

Thèse présentée
devant l'Université Paris XIII le 17 décembre 2009
pour obtenir
le grade de DOCTEUR DE L'UNIVERSITÉ PARIS XIII
Mention MATHÉMATIQUES APPLIQUÉES
par
Ludovic MÉTIVIER

Laboratoire d'accueil : Direction Géologie Géochimie Géophysique IFP
Ecole Doctorale : Institut Galilée Université Paris XIII

Sujet de la thèse :

Une méthode d'inversion non linéaire pour l'imagerie sismique haute résolution

Rapporteurs

Habib	AMMARI
Stéphane	OPERTO

Directrice de thèse

Laurence	HALPERN
----------	---------

Responsable scientifique IFP

Florence	DELPRAT-JANNAUD
----------	-----------------

Examinateurs

Patrick	LAILLY
Olivier	LAFITTE
Gilles	LAMBARÉ

A ma grand-mère,

Remerciements

Je remercie en premier lieu, avec chaleur, ma directrice de thèse Laurence Halpern, pour son encadrement rigoureux, et pour avoir su, à chacune de nos entrevues, m'ouvrir aux problèmes mathématiques généraux que mes préoccupations plus particulières soulevaient. J'ai pu, de ce fait, sentir la richesse des mathématiques et de leurs applications, et apprendre qu'une part primordiale du métier de chercheur consiste à toujours rester en éveil, curieux du travail des autres.

Je remercie avec la même chaleur Florence Delprat-Jannaud, qui m'a guidé avec attention tout au long de mes travaux de thèse, et m'a initié à la géophysique. Son regard de physicienne porté sur les problèmes que j'abordais sous l'angle des mathématiques et de l'informatique m'a également apporté une ouverture et une complémentarité que j'ai beaucoup appréciées.

Mes sincères remerciements vont également à Patrick Lailly, pour son implication, ses idées, ses conseils, toujours très précieux, et l'enseignement qu'il m'a donné avec ferveur de l'imagerie sismique.

Je remercie également mes rapporteurs Habib Ammari et Stéphane Operto, pour leur lecture attentive de ce manuscrit, les conseils qu'ils ont pu me donner et les discussions qui sont nées de ces lectures.

Mes remerciements s'adressent aussi à Gilles Lambaré, pour l'attention qu'il a portée à l'examen de ces travaux, et pour avoir accepté de participer au jury de soutenance.

Je remercie chaleureusement Olivier Lafitte pour les discussions que nous avons pu avoir au cours de ma thèse et ses conseils, ainsi que pour avoir accepté de présider le jury de soutenance.

Je me dois également de saluer toute l'équipe de géophysique de l'IFP, pour son accueil, sa convivialité, et sa bonne humeur, dans le désordre : Jorge, Jean-Marc, Jean-Philippe, Judith, Nicolas, Vincent, Karine et Karine, Perrine, Zhaniya, Caroline, Patrick, Nolwenn, Marc, Giselle, Christian, Michel, Kazem, Christine, Martine, Josette, Louis, Frédéric, Delphine, et bien sûr le canard.

Je salue également tous les camarades de l'université que j'ai pu rencontrer, croiser, recroiser, et connaître, au cours de mes passages le vendredi, pour les séminaires ; la liste est longue, ils se reconnaîtront, parmi eux cependant : Sabrina, Ibrahima, Benjamin et Thuy.

Je ne peux terminer sans un mot pour tous ceux qui me sont chers. Merci donc

les amis, à Paris, à Toulouse, et à Rennes, pour votre soutien constant. Et merci à toute ma famille, sur qui je sais pouvoir toujours compter, une chance dont je ne cesserai de m'émerveiller.

Une dernière et tendre pensée pour celle qui partage mes jours et mes nuits.

Table des matières

1 Introduction, contexte et problématique	11
Introduction générale	11
1.1 Inversion linéaire de données sismiques	17
1.1.1 Résolution mathématique et approximation linéaire	17
1.1.2 Limitations	19
1.1.3 Le problème inverse non linéaire	20
1.2 Inversion de données sismiques de puits 1D	22
1.2.1 Particularité de l'acquisition	22
1.2.2 Méthode mathématique de résolution	22
1.2.3 Résultats numériques	27
1.2.4 Limitations	36
Plan du manuscrit	38
2 Formulation d'un problème inverse bidimensionnel	41
Introduction	41
2.1 Description des données	43
2.2 Estimation du modèle d'impédance	45
2.2.1 Présentation du problème	45
2.2.2 Problème direct 2D	46
2.2.3 Problème inverse 2D	47
2.3 Indétermination et instabilité	48
2.3.1 Définitions	48
2.3.2 Origine de l'indétermination	49
2.3.3 Information <i>a priori</i> sur le modèle d'impédance	51
2.3.4 Information <i>a priori</i> sur les conditions de pression	52
2.3.5 Problème inverse régularisé	56
Conclusion	58
3 Résolution numérique du problème inverse	59
Introduction	59
3.1 Définition du problème numérique	61
3.1.1 Contexte et définition d'une couche PML	61
3.1.2 Définition du problème direct	63
3.1.3 Définition du problème inverse (P)	64

3.2	Méthode d'optimisation	66
3.2.1	Algorithmes de Newton et Quasi-Newton	66
3.2.2	Optimisation emboîtée	68
3.3	Calcul de $\nabla J(v)$, $R'_1(\alpha)$ et $R'_2(\alpha)$ dans un formalisme continu	78
3.3.1	Calcul du gradient de la fonctionnelle coût sismique J	78
3.3.2	Calcul du gradient des termes de régularisation	84
3.4	Discrétisation	86
3.4.1	Méthodologie de l'implémentation numérique	86
3.4.2	Notations	86
3.4.3	Problème direct discret	89
3.4.4	Problème inverse discret	92
3.4.5	Calcul du gradient de la fonctionnelle $\mathcal{J}(\alpha, h_1, \dots, h_S)$	93
3.4.6	Calcul du gradient des termes de régularisation	101
3.4.7	Expression finale du gradient discret	102
3.5	Algorithmes pour le calcul du gradient	104
3.5.1	Calcul de $\partial_{h_s} \mathcal{J}(\alpha, h_1, \dots, h_S)$	104
3.5.2	Calcul de $\partial_\alpha \mathcal{J}(\alpha, h_1, \dots, h_S)$	105
3.6	Parallélisation des algorithmes de calcul du gradient	108
3.6.1	Méthodes de décomposition de domaines	108
3.6.2	Implémentation	109
3.6.3	Utilisation de l'interface MPI	114
	Conclusion	117
3.7	Annexes	119
	Annexe 1. Introduction d'un terme de régularisation sur les CP après interpolation	119
	Annexe 2. Calcul des matrices d'interpolation B_s	121
	Annexe 3. Calcul de $\partial_\alpha \mathcal{J}(\alpha, h)$ pour une mémoire limitée	125
4	Evaluation de la méthode d'inversion	135
	Introduction	135
4.1	Description du cas test	137
4.1.1	Modèle exact et acquisition	137
4.1.2	Discrétisation	140
4.2	Scalabilité de la parallélisation par décomposition de domaine	143
4.2.1	Description du cluster de l'IFP	143
4.2.2	Tests de performance	143
4.2.3	Améliorations possibles	147
4.3	Analyse de la résolution dans un cas idéal (CP connues)	149
4.3.1	Présentation du résultat	149
4.3.2	Résolution verticale	153
4.3.3	Résolution latérale	158
4.3.4	Premières conclusions	166
4.4	Analyse de la résolution réelle (CP inconnues)	168

4.4.1	Poids de régularisation	168
4.4.2	Interpolation des CP	169
4.4.3	Modèle d'impédance <i>a priori</i> dans le puits	173
4.4.4	Présentation du résultat	174
4.4.5	Indétermination et optimisation emboîtée	182
4.4.6	Résolution verticale	186
4.4.7	Résolution latérale	189
4.4.8	Conclusion sur l'analyse de la résolution	198
4.5	Analyses de sensibilité	200
4.5.1	Analyse de sensibilité par rapport au modèle de vitesse	200
4.5.2	Analyse de sensibilité au bruit sur les données	211
	Conclusion	217
5	Analyse des équations PML	221
	Introduction	221
5.1	Analyse mathématique	224
5.1.1	Contexte et résultats	224
5.1.2	Définition du système étendu	226
5.1.3	Définition d'un symétriseur pour le système étendu	228
5.1.4	Théorème 1 : cas du demi-espace pour des coefficients d'impédance lisses	229
5.1.5	Théorème 2 : Problème de transmission pour un modèle d'impédance discontinu 1D	237
5.1.6	Conclusion de l'analyse mathématique	241
5.2	Résultats numériques	242
5.2.1	Description des expériences	242
5.2.2	Figures résultats	244
	Conclusion	256
	Conclusion et perspectives	257
	Table des figures	261
	Références	267

Chapitre 1

Introduction, contexte et problématique

Introduction générale

L'imagerie sismique est une discipline issue de la géophysique. Elle s'intéresse à la conception de méthodes servant à la détermination de la structure et la composition internes de la croûte terrestre, et plus particulièrement de ses couches les moins profondes (les investigations dépassent très rarement la dizaine de kilomètres de profondeur). Différents objectifs guident la mise au point de telles méthodes. Le premier d'entre eux, pouvant être qualifié d'objectif historique, est la localisation de ressources naturelles au sein du sous-sol, telles que le gaz, ou le pétrole. De cette activité de recherche pétrolière, dont les prémisses remontent à la seconde moitié du 19^e siècle, sont en effet nées les bases de l'imagerie sismique, sous l'impulsion de chercheurs tels que Ludger Mintrop [Mintrop, 1931]. Les premiers gisements découverts ont été repérés par des observations géologiques réalisées en surface. C'est ainsi que fut foré le premier puits de pétrole en 1859 au Texas, à quelques dizaines de mètres de profondeur. Pour l'essentiel, les réservoirs sont cependant situés dans des zones plus profondes, et il est par conséquent impossible de recueillir suffisamment d'indices en surface pour les localiser. Il est donc nécessaire de recourir aux méthodes d'imagerie sismique, donnant une information sur la structure du sous-sol, à partir de laquelle une interprétation géologique et géophysique permet d'évaluer la probabilité de la présence de pétrole et de gaz. L'imagerie sismique joue dans ce cas le rôle d'un outil de prospection. Son utilité dans l'industrie pétrolière ne se limite pas, pour autant, à cet aspect. En effet, lorsque les réservoirs sont mis en production, par l'installation de puits grâce auxquels les ressources sont extraites, l'évolution du gisement au sein du sous-sol est surveillée via des campagnes d'études sismiques, qui procurent des informations essentielles pour tenter de parvenir à une exploitation optimale de ceux-ci. L'imagerie sismique est donc indissociable du développement de l'industrie pétrolière en ce qui concerne l'exploration et la production. Plus récemment, l'éveil des consciences politiques à la nécessité d'efforts en matière de sauvegarde de l'environnement.

ronnement, et plus particulièrement dans la lutte contre le réchauffement climatique, a conduit à une recherche de procédés permettant de réduire les émissions de gaz à effet de serre tels que le CO₂. Des travaux ont été menés en ce sens sur la possibilité de capter le CO₂ à la source des émissions (centrales à charbon, industries), le compresser jusqu'à l'état liquide, pour le séquestrer dans les cavités du sous-sol. Cette chaîne de traitement requiert une expertise scientifique dans de nombreux domaines, dont celui de l'imagerie sismique. Dans ce contexte, celle-ci est mise à contribution à la fois pour la localisation de sites propices au stockage, et pour le contrôle de leur évolution, dès lorsqu'ils accueillent le CO₂. La surveillance de l'état du site de stockage est en effet un aspect crucial du problème. Un relâchement brutal de gaz, ou la contamination de nappes phréatiques, sont des risques dont il faut se prémunir, et une veille permanente doit être assurée.

Quels moyens l'imagerie sismique déploie-t-elle pour parvenir à déterminer la structure du sous-sol ? Le problème posé est celui de l'estimation de paramètres physiques non accessibles directement à la mesure, ou pour le moins très localement. Les puits de forage ou le carottage permettent en effet de posséder une information fiable de la composition du sous-sol, mais cette information n'est valable que dans la zone très restreinte du forage. Ceux-ci peuvent atteindre plusieurs kilomètres de profondeur, mais ne donnent qu'une information locale au niveau du puits. Pour disposer d'une information à l'échelle de plusieurs kilomètres carrés en surface, et plusieurs kilomètres de profondeur, le géophysicien emploie une mesure indirecte de ces paramètres par le biais d'expériences dites de "sismique réflexion", dont une description sommaire est donnée ici.

Le sous-sol terrestre peut être assimilé à un milieu de propagation d'ondes plus ou moins complexe : acoustique (seules les ondes de compression sont prises en compte), élastique (les ondes de compression et les ondes transverses, ou de cisaillement, sont prises en compte), élastique avec divers degrés d'anisotropie, poroélastique avec effets d'anisotropie. La composition du sous-sol influe directement sur la manière dont ces ondes se propagent en son sein. Ainsi, à l'inverse, une étude de leur propagation peut fournir d'importantes indications sur sa composition. Cette étude est menée de la manière suivante. En surface, une impulsion est générée par des camions vibrants ou des explosifs en sismique terrestre, ou des pistolets à air comprimés en sismique marine. Cette impulsion provoque la propagation d'un champ d'ondes dans le sous-sol. Ces ondes subissent divers phénomènes liés à la propagation tels que des réflexions, réfractions, diffractions, transmissions, atténuations, ou conversions, provoqués par les discontinuités des paramètres mécaniques du sous-sol. Une partie du signal émis par les sources est donc réfléchi vers la surface. Des récepteurs y sont disposés et procèdent à son enregistrement. Cette acquisition est ensuite répétée pour différentes positions de la source et des récepteurs. Cette succession d'enregistrements constitue les données à partir desquelles une estimation de la structure du sous-sol est recherchée. L'imagerie sismique cherche donc à recueillir des informations sur un

milieu de propagation (le sous-sol), à partir de l'enregistrement de la réponse de ce milieu à une excitation en surface. Par là, elle s'inscrit dans la catégorie très vaste des problèmes inverses, que l'on retrouve dans de nombreux domaines : gravimétrie, imagerie médicale, météorologie.

Différentes méthodes d'imagerie existent, plus ou moins complexes, et plus ou moins faciles à mettre en oeuvre, qui reposent toutes sur l'interprétation de données sismiques dont l'acquisition vient d'être décrite. Celles-ci peuvent donner une première idée de la structure du sous-sol. Pour se forger une intuition, il est intéressant de réfléchir au cas simple d'un milieu de propagation à deux couches homogènes, présentant une discontinuité à leur interface. En première approximation, la propagation des ondes dans le sous-sol suit les lois de l'optique géométrique. Ainsi, pour une source située en surface entourée de plusieurs récepteurs, et une discontinuité plane (parallèle à la surface), le signal en temps enregistré par chacun des récepteurs (ou trace sismique) est constitué de la réflexion du signal émis par la source, au niveau du point milieu, c'est-à-dire le point de la discontinuité situé à égale distance de la source et du récepteur (fig. 1.1).

La représentation conventionnelle des données sismiques consiste à disposer les traces sismiques les unes à côté des autres, l'échelle verticale représentant le temps d'enregistrement, tandis que l'échelle horizontale représente la distance du récepteur par rapport à la source, ou "offset". Pour l'expérience décrite plus haut, une telle représentation conduit à la visualisation d'une figure de type hyperbolique. Une telle figure contient en elle-même une première forme d'information, témoignant d'un important contraste dans la structure du sous-sol, à l'origine de l'enregistrement d'une réflexion conséquente. Cependant, la visualisation des données brutes ne procure pas d'information sur la géométrie ou sur l'amplitude du contraste. Les méthodes d'imagerie sismique s'attachent à cette reconstruction.

La première grande famille de méthodes d'imagerie sismique est celle des méthodes de migration. Elle permettent de transformer un jeu de données sismiques en une représentation spatiale du sous-sol, qui restitue la profondeur et la géométrie des discontinuités. Pour cela, une estimation de la vitesse de propagation des ondes est réalisée. De nombreuses méthodes existent pour déterminer une telle estimation, les plus simples étant basées sur l'analyse des temps d'arrivée des ondes principales, à partir des données sismiques. Cette estimation du champ de vitesse permet de procéder à la "migration" des données, en une image en profondeur du sous-sol. Dans l'exemple précédent, ceci correspond à l'horizontalisation de la figure hyperbolique observée sur les données sismiques, qui permet de retrouver la structure horizontale de la discontinuité, et sa profondeur. De manière générale, la famille des méthodes de migration regroupe tous les procédés permettant de passer d'une

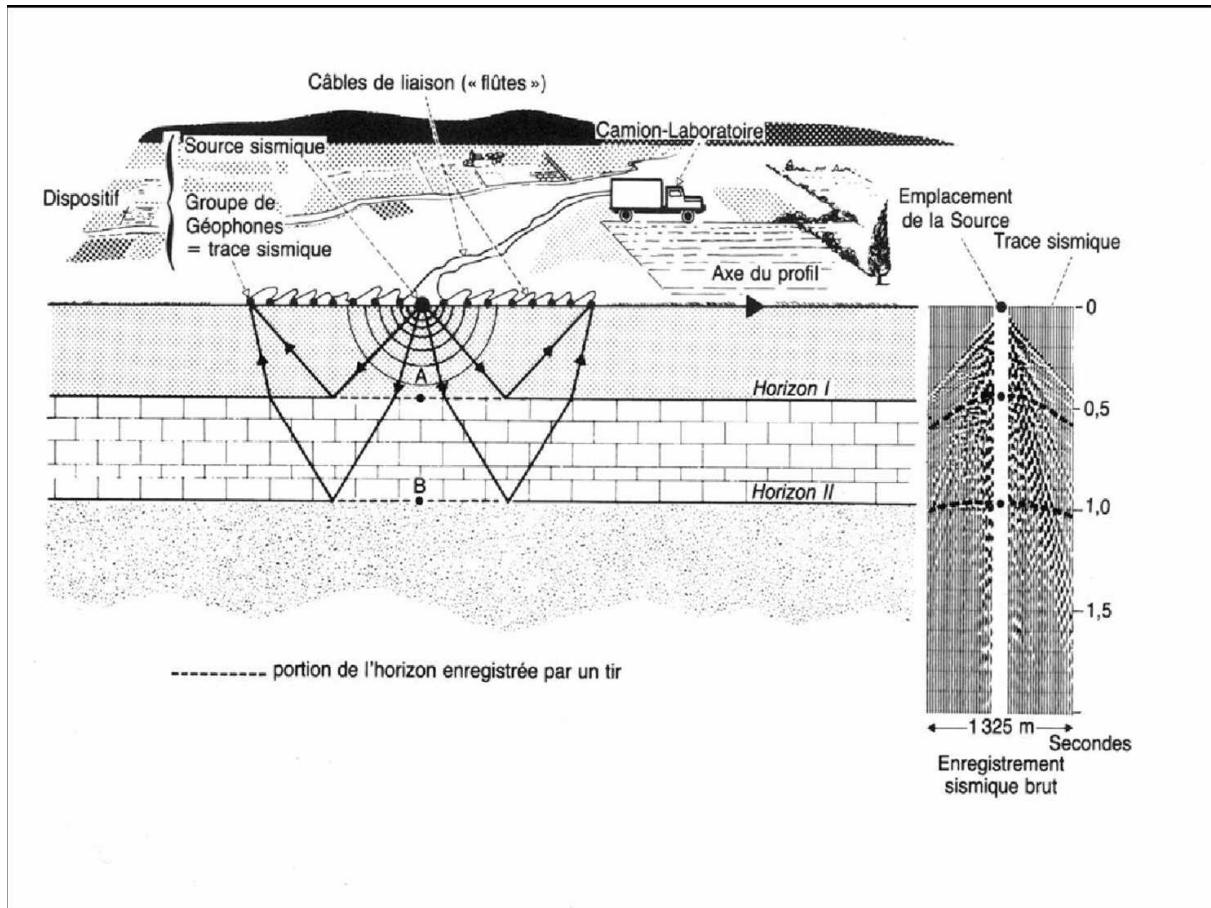


FIG. 1.1 – Principe d'une expérience de sismique réflexion.

image en temps de parcours à une image en profondeur du sous-sol. L'avantage de ces méthodes est la simplicité de leur mise en oeuvre, et l'expérience acquise au cours du temps sur leur utilisation. Ce sont en effet des méthodes généralement bien éprouvées, dont les comportements sont connus. Une description plus détaillée des méthodes de migration est donnée dans [Claerbout, 1976] [Yilmaz, 1987]. Leur principale limitation consiste en la construction d'une image purement structurale, ne donnant qu'une estimation de la géométrie du sous-sol. Un géophysicien qui ambitionne de déterminer une estimation quantitative de paramètres physiques décrivant la structure et la composition du sous-sol doit se tourner vers des méthodes plus sophistiquées, appelées méthodes d'inversion.

Ces méthodes d'inversion font d'abord appel à une modélisation de la propagation des ondes dans le sous-sol, par la définition de ce qui est appelé le problème direct. Celui-ci correspond au calcul du signal enregistré par les récepteurs, connaissant la position et la nature de la source s , appartenant à l'espace des sources \mathbb{S} , la position des récepteurs, et la structure du sous-sol m , appartenant à l'espace des modèles \mathbb{M} . Cette modélisation de la propagation peut-être plus ou moins précise,

suivant les phénomènes physiques qui désirent être pris en compte. Ces modèles sont décrits par des équations de propagation, qui permettent de définir un opérateur de modélisation $\mathcal{M}(m, s)$. Celui-ci associe aux paramètres du sous-sol $m(x, y, z)$, et à une source $s(x, y, z, t)$, la réponse enregistrée aux récepteurs, définie dans l'espace des données \mathbb{D} .

$$\mathcal{M} : (m, s) \longrightarrow \mathcal{M}(m, s),$$

$$\mathbb{M} \times \mathbb{S} \longrightarrow \mathbb{D}.$$

Cet outil de modélisation peut être utilisé pour déterminer une estimation de paramètres mécaniques du sous-sol $m(x, y, z)$ de manière simple, au sein d'une procédure de type essai-erreur. En effet, à partir des connaissances acquises sur le terrain par les géologues, et leur corrélation avec l'information locale obtenue au niveau des forages, il est possible de bâtir une estimation *a priori* de ces paramètres. Les données sismiques synthétiques associées à ce modèle, étant donnée une estimation de la source $s(x, y, z, t)$, par l'entremise du problème direct, sont ensuite comparées au données réelles. Cette comparaison induit un critère qui offre une possibilité de contrôle sur la validité du modèle qui a été construit. Cette méthode n'entre pas véritablement dans la catégorie des méthodes d'inversion, néanmoins, elle est souvent utilisée par les gens de métier, leur permettant d'inclure l'information sismique dans leur recherche d'un modèle de sous-sol .

Une première approche de la définition d'une méthode d'inversion consiste en la construction de l'opérateur inverse de l'opérateur de modélisation. Cet opérateur associe à un jeu de données sismiques la distribution de paramètres du sous-sol produisant ces données. Dans des contextes de modélisation 1D, et pour des données non bruitées, la théorie permet de calculer cet opérateur inverse [Sondhi, Gopinath, 1971], [Bamberger et al., 1977], plus exactement de définir les sous-espaces dans l'espace des données et l'espace des modèles qui permettent de définir un opérateur inverse continu. Cependant, dans la pratique, les données sont bruitées, et la modélisation prend en compte des contextes multi-dimensionnels. Le problème inverse se voit donc posé comme un problème de minimisation de la distance, au sens d'une certaine norme $\|\cdot\|$, entre les données synthétiques $\mathcal{M}(m, s)$, et les données réelles $D(t)$, par rapport aux paramètres du sous-sol,

$$\min_m \|D - \mathcal{M}(m, s)\|^2, \quad (1.1)$$

une estimation de la source $s(x, y, z, t)$ étant effectuée au préalable. Ce problème est résolu par une méthode d'optimisation appropriée. Ceci peut-être vu comme l'automatisation de la procédure d'essai-erreur, les améliorations à apporter au modèle de sous-sol étant calculées par un processus mathématique tirant profit de l'information contenu dans le gradient de l'opérateur de modélisation. La norme la plus

souvent utilisée est la norme L^2 . Cependant, dans une interprétation statistique, ceci suppose que le bruit enregistré sur les données possède une distribution gaussienne, ce qui n'est pas toujours adéquat. D'autres normes peuvent ainsi être utilisées (norme L^1 , norme L^∞ , et plus généralement les normes L^p [Tarantola, 2005]).

Un autre problème se pose également, celui de l'unicité de la solution. Dans le cadre d'un opérateur de modélisation $\mathcal{M}(m, s)$ linéaire, la fonctionnelle à minimiser, qui mesure la distance entre les données, est convexe, et le minimum unique. Néanmoins, une modélisation précise de la propagation des ondes induit une dépendance non linéaire de l'opérateur de modélisation par rapport aux paramètres du sous-sol $m(x, y, z)$, ce qui engendre le plus souvent la définition d'une fonctionnelle non convexe. Pour cette raison, les méthodes de résolution du problème d'inversion de données sismiques procèdent dans un premier temps à une linéarisation de l'opérateur de modélisation $\mathcal{M}(m, s)$, autour d'un modèle de référence $m_0(x, y, z)$. Cette procédure de linéarisation est décrite de manière détaillée dans la partie suivante. Ces méthodes d'inversion permettent d'obtenir à la fois une information sur la structure géométrique du sous-sol, et une information quantitative sur les valeurs du modèle $m(x, y, z)$ estimé, et sont très utilisées aujourd'hui. Néanmoins, elles souffrent de certaines limitations, notamment en ce qui concerne la résolution de l'image du sous-sol qu'elles procurent, limitée à plusieurs dizaines de mètres, et l'imagerie de milieux faisant apparaître de très fortes hétérogénéités.

1.1 Inversion linéaire de données sismiques

1.1.1 Résolution mathématique et approximation linéaire

Nous revenons dans cette partie sur la définition des méthodes d'inversion de données sismiques classiques, posées sous la forme du problème de minimisation

$$\min_m \|D - \mathcal{M}(m, s)\|^2. \quad (1.2)$$

une estimation de la source $s(x, y, z, t)$ étant effectuée au préalable.

Ce problème est un problème inverse intrinsèquement non linéaire, dont la résolution est par conséquent difficile. En effet, une modélisation précise de la propagation des ondes dans le sous-sol implique la non-linéarité de l'opérateur de modélisation $\mathcal{M}(m, s)$ par rapport aux paramètres du sous-sol $m(x, y, z)$. Dans la pratique, pour parvenir à des méthodes de résolution plus simples à mettre en oeuvre, ce problème est linéarisé. Soit un modèle de référence $m_0(x, y, z)$ du modèle de sous-sol recherché $m(x, y, z)$, par exemple une approximation basse fréquence de ce modèle. On s'intéresse à la reconstruction de la perturbation $dm(x, y, z)$, telle que

$$m = m_0 + dm.$$

Considérons le développement de l'opérateur \mathcal{M} autour du modèle de référence $m_0(x, y, z)$. Ce développement s'écrit

$$\mathcal{M}(m, s) = \mathcal{M}(m_0, s) + \partial_m \mathcal{M}(m_0, s).dm + o(dm).$$

Quel est le sens physique de ce développement de Taylor associé à l'opérateur de modélisation $\mathcal{M}(m, s)$? Bleistein [Bleistein et al., 2000] propose l'interprétation suivante. Le milieu de propagation est supposé pouvoir correspondre à la somme d'une partie lisse, et connue, notée $m_0(x, y, z)$ et d'une perturbation haute fréquence de ce milieu de référence, notée $dm(x, y, z)$. Le champ d'onde qui se propage dans ce milieu, se décompose pour sa part en une partie qui dépend linéairement de la perturbation $dm(x, y, z)$, $\mathcal{M}'(m_0)dm$, et une partie qui en dépend non linéairement, $\mathcal{M}(m_0) + o(dm)$. La partie qui dépend linéairement de la perturbation peut-être interprétée physiquement comme la partie correspondant aux réflexions primaires sur les discontinuités du sous-sol [Bleistein et al., 2000]. Ces premières réflexions constituent le plus souvent la majeure partie du signal enregistré, les réflexions multiples enregistrées par la suite étant d'intensité plus faible. La définition du problème inverse linéaire découle de l'approximation au premier ordre consistant à considérer ces phénomènes non-linéaires de réflexions multiples négligeables devant les réflexions primaires. Ceci mène à la définition du problème inverse linéarisé

$$\min_{dm} \|D - \partial_m \mathcal{M}(m_0, s)dm\|^2. \quad (1.3)$$

Dans la pratique, ce problème, avant d'être résolu, nécessite deux étapes préalables. La première consiste, comme cela a déjà été évoqué, en l'estimation du signal émis

par les sources $s(x, y, z, t) \in \mathbb{S}$ ayant servi à enregistrer les données . Cette estimation est en effet nécessaire à la définition de l'opérateur de modélisation. Cette opération rentre dans le processus classique d'inversion linéarisée et est globalement bien maîtrisée.

La seconde étape consiste en la définition de l'approximation basse fréquence $m_0(x, y, z)$ du modèle de sous-sol $m(x, y, z)$. Celle-ci est établie à partir de différentes sources d'information. D'une part, les géologues sont capables de déterminer, d'après des études de terrain, les variations à très grandes échelles du sous-sol. D'autre part, des relevés stratigraphiques effectués dans des puits donnent une information locale très précise. La détermination du modèle *a priori* $m_0(x, y, z)$ découle de la combinaison de ces informations, qui pose des problèmes importants de mise à l'échelle, la précision de l'information acquise au niveau des puits pouvant être inférieure au mètre, tandis que les variations données par les géologues s'accordent sur des distances de l'ordre du kilomètre.

L'estimation des sources permet donc de définir le problème direct, puis celle du modèle *a priori* $m_0(x, y, z)$ permet de définir le problème inverse linéarisé (1.3), qui peut ensuite être résolu numériquement. Celle-ci fait intervenir la résolution du système linéaire

$$(\partial_m \mathcal{M}(m_0, s))^T \partial_m \mathcal{M}(m_0, s) dm = (\partial_m \mathcal{M}(m_0, s))^T D, \quad (1.4)$$

équivalent au problème (1.3). Suivant la complexité de l'opérateur de modélisation, ce système peut-être plus ou moins difficile à résoudre. Certaines techniques, appartenant à la famille des méthodes de migration, procèdent simplement à l'évaluation du second membre

$$(\partial_m \mathcal{M}(m_0, s))^T D,$$

plutôt qu'à la résolution du système linéaire.

L'application aux données de l'opérateur adjoint de modélisation permet en effet d'effectuer une conversion des données dans ce qui est qualifié le domaine profondeur. Cette technique permet de fournir, à moindre coût, une première image de la structure du sous-sol. Cependant, en tant que méthode de migration, ceci n'autorise aucune estimation quantitative de la composante haute fréquence du paramètre de sous-sol $dm(x, y, z)$. Les thèses de [Duquet, 1996] et [Pion, 2007] montrent que la résolution du système linéaire permet de fournir une telle estimation, tout en améliorant la qualité de l'image du sous-sol qui est obtenue en terme de précision.

1.1.2 Limitations

Malgré des résultats probants dans certains cas d'application, cette méthode d'inversion linéarisée présente des inconvénients, et peut rencontrer des difficultés.

Tout d'abord, comme cela a été évoqué, elle nécessite la détermination d'un modèle de sous-sol *a priori* $m_0(x, y, z)$. Celui-ci doit rendre suffisamment compte de la réalité du terrain pour que la méthode fonctionne. Cependant, cette détermination se révèle souvent malaisée. Le travail des géologues doit être mené avec attention pour relever les profils des couches sédimentaires, et les éventuelles failles qui les traversent. De plus, les informations disponibles au niveau des puits étant très localisées, elles ne permettent pas toujours de corroborer les hypothèses émises quant à la constitution du milieu. Une partie très importante du travail d'inversion consiste donc en la détermination du meilleur modèle *a priori* $m_0(x, y, z)$ possible. La détermination de ce modèle *a priori* est d'autant plus importante que la solution du problème inverse linéarisé dépend très fortement de ce modèle. Dans la chaîne de travail classique de l'inversion linéaire de données sismiques de surface, cette étape est donc particulièrement cruciale.

Une seconde limitation, également très importante, provient du fait de l'hypothèse correspondant à la linéarisation du problème inverse. En effet, l'amplitude de la partie du signal qui dépend non linéairement de la perturbation $dm(x, y, z)$ recherchée est considérée comme négligeable devant l'amplitude des réflexions primaires. Cependant, plus le sous-sol présente d'hétérogénéités, moins cette hypothèse est vérifiée. Dans le cas de milieux très hétérogènes, la part due aux réflexions multiples devient fortement non négligeable [Delprat-Jannaud, Lailly, 2004, 2008]. La méthode d'inversion linéaire se trouve alors impuissante à expliquer les données, et le modèle de perturbation $dm(x, y, z)$ calculé ne correspond pas aux discontinuités du sous-sol.

La dernière limitation de la méthode est liée à la résolution maximale qu'il est possible d'espérer pour une telle méthode d'imagerie sismique. Le critère de Rayleigh employé dans le cadre de la propagation d'ondes sismiques dans le sous-sol, permet d'établir une rapide estimation de cette résolution. Il indique que deux discontinuités du sous-sol peuvent être dissociées par la méthode d'imagerie si la distance qui les sépare est de l'ordre du quart de la longueur d'onde moyenne. Dans le contexte de la sismique, les sources émettent des signaux centrés autour d'une fréquence comprise entre 20 Hz et 30 Hz. De plus, la vitesse de propagation moyenne typique dans la croûte supérieure est située autour de 3000 m.s^{-1} . La longueur d'onde moyenne prise en compte par une méthode inverse linéarisée est donc comprise entre 100 m et 150 m, ce qui conduit à une résolution moyenne comprise entre 25 m et 40 m. Des travaux plus détaillés sur ce sujet permettent d'obtenir des estimations plus fines et adaptées au contexte d'utilisation [Garnier, Papanicolaou, 2009]. Cependant, cette première approche par l'utilisation du critère de Rayleigh permet déjà de souligner

qu'une telle méthode ne peut fournir d'information sur les variations du sous-sol à une échelle inférieure à la vingtaine de mètres.

Ces limitations conduisent à s'intéresser à une méthodologie permettant d'imager des milieux présentant de fortes hétérogénéités. Ceci requiert l'abandon de l'hypothèse de linéarisation, et c'est pourquoi l'on s'intéresse à la résolution du problème inverse non linéaire.

1.1.3 Le problème inverse non linéaire

Quels sont les bénéfices que la résolution du problème inverse non linéaire peut apporter ? Tout d'abord, la résolution de ce problème permet la prise en compte des phénomènes non linéaires de la propagation des ondes, dont la présence de réflexions multiples. Ceci permet de surmonter la deuxième des limitations présentées dans le paragraphe précédent. De plus, une remarque de [Delprat-Jannaud, Lailly, 2004], formulée à partir des observations de [Fink, 1997], indique que cette prise en compte des réflexions multiples laisse espérer une amélioration importante de la résolution de la méthode d'imagerie, ce qui permettrait de s'affranchir de la dernière des limitations évoquées pour le problème inverse linéarisé.

Cependant, la résolution d'un tel problème pose de nombreuses difficultés. Tout d'abord, à la différence du problème inverse linéarisé qui consiste en la minimisation d'une fonctionnelle quadratique, possédant par conséquent un minimum global, la minimisation d'une fonctionnelle non linéaire pose le problème de l'existence de minima locaux. L'unicité de la solution est donc sérieusement remise en cause, ce qui laisse surgir toutes sortes de problèmes liés au caractère mal posé du problème inverse non linéaire, telle l'indétermination. D'autre part, d'un point de vue numérique, la résolution d'un problème d'optimisation non linéaire requiert l'utilisation d'algorithmes plus complexes, et plus coûteux en temps de calcul. Il en va de même pour la modélisation de la propagation des ondes. Tandis que la modélisation de la propagation limitée aux réflexions primaires peut être avantageusement calculée par des méthodes convolutives, la modélisation incluant les non linéarités de la propagation requiert la résolution de l'équation des ondes dans un domaine borné, par des méthodes de discréétisation (éléments finis, différences finies) plus coûteuses. Les moyens de calcul à mettre en oeuvre sont dès lors beaucoup plus conséquents.

