological Data Centre (1997). Rete sismica nazionale (rsn), istituto nazionale di geofisica e vulcanologia (ingv), italy.
- Jeffreys, H. (1926). The rigidity of the earth's central core. *Geophysical Journal International*, 1(s7), 371–383.
- Jin, S., Madariaga, R., Virieux, J., & Lambaré, G. (1992). Two-dimensional asymptotic iterative elastic inversion. *Geophysical Journal International*, 108, 575–588.
- Julià, J., Ammon, C. J., Herrmann, R. B., & Correig, A. M. (2000). Joint inversion of receiver function and surface wave dispersion observations. *Geophysical Journal International*, (pp. 99–112).
- Kamei, R. & Pratt, R. G. (2013). Inversion strategies for visco-acoustic waveform inversion. *Geophysical Journal International*, 194, 859–894.
- 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.
- Kausel, E. (2006). *Fundamental solutions in elastodynamics : A compendium*. Cambridge University Press.
- Kausel, E. (2012). Lamb's problem at its simplest. In *Proc. R. Soc. A : The Royal Society*.
- Kelley, C. (1999). *Iterative Methods for Optimization*. SIAM.
- Kelly, K., Ward, R., Treitel, S., & Alford, R. (1976). Synthetic seismograms - a finite-difference approach. *Geophysics*, 41, 2–27.
- Kennett, B. L. N. (1983). *Seismic wave propagation in stratified media*. Cambridge : Cambridge University Press.
- Kennett, B. L. N. & Engdahl, E. R. (1991). Travel times for global earthquake location and phase association. *Geophysical Journal International*, 105, 429–465.
- Kennett, B. L. N. & Kerry, N. J. (1979). Seismic waves in a stratified half space. *Geophysical Journal of the Royal Astronomical Society*, 57, 557–583.
- Kern, M. (2002). *Problèmes inverses*. Ecole Supérieure d'ingénieurs Léonard de Vinci.
- Kim, Y., Liu, Q., & Tromp, J. (2011). Adjoint centroid-moment tensor inversions. *Geophysical Journal International*, 186(1), 264–278.

- Kirkpatrick, S., Gelatt, C. D., & Vecchi, M. P. (1983). Optimization by simulated annealing. *Science*, 220(4598), 671–680.
- Kissling, E. (1993). Deep structure of the alps—what do we really know ? *Physics of the Earth and Planetary Interiors*, 79(1), 87–112.
- Kissling, E., Schmid, S. M., Lippitsch, R., Ansorge, J., & Fügenschuh, B. (2006). Lithosphere structure and tectonic evolution of the alpine arc : new evidence from high-resolution teleseismic tomography. *Geological Society, London, Memoirs*, 32(1), 129–145.
- Kissling, E. & Spakman, W. (1996). Interpretation of tomographic images of uppermost mantle structure : examples from the western and central alps. *Journal of Geodynamics*, 21(1), 97–111.
- Köhn, D., De Nil, D., Kurzmann, A., Przebindowska, A., & Bohlen, T. (2012). On the influence of model parametrization in elastic full waveform tomography. *Geophysical Journal International*, 191, 325–345.
- Komatitsch, D., Tromp, J., & Vilotte, J. P. (1998). The spectral element method for elastic wave equations : application to 2D and 3D seismic problems. In *Expanded abstracts*, volume II (pp. 1460–1463).
- Komatitsch, D. (1997). *Méthodes spectrales et éléments spectraux pour l'équation de l'élastodynamique 2D et 3D en milieu hétérogène*. PhD thesis, Institut de Géophysique du Globe de Paris.
- 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., Ritsema, J., & Tromp, J. (2002). The spectral-element method, Beowulf computing, and global seismology. *Science*, 298(5599), 1737–1742.
- Komatitsch, D. & Tromp, J. (1999a). Introduction to the spectral element method for 3D seismic wave propagation. *Geophysical Journal International*, 139, 806–822.
- Komatitsch, D. & Tromp, J. (1999b). Introduction to the spectral element method for three-dimensional seismic wave propagation. *Geophysical Journal International*, 139(3), 806–822.
- Komatitsch, D. & Tromp, J. (2002a). Spectral-element simulations of global seismic wave propagation—i. validation. *Geophysical Journal International*, 149(2), 390–412.
- Komatitsch, D. & Tromp, J. (2002b). Spectral-element simulations of global seismic wave propagation, II. 3D models, oceans, rotation, and self-gravitation. *Geophysical Journal International*, 150, 303–318.
- Komatitsch, D., Tsuboi, S., & Tromp, J. (2005). The spectral-element method in seismology. *Seismic Earth : Array Analysis of Broadband Seismograms*, (pp. 205–227).
- Komatitsch, D. & Vilotte, J. P. (1998). The spectral element method : an efficient tool to simulate the seismic response of 2D and 3D geological structures. *Bulletin of the Seismological Society of America*, 88, 368–392.
- Kurzmann, A., Przebindowska, A., Kohn, D., & Bohlen, T. (2013). Acoustic full waveform tomography in the presence of attenuation : a sensitivity analysis. *Geophysical Journal International*, 195(2), 985–1000.
- Lailly, P. (1983a). The seismic inverse problem as a sequence of before stack migrations. In R. Bednar & Weglein (Eds.), *Conference on Inverse Scattering, Theory and application*, Society for Industrial and Applied Mathematics, Philadelphia (pp. 206–220).