D'autre part, une nouvelle difficulté se pose sur le plan théorique. En effet, la détermination d'une solution du problème (1.2) est extrêmement sensible à la donnée *a priori* du signal émis par les sources. Une très faible perturbation commise sur la source, se répercute au sein de la procédure d'inversion sur le modèle de sous-sol déterminé $m(x, y, z)$ par une accumulation d'erreurs, qui deviennent très importantes pour la partie profonde de celui-ci [Delprat-Jannaud, Lailly, 2005]. Les technologies actuelles de calibration des sources ne permettent pas d'obtenir une

précision suffisante sur leur détermination pour pouvoir les employer dans le cadre d'une inversion non linéaire. La prise en compte des non-linéarités de la propagation doit donc se faire d'une manière alternative. Pour cela, il est intéressant de s'intéresser à une méthode particulière d'imagerie sismique, l'inversion de données sismiques de puits. Elle permet, comme cela est expliqué dans la partie suivante, la prise en compte du problème inverse non linéaire.

1.2 Inversion de données sismiques de puits 1D

L'inversion de données sismiques de puits (PSV, pour profil sismique vertical) a été étudiée dès le début des années 80, dans le cadre d'une modélisation du sous-sol et de la propagation des ondes monodimensionnelle [Macé, Lailly, 1986] [Bamberger, 1978]. Par la configuration particulière de l'acquisition, cette méthode permet de procéder à la résolution du problème inverse non linéaire correspondant. A la différence de l'inversion de données sismiques de surface, qui permet de construire une image du sous-sol sur des étendues de l'ordre du km en surface, et jusqu'à presque 10 km en profondeur, l'inversion de données sismiques de puits procure une image du sous-sol en profondeur, mais entièrement localisée autour du puits. L'information obtenue est en revanche beaucoup plus précise que celle obtenue par l'inversion de données sismiques de surface. Cette partie est consacrée à la description de cette méthode, des avantages qu'elle procure, et des résultats numériques obtenus dans le cadre d'une modélisation 1D acoustique. L'effet haute résolution attendu lorsque la méthode est utilisée pour l'imagerie de milieux très hétérogènes est mis en avant.

1.2.1 Particularité de l'acquisition

L'acquisition de données sismiques de puits se fait en disposant verticalement les récepteurs chargés d'enregistrer les données dans un puits, à la différence des méthodes conventionnelles, pour lesquelles ceux-ci sont placés en surface. Une source est ensuite placée très proche de la tête du puits, et l'ébranlement suscité par celle-ci est enregistré par les récepteurs. Ce dispositif particulier permet, à partir des données, de reconstruire, en plus du modèle de sous-sol recherché, l'amplitude de la condition de pression à la profondeur du premier récepteur au cours du temps d'enregistrement. Ce signal, noté par la suite $h(t)$, permet de s'affranchir de la connaissance *a priori* de la source. La prise en compte des non-linéarités dans l'inversion est donc rendue possible, les problèmes de sensibilité de la solution à une erreur sur la source ne se posant plus. La partie suivante présente le modèle mathématique utilisé. L'opérateur de modélisation est d'abord introduit via la définition du problème direct, ce qui permet d'aborder ensuite la définition du problème inverse.

1.2.2 Méthode mathématique de résolution

1.2.2.1 Problème direct

Le contexte d'application retenu est celui de la modélisation acoustique 1D. Les récepteurs, au nombre de Q , sont disposés verticalement aux profondeurs

$$(z_r)_q, \quad q = 1, \dots, Q.$$

Le domaine dans lequel se propage les ondes est

$$\Omega = [(z_r)_1, z_{max}].$$

Ainsi, la zone située entre la surface et le premier récepteur n'est pas prise en compte. De plus, la profondeur z_{max} est choisie telle que

$$z_{max} > (z_r)_Q.$$

Le problème direct modélise donc la propagation des onde à l'aplomb du puits, à partir de la profondeur du premier récepteur $(z_r)_1$. De plus, la profondeur z_{max} dépend du temps d'enregistrement T qui définit la durée de l'acquisition.

Soit $p(z, t)$ le champ de pression défini sur $\Omega \times [0, T]$. Soit $I(z)$ et $c(z)$, respectivement l'impédance acoustique et la vitesse du milieu, définies sur Ω . Ces deux quantités décrivent la propagation des ondes dans Ω , régie par les équations

$$\frac{1}{Ic} \partial_{tt} p - \partial_z \left(\frac{c}{I} \partial_z p \right) = 0, \quad \text{sur } \Omega \times [0, T], \quad (1.5)$$

avec les conditions au bord

$$p((z_r)_1, t) = h(t), \quad p(z_{max}, t) = 0, \quad (1.6)$$

et des conditions initiales nulles. Physiquement, ceci correspond à la propagation d'une onde plane à incidence normale dans le sous-sol, dans un milieu invariant horizontalement.

L'impédance acoustique est la racine du produit du module d'élasticité isostatique $\kappa(z)$ ("bulk modulus" en anglais) et de la densité volumique du matériau $\rho(z)$,

$$I(z) = \sqrt{(\rho(z)\kappa(z))}.$$

Cette quantité caractérise la résistance que le milieu oppose au déplacement suscité par le passage d'une onde de compression. Sa reconstruction permet aux géophysiciens intéressés par la caractérisation du milieu d'acquérir une information structurale précieuse. Elle permet également de fournir des informations complémentaires à partir desquelles des propriétés plus spécifiques, telles que la perméabilité et la porosité des roches, peuvent être établies. C'est la distribution d'impédance du milieu qui va être calculée à partir des données sismiques.

Pour ce faire, en l'état, ceci suppose que la distribution de vitesse soit connue. Le cadre particulier de la 1D permet de s'affranchir de ce pré-requis, en autorisant un changement de variable temps-profondeur. Soit en effet τ la variable de temps de parcours, telle que

$$d\tau = \frac{dz}{c}, \quad \tau(z) = \int_0^z \frac{ds}{c(s)}.$$

Alors, les dérivées partielles suivant z sont remplacées par des dérivées partielles suivant τ en utilisant la relation

$$\partial_z = \frac{1}{c} \partial_\tau.$$

Les temps de parcours correspondants à la profondeur des récepteurs sont notés

$$(\tau_r)_q = \tau((z_r)_q), \quad q = 1, \dots, Q,$$

et

$$\tau_{max} = \tau(z_{max}).$$

Soit alors Ω_τ le domaine en temps de parcours,

$$\Omega_\tau = [(\tau_r)_1, \tau_{max}].$$

Le champ de pression $p(\tau, t)$ en temps de parcours est donné par les équations

$$\frac{1}{I} \partial_{tt} p - \partial_\tau \left(\frac{1}{I} \partial_\tau p \right) = 0, \quad \text{sur } \Omega_\tau \times [0, T]. \quad (1.7)$$

avec les conditions au bord

$$p((\tau_r)_1, t) = h(t), \quad p(\tau_{max}, t) = 0, \quad (1.8)$$

et des conditions initiales nulles.

Ainsi, la conversion temps-profondeur permet de s'affranchir de la connaissance du modèle de vitesse. Etant donnés une condition de pression $h(t)$ à la profondeur du premier récepteur et un modèle d'impédance $I(\tau)$ décrit dans le domaine Ω_τ , il est donc possible de calculer le champ de pression dans tout le domaine. Cependant le signal enregistré par les récepteurs est la vitesse de déplacement u_z , et non pas directement la pression. Ces deux quantités sont liées par la relation

$$\partial_t u_z = -\frac{1}{I} \partial_\tau p. \quad (1.9)$$

Cette dernière relation, ajoutée aux précédentes, permet de définir l'opérateur de modélisation $\mathcal{M}(I, h)$, tel que

$$\mathcal{M}(I, h) : (I, h) \longrightarrow (u_z(\tau_1, t), \dots, u_z(\tau_Q, t)), \quad (1.10)$$

où u_z est calculé à partir des équations (1.7), (1.8) et (1.9).

Remarque 1. La condition de Dirichlet homogène en $\tau = \tau_{max}$ induit une réflexion totale sur le fond du modèle qui n'a pas d'origine physique. Cependant, les faibles coûts de calcul liés au cas 1D rendent possible la solution qui consiste à allonger le

modèle en profondeur de manière artificielle, de sorte que durant le temps de propagation considéré, les ondes réfléchies sur le fond du modèle n'aient pas le temps d'être enregistrées par les récepteurs.

Remarque 2. *Le choix de résoudre ce problème de propagation d'ondes acoustiques 1D par la résolution de l'équation des ondes n'est pas judicieux en termes de temps de calcul. En effet, dans les applications développées par [Macé, Lailly, 1986], ce problème est résolu par la méthode des caractéristiques, qui se trouve être beaucoup plus efficace. Cependant, cette méthode est restreinte à la résolution de problèmes 1D. Aussi, la résolution de l'équation des ondes pour le problème direct 1D permet de conférer à celui-ci une portée beaucoup plus générale, en offrant la possibilité d'une transposition à un contexte multidimensionnel.*

Il est alors possible de définir le problème inverse de données sismiques de puits 1D.

1.2.2.2 Problème inverse

Le problème inverse de données sismiques de puits 1D consiste à calculer, à partir des données sismiques enregistrées $D(t)$, un profil d'impédance $I(\tau)$, et une condition de pression à la profondeur du premier récepteur $h(t)$, tels que le simsogramme synthétique leur correspondant par le problème direct soit aussi proche que possible des données, au sens d'une norme définie dans l'espace des données. Ceci se traduit mathématiquement par la définition du problème (P_{1D})

$$(P_{1D}) \quad \min_{I,h} J(I, h),$$

où

$$J(I, h) = \|D - \mathcal{M}(I, h)\|_{(L^2(0,T))^Q}^2.$$

Le choix de la norme utilisée dans l'espace des données est d'une importance cruciale pour la recherche de la solution. Dans les cas d'inversion de données réelles, il faut en effet prendre en compte la problématique du bruit, présent sur les données, et à toutes les étapes de la modélisation. La prise en compte de ce bruit amène à hiérarchiser les données suivant la confiance qu'il est possible de leur accorder. Cette hiérarchie est prise en compte dans l'inversion par la définition de matrices de covariance [Tarantola, 2005]. Ces matrices de covariance permettent de définir des normes particulières, au sens desquelles le problème inverse doit être résolu, afin de pénaliser l'information à laquelle le moins de crédit est porté, et favoriser celle qui semble la plus fiable. Cette problématique n'est pas abordée dans le cadre de cette thèse, l'essentiel du problème étant focalisé sur la définition d'une méthodologie

d'imagerie sismique non linéaire multi-dimensionnelle, qui est testée dans un contexte de données synthétiques.

D'un point de vue purement numérique, le problème inverse est donc ramené à un problème d'optimisation non linéaire difficile et coûteux en temps de calcul, l'évaluation de la fonctionnelle $J(I, h)$ faisant intervenir la résolution d'un problème de propagation d'ondes. La méthode de résolution numérique utilisée pour résoudre le problème (P_{1D}) est une version simplifiée de celle qui est utilisée pour résoudre le problème inverse 2D de données sismiques de puits, méthode dont une description exhaustive est donnée dans le chapitre 3. Aussi, on se contentera ici de décrire dans les grandes lignes la résolution numérique du problème (P_{1D}).

Tout d'abord, l'espace des modèles \mathbb{M} et l'espace des données \mathbb{D} sont discrétisés. Ceci permet l'utilisation du schéma de Yee centré d'ordre 2 pour la discrétisation des équations (1.7), (1.8) et (1.9) [Yee, 1966]. De ceci découle la définition d'un opérateur de modélisation discret $\mathcal{M}_d(I, h)$, qui sert lui même à la définition de la fonctionnelle discrète $J_d(I, h)$, telle que

$$J_d(I, h) = \|D - \mathcal{M}_d(I, h)\|^2,$$

où $\|\cdot\|$ désigne une norme dans l'espace discrétisé des données (les données D discrètes sont assimilées à leurs homologues continues).

La minimisation de la fonctionnelle s'effectue ensuite par l'utilisation d'un l'algorithme Quasi-Newton, utilisant une formule de mise à jour de la Hessienne de type BFGS (des noms de ses inventeurs Broyden, Fletcher, Goldfarb et Shanno) [Shanno, 1970]. Les lecteurs intéressés par la problématique des algorithmes d'optimisation peuvent se rapporter au chapitre 2 où un rappel des méthodes de Newton est effectué, et peuvent également consulter [Gilbert and al., 2006].

L'utilisation de cet algorithme d'optimisation nécessite le calcul du gradient de la fonctionnelle discrète $\nabla J_d(I, h)$, qui est effectué par la méthode de l'état adjoint. Cette méthode permet de définir un système d'équations correspondant à un problème de propagation d'onde rétrograde en temps. Ce système est le système adjoint des équations (1.7) et (1.8). Le gradient de la fonctionnelle discrète $\nabla J_d(I, h)$ se détermine à partir de la solution de ces équations et de leurs adjointes.

Cette méthode numérique permet de résoudre efficacement le problème (P_{1D}). L'intérêt de la méthode est la prise en compte des réflexions multiples au sein de l'inversion, qui, dans le cas de milieux fortement hétérogènes, ne peuvent plus être négligées. De plus, cette prise en compte résulte en une nette amélioration de la résolution du modèle d'impédance calculé. Ceci est au cœur de la partie suivante, qui présente les résultats numériques obtenus.

1.2.3 Résultats numériques

Les résultats présentés ici sont des résultats d'inversion sur données synthétiques. Le modèle d'impédance exact est issu de données réelles (fig. 1.2), et est choisi, conformément à ce qui a été évoqué précédemment, très hétérogène. La source employée pour générer les données est un signal centré sur une fréquence de 30 Hz, d'usage courant en géophysique, appelé signal de Ricker (fig.1.3). Le choix est fait de travailler sur un jeu de données synthétiques non bruitées, afin de se placer dans un cadre idéal dans lequel l'inversion peut se faire à un degré de précision très élevé. Ceci permet d'illustrer le potentiel maximal de la méthode en termes de résolution de l'image d'impédance construite.

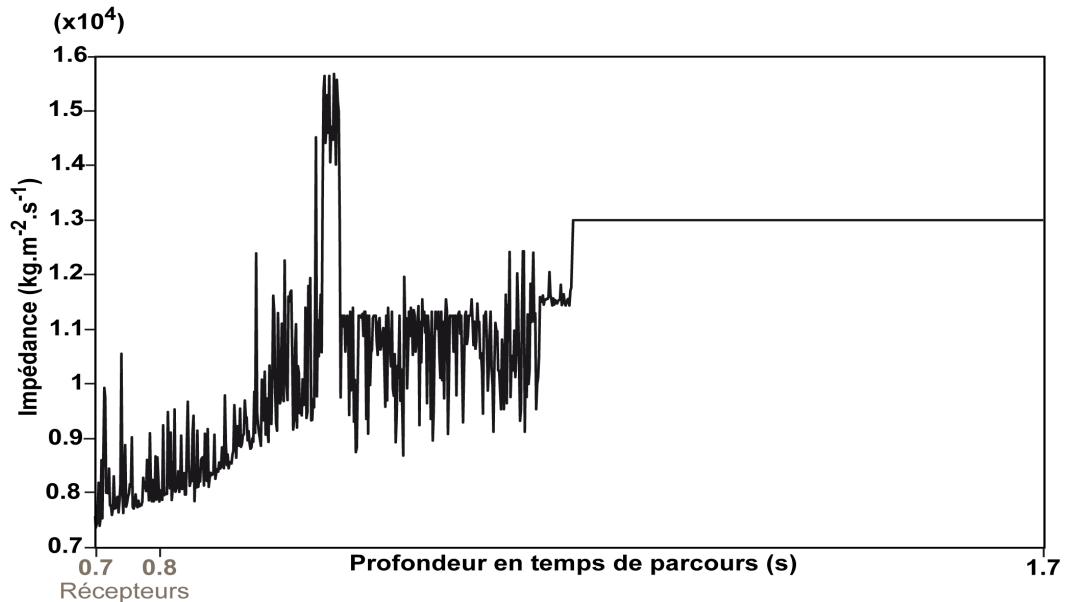


FIG. 1.2 – Modèle d'impédance exact.

Pour construire le jeu de données synthétiques, la source est placée à une profondeur de 500 ms. La vitesse de déplacement est enregistrée au niveau de 100 récepteurs disposés à la verticale de la source, à intervalle régulier (toutes les millisecondes), à partir de la profondeur de 700 ms. Les pas d'échantillonage en temps Δt et en temps de parcours $\Delta\tau$ valent tous deux 1 ms. Le temps d'enregistrement choisi est $T = 2.1$ s. Conformément à la remarque formulée sur l'existence d'une réflexion totale sur le fond du modèle, celui-ci étant situé à $\tau = 1.7$ s, le temps d'enregistrement T correspond au temps nécessaire au signal émis par la source pour traverser le modèle, se réfléchir sur le fond et arriver au dernier récepteur, situé à $\tau = 0.8$ s. Les réflexions parasites sur le fond du modèle ne sont donc pas prises en compte.

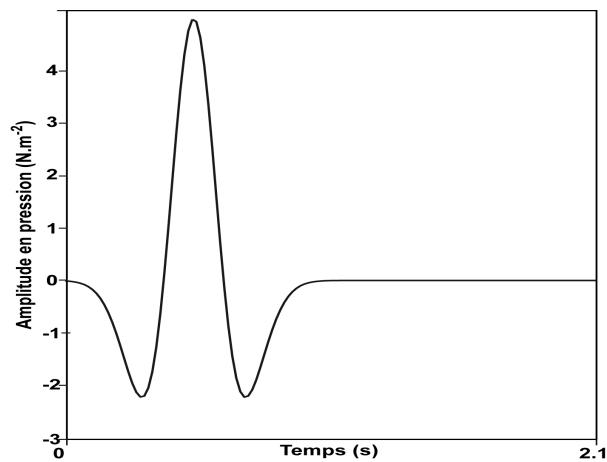


FIG. 1.3 – Signal de Ricker.

La vitesse de déplacement vertical enregistrée au niveau des récepteurs constitue le jeu de données synthétiques. Celui-ci est représenté fig.1.4. L'échelle verticale est en temps, et l'échelle latérale va du récepteur le moins profond au récepteur le plus profond. Ces données sont constituées de trains d'ondes descendantes (de la gauche vers la droite) issues de la source, et d'ondes montantes (de la droite vers la gauche) générées par les réflexions sur les discontinuités d'impédance. L'intensité des trains d'ondes est représentée en utilisant le codage couleur suivant : les intensités positives sont colorées en blanc, tandis que les intensités négatives sont en noir.

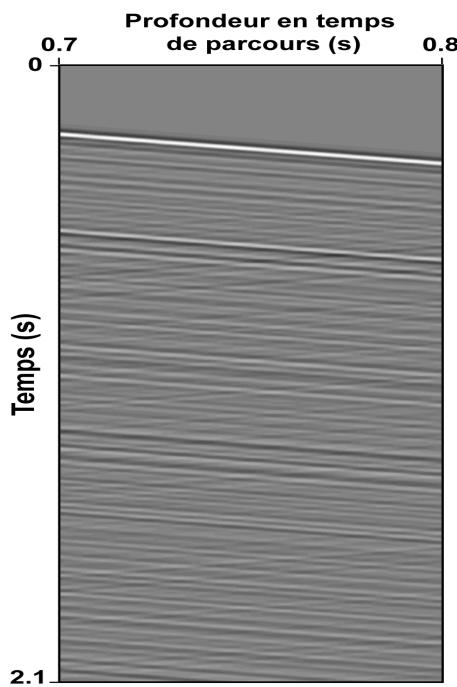


FIG. 1.4 – Données PSV synthétiques.

Deux signaux enregistrés respectivement au niveau du premier et du cinquantième récepteurs sont isolés (fig. 1.5, 1.6). Le premier pic qui peut être observé sur ces figures correspond à l'arrivée de l'onde directe, en blanc au sommet de la figure 1.4.

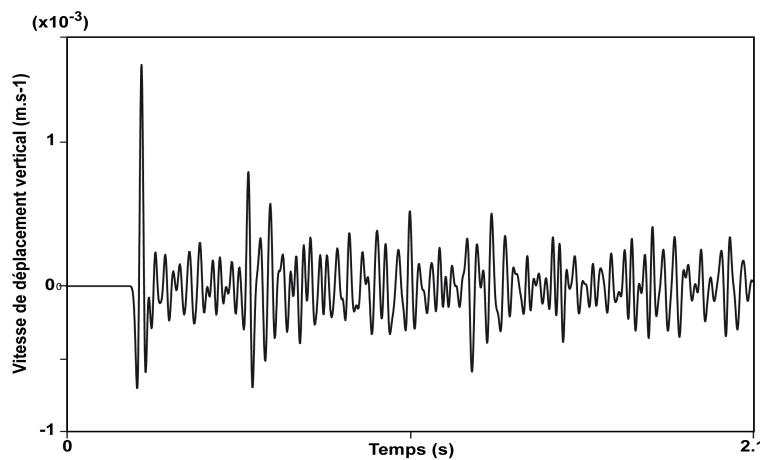


FIG. 1.5 – Enregistrement sur le premier récepteur.

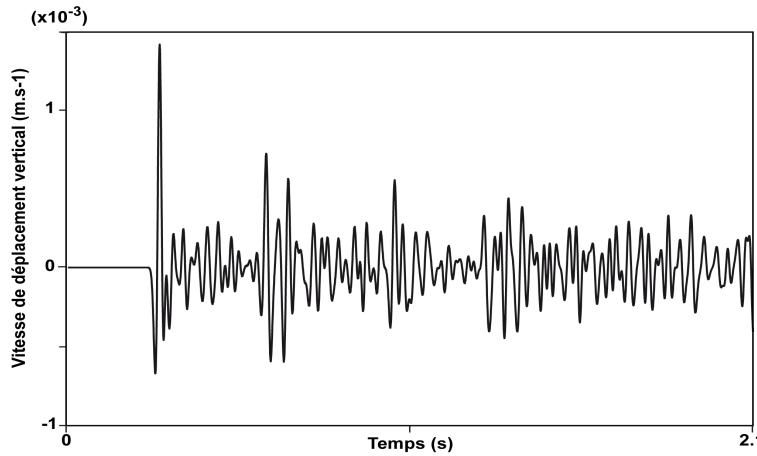


FIG. 1.6 – Enregistrement sur le 50^e récepteur.

Les très nombreuses réflexions multiples qui peuvent être observées sur les données témoignent du caractère très hétérogène du modèle d’impédance utilisé. Ces données synthétiques, une fois constituées, sont inversées par la méthode qui a été décrite plus haut.

La condition de pression $h(t)$ choisie au début de l’inversion est nulle sur l’intervalle $[0, T]$. Le modèle d’impédance en entrée est lui choisi constant. La valeur de cette constante est prise égale à la valeur d’impédance du modèle exact au niveau du premier récepteur $I((\tau_r)_1)$, de manière à s’affranchir de l’indétermination de couplage entre la condition de pression $h(t)$ et le modèle d’impédance $I(\tau)$. Celle-ci se manifeste dans le cas 1D par un simple problème de mise à l’échelle : la solution du problème inverse posé est la même, quels que soient les modèles initiaux d’impédance et de condition de pression, à un facteur multiplicatif près cependant. Le choix d’initialiser la première valeur du modèle d’impédance à la valeur exacte permet de s’affranchir de cette indétermination. Le chapitre 2 revient plus en détail sur cet aspect du problème inverse.

L’analyse des différences entre les données exactes et les données calculées par résolution du problème direct à l’issue de l’inversion (ou résidus) confirme le bon fonctionnement de la méthode numérique de résolution (fig. 1.8, 1.9 et 1.10). Les résidus sur l’ensemble des récepteurs, représentés à la même échelle que les données sismiques, ne laissent apparaître aucun signal. Les résidus isolés au premier et cinquantième récepteur confirment cette observation : la superposition du signal exact et du signal calculé (en noir) est parfaite, et la différence entre les deux signaux (en gris) est identiquement nulle. La fonction coût $J_d(I, h)$, initialement normalisée à 1, décroît rapidement de 4 ordres de grandeur (75 itérations pour 15 mn de calcul). Néanmoins, afin d’expérimenter le potentiel de la méthode d’inversion, les calculs sont prolongés afin atteindre la meilleure précision possible. A l’issue de 31 heures de

calcul, et environ 10000 itérations de l'algorithme de Quasi-Newton, sur une machine de bureau standard (2.5 Go de mémoire RAM, processeur cadencé à 2.80 Ghz), la fonctionnelle décroît de presque 8 ordres de grandeurs. La décroissance la fonction coût est représentée en figure 1.7 sur une échelle logarithmique.

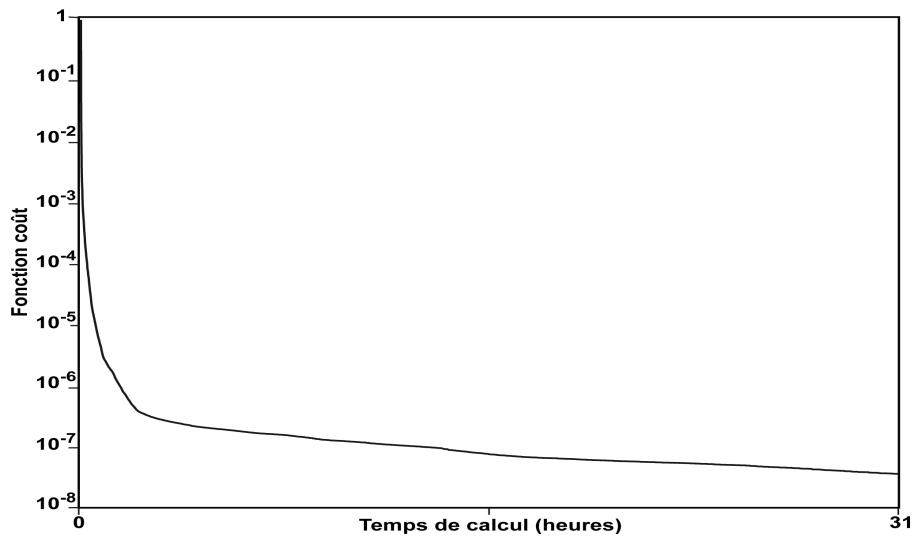


FIG. 1.7 – Décroissance de la fonction coût (cas 1D).

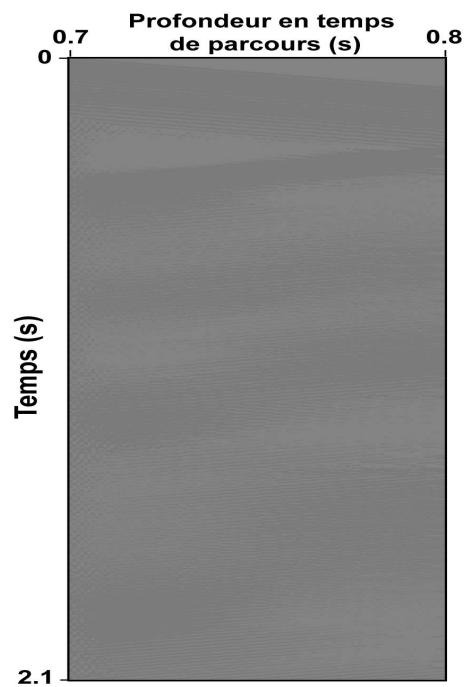


FIG. 1.8 – Résidus obtenus sur l'ensemble des récepteurs.

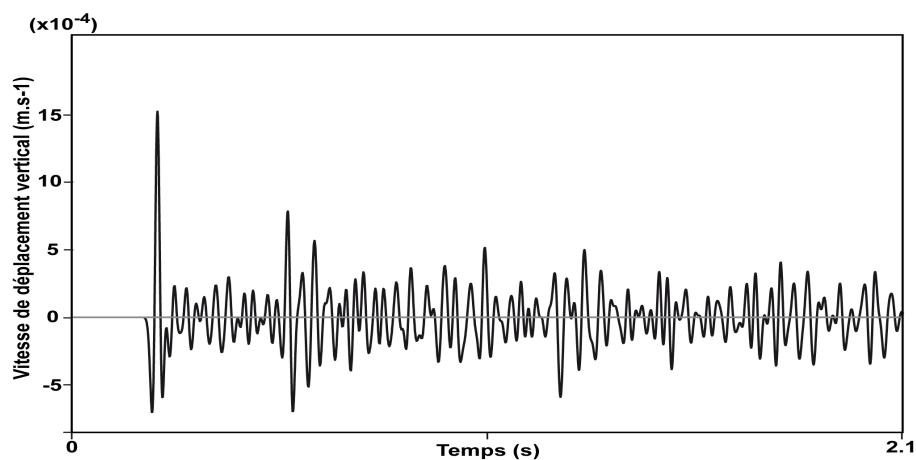


FIG. 1.9 – Résidus obtenus pour le premier récepteur. En noir, données exactes et données synthétiques, en gris, différence entre ces deux signaux.

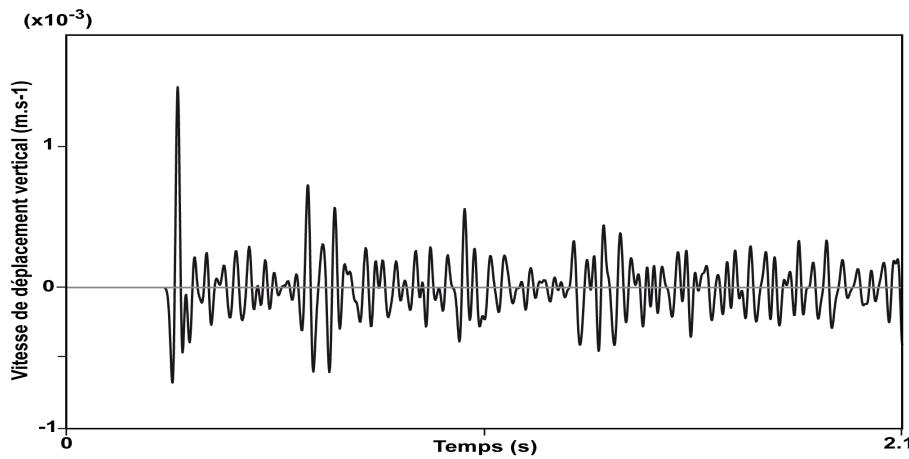


FIG. 1.10 – Résidus obtenus pour le 50^e récepteur (bas). En noir, données exactes et données synthétiques, en gris, différence entre ces deux signaux.

La distribution d'impédance et la condition de pression calculées ajustent donc les observations avec une très grande précision. Le profil de la condition de pression calculée, représenté fig. 1.11 atteste que la condition de pression calculée est presque exacte. La différence entre la condition de pression exacte et la condition de pression calculée (représentée en gris) est en effet presque nulle sur l'intervalle $[0, T]$.

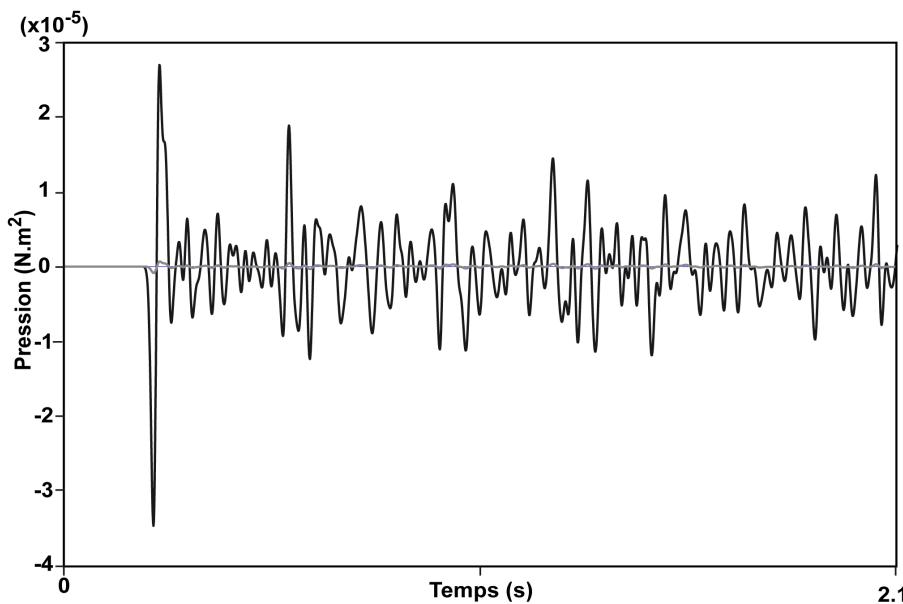


FIG. 1.11 – Résultat d'inversion 1D par rapport à la condition de pression au premier récepteur. En noir, condition exacte et condition calculée, en gris, différence.

Le modèle d'impédance calculé est, quant à lui, à peu près centré sur l'itéré initial (fig.1.12). La très basse fréquence est filtrée par l'acquisition des données, le signal servant à les constituer étant centré autour de 30 Hz. La basse fréquence du modèle d'impédance n'est donc pas recouvrée. En revanche, les variations verticales le sont, et avec une excellente précision. Il faut distinguer en réalité deux zones distinctes. La première correspond à celle où sont localisés les récepteurs, qui s'étend sur les 100 premières millisecondes. Dans cette zone, qui concentre l'information, le modèle d'impédance est retrouvé de manière quasi-exacte. La seconde zone, que l'on appellera par la suite "zone profonde", est la zone située sous les récepteurs. Dans cette zone, même si l'information basse fréquence est perdue, les variations d'impédance sont calculées avec une très grande précision. Des détails à l'échelle de la milliseconde sont retrouvés, ce qui correspond, pour une vitesse de 3000 m.s^{-1} , à des variations de 3 m. Ceci est à comparer avec la résolution moyenne des méthodes d'imagerie sismique classiques, de l'ordre de 20 m, dans le meilleur des cas. Cette expérience met donc en évidence le phénomène de super-résolution [Fink, 1997] prédit par la théorie.

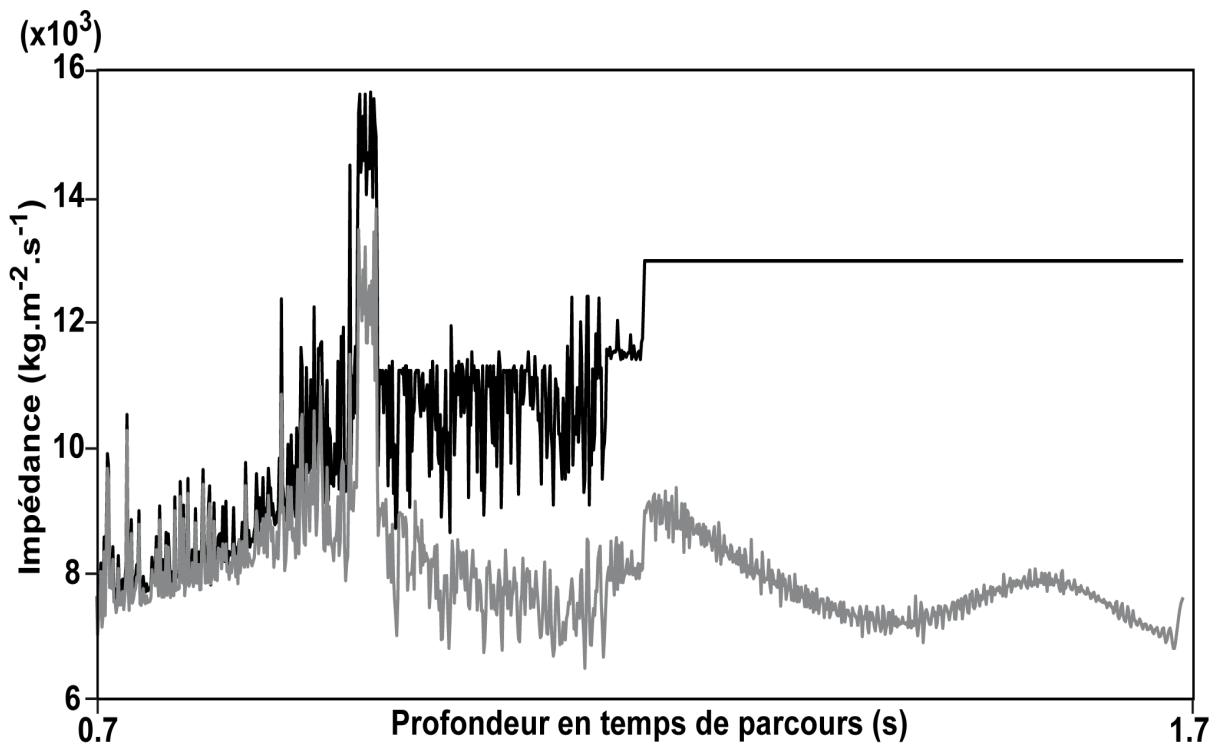


FIG. 1.12 – Résultat d'inversion 1D par rapport à l'impédance.

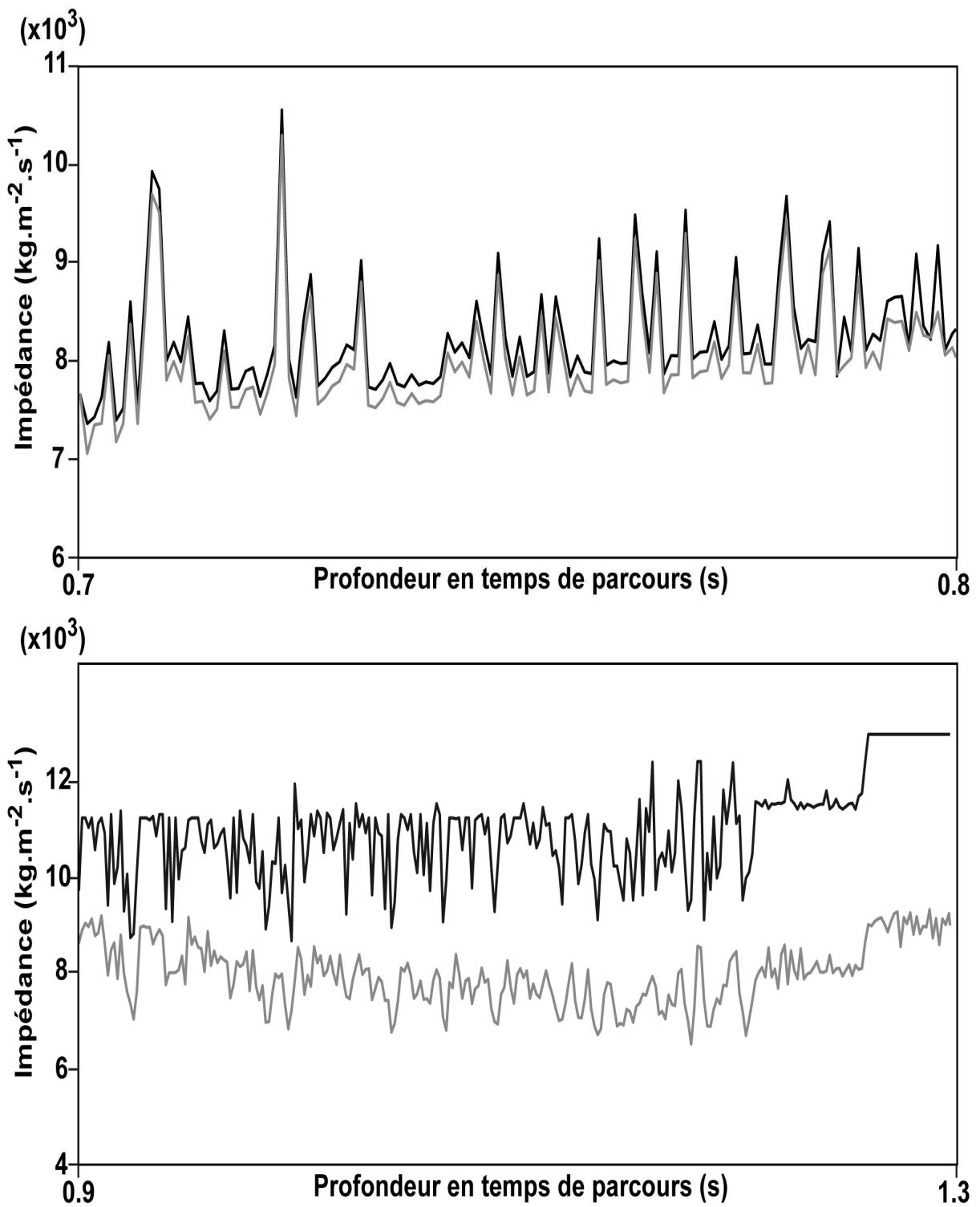


FIG. 1.13 – Résultat d'inversion 1D par rapport à l'impédance. Zoom sur la zone des récepteurs (haut), zoom sur la partie profonde (bas).

1.2.4 Limitations

Les expérimentations numériques qui ont pu être menées dans le cadre 1D sont donc très encourageantes. Il semble en effet que l'inversion de données sismiques de puits est une voie pour surmonter les difficultés que la présence d'hétérogénéités à très fines échelles pose aux méthodes d'inversion classiques. Le calcul au cours de l'inversion d'un intermédiaire de calcul (la condition de pression au premier récepteur $h(t)$) permet en effet de s'affranchir de la détermination *a priori* des sources servant à générer les données, et par là, de prendre en compte dans l'inversion tous les multiples issues des hétérogénéités du milieu. Ceci requiert la mise en place d'une méthode numérique robuste, toutefois, les résultats montrent que cela permet de reconstruire une distribution d'impédance qui corrèle avec beaucoup de détails le modèle recherché. De la prise en compte des réflexions multiples naît donc une amélioration de la résolution de la méthode d'imagerie.

Cependant, le cadre 1D dans lequel l'approche est développée est très restrictif. L'hypothèse de la propagation d'ondes planes et la reconstruction d'un modèle d'impédance 1D limitent la portée de cette méthode à un voisinage de quelques mètres autour du puits. Même si la méthode permet d'obtenir une image en zone profonde, la limitation de la zone imagée à un voisinage très proche du puits restreint fortement la portée de l'application, et ce d'autant plus si cela est rapporté aux zones couvertes en exploration sismique par des méthodes d'inversion classiques, qui peuvent s'étendre sur plusieurs dizaines de kilomètres carrés.

De ces considérations naît naturellement la question suivante : est-il possible d'adapter la méthode d'imagerie sismique de puits à un cadre multi-dimensionnel ? Une telle adaptation permettrait d'étendre la zone imagée autour de puits en prenant en compte des perturbations latérales d'impédance, et permettrait d'utiliser un formalisme plus réaliste pour modéliser la propagation des ondes. L'étendue de la zone couverte ne pourrait évidemment rivaliser avec celle des zones couvertes par l'imagerie de données sismiques de surface, en raison des problèmes d'indétermination liés à la concentration de l'information sismique au niveau du puits, et la nécessité de posséder une estimation du champ de vitesse au voisinage du puits. Cependant, la possibilité de posséder une information haute résolution sur la structure du sous-sol, sur une zone de plusieurs kilomètres de profondeur et plusieurs centaines de mètres de rayon présente un intérêt certain, notamment pour la problématique de monitoring de réservoir, et ses récents développements dans le cadre du stockage souterrain de CO₂.

Le travail mené au cours de cette thèse tente d'apporter une réponse à cette question. L'objectif est de mettre en évidence les difficultés soulevées par une telle extension, et de montrer quels outils peuvent aider à les surmonter. Le travail présenté ici s'intéresse donc au développement de la méthode d'inversion de données sismiques de puits dans un cadre 2D, pour une modélisation acoustique de la propagation

des ondes. L'objectif visé à terme est celui d'une application 3D acoustique, et nous avons veillé à ce que l'essentiel des aspects développés au cours de la thèse s'étendent sans difficultés à ce cadre.

Plan du manuscrit

Dans un premier temps, le chapitre 2 présente la formulation du problème inverse adaptée à l'introduction d'un cadre de modélisation multidimensionnel. En particulier, se posent d'importants problèmes d'indétermination, dûs à l'introduction massive de nouvelles inconnues. L'introduction d'information *a priori* devient alors inévitable, mais il convient de savoir quelle information peut être utilisée à cette fin.

Le problème d'inversion multidimensionnel étant défini, le chapitre 3 expose en détail la méthode de résolution numérique. Le problème posé est un imposant problème d'optimisation numérique, qui nécessite à chaque étape de sa résolution la définition de méthodes adaptées, permettant d'optimiser les temps de calcul et la taille des données manipulées. Une définition du problème est d'abord donnée. Celle-ci fait l'usage de conditions absorbantes pour décrire la propagation des ondes dans le milieu. En effet, si l'utilisation de celles-ci a pu être évitée dans le cadre 1D en définissant un domaine de calcul suffisamment étendu, ce recours n'est plus possible dans un cadre multi-dimensionnel, sous peine d'alourdir les calculs à tel point que l'inversion des données soit rendue impossible. La seconde partie de ce chapitre présente ensuite la procédure d'optimisation développée pour notre problème, qui répond à un besoin d'efficacité des calculs, et à celui de définir une construction de la solution minimisant l'indétermination présentée au chapitre 2. Au sein de la partie 3, le calcul du gradient de la fonctionnelle associée au problème est donné dans un cadre continu. La partie 4 présente la discréétisation du problème, et les parties 5 et 6 donnent les algorithmes mis en place pour l'implémentation de la méthode sur une architecture de calcul parallèle.

Le chapitre 4 est ensuite consacré à une évaluation de la méthode développée, tant d'un point de vue numérique (convergence de la méthode, temps de calcul requis), que d'un point de vue géophysique (résolution du résultat, indétermination, sensibilité). La première partie de ce chapitre décrit exhaustivement le cas test utilisé pour toutes les expériences numériques. Puis, dans la seconde partie, on s'intéresse à l'efficacité de la parallélisation des calculs. La troisième et la quatrième partie présentent quant à elles une analyse détaillée de la résolution de la méthode d'imagerie 2D. La cinquième partie présente finalement quelques cas d'analyse de sensibilité de la méthode ; sensibilité au bruit sur les données, et sensibilité à une erreur commise sur l'estimation du modèle de vitesse, qui est nécessaire à la mise en place de l'inversion 2D.

Finalement, le chapitre 5 revient sur la définition des équations de propagation d'ondes en milieu borné faisant intervenir des conditions absorbantes. Les conditions absorbantes qui ont été utilisées sont des couches PML, et les équations correspondantes ont été dérivées de celles introduites par Bérenger pour l'électromagnétisme

[Bérenger, 1994]. Une analyse mathématique de ces équations est formulée dans la première partie de ce chapitre, sur la base des travaux de thèse de Petit-Bergez [Petit-Bergez, 2006]. De telles équations sont en effet couramment utilisées pour des milieux homogènes, mais leur analyse dans le cadre de milieux hétérogènes n'avait pas été proposée auparavant. Dans le cadre d'application de cette thèse, certains paramètres décrivant le milieu de propagation (comme l'impédance) présentent même des discontinuités. Des estimations d'énergie ont donc été déterminées, qui montrent que le problème, sous certaines conditions de régularité, reste faiblement bien posé de défaut 1. Dans la deuxième partie, des tests numériques ont été entrepris pour s'assurer du bon comportement de la couche PML, mieux comprendre l'effet que les hétérogénéités du milieu peuvent avoir sur leur caractère absorbant, et s'assurer que le biais de modélisation introduit par l'utilisation de cette condition absorbante est suffisamment faible pour que celle-ci puisse être employée dans un contexte d'inversion non linéaire de données sismiques de puits. Cette partie a fait l'objet de la publication d'un article [Métivier, 2009].

Chapitre 2

Formulation d'un problème inverse bidimensionnel

Introduction

L'extension à deux dimensions d'espace de l'inversion 1D non linéaire de données sismiques a pour objectif de s'affranchir du cadre restrictif de la modélisation 1D. Elle doit permettre de déterminer une estimation des paramètres du sous-sol dans une zone étendue latéralement autour du puits. Cette nouvelle méthode reprend les principes fondateurs de l'inversion 1D. L'information est acquise pour des récepteurs alignés verticalement dans un puits, et la résolution du problème consiste en la minimisation d'une fonctionnelle mesurant l'écart entre les données et un sismogramme synthétique.

La première différence consiste en l'acquisition de plusieurs jeux de données, pour des positions de sources différentes, afin d'acquérir une information sur les variations latérales du milieu, à la différence du problème 1D qui ne considère qu'un seul jeu de données, pour une source positionnée à la verticale du puits.

La seconde différence consiste en l'utilisation d'un modèle de propagation d'ondes 2D pour le calcul des sismogrammes synthétiques, ce qui permet de s'affranchir de l'hypothèse de propagation d'ondes planes. Cependant, ceci requiert l'utilisation d'une estimation du modèle de vitesse de propagation, dans le domaine où la propagation des ondes est simulée. En effet, l'utilisation de la conversion profondeur-temps, qui permet de s'affranchir de ce modèle de vitesse, en travaillant en temps de parcours, et qui est utilisée pour les équations de propagation d'ondes acoustiques 1D, ne s'adapte pas au contexte de modélisation 2D. L'estimation du champ de vitesse de propagation dans le sous-sol est une problématique à part, qui dépasse le cadre des travaux qui sont présentés ici. Ceux-ci se focalisent sur l'extension multidimensionnelle de l'inversion non linéaire de données sismiques de puits, qui permet de reconstruire un modèle d'impédance acoustique à partir de données sismiques.

Ainsi, le modèle de vitesse est toujours supposé connu. Néanmoins, il n'est pas possible de faire totalement abstraction de la difficulté à réaliser une estimation fiable du champ de vitesse. Notamment, à partir de données sismiques de puits, l'estimation du champ de vitesse ne peut-être réalisée que sur un domaine d'étendue limitée autour du puits, de l'ordre de quelques centaines de mètres. De plus, cette estimation est d'autant plus difficile que le sous-sol présente d'importantes variations latérales. Ceci est donc la motivation principale pour limiter le contexte de cette étude à un milieu 2D variant peu latéralement. Cette hypothèse joue par la suite un rôle très important dans les techniques mises en place pour limiter l'indétermination du problème.

Par analogie avec le problème 1D, des inconnues auxiliaires sont introduites pour s'affranchir de la connaissance *a priori* des sources servant à générer les données. Ces inconnues correspondent au champ de pression à la profondeur du premier récepteur. Ce champ de pression dépend du temps, sur l'intervalle correspondant au temps d'enregistrement des données. De plus, dans le cadre de l'extension 2D, il dépend de l'extension latérale du domaine imagé, à la différence du problème 1D, pour lequel il ne dépend que du temps. Pour finir, à chaque jeu de données enregistré correspond un champ de pression à la profondeur du premier récepteur différent, il ne s'agit donc plus de reconstruire, en plus du modèle d'impédance, un seul champ de pression, mais bien chacun d'entre eux, pour l'ensemble des jeux de données.

Aussi, l'introduction d'une nouvelle dimension d'espace se traduit par une explosion du nombre d'inconnues prises en compte dans l'inversion. Ceci rend la résolution du problème inverse 2D fortement indéterminée : l'information sismique acquise n'est pas suffisante pour permettre de bien déterminer la solution du problème dans tout le domaine. Cette notion d'indétermination est largement développée dans ce chapitre, et constitue le cœur du problème méthodologique posé par l'extension 2D du problème d'inversion de données sismiques de puits. Les solutions qui lui sont apportées sont présentées ici. Elles consistent, pour l'essentiel, à introduire dans le processus d'inversion toute l'information *a priori* disponible sur les inconnues recherchées pour compenser le manque d'information sismique.

Dans un premier temps, une description des données sismiques utilisées pour le problème inverse 2D est donnée. La partie suivante présente le cadre précis de l'application, et la formalisation mathématique du problème inverse. La troisième partie aborde le problème de l'indétermination. Après en avoir donné une définition, elle explique en quoi le problème traité est indéterminé, et quels sont les moyens utilisés pour faire face à ce problème. La dernière partie conclut en présentant la définition mathématique du problème d'inversion 2D.

2.1 Description des données

Comme pour le problème 1D, Q récepteurs sont installés dans un puits vertical. Leurs coordonnées sont notées $((x_r)_q, (z_r)_q)$, $1 \leq q \leq Q$. L'alignement vertical du puits et la position centrale de celui-ci impliquent que

$$\forall q, \quad 1 \leq q \leq Q, \quad (x_r)_q = 0.$$

La notation $(x_r)_q$ sera cependant conservée afin de ne pas nuire à la généralité des équations. De plus, dans la pratique, l'alignement vertical n'est jamais exactement respecté : le puits peut être dévié, et le positionnement des récepteurs subir des contraintes extérieures, qui peuvent gêner l'alignement exact. Ces récepteurs enregistrent les champs de vitesse de déplacement vertical et horizontal, respectivement $u_z(x, z, t)$ et $u_x(x, z, t)$. L'enregistrement de ce dernier est une nouveauté par rapport au problème 1D. Ce mode latéral de vibration est pris en compte pour maximiser l'information sismique disponible. Il apporte une information supplémentaire essentielle sur les variations latérales du milieu. De plus, ceci est conforme avec les données relevées en pratique dans les campagnes sismiques de puits. Les capteurs utilisés permettent en effet d'enregistrer les modes de vibration verticale et horizontale.

La volonté de reconstruire le modèle d'impédance dans un domaine étendu latéralement implique la nécessité d'acquérir de l'information sismique sur l'ensemble de ce domaine, et non seulement au niveau du puits. A cette fin, S sources servant à constituer les données sont réparties autour de la tête du puits. Suivant la distance à la tête du puits, chaque source éclaire une partie différente du milieu. Le cumul des données permet d'acquérir une information sur l'ensemble de l'extension latérale du milieu. L'éclairage latéral maximal peut être évalué grossièrement en utilisant l'approximation de l'optique géométrique, et en faisant l'hypothèse que le milieu de propagation possède une structure 1D verticale. Dans le cadre de l'approximation géométrique, seules les réflexions primaires sont prises en compte. Le temps d'enregistrement et la vitesse moyenne du milieu donnent une évaluation de la longueur maximale du trajet des ondes enregistrées. En utilisant cette longueur maximale de trajet, la position de la source la plus éloignée du puits, l'approximation de l'optique géométrique et l'hypothèse de structuration 1D du milieu, il est possible de donner une idée de la zone qui va pouvoir être imagée (cf fig. 2.1).

La taille de la zone imagée dépend donc du temps d'enregistrement, et du déport par rapport à la tête du puits (ou offset) maximal utilisé pour les sources. Plus l'offset utilisé est grand, plus l'extension latérale de la zone imagée est grande. Dans la pratique, la zone éclairée est en fait plus large, car l'optique géométrique ne prend pas en compte les réflexions multiples qui naissent dans le milieu de propagation, et

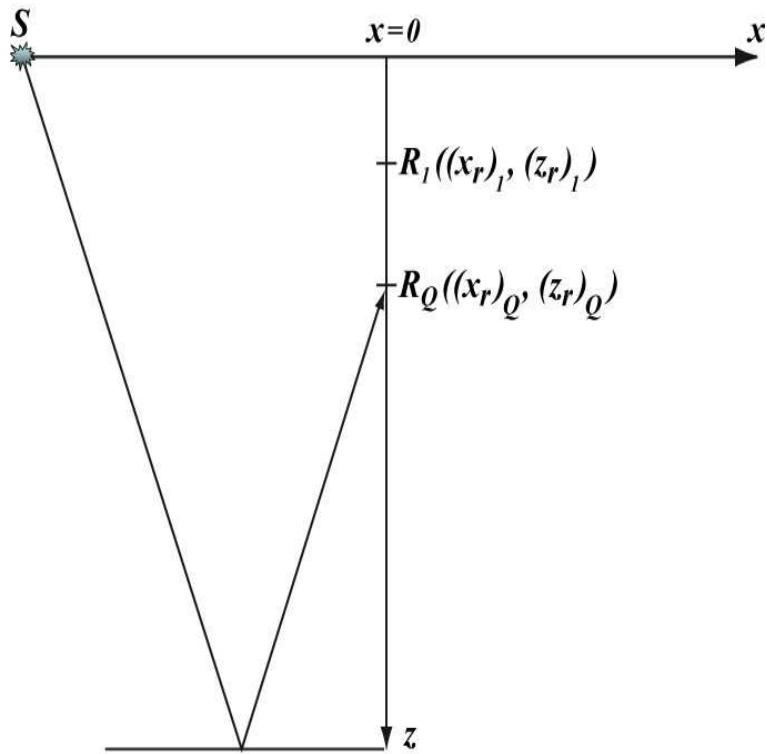


FIG. 2.1 – Zone d'éclairage suivant l'approximation géométrique.

qui sont cependant enregistrées par les récepteurs (cf fig.2.2).

La méthode d'acquisition des données permet donc de constituer S jeux de données indépendants, obtenus par l'enregistrement du champ de vitesse de déplacement vertical et horizontal au niveau de Q récepteurs, à la suite de l'explosion des sources en surface, à différentes distances de la tête du puits. C'est à partir de ces données qu'une information sur le milieu de propagation est recherchée, plus précisément une évaluation quantitative du modèle d'impédance.

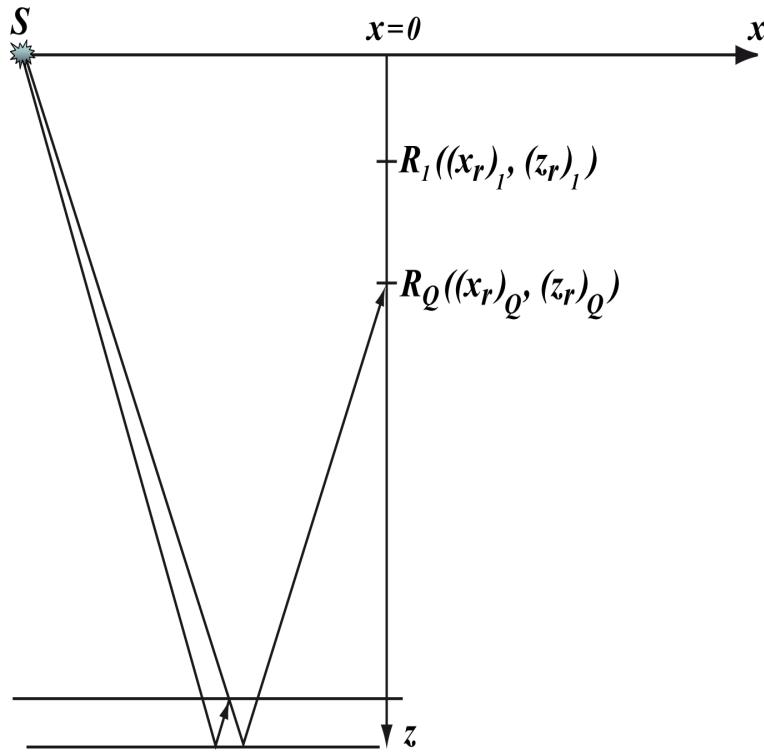


FIG. 2.2 – Prise en compte des multiples pour la zone d'éclairage.

2.2 Estimation du modèle d'impédance

2.2.1 Présentation du problème

Nous nous plaçons dans le cadre de la propagation d'ondes acoustiques. Pour le problème 1D, il est possible de décrire le milieu de propagation à l'aide du seul modèle d'impédance. Ceci s'effectue par un changement de variable correspondant à une conversion temps/profondeur qui permet de masquer la variable de vitesse de propagation des ondes dans les équations. Ceci est impossible pour le modèle 2D. Le milieu de propagation est donc décrit par deux paramètres indépendants. Ceux qui sont retenus ici sont la vitesse de propagation $c(x, z)$ et l'impédance $I(x, z)$. Dans toute la suite, le champ de vitesse de propagation $c(x, z)$ est supposé connu, comme cela a été annoncé en introduction de ce chapitre. Une estimation quantitative du modèle d'impédance $I(x, z)$ est recherchée, à partir des S jeux de données sismiques décrits plus haut. Les S champs de pression à la profondeur du premier récepteur correspondant à chaque jeu de données sont également reconstruits. Ces inconnues auxiliaires sont désignées par CP (Conditions de Pression) dans toute la suite du document. La procédure globale d'inversion est la même que pour le problème 1D : une fonctionnelle mesure l'écart entre les données et les sismogrammes synthétiques 2D. Ces sismogrammes synthétiques sont obtenus par la résolution d'un problème

de propagation d'ondes 2D, le problème direct. La résolution du problème inverse consiste à trouver une distribution d'impédance et des CP qui minimisent cette fonctionnelle. La formalisation mathématique de ces deux problèmes est présentée dans les parties suivantes.

2.2.2 Problème direct 2D

Soit Ω le demi-espace $\mathbb{R} \times [z_1, +\infty[$, où z_1 est la profondeur du premier récepteur. A la différence du modèle de vitesse, qui présente en principe une certaine régularité, le modèle d'impédance $I(x, z)$ est intrinsèquement discontinu. La méthode d'inversion de données sismiques repose sur le calcul d'un modèle d'impédance pouvant expliquer les réflexions enregistrées dans les données, qui sont issues de ces discontinuités. Le modèle d'impédance est donc défini dans $C_m^\infty(\Omega)$ (espace des fonctions $C^\infty(\Omega)$ par morceaux définies sur Ω). La question de la régularité de la CP à la frontière ne sera pas abordée ici, ceci renvoyant au problème abordé dans le chapitre 5. La propagation des ondes acoustiques dans Ω sur l'intervalle de temps $[0, T]$ est décrite par l'équation scalaire en pression

$$\partial_{tt} p - c I \operatorname{div} \left(\frac{c}{I} \nabla p \right) = 0, \quad (2.1)$$

avec la CP

$$p(x, z_1, t) = h(x, t), \quad (2.2)$$

et des conditions initiales nulles. La variable $p(x, z, t)$ correspond au champ de pression dans Ω à l'instant t . Les symboles ∂_{tt} et ∇ désignent respectivement l'opérateur de dérivation partielle d'ordre 2 en temps et l'opérateur gradient. De plus, il est possible de relier les vitesses de déplacement vertical et horizontal à la pression par les relations

$$\partial_t u_x = \frac{c}{I} \partial_x p \quad \partial_t u_z = \frac{c}{I} \partial_z p. \quad (2.3)$$

où ∂_t , ∂_x et ∂_z désignent les opérateurs de dérivation partielle d'ordre 1 respectivement en temps, en espace latéralement et en espace verticalement.

La résolution du problème direct consiste, pour un modèle d'impédance $I(x, z)$ et une CP $h(x, t)$, à calculer le sismogramme synthétique $\tilde{u}(t) = (\tilde{u}_x(t), \tilde{u}_z(t))$ avec

$$\begin{aligned} \tilde{u}_x(t) &= (u_x((x_r)_1, (z_r)_1, t), \dots, u_x((x_r)_Q, (z_r)_Q, t)) \\ \tilde{u}_z(t) &= (u_z((x_r)_1, (z_r)_1, t), \dots, u_z((x_r)_Q, (z_r)_Q, t)), \end{aligned}$$

où u_x et u_z sont calculés à partir des équations (2.1), (2.2) et (2.3). Ceci est formalisé par l'opérateur

$$\mathcal{M} : (I, h) \longrightarrow \tilde{u}(t),$$

où $\tilde{u}(t)$ appartient à $(L^2(0, T))^{2Q}$.

La résolution du problème direct consiste donc à calculer la réponse en vitesse de déplacement au niveau des récepteurs, à partir d'un modèle 2D d'impédance acoustique et du champ de pression à la profondeur du premier récepteur, défini sur le domaine $[0, T] \times \mathbb{R}$, le champ de vitesse $c(x, z)$ étant connu. Cette définition du problème direct permet d'aborder celle du problème inverse.

2.2.3 Problème inverse 2D

Considérant S jeux de données $D_s(t) \in (L^2(0, T))^{2Q}$ obtenus pour S sources placées à différents offsets, la fonctionnelle mesurant l'écart entre le sismogramme synthétique et les données est la fonctionnelle suivante

$$J(I, h_1, \dots, h_S) = \sum_{s=1}^S \|D_s - \mathcal{M}(I, h_s)\|_{(L^2(0, T))^{2Q}}^2.$$

Dans cette formulation, la norme $\|\cdot\|_{(L^2(0, T))^{2Q}}$ est celle associée au produit scalaire défini sur $(L^2(0, T))^{2Q}$ par

$$\forall (u, v) \in (L^2(0, T))^{2Q} \times (L^2(0, T))^{2Q}, \quad (u, v) = \sum_{q=1}^{2Q} \int_0^T u_q(t)v_q(t)dt.$$

Formellement, le problème inverse 2D s'écrit donc

$$(P_0) \quad \min_{I, h_1, \dots, h_S} J(I, h_1, \dots, h_S).$$

La résolution de ce problème inverse consiste à déterminer un ensemble de fonctions

$$\bar{I}(x, z), \bar{h}_1(x, t), \dots, \bar{h}_S(x, t),$$

produisant un sismogramme synthétique le plus proche possible, au sens de la norme $L^2(0, T)$, des données enregistrées $D_s(t)$. Cependant, la résolution de ce problème est assujettie à d'importants problèmes d'indétermination. Notamment, l'information acquise avec le dispositif sources/récepteurs n'est pas suffisante pour déterminer de manière unique la solution du problème. Ceci est le problème fondamental posé par l'extension multidimensionnelle de l'inversion de données sismiques de puits. Une méthodologie reposant sur des hypothèses précises doit être mise en place pour surmonter cette difficulté. Elle est détaillée dans la partie suivante.

2.3 Indétermination et instabilité

2.3.1 Définitions

L'extension multidimensionnelle de l'inversion 1D de données sismiques de puits suscite des problèmes de détermination et de stabilité de la solution. Il convient dans un premier temps de donner une définition claire de ces notions. Ces définitions s'appuient sur la caractérisation des problèmes bien posés introduite par Hadamard [Hadamard, 1902]. Les critères définis par Hadamard pour qu'un problème soit bien posé sont

- { - l'existence d'une solution
- l'unicité de la solution
- la dépendance continue de la solution du problème par rapport aux données.

Fréquemment, les problèmes inverses sont des problèmes mal posés, dans le sens où l'une ou plusieurs de ces trois propriétés n'est pas respectée. L'exemple le plus connu est celui de l'équation de la chaleur. Tandis que la détermination à l'aide de l'équation de la chaleur d'une distribution finale de température, dans un matériau donné, à partir d'une condition initiale, est un problème bien posé au sens d'Hadamard, le problème inverse consistant à déduire la condition initiale de température à partir de la distribution finale est un problème mal posé : de très faibles perturbations sur la condition finale mènent à des distributions initiales très différentes. Dans ce cas particulier, la dernière propriété n'est pas satisfaite : la solution du problème inverse (la distribution initiale de température) ne dépend pas continûment des données (la distribution finale de température).

Nous allons dans la suite distinguer deux caractéristiques des problèmes mal posés, reliées chacune à un critère d'Hadamard, qui font souvent l'objet de confusions dans la littérature. La première d'entre elles est l'instabilité, qui vient d'être évoquée pour l'exemple de l'équation de la chaleur. L'instabilité traduit la non-continuité de la dépendance de la solution par rapport aux données du problème. Une faible perturbation des données engendre une importante modification de la solution du problème. La deuxième d'entre elles est l'indétermination. Cette notion est liée au non respect de la seconde propriété d'Hadamard, et caractérise la non-unicité de la solution du problème. Plusieurs solutions au problème existent, et peuvent être très différentes l'une de l'autre. La question se pose alors de pouvoir discriminer parmi les solutions au problème les "bonnes" solutions des "mauvaises".

Les travaux menés dans cette thèse ont principalement porté sur le problème

de l'indétermination. Pour s'affranchir du problème d'instabilité, les expériences numériques ont été effectuées sur des données synthétiques non bruitées, modélisées par les équations définissant le problème direct. Le bruit de mesure et le bruit de modélisation sont donc ainsi évités. Cependant, l'instabilité du problème a été mise en évidence par les expériences numériques qui ont été menées (chapitre 4). En effet, celles-ci ont montré que des modèles de paramètre (I, h_1, \dots, h_S) très différents correspondent à des données très proches. Cette instabilité du problème mérite un traitement spécifique, dont une première approche pourrait être l'introduction de termes de régularisation L^1 . Cette idée est développée plus avant dans les perspectives.

2.3.2 Origine de l'indétermination

L'indétermination est donc liée au problème de l'unicité de la solution. Dans le cas de la résolution du problème (P_0) , la question de l'unicité est extrêmement délicate à trancher. Il s'agit en effet de la recherche du minimum d'une fonctionnelle non-linéaire. Cette fonctionnelle peut présenter des minima locaux, des plateaux, et rien n'assure de l'existence d'un minimum global. Il est donc tout à fait possible que le problème (P_0) présente un ensemble important de solutions.

Il est intéressant de s'interroger sur l'origine physique de cette potentielle non-unicité de la solution. Celle-ci découle directement de deux facteurs, qui sont l'insensibilité du problème direct aux variations de certains des paramètres dont il dépend, et le couplage dans l'inversion des inconnues d'impédance $I(x, z)$ d'une part et des CP $h_s(x, t)$ d'autre part.

L'insensibilité du problème direct s'interprète facilement d'un point de vue physique. Ce problème consistant à calculer le champ d'ondes acoustiques au niveau des récepteurs, une perturbation, si grande soit-elle, sur les paramètres d'impédance $I(x, z)$ ou du champ de pression au premier récepteur $h(x, t)$ situés à une grande distance des récepteurs latéralement, n'a que peu ou pas d'effet sur le sismogramme synthétique $\tilde{u}(t)$. Ainsi, le problème direct est insensible aux variations des paramètres éloignés latéralement. La résolution du problème inverse consistant à déterminer un jeu de paramètres permettant de produire un sismogramme synthétique aussi proche que possible des données enregistrées, l'insensibilité du problème direct résulte en un manque d'information sur les paramètres éloignés latéralement. L'ensemble des valeurs que ceux-ci peuvent prendre, tout en produisant une valeur de la fonction coût sismique $J(\alpha, h_1, \dots, h_S)$ similaire, est très vaste. Ce problème est donc lié au faible éclairage des paramètres éloignés latéralement des récepteurs.

Le problème du couplage des inconnues posé par la résolution du problème inverse est un peu plus complexe. Pour mieux l'appréhender, il est intéressant de revenir au problème inverse de données de puits 1D. Dans ce contexte, dans le cadre de

l'approximation géométrique de la propagation des ondes, les données enregistrées correspondant à un réflecteur $r(z)$ situé sous le puits (un réflecteur désigne un contraste d'impédance) sont constituées d'abord de l'onde descendante issue des sources, puis de l'onde montante réfléchie par $r(z)$. Il est donc assez simple de dissocier dans le signal enregistré par les récepteurs la contribution de la condition de pression $h(t)$ (onde descendante) de celle du réflecteur $r(z)$ (onde montante). Le couplage des inconnues $h(t)$ et $I(z)$ ne pose donc pas véritablement de problème pour l'inversion. L'étude mathématique du problème 1D révèle que l'indétermination relative au couplage entre les deux inconnues n'est liée qu'à leur amplitude. En effet, il est prouvé de manière théorique [Macé, Lailly, 1986], [Bamberger, 1978] que pour $\bar{h}(t)$, $\bar{I}(z)$ solution du problème 1D, et pour $\lambda \in \mathbb{R}_+^*$ suffisamment petit, tout couple

$$(\lambda \bar{h}(t), \lambda \bar{I}(z))$$

est solution du même problème. Ainsi, pour le problème 1D, le problème de couplage des inconnues est résolu en fixant la valeur de l'impédance au premier récepteur $I((x_r)_1, (z_r)_1)$, valeur qui de plus est accessible directement à la mesure.

En revanche, dans le cadre de l'extension 2D, pour un réflecteur $r(x, z)$ éloigné latéralement des récepteurs, cette dissociation devient plus délicate. En effet, dans ce cas, l'onde descendante correspondant à $h(x, t)$ n'est pas enregistrée par les récepteurs. Ainsi, les influences de $h(x, t)$ et de l'impédance $I(x, z)$ sur le signal enregistré par les récepteurs correspondant au réflecteur $r(x, z)$ ne peuvent être dissociées. De ceci résulte une forte indétermination du problème d'inversion 2D. En effet, de nombreuses combinaisons entre les CP et l'impédance $I(x, z)$ peuvent être à l'origine des signaux enregistrés par les récepteurs.