- Lailly, P. (1983b). The seismic problem as a sequence of before-stack migrations. In J. Bednar (Ed.), *Conference on Inverse Scattering : Theory and Applications* : SIAM, Philadelphia.
- Lailly, P. (1984). The seismic inverse problem as a sequence of before stack migrations. In R. Bednar & Weglein (Eds.), *Conference on Inverse Scattering, SIAM, Philadelphia* (pp. 206–220). : Soc. Ind. appl. Math.
- Lamb, H. (1904). On the propagation of tremors over the surface of an elastic solid. *Philosophical Transactions of the Royal Society A*, A 203, 1–42.
- Lambaré, G., Lucio, P. S., & Hanyga, A. (1996). Two-dimensional multivalued travelttime and amplitude maps by uniform sampling of ray field. *Geophysical Journal International*, 125, 584–598.
- Lambaré, G., Operto, S., Podvin, P., Thierry, P., & Noble, M. (2003). 3-D ray+Born migration/inversion - part 1 : theory. *Geophysics*, 68, 1348–1356.
- Langston, C. A. (1979). Structure under mount rainier, washington, inferred from teleseismic body waves. *Journal of Geophysical Research*, 84, 4749–4762.
- Lardeaux, J.-M., Schwartz, S., Tricart, P., Paul, A., Guillot, S., Béthoux, N., & Masson, F. (2006). A crustal-scale cross-section of the south-western alps combining geophysical and geological imagery. *Terra Nova*, 18(6), 412–422.
- Laubscher, H. (1990). The problem of the moho in the alps. *Tectonophysics*, 182(1), 9–20.
- Laubscher, H. (1991). The arc of the western alps today. *Eclogae Geologicae Helvetiae*, 84(3), 359–631.
- Lay, T. & Wallace, T. C. (1995). *Modern global seismology*. Academic Press.
- Lehmann, I. (1936). P'. *Publications du Bureau Central Séismologique International*, A14, 87–115.
- Lemoine, M., Graciansky, P. C. D., & Tricart, P. (2000). *De l'océan à la chaîne de montagnes : tectonique des plaques dans les Alpes*. Editions scientifiques Gordon-Breach.
- Levander, A. R. (1988). Fourth-order finite-difference P-SV seismograms. *Geophysics*, 53(11), 1425–1436.
- Lévêque, J.-J., Rivera, L., & Wittlinger, G. (1993). On the use of the checker-board test to assess the resolution of tomographic inversions. *Geophysical Journal International*, 115(1), 313–318.
- LeVeque, R. J. (2002). *Finite volume methods for hyperbolic problems*. New-York, USA : Cambridge University Press.
- Ligorria, J. P. & Ammon, C. J. (1999). Iterative deconvolution and receiver-function estimation. *Bulletin of Seismological Society of America*, 89, 1395 – 1400.
- Lions, J. L. (1968). *Contrôle optimal de systèmes gouvernés par des équations aux dérivées partielles*. Dunod, Paris.
- Lippitsch, R., Kissling, E., & Ansorge, J. (2003). Upper mantle structure beneath the alpine orogen from high-resolution teleseismic tomography. *Journal of Geophysical Research : Solid Earth*, 108(B8).
- Liu, Q. & Gu, Y. (2012). Seismic imaging : from classical to adjoint tomography. *Tectonophysics*, 566, 31–66.

- 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.
- Lombardi, D., Braunmiller, J., Kissling, E., & Giardini, D. (2008). Moho depth and poisson's ratio in the western-central alps from receiver functions. *Geophysical Journal International*, 173(1), 249–264.
- 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)*, (pp. 144–179).
- Lucio, P. S., Lambaré, G., & Hanyga, A. (1995). 3D multivalued dynamic ray tracing for migration/inversion. In *Expanded Abstracts* (pp. 442–446). : Bresilian Union of Geophysics and Latino-American Geophysical Union.
- Luenberger, D. G. (1973). *Introduction to linear and nonlinear programming*, volume 28. Addison-Wesley Reading, MA.
- Lumley, D. E., Claerbout, J. F., & Bevc, D. (1994). Anti-aliased kirchhoff 3-D migration. In *Expanded Abstracts, 64th Annual SEG Meeting and Exposition* (pp. 1282–1285). : Society of Exploration Geophysics.
- Luo, Y. & Schuster, G. T. (1991). Wave-equation travelttime inversion. *Geophysics*, 56(5), 645–653.
- Maggi, A., Tape, C., Chen, M., Chao, D., & Tromp, J. (2009). An automated time-window selection algorithm for seismic tomography. *Geophysical Journal International*, 178, 257–281.
- Malinowski, M., Operto, S., & Ribodetti, A. (2011). High-resolution seismic attenuation imaging from wide-aperture onshore data by visco-acoustic frequency-domain full waveform inversion. *Geophysical Journal International*, 186(3), 1179–1204.
- Marchant, R. & Stampfli, G. (1997). Subduction of continental crust in the western alps. *Tectonophysics*, 269(3), 217–235.
- Marfurt, K. (1984). Accuracy of finite-difference and finite-element modeling of the scalar and elastic wave equations. *Geophysics*, 49, 533–549.
- Marquering, H., Dahlen, F. A., & Nolet, G. (1999). Three-dimensional sensitivity kernels for finite-frequency traveltimes : the banana-doughnut paradox. *Geophysical Journal International*, 137, 805–815.
- Marquering, H., Nolet, G., & Dahlen, F. A. (1998). Three-dimensional waveform sensitivity kernels. *Geophysical Journal International*, 132, 521–534.
- Masson, F., Verdun, J., Bayer, R., & Debeglia, N. (1999). Une nouvelle carte gravimétrique des alpes occidentales et ses conséquences structurales et tectoniques. *Comptes Rendus de l'Académie des Sciences-Series IIA-Earth and Planetary Science*, 329(12), 865–871.
- Masson, Y., Cupillard, P., Capdeville, Y., & Romanowicz, B. (2014). On the numerical implementation of time-reversal mirrors for tomographic imaging. *Geophysical Journal International*, 196, 1580–1599.

- Masson, Y. & Romanowicz, B. (2016). Fast computation of synthetic seismograms within a medium containing remote localized perturbations : A numerical solution to the scattering problem. *Geophysical Journal International*, (pp. ggw412).
- 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.
- Menard, G. & Thouvenot, F. (1984). Ecaillage de la lithosphere europeenne sous les alpes occidentales ; arguments gravimetriques et sismiques lies a l'anomalie d'ivrea. *Bulletin de la Société Géologique de France*, 7(5), 875–884.
- Menke, W. (2012). *Geophysical data analysis : Discrete inverse theory*, volume 45. Academic press.
- Métivier, L., Bretaudeau, F., Brossier, R., Operto, S., & Virieux, J. (2014). Full waveform inversion and the truncated Newton method : quantitative imaging of complex subsurface structures. *Geophysical Prospecting*, 62, 1353–1375.
- Métivier, L. & Brossier, R. (2016a). The SEISCOPE optimization toolbox : A large-scale non-linear optimization library based on reverse communication. *Geophysics*, 81(2), F11–F25.
- Métivier, L. & Brossier, R. (2016b). The seiscope optimization toolbox : A large-scale nonlinear optimization library based on reverse communication. *Geophysics*, 81(2), F11–F25.
- Métivier, L., Brossier, R., Virieux, J., & Operto, S. (2013). Full Waveform Inversion and the truncated Newton method. *SIAM Journal On Scientific Computing*, 35(2), B401–B437.
- 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.
- Moczo, P., Robertsson, J. O. A., & Eisner, L. (2006). The finite-difference time-domain method for modelling of seismic wave propagation. In *Advances in wave propagation in heterogeneous earth*. Elsevier Academic Press.
- Mohorovicic, A. (1909). Das beben vom 8. x. 1909. *Jb. Met. Obs. Zagreb (Agram)*, 9, 1–63.
- Monteiller, V., Chevrot, S., Komatitsch, D., & Trom, J. (2013). A hybrid method to compute short period synthetic seismograms of teleseismic body waves in a 3-d regional model. *Geophysical Journal International*, 192(1), 230–247.
- Monteiller, V., Chevrot, S., Komatitsch, D., & Wang, Y. (2015). Three-dimensional full waveform inversion of short-period teleseismic wavefields based upon the sem-dsm hybrid method. *Geophysical Journal International*, 202(2), 811–827.
- 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. (2004a). Finite-frequency tomography reveals a variety of plumes in the mantle. *Science*, 303, 338–343.
- Montelli, R., Nolet, G., Masters, G., Dahlen, F. A., & Hung, S. H. (2004b). P and PP global travel time tomography. *Geophysical Journal International*, 158, 637–654.