L'ensemble des solutions du problème (P_0) est donc potentiellement très vaste, de par les problèmes d'éclairage du milieu dans les zones éloignée latéralement, et surtout en raison de la difficulté à dissocier les contributions de l'impédance $I(x, z)$ et des CP aux données enregistrées au niveau des récepteurs. Pour surmonter cette difficulté, il convient de réduire l'ensemble des solutions.

Cette réduction de l'ensemble des solutions ne peut se faire que par l'ajout de critères supplémentaires sur les solutions qui doivent être retenues. Ceci suppose qu'une certaine information existe au préalable sur le résultat, et qu'il est possible de discriminer parmi les solutions celles qui possèdent un sens physique de celles qui n'en ont pas. Un certain effort de formalisation de la connaissance doit donc être effectué. Cette information *a priori* est introduite dans le problème via l'utilisation de termes de régularisation, qui sont décrits dans la partie suivante. Ce sont des termes de régularisation qui s'inspirent des travaux initiés par [Tikhonov, 1943], sur la régularisation des problèmes inverses mal posés. Dans un premier temps, les termes de régularisation correspondant à l'information disponible sur le modèle d'impédance sont donnés.

2.3.3 Information *a priori* sur le modèle d'impédance

2.3.3.1 Information au niveau du puits

Lors du forage d'un puits, des mesures de diagraphie permettent d'obtenir une idée assez précise du modèle d'impédance localement au niveau du puits. Ces mesures à très fines échelles (< 1 m) sont homogénéisées pour pouvoir être introduites à l'échelle de l'inversion de données sismiques (≤ 10 m). Une information *a priori* sur le modèle d'impédance dans le puits est donc disponible. Cette information pourrait être imposée comme modèle d'impédance exact dans le puits. Cependant, l'estimation de ce modèle local est sujette à des erreurs qui peuvent être liées à la mesure ou la mise à l'échelle de l'information. Il est donc préférable d'introduire cette information via la définition d'un terme de régularisation $R_1(I)$ dans la fonction coût, qui mesure l'écart en norme L^2 entre le modèle d'impédance calculé au niveau du puits, et le modèle *a priori* disponible $I_0(z)$,

$$R_1(I) = \int_{(z_r)_1}^{(z_r)_Q} (I((x_r)_1, z) - I_0(z))^2 dz.$$

Cette information *a priori* permet d'aider la détermination des valeurs d'impédance clés, au niveau du puits. Celles-ci possèdent une très forte influence sur le calcul du sismogramme synthétique, puisqu'elles correspondent à l'emplacement des récepteurs. L'introduction de cette information a pour effet d'accélérer la convergence de l'algorithme et permet de caler l'information basse fréquence du modèle. De plus, elle permet de lutter contre l'indétermination due au couplage entre les CP et l'impédance $I(x, z)$ au niveau du puits. Tout comme pour le problème 1D, les premières valeurs d'impédance au niveau du puits sont nécessaires à la détermination des bonnes amplitudes des CP. En revanche, cette information n'a pas d'effet sur l'indétermination latérale du modèle. Un second terme de régularisation est donc également employé pour le modèle d'impédance.

2.3.3.2 Variations latérales

Le contexte d'application de la méthode, défini dans l'introduction de ce chapitre, impose de s'intéresser à des milieux présentant de faibles variations latérales. Cette hypothèse constitue un critère efficace pour réduire l'indétermination latérale liée aux paramètres d'impédance acoustique. Cette information *a priori* est prise en compte par l'introduction d'un terme de régularisation $R_2(I)$ dans la fonction coût, qui mesure le carré de la norme $L^2(\Omega)$ des dérivées latérales du modèle d'impédance acoustique $I(x, z)$ recherché.

$$R_2(I) = \|\partial_x I\|_{L^2(\Omega)}^2.$$

Les paramètres d'impédance ayant peu d'influence sur le sismogramme synthétique se voient donc contraints de respecter un critère de régularité latérale. Cette information se substitue à l'information sismique à mesure que l'influence des paramètres d'impédance sur le sismogramme synthétique diminue. Cette approche s'étend

aisément au cas de milieux présentant une régularité latérale suivant une certaine déclivité. Le terme de régularisation consiste alors à minimiser la norme L^2 des dérivées spatiales de l'impédance suivant la direction correspondant à cette déclivité.

Ce critère de régularité permet donc d'introduire l'information *a priori* nécessaire pour restreindre l'espace des solutions, en ce qui concerne le modèle d'impédance. En outre, cette information sur la structuration générale du milieu peut être donnée par des études géologiques de terrain. Elle joue également un rôle fondamental dans la partie à suivre, qui concerne l'information *a priori* introduite pour lever l'indétermination sur le calcul des CP. En effet, les termes de régularisation $R_1(I)$ et $R_2(I)$ ne suffisent pas seuls à combattre toute l'indétermination du problème, et une information supplémentaire pour contraindre les CP doit être ajoutée.

2.3.4 Information *a priori* sur les conditions de pression

2.3.4.1 Profil d'une condition de pression pour des données générées par une source : difficultés de régularisation

Le champ de pression généré par une source ponctuelle placée en surface et se propageant dans un milieu 2D possède un profil spatial de type hyperbolique. Ainsi, le profil du champ de pression pris dans le domaine temps/extension latérale, enregistré à la profondeur du premier récepteur $z = (z_r)_1$, possède également un profil courbe, de type hyperbolique. La courbure de cette hyperbole est liée au champ de vitesse dans la zone supérieure du modèle, ou couverture, qui s'étend de la surface, à la profondeur du premier récepteur $z = (z_r)_1$. Il n'existe donc pas de critère de régularité simple pouvant servir d'information *a priori* pour les CP correspondant à des données obtenues par l'explosion de sources uniques. Pourtant, l'extension latérale des CP, comme celle du modèle d'impédance, souffre d'indétermination, de par la concentration de l'information autour du puits et l'insensibilité du problème direct aux variations des paramètres éloignés latéralement. Une alternative doit donc être proposée, afin de réduire l'indétermination par rapport aux CP.

2.3.4.2 Approximation ondes planes et recombinaison des données

Pour pallier à ce problème, une solution consiste à se rapprocher d'une situation plus favorable à l'introduction de critères de régularisation sur les CP.

Considérons l'exemple simple d'un milieu parfaitement 1D (l'impédance et la vitesse ne dépendent que de la variable z : $I(x, z) = I(z)$ et $c(x, z) = c(z)$). Pour un tel milieu, la propagation d'une onde plane à incidence normale produit un champ de pression à la profondeur du premier récepteur $(z_r)_1$ invariant dans la direction x dans le plan (x, t) . De plus, la propagation d'ondes planes avec des angles d'incidence suffisamment petits produit également des champs de pression à la profondeur $(z_r)_1$ présentant une régularité suivant la direction normale à la direction de propagation

dans le plan (x, t) , du moins pour les premières arrivées. En effet, les temps plus longs voient l'arrivée d'ondes réfléchies se propageant dans la direction opposée à celle de l'onde plane initiale. Néanmoins, l'amplitude de ces ondes réfléchies est bien moins intense que celle de l'arrivée directe. Ainsi, la CP correspondant à un tel éclairage possède une régularité importante suivant une direction dans le plan (x, t) qui peut être reliée à l'angle d'éclairage. Si le milieu de propagation est un milieu 2D proche d'un milieu 1D (faibles variations latérales), ces résultats de régularité demeurent pour l'essentiel. Les variations latérales du champ de vitesse $c(x, z)$ dans la partie superficielle, ou couverture (zone allant de la surface $z = 0$ à la profondeur du premier récepteur $z = (z_r)_1$) produisent quelques perturbations sur les CP mais la régularité générale du profil est conservée. Dans l'hypothèse d'un milieu proche 1D, l'éclairage par ondes planes suivant différents angles d'incidence, pourvu que ces angles ne soient pas trop élevés ($< 10^\circ$ par exemple) permet donc d'obtenir des CP présentant un profil assez régulier suivant la direction normale à la direction de propagation. Pour de telles CP, il est alors aisément d'introduire un terme de régularisation, qui pénalise les dérivées de la CP suivant cette direction, de la même manière que l'impédance $I(x, z)$ est régularisée suivant la direction de stratification dans le paragraphe précédent. La question qu'il faut alors résoudre est la suivante : comment obtenir les données correspondant à un éclairage par de telles ondes planes à partir des S jeux de données obtenus par un éclairage par points sources ?

Le cas le plus simple à traiter est celui de l'onde plane se propageant à incidence normale. L'explosion simultanée des sources en surface, réparties sur la largeur du domaine imagé, produit la propagation d'un front d'onde dont le profil correspond à la superposition des fronts d'ondes hyperboliques. Pour un nombre de sources suffisant, réparties autour du puits à intervalle suffisamment réduit, ce profil correspond localement autour du puits à celui d'une onde plane se propageant à incidence normale. De plus, le calcul du champ d'onde acoustique étant linéaire par rapport aux sources, les données acquises pour l'explosion simultanée des sources correspondent à la sommation des données acquises indépendamment pour chaque source. Il convient donc simplement d'effectuer la sommation des S jeux de données obtenus pour chaque source pour obtenir un jeu de données correspondant à la propagation d'une onde présentant une régularité latérale s'approchant de celle d'une onde plane se propageant à incidence normale.

Pour s'approcher de l'éclairage du milieu par une onde plane oblique, la méthode est un peu plus complexe. Considérons l'explosion de deux sources voisines S_k et S_{k+1} en surface, situées respectivement aux abscisses x_k et x_{k+1} . Contrairement au cas précédent, si l'explosion des sources n'est pas simultanée, mais qu'un certain retard Δt est introduit d'une source par rapport à l'autre, le front d'onde généré localement au niveau du puits présente un profil oblique. L'angle d'éclairage de cette pseudo-onde plane est l'angle θ que fait la direction de propagation avec la verticale du milieu (fig. 2.3). De plus, il existe une relation entre cet angle, la distance entre les

sources, et le décalage temporel introduit entre les sources S_k et S_{k+1} , telle que, pour une vitesse moyenne c_{moy} de la zone de couverture,

$$\Delta t(S_k) \simeq (x_{k+1} - x_k) \sin(\theta) c_{moy}$$

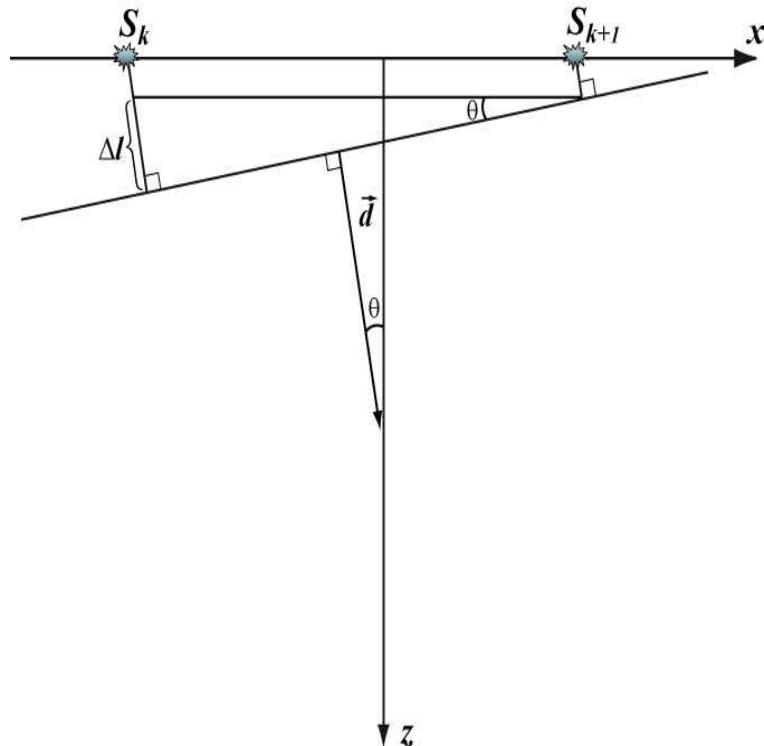


FIG. 2.3 – Construction d'une onde plane à incidence oblique

Ces décalages étant relatifs d'une source à l'autre, il convient de fixer le décalage d'au moins l'une d'entre elles. Dans le cas d'un angle $\theta > 0$, c'est la source la plus à droite qui sert de référence, et son décalage est fixé à $\Delta t = 0$ s. Dans le cas contraire ($\theta < 0$), c'est la source la plus à gauche qui joue ce rôle et qui voit, de la même façon, son décalage fixé à $\Delta t = 0$ s. En utilisant cet artifice de décalages en temps calculés à partir d'une évaluation de la vitesse moyenne de la couverture, il est donc possible de générer des ondes planes à incidence oblique. L'amplitude des décalages en temps est donnée par l'angle d'éclairage que l'on désire obtenir. L'angle maximum utilisé doit rester suffisamment petit pour garantir la bonne régularité de la pseudo-onde plane correspondante, et doit également correspondre à la géométrie de l'acquisition. A celle-ci correspond en effet une zone maximale d'éclairage, dont une évaluation peut-être faite, et qui donne un angle maximal d'éclairage (cf paragraphe 2.1).

Deux CP obtenues par l'explosion de sources décalées en temps dans un milieu proche 1D sont présentées fig 2.4, pour une pseudo-onde plane se propageant à

incidence normale, et une autre se propageant avec un angle de $\theta = 6^\circ$. Le temps est représenté suivant l'axe vertical, et l'extension latérale du milieu suivant l'axe des abscisses x . Les niveaux de gris indiquent l'intensité du signal, les signaux clairs (vers le blanc) étant d'intensité positive, et les signaux foncés (vers le noir) étant d'intensité négative. Comme cela a été annoncé, ces CP possèdent effectivement une certaine régularité suivant une pente relative à la direction de propagation, même si la superposition imparfaite des fronts d'ondes hyperboliques issus de chaque source, et les perturbations engendrées par les hétérogénéités latérales de vitesse dans la zone de couverture dégradent cette hypothèse de régularité.

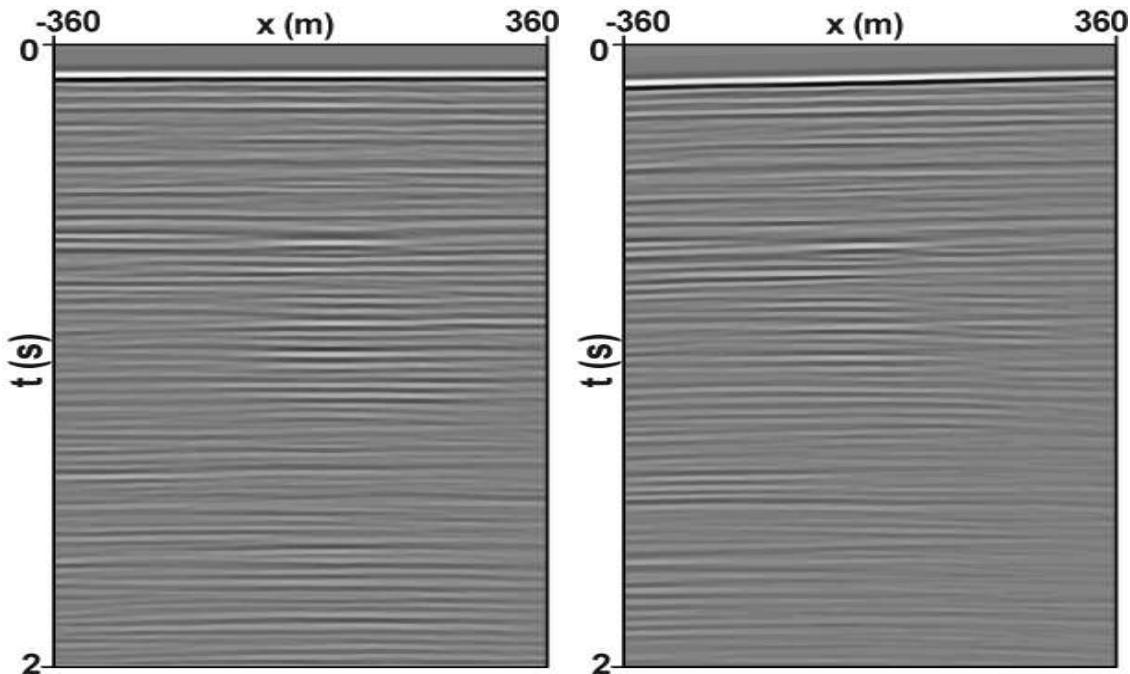


FIG. 2.4 – CP exactes incidence normale (gauche), incidence oblique 6° (droite).

La linéarité du problème de propagation d'ondes par rapport aux sources permet d'introduire *a posteriori* ces décalages en temps directement dans les données, ce qui permet de ne pas revenir sur la méthode d'acquisition. Les données correspondant à la propagation d'ondes planes obliques sont donc obtenues à partir de la sommation et de l'introduction des décalages en temps appropriés des S jeux de données initiaux, obtenus pour des sources indépendantes.

Il est donc possible, à partir des S jeux de données initiaux, de constituer S nouveaux jeux de données, par sommation et introduction de décalages temporels, qui correspondent à l'illumination du milieu de propagation par des quasi-

ondes planes possédant divers angles de propagation. Ce concept a été introduit par [Claerbout, Schultz, 1978] dans le domaine de la géophysique sous le nom de "slant-stacking", et est en fait équivalent, sur le plan mathématique, à une transformation de Radon opérée sur les S jeux de données, en utilisant une estimation de la vitesse moyenne c_{moy} pour passer du domaine profondeur (x, z) au domaine temps (x, t) [Radon, 1917]. Les CP correspondant à ces données recombinées présentent des profils réguliers dans le domaine (x, t) , suivant une droite dont le vecteur directeur γ_s est donnée par

$$\gamma_s = \begin{bmatrix} (\gamma_s)_x \\ (\gamma_s)_t \end{bmatrix} = \begin{bmatrix} 1 \\ \frac{\sin(\theta_s)}{c_{moy}} \end{bmatrix}.$$

Ceci permet d'introduire une information *a priori* sur la régularité des CP afin de lever l'indétermination de leur extension latérale. Cette information se traduit mathématiquement par l'ajout d'un terme de régularisation R_3 dans la fonctionnelle J , sous la forme

$$R_3(h_1, \dots, h_S) = \sum_{s=1}^S \|(\gamma_s)_x \partial_x h_s + (\gamma_s)_t \partial_t h_s\|_{L^2([0,T] \times \mathbb{R})}^2.$$

Les paragraphes précédents ont donc permis d'illustrer l'origine de l'indétermination et la difficulté posée par celle-ci pour la résolution du problème inverse (P_0), ainsi que les moyens utilisés pour la réduire. Cette indétermination est la traduction de la non-unicité de la solution du problème résolu, et sa réduction consiste en l'introduction d'information supplémentaire afin de restreindre l'espace des solutions. Cette information est prise en compte par l'introduction de termes de régularisation, qui amènent à définir un nouveau problème inverse, le problème inverse régularisé.

2.3.5 Problème inverse régularisé

2.3.5.1 Expression de la fonctionnelle régularisée

A l'issue de la définition des termes de régularisation, la fonctionnelle J se voit augmentée de trois termes, les deux premiers portant sur l'impédance $I(x, z)$ et le troisième sur les CP $h_s(x, t)$. Soit $\varepsilon \in \mathbb{R}^3 = (\varepsilon_1, \varepsilon_2, \varepsilon_3)$, et J^ε la fonctionnelle régularisée, alors

$$\begin{aligned} J^\varepsilon(I, h_1, \dots, h_S) &= J(I, h_1, \dots, h_S) + \\ &\varepsilon_1 \|\partial_x I\|_{L^2(\Omega)}^2 + \varepsilon_2 \int_{z_1}^{z_N} (I(0, z) - I_0(z))^2 dz + \varepsilon_3 \sum_{s=1}^S \|(k_s)_x \partial_x h_s + (k_s)_t \partial_t h_s\|_{L^2([0,T] \times \mathbb{R})}^2. \end{aligned} \tag{2.4}$$

Le problème d'optimisation à résoudre devient alors

$$\min_{I, h_1, \dots, h_S} J^\varepsilon(I, h_1, \dots, h_S) \quad (2.5)$$

2.3.5.2 Une information *a priori* inexacte

Il résulte de cette introduction d'information *a priori* approximative la nécessité d'un réglage attentif des poids attribués aux termes de régularisation. En effet, il est primordial de noter que l'information introduite n'est pas exacte, et ce à plusieurs niveaux. Le profil d'impédance mesuré au puits est sujet à des erreurs de mesure et de mise à l'échelle, il est donc probable que l'information basse fréquence qu'il contient soit fiable, mais il est possible de s'interroger sur la confiance à accorder à ses variations hautes fréquences. D'autre part, le modèle d'impédance recherché est considéré "proche" d'un milieu 1D, cependant il possède des hétérogénéités latérales qui contredisent l'information *a priori* introduite par le terme de régularisations $R_2(I)$. Celui-ci tend en effet à minimiser les variations latérales du modèle d'impédance $I(x, z)$. Finalement, les CP correspondant aux données recombinées par "slant-stacking" présentent une certaine régularité suivant les pentes γ_s , cependant, ces données ne correspondent qu'à la propagation de pseudo-ondes planes, le champ de vitesse dans la zone de couverture $\mathbb{R} \times [0, (z_r)_1]$ présente des variations latérales, et les réflexions sur le modèle d'impédance pour des angles de propagation non nuls produisent des ondes se propageant dans la direction de propagation opposée de celle que l'on désire imposer. L'information *a priori* sur les CP est donc particulièrement sujette à caution, même si elle permet d'exprimer les tendances générales du modèle.

En conséquence, un réglage fin et précis des poids de régularisation $\varepsilon_i, i = 1, \dots, 3$ est nécessaire pour parvenir à déterminer une solution satisfaisant les critères de régularité attendus et minimisant la fonction coût sismique $J(I, h_1, \dots, h_S)$. Un compromis entre l'information sismique et l'information *a priori* doit ainsi être trouvé. Dans la pratique, ces poids sont réglés par une méthode essai-erreur. Leur validation repose sur l'analyse du résultat, et la détection de conflits entre l'information sismique et l'information de régularité, à partir des variations des valeurs du coût sismique et du coût de régularisation au cours de l'inversion.

Conclusion

Le principal problème posé par la définition du problème 2D d'inversion non linéaire de données de puits est l'indétermination de la solution. Pour surmonter cette difficulté, tout d'abord, l'information sismique est multipliée, par rapport au problème 1D, par l'utilisation de plusieurs sources placées à différentes distances de la tête du puits. De plus, la composante horizontale du champ de vitesse est enregistrée au niveau des récepteurs, en plus de la composante verticale, seule information enregistrée dans le cas 1D.

Cependant, cette inflation de l'information sismique ne suffit pas à compenser celle du nombre d'inconnues. Le problème direct reste insensible aux variations des paramètres éloignés latéralement de l'axe du puits, et l'extension 2D fait apparaître un redoutable problème d'indétermination lié au couplage entre les inconnues d'impédance $I(x, z)$ et celles liées aux CP. Pour lutter contre cette indétermination, l'approche est restreinte à l'imagerie de milieux proches de milieux 1D, présentant de faibles variations latérales. D'une part, cette hypothèse est cohérente avec celle faite sur la connaissance du champ de vitesse de propagation des ondes acoustiques dans le milieu imaginé. En effet, les méthodes utilisées pour déterminer cette dernière ne permettent pas d'obtenir une information très précise sur le champ de vitesse éloigné latéralement pour des milieux fortement 2D. D'autre part, cette hypothèse permet d'introduire directement une information *a priori* sur l'extension latérale du modèle d'impédance et elle fournit les conditions nécessaires à la recombinaison des données par "slant-stacking" afin de se rapporter à un éclairage du milieu par ondes planes, pour lequel une information *a priori* sur la régularité latérale des CP est disponible.

L'introduction de ces critères de régularité permet de restreindre efficacement l'espace des solutions. Pour cela, il est néanmoins nécessaire de procéder à un réglage soigné des poids de régularisation $\varepsilon_i, i = 1, \dots, 3$ dans la fonctionnelle régularisée $J^\varepsilon(I, h_1, \dots, h_S)$, l'information *a priori* utilisée n'étant pas une information exacte. D'autre part, même si l'espace des solutions est restreint, l'indétermination due au couplage des inconnues reste forte. Ceci va induire l'édition d'une méthode de résolution numérique spécifique au problème, qui, dans sa manière de construire la solution du problème inverse, permettra de lutter efficacement contre cette partie de l'indétermination inhérente au problème inverse 2D. La définition de cette méthode numérique fait l'objet du chapitre suivant.

Chapitre 3

Résolution numérique du problème inverse

Introduction

La résolution du problème d'inversion 2D de données sismiques de puits, tel qu'il a été défini dans la partie précédente, est un problème coûteux, tant d'un point de vue du nombre d'opérations à réaliser, que du point de vue de la taille des données manipulées. Deux facteurs en sont principalement responsables. Le premier tient à l'application de la méthode : la volonté d'imager le sous-sol à haute résolution implique de décrire celui-ci finement au niveau discret. Pour le cas d'application considéré dans le chapitre 4, le domaine sur lequel le modèle d'impédance est défini est un rectangle de 720 m de large sur 3400 m de profondeur, et le temps d'enregistrement T est fixé à 2 s. Les pas de discrétisation choisis,

$$\Delta x = 12 \text{ m}, \Delta z = 8 \text{ m}, \Delta t = 0.9 \times 10^{-3} \text{ s},$$

pour modéliser la propagation des ondes acoustiques, sont petits en regard de la taille du domaine et du temps d'enregistrement, et de ceci découle un nombre très important d'inconnues (plusieurs centaines de milliers). Le deuxième facteur tient à la nature du problème : c'est un problème d'optimisation pour lequel l'évaluation du critère à minimiser requiert la résolution d'un problème de propagation d'ondes dans un milieu 2D finement discrétilisé. Ce problème d'optimisation devant être résolu avec une grande précision, le nombre d'évaluations du critère et de son gradient dépasse la centaine de milliers. Les méthodes numériques utilisées sont donc choisies dans un souci constant de minimiser le temps de calcul et le volume des données. Ce chapitre est consacré à la description de ces méthodes et de leur mise en oeuvre numérique.

Dans une première partie, une définition détaillée du problème numérique, ainsi que les notations utilisées dans tout le chapitre sont données. Le problème inverse dont la résolution est implémentée diffère en effet quelque peu du problème inverse défini dans le chapitre précédent. Dans celui-ci, le problème direct est résolu sur

le demi-espace. Cependant, numériquement, la modélisation de la propagation des ondes ne peut se faire que sur un domaine borné. Ceci fait apparaître le problème du choix des conditions aux limites à imposer aux frontières de ce domaine fictif. En particulier, il faut éviter la création de réflexions parasites sur les bords de ce domaine, qui induiraient une erreur dans la modélisation de la propagation des ondes dans le sous-sol. Des couches absorbantes de type PML [Bérenger, 1994] sont donc utilisées, comme ceci est expliqué dans la première partie de ce chapitre. La résolution du problème inverse correspondant est ensuite décomposée en cinq parties. La première concerne le choix d'une méthode d'optimisation adaptée. L'objectif est de réduire l'indétermination de couplage entre les inconnues afin de présenter une convergence satisfaisante. La seconde et la troisième parties exposent la méthode utilisée pour calculer le gradient de la fonctionnelle minimisée, requis par la procédure d'optimisation. Cette méthode est celle de l'état adjoint, adaptée aux équations PML décrites dans la première partie. Tandis que la seconde partie présente la méthode dans un cadre continu, la troisième est consacrée à la présentation du schéma numérique utilisé pour résoudre le problème discret et à l'adaptation du calcul du gradient à ce cadre discret. La quatrième partie décrit ensuite les algorithmes employés pour une implémentation optimale de ce calcul. Finalement, la réduction efficace du temps de calcul de la méthode d'inversion numérique développée requiert l'utilisation d'une méthode de décomposition de domaines, décrite dans une cinquième partie. Elle permet de bénéficier de la puissance de calcul des machines à architecture parallèle.

3.1 Définition du problème numérique

3.1.1 Contexte et définition d'une couche PML

Soit le domaine dans lequel une estimation de l'impédance est recherchée,

$$\mathcal{O} = [-a, a] \times [z_{min}, z_{max}].$$

Les notations concernant les récepteurs introduites dans le chapitre précédent sont conservées. Le puits vertical est placé au milieu du domaine \mathcal{O} , en $x = 0$, et Q récepteurs sont placés dans ce puits. Leurs coordonnées sont notées

$$(x_r)_q, (z_r)_q, \quad 1 \leq q \leq Q.$$

Comme précédemment, le domaine de calcul commence à la profondeur du premier récepteur, si bien que

$$z_{min} = (z_r)_1.$$

De même, le domaine \mathcal{O} s'étend au delà du dernier récepteur, si bien que

$$z_{max} > (z_r)_Q.$$

Les données utilisées sont les vitesses de déplacement horizontal et vertical au niveau de ces récepteurs, enregistrées sur l'intervalle de temps $[0, T]$, à la suite de l'explosion de S sources situées en surface. Une recombinaison linéaire est ensuite opérée, comme expliqué au chapitre 2, de manière à obtenir S jeux de données indépendants, correspondant à la propagation de quasi-ondes planes suivant différents angles d'incidence. Un enregistrement au récepteur q dans le jeu de données s est noté

$$(D_{q,s}^x(t), D_{q,s}^z(t)) \in (L^2(0, T))^2,$$

l'indice x (respectivement z) faisant référence à la vitesse de déplacement horizontal, (respectivement à la vitesse de déplacement vertical). Un jeu entier de données est noté, comme au chapitre 2,

$$D_s(t) = (D_{1,s}^x(t), \dots, D_{Q,s}^x(t), D_{1,s}^z(t), \dots, D_{Q,s}^z(t)) \in (L^2(0, T))^{2Q}.$$

Le modèle de vitesse de propagation des ondes acoustiques $c(x, z)$ étant supposé connu, le problème direct consiste à calculer les vitesses de déplacement horizontal et vertical au niveau des récepteurs, à partir d'un modèle d'impédance $I(x, z)$ et d'une condition de pression à la profondeur du premier récepteur suffisamment régulière, notée $h(x, t)$ définie sur $[-a, a] \times [0, T]$, suivant les équations de l'acoustique. Cependant, le domaine \mathcal{O} est borné, à l'inverse du sous-sol, assimilé à un demi-espace d'extension latérale et de profondeur infinies. Ceci pose un problème évident : la résolution du problème direct sur \mathcal{O} peut entraîner l'apparition de réflexions sur ses frontières, qui viendraient perturber le signal enregistré au niveau des récepteurs, si les conditions au bord de \mathcal{O} sont mal choisies. Il est donc nécessaire d'utiliser des

conditions aux limites absorbantes. Parmi les nombreuses méthodes existantes, les couches absorbantes PML (Perfectly Matched Layers), très populaires depuis leur introduction au milieu des années 90 par Bérenger [Bérenger, 1994], sont retenues. Le principe de cette méthode est de borner le domaine de calcul par une couche d'épaisseur non nulle, qui absorbe les ondes. Dans le cas de milieux homogènes, cette absorption est réalisée quelque soit l'angle d'incidence des ondes sur la couche PML, ce qui empêche leur réflexion dans le domaine d'intérêt. La formulation employée ici est une adaptation aux équations d'Euler sans convection dans un milieu hétérogène des équations de Bérenger pour le mode transverse électrique 2D des ondes électromagnétiques. Le comportement de la solution de l'équation des ondes à l'intérieur du domaine s'approche donc de celui qui correspondrait à leur propagation dans un domaine d'extension infini.

Soit $l \in \mathbb{R}_+$ l'épaisseur de la couche PML, et

$$\Omega = [-(a + l), a + l] \times [z_{min}, z_{max} + l].$$

Le domaine Ω est le domaine sur lequel les équations PML sont résolues. La couche PML englobe le domaine d'intérêt \mathcal{O} au niveau de son extension latérale et en profondeur, mais pas au niveau du premier récepteur, où la condition de pression $h(x, t)$ est connue. Le modèle d'impédance $I(x, z)$ est prolongé par continuité dans les couches PML de la manière suivante

$$I(x, z) = I(-a, z), \quad (x, z) \in [-(a + l), -a] \times [z_{min}, z_{max}]$$

$$I(x, z) = I(a, z), \quad (x, z) \in [a, a + l] \times [z_{min}, z_{max}]$$

$$I(x, z) = I(x, z_{max}), \quad (x, z) \in [-a, a] \times [z_{max}, z_{max} + l]$$

$$I(x, z) = I(-a, z_{max}), \quad (x, z) \in [-(a + l), -a] \times [z_{max}, z_{max} + l]$$

$$I(x, z) = I(a, z_{max}), \quad (x, z) \in [a, a + l] \times [z_{max}, z_{max} + l].$$

Il en va de même pour le modèle de vitesse $c(x, z)$, qui est prolongé par continuité de la même manière. Cette prolongation par continuité est effectuée afin de ne pas introduire de discontinuités dans la couche PML, ce qui nuirait à son efficacité. Ces notions permettent d'aborder maintenant la formulation du problème direct en milieu PML.

3.1.2 Définition du problème direct

Soient

$$\Gamma_x = [-a - l, a + l], \quad \Gamma_z = [z_{min}, z_{max} + l], \quad \Gamma_t = [0, T]$$

$$\mathcal{Q} = \Omega \times \Gamma_t, \quad \Sigma_x = \Gamma_x \times \Gamma_t, \quad \Sigma_z = \Gamma_z \times \Gamma_t.$$

Donnons d'abord la formulation des équations d'Euler 2D linéarisées, sans terme de convection, décrivant la propagation des ondes acoustiques dans le domaine d'intérêt \mathcal{O} . Ces équations sont écrites, pour des raisons de commodité, en fonction de l'inverse de l'impédance $\alpha(x, z) \in C_m^\infty(\Omega)$. Dans toute la suite du problème, c'est cette quantité qui est considérée. Le problème inverse résolu consiste à estimer ce paramètre, plutôt que l'impédance. Ceci ne change rien au problème, néanmoins, cela permet une meilleure lisibilité des équations. Le système des équations d'Euler pour l'acoustique s'écrit donc

$$\begin{cases} \partial_t u_x - \alpha c \partial_x p = 0 \\ \partial_t u_z - \alpha c \partial_z p = 0 \\ \partial_t p - \frac{c}{\alpha} \partial_x u_x - \frac{c}{\alpha} \partial_z u_z = 0, \end{cases} \quad (3.1)$$

avec les condition aux limites

$$\begin{cases} p(-a, z, t) = 0, & p(x, z_{min}, t) = h(x, t), \\ p(a, z, t) = 0, & p(x, z_{max}, t) = 0. \end{cases} \quad (3.2)$$

La formulation PML de type Bérenger de ces équations est

$$\begin{cases} \partial_t u_x - \alpha c \partial_x (p_x + p_z) + c \sigma_x u_x = 0 \\ \partial_t u_z - \alpha c \partial_z (p_x + p_z) + c \sigma_z u_z = 0 \\ \partial_t p_x - \frac{c}{\alpha} \partial_x u_x + c \sigma_x p_x = 0 \\ \partial_t p_z - \frac{c}{\alpha} \partial_z u_z + c \sigma_z p_z = 0, \end{cases} \quad (3.3)$$

avec les conditions aux limites

$$\begin{cases} (p_x + p_z)(-(a + l), z, t) = 0, & (p_x + p_z)(x, z_{min}, t) = h(x, t) \\ (p_x + p_z)(a + l, z, t) = 0, & (p_x + p_z)(x, z_{max} + l, t) = 0. \end{cases} \quad (3.4)$$

Les coefficients d'absorption $\sigma_x(x)$ et $\sigma_z(z)$ sont définis par

$$\sigma_x(x) = \begin{cases} 0 & \text{si } x \in [-a, a] \\ -\sigma_{0x} \frac{(x+a)^3}{l} & \text{si } x \in [-(a+l), -a] \\ \sigma_{0x} \frac{(x-a)^3}{l} & \text{si } x \in [a, a+l] \end{cases}$$

$$\sigma_z(z) = \begin{cases} 0 & \text{si } z \in [z_{min}, z_{max}] \\ \sigma_{0z} \frac{(z-z_{max})^3}{l} & \text{si } z \in [z_{max}, z_{max}+l], \end{cases}$$

où σ_{0x}, σ_{0z} sont des constantes calibrées de manière empirique pour optimiser le comportement absorbant de la couche PML (cf chapitre 5). Ces coefficients d'absorption sont nuls dans le domaine d'intérêt \mathcal{O} et croissent polynomialement depuis la frontière de \mathcal{O} vers la frontière extérieure du domaine Ω .