- Mora, P. R. (1987). Nonlinear two-dimensional elastic inversion of multi-offset seismic data. *Geophysics*, 52, 1211–1228.
- Mosegaard, K. & Tarantola, A. (1995). Monte Carlo sampling of solutions to inverse problems. *Journal of Geophysical Research*, 100 B7, 12431–12447.
- Mulder, W. A. & Plessix, R. E. (2004). How to choose a subset of frequencies in frequency-domain finite-difference migration. *Geophysical Journal International*, 158, 801–812.
- Nicolas, A., Hirn, A., Nicolich, R., & Polino, R. (1990). Lithospheric wedging in the western alps inferred from the ecors-crop traverse. *Geology*, 18(7), 587–590.
- Nissen-Meyer, T., Dahlen, F. A., & Fournier, A. (2007a). Spherical-earth frechet sensitivity kernels. *Geophysical Journal International*, 168, 1051–1066.
- Nissen-Meyer, T., Fournier, A., & Dahlen, F. A. (2007b). A two-dimensional spectral-element method for computing spherical-earth seismograms - i. moment-tensor source. *Geophysical Journal International*, 168, 1067–1092.
- Nissen-Meyer, T., Fournier, A., & Dahlen, F. A. (2008). A 2-d spectral-element method for computing spherical-earth seismograms - ii. waves in solid-fluid media. *Geophysical Journal International*, 174, 873–888.
- Nissen-Meyer, T., van Driel, M., Stähler, S., Hosseini, K., Hempel, S., Auer, L., Colombi, A., & Fournier, A. (2014). Axisem : broadband 3-d seismic wavefields in axisymmetric media. *Solid Earth*, 5(1), 425–445.
- Noble, M., Gesret, A., & Belayouni, N. (2014). Accurate 3-d finite difference computation of travel time in strongly heterogeneous media. *Geophysical Journal International*, 199, 1572–1585.
- Nocedal, J. (1980). Updating Quasi-Newton Matrices With Limited Storage. *Mathematics of Computation*, 35(151), 773–782.
- Nocedal, J. & Wright, S. J. (2006a). *Numerical Optimization*. Springer, 2nd edition.
- Nocedal, J. & Wright, S. J. (2006b). *Numerical Optimization*. Springer, 2nd edition.
- Nocquet, J.-M., Sue, C., Walpersdorf, A., Tran, T., Lenôtre, N., Vernant, P., Cushing, M., Jouanne, F., Masson, F., Baize, S., et al. (2016). Present-day uplift of the western alps. *Scientific Reports*, 6.
- Nolet, G. (1985). Solving or resolving inadequate and noisy tomographic systems. *Journal of Computational Physics*, 61, 463–482.
- Nolet, G. (1987). *Seismic tomography with applications in global seismology and exploration geophysics*. D. Reidel publishing Company.
- Nolet, G. (2008). *A Breviary of Seismic Tomography*. Cambridge, UK : Cambridge University Press.
- O’Brien, G. S. (2014). Elastic lattice modelling of seismic waves including a free surface. *Computers & Geosciences*, 67, 117–124.
- Ohminato, T. & Chouet, B. A. (1997). A free-surface boundary condition for including 3D topography in the finite-difference method. *Bulletin of the Seismological Society of America*, 87, 494–515.

- Oldham, R. (1906). The constitution of the earth. *Quarterly Journal of the Geological Society of London*, 62, 456–475.
- Operto, S., Brossier, R., Gholami, Y., Métivier, L., Prioux, V., Ribodetti, A., & Virieux, J. (2013). A guided tour of multiparameter full waveform inversion for multicomponent data : from theory to practice. *The Leading Edge*, Special section Full Waveform Inversion(September), 1040–1054.
- Operto, S., Miniussi, A., Brossier, R., Combe, L., Métivier, L., Monteiller, V., Ribodetti, A., & Virieux, J. (2015). Efficient 3-D frequency-domain mono-parameter full-waveform inversion of ocean-bottom cable data : application to Valhall in the visco-acoustic vertical transverse isotropic approximation. *Geophysical Journal International*, 202(2), 1362–1391.
- Owens, T. J., Zandt, G., & Taylor, S. R. (1984). Seismic evidence for an ancient rift beneath the Cumberland Plateau, Tennessee : A detailed analysis of broadband teleseismic P waveforms. *Journal of Geophysical Research*, 89, 7783–7795.
- O'Brien, G., Nissen-Meyer, T., & Bean, C. (2012). A lattice boltzmann method for elastic wave propagation in a poisson solid. *Bulletin of the Seismological Society of America*, 102, 1224–1234.
- Pageot, D. (2012). *Imagerie lithosphérique par inversion de formes d'ondes de données télé-sismiques*. PhD thesis, Université Nice Sophia-Antipolis.
- Pageot, D., Operto, S., Vallée, M., Brossier, R., & Virieux, J. (2010). Lithospheric imaging from teleseismic data by frequency-domain elastic full-waveform tomography. In *Expanded abstracts* (pp. WS6). : EAGE.
- Pageot, D., Operto, S., Vallée, M., Brossier, R., & Virieux, J. (2013). A parametric analysis of two-dimensional elastic full waveform inversion of teleseismic data for lithospheric imaging. *Geophysical Journal International*, 193(3), 1479–1505.
- Paige, C. C. & Saunders, M. A. (1982). ALGORITHM 583 LSQR : Sparse linear equations and least squares problems. *ACM Transactions on Mathematical Software*, 8(2), 195–209.
- Park, J. & Levin, V. (2000). Receiver functions from multiple-taper spectral correlation estimates. *Bulletin of Seismological Society of America*, 90, 1507 – 1520.
- Paul, A., Cattaneo, M., Thouveot, F., Spallarossa, D., Béthoux, N., & Fréchet, J. (2001). A three-dimensional crustal velocity model of the southwestern alps from local earthquake tomography. *Journal of Geophysical Research*, 106(19), 367–390.
- Pfiffner, O. A. (2014). *Geology of the Alps*. John Wiley & Sons.
- Piromallo, C. & Faccenna, C. (2004). How deep can we find the traces of alpine subduction ? *Geophysical Research Letters*, 31(6).
- Piromallo, C. & Morelli, A. (2003). P wave tomography of the mantle under the alpine-mediterranean area. *Journal of Geophysical Research : Solid Earth*, 108(B2).
- Plessix, R. E. (2006a). 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. (2006b). A robust three-dimensional iterative solver for the time-harmonic wave equation. In *Extended Abstracts* : Eur. Ass. Expl.Geophys.