Les équations PML résultent de la séparation de la variable de pression en deux inconnues p_x et p_z , telles que $p_x + p_z = p$. Ainsi, le système initial d'Euler est augmenté d'une nouvelle équation, et pour chacune d'entre elles un terme d'absorption est introduit. Comme ces coefficients d'absorption sont nuls dans le domaine d'intérêt \mathcal{O} , les équations PML (3.3) sont équivalentes aux équations originelles (3.1) sur \mathcal{O} .

Soit

$$\tilde{u}_x(t) = (u_x((x_r)_1, (z_r)_1, t), \dots, u_x((x_r)_Q, (z_r)_Q, t))$$

$$\tilde{u}_z(t) = (u_z((x_r)_1, (z_r)_1, t), \dots, u_z((x_r)_Q, (z_r)_Q, t)),$$

où u_x et u_z sont solution des équations (3.3) et (3.4). Alors le problème direct associé aux équations PML est représenté par l'opérateur

$$\widetilde{\mathcal{M}} : (\alpha, h) \longrightarrow (\tilde{u}_x(t), \tilde{u}_z(t)), \quad (3.5)$$

à valeur dans $(L^2(0, T))^{2Q}$.

Cet opérateur diffère donc de celui qui a été défini dans le chapitre 2 par la prise en compte d'un milieu borné, adapté à la résolution numérique des équations, entouré d'une couche PML, dans le but de ne pas introduire de réflexions parasites. Sa définition permet celle du problème inverse (P).

3.1.3 Définition du problème inverse (P)

Soit v le vecteur contenant toutes les inconnues du problème

$$v = (\alpha, h_1, \dots, h_S).$$

Cette notation est utilisée dans toute la suite de l'exposé. La fonctionnelle mesurant le "coût sismique", c'est à dire l'écart entre les données et les réponses aux récepteurs calculées par la résolution du problème direct pour un certain jeu d'inconnues v est la fonctionnelle

$$J(v) = \sum_{s=1}^S \|D_s - \widetilde{\mathcal{M}}(\alpha, h_s)\|_{(L^2(0,T))^{2Q}}^2.$$

Comme cela a été évoqué dans le chapitre précédent, il est nécessaire d'employer des termes de régularisation pour réduire l'indétermination inhérente au problème. Leur expression est

$$\left\{ \begin{array}{l} R_1(\alpha) = \|\partial_x \alpha\|_{L^2(\Omega)}^2 \\ R_2(\alpha) = \int_{(z_r)_1}^{(z_r)_Q} (\alpha((x_r)_1, z) - \alpha_0(z))^2 dz \\ R_3(h_1, \dots, h_S) = \sum_{s=1}^S \int_0^T \int_{-a-l}^{a+l} ((k_s)_x \partial_x h_s + (k_s)_t \partial_t h_s)^2 dx dt. \end{array} \right. \quad (3.6)$$

La fonctionnelle régularisée dont le minimum correspond à la solution du problème inverse défini dans le chapitre 2 est donc $J^\varepsilon(v)$ telle que

$$J^\varepsilon(v) = J(v) + \varepsilon_1 R_1(\alpha) + \varepsilon_2 R_2(\alpha) + \varepsilon_3 R_3(h_1, \dots, h_S)$$

Le problème inverse (P) est donc

$$(P) \quad \min_v J^\varepsilon(v).$$

Après discréttisation, suivant le cas d'application, le nombre d'inconnues discrètes peut atteindre plusieurs centaines de milliers et le nombre de données discrètes plusieurs millions. Ce problème inverse est donc en pratique un problème de grande taille, coûteux en temps de calcul, qui nécessite pour sa résolution l'emploi de méthodes numériques adaptées. Ces méthodes sont présentées dans les parties suivantes, en commençant par la présentation de la méthode d'optimisation employée.

3.2 Méthode d'optimisation

3.2.1 Algorithmes de Newton et Quasi-Newton

Après discréttisation, le problème (P), défini dans la partie précédente, est un problème d'optimisation non linéaire, différentiable, et de grande taille. Les méthodes les plus utilisées pour résoudre ce type de problème sont basées sur l'algorithme Quasi-Newton, qui dérive lui-même de l'algorithme de Newton. Ces algorithmes appartiennent tous deux à la famille des algorithmes de descente, qui reposent sur le principe suivant. Soit $x \in \mathbb{R}^m$, $m \in \mathbb{N}$. Soit

$$\begin{aligned} f : & \quad x \rightarrow f(x) \\ & \mathbb{R}^m \rightarrow \mathbb{R}_+^*, \end{aligned}$$

telle que f est de classe \mathcal{C}^2 . Le problème suivant

$$\min_{x \in \mathbb{R}^m} f(x),$$

est résolu par un algorithme de descente par la construction d'une suite d'itérés x^k , à partir d'un point initial x^0 , telle que

$$x_{k+1} = x_k + a_k d_k, \quad f(x_{k+1}) < f(x_k).$$

La quantité d_k est la direction de descente, calculée à partir d'informations sur les dérivées de f , et a_k un réel positif qui correspond au pas effectué dans cette direction. Celui-ci peut-être calculé de manière exacte (recherche linéaire exacte), mais cette solution est généralement délaissée, car trop coûteuse, au profit des règles de Wolfe, qui permettent une convergence plus rapide.

Remarque. Il existe également une autre classe de méthodes, appelées méthodes de régions de confiance, qui permettent de déterminer directement le terme $a_k d_k$, sans dissocier le calcul de a_k et d_k [Gilbert and al., 2006].

Si f est convexe, ou localement convexe, cet algorithme converge vers un minimum local. La spécificité des algorithmes de descente tient à la manière dont cette direction est calculée. Dans le cas de l'algorithme de Newton, elle est calculée à partir du calcul du gradient $g_k = \nabla f(x_k)$ et de la matrice Hessienne $H_k = \nabla^2 f(x_k)$. Cet algorithme peut être présenté comme suit :

```

Data :  $x^0, r, N_{iter}$ 
Result :  $\underset{x}{\operatorname{argmin}} f(x)$ 
calculer  $g_0$ ;
calculer  $H_0$ ;
 $k = 0$ ;
while  $f(x) > r$  et  $k < N_{iter}$  do
| calculer la direction de descente  $d = -H^{-1}g$ ;
| calculer le pas  $a \in \mathbb{R}_+^*$  (recherche linéaire);
| calculer  $x = x + ad$ ;
| calculer  $g(x)$ ;
| calculer  $H(x)$ ;
|  $k = k + 1$ ;
end

```

Algorithme 1 : Algorithme de Newton

Cet algorithme correspond à la construction successive d'approximations quadratiques de la fonction $f(x)$, et la minimisation de ces approximations à chaque itération. Celles-ci utilisent le pas de Newton, calculé à partir de l'inverse de la Hessienne H^{-1} et du gradient g . Les itérations s'interrompent quand la fonctionnelle devient inférieure à un résidu r défini par l'utilisateur, ou lorsque le nombre d'itérations k dépasse le nombre maximal autorisé N_{iter} . Le pas de descente a_k est quant à lui calculé par les méthodes spécifiées plus haut. Couramment, il s'agit de l'implémentation des règles de Wolfe, qui définissent un bon compromis entre la détermination exacte du pas optimal dans la direction d_k , coûteuse en temps de calcul, et celle d'un pas non optimal, mais satisfaisant des critères de décroissance de la fonction coût, et plus rapidement réalisée. L'exposé détaillé de ces méthodes dépasse le cadre de ce rappel ; on peut néanmoins se référer à [Gilbert and al., 2006] pour des considérations plus exhaustives.

L'application d'un tel algorithme requiert le calcul de la Hessienne de la fonction coût à minimiser (ou directement de son inverse), ce qui s'avère très coûteux dans le cadre de problèmes complexes de grande taille. Une nouvelle classe d'algorithme, dénommée Quasi-Newton, a donc vu le jour, basée sur l'approche de l'algorithme de Newton, mais utilisant une approximation M_k de l'inverse de la Hessienne $(H_k)^{-1}$ pour calculer le pas de descente, plutôt que la Hessienne exacte. Cette approximation M_k est calculée sur la base du gradient $g_i, i = 1, \dots, k$ calculé au cours des précédentes itérations, à partir d'une approximation initiale M_0 . Cette approximation initiale, dans le cas où aucune information n'est disponible, peut être par exemple la matrice identité. Plusieurs formules de mise à jour de l'approximation M_k à partir du gradient g_k existent, la plus populaire d'entre elles étant la mise à jour BFGS, qui tient son nom des initiales de ses concepteurs (Broyden, Fletcher, Goldfarb, Shanno) [Shanno, 1970]. Cette formulation définit l'algorithme Quasi-Newton BFGS, et son expression est

$$M_{k+1} = M_k + \frac{(d_k d_k^T) (d_k y_k + y_k^T M_k y_k)}{(d_k^T y_k)^2} - \frac{M_k y_k d_k^T + d_k y_k^T M_k}{d_k^T y_k},$$

où

$$y_k = g_{k+1} - g_k.$$

Une adaptation de cet algorithme aux problèmes de très grande taille a ensuite été proposée, initialement par Nocedal [Nocedal and al., 1995], [Nocedal and al., 1997]. C'est la formulation l-BFGS ("l" pour "limited memory"). Cet algorithme n'utilise qu'un nombre limité m de gradients pour calculer les mises à jour BFGS de M . De plus, il ne construit pas explicitement l'approximation de l'inverse de la Hessienne M (matrice de grande taille pour un grand nombre de variables), mais stocke la seule information nécessaire au calcul du produit matrice vecteur $M_k g_k$. Ainsi, à l'itération k , la mise à jour de M_k est calculée à partir des gradients g_i , $i = k - m, \dots, k$. Ceci permet d'accélérer les calculs du pas de descente et de limiter l'espace mémoire requis. Communément, la valeur choisie pour m vaut 20, à comparer avec les milliers d'itérations qui peuvent être effectués dans le cadre de problème de grande taille, ce qui pour un algorithme BFGS classique implique la prise en compte d'autant de mises à jours de la matrice M_k . Cette formulation est donc particulièrement adaptée pour les problèmes de très grande taille. C'est l'option la plus généralement retenue pour traiter des problèmes d'optimisation non linéaires, sans contraintes, faisant appel à beaucoup d'inconnues, pour lesquels le gradient $g(x)$ de la fonction à minimiser $f(x)$ existe et peut être calculé.

L'utilisation de cet algorithme semble donc particulièrement recommandée pour la résolution du problème inverse (P). Cependant, comme cela a été évoqué au chapitre 2, l'indétermination due au couplage des inconnues (impédance et CP) dans la procédure d'inversion est très importante. L'utilisation directe de cet algorithme sur l'ensemble des inconnues v n'est en effet pas efficace, et il convient donc d'adapter l'usage qu'il est possible d'en faire au cas particulier du problème traité.

3.2.2 Optimisation emboîtée

3.2.2.1 Intérêt

Le couplage des variables d'impédance et des CP dans l'optimisation est source d'indétermination, comme cela a été évoqué dans la partie 2. De plus, l'utilisation d'une variable globale contenant des inconnues de nature physique, de sensibilité, et de grandeur radicalement différentes génère d'importants problèmes de conditionnements. Le calcul, en certains points, par différences finies, des valeurs propres de la Hessienne $\nabla^2 J^\varepsilon(v)$, fait apparaître des différences de plusieurs ordres de grandeur entre celles-ci. Ceci n'est en fait qu'une traduction numérique de l'indétermination :

la disparité des valeurs propres réflète la disparité de la sensibilité de la fonctionnelle par rapport à ses variables. L'application de l'algorithme Quasi-Newton l-BFGS à ce problème très mal conditionné se traduit par une convergence extrêmement lente. La possibilité de séparer les inconnues portant sur les CP de celles portant sur le modèle d'impédance dans la procédure d'optimisation présente donc le double intérêt de limiter l'indétermination de couplage et d'améliorer la vitesse de convergence de l'algorithme.

La structure particulière de la fonctionnelle $J^\varepsilon(v)$ favorise cette séparation. En effet, par la linéarité du problème direct par rapport aux CP, la fonctionnelle "coût systémique" $J(v)$ est quadratique par rapport aux CP. De plus, le terme de régularisation sur les CP introduit dans $J^\varepsilon(v)$ est lui-même quadratique en h . Ainsi la partie fortement non linéaire de la fonctionnelle porte exclusivement sur le modèle d'impédance. Ceci incite à résérer l'usage de la méthode Quasi-Newton l-BFGS décrite plus haut à la détermination du modèle $\alpha(x, z)$, tandis que les CP optimales peuvent être déterminées par l'utilisation d'une méthode de gradient conjugué. Le gradient conjugué est en effet la méthode la plus rapide pour la résolution de problèmes quadratiques. L'interprétation mathématique de cette séparation est présentée dans le paragraphe suivant.

3.2.2.2 Mise en oeuvre mathématique

Le problème d'optimisation (P) est équivalent au problème d'optimisation (P_2)

$$(P_2) \quad \min_{\alpha} \tilde{J}(\alpha),$$

avec
$$\begin{cases} \tilde{J}(\alpha) = J^\varepsilon(\alpha, \bar{h}(\alpha)), \\ \bar{h}(\alpha) = \arg \min_h J^\varepsilon(\alpha, h), \\ h = (h_1, \dots, h_S). \end{cases}$$

En notant formellement $\partial_i J^\varepsilon(\alpha, h_1, \dots, h_S)$ la dérivée partielle de $J^\varepsilon(\alpha, h_1, \dots, h_S)$ par rapport à sa i ème variable, $i = 1, \dots, S+1$, et ∂_h la dérivation partielle par rapport à l'ensemble des CP $h_s(x, t)$,

$$\tilde{J}'(\alpha) = \partial_1 J^\varepsilon(\alpha, \bar{h}) + \partial_h J^\varepsilon(\alpha, \bar{h}(\alpha)). \bar{h}'(\alpha)$$

Or, $\bar{h}(\alpha)$ est un minimum en h de $J^\varepsilon(\alpha, h)$, ce qui implique que $\partial_h J^\varepsilon(\alpha, \bar{h}(\alpha)) = 0$. Ainsi

$$\tilde{J}'(\alpha) = \partial_1 J^\varepsilon(\alpha, \bar{h}_1(\alpha), \dots, \bar{h}_S(\alpha)).$$

Cette quantité correspond à la première composante de la transposée du gradient $(\nabla J^\varepsilon(\alpha, \bar{h}(\alpha)))^T$. En outre, le calcul des CP optimales $\bar{h}(\alpha) = (\bar{h}_1(\alpha), \dots, \bar{h}_S(\alpha))$ peut s'effectuer par l'utilisation du gradient conjugué, qui requiert simplement la

connaissance des S composantes supplémentaires du gradient $\nabla J^\varepsilon(\alpha, h)$.

En effet, la fonctionnelle $J^\varepsilon(v)$ dépend de manière quadratique des CP, indépendamment les unes des autres. En posant

$$J_s(\alpha, h_s) = \|D_s(t) - \widetilde{\mathcal{M}}(\alpha, h_s)\|_{(L^2(0,T))^{2Q}}^2$$

$$(R_3)_s(h_s) = \int_0^T \int_{-a-l}^{a+l} ((k_s)_x \partial_x h_s + (k_s)_t \partial_t h_s)^2 dx dt,$$

il vient

$$J^\varepsilon(\alpha, h_1, \dots, h_S) = \varepsilon_1 R_1(\alpha) + \varepsilon_2 R_2(\alpha) + \sum_{s=1}^S J_s(\alpha, h_s) + \varepsilon_3 (R_3)_s(h_s).$$

La détermination de $\bar{h}(\alpha)$ consiste donc en la résolution des problèmes

$$\bar{h}_s(\alpha) = \arg \min_{h_s} J_s(\alpha, h_s) + \varepsilon_3 (R_3)_s(h_s), \quad s = 1, \dots, S.$$

Après discréétisation, une réécriture formelle de ce problème est possible, sous la forme

$$\bar{h}_s(\alpha) = \arg \min_{h_s} \|D_s - F(\alpha)h_s\|^2 + \varepsilon_3 \|Gh_s\|^2,$$

où $\|\cdot\|$ désigne la norme euclidienne, avec

$$F(\alpha)h_s = \widetilde{\mathcal{M}}(\alpha, h_s), \quad \|Gh_s\|^2 = (R_3)_s(h_s),$$

ce qui est équivalent à résoudre le système linéaire

$$(F(\alpha)^T F(\alpha) + \varepsilon_3 G^T G) h_s = F(\alpha)^T D_s.$$

La matrice $F(\alpha)^T F(\alpha) + \varepsilon_3 G^T G$ est symétrique positive par construction. Son caractère défini repose sur l'inversibilité des matrices F et G , c'est-à-dire,

$$\forall h_s, \quad F(\alpha)h_s = 0 \Rightarrow h_s = 0, \quad Gh_s = 0 \Rightarrow h_s = 0.$$

Cette propriété, relativement à la matrice $F(\alpha)$, découle du fait que le problème direct défini dans la partie précédente est un problème bien posé. La preuve en est apportée par l'analyse précise des équations PML effectuée dans le chapitre 5. Quant à la matrice G , cette propriété se vérifie automatiquement pour

$$(k_s)_x \neq 0, \text{ et } (k_s)_t \neq 0.$$

Remarque. *Dans le cas où l'un des deux termes est nul, cette propriété n'est*

plus garantie, ce qui pose problème pour la résolution par le gradient conjugué. Notamment, la régularisation employée pour la régularisation de la CP correspondant à la propagation d'une onde plane à incidence normale dans le milieu est telle que $(k_s)_t = 0$. Néanmoins, comme cela va être expliqué par la suite, le terme de régularisation $R_3(h_1, \dots, h_S)$ va être abandonné au profit d'une autre prise en compte de l'information a priori, ce qui permet de s'affranchir de cette difficulté.

L'implémentation du gradient conjugué requiert simplement le calcul du second membre $F(\alpha)^T D_s$ et des produits matrices vecteurs $(F(\alpha)^T F(\alpha) + \varepsilon_3 G^T G) h_s$, tels que

$$\begin{cases} F(\alpha)^T D_s = \partial_{h_s} J^\varepsilon(\alpha, h_1, \dots, h_{s-1}, 0, h_{s+1}, \dots, h_S) \\ (F(\alpha)^T F(\alpha) + \varepsilon_3 G^T G) h_s = \partial_{h_s} J^\varepsilon(\alpha, h_1, \dots, h_S) - F(\alpha)^T D_s, \end{cases}$$

La donnée du gradient $\nabla J^\varepsilon(\alpha, h_1, \dots, h_S)$ permet donc de mettre en place la procédure d'optimisation emboîtée. Son application est simple : l'algorithme d'optimisation de type quasi-Newton s'applique à la fonctionnelle $\tilde{J}(\alpha)$. A chaque itération de cet algorithme, le calcul des arguments $\bar{h}_1(\alpha), \dots, \bar{h}_S(\alpha)$ est effectué par la méthode du gradient conjugué, afin d'évaluer la fonction coût $\tilde{J}(\alpha)$ et son gradient $\tilde{J}'(\alpha)$. Ceci est résumé dans l'algorithme suivant.

```

Data :  $\alpha^0, h^0, r, N_{iter}$ 
Result :  $\operatorname{argmin}_v J^\varepsilon(v)$ 
 $\alpha = \alpha^0;$ 
for  $s = 1, \dots, S$  do
| calculer  $\bar{h}_s(\alpha)$  par gradient conjugué;
end
calculer  $g = \tilde{J}'(\alpha) = \partial_1 J^\varepsilon(\alpha, \bar{h}_1(\alpha), \dots, \bar{h}_S(\alpha));$ 
initialiser  $M = Id$ ;
 $k = 0;$ 
while  $\tilde{J}(\alpha) > r$  et  $k < N_{iter}$  do
| calculer la direction de descente  $d = -Mg$ ;
| calculer le pas  $a \in \mathbb{R}_+^*$  (recherche linéaire);
| calculer  $\alpha = \alpha + ad$ ;
for  $s = 1, \dots, S$  do
| calculer  $\bar{h}_s(\alpha)$  par gradient conjugué;
end
calculer  $g = \tilde{J}'(\alpha) = \partial_1 J^\varepsilon(\alpha, \bar{h}_1(\alpha), \dots, \bar{h}_S(\alpha));$ 
calculer  $M$  par mise à jour l-BFGS;
 $k = k + 1;$ 
end
```

Algorithme 2 : Algorithme d'optimisation emboîtée

Cet algorithme permet de découpler le calcul des différentes inconnues. Pour chaque itéré α_k , les conditions de pression optimales correspondantes

$$\bar{h}_{s,k}, \quad 1 \leq s \leq S,$$

sont calculées, ce qui permet de lever l'indétermination de couplage. Cependant, ceci implique une longue suite de résolutions de problèmes quadratiques par gradient conjugué. En effet, un tel problème est résolu pour chaque CP, et ceci à chaque itération de l'algorithme Quasi-Newton l-BFGS. Dans la pratique, il est donc important de contrôler la précision avec laquelle ces problèmes sont résolus. Si celle-ci est grande, le temps de calcul en est augmenté d'autant, car le nombre d'itérations de l'algorithme du gradient conjugué augmente, mais le calcul du gradient $\tilde{J}'(\alpha)$ est plus précis, et l'algorithme Quasi-Newton converge plus rapidement. A l'inverse, si cette précision est faible, les temps de calculs sont considérablement allégés, mais la précision du calcul de $\tilde{J}'(\alpha)$ est moins bonne, ce qui entraîne une convergence plus lente, voire un échec de l'algorithme Quasi-Newton. Il est donc nécessaire de trouver un bon compromis entre ces deux extrêmes. Celui-ci est déterminé par une méthode de type essai-erreur. Cependant, même une fois ce compromis établi, il est clair que la part la plus importante du temps de calcul est liée aux résolutions du gradient conjugué. Deux améliorations permettant d'accélérer ces résolutions ont donc été définies. La première permet d'utiliser des préconditionnements adéquats pour accélérer la convergence. La seconde consiste en la paramétrisation des CP par interpolation linéaire afin de réduire le nombre d'inconnues dans la procédure de gradient conjugué, ce qui conduit à la définition d'un problème inverse (P_3) voisin de (P), faisant intervenir un nombre réduit d'inconnues, donc plus rapide à résoudre, et n'utilisant pas le terme de régularisation $R_3(h_1, \dots, h_S)$.

3.2.2.3 Préconditionnement du gradient conjugué

La vitesse de convergence du gradient conjugué peut être accélérée par l'utilisation de préconditionneurs adéquats. Rappelons brièvement la définition d'un préconditionneur.

Soit un système linéaire de type $Ax = b$. La vitesse de convergence d'un solveur linéaire appliquée à ce système dépend du nombre de conditionnement de la matrice A , qui est égal au rapport de sa valeur singulière maximale sur sa valeur singulière minimale. Plus ce nombre est élevé, moins bien le système est conditionné, et plus longue sera sa résolution. Ce nombre de conditionnement peut être amélioré par la multiplication par une matrice de préconditionnement C , ou préconditionneur. Le système résolu est alors $CAx = Cb$.

La résolution du gradient conjugué intervenant un grand nombre de fois dans la procédure d'optimisation emboîtée, l'utilisation de préconditionneurs adéquats est particulièrement indiquée, afin d'accélérer la convergence. Aussi, l'implémentation retenue pour l'algorithme du gradient conjugué permet le calcul simultané de la solution du problème linéaire et d'un préconditionneur adapté à un problème voisin, $A'x = b'$ c'est à dire pour lequel la matrice A' et le second membre b' sont "proches" de A et b , en un certain sens. Dans le cas de l'optimisation emboîtée, cette hypothèse de "proximité" peut être retenue, de manière intuitive. En effet, à l'itération k de l'algorithme de Quasi-Newton, pour une CP

$$h_s(x, t), \quad 1 \leq s \leq S,$$

le gradient conjugué est utilisé pour résoudre le problème

$$(F(\alpha)^T F(\alpha) + \varepsilon_3 G^T G) h_s = F^T(\alpha^k) D_s.$$

Comme $\alpha_{k+1} = \alpha_k + a_k d_k$ est "proche" de α_k , pour un pas a_k suffisamment petit, le problème résolu par gradient conjugué à l'itération $k+1$ peut être supposé assez proche de celui résolu à l'itération k . Ceci pourrait laisser entendre cependant que la matrice $F(\alpha_k)$ varie continûment en fonction de α ce qui n'est vraisemblablement pas le cas. Aucune analyse mathématique n'est entreprise ici ; il semble cependant dans la pratique que les préconditionnements calculés à une itération k s'appliquent bien à l'itération $k+1$ et permettent d'accélérer la convergence du gradient conjugué. Ceci n'a donc pas valeur de démonstration, mais sert simplement à illustrer l'intuition que l'on peut avoir de la ressemblance de la succession de problèmes quadratiques résolus par gradient conjugué au cours des itérations de l'algorithme Quasi-Newton.

3.2.2.4 Paramétrisation des CP et définition du problème inverse (P_3)

Il a été montré dans le chapitre précédent qu'une recombinaison linéaire des données était effectuée afin d'imposer une structure particulière aux CP dans le plan (x, t) , qui permet d'introduire une information *a priori* sur leur régularité. Celle-ci se traduit par de faibles variations suivant une direction $\gamma_s \in \mathbb{R}^2$, $1 \leq s \leq S$, dans le plan (x, t) , connue. Cette information est pour l'instant prise en compte par un terme de régularisation qui permet de minimiser les variations des CP suivant la direction γ_s qui leur correspond, ce qui, on l'a vu, peut mener à la définition d'un problème non convexe pour le gradient conjugué, si la direction de régularisation est parallèle à l'un des deux axes du plan (x, t) . Cependant, elle peut être prise en compte différemment, en utilisation une paramétrisation adéquate des CP qui permet de réduire le nombre d'inconnues associée à chaque CP. En effet, chacune d'entre elles peut être paramétrée par la sélection d'un ensemble de p traces en temps,

$$h_s(x_{i_k}, t), \quad \text{pour } k = 1, \dots, p, \quad p \in \mathbb{N}.$$

Les CP sont ensuite calculées par interpolation linéaire de cet ensemble de traces suivant la direction γ_s . Ainsi, le nombre d'inconnues associées à la reconstruction des CP diminue, ce qui a pour effet de réduire à la fois l'indétermination

latérale et d'accélérer la convergence du gradient conjugué. Ceci permet également de se dispenser du terme de régularisation sur les CP, $R_3(h_1, \dots, h_S)$, l'information *a priori* étant prise en compte par l'interpolation. Un nouveau terme peut-être éventuellement introduit, pour aider à la stabilisation du problème, mais afin de clarifier l'exposé, l'hypothèse sera faite dans la suite que l'interpolation suffit à prendre en compte l'information *a priori* disponible sur les CP. Pour plus de détails, les lecteurs se reporteront à l'annexe 1 de ce chapitre.

Soit

$$q_s(t) = (h_s(x_{i_1}, t), \dots, h_s(x_{i_p}, t)), \quad (3.7)$$

le vecteur des traces sélectionnées. Soient $N \in \mathbb{N}$ le nombre de pas de temps et $K \in \mathbb{N}$ le nombre de pas d'espace dans la direction x obtenus après discréétisation. Alors

$$q_s \in \mathbb{R}^{Np}, \quad h_s \in \mathbb{R}^{NK}.$$

L'interpolation consiste en la construction d'une matrice $B_s \in \mathbb{M}_{NK, Np}$ s'exprimant en fonction de la direction γ_s telle que

$$h_s = B_s q_s.$$

La détermination des matrices B_s est donnée dans l'annexe 2 de ce chapitre. Soit alors la fonctionnelle $\widehat{J}^\varepsilon(\alpha, q_1, \dots, q_S)$ telle que

$$\widehat{J}^\varepsilon(\alpha, q_1, \dots, q_S) = J(\alpha, B_1 q_1, \dots, B_S q_S) + \varepsilon_1 R_1(\alpha) + \varepsilon_2 R_2(\alpha).$$

Ceci permet de définir le problème inverse

$$(P_3) \quad \min_{\alpha, q_1, \dots, q_S} \widehat{J}^\varepsilon(\alpha, q_1, \dots, q_S).$$

Ce problème peut être résolu par la procédure d'optimisation emboîtée, tout comme (P) , car $\widehat{J}^\varepsilon(\alpha, q_1, \dots, q_S)$ dépend de manière quadratique des inconnues q_s , $1 \leq s \leq S$. L'intérêt de la résolution de ce problème est la diminution de la taille des inconnues $q_s(t)$ après discréétisation par rapport à celle des CP, qui permet une convergence plus rapide de l'algorithme du gradient conjugué.

Ceci requiert cependant de pouvoir déterminer les dérivées $\partial_{q_s} \widehat{J}^\varepsilon(\alpha, q_1, \dots, q_S)$. Or, ces dérivées s'expriment très facilement en fonction de $\partial_{h_s} J(\alpha, h_1, \dots, h_S)$ et des matrices B_s . En effet, en utilisant les notations discrètes introduites dans le paragraphe précédent, il vient

$$\begin{aligned} \partial_{q_s} \widehat{J}^\varepsilon(\alpha, q_1, \dots, q_S) &= \partial_{q_s} \|F(\alpha) B_s q_s - D_s\|^2 = B_s^T F(\alpha)^T F(\alpha) B_s q_s - B_s^T F(\alpha)^T D_s = \\ &= B_s^T (F^T(\alpha) F(\alpha) h_s - F(\alpha)^T D_s) = B_s^T \partial_{h_s} J(\alpha, h_1, \dots, h_S). \end{aligned}$$

D'autre part

$$\partial_1 \widehat{J}^\varepsilon(\alpha, \bar{q}_1(\alpha), \dots, \bar{q}_S(\alpha)) = \partial_1 J(\alpha, \bar{h}_1(\alpha), \dots, \bar{h}_S(\alpha)) + \varepsilon_1 R'_1(\alpha) + \varepsilon_2 R'_2(\alpha),$$

avec

$$\bar{h}_s(\alpha) = B_s \bar{q}_s(\alpha), \quad s = 1, \dots, S.$$

Ainsi la connaissance des dérivées partielles

$$\partial_s J(\alpha, h_1, \dots, h_S), \quad \text{pour } s = 1, \dots, S+1,$$

ainsi que celle des dérivées des termes de régularisation $R'_1(\alpha)$ et $R'_2(\alpha)$, et la connaissance des matrices d'interpolation B_s , suffisent pour résoudre le problème (P_3) par la méthode d'optimisation emboîtée.

Cette interpolation consiste en réalité à réduire le nombre de degrés de liberté sur lesquels l'algorithme d'optimisation peut jouer, et donc en ce sens à réduire l'indétermination du problème, tout en accélérant la convergence. Les degrés de liberté choisis correspondent aux traces sélectionnées. Celles-ci doivent être cependant suffisamment nombreuses pour permettre à l'algorithme de calculer des CP assez proches de celles recherchées. Si cela n'est pas le cas, les dérivées

$$\partial_1 \widehat{J}^\varepsilon(\alpha, \bar{q}_1(\alpha), \dots, \bar{q}_S(\alpha)),$$

deviennent trop approximatives pour permettre à l'algorithme de Quasi-Newton de converger rapidement. On retrouve à nouveau le compromis entre une résolution fine ou une résolution grossière du problème de calcul des CP optimales, qui se traduit dans le premier cas par une bonne convergence de Quasi-Newton mais des temps de calcul importants, et dans le deuxième cas par une mauvaise convergence de Quasi-Newton mais des temps de calcul bien moindres.

3.2.2.5 Algorithme d'optimisation emboîtée final

L'algorithme suivant résume la méthode d'optimisation mise en place pour résoudre le problème inverse (P_3).

```

Data :  $\alpha^0, q_1^0, \dots, q_S^0, r, N_{iter}$ 
Résultat :  $\arg \min_{\alpha, q_1, \dots, q_S} \widehat{J}(\alpha, q_1, \dots, q_S)$ 
 $\alpha = \alpha^0;$ 
for  $s = 1, \dots, S$  do
    | calculer  $\bar{q}_s(\alpha)$  par gradient conjugué sans préconditionnement;
    | calculer le préconditionnement  $C_s$ ;
    | calculer la CP entière  $\bar{h}_s = B_s \bar{q}_s$ ;
end
calculer  $g = \partial_1 J(\alpha, \bar{h}_1, \dots, \bar{h}_S) + R'_1(\alpha) + R'_2(\alpha)$ ;
initialiser  $M = Id$ ;
 $k = 0$ ;
while  $\widehat{J}^\varepsilon(\alpha^k, \bar{q}_1, \dots, \bar{q}_S) > r$  et  $k < N_{iter}$  do
    | calculer la direction de descente  $d = -Mg$ ;
    | calculer le pas  $a \in \mathbb{R}_+^*$  (recherche linéaire);
    | calculer  $\alpha = \alpha + ad$ ;
    for  $s = 1, \dots, S$  do
        | calculer  $\bar{q}_s(\alpha)$  par gradient conjugué avec préconditionnement  $C_s$ ;
        | calculer nouveau préconditionnement  $C_s$  ;
        | calculer CP entière  $\bar{h}_s = B_s \bar{q}_s$ ;
    end
    calculer  $g = \partial_1 J(\alpha, \bar{h}_1, \dots, \bar{h}_S) + R'_1(\alpha) + R'_2(\alpha)$ ;
    calculer  $M$  par mise à jour l-BFGS;
     $k = k + 1$ ;
end
```

Algorithm 3 : Algorithme d'optimisation emboîtée

La principale caractéristique de cet algorithme est l'utilisation de l'algorithme du gradient conjugué au sein de la boucle d'optimisation dévolue à l'algorithme Quasi-Newton. Ceci est utile à la réduction de l'indétermination et l'accélération de la convergence de la méthode. En contrepartie, un grand nombre de problèmes quadratiques sur les CP sont résolus par gradient conjugué au cours des itérations Quasi-Newton l-BFGS. Il est donc nécessaire d'optimiser la convergence du gradient conjugué sur ces problèmes afin de réduire les temps de calcul. Ceci est effectué via le calcul de préconditionnements successifs et la paramétrisation des CP, qui conduit à résoudre le problème inverse (P_3), qui fait intervenir un nombre moins important d'inconnues que le problème (P). Pour l'implémentation numérique de cette méthode d'optimisation, les algorithmes utilisés sont, pour la partie Quasi-Newton l-BFGS, l'algorithme l-BFGS-b [Nocedal and al., 1995], [Nocedal and al., 1997], et pour la partie gradient conjugué, l'algorithme N1CG1 [Gilbert, 1999].

Comme cela a été montré, la résolution du problème (P_3) ne nécessite que le calcul des dérivées

$$\nabla J(v) = (\partial_\alpha J(v), \partial_{h_1} J(v), \dots, \partial_{h_S} J(v)), \quad R'_1(\alpha), \quad R'_2(\alpha),$$

les matrices d'interpolation $B_s, s = 1, \dots, S$, étant connues.

La partie suivante s'attache au calcul de ces dérivées. Les dérivées des termes de régularisation $R_1(\alpha)$ et $R_2(\alpha)$ sont assez facilement déterminées. En revanche, le calcul du gradient $\nabla J(v)$ ne peut être fait de manière directe, en raison des coûts de calcul qui seraient alors engendrés. Il est déterminé par la méthode de l'état adjoint, qui est exposée dans la partie suivante.

3.3 Calcul de $\nabla J(v)$, $R'_1(\alpha)$ et $R'_2(\alpha)$ dans un formalisme continu

Les calculs de $\nabla J(v)$, $R'_1(\alpha)$, $R'_2(\alpha)$ sont décomposés en deux parties. Le cœur du problème est la détermination du gradient de la fonctionnelle coût sismique $\nabla J(v)$. Ceci est effectué par la méthode de l'état adjoint, qui permet de déterminer $\nabla J(v)$ par la définition d'un système d'équations adjoint au système (3.3), (3.4). Le calcul du gradient des termes de régularisation est ensuite exposé.