- Plessix, R. E. & Cao, Q. (2011). A parametrization study for surface seismic full waveform inversion in an acoustic vertical transversely isotropic medium. *Geophysical Journal International*, 185, 539–556.
- Plonka, A., Blom, N., & Fichtner, A. (2016). The imprint of crustal density heterogeneities on regional seismic wave propagation. *Solid Earth Discussions*, 2016, 1–28.
- Podvin, P. & Lecomte, I. (1991). Finite difference computation of traveltimes in very contrasted velocity model : a massively parallel approach and its associated tools. *Geophysical Journal International*, 105, 271–284.
- Potin, B., Valette, B., Thouvenot, F., & Monteiller, V. (2015). From local earthquakes tomography to moho topography beneath the western alps. In *EGU General Assembly Conference Abstracts*, volume 17 (pp. 2690).
- Pozrikidis, C. (2005). *Introduction to finite and spectral element methods using MATLAB*. CRC Press.
- Pratt, R. G. (1990a). Frequency-domain elastic modeling by finite differences : a tool for cross-hole seismic imaging. *Geophysics*, 55(5), 626–632.
- Pratt, R. G. (1990b). 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 physical scale model. *Geophysics*, 64, 888–901.
- Pratt, R. G. & Sams, M. S. (1996). Reconciliation of crosshole seismic velocities with well information in a layered sedimentary environment. *Geophysics*, 61, 549–560.
- 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. & Worthington, M. H. (1990). Inverse theory applied to multi-source cross-hole tomography. Part I : acoustic wave-equation method. *Geophysical Prospecting*, 38, 287–310.
- Prieux, V., Brossier, R., Operto, S., & Virieux, J. (2013a). Multiparameter full waveform inversion of multicomponent OBC data from valhall. Part 1 : imaging compressional wavespeed, density and attenuation. *Geophysical Journal International*, 194(3), 1640–1664.
- Prieux, V., Brossier, R., Operto, S., & Virieux, J. (2013b). Multiparameter full waveform inversion of multicomponent OBC data from valhall. Part 2 : imaging compressional and shear-wave velocities. *Geophysical Journal International*, 194(3), 1665–1681.
- Prieux, V., Brossier, R., Operto, S., & Virieux, J. (2013c). Multiparameter full waveform inversion of multicomponent ocean-bottom-cable data from Valhall field. Part 1 : imaging compressional wave speed, density and attenuation. *Geophysical Journal International*, 194(3), 1640–1664.
- Priolo, E., Carcione, J. M., & Seriani, G. (1994). Numerical simulation of interface waves by high-order spectral modeling techniques. *Journal of the Acoustical Society of America*, 95, 681–693.
- Przebindowska, A., Kurzmann, A., Köhn, D., & Bohlen, T. (2012). The role of density in acoustic full waveform inversion of marine reflection seismics. In *74th EAGE Conference and Exhibition incorporating EUROPEC 2012*.

- Randall, G. (1989). Efficient calculation of differential seismograms for lithospheric receiver functions. *Geophysical Journal International*, 99(3), 469–481.
- 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.
- Rawlinson, N., Pozgay, S., & Fishwick, S. (2010). Seismic tomography : a window into deep earth. *Physics of the Earth and Planetary Interiors*, 178(3), 101–135.
- Rawlinson, N. & Spakman, W. (2016). On the use of sensitivity tests in seismic tomography. *Geophysical Journal International*, 205(2), 1221–1243.
- 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.
- Renard, M., Lagabrielle, Y., Martin, E., et al. (2015). *Eléments de géologie*.
- RESIF (1995). Resif-rlbp french broad-band network, resif-rap strong motion network and other seismic stations in metropolitan france, resif – réseau sismologique et géodésique français.
- 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.
- Robertsson, J. & Chapman, C. (2000). An efficient method for calculating finite-difference seismograms after model alterations. *Geophysics*, 65(3), 907–918.
- Roecker, S., Baker, B., & McLaughlin, J. (2010). A finite-difference algorithm for full waveform teleseismic tomography. *Geophysical Journal International*, 181, 1017–1040.
- Rondenay, S., Bostock, M. G., & Fischer, K. M. (2005). Multichannel inversion of scattered teleseismic body waves : practical considerations and applicability. In G. N. Alan Levander (Ed.), *Seismic Earth : array analysis of broadband seismograms*, *Geophysical Monograph 157* (pp. 187–203). American Geophysical Union Washington, DC. In *Seismic Earth : array analysis of broadband seismograms*, *Geophysical Monograph 157*.
- Rondenay, S., Bostock, M. G., & Shragge, J. (2001). Multiparameter two-dimensional inversion of scattered teleseismic body waves. 3. Application to the Cascadia 1993 data set. *Journal of Geophysical Research*, 106(12), 30795–30807.
- Roure, F. (1996). *The ECORS-CROP Alpine seismic traverse*. Number 170. Société géologique de France.
- Ryberg, T. & Weber, M. (2000). Receiver function arrays : a reflection seismic approach. *Geophysical Journal International*, 141(1), 1–11.
- 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. & Mosegaard, K. (2002). Monte Carlo methods in geophysical inverse problems. *Reviews of Geophysics*, 40(3), 1–29.
- Sambridge, M. S. (1999a). Geophysical inversion with a neighbourhood algorithm - I. Searching a parameter space. *Geophysical Journal International*, 138, 479–494.