Les termes

$$\partial_\alpha J(v), \quad R'_1(\alpha), \quad R'_2(\alpha),$$

sont calculés pour le produit scalaire $(., .)_{L^2(\Omega)}$, tandis que les termes

$$\partial_{h_s} J(v), \quad 1 \leq s \leq S,$$

sont calculés pour le produit scalaire $(., .)_{L^2(\Sigma_x)}$.

3.3.1 Calcul du gradient de la fonctionnelle coût sismique J

3.3.1.1 Réécriture des équations d'Euler PML

Dans un premier temps, une réécriture des équations PML sous la forme d'un système est donnée. Soit $U = (u_x, u_z, p_x, p_z) \in L^2(\mathcal{Q})$,

$$A = \begin{pmatrix} 0 & 0 & \alpha c & \alpha c \\ 0 & 0 & 0 & 0 \\ \frac{c}{\alpha} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad B = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & \alpha c & \alpha c \\ 0 & 0 & 0 & 0 \\ 0 & \frac{c}{\alpha} & 0 & 0 \end{pmatrix}, \quad S = \begin{pmatrix} c\sigma_x & 0 & 0 & 0 \\ 0 & c\sigma_z & 0 & 0 \\ 0 & 0 & c\sigma_x & 0 \\ 0 & 0 & 0 & c\sigma_z \end{pmatrix}, \quad (3.8)$$

alors les équations (3.3) et (3.4) se récrivent

$$\partial_t U - A(\alpha) \partial_x U - B(\alpha) \partial_z U + SU = 0, \quad (3.9)$$

avec

$$\begin{cases} (U_3 + U_4)(x, z_{min}, t) = h(x, t), & (U_3 + U_4)(x, z_{max}, t) = 0, \\ (U_3 + U_4)(-a - l, z, t) = 0, & (U_3 + U_4)(a + l, z, t) = 0, \\ U(x, z, 0) = 0. \end{cases} \quad (3.10)$$

3.3.1.2 Simplification du problème

La fonctionnelle coût sismique $J(v)$ peut être décomposée en S sous-fonctionnelles

$$J_s(\alpha, h_s), \quad 1 \leq s \leq S,$$

correspondant chacune à un jeu de données D_s et une CP h_s ,

$$\begin{cases} J_s(\alpha, h_s) = \|D_s - \tilde{\mathcal{M}}(\alpha, h_s)\|_{(L^2(0,T))^{2Q}}^2, \\ J(v) = \sum_{s=1}^S J_s(\alpha, h_s). \end{cases}$$

Ainsi, le gradient de $J(v)$ se déduit facilement de celui des fonctionnelles $J_s(\alpha, h_s)$. En effet,

$$\begin{cases} \partial_\alpha J(v) = \sum_{s=1}^S \partial_\alpha J_s(\alpha, h_s), \\ \partial_{h_s} J(v) = \partial_{h_s} J_s(\alpha, h_s), \quad 1 \leq s \leq S. \end{cases} \quad (3.11)$$

La méthode de l'état adjoint va être appliquée au calcul du gradient d'une sous-fonctionnelle particulière, $J_1(\alpha, h_1)$. Afin de simplifier les notations, on considère la fonctionnelle $\check{J}(\alpha, h) = J_1(\alpha, h_1)$ dont l'expression est

$$\check{J}(\alpha, h) = \sum_{q=1}^Q \| (D_{1,q}^x(t) - U_1((x_r)_q, (z_r)_q, t)) \|_{L^2(0,T)}^2 + \| (D_{1,q}^z(t) - U_2((x_r)_q, (z_r)_q, t)) \|_{L^2(0,T)}^2,$$

avec $U(x, z, t)$ solution des équations (3.9) et (3.10), pour la CP $h(x, t) = h_1(x, t)$.

3.3.1.3 Méthode de l'état adjoint

Soit le problème d'optimisation

$$(P_4) \quad \min_{\alpha, h} \check{J}(\alpha, h).$$

Le Lagrangien associé au problème (P_4) est défini en prenant la condition " $U(x, z, t)$ solution des équations (3.9) et (3.10)" , qui entre dans la définition de $\check{J}(\alpha, h)$, comme contrainte du problème de minimisation (P_4) . Soit

$$G(U) = \sum_{q=1}^Q \| (D_{1,q}^x(t) - U_1((x_r)_q, (z_r)_q, t)) \|_{L^2(0,T)}^2 + \| (D_{1,q}^z(t) - U_2((x_r)_q, (z_r)_q, t)) \|_{L^2(0,T)}^2,$$

pour U quelconque.

Soient également

$$\lambda \in (L^2(\mathcal{Q}))^4, \mu_1 \in L^2(\Sigma_x), \mu_2 \in L^2(\Sigma_x), \mu_3 \in L^2(\Sigma_z), \mu_4 \in L^2(\Sigma_z), \mu_5 \in (L^2(\Omega))^4,$$

les multiplicateurs de Lagrange associés aux contraintes. On note $\nu = (\lambda, \mu_1, \dots, \mu_5)$. Alors, le Lagrangien $L(\alpha, h, U, \nu)$ associé au problème de minimisation (P_4) , s'écrit

$$\begin{aligned}
L(\alpha, h, U, \nu) &= G(U) + (\lambda, \partial_t U - A(\alpha)\partial_x U - B(\alpha)\partial_z U + SU)_{(L^2(\mathcal{Q}))^4} \\
&\quad + (\mu_1, U_3(x, z_{min}, t) + U_4(x, z_{min}, t) - h)_{L^2(\Sigma_x)} \\
&\quad + (\mu_2, U_3(x, z_{max} + l, t) + U_4(x, z_{max} + l, t))_{L^2(\Sigma_x)} \\
&\quad + (\mu_3, U_3(-a - l, z, t) + U_4(-a - l, z, t))_{L^2(\Sigma_z)} \\
&\quad + (\mu_4, U_3(a + l, z, t) + U_4(a + l, z, t))_{L^2(\Sigma_z)} \\
&\quad + (\mu_5, U(x, z, 0))_{(L^2(\Omega))^4}.
\end{aligned} \tag{3.12}$$

Le gradient $\nabla \check{J}(\alpha, h)$ peut s'exprimer en fonction du lagrangien $L(\alpha, h, U, \nu)$. En effet, pour $\overline{U}(\alpha, h)$ solution de (3.9), (3.10)

$$\left\{
\begin{array}{l}
L(\alpha, h, \overline{U}(\alpha, h), \nu) = \check{J}(\alpha, h) \\
\partial_\alpha L(\alpha, h, \overline{U}(\alpha, h), \nu) = \partial_\alpha \check{J}(\alpha, h) \\
\partial_h L(\alpha, h, \overline{U}(\alpha, h), \nu) = \partial_h \check{J}(\alpha, h).
\end{array}
\right.$$

Or, en notant $\partial_i L$, $i = 1, \dots, 4$, les dérivées partielles de $L(\alpha, h, U, \nu)$ par rapport à ses arguments i ,

$$\left\{
\begin{array}{l}
\partial_\alpha L(\alpha, h, \overline{U}(\alpha, h), \nu) = \partial_3 L(\alpha, h, \overline{U}(\alpha, h), \nu). \partial_\alpha \overline{U}(\alpha, h) + \partial_1 L(\alpha, h, \overline{U}(\alpha, h), \nu) \\
\partial_h L(\alpha, h, \overline{U}(\alpha, h), \nu) = \partial_3 L(\alpha, h, \overline{U}(\alpha, h), \nu). \partial_h \overline{U}(\alpha, h) + \partial_2 L(\alpha, h, \overline{U}(\alpha, h), \nu).
\end{array}
\right.$$

Ainsi, pour $\overline{\nu}$ solution de

$$\forall (\alpha, h), \quad \partial_3 L(\alpha, h, \overline{U}(\alpha, h), \overline{\nu}) = 0, \tag{3.13}$$

$$\text{alors } \left\{
\begin{array}{l}
\partial_\alpha \check{J}(\alpha, h) = \partial_1 L(\alpha, h, \overline{U}(\alpha, h), \overline{\nu}) \\
\partial_h \check{J}(\alpha, h) = \partial_2 L(\alpha, h, \overline{U}(\alpha, h), \overline{\nu}).
\end{array}
\right. \tag{3.14}$$

La définition de $\overline{\nu}$ permet donc de donner une expression simple du gradient $\nabla \check{J}(\alpha, h)$ ne faisant pas intervenir les termes d'évaluation coûteuse

$$\partial_\alpha \widetilde{\mathcal{M}}(\alpha, h), \quad \partial_h \widetilde{\mathcal{M}}(\alpha, h),$$

comme l'aurait fait un calcul direct du gradient.

3.3.1.4 Expression du gradient en fonction de l'état adjoint

Le développement de l'expression (3.14) donne, pour $\tilde{\alpha} \in L^2(\Omega)$,

$$\begin{aligned} \partial_1 L(\alpha, h, \bar{U}(\alpha, h), \bar{\nu}).\tilde{\alpha} &= \int_0^T (\bar{\lambda}, -\tilde{\alpha} A'(\alpha) \partial_x \bar{U}(\alpha, h) - \tilde{\alpha} B'(\alpha) \partial_z \bar{U}(\alpha, h))_{(L^2(\Omega))^4} dt \\ &\quad - \int_0^T (\bar{\lambda}_1, \tilde{\alpha} c \partial_x (\bar{U}_3 + \bar{U}_4))_{L^2(\Omega)} dt + \int_0^T \left(\bar{\lambda}_3, \frac{\tilde{\alpha} c}{\alpha^2} \partial_x \bar{U}_1 \right)_{L^2(\Omega)} dt \\ &\quad - \int_0^T (\bar{\lambda}_2, \tilde{\alpha} c \partial_z (\bar{U}_3 + \bar{U}_4))_{L^2(\Omega)} dt + \int_0^T \left(\bar{\lambda}_4, \frac{\tilde{\alpha} c}{\alpha^2} \partial_z \bar{U}_2 \right)_{L^2(\Omega)} dt, \end{aligned} \quad (3.15)$$

d'après la définition des matrices A et B données en (3.8). Aussi,

$$\partial_\alpha \check{J}(\alpha, h) = - \int_0^T \bar{\lambda}_1 c \partial_x (\bar{p}_x + \bar{p}_z) + \bar{\lambda}_2 c \partial_z (\bar{p}_x + \bar{p}_z) dt + \int_0^T \bar{\lambda}_3 \frac{c}{\alpha^2} \partial_x \bar{u}_x + \bar{\lambda}_4 \frac{c}{\alpha^2} \partial_z \bar{u}_z dt. \quad (3.16)$$

D'autre part, $\tilde{h} \in L^2(\Sigma_x)$,

$$\partial_2 L(\alpha, h, \bar{U}(\alpha, h), \bar{\nu}).\tilde{h} = (\bar{\mu}_1, -\tilde{h})_{L^2(\Sigma_x)},$$

donc

$$\partial_h \check{J}(\alpha, h) = -\bar{\mu}_1(x, t).$$

3.3.1.5 Définition du système adjoint

Il reste maintenant à déterminer le système adjoint dont $\bar{\nu}$ est solution. Ce système est dérivé de l'équation (3.13) de la manière suivante. Pour \tilde{U} dans $L^2(\mathcal{Q})^4$,

$$\partial_3 L(\alpha, h, U, \lambda).\tilde{U} = I_1(\alpha, h, U).\tilde{U} + I_2(\alpha, h, U, \nu).\tilde{U} + I_3(\alpha, h, U, \nu).\tilde{U},$$

avec

$$\begin{aligned} I_1(\alpha, h, U).\tilde{U} &= G'(U).\tilde{U} = 2 \sum_{q=1}^Q \left(U_1((x_r)_q, (z_r)_q, t) - D_q^x(t), \tilde{U}_1 \right)_{L^2(\Gamma_t)} + \\ &\quad 2 \sum_{q=1}^Q \left(U_2((x_r)_q, (z_r)_q, t) - D_q^z(t), \tilde{U}_2 \right)_{L^2(\Gamma_t)} \\ &= (f^x, \tilde{U}_1)_{L^2(\mathcal{Q})} + (f^z, \tilde{U}_2)_{L^2(\mathcal{Q})}, \end{aligned}$$

où

$$\begin{cases} f_x(x, z, t) = 2 \sum_{q=1}^Q (U_1(x, z, t) - D_q^x(t)) \delta_{Mr_q} \\ f_z(x, z, t) = 2 \sum_{q=1}^Q (U_2(x, z, t) - D_q^z(t)) \delta_{Mr_q}, \end{cases}$$

et $\delta_{Mr_q} = \delta((x_r)_q, (z_r)_q)$, avec $\delta(x, z)$ fonction de Dirac définie sur \mathbb{R}^2 , et

$$\left\{ \begin{array}{l} I_2(\alpha, h, U, \nu). \tilde{U} = \left(\lambda, \partial_t \tilde{U} - A \partial_x \tilde{U} - B \partial_z \tilde{U} + S \tilde{U} \right)_{(L^2(\mathcal{Q}))^4}, \\ \\ I_3(\alpha, h, U, \nu). \tilde{U} = \left(\mu_1, \tilde{U}_3(x, z_{min}, t) + \tilde{U}_4(x, z_{min}, t) \right)_{L^2(\Sigma_x)} + \\ \quad \left(\mu_2, \tilde{U}_3(x, z_{max} + l, t) + \tilde{U}_4(x, z_{max} + l, t) \right)_{L^2(\Sigma_x)} + \\ \quad \left(\mu_3, \tilde{U}_3(-a - l, z, t) + \tilde{U}_4(-a - l, z, t) \right)_{L^2(\Sigma_z)} + \\ \quad \left(\mu_4, \tilde{U}_3(a + l, z, t) + \tilde{U}_4(a + l, z, t) \right)_{L^2(\Sigma_z)} + (\mu_5, U(x, z, 0))_{L^2(\Omega)^4}. \end{array} \right.$$

Les intégrations par parties en temps et en espace dans $I_2(\alpha, h, U, \lambda). \tilde{U}$ donnent

$$\begin{aligned} I_2(\alpha, h, U, \lambda). \tilde{U} &= \left(\lambda, \partial_t \tilde{U} - A \partial_x \tilde{U} - B \partial_z \tilde{U} + S \tilde{U} \right)_{(L^2(\mathcal{Q}))^4} = \\ &\quad - \left(\partial_t \lambda, \tilde{U} \right)_{(L^2(\mathcal{Q}))^4} + \left(\lambda(x, z, T), \tilde{U}(x, z, T) \right)_{(L^2(\Omega))^4} \\ &\quad - \left(\lambda(x, z, 0), \tilde{U}(x, z, 0) \right)_{(L^2(\Omega))^4} + \left(\partial_x (A^T \lambda), \tilde{U} \right)_{(L^2(\mathcal{Q}))^4} \\ &\quad - \left(A^T \lambda(a + l, z, t), \tilde{U}(a + l, z, t) \right)_{(L^2(\Sigma_z))^4} \\ &\quad + \left(A^T \lambda(-a - l, z, t), \tilde{U}(-a - l, z, t) \right)_{(L^2(\Sigma_z))^4} \\ &\quad + \left(\partial_z (B^T \lambda), \tilde{U} \right)_{(L^2(\mathcal{Q}))^4} - \left(B^T \lambda(x, z_{max} + l, t), \tilde{U}(x, z_{max} + l, t) \right)_{(L^2(\Sigma_x))^4} \\ &\quad + \left(B^T \lambda(x, z_{min}, t), \tilde{U}(x, z_{min}, t) \right)_{(L^2(\Sigma_x))^4} + \left(S^T \lambda, \tilde{U} \right)_{(L^2(\mathcal{Q}))^4}. \end{aligned}$$

Ainsi, l'équation (3.13) se traduit par

$$\begin{aligned} \forall \tilde{U}, \quad & \left(-\partial_t \lambda + \partial_x (A^T \lambda) + \partial_z (B^T \lambda) + S^T \lambda, \tilde{U} \right)_{(L^2(\mathcal{Q}))^4} + \left(f^x, \tilde{U}_1 \right)_{L^2(\mathcal{Q})} + \left(f^z, \tilde{U}_2 \right)_{L^2(\mathcal{Q})} + \\ & \left(\lambda(x, z, T), \tilde{U}(x, z, T) \right)_{(L^2(\Omega))^4} - \left(\lambda(x, z, 0), \tilde{U}(x, z, 0) \right)_{(L^2(\Omega))^4} - \\ & \left(A^T \lambda(a + l, z, t), \tilde{U}(a + l, z, t) \right)_{(L^2(\Sigma_z))^4} + \left(A^T \lambda(-a - l, z, t), \tilde{U}(-a - l, z, t) \right)_{(L^2(\Sigma_z))^4} - \\ & \left(B^T \lambda(x, z_{max} + l, t), \tilde{U}(x, z_{max} + l, t) \right)_{(L^2(\Sigma_x))^4} + \left(B^T \lambda(x, z_{min}, t), \tilde{U}(x, z_{min}, t) \right)_{(L^2(\Sigma_x))^4} + \\ & \left(\mu_1, \tilde{U}_3(x, z_{min}, t) + \tilde{U}_4(x, z_{min}, t) \right)_{L^2(\Sigma_x)} + \left(\mu_2, \tilde{U}_3(x, z_{max} + l, t) + \tilde{U}_4(x, z_{max} + l, t) \right)_{L^2(\Sigma_x)} + \\ & \left(\mu_3, \tilde{U}_3(-a - l, z, t) + \tilde{U}_4(-a - l, z, t) \right)_{L^2(\Sigma_z)} + \left(\mu_4, \tilde{U}_3(a + l, z, t) + \tilde{U}_4(a + l, z, t) \right)_{L^2(\Sigma_z)} + \\ & (\mu_5, U(x, z, 0))_{(L^2(\Omega))^4} = 0. \end{aligned}$$

Ceci n'est possible que si les termes en facteur des mêmes \tilde{U}_i dans les produits scalaires sont nuls, cette égalité étant valable pour tout \tilde{U} . Aussi,

$$\left\{ \begin{array}{l} -\partial_t \lambda + \partial_x (A^T \lambda) + \partial_z (B^T \lambda) + S^T \lambda = (f^x, f^z, 0, 0) \\ \lambda(x, z, T) = 0, \quad \lambda(x, z, 0) = \mu_5(x, z) \\ A^T \lambda(a + l, z, t) = \mu_4(z, t) (0, 0, 1, 1) \\ A^T \lambda(-a - l, z, t) = \mu_3(z, t) (0, 0, 1, 1) \\ B^T \lambda(x, z_{max} + l, t) = \mu_2(z, t) (0, 0, 1, 1) \\ B^T \lambda(x, z_{min}, t) = \mu_1(z, t) (0, 0, 1, 1). \end{array} \right.$$

Le développement de ces expressions à partir de la définition des matrices $A(\alpha)$ et $B(\alpha)$ donnée en (3.8) mène à la définition du système adjoint,

$$\left\{ \begin{array}{l} \partial_t \lambda_1 - \partial_x \left(\frac{c}{\alpha} \lambda_3 \right) - c \sigma_x \lambda_1 = f_x(x, z, t) \\ \partial_t \lambda_2 - \partial_x \left(\frac{c}{\alpha} \lambda_4 \right) - c \sigma_z \lambda_2 = f_z(x, z, t) \\ \partial_t \lambda_3 - \partial_x (\alpha c \lambda_1) - \partial_z (\alpha c \lambda_2) - c \sigma_x \lambda_3 = 0 \\ \partial_t \lambda_4 - \partial_x (\alpha c \lambda_1) - \partial_z (\alpha c \lambda_2) - c \sigma_z \lambda_4 = 0, \end{array} \right. \quad (3.17)$$

avec la condition finale $\lambda(x, z, T) = 0$, les conditions aux limites

$$\begin{cases} \lambda_3(-a - l, z, t) = 0 & \lambda_4(x, z_{min}, t) = 0 \\ \lambda_3(a + l, z, t) = 0 & \lambda_4(x, z_{max} + l, t) = 0, \end{cases}$$

et les conditions sur les autres variables adjointes

$$\begin{cases} \mu_5(x, z) = \lambda(x, z, 0) \\ \mu_1(z, t) = -\alpha(-a - l, z)c(-a - l, z)\lambda_1(-a - l, z, t) \\ \mu_2(z, t) = \alpha(a + l, z)c(a + l, z)\lambda_1(a + l, z, t) \\ \mu_3(x, t) = -\alpha(x, z_{min})c(x, z_{min})\lambda_2(x, z_{min}, t) \\ \mu_4(x, t) = \alpha(x, z_{max} + l)c(x, z_{max} + l)\lambda_2(x, z_{max} + l, t). \end{cases}$$

Ce système définit l'état adjoint, qui permet de calculer les valeurs du gradient $\nabla \check{J}(\alpha, h)$, par les équations (3.16), puis celui du coût sismique total $\nabla J(v)$ par les équations (3.11). Il reste alors à calculer le gradient des termes de régularisation $R_1(\alpha)$ et $R_2(\alpha)$ définis par l'équation (3.6).

3.3.2 Calcul du gradient des termes de régularisation

Pour $\tilde{\alpha}$ dans $L^2(\Omega)$,

$$\begin{aligned} R'_1(\alpha).\tilde{\alpha} &= \lim_{t \rightarrow 0} \frac{R_1(\alpha + t\tilde{\alpha}) - R_1(\alpha)}{t} = \lim_{t \rightarrow 0} tR_1(\tilde{\alpha}) + 2(\partial_x \alpha, \partial_x \tilde{\alpha}) \\ &= -2(\partial_{xx} \alpha, \tilde{\alpha})_{L^2(\Omega)} + \int_{\Gamma_z} [\partial_x \alpha \tilde{\alpha}]_{-a-l}^{a+l} dz. \end{aligned}$$

Ainsi,

$$R'_1(\alpha)(x, z) = \begin{cases} -2\partial_{xx} \alpha \text{ pour } x \in]-a - l, a + l[\times \Gamma_z \\ -2\partial_{xx}(a + l, z) + 2\partial_x \alpha(a + l, z) \text{ pour } x = a + l, z \in \Gamma_z \\ -2\partial_{xx}(-a - l, z) - 2\partial_x \alpha(-a - l, z) \text{ pour } x = -a - l, z \in \Gamma_z. \end{cases}$$

De la même façon

$$\begin{aligned}
 R'_2(\alpha) \cdot \tilde{\alpha} &= \lim_{t \rightarrow 0} \frac{R_2(\alpha + t\tilde{\alpha}) - R_2(\alpha)}{t} \\
 &= \lim_{t \rightarrow 0} t \int_{(z_r)_1}^{(z_r)_Q} \tilde{\alpha}^2(0, z) dz + 2 \int_{(z_r)_1}^{(z_r)_Q} \tilde{\alpha}(0, z) (\alpha(0, z) - \alpha_0(z)) dz \\
 &= 2 \int_{(z_r)_1}^{(z_r)_Q} \tilde{\alpha}(0, z) (\alpha(0, z) - \alpha_0(z)) dz,
 \end{aligned}$$

et

$$R'_2(\alpha)(x, z) = \begin{cases} 2(\alpha(0, z) - \alpha_0(z)) & \text{pour } x = 0, \quad z \in [(z_r)_1, (z_r)_Q] \\ 0 & \text{sinon.} \end{cases}$$

Le gradient des termes de régularisation $R_1(\alpha)$ et $R_2(\alpha)$ est donc déterminé.

Les calculs qui ont été menés dans cette partie permettent de disposer d'une expression du gradient de la fonctionnelle $J(v)$ et des termes de régularisation $R_1(\alpha)$ et $R_2(\alpha)$, nécessaires à l'emploi de la méthode d'optimisation numérique présentée dans la partie 2 de ce chapitre, dans un cadre continu. Conformément à ce qui a été indiqué dans la partie présentant la méthode d'optimisation, le terme de régularisation $R_3(h_1, \dots, h_S)$ n'est pas utilisé, l'information *a priori* sur les CP étant principalement prise en compte par une méthode d'interpolation. C'est pourquoi il n'est pas nécessaire de calculer l'expression de son gradient.

Pour passer à l'étape d'implémentation numérique de la méthode, il est maintenant nécessaire de détenir une version discrète de cette expression. La partie suivante introduit donc le cadre dans lequel s'effectue la discréttisation, utilisé d'abord pour la résolution du problème direct, puis l'adaptation du calcul du gradient à ce cadre discret.

3.4 Discrétisation

3.4.1 Méthodologie de l'implémentation numérique

L'implémentation numérique de la méthode nécessite dans un premier temps la définition d'un schéma numérique discret adapté au problème direct. Ceci permet de définir une fonctionnelle et un problème inverse discrets, ce dernier consistant à minimiser cette fonctionnelle. La question se pose alors du calcul du gradient de cette fonctionnelle. Celui-ci pourrait être calculé directement par la discrétisation des formules déterminées dans la partie précédente. Cependant, cette solution doit être évitée. En effet, elle ne permet pas un calcul suffisamment précis du gradient de la fonctionnelle discrète, ce qui rend la méthode d'optimisation numérique caduque.

L'alternative consiste à suivre la même démarche que celle proposée dans la partie précédente en adaptant chaque étape au cadre discret introduit pour le problème direct. Ainsi, un lagrangien discret est défini, ainsi qu'un état adjoint discret, solution des équations adjointes du problème direct discrétisé. Il est d'ailleurs possible de voir que ces équations diffèrent de celles qui auraient été obtenues par discrétisation directe du problème adjoint continu déterminé dans la partie précédente.

Cette partie est donc consacrée à la discrétisation du problème (P_3), par la définition d'un problème direct $\widetilde{\mathcal{M}}_d(\alpha, h)$ et d'une fonctionnelle $\mathcal{J}(\alpha, h_1, \dots, h_S)$ discrets. Le minimum de cette fonctionnelle est déterminé par la méthode d'optimisation décrite dans la partie 2 de ce chapitre, qui nécessite le calcul de son gradient, calculé par la méthode de l'état adjoint, adaptée au cadre discret.

3.4.2 Notations

Dans un premier temps, il convient d'introduire les notations utilisées pour la description du problème discret. Soit $(K, M, N) \in \mathbb{N}^3$ le nombre de points de discrétisation utilisés. Ceci permet de définir les pas de discrétisation $(\Delta x, \Delta z, \Delta t) \in \mathbb{R}^3$, tels que

$$\Delta x = \frac{2(a + l)}{K + 1}, \quad \Delta z = \frac{z_{max} + l - z_{min}}{M + 1}, \quad \Delta t = \frac{T}{N + 1}.$$

Les points du maillage (x_i, z_j, t^n) en espace et en temps sont donnés par

$$\begin{cases} x_i = -a - l + i\Delta x, & \forall i = 0 \dots K + 1 \\ z_j = z_{min} + j\Delta z, & \forall j = 0 \dots M + 1 \\ t^n = n\Delta t, & \forall n = 0 \dots N + 1. \end{cases}$$

Les espaces discrets suivants sont définis, analogues des espaces introduits dans

la formulation continue du problème

$$\left\{ \begin{array}{l} \mathbb{G}_x = \{x_i, \quad 0 \leq i \leq K+1\}, \quad \mathbb{G}_t = \{t^n, \quad 0 \leq t \leq N+1\} \\ \mathbb{G}_z = \{z_j, \quad 0 \leq j \leq M+1\}, \quad \mathbb{S}_x = \mathbb{G}_x \times \mathbb{G}_t, \\ \mathbb{S}_z = \mathbb{G}_z \times \mathbb{G}_t, \quad \Omega_d = \mathbb{G}_x \times \mathbb{G}_z, \quad \mathcal{Q}_d = \Omega_d \times \mathbb{G}_t. \end{array} \right.$$

Il est également nécessaire de définir l'équivalent discret \mathcal{O}_d du domaine d'intérêt \mathcal{O} . Pour cela sont définis $(K', M') \in \mathbb{N}^2$, correspondant à l'épaisseur discrète de la couche PML, tels que

$$K' = \frac{l}{\Delta x}, \quad M' = \frac{l}{\Delta z}.$$

Alors, $\mathcal{O}_d = \{x_i, z_j, \quad K' \leq i \leq K - K', \quad 0 \leq j \leq M - M'\}$.

Une fonction est définie sur ces espaces par l'ensemble fini de valeurs prises aux points du maillage. La notation suivante est utilisée pour $f \in \mathbb{R}^{(K+2) \times (M+2) \times (N+2)}$,

$$f(x_i, z_j, t^n) = f_{ij}^n.$$

Les produits scalaires suivants sont également définis sur les espaces discrets introduits.

$$\begin{aligned} \text{Pour } (f, g) \in (\mathbb{R}^{K+2})^2, \quad (f, g)_x &= \sum_{i=1}^K f_i g_i \Delta x \\ (f, g) \in (\mathbb{R}^{M+2})^2, \quad (f, g)_z &= \sum_{j=1}^M f_j g_j \Delta z \\ (f, g) \in (\mathbb{R}^{N+2})^2, \quad (f, g)_t &= \sum_{n=1}^N f^n g^n \Delta t \\ (f, g) \in (\mathbb{R}^{(K+2) \times (N+2)})^2, \quad (f, g)_{x,t} &= \sum_{i=1}^K \sum_{n=1}^N f_i^n g_i^n \Delta x \Delta t \\ (f, g) \in (\mathbb{R}^{(M+2) \times (N+2)})^2, \quad (f, g)_{z,t} &= \sum_{j=1}^M \sum_{n=1}^N f_j^n g_j^n \Delta z \Delta t \\ (f, g) \in (\mathbb{R}^{(K+2) \times (M+2)})^2, \quad (f, g)_{x,z} &= \sum_{i=1}^K \sum_{j=1}^M f_{ij} g_{ij} \Delta x \Delta z \\ (f, g) \in (\mathbb{R}^{(K+2) \times (M+2) \times (N+2)})^2, \quad (f, g)_{x,z,t} &= \sum_{i=1}^K \sum_{j=1}^M \sum_{n=1}^N f_{ij}^n g_{ij}^n \Delta x \Delta z \Delta t \end{aligned}$$

Les normes associées sont indexées de manière similaire. Pour

$$f \in \mathbb{R}^{(K+2) \times (M+2) \times (N+2)}, \quad \|f\|_{x,z,t}^2 = (f, f)_{x,z,t}, \quad \text{etc...}$$

De plus, par la suite, pour $P \in \mathbb{N}$, et

$$(f, g) \in (\mathbb{R}^{(K+2) \times (M+2) \times (N+2)})^P \times (\mathbb{R}^{(K+2) \times (M+2) \times (N+2)})^P$$

la notation $(f, g)_{x,z,t}$ désigne la quantité

$$(f, g)_{x,z,t} = \sum_{p=1}^P (f_p, g_p)_{x,z,t},$$

ce qui correspond simplement au produit scalaire induit par le produit scalaire euclidien sur les espaces produits. Ceci s'étend à tous les produits scalaires définis ici, et à leur normes associées.

Soit $f \in \mathbb{R}^{(K+2) \times (M+2) \times (N+2)}$. Les opérateurs de dérivation discrets classiques sont introduits

$$\begin{cases} D_x^+ f_{ij}^n = \frac{f_{i+1,j}^n - f_{ij}^n}{\Delta x}, & D_x^- f_{ij}^n = \frac{f_{ij}^n - f_{i-1,j}^n}{\Delta x} \\ D_z^+ f_{ij}^n = \frac{f_{i,j+1}^n - f_{ij}^n}{\Delta z}, & D_z^- f_{ij}^n = \frac{f_{ij}^n - f_{i,j-1}^n}{\Delta z} \\ D_t^+ f_{ij}^n = \frac{f_{ij}^{n+1} - f_{ij}^n}{\Delta t}, & D_t^- f_{ij}^n = \frac{f_{ij}^n - f_{ij}^{n-1}}{\Delta t}, \end{cases}$$

ainsi que des opérateurs de translation,

$$\begin{cases} T_x^+ f_{ij}^n = f_{i+1,j}^n, & T_x^- f_{ij}^n = f_{i-1,j}^n \\ T_z^+ f_{ij}^n = f_{i,j+1}^n, & T_z^- f_{ij}^n = f_{i,j-1}^n \\ T_n^+ f_{ij}^n = f_{ij}^{n+1}, & T_n^- f_{ij}^n = f_{ij}^{n-1}. \end{cases}$$

Finalement, le symbole de Kronecker δ_{ij} est également utilisé, tel que

$$\delta_{ij} = \begin{cases} 1 & \text{si } i = j \\ 0 & \text{sinon.} \end{cases}$$

Ces notations permettent d'introduire le schéma numérique défini pour discréteriser le problème direct de propagation d'ondes acoustiques dans un milieu bordé d'une couche PML.

3.4.3 Problème direct discret

Le contexte d'utilisation de la méthode numérique développée est l'inversion non linéaire de données sismiques de puits. Dans ce contexte, une information à fine échelle est recherchée sur les hétérogénéités du sous-sol. Ceci se traduit par une discrétisation fine du sous-sol et le choix de pas de discrétisation spatiaux très petits. De plus, le problème ne fait pas apparaître une géométrie particulière nécessitant l'emploi de maillages complexes. C'est donc une méthode de discrétisation par différences finies d'ordre 2 qui est retenue. Plus précisément, le schéma retenu est le schéma de Yee [Yee, 1966], centré, d'ordre 2.

Ce choix d'une méthode de différences finies d'ordre 2 permet en outre de privilégier la rapidité d'exécution, par le faible nombre d'opérations utilisées en regard d'une méthode plus complexe. Cette recherche d'efficacité est la véritable problématique de l'implémentation numérique de la méthode.