- Sambridge, M. S. (1999b). Geophysical inversion with a neighbourhood algorithm - II. Appraising the ensemble. *Geophysical Journal International*, 138, 727–746.
- Scales, J. A. & Snieder, R. (2000). The anatomy of inverse problems. *Geophysics*, 65, 1708–1710.
- Schmid, S. & Kissling, E. (2000). The arc of the western alps in the light of geophysical data on deep crustal structure. *Tectonics*, 19(1), 62–85.
- Schubert, B. (2003). *The spectral-element method for seismic wave propagation : theory, implementation and comparison to finite difference methods*. PhD thesis, LMU.
- Sen, M. & Stoffa, P. (1991). Non-linear one-dimensional seismic waveform inversion using simulated annealing. *Geophysics*, 56, 1624–1638.
- Sen, M. K. & Stoffa, P. L. (1995). *Global Optimization Methods in Geophysical Inversion*. Elsevier Science Publishing Co.
- Seriani, G. & Oliveira, S. (2008). Dispersion analysis of spectral element methods for elastic wave propagation. *Wave Motion*, 45(6), 729–744.
- Shang, X., de Hoop, M. V., & van der Hilst, R. D. (2012). Beyond receiver functions : Passive source reverse time migration and inverse scattering of converted waves. *Geophysical Research Letters*, 39(15).
- Sheen, D.-H., Tuncay, K., Baag, C.-E., & Ortoleva, P. J. (2006). Time domain gauss—newton seismic waveform inversion in elastic media. *Geophysical Journal International*, 167(3), 1373–1384.
- Shin, C., Yoon, K., Marfurt, K. J., Park, K., Yang, D., Lim, H. Y., Chung, S., & Shin, S. (2001). Efficient calculation of a partial derivative wavefield using reciprocity for seismic imaging and inversion. *Geophysics*, 66(6), 1856–1863.
- Shragge, J., Bostock, M. G., & Rondenay, S. (2001). Multiparameter two-dimensional inversion of scattered teleseismic body waves 2. numerical examples. *Journal of Geophysical Research*, 106(12), 30783–30793.
- Sieminski, A., Trampert, J., & Tromp, J. (2009). Principal component analysis of anisotropic finite-frequency sensitivity kernels. *Geophysical Journal International*, 179(2), 1186–1198.
- Simons, F., Nolet, G., Babcock, J., Davis, R., & Orcutt, J. (2006). A future for drifting seismic networks. *Eos*, 87(31), 305–307.
- Simons, F. J., Nolet, G., Georgief, P., Babcock, J. M., Regier, L. A., & Davis, R. E. (2009). On the potential of recording earthquakes for global seismic tomography by low-cost autonomous instruments in the oceans. *Journal of Geophysical Research : Solid Earth*, 114(B5).
- Sirgue, L. (2003). *Inversion de la forme d'onde dans le domaine fréquentiel de données sismiques grand offset*. PhD thesis, Université Paris 11, France - Queen's University, Canada.
- Sirgue, L. & Pratt, R. G. (2004). Efficient waveform inversion and imaging : a strategy for selecting temporal frequencies. *Geophysics*, 69(1), 231–248.
- Snieder, R. (1986). 3-d linearized scattering of surface waves and a formalism for surface wave holography. *Geophysical Journal International*, 84(3), 581–605.
- Snieder, R. & Nolet, G. (1987). Linearized scattering of surface-waves on a spherical earth. *Journal of Geophysics - Zeitschrift für Geophysik*, 61(1), 55–63.

- Spada, M., Bianchi, I., Kissling, E., Agostinetti, N. P., & Wiemer, S. (2013). Combining controlled-source seismology and receiver function information to derive 3-d moho topography for italy. *Geophysical Journal International*, 194(2), 1050–1068.
- Spakman, W., van der Lee, S., & van der Hilst, R. (1993). Travel-time tomography of the european-mediterranean mantle down to 1400 km. *Physics of the Earth and Planetary Interiors*, 79(1), 3–74.
- Stacey, R. (1988). Improved transparent boundary formulations for the elastic-wave equation. *Bulletin of the Seismological Society of America*, 78(6), 2089–2097.
- Stampfli, G., Borel, G. D., Marchant, R., & Mosar, J. (2002). Western alps geological constraints on western tethyan reconstructions. *Journal of the Virtual Explorer*, 8, 77.
- Stehly, L., Fry, B., Campillo, M., Shapiro, N., Guilbert, J., Boschi, L., & Giardini, D. (2009). Tomography of the Alpine region from observations of seismic ambient noise. *Geophysical Journal International*, 178, 338–350.
- Stein, S. & Wysession, M. (2009). *An introduction to seismology, earthquakes, and earth structure*. John Wiley & Sons.
- Stoneley, R. (1926). The effect of the ocean on rayleigh waves. *Geophysical Journal International*, 1(s7), 349–356.
- Sue, C., Thouvenot, F., Fréchet, J., & Tricart, P. (1999). Widespread extension in the core of the western alps revealed by earthquake analysis. *Journal of Geophysical Research*, 104(2561125622).
- Symes, W. W. (2007). Reverse time migration with optimal checkpointing. *Geophysics*, 72(5), SM213–SM221.
- Taflove, A. & Hagness, S. C. (2000). *Computational electrodynamics : the finite-difference time-domain method*. London, United Kindom : Artech House.
- Taflove, A. & Hagness, S. C. (2005). *Computational Electrodynamics : The Finite-Difference Time-Domain Method*. Artech House, 3rd edition.
- Tago, J., Métivier, L., & Virieux, J. (2014). SMART layers : a simple and robust alternative to PML approaches for elastodynamics. *Geophysical Journal International*, 199(2), 700–706.
- Takeuchi, N. & Geller, R. J. (2000). Optimally accurate second-order time-domain finite-difference scheme for computing synthetic seismograms in 2-D and 3-D media. *Physics of the Earth and Planetary Interiors*, 119, 99–131.
- Takeuchi, N., Geller, R. J., & Cummins, P. R. (2000). Complete synthetic seismograms for 3-d heterogeneous earth models computed using modified dsm operators and their applicability to inversion for earth structure. *Physics of the Earth and Planetary Interiors*, 119, 25–36.
- Tape, C., Liu, Q., Maggi, A., & Tromp, J. (2009). Adjoint tomography of the southern california crust. *Science*, 325, 988–992.
- Tape, C., Liu, Q., Maggi, A., & Tromp, J. (2010). Seismic tomography of the southern California crust based on spectral-element and adjoint methods. *Geophysical Journal International*, 180, 433–462.
- Tarantola, A. (1984a). Inversion of seismic reflection data in the acoustic approximation. *Geophysics*, 49(8), 1259–1266.

- Tarantola, A. (1984b). Linearized inversion of seismic reflection data. *Geophysical Prospecting*, 32, 998–1015.
- 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*. New York : Elsevier.
- Tarantola, A. (2005). *Inverse Problem theory and methods for model parameter estimation*. Philadelphia : Society for Industrial and Applied Mathematics.
- Tarantola, A. & Nercessian, A. (1984). Three-dimensional inversion without blocks. *Geophysical Journal of the Royal Astronomical Society*, 76, 299–306.
- Tarantola, A. & Valette, B. (1982). Inverse problems = quest for information. *Journal of Geophysics*, 50, 159–170.
- Tarrass, I., Giraud, L., & Thore, P. (2011). New curvilinear scheme for elastic wave propagation in presence of curved topography. *Geophysical Prospecting*, 59(5), 889–906.
- Thierry, P., Lambaré, G., Podvin, P., & Noble, M. (1996). 3D prestack preserved amplitude migration : application to real data. In *Expanded Abstracts* (pp. 555–558). : Society of Exploration Geophysics.
- Thouvenot, F., Paul, A., Frechet, J., Béthoux, N., Jenatton, L., & Guiguet, R. (2007). Are there really superposed mohos in the southwestern alps ? new seismic data from fan-profiling reflections. *Geophysical Journal International*, 170(3), 1180–1194.
- Tiberi, C., Diament, M., Déverchère, J., Petit-Mariani, C., Mikhailov, V., Tikhotsky, S., & Achauer, U. (2003). Deep structure of the baikal rift zone revealed by joint inversion of gravity and seismology. *Journal of Geophysical Research : Solid Earth*, 108(B3).
- Tikhonov, A. & Arsenin, V. (1977). *Solution of ill-posed problems*. Winston, Washington, DC.
- Tikhonov, A. N. (1963). Resolution of ill-posed problems and the regularization method (in russian, french translation, mir, moscow, 1976). *Dokl. Akad. Nauk SSSR*, 151, 501–504.
- Tong, P., Chen, C.-w., Komatitsch, D., Basini, P., & Liu, Q. (2014a). High-resolution seismic array imaging based on an sem-fk hybrid method. *Geophysical Journal International*, 197, 369–395.
- Tong, P., Komatitsch, D., Tseng, T.-L., Hung, S.-H., Chen, C.-W., Basini, P., & Liu, Q. (2014b). A 3-d spectral-element and frequency-wave number hybrid method for high-resolution seismic array imaging. *Geophysical Research Letters*, 41, 7025–7034.
- Trampert, J., Fichtner, A., & Ritsema, J. (2013). Resolution tests revisited : the power of random numbers. *Geophysical Journal International*, 192(2), 676–680.
- Tromp, J., Tape, C., & Liu, Q. (2005). Seismic tomography, adjoint methods, time reversal and banana-doughnut kernels. *Geophysical Journal International*, 160, 195–216.
- University of Genova (1967). Regional seismic network of north western italy, international federation of digital seismograph networks, other/seismic network,.
- Unser, M. (2000). Sampling-50 years after shannon. *Proceedings of the IEEE*, 88(4), 569–587.

- Vallée, M., Charléty, M., Ferreira, A., Delouis, B., & Vergoz, J. (2011). Scardec : a new technique for the rapid determination of seismic moment magnitude, focal mechanism and source time functions for large earthquakes using body wave deconvolution. *Geophysical Journal International*, 184, 338–358.
- Vallée, M. & Douet, V. (2016). A new database of source time functions (stfs) extracted from the scardecmethod. *Physics of the Earth and Planetary Interiors*.
- Vanorio, T., Virieux, J., Capuano, P., & Russo, G. (2005). Three-dimensional seismic tomography from p wave and s wave microearthquake travel times and rock physics characterization of the campi flegrei caldera. *Journal of Geophysical Research*, 110, B03201.
- Vernant, P., Masson, F., Bayer, R., & Paul, A. (2002). Sequential inversion of local earthquake traveltimes and gravity anomaly—the example of the western alps. *Geophysical Journal International*, 150(1), 79–90.
- Vignaroli, G., Faccenna, C., Jolivet, L., Piromallo, C., & Rossetti, F. (2008). Subduction polarity reversal at the junction between the western alps and the northern apennines, italy. *Tectonophysics*, 450(1), 34–50.
- Vinnik, L. (1977). Detection of waves converted from p to sv in the mantle. *Physics of the Earth and planetary interiors*, 15(1), 39–45.
- Virieux, J. (1986). P-SV wave propagation in heterogeneous media : Velocity-stress finite difference method. *Geophysics*, 51, 889–901.
- Virieux, J., Etienne, V., Cruz-Atienza, V., Brossier, R., Chaljub, E., Coutant, O., Garambois, S., Mercerat, D., Prioux, V., Operto, S., Ribodetti, A., & Tago, J. (2012). Modelling seismic wave propagation for geophysical imaging. In M. Kanao (Ed.), *Seismic Waves - Research and Analysis* (pp. 253–304). InTech.
- Virieux, J. & Operto, S. (2009). An overview of full waveform inversion in exploration geophysics. *Geophysics*, 74(6), WCC1–WCC26.
- Voigt, W. (1889). Über die beziehung zwischen den beiden elastizitatkonstanten isotroper körper. *Annalen der Physik*, 38, 573–587.
- von Blanckenburg, F. & Davies, J. H. (1995). Slab breakoff : a model for syncollisional magmatism and tectonics in the alps. *Tectonics*, 14(1), 120–131.
- von Rebeur-Paschwitz, E. (1889). The earthquake of tokiyo, april 18, 1889. *Nature*, 40, 294–295.
- Wagner, M., Kissling, E., & Husen, S. (2012). Combining controlled-source seismology and local earthquake tomography to derive a 3-d crustal model of the western alpine region. *Geophysical Journal International*, 191(2), 789–802.
- Waldhauser, F., Lippitsch, R., Kissling, E., & Ansorge, J. (2002). High-resolution teleseismic tomography of upper-mantle structure using an a priori three-dimensional crustal model. *Geophysical Journal International*, 150(2), 403–414.
- Wang, Y., Chevrot, S., Monteiller, V., Komatitsch, D., Mouthereau, F., Manatschal, G., Sylvander, M., Diaz, J., Ruiz, M., Grimaud, F., et al. (2016). The deep roots of the western pyrenees revealed by full waveform inversion of teleseismic p waves. *Geology*, 44(6), 475–478.
- Warner, M., Ratcliffe, A., Nangoo, T., Morgan, J., Umpleby, A., Shah, N., Vinje, V., Stekl, I., Guasch, L., Win, C., Conroy, G., & Bertrand, A. (2013). Anisotropic 3D full-waveform inversion. *Geophysics*, 78(2), R59–R80.