Le schéma numérique utilisé pour la discrétisation des équations PML (3.3) est donc, pour (u_x, u_z, p_x, p_z) dans $(\mathbb{R}^{(K+2) \times (M+2) \times (N+2)})^4$,

$$\left\{ \begin{array}{l} \frac{(u_x)_{i-1/2,j}^{n-1/2} - (u_x)_{i-1/2,j}^{n-3/2}}{\Delta t} - \alpha_{i-1/2,j} c_{i-1/2,j} \left[\frac{(p_x + p_z)_{ij}^{n-1} - (p_x + p_z)_{i-1,j}^{n-1}}{\Delta x} \right] \\ \quad + c_{i-1/2,j} (\sigma_x)_{i-1/2} \frac{(u_x)_{i-1/2,j}^{n-1/2} + (u_x)_{i-1/2,j}^{n-3/2}}{2} = 0 \quad i = 1 \dots K + 1, j = 0 \dots M + 1, n = 1 \dots N, \\ \\ \frac{(u_z)_{i,j-1/2}^{n-1/2} - (u_z)_{i,j-1/2}^{n-3/2}}{\Delta t} - \alpha_{i,j-1/2} c_{i,j-1/2} \left[\frac{(p_x + p_z)_{ij}^{n-1} - (p_x + p_z)_{i,j-1}^{n-1}}{\Delta z} \right] \\ \quad + c_{i,j-1/2} (\sigma_z)_{j-1/2} \frac{(u_z)_{i,j-1/2}^{n-1/2} + (u_z)_{i,j-1/2}^{n-3/2}}{2} = 0 \quad i = 0 \dots K + 1, j = 1 \dots M + 1, n = 1 \dots N, \\ \\ \frac{(p_x)_{ij}^n - (p_x)_{ij}^{n-1}}{\Delta t} - \frac{c_{ij}}{\alpha_{ij}} \frac{(u_x)_{i+1/2,j}^{n-1/2} - (u_x)_{i-1/2,j}^{n-1/2}}{\Delta x} + c_{ij} (\sigma_x)_i \frac{(p_x)_{ij}^n + (p_x)_{ij}^{n-1}}{2} = 0 \\ \quad i = 1 \dots K, j = 1 \dots M, n = 1 \dots N, \\ \\ \frac{(p_z)_{ij}^n - (p_z)_{ij}^{n-1}}{\Delta t} - \frac{c_{ij}}{\alpha_{ij}} \frac{(u_z)_{i,j+1/2}^{n-1/2} - (u_z)_{i,j-1/2}^{n-1/2}}{\Delta z} + c_{ij} (\sigma_z)_j \frac{(p_z)_{ij}^n + (p_z)_{ij}^{n-1}}{2} = 0 \\ \quad i = 1 \dots K, j = 1 \dots M, n = 1 \dots N. \end{array} \right. \tag{3.18}$$

avec les conditions, $\forall i = 0 \dots K + 1, j = 0 \dots M + 1, n = 0 \dots N + 1,$

$$\left\{ \begin{array}{l} (u_x)_{i+1/2,j}^{-1/2} = 0 \\ (u_z)_{i,j+1/2}^{-1/2} = 0 \\ (p_x)_{ij}^0 = 0 \\ (p_z)_{ij}^0 = 0, \end{array} \right. \left\{ \begin{array}{l} (p_x + p_z)_{0j}^n = 0 \\ (p_x + p_z)_{K+1,j}^n = 0, \end{array} \right. \left\{ \begin{array}{l} (p_x + p_z)_{i0}^n = h_i^n \\ (p_x + p_z)_{i,M+1}^n = 0. \end{array} \right. \quad (3.19)$$

La définition d'un schéma centré conduit à utiliser des grilles décalées en espace et en temps pour certaines des variables. Plus précisément, ici, le choix a été fait de discréteriser les variables de pression p_x et p_z sur la grille initiale, tandis que les variables de vitesse de déplacement u_x (respectivement u_z) sont définies sur une grille décalée en temps et en espace suivant la direction x (respectivement en temps et en espace suivant la direction z). Ceci implique cependant le calcul des valeurs d'impédance et de vitesse sur les noeuds intermédiaires par interpolation des valeurs aux deux noeuds les plus proches sur la grille initiale. Ainsi, par exemple

$$\alpha_{i+1/2,j} = \frac{\alpha_{i+1,j} + \alpha_{ij}}{2} = \frac{T_x^+ \alpha_{ij} + \alpha_{ij}}{2}.$$

Il est possible de se ramener à un système discret, plus aisé à manipuler, équivalent discret du système (3.9). Soit

$$u_{ij}^n = \left((u_x)_{i-1/2,j}^{n-1/2}, (u_z)_{i,j-1/2}^{n-1/2}, (p_x)_{ij}^n, (p_z)_{ij}^n \right)^T,$$

et

$$a_{ij} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ \frac{c_{ij}}{\alpha_{ij}} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad b_{ij} = \begin{pmatrix} 0 & 0 & \frac{(\alpha_{ij} + T_x^- \alpha_{ij})(c_{ij} + T_x^- c_{ij})}{4} & \frac{(\alpha_{ij} + T_x^- \alpha_{ij})(c_{ij} + T_x^- c_{ij})}{4} \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix},$$

$$c_{ij} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & \frac{c_{ij}}{\alpha_{ij}} & 0 & 0 \end{pmatrix}, \quad d_{ij} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & \frac{(\alpha_{ij} + T_x^- \alpha_{ij})(c_{ij} + T_x^- c_{ij})}{4} & \frac{(\alpha_{ij} + T_x^- \alpha_{ij})(c_{ij} + T_x^- c_{ij})}{4} \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix},$$

$$e_{ij} = \begin{pmatrix} \frac{((\sigma_x)_i + T_x^- (\sigma_x)_i)(c_{ij} + T_x^- c_{ij})}{4} & 0 & 0 & 0 \\ 0 & \frac{((\sigma_z)_j + T_z^- (\sigma_z)_j)(c_{ij} + T_z^- c_{ij})}{4} & 0 & 0 \\ 0 & 0 & c_{ij}(\sigma_x)_i & 0 \\ 0 & 0 & 0 & c_{ij}(\sigma_z)_j \end{pmatrix}.$$

Soit alors

$$u^n = \begin{bmatrix} u_{11}^n \\ \vdots \\ u_{1,M+1}^n \\ \vdots \\ u_{K+1,1}^n \\ \vdots \\ u_{K+1,M+1}^n \end{bmatrix}, \quad \left\{ \begin{array}{l} a = \text{diag}(a_{11}, \dots, a_{1,M}, \dots, a_{K,1}, \dots, a_{K,M}) \\ b = \text{diag}(b_{11}, \dots, b_{1,M}, \dots, b_{K,1}, \dots, b_{K,M}) \\ c = \text{diag}(c_{11}, \dots, c_{1,M}, \dots, c_{K,1}, \dots, c_{K,M}) \\ d = \text{diag}(d_{11}, \dots, d_{1,M}, \dots, d_{K,1}, \dots, d_{K,M}) \\ e = \text{diag}(e_{11}, \dots, e_{1,M}, \dots, e_{K,1}, \dots, e_{K,M}), \end{array} \right.$$

et

$$U = \begin{bmatrix} u^1 \\ \vdots \\ u^N \end{bmatrix}, \quad \left\{ \begin{array}{l} A = \text{diag}(\underbrace{a, \dots, a}_N) \\ B = \text{diag}(\underbrace{b, \dots, b}_N) \\ C = \text{diag}(\underbrace{c, \dots, c}_N) \end{array} \right. , \quad \left\{ \begin{array}{l} D = \text{diag}(\underbrace{d, \dots, d}_N) \\ E = \text{diag}(\underbrace{e, \dots, e}_N) \end{array} \right.$$

Les équations 3.18 sont équivalentes au système

$$D_t^- U - AD_x^+ U - BD_x^- T_n^- U - CD_z^+ U - DD_z^- T_n^- U + \frac{1}{2}E(U + T_n^- U) = 0, \quad (3.20)$$

avec les conditions aux limites et les conditions initiales

$$\forall(i, j, n), \quad 0 \leq i \leq K + 1, \quad 0 \leq j \leq M + 1, \quad 0 \leq n \leq N + 1,$$

$$\left\{ \begin{array}{l} U_{ij}^0 = 0, \quad (U_3 + U_4)_{0j}^n = 0, \quad (U_3 + U_4)_{K+1,j}^n = 0, \\ (U_3 + U_4)_{i0}^n = h_i^n, \quad (U_3 + U_4)_{i,M+1}^n = 0. \end{array} \right. \quad (3.21)$$

où les notations suivantes sont employées

$$\forall(i, j, n), \quad 0 \leq i \leq K + 1, \quad 0 \leq j \leq M + 1, \quad 0 \leq n \leq N + 1,$$

$$U_{ij}^n = u_{ij}^n, \quad (U_1)_{ij}^n = (u_x)_{ij}^n, \quad (U_2)_{ij}^n = (u_z)_{ij}^n, \quad (U_3)_{ij}^n = (p_x)_{ij}^n, \quad (U_4)_{ij}^n = (p_z)_{ij}^n.$$

Les récepteurs étant placés aux coordonnées

$$(x_r)_q, (z_r)_q, \quad 1 \leq q \leq Q,$$

les indices correspondants de la grille sont notés $(i_r)_q, (j_r)_q$.

Le problème direct discret peut alors se formaliser par la définition d'un opérateur $\widetilde{\mathcal{M}}_d$, analogue discret de l'opérateur continu $\widetilde{\mathcal{M}}$. Soit

$$\begin{aligned} (\widetilde{U}_x) \in (\mathbb{R}^{N+2})^Q, \quad & (\widetilde{U}_x) = \left((U_1)_{(i_r)_1, (j_r)_1}^n, \dots, (U_1)_{(i_r)_Q, (j_r)_Q}^n \right), \quad n = 0, \dots, N+1, \\ (\widetilde{U}_z) \in (\mathbb{R}^{N+2})^Q, \quad & (\widetilde{U}_z) = \left((U_2)_{(i_r)_1, (j_r)_1}^n, \dots, (U_2)_{(i_r)_Q, (j_r)_Q}^n \right), \quad n = 0, \dots, N+1, \end{aligned}$$

où U est solution du système discret (3.20),(3.21). Alors

$$\widetilde{\mathcal{M}}_d : (\alpha, h) \longrightarrow \left((\widetilde{U}_x), (\widetilde{U}_z) \right)$$

est l'opérateur associé au problème direct discret, à valeur dans $(\mathbb{R}^{N+2})^{2Q}$.

Remarque. *Les variables α et h sont à prendre dans leur version discrète : dans cette définition et toute la suite de cette partie, α est pris dans $\mathbb{R}^{(K+2) \times (M+2)}$ et h dans $\mathbb{R}^{(K+2) \times (N+2)}$.*

La définition du problème direct discret permet d'aborder celle du problème inverse discret.

3.4.4 Problème inverse discret

La discréttisation du problème direct et la définition de l'opérateur $\widetilde{\mathcal{M}}_d$ permettent de définir la fonctionnelle discrète associée au coût sismique

$$\mathcal{J}(\alpha, h_1, \dots, h_S) = \sum_{s=1}^S \|\mathcal{D}_s - \widetilde{\mathcal{M}}_d(\alpha, h_s)\|_t^2.$$

où \mathcal{D}_s , défini dans $(\mathbb{R}^{N+2})^{2Q}$, correspond au jeu de données D_s discréttisé.

Les termes de pénalisation introduits pour stabiliser le problème inverse sont également discréttisés. Les analogues discrets $\mathcal{R}_1(\alpha), \mathcal{R}_2(\alpha)$ sont définis par

$$\mathcal{R}_1(\alpha) = \|D_x^- \alpha\|_{x,z}^2,$$

$$\mathcal{R}_2(\alpha) = \sum_{j=(j_r)_1}^{(j_r)_Q} ((\alpha_0)_j - \alpha_{(i_r)_q, j})^2 \Delta z.$$

Aussi la fonctionnelle discrète dont le minimum est recherché est la fonctionnelle régularisée

$$\widehat{\mathcal{J}}^\varepsilon(\alpha, q_1, \dots, q_S) = \mathcal{J}(\alpha, B_1 q_1, \dots, B_S q_S) + \varepsilon_1 \mathcal{R}_1(\alpha) + \varepsilon_2 \mathcal{R}_2(\alpha),$$

où les termes q_s , $1 \leq s \leq S$, sont définis par l'équation (3.7).

Ceci permet de définir le problème inverse discret (\mathcal{P}_3)

$$(\mathcal{P}_3) \quad \min_{\substack{\alpha_{ij}, (q_1)_{i_k}^n, \dots, (q_S)_{i_k}^n \\ i = 0, \dots, K+1 \\ j = 0, \dots, M+1 \\ n = 0, \dots, N+1 \\ k = 0, \dots, p}} \hat{\mathcal{J}}^\varepsilon(\alpha, q_1, \dots, q_S).$$

Pour résoudre le problème (\mathcal{P}_3) , l'algorithme d'optimisation emboîtée présenté dans la première partie est utilisé. Comme cela a été vu, il nécessite le calcul du gradient de la fonctionnelle $\mathcal{J}(\alpha, h_1, \dots, h_S)$, et de celui des termes de régularisation $\mathcal{R}_1(\alpha)$ et $\mathcal{R}_2(\alpha)$. Comme pour le problème continu, le gradient de la partie coût sismique est d'abord déterminé, par la méthode de l'état adjoint, adaptée au cadre discret introduit via les équations (3.20), (3.21). Le gradient des termes de régularisation est calculé par la suite.

3.4.5 Calcul du gradient de la fonctionnelle $\mathcal{J}(\alpha, h_1, \dots, h_S)$

3.4.5.1 Expression du gradient en fonction de l'état adjoint

Le procédé utilisé est analogue à celui utilisé dans le cas continu, dont les difficultés ont déjà été abordées dans la partie précédente. Le problème est d'abord simplifié par la définition de

$$\mathcal{J}_s(\alpha, h_s) = \|\mathcal{D}_s - \widetilde{\mathcal{M}}_d(\alpha, h_s)\|_t^2. \quad (3.22)$$

Ainsi,

$$\begin{cases} \partial_\alpha \mathcal{J}(\alpha, h_1, \dots, h_S) = \sum_{s=1}^S \partial_\alpha \mathcal{J}_s(\alpha, h_s) \\ \partial_{h_s} \mathcal{J}(\alpha, h_1, \dots, h_S) = \partial_{h_s} \mathcal{J}_s(\alpha, h_s). \end{cases} \quad (3.23)$$

Comme précédemment pour le problème continu, on va s'intéresser au calcul du gradient de la fonctionnelle particulière $\mathcal{J}_1(\alpha, h_1)$ qui est renotée $\check{\mathcal{J}}(\alpha, h)$ pour simplifier les notations. Cette fonctionnelle est telle que

$$\check{\mathcal{J}}(\alpha, h) = \sum_{q=1}^Q \|(\mathcal{D}_{1,q}^x)^n - (U_1)_{(i_r)_q, (j_r)_q}^n\|_t^2 + \|(\mathcal{D}_{1,q}^z)^n - (U_2)_{(i_r)_q, (j_r)_q}^n\|_t^2,$$

où U est solution des équations (3.20), (3.21), pour la CP $h(x, t) = h_1(x, t)$.

Soit

$$\mathcal{G}(U) = \sum_{q=1}^Q \|(\mathcal{D}_{1,q}^x)^n - (U_1)_{(i_r)_q, (j_r)_q}^n\|_t^2 + \|(\mathcal{D}_{1,q}^z)^n - (U_2)_{(i_r)_q, (j_r)_q}^n\|_t^2,$$

pour $U \in (\mathbb{R}^{(K+2) \times (M+2) \times (N+2)})^4$ quelconque.

Soient

$$\begin{cases} \lambda \in (\mathbb{R}^{(K+2) \times (M+2) \times (N+2)})^4, \mu_1 \in \mathbb{R}^{(K+2) \times (N+2)}, \\ \mu_2 \in \mathbb{R}^{(K+2) \times (N+2)}, \mu_3 \in \mathbb{R}^{(M+2) \times (N+2)}, \\ \mu_4 \in \mathbb{R}^{(M+2) \times (N+2)}, \mu_5 \in (\mathbb{R}^{(K+2) \times (M+2)})^4, \end{cases}$$

et $\nu = (\lambda, \mu_1, \dots, \mu_5)$.

Le Lagrangien $\mathcal{L}(\alpha, h, U, \nu)$ associé au problème (\mathcal{P}_4)

$$(\mathcal{P}_4) \quad \min_{\alpha_{ij}, h_i^n} \quad \check{\mathcal{J}}(\alpha, h),$$

$$\begin{cases} i = 0, \dots, K+1 \\ j = 0, \dots, M+1 \\ n = 0, \dots, N+1 \end{cases}$$

est défini par

$$\mathcal{L}(\alpha, h, U, \nu) =$$

$$\begin{aligned} & \mathcal{G}(U) + (\mu_1, (U_3 + U_4)_{i0}^n - h)_x + (\mu_2, (U_3 + U_4)_{i,M+1}^n)_x + \\ & (\mu_3, (U_3 + U_4)_{0j}^n)_z + (\mu_4, (U_3 + U_4)_{K+1,j}^n)_z + (\mu_5, U_{ij}^0)_{x,z} + \\ & \left(\lambda, D_t^+ U - A(\alpha) D_x^+ U - B(\alpha) D_x^- T_n^- U - C(\alpha) D_z^+ U - D(\alpha) D_z^- T_n^- U + \frac{1}{2} E (U + T_n^- U) \right)_{x,z,t}. \end{aligned}$$

Il est à noter que les composantes de la variable λ sont discrétisées sur les mêmes grilles que les composantes de la variable U dont elles sont les adjointes. Ainsi, la variables λ_1 est définie sur la grille décalée en temps et latéralement en espace, tandis que la variable λ_2 est définie sur la grille décalée en temps et verticalement en espace. Les variables λ_3 et λ_4 sont définies sur la grille initiale.

Comme pour le problème continu, pour $\bar{U}(\alpha, h)$ solution de (3.20) et (3.21) et pour $\bar{\nu}$ solution de

$$\forall (\alpha, h) \quad \partial_3 \mathcal{L}(\alpha, h, \bar{U}(\alpha, h), \bar{\nu}) = 0, \quad (3.24)$$

le gradient de $\check{\mathcal{J}}(\alpha, h)$ est donné par

$$\partial_\alpha \check{\mathcal{J}}(\alpha, h) = \partial_1 \mathcal{L}(\alpha, h, \bar{U}(\alpha, h), \bar{\nu}), \quad \partial_h \check{\mathcal{J}}(\alpha, h) = \partial_2 \mathcal{L}(\alpha, h, \bar{U}(\alpha, h), \bar{\nu}).$$

Ainsi, pour $\tilde{\alpha} \in \mathbb{R}^{(K+2) \times (M+2)}$,

$$\begin{aligned}
& \partial_\alpha \check{\mathcal{J}}(\alpha, h) \cdot \tilde{\alpha} = \\
& - \sum_{n=1}^N \Delta t \left(\bar{\lambda}^n, \tilde{\alpha} A'(\alpha) D_x^+ \bar{U}^n(\alpha, h) + \tilde{\alpha} B'(\alpha) D_x^- T_n^- \bar{U}^n(\alpha, h) \right)_{x,z} \\
& - \sum_{n=1}^N \Delta t \left(\bar{\lambda}^n, \tilde{\alpha} C'(\alpha) D_z^+ \bar{U}^n(\alpha, h) + \tilde{\alpha} D'(\alpha) D_z^- T_n^- \bar{U}^n(\alpha, h) \right)_{x,z} = \\
& - \sum_{n=1}^N \Delta t \left((\bar{\lambda}_3)^n, \frac{c\tilde{\alpha}}{\alpha^2} D_x^+ (\bar{U}_1)^n(\alpha, h) \right)_{x,z} - \sum_{n=1}^N \Delta t \left((\bar{\lambda}_4)^n, \frac{c\tilde{\alpha}}{\alpha^2} D_z^+ (\bar{U}_2)^n(\alpha, h) \right)_{x,z} \\
& - \sum_{n=1}^N \Delta t \left(\frac{(\bar{\lambda}_1)^{n-1/2}}{4}, ((c + T_x^- c)\tilde{\alpha} + (c + T_x^- c)T_x^- \tilde{\alpha}) D_x^- T_n^- (\bar{U}_3 + \bar{U}_4)^n(\alpha, h) \right)_{x,z} \\
& - \sum_{n=1}^N \Delta t \left(\frac{(\bar{\lambda}_2)^{n-1/2}}{4}, ((c + T_z^- c)\tilde{\alpha} + (c + T_z^- c)T_z^- \tilde{\alpha}) D_z^- T_n^- (\bar{U}_3 + \bar{U}_4)^n(\alpha, h) \right)_{x,z}.
\end{aligned}$$

$$\begin{aligned}
\text{Or, } & \left(\frac{(\bar{\lambda}_2)^{n-1/2}}{4}, ((c + T_z^- c)\tilde{\alpha} + (c + T_z^- c)T_z^- \tilde{\alpha}) D_z^- T_n^- (\bar{U}_3 + \bar{U}_4)^n(\alpha, h) \right)_{x,z} = \\
& \left(\frac{(\bar{\lambda}_2)^{n-1/2}}{4} (c + T_z^- c) D_z^- T_n^- (\bar{U}_3 + \bar{U}_4)^n(\alpha, h), \tilde{\alpha} \right)_{x,z} \\
& + \left(\frac{(\bar{\lambda}_2)^{n-1/2}}{4} (c + T_z^- c) D_z^- T_n^- (\bar{U}_3 + \bar{U}_4)^n(\alpha, h), T_z^- \tilde{\alpha} \right)_{x,z},
\end{aligned}$$

$$\begin{aligned}
\text{et } & \left(\frac{(\bar{\lambda}_2)^{n-1/2}}{4} (c + T_z^- c) D_z^- T_n^- (\bar{U}_3 + \bar{U}_4)^n(\alpha, h), T_z^- \tilde{\alpha} \right)_{x,z} = \\
& \left(T_z^+ \left(\frac{(\bar{\lambda}_2)^{n-1/2}}{4} (c + T_z^- c) D_z^- T_n^- (\bar{U}_3 + \bar{U}_4)^n(\alpha, h) \right), \tilde{\alpha} \right)_{x,z} \\
& + \left(\frac{(\bar{\lambda}_2)_{i1}^{n-1/2}}{4} (c + T_z^- c)_{i1} D_z^- T_n^- (\bar{U}_3 + \bar{U}_4)_{i1}^n(\alpha, h), \tilde{\alpha}_{i0} \right)_x \\
& - \left(\frac{(\bar{\lambda}_2)_{i,M+1}^{n-1/2}}{4} (c + T_z^- c)_{i,M+1} D_z^- T_n^- (\bar{U}_3 + \bar{U}_4)_{i,M+1}^n(\alpha, h), \tilde{\alpha}_{iM} \right)_x.
\end{aligned}$$

De la même façon,

$$\begin{aligned}
& \left(\frac{(\bar{\lambda}_1)^{n-1/2}}{4}, ((c + T_x^- c)\tilde{\alpha} + (c + T_x^- c)T_x^- \tilde{\alpha})D_x^- T_n^- (\bar{U}_3 + \bar{U}_4)^n(\alpha, h) \right)_{x,z} = \\
& \left(\frac{(\bar{\lambda}_1)^{n-1/2}}{4}(c + T_x^- c)D_x^- T_n^- (\bar{U}_3 + \bar{U}_4)^n(\alpha, h), \tilde{\alpha} \right)_{x,z} \\
& + \left(T_x^+ \left(\frac{(\bar{\lambda}_1)^{n-1/2}}{4}(c + T_x^- c)D_x^- T_n^- (\bar{U}_3 + \bar{U}_4)^n(\alpha, h) \right), \tilde{\alpha} \right)_{x,z} \\
& + \left(\frac{(\bar{\lambda}_1)_{1j}^{n-1/2}}{4}(c + T_x^- c)_{i1} D_x^- T_n^- (\bar{U}_3 + \bar{U}_4)_{1j}^n(\alpha, h), \tilde{\alpha}_{0j} \right)_z \\
& - \left(\frac{(\bar{\lambda}_1)_{K+1,j}^{n-1/2}}{4}(c + T_x^- c)_{K+1,j} D_x^- T_n^- (\bar{U}_3 + \bar{U}_4)_{K+1,j}^n(\alpha, h), \tilde{\alpha}_{Kj} \right)_z.
\end{aligned}$$

Comme $\bar{U}(\alpha, h)$ est solution des équations (3.21),

$$D_x^- T_n^- (\bar{U}_3 + \bar{U}_4)_{K+1,j}^n(\alpha, h) = 0, \quad D_z^- T_n^- (\bar{U}_3 + \bar{U}_4)_{i,M+1}^n(\alpha, h) = 0.$$

Finalement, pour $1 \leq i \leq K$, $1 \leq j \leq M$

$$\begin{aligned}
\partial_\alpha \check{\mathcal{J}}(\alpha, h)_{ij} = & - \sum_{n=1}^N (\bar{\lambda}_3)_{ij}^n(\alpha, h) \frac{c_{ij}}{\alpha_{ij}^2} D_x^+ (\bar{U}_1)_{ij}^n(\alpha, h) \Delta t \\
& - \sum_{n=1}^N (\bar{\lambda}_4)_{ij}^n(\alpha, h) \frac{c_{ij}}{\alpha_{ij}^2} D_z^+ (\bar{U}_2)_{ij}^n(\alpha, h) \Delta t \\
& + \sum_{n=1}^N \frac{(\bar{\lambda}_1)_{i-1/2,j}^{n-1/2}}{2} c_{i-1/2,j} D_x^- (\bar{U}_3 + \bar{U}_4)_{ij}^{n-1}(\alpha, h) \Delta t \\
& + \sum_{n=1}^N \frac{(\bar{\lambda}_1)_{i+1/2,j}^{n-1/2}}{2} c_{i+1/2,j} D_x^+ (\bar{U}_3 + \bar{U}_4)_{ij}^{n-1}(\alpha, h) \Delta t \\
& + \sum_{n=1}^N \frac{(\bar{\lambda}_2)_{i,j-1/2}^{n-1/2}}{2} c_{i,j-1/2} D_z^- (\bar{U}_3 + \bar{U}_4)_{ij}^{n-1}(\alpha, h) \Delta t \\
& + \sum_{n=1}^N \frac{(\bar{\lambda}_2)_{i,j+1/2}^{n-1/2}}{2} c_{i,j+1/2} D_z^+ (\bar{U}_3 + \bar{U}_4)_{ij}^{n-1}(\alpha, h) \Delta t.
\end{aligned}$$

Pour $1 \leq j \leq M$

$$\partial_\alpha \check{\mathcal{J}}(\alpha, h)_{0j} = \sum_{n=1}^N \frac{(\bar{\lambda}_1)_{1j}^{n-1/2}}{2} c_{1j} D_x^- (\bar{U}_3 + \bar{U}_4)_{1j}^n(\alpha, h) \Delta t,$$

$$\partial_\alpha \check{\mathcal{J}}(\alpha, h)_{K+1,j} = 0.$$

Pour $1 \leq i \leq K$

$$\partial_\alpha \check{\mathcal{J}}(\alpha, h)_{i0} = \sum_{n=1}^N \frac{(\bar{\lambda}_2)_{i1}^{n-1/2}}{2} c_{i1} D_z^- (\bar{U}_3 + \bar{U}_4)_{i1}^n(\alpha, h) \Delta t,$$

$$\partial_\alpha \check{\mathcal{J}}(\alpha, h)_{i,M+1} = 0.$$

Il reste à définir le gradient discret $\partial_h \check{\mathcal{J}}(\alpha, h)$. Or, pour \tilde{h} dans $\mathbb{R}^{(K+2) \times (N+2)}$,

$$\partial_2 \mathcal{L}(\alpha, h, \bar{U}(\alpha, h), \bar{\nu}) \tilde{h} = \left(\mu_1, -\tilde{h} \right)_{x,t}.$$

De ceci se déduit,

$$\partial_h \check{\mathcal{J}}(\alpha, h) = -(\mu_1)_i^n, \quad \text{pour } i = 1, \dots, K, \quad n = 1 \dots N.$$

3.4.5.2 Système adjoint

Il reste maintenant à expliciter le système discret dont l'état adjoint est solution. Celui-ci se dérive de l'équation (3.24). Le procédé est le même que pour le cas continu. Soit \tilde{U} dans $(\mathbb{R}^{(K+2) \times (M+2) \times (N+2)})^4$ alors,

$$\partial_3 \mathcal{L}(\alpha, h, U, \nu) \cdot \tilde{U} = \mathcal{I}_1(\alpha, h, U) \cdot \tilde{U} + \mathcal{I}_2(\alpha, h, U, \nu) \cdot \tilde{U} + \mathcal{I}_3(\alpha, h, U, \nu) \cdot \tilde{U},$$

$$\text{avec } \mathcal{I}_1(\alpha, h, U) \cdot \tilde{U} = \mathcal{G}'(U) \cdot \tilde{U} = -2 \sum_{q=1}^Q \left(\mathcal{D}_q^x - (U_1)_{(i_r)_q, (j_r)_q}, (\tilde{U}_1)_{(i_r)_q, (j_r)_q} \right)_t$$

$$-2 \sum_{q=1}^Q \left(\mathcal{D}_q^z - (U_2)_{(i_r)_q, (j_r)_q}, (\tilde{U}_2)_{(i_r)_q, (j_r)_q} \right)_t$$

$$= \left(F_x, \tilde{U}_1 \right)_{L^2(\mathcal{Q}_d)} + \left(F_z, \tilde{U}_2 \right)_{x,z,t}$$

où

$$\begin{cases} (F_x)_{ij}^n = \sum_{q=1}^Q ((\mathcal{D}_q^x)^n - (U_1)_{ij}^n) \delta_{i,(i_r)_q} \delta_{j,(j_r)_q} \\ (F_z)_{ij}^n = \sum_{q=1}^Q ((\mathcal{D}_q^z)^n - (U_2)_{ij}^n) \delta_{i,(i_r)_q} \delta_{j,(j_r)_q} \end{cases}$$

et

$$\begin{aligned}\mathcal{I}_2(\alpha, h, U, \nu) \cdot \tilde{U} &= \\ \left(\lambda, D_t^- \tilde{U} - AD_x^+ \tilde{U} - BD_x^- T_n^- \tilde{U} - CD_z^+ \tilde{U} - DD_z^- T_n^- \tilde{U} + \frac{1}{2} E \left(\tilde{U} + T_n^- \tilde{U} \right) \right)_{x,z,t}, \\ \mathcal{I}_3(\alpha, h, U, \nu) \cdot \tilde{U} &= \left(\mu_1, (\tilde{U}_3 + \tilde{U}_4)_{i0}^n - h \right)_{x,t} + \left(\mu_2, (\tilde{U}_3 + \tilde{U}_4)_{i,M+1}^n \right)_{x,t} + \\ &\quad \left(\mu_3, (\tilde{U}_3 + \tilde{U}_4)_{0j}^n \right)_{z,t} + \left(\mu_4, (\tilde{U}_3 + \tilde{U}_4)_{K+1,j}^n \right)_{z,t} + \left(\mu_5, \tilde{U}_{ij}^0 \right)_{x,z}.\end{aligned}$$

En utilisant les propriétés des opérateurs de translation et en réalisant des intégrations par parties dans le terme \mathcal{I}_2 , il vient la relation suivante

$$\begin{aligned}
& \forall \tilde{U}, \quad \partial_3 \mathcal{L}(\alpha, h, U, \nu) \cdot \tilde{U} = 0 \iff \forall \tilde{U}, \quad \left(F_x, \tilde{U}_1 \right)_{x,z,t} + \left(F_z, \tilde{U}_2 \right)_{x,z,t} + \\
& \quad \left(-D_t^+ \lambda + D_x^- A^T \lambda + D_z^- C^T \lambda + T_n^+ D_x^+ B^T \lambda + T_n^+ D_z^+ D^T \lambda + \frac{1}{2} (E\lambda + ET_n^+ \lambda), \tilde{U} \right)_{x,z,t} \\
& \quad + \left(\lambda^{N+1}, \tilde{U}^N \right)_{x,z} - \left(\lambda^1, \tilde{U}^0 \right)_{x,z} - \left((A^T \lambda)_K^n, \tilde{U}_{K+1,j}^n \right)_{z,t} \\
& \quad + \left((A^T \lambda)_{0j}^n, \tilde{U}_{1j}^n \right)_{z,t} + \left((C^T \lambda)_{iM}^n, \tilde{U}_{i,M+1}^n \right)_{z,t} - \left((C^T \lambda)_{i0}^n, \tilde{U}_{i1}^n \right)_{z,t} \\
& \quad - \left(D_x^+ B^T \lambda^{N+1}, \tilde{U}^N \right)_{x,z} + \left(D_x^+ B^T \lambda^1, \tilde{U}^0 \right)_{x,z} - \left(T_n^+ (B^T \lambda)_{K+1,j}^n, \tilde{U}_{K,j}^n \right)_{z,t} \\
& \quad + \left((B^T \lambda)_{K+1,j}^{N+1}, \tilde{U}_{K,j}^N \right)_z - \left((B^T \lambda)_{K+1,j}^1, \tilde{U}_{K,j}^0 \right)_z \\
& \quad + \left(T_n^+ (B^T \lambda)_{1j}^n, \tilde{U}_{0j}^n \right)_{z,t} - \left((B^T \lambda)_{1j}^{N+1}, \tilde{U}_{0j}^N \right)_z + \left((B^T \lambda)_{1j}^1, \tilde{U}_{0j}^0 \right)_z \\
& \quad - \left(D_z^+ D^T \lambda^{N+1}, \tilde{U}^N \right)_{x,z} + \left(D_z^+ D^T \lambda^1, \tilde{U}^0 \right)_{x,z} \\
& \quad - \left(T_n^+ (D^T \lambda)_{i1}^n, \tilde{U}_{i0}^n \right)_{x,t} + \left((D^T \lambda)_{i1}^{N+1}, \tilde{U}_{i0}^N \right)_x - \left((D^T \lambda)_{i1}^1, \tilde{U}_{i0}^0 \right)_x \\
& \quad + \left(T_n^+ (D^T \lambda)_{i,M+1}^n, \tilde{U}_{iM}^n \right)_{(x,t)} - \left((D^T \lambda)_{i,M+1}^{N+1}, \tilde{U}_{iM}^N \right)_x + \left((D^T \lambda)_{i,M+1}^1, \tilde{U}_{iM}^0 \right)_x \\
& \quad - \left(E \frac{\lambda^{N+1}}{2}, \tilde{U}^N \right)_{x,z} + \left(E \frac{\lambda^1}{2}, \tilde{U}^0 \right)_{x,z} + \left(\mu_1, (\tilde{U}_3 + \tilde{U}_4)_{i0}^n - h \right)_{x,t} \\
& \quad + \left(\mu_2, (\tilde{U}_3 + \tilde{U}_4)_{i,M+1}^n \right)_{x,t} + \left(\mu_3, (\tilde{U}_3 + \tilde{U}_4)_{0j}^n \right)_{z,t} \\
& \quad + \left(\mu_4, (\tilde{U}_3 + \tilde{U}_4)_{K+1,j}^n \right)_{z,t} + \left(\mu_5, \tilde{U}_{ij}^0 \right)_{x,z} = 0.
\end{aligned}$$

Les termes de bord correspondants au même produit scalaire, en facteur du même \tilde{U} , s'annulent donc, puisque l'expression est nulle pour tout accroissement \tilde{U} . Aussi,

$$\left\{ \begin{array}{l} -D_t^+ \lambda_{ij}^n + D_x^- A^T \lambda_{ij}^n + D_z^- C^T \lambda_{ij}^n + T_n^+ D_x^+ B^T \lambda_{ij}^n + T_n^+ D_z^+ D^T \lambda_{ij}^n + \frac{1}{2} (E + T_n^+ E) \lambda_{ij}^n = \\ ((F_x)_{ij}^n, (F_z)_{ij}^n, 0, 0), \\ (I - \frac{1}{2} E - D_x^+ B^T - D_z^+ D^T) \lambda_{ij}^{N+1} = 0, \\ (I - \frac{1}{2} E - D_x^+ B^T - D_z^+ D^T) \lambda_{ij}^1 = (\mu_5)_{ij}, \\ (A^T \lambda)_{Kj}^n = 0, \quad (A^T \lambda)_{0j}^n = 0, \quad (C^T \lambda)_{iM}^n = 0, \quad (C^T \lambda)_{i0}^n = 0, \\ (B^T \lambda)_{1j}^n = (\mu_3)_j^n, \quad (B^T \lambda)_{K+1,j}^n = 0, \quad (D^T \lambda)_{i1}^n = (\mu_1)_i^n, \quad (D^T \lambda)_{i,M+1}^n = 0, \end{array} \right.$$

I désignant la matrice identité de $\mathcal{M}_4(\mathbb{R})$.