- Weidle, C. & Widiyantoro, S. (2005). Improving depth resolution of teleseismic tomography by simultaneous inversion of teleseismic and global p-wave traveltime data—application to the vrancea region in southeastern europe. *Geophysical Journal International*, 162(3), 811–823.
- Wiechert, E. (1897). Über die massenverteilung im inneren der erde. *Nachrichten von der Gesellschaft der Wissenschaften zu Göttingen, Mathematisch-Physikalische Klasse*, 1897, 221–243.
- Wilcox, L. C., Stadler, G., Burstedde, C., & Ghattas, O. (2010). A High-Order Discontinuous Galerkin Method for Wave Propagation Through Coupled Elastic-Acoustic Media. *Journal of Computational Physics*, 229(24), 9373–9396.
- Williamson, P. (1991). A guide to the limits of resolution imposed by scattering in ray tomography. *Geophysics*, 56, 202–207.
- Wilson, C. & Guitton, A. (2007). Teleseismic wavefield interpolation and signal extraction using high-resolution linear radon transforms. *Geophysical Journal International*, 168(1), 171–181.
- Woodward, M. J. (1992). Wave-equation tomography. *Geophysics*, 57, 15–26.
- Wortel, M. & Spakman, W. (2000). Subduction and slab detachment in the mediterranean-carpathian region. *Science*, 290(5498), 1910–1917.
- Wu, R. S. & Aki, K. (1985). Scattering characteristics of elastic waves by an elastic heterogeneity. *Geophysics*, 50(4), 582–595.
- Wu, R.-S. & Aki, K. (1988). Introduction : Seismic wave scattering in three-dimensionally heterogeneous earth. In *Scattering and Attenuations of Seismic Waves, Part I* (pp. 1–6). Springer.
- Wu, R. S. & Toksöz, M. N. (1987). Diffraction tomography and multisource holography applied to seismic imaging. *Geophysics*, 52, 11–25.
- Ye, R., de Hoop, M. V., Petrovitch, C. L., Pyrak-Nolte, L. J., & Wilcox, L. C. (2016). A Discontinuous Galerkin method with a modified penalty flux for the propagation and scattering of acousto-elastic waves. *Geophysical Journal International*, 205(2), 1267–1289.
- Yuan, X., Kind, R., Li, X., & Wang, R. (2006). The s receiver functions : synthetics and data example. *Geophysical Journal International*, 165(2), 555–564.
- Yuan, Y., Simons, F., & Bozdogan, E. (2015). Multiscale adjoint waveform tomography for surface and body waves. *Geophysics*, 80, R281–R302.
- Yuan, Y. O., Simons, F. J., & Tromp, J. (2016). Double-difference adjoint seismic tomography. *Geophysical Journal International*, 206(3), 1599–1618.
- 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.
- Zeng, C., Xia, J., Miller, R. D., & Tsoflias, G. P. (2012). An improved vacuum formulation for 2d finite-difference modeling of rayleigh waves including surface topography and internal discontinuities. *Geophysics*, 77(1), T1–T9.
- Zhang, Q., Zhou, H., Li, Q., Chen, H., & Wang, J. (2016). Robust source-independent elastic full-waveform inversion in the time domain. *Geophysics*, 81(2), R29–R44.

- Zhang, W. & Chen, X. (2006). Traction image method for irregular free surface boundaries in finite difference seismic wave simulation. *Geophysical Journal International*, 167(1), 337–353.
- Zhao, L., Paul, A., Guillot, S., Solarino, S., Malusà, M. G., Zheng, T., Aubert, C., Salimbeni, S., Dumont, T., Schwartz, S., et al. (2015). First seismic evidence for continental subduction beneath the western alps. *Geology*, 43(9), 815–818.
- Zhao, L., Paul, A., Malusà, M. G., Xu, X., Zheng, T., Solarino, S., Guillot, S., Schwartz, S., Dumont, T., Salimbeni, S., Aubert, C., Pondrelli, S., Wang, Q., & Rixiang (2016a). Continuity of the alpine slab unraveled by high-resolution P wave tomography. *Journal of Geophysical Research : Solid Earth*, (pp. n/a–n/a). 2016JB013310.
- Zhao, L., Paul, A., Solarino, S., Aubert, C., Zheng, T., Salimbeni, S., Guillot, S., Zhu, R., & Wang, Q. (2013). The civalps seismic experiment : first high-resolution data on the crust and upper mantle structures of the southwestern alps. In *AGU Fall Meeting Abstracts*, volume 1 (pp. 2586).
- Zhao, L., Paul, A., Solarino, S., & RESIF (2016b). Seismic network yp : Civalps temporary experiment (china-italy-france alps seismic transect) ; resif - réseau sismologique et géodésique français.
- Zhinan, X., Komatitsch, D., Martin, R., & Matzen, R. (2014). Improved forward wave propagation and adjoint-based sensitivity kernel calculations using a numerically stable finite-element PML. *Geophysical Journal International*, in press.
- Zhu, H., Bozdağ, E., Peter, D., & Tromp, J. (2012). Structure of the european upper mantle revealed by adjoint tomography. *Nature Geoscience*, 5, 493–498.
- Zhu, H., Bozdağ, E., & Tromp, J. (2015). Seismic structure of the European upper mantle based on adjoint tomography. *Geophysical Journal International*, 201(1), 18–52.
- Zhu, L. (2000). Crustal structure across the san andreas fault, southern california from teleseismic converted waves. *Earth and Planetary Science Letters*, 179(1), 183–190.
- Zhu, L. & Kanamori, H. (2000). Moho depth variation in southern california from teleseismic receiver functions. *Journal of Geophysical Research : Solid Earth*, 105(B2), 2969–2980.
- Zhu, L. & Rivera, L. A. (2002). A note on the dynamic and static displacements from a point source in multilayered media. *Geophysical Journal International*, 148(3), 619–627.
- Ziegler, P. & Dèzes, P. (2006). Crustal evolution of western and central europe. *Geological Society, London, Memoirs*, 32(1), 43–56.
- Zienkiewicz, O. & Taylor, R. (1967). *The Finite Element Method for Solid and Structural Mechanics*. McGraw Hill, New York.
- Zienkiewicz, O. C., Taylor, R. L., & Zhu, J. Z. (2005). *The Finite Element Method : Its Basis and Fundamentals*. London : Elsevier. 6th edition.

Résumé

L'imagerie fine de la lithosphère est essentielle pour comprendre les mécanismes profonds qui gouvernent la dynamique terrestre et valider les modèles géodynamiques déduits des observations de surface. En sismologie, les méthodes de tomographie ou de migration de fonctions récepteur n'ont pas toujours la résolution suffisante pour proposer une interprétation géodynamique univoque. Les déploiements récents de réseaux denses de capteurs ainsi que l'augmentation des capacités de calcul permettent aujourd'hui l'inversion de formes d'ondes complètes pour l'imagerie de la lithosphère à partir de données téléseismiques. Cette technique permet d'obtenir des modèles quantitatifs du sous-sol avec une résolution spatiale proche de la longueur d'onde en exploitant la totalité de l'information contenue dans les sismogrammes.

Dans cette thèse, un algorithme massivement parallèle d'inversion de formes d'ondes téléseismiques est développé pour des milieux élastiques isotropes 3D. La modélisation sismique est effectuée efficacement avec une méthode hybride d'injection de champ d'onde. Une première modélisation est effectuée dans une Terre globale à symétrie sphérique avec le logiciel AxiSEM pour déterminer les champs d'ondes aux bords de la cible lithosphérique. Ces solutions sont ensuite propagées dans cette cible avec une méthode aux éléments finis spectraux. Le problème inverse est résolu avec un algorithme d'optimisation locale de type quasi-Newton (I-BFGS).

La sensibilité de la méthode à la configuration expérimentale (paramétrisation du milieu, modèle initial, géométrie et échantillonnage du dispositif de capteurs) est tout d'abord analysée avec un modèle synthétique réaliste des Alpes occidentales. Les trois principales conclusions sont : (i) Des modèles précis et bien résolus sont obtenus en reconstruisant la densité et les paramètres de Lamé (ou les vitesses P et S) durant l'inversion et en recombinaison des impédances P et S a posteriori. (ii) Des modèles terrestres de référence tels que PREM fournissent des modèles initiaux acceptables. (iii) Des déploiements surfaciques de capteurs par opposition à des lignes denses améliorent significativement la résolution verticale et horizontale.

L'algorithme est finalement appliqué à neuf événements de la campagne CIFALPS dans les Alpes occidentales jusqu'à une fréquence de 0.2Hz. Les modèles de vitesses P et S et de densité révèlent les grandes structures lithosphériques de la chaîne alpine avec une résolution encore jamais atteinte, en particulier la géométrie des grands bassins sédimentaires, du corps d'Ivrée et des Moho européen et adriatique. Plus profondément, deux anomalies de vitesses lentes sous les Alpes occidentales sont imagées dans le manteau. La première à l'Ouest de la chaîne est interprétée comme la signature d'une remontée asthénosphérique, la seconde au niveau du corps d'Ivrée indique le détachement du panneau plongeant européen. L'étude valide ainsi l'hypothèse d'une subduction continentale de la croûte européenne et confirme un éventuel déchirement du panneau plongeant européen au cours de la collision continentale.

Abstract

Fine-scale imaging of lithospheric structures is crucial to understand the deep mechanisms that govern the Earth dynamics and to appraise geodynamical models inferred from surface observations. In seismology, traveltime tomography or receiver functions migration have often an insufficient resolution to propose an unique geodynamical interpretation of lithospheric structures. However, the recent deployments of dense seismic arrays, together with the continuous growth in computing facilities, are amenable to full-waveform inversion (FWI) of teleseismic data for high-resolution lithospheric imaging. This technique can produce quantitative multi-parameter models of the subsurface with a resolution close to the wavelength by handling the full information carried out by very general wave types.

In this thesis, a massively-parallel teleseismic waveform inversion algorithm is developed for 3D elastic isotropic media. Efficient seismic modeling is performed with a wavefield injection hybrid approach. A first simulation is performed in a symmetric global Earth with the AxiSEM code to compute the wavefields on the borders of the lithospheric target. Then, these wavefields are propagated in the target with the spectral finite-element method. After linearization, the inverse problem is solved with a quasi-Newton (I-BFGS) local optimization algorithm.

The sensitivity of the teleseismic FWI to the experimental setup (subsurface parametrization, initial model, sampling and geometry of the station layout) is first assessed with a realistic synthetic model of the Western Alps. The main conclusions are : (i) Reliable models are obtained by updating the density and Lamé parameters (or the P and S wavespeeds) during inversion before recombining the P and S impedances for visualization and interpretation. These numerical results are consistent with the radiation pattern of the different investigated parametrizations. (ii) One-dimensional reference Earth models, such as PREM, are sufficiently-accurate to drive the inversion toward acceptable models. (iii) Coarse areal deployment of stations provide subsurface models of higher vertical and horizontal resolutions compared to dense linear deployments.

The method is finally applied to nine events of the CIFALPS experiment carried out in the Western Alps, up to a frequency of 0.2Hz. Reliable models of P and S wavespeeds and density reveal with an unprecedented resolution the crustal and lithospheric structures of the Alpine Belt, in particular the geometry of the large sedimentary basins from the Po plain and Rhone valley, the Ivrea body, and the European and Adriatic Mohos. Deeper, two slow velocity anomalies beneath the Western Alps are imaged in the mantle. The first, to the west of the chain, is interpreted as the signature of an asthenospheric upwelling, the second near the location of the Ivrea body indicates the European slab break-off. The study validates the hypothesis of the European continental crust subduction and confirms the possible tearing of the European slab.