Le développement de ces expressions mène au système adjoint discret

$$\left\{ \begin{array}{l} \frac{(\lambda_1)_{i-1/2,j}^{n+1/2} - (\lambda_1)_{i-1/2,j}^{n-1/2}}{\Delta t} - \frac{\frac{c_{ij}}{\alpha_{ij}} (\lambda_3)_{ij}^n - \frac{c_{i-1,j}}{\alpha_{i-1,j}} (\lambda_3)_{i-1,j}^n}{\Delta x} - \frac{1}{2} c_{i-1/2,j} (\sigma_x)_{i-1/2} \left((\lambda_1)_{i-1/2,j}^{n+1/2} + (\lambda_1)_{i-1/2,j}^{n-1/2} \right) \\ = \sum_{q=1}^Q ((D_q^x)^n - (U_1)_{ij}^n) \delta_{(i_r)_q} \delta_{(j_r)_q}, \quad 1 \leq i \leq K+1, \quad 0 \leq j \leq M+1, \quad 0 \leq n \leq N, \\ \frac{(\lambda_2)_{i,j-1/2}^{n+1/2} - (\lambda_2)_{i,j-1/2}^{n-1/2}}{\Delta t} - \frac{\frac{c_{ij}}{\alpha_{ij}} (\lambda_4)_{ij}^n - \frac{c_{i,j-1}}{\alpha_{i,j-1}} (\lambda_4)_{i,j-1}^n}{\Delta z} - \frac{1}{2} c_{i,j-1/2} (\sigma_z)_{j-1/2} \left((\lambda_2)_{i,j-1/2}^{n+1/2} + (\lambda_2)_{i,j-1/2}^{n-1/2} \right) \\ = \sum_{q=1}^Q ((D_q^z)^n - (U_2)_{ij}^n) \delta_{(i_r)_q} \delta_{(j_r)_q}, \quad 0 \leq i \leq K+1, \quad 1 \leq j \leq M+1, \quad 0 \leq n \leq N, \\ \frac{(\lambda_3)_{ij}^{n+1} - (\lambda_3)_{ij}^n}{\Delta t} - \frac{(\alpha c \lambda_1)_{i+1/2,j}^{n+1/2} - (\alpha c \lambda_1)_{i-1/2,j}^{n+1/2}}{\Delta x} - \frac{(\alpha c \lambda_2)_{i,j+1/2}^{n+1/2} - (\alpha c \lambda_2)_{i,j-1/2}^{n+1/2}}{\Delta z} \\ - \frac{1}{2} c_{ij} (\sigma_x)_i ((\lambda_3)_{ij}^{n+1} + (\lambda_3)_{ij}^n) = 0, \quad 1 \leq i \leq K, \quad 1 \leq j \leq M, \quad 0 \leq n \leq N, \\ \frac{(\lambda_4)_{ij}^{n+1} - (\lambda_4)_{ij}^n}{\Delta t} - \frac{(\alpha c \lambda_1)_{i+1/2,j}^{n+1/2} - (\alpha c \lambda_1)_{i-1/2,j}^{n+1/2}}{\Delta x} - \frac{(\alpha c \lambda_2)_{i,j+1/2}^{n+1/2} - (\alpha c \lambda_2)_{i,j-1/2}^{n+1/2}}{\Delta z} \\ - \frac{1}{2} c_{ij} (\sigma_z)_j ((\lambda_4)_{ij}^{n+1} + (\lambda_4)_{ij}^n) = 0, \quad 1 \leq i \leq K, \quad 1 \leq j \leq M, \quad 0 \leq n \leq N, \end{array} \right.$$

avec la condition finale

$$(\lambda_1)_{i-1/2,j}^{N+1/2} - \frac{c_{i-1/2,j}(\sigma_x)_{i-1/2}}{2} (\lambda_1)_{i-1/2,j}^{N+1/2} = 0$$

$$(\lambda_2)_{i,j-1/2}^{N+1/2} - \frac{c_{i,j-1/2}(\sigma_z)_{j-1/2}}{2} (\lambda_2)_{i,j-1/2}^{N+1/2} = 0$$

$$(\lambda_3)_{ij}^{N+1} - \frac{c_{ij}(\sigma_x)_i}{2} (\lambda_3)_{ij}^{N+1} = D_x^+ \left((\alpha c)_{i+1/2,j} (\lambda_1)_{i-1/2,j}^{N+1/2} \right) + D_z^+ \left((\alpha c)_{j+1/2,j} (\lambda_2)_{i,j-1/2}^{N+1/2} \right)$$

$$(\lambda_4)_{ij}^{N+1} - \frac{c_{ij}(\sigma_z)_j}{2} (\lambda_4)_{ij}^{N+1} = D_x^+ \left((\alpha c)_{i+1/2,j} (\lambda_1)_{i-1/2,j}^{N+1/2} \right) + D_z^+ \left((\alpha c)_{j+1/2,j} (\lambda_2)_{i,j-1/2}^{N+1/2} \right),$$

équivalente à

$$\lambda_{ij}^{N+1} = 0,$$

et les conditions aux limites

$$(\lambda_3)_{K,j}^n = 0, \quad (\lambda_3)_{0,j}^n = 0, \quad (\lambda_4)_{i,M}^n = 0, \quad (\lambda_4)_{i,0}^n = 0,$$

$$(\alpha c)_{1/2,j} (\lambda_1)_j^n = (\mu_3)_j^n, \quad (\lambda_1)_{K+1,j}^n = 0, \quad (\alpha c)_{i,1/2} (\lambda_2)_{i,1}^n = (\mu_1)_i^n, \quad (\lambda_2)_{i,M+1}^n = 0.$$

Tout comme pour le problème posé en continu, il est donc possible de dériver de l'équation (3.24) le système discret adjoint du système discret utilisé pour résoudre le problème direct (3.20),(3.21). Ceci nécessite quelques efforts de calculs, cependant, comme cela a déjà été indiqué, le système obtenu n'est pas équivalent à la discrétisation du système adjoint obtenu en continu, et cet effort est absolument nécessaire pour pouvoir calculer le gradient discret avec une bonne précision.

3.4.6 Calcul du gradient des termes de régularisation

Il reste maintenant à calculer le gradient des termes de régularisation. Le principe du calcul est similaire à celui du cas continu. Pour $\tilde{\alpha}$ dans $\mathbb{R}^{(K+2) \times (M+2)}$,

$$\begin{aligned} \mathcal{R}'_1(\alpha).\tilde{\alpha} &= \lim_{t \rightarrow 0} \frac{\mathcal{R}_1(\alpha + t\tilde{\alpha}) - \mathcal{R}_1(\alpha)}{t} = \lim_{t \rightarrow 0} t \|D_x^- \tilde{\alpha}\|_{L^2(\Omega_d)}^2 + 2(D_x^- \alpha, D_x^- \tilde{\alpha})_{x,z} \\ &= -2(D_x^+ D_x^- \alpha, \tilde{\alpha})_{L^2(\Omega_d)} + 2(D_x^- \alpha_{K+1,j}, \tilde{\alpha}_{K,j})_{L^2(\Omega_d)} - 2(D_x^- \alpha_{1,j}, \tilde{\alpha}_{0,j})_{L^2(\Omega_d)}. \end{aligned}$$

Ainsi

$$(\mathcal{R}_1)'(\alpha)_{ij} = -2D_x^+ D_x^- \alpha_{ij}, \quad \forall (i,j) \in \mathbb{N}^2, \quad 1 \leq i \leq K-1, \quad 1 \leq j \leq M,$$

$$(\mathcal{R}_1)'(\alpha)_{0j} = -2D_x^- \alpha_{1,j}, \quad \forall j \in \mathbb{N}, \quad 1 \leq j \leq M,$$

$$(\mathcal{R}_1)'(\alpha)_{Kj} = 2D_x^- \alpha_{K+1,j}, \quad \forall j \in \mathbb{N}, \quad 1 \leq j \leq M.$$

D'autre part, comme

$$\mathcal{R}_2(\alpha) = \sum_{j=(j_r)_1}^{(j_r)_Q} ((\alpha_0)_j - \alpha_{(i_r)_1, j})^2 \Delta z,$$

un calcul direct du gradient discret pour le produit scalaire $L^2(\Omega_d)$ donne

$$\partial_\alpha(\mathcal{R}_2(\alpha))_{ij} = \begin{cases} -2((\alpha_0)_j - \alpha_{ij}) & \text{si } i = (i_r)_1, \quad (j_r)_1 \leq j \leq (j_r)_Q, \\ 0 & \text{sinon} \end{cases}$$

3.4.7 Expression finale du gradient discret

Il est possible de proposer une expression du gradient qui synthétise les résultats obtenus. Soient $\bar{U}^s(\alpha, h_s)$ la solution du problème direct discret (3.20),(3.21), associée à la CP $h_s = B_s q_s$ et au modèle α , $\bar{\lambda}^s$ la solution du problème adjoint à ce problème direct, alors le gradient discret de la fonctionnelle $\mathcal{J}^\varepsilon(\alpha, q_1, \dots, q_s)$ s'exprime

$\forall(i, j) \in \mathbb{N}^2 \quad 1 \leq i \leq K, \ 0 \leq j \leq M,$

$$\partial_\alpha \widehat{\mathcal{J}}^\varepsilon(\alpha, q_1, \dots, q_S)_{ij} = (1 - \delta_{i,0})(1 - \delta_{j,0})$$

$$\begin{aligned} & \left[- \sum_{s=1}^S \sum_{n=1}^N (\bar{\lambda}_3)_{ij}^n \frac{c_{ij}}{\alpha_{ij}^2} D_x^+(U_1^s)_{ij}^{n+1} \Delta t - \sum_{s=1}^S \sum_{n=1}^N (\bar{\lambda}_4)_{ij}^n \frac{c_{ij}}{\alpha_{ij}^2} D_z^+(U_2^s)_{ij}^{n+1} \Delta t \right. \\ & + \sum_{s=1}^S \sum_{n=1}^N \frac{(\bar{\lambda}_1)_{i-1/2,j}^n}{2} c_{i-1/2,j} D_x^-(U_3^s + U_4^s)_{ij}^n \Delta t + \sum_{s=1}^S \sum_{n=1}^N \frac{(\bar{\lambda}_1)_{i+1/2,j}^n}{2} c_{i+1/2,j} D_x^+(U_3^s + U_4^s)_{ij}^n \Delta t \\ & + \sum_{s=1}^S \sum_{n=1}^N \frac{(\bar{\lambda}_2)_{i,j-1/2}^n}{2} c_{i,j-1/2} D_z^-(U_3^s + U_4^s)_{ij}^n \Delta t + \sum_{s=1}^S \sum_{n=1}^N \frac{(\bar{\lambda}_2)_{i,j+1/2}^n}{2} c_{i,j+1/2} D_z^+(U_3^s + U_4^s)_{ij}^n \Delta t \Big] \\ & + \sum_{s=1}^S \sum_{n=1}^N \frac{(\bar{\lambda}_1)_{1/2,j}^n}{2} c_{1/2,j} D_x^-(U_3^s + U_4^s)_{1j}^n \Delta t \delta_{i,0} \sum_{s=1}^S \sum_{n=1}^N \frac{(\bar{\lambda}_2)_{i,1/2}^n}{2} c_{i,1/2} D_z^-(U_3^s + U_4^s)_{i1}^n \Delta t \delta_{j,0} \\ & - 2\varepsilon_1 D_x^+ D_x^- \alpha_{ij} (1 - \delta_{i,K}) (1 - \delta_{i,0}) (1 - \delta_{j,M}) (1 - \delta_{j,0}) \\ & - 2\varepsilon_1 D_x^- \alpha_{1j} \delta_{i,0} (1 - \delta_{j,0}) + 2\varepsilon_1 D_x^- \alpha_{K+1,j} \delta_{i,K} (1 - \delta_{j,0}) \\ & \left. - 2 \sum_{q=1}^Q \varepsilon_2 \left((\alpha_0)_j - \alpha_{(ij)} \right) \delta_{i,(i_r)_1} \delta_{j,(j_r)_q}, \right. \end{aligned}$$

et $\forall(i, n) \in \mathbb{N}^2, \ 0 \leq i \leq K, \ 0 \leq n \leq N,$

$$\partial_{q_s} \widehat{\mathcal{J}}^\varepsilon(\alpha, q_1, \dots, q_S)_i^n = \left(B_s^T \partial_{h_s} \mathcal{J}(\alpha, h_1, \dots, h_S) \right)_i^n,$$

$$\partial_{h_s} \mathcal{J}(\alpha, h_1, \dots, h_S)_i^n = \alpha_{i,1/2} c_{i,1/2} (\lambda_2^s)_i^n (1 - \delta_{i,0}) (1 - \delta_{n,0}).$$

Le problème posé à ce stade est celui de l'implémentation algorithmique de ce calcul. Celle-ci doit répondre à deux exigences qui s'opposent : minimiser le nombre d'opérations effectuées et minimiser la taille des données stockées en mémoire. La partie suivante est consacrée à la description des algorithmes mis en jeu pour parvenir à un tel compromis.

3.5 Algorithmes pour le calcul du gradient

3.5.1 Calcul de $\partial_{h_s} \mathcal{J}(\alpha, h_1, \dots, h_S)$

L'égalité (3.23) conduit à s'intéresser à l'évaluation de

$$\partial_{h_s} \mathcal{J}_s(\alpha, h_s), \quad s = 1, \dots, S.$$

On s'intéresse à la sous-fonctionnelle $\mathcal{J}_1(\alpha, h_1)$, renotée $\check{\mathcal{J}}(\alpha, h)$ pour plus de concision. L'évaluation du terme $\partial_h \check{\mathcal{J}}(\alpha, h)$ fait intervenir le calcul de l'état adjoint, dont l'implémentation doit être effectuée avec attention, dans l'optique de minimiser les temps de calcul.

La partie précédente donne

$$\partial_h \check{\mathcal{J}}(\alpha, h)_i^n = \alpha_{i,1/2} c_{i,1/2} (\lambda_2)_i^n (1 - \delta_{i,0}) (1 - \delta_{n,0}).$$

Ceci nécessite donc exclusivement le calcul de l'état adjoint λ à tous les pas de temps $n = 1, \dots, N$. La structure du système adjoint implique que λ se calcule à partir d'un état final nul ($\lambda^{N+1} = 0$), de manière rétrograde en temps (rétrorépropagation), par la résolution d'équations dans lesquelles U apparaît comme terme source. Pour déterminer λ au temps t^n , il est nécessaire de connaître U^n et λ^{n+1} . Par récurrence, l'état final λ^{N+1} étant nul, pour calculer λ au temps t^n , il faut donc connaître U^{n+1}, \dots, U^N . Cependant, les valeurs de l'état direct intervenant comme terme source sont les valeurs des vitesses de déplacement horizontal et vertical au niveau des récepteurs. Il n'est donc pas nécessaire de connaître les valeurs de l'état direct U dans tout le domaine pour déterminer l'état adjoint. L'algorithme de calcul qui s'ensuit repose donc sur le principe suivant. Dans un premier temps, l'état direct est calculé depuis son état initial $U^0 = 0$, jusqu'à l'état U^N , en stockant à chaque pas de temps les valeurs

$$(U_1)_{(i_r)_q, (j_r)_q}^n, (U_2)_{(i_r)_q, (j_r)_q}^n, \quad q = 1, \dots, Q.$$

Dans un second temps, l'état adjoint est calculé depuis l'état final $\lambda^{N+1} = 0$ jusqu'à l'état initial λ^0 en utilisant les valeurs de l'état direct nécessaires à la résolution du système adjoint. A chaque itération, la composante correspondante de $\partial_h \check{\mathcal{J}}(\alpha, h)_i^n$ est mise à jour.

Soit ϕ_n la fonction de propagation de n pas de temps $n = 1, \dots, N + 1$ pour l'état direct U , telle que

$$\forall k \in \mathbb{N}, \quad k = n, \dots, N + 1, \quad U^k = \phi_n(U^{k-n}, \alpha, h).$$

Cette fonction correspond à la résolution du problème direct discret pour n itérations en temps, à partir d'un état initial donné.

De plus, soit ψ_{-n} la fonction de retropropagation de n pas de temps pour l'état adjoint λ , telle que

$$\forall k \in \mathbb{N}, \quad k = 0, \dots, N + 1 - n, \quad \lambda^k = \psi_{-n}(U^{k+n}, \dots, U^{k+1}, \lambda^{k+n}, \alpha, h).$$

Cette fonction correspond à la résolution du système adjoint discret pour n itérations en temps, à partir d'un instant final donné, et des valeurs de l'état direct U au niveau des récepteurs pour les n pas de temps considérés. L'algorithme de calcul de $\partial_h \check{\mathcal{J}}(\alpha, h)$ peut être décrit comme suit.

```

Data :  $\alpha, h$ 
Result :  $\partial_h \check{\mathcal{J}}(\alpha, h)$ 
 $U^0 = 0;$ 
for  $n = 1, N$  do
   $U^n = \phi_1(U^{n-1}, \alpha, h);$ 
  stockage de  $(U_1)_{(i_r)_q, (j_r)_q}^n, (U_2)_{(i_r)_q, (j_r)_q}^n, q = 1, \dots, Q;$ 
end
 $\lambda^{N+1} = 0;$ 
for  $n = N, 0$  do
   $\lambda^n = \psi_{-1}(\lambda^{n+1}, U^n, \alpha, h);$ 
   $\partial_h \check{\mathcal{J}}(\alpha, h)_i^n = \alpha_{i,1/2} c_{i,1/2} (\lambda_2)_i^n (1 - \delta_{i,0}) (1 - \delta_{n,0});$ 
end
```

Algorithme 4 : Algorithme de rétropropagation pour $\partial_h \check{\mathcal{J}}(\alpha, h)$.

Cet algorithme permet un calcul optimal en temps du terme $\partial_h \check{\mathcal{J}}(\alpha, h)$, par la résolution successive du système direct et du système adjoint, en stockant les valeurs de l'état direct au niveau des récepteurs dans la première boucle en temps.

3.5.2 Calcul de $\partial_\alpha \mathcal{J}(\alpha, h_1, \dots, h_S)$

Le calcul du gradient suivant α de la fonctionnelle $\mathcal{J}(\alpha, h_1, \dots, h_S)$ possède une structure sensiblement différente du gradient suivant les CP. D'après l'équation (3.23), il doit en effet prendre en compte toutes les contributions des sous fonctionnelles $\mathcal{J}_s(\alpha, h_s)$, $s = 1, \dots, S$. La structure de l'algorithme de calcul correspondant est donc

```

Data :  $\alpha, h_1, \dots, h_S$ 
Result :  $\partial_\alpha \mathcal{J}(\alpha, h_1, \dots, h_S)$ 
 $\partial_\alpha \mathcal{J}(\alpha, h_1, \dots, h_S) = 0;$ 
for  $s = 1, S$  do
  calculer  $\partial_\alpha \mathcal{J}_s(\alpha, h_s);$ 
   $\partial_\alpha \mathcal{J}(\alpha, h_1, \dots, h_S) = \partial_\alpha \mathcal{J}(\alpha, h_1, \dots, h_S) + \partial_\alpha \mathcal{J}_s(\alpha, h_s);$ 
end
```

Algorithme 5 : Structure de l'algorithme de calcul du gradient suivant α .

Comme précédemment, la difficulté est ramenée au calcul des termes $\partial_\alpha \mathcal{J}_s(\alpha, h_s)$, qui nécessitent le calcul de l'état adjoint. Afin de détailler comment cette difficulté est appréhendée, le calcul du gradient suivant α de la fonctionnelle $\mathcal{J}_1(\alpha, h_1)$, à nouveau renotée $\check{\mathcal{J}}(\alpha, h)$, est présenté.

L'algorithme mis en place pour le calcul de $\partial_\alpha \check{\mathcal{J}}(\alpha, h)$ est assez similaire à celui utilisé dans la partie précédente pour le calcul de $\partial_h \check{\mathcal{J}}(\alpha, h)$. La principale différence est maintenant la nécessité de stocker l'état direct à chaque pas de temps non plus simplement au niveau des récepteurs mais dans tout le domaine. En effet, le gradient de la fonctionnelle suivant α est une somme en temps

$$\sum_{n=1}^N I_{ij}^n(U^n, \lambda^n),$$

où les termes I_{ij}^n sont définis par

$$\begin{aligned} I_{ij}^n(U^n, \lambda^n) = & (1 - \delta_{i,0})(1 - \delta_{j,0}) \left[-(\lambda_3)_{ij}^n \frac{c_{ij}}{\alpha_{ij}^2} D_x^+(U_1)_{ij}^{n+1} \Delta t - (\lambda_4)_{ij}^n \frac{c_{ij}}{\alpha_{ij}^2} D_z^+(U_2)_{ij}^{n+1} \Delta t \right. \\ & + \frac{(\lambda_1)_{i-1/2,j}^n}{2} c_{i-1/2,j} D_x^-(U_3 + U_4)_{ij}^n \Delta t + \frac{(\lambda_1)_{i+1/2,j}^n}{2} c_{i+1/2,j} D_x^+(U_3 + U_4)_{ij}^n \Delta t \\ & + \frac{(\lambda_2)_{i,j-1/2}^n}{2} c_{i,j-1/2} D_z^-(U_3 + U_4)_{ij}^n \Delta t + \frac{(\lambda_2)_{i,j+1/2}^n}{2} c_{i,j+1/2} D_z^+(U_3 + U_4)_{ij}^n \Delta t \left. \right] \\ & + \frac{(\lambda_1)_{1/2,j}^n}{2} c_{1/2,j} D_x^-(U_3 + U_4)_{1j}^n \Delta t \delta_{i,0} + \frac{(\lambda_2)_{i,1/2}^n}{2} c_{i,1/2} D_z^-(U_3 + U_4)_{i1}^n \Delta t \delta_{j,0}. \end{aligned} \quad (3.25)$$

La supposition est faite ici que les moyens de calcul utilisés permettent le stockage en mémoire de l'état direct U à tous les pas de temps et dans tout le domaine. Cette supposition est justifiée *a posteriori* par l'emploi d'une méthode de décomposition de domaines, qui permet de répartir la charge mémoire sur différents processeurs. Néanmoins, des algorithmes spécifiques ont été définis dans le cas où les capacités mémoires ne permettent pas d'effectuer ce stockage massif de données. Les lecteurs intéressés par ces questions se reporteront à l'annexe 3, qui fournit un éclairage sur la possibilité de stocker dans son intégralité l'état direct en mémoire, selon les cas d'application, et présente ces algorithmes. En supposant le stockage possible, l'algorithme utilisé pour le calcul de $\partial_\alpha \check{\mathcal{J}}(\alpha, h)$ est presque similaire à celui utilisé pour calculer $\partial_h \check{\mathcal{J}}(\alpha, h)$

```

Data :  $\alpha, h$ 
Result :  $\partial_\alpha \check{\mathcal{J}}(\alpha, h)$ 
 $U^0 = 0;$ 
for  $n = 1, N$  do
|    $U^n = \phi_1(U^{n-1}, \alpha, h);$ 
|   stockage de  $U^n$  ;
end
 $\lambda^{N+1} = 0;$ 
 $\partial_\alpha \check{\mathcal{J}}(\alpha, h) = 0;$ 
for  $n = N, 0$  do
|    $\lambda^n = \psi_{-1}(\lambda^{n+1}, U^n, \alpha, h);$ 
|    $\partial_\alpha \check{\mathcal{J}}(\alpha, h)_{ij} = \partial_\alpha \check{\mathcal{J}}(\alpha, h)_{ij} + I_{ij}^n(U^n, \lambda^n);$ 
end

```

Algorithme 6 : Algorithme de rétropropagation pour $\partial_\alpha \check{\mathcal{J}}(\alpha, h)$.

Cet algorithme permet un calcul efficace du gradient suivant α de la fonctionnelle $\partial_\alpha \check{\mathcal{J}}(\alpha, h)$, tout comme le précédent.

A ce stade, la méthode numérique peut d'ores et déjà être implémentée. Le calcul du gradient de la fonctionnelle $\check{\mathcal{J}}^\varepsilon(\alpha, q_1, \dots, q_S)$ est adapté au cadre discret nécessaire à la modélisation de la propagation des ondes acoustiques, et les algorithmes de calcul mis en jeu pour l'évaluation de ces quantités sont donnés. Néanmoins, une difficulté subsiste, celle du temps de calcul. En effet, pour les cas d'application visés, l'implémentation séquentielle de la méthode numérique donnerait lieu à des expériences numériques durant plusieurs mois. La seule solution pour réduire ces temps de calcul réside dans la parallélisation sur plusieurs processeurs de l'implémentation. Les deux algorithmes qui viennent d'être présentés sont en effet aisément parallélisables. Les fonctions ϕ et ψ au coeur de l'algorithme correspondent à la résolution d'une équation des ondes, respectivement directe et adjointe, résolution qui peut être parallélisée par une méthode de décomposition de domaines. Cette méthode, très utile pour paralléliser de coûteuses séquences de calcul, est exposée dans la partie suivante.

3.6 Parallélisation des algorithmes de calcul du gradient

3.6.1 Méthodes de décomposition de domaines

Une méthode de décomposition de domaines consiste en la réduction d'un problème de calcul posé sur un espace et un intervalle de temps donnés à un ensemble de sous-problèmes définis sur une partie de cet espace, et/ou de cet intervalle de temps. Ces problèmes sont dépendants les uns des autres, dans la mesure où la résolution de chacun d'entre eux nécessite la résolution des autres, néanmoins ces résolutions peuvent être opérées simultanément sur une architecture de calcul distribué. La force de ces méthodes, et l'engouement qu'elles suscitent, résident donc dans leur faculté à ramener la résolution, en principe séquentielle, d'un problème, à la résolution parallèle de sous-problèmes, qui nécessitent des coûts de calcul et un espace mémoire moindres.

Une application plus particulière des méthodes de décomposition de domaines à des problèmes de contrôle optimal régis par des équations aux dérivées partielles a été développée dès l'émergence de ces méthodes. L'idée sous-jacente est de décomposer non pas simplement la résolution d'équations aux dérivées partielles sur un domaine donné, mais le problème général de contrôle, en sous-problèmes de contrôle définis sur les sous-domaines, et de les résoudre simultanément. Les bases de cette méthode ont été posées par [Bensoussan et al., 1973], puis elles ont été approfondies notamment par Benamou [Benamou, 1994], [Benamou, Després, 1997]. Un travail assez exhaustif sur ces méthodes a été également effectué par [Bounaïm, 1999], au cours de ses travaux de thèse. De telles méthodes pourraient *a priori* s'appliquer au problème d'inversion de données sismiques de puits 2D. En adoptant le point de vue du contrôle optimal, l'état du système est décrit par les vitesses de déplacement au niveau des récepteurs, les équations PML pour le système d'Euler jouant le rôle des équations d'état, tandis que le modèle $\alpha(x, z)$ et les CP $h_s(x, t)$ jouent le rôle de la commande. Un aspect spécifique de notre méthode d'inversion est la dépendance non-linéaire du problème de contrôle par rapport à la commande α . Ces méthodes de décomposition de domaines pour des problèmes de contrôle optimal sont en effet définies le plus souvent pour des problèmes dépendant linéairement de la commande. De plus, dans notre cas d'application, la décomposition en sous-problèmes de contrôle du problème global requiert que chaque sous-domaine contienne au moins un récepteur, afin de pouvoir y définir une sous-fonctionnelle calculant un coût sismique local. En raison de la géométrie du problème, ceci ne mène pas à des découpages judicieux du domaine. Celui-ci doit en effet pouvoir être découpé en domaines de taille équivalente pour que la parallélisation soit efficace, car le nombre d'opérations à effectuer par sous-domaines est proportionnel à la taille de celui-ci, et le temps de calcul global est toujours donné par le processeur le plus lent, c'est à dire celui associé au sous-domaine dont la taille est la plus grande. Or, les récepteurs ne sont présents qu'au

centre de la partie supérieure du domaine. Le fait que l'état du système soit mesuré localement empêche donc l'utilisation directe de ces méthodes.

La parallélisation de l'inversion est donc introduite au niveau du calcul du gradient, le problème d'optimisation en lui-même (recherche linéaire, calcul de la direction de descente, mise à jour de l'approximation de la Hesse) étant résolu de manière séquentielle par un seul processeur. Cette approche n'est pas pénalisante, car le temps de calcul nécessaire à ces opérations est largement négligeable devant le temps requis par celui du gradient. Cette parallélisation est introduite via l'utilisation d'une méthode de décomposition de domaines, appliquée au problème de résolution de l'équation des ondes, et de son équation adjointe.

Dans les algorithmes présentés précédemment, ceci correspond à la parallélisation des fonctions $\phi_k(U^{n-1}, \alpha, h)$ et $\psi_k(\lambda^{n+1}, U^n, \alpha, h)$. Les schémas numériques directs et adjoints étant tout à fait similaires, seule la parallélisation du problème direct est présentée (c'est à dire celle de la fonction $\phi_k(U^{n-1}, \alpha, h)$), celle du problème adjoint suivant exactement le même principe.

3.6.2 Implémentation

Le domaine initial Ω_d est divisé en P sous-domaines $\Omega_p, p = 0, \dots, P$. La restriction d'une fonction f définie sur Ω_d à un domaine Ω_p est notée $f|_{\Omega_p}$. Pour p fixé, pour toute fonction f définie sur Ω_d , la notation $\tilde{f} = f|_{\Omega_p}$ est utilisée. La frontière de Ω_p est notée Γ_p . Soient k_p, K_p et m_p, M_p les bornes discrètes de Ω_p , telles que

$$\Omega_p = \{(x_i, z_j) \in \mathbb{R}^2, k_p \leq i \leq K_p, m_p \leq j \leq M_p\}$$

Soit g_p la condition au limite en pression, définie sur le bord du domaine Γ_p . Le problème discret résolu sur le sous-domaine p est alors, à une étape de temps n donnée,

$$\left\{
\begin{aligned}
& \frac{(\tilde{u}_x)_{i+1/2,j}^{n-1/2} - (\tilde{u}_x)_{i+1/2,j}^{n-3/2}}{\Delta t} - \tilde{\alpha}_{i+1/2,j} \tilde{c}_{i+1/2,j} \left[\frac{(\tilde{p}_x + \tilde{p}_z)_{i+1,j}^{n-1} - (\tilde{p}_x + \tilde{p}_z)_{ij}^{n-1}}{\Delta x} \right] \\
& + \frac{\tilde{c}_{i+1/2,j} (\tilde{\sigma}_x)_{i+1/2}}{2} \left((\tilde{u}_x)_{i+1/2,j}^{n-1/2} + (\tilde{u}_x)_{i+1/2,j}^{n-3/2} \right) = 0, \quad i = k_p + 1, K_p, \quad j = m_p, M_p, \\
& \frac{(\tilde{u}_z)_{i,j+1/2}^{n-1/2} - (\tilde{u}_z)_{i,j+1/2}^{n-3/2}}{\Delta t} - \tilde{\alpha}_{i,j+1/2} \tilde{c}_{i,j+1/2} \left[\frac{(\tilde{p}_x + \tilde{p}_z)_{i,j+1}^{n-1} - (\tilde{p}_x + \tilde{p}_z)_{ij}^{n-1}}{\Delta z} \right] \\
& + \frac{\tilde{c}_{i,j+1/2} (\tilde{\sigma}_z)_{j+1/2}}{2} \left((\tilde{u}_z)_{i,j+1/2}^{n-1/2} + (\tilde{u}_z)_{i,j+1/2}^{n-3/2} \right) = 0, \quad i = k_p, K_p, \quad j = m_p + 1, M_p, \\
& \frac{(\tilde{p}_x)_{ij}^n - (\tilde{p}_x)_{ij}^{n-1}}{\Delta t} - \frac{\tilde{c}_{ij}}{\tilde{\alpha}_{ij}} \frac{(\tilde{u}_x)_{i+1/2,j}^{n-1/2} - (\tilde{u}_x)_{i-1/2,j}^{n-1/2}}{\Delta x} + \frac{(\tilde{c}_{ij} \tilde{\sigma}_x)_i}{2} ((\tilde{p}_x)_{ij}^n + (\tilde{p}_x)_{ij}^{n-1}) = 0, \\
& \quad i = k_p + 1, K_p - 1, \quad j = m_p + 1, M_p - 1, \\
& \frac{(\tilde{p}_z)_{ij}^n - (\tilde{p}_z)_{ij}^{n-1}}{\Delta t} - \frac{\tilde{c}_{ij}}{\tilde{\alpha}_{ij}} \frac{(\tilde{u}_z)_{i,j+1/2}^{n-1/2} - (\tilde{u}_z)_{i,j-1/2}^{n-1/2}}{\Delta z} + \frac{(\tilde{c}_{ij} \tilde{\sigma}_z)_j}{2} ((\tilde{p}_z)_{ij}^n + (\tilde{p}_z)_{ij}^{n-1}) = 0, \\
& \quad i = k_p + 1, K_p - 1, \quad j = m_p + 1, M_p - 1,
\end{aligned} \tag{3.26}
\right.$$

avec les conditions aux limites

$$\left\{
\begin{aligned}
& (p_x + p_z)_{k_q,j}^n = (g_q^1)_j^n, \quad (p_x + p_z)_{K_q,j}^n = (g_q^2)_j^n, \\
& (p_x + p_z)_{i,m_q}^n = (g_q^3)_i^n, \quad (p_x + p_z)_{i,M_q}^n = (g_q^4)_i^n,
\end{aligned} \tag{3.27}
\right.$$

et des conditions initiales nulles.

La différence avec le problème posé sur le domaine entier Ω_d consiste donc simplement en la définition des conditions aux limites en pression. En effet, la décomposition du domaine Ω_d en sous-domaines Ω_p introduit des frontières internes artificielles, pour lesquelles les conditions de pression à imposer ne sont pas définies *a priori*. L'idée de base pour résoudre ce problème consiste donc à définir des recouvrements entre les sous-domaines, de manière à ce que la condition au bord pour un sous-domaine corresponde à une succession de points internes à un sous-domaine voisin. Celui-ci peut alors adresser les données de bord nécessaires à l'autre sous-domaine, par un échange de données standard.

Seules les données de pression sur les bords des sous-domaines étant nécessaires à la définition des schémas sur chaque sous-domaine, les recouvrements se limitent à

une bande d'une épaisseur de un point de maillage. Pour donner un exemple concret, dans le cas de la décomposition horizontale du domaine Ω_d en 2 sous-domaines Ω_1 et Ω_2 , les bornes choisies pour ces sous-domaines sont les suivantes

$$\begin{cases} k_1 = 0, \quad K_1 = K + 1, \quad m_1 = 0, \quad M_1 = m_2 + 1, \\ k_2 = 0, \quad K_2 = K + 1, \quad m_2 = M_1 - 1, \quad M_2 = M + 1, \end{cases}$$

Pour un découpage en deux sous-domaines de même taille, il reste à définir M_1 où m_2 , qui peuvent être pris égaux à $E(\frac{M+1}{2})$, où $E(x)$ désigne la fonction partie entière. En conséquence, les deux sous-domaines partagent la bande

$$\{(x_i, z_j), \quad i = 0, K + 1, \quad j = m_2, m_2 + 1\}$$

Pour un découpage plus complexe, le principe est exactement le même (cf fig.3.1).

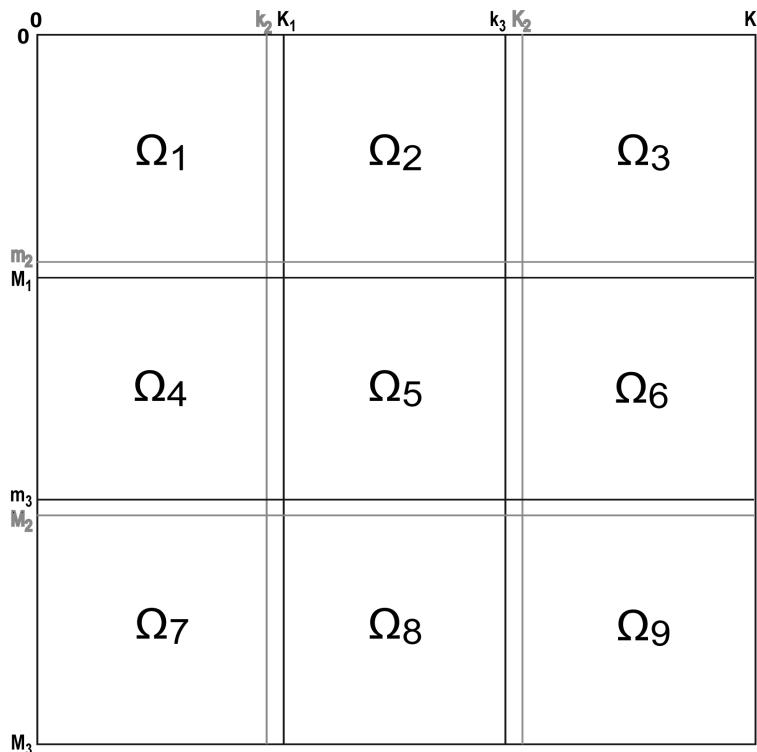


FIG. 3.1 – Découpage en 9 sous-domaines.

La version parallèle $\phi_k^{par}(U^n, \alpha, h)$ de la fonction $\phi_k(U^n, \alpha, h)$ repose donc sur l'idée suivante : à chaque itération en temps, chaque processeur résout le sous-problème discret associé au sous-domaine auquel il est rattaché, puis il échange avec ses voisins les données de pression correspondant aux conditions au bord utilisées à l'itération suivante.

Les découpages de Ω_d utilisés étant réguliers, chaque processeur a au plus 4 voisins. Un vecteur $V \in \mathbb{N}^4$ contient les identifiants des éventuels voisins dans l'ordre conventionnel suivant : Nord - Sud - Ouest - Est. Si le processeur considéré ne possède pas de voisins dans une direction donnée, la composante de V correspondante est initialisée à -1 . L'algorithme parallèle mis en place est décrit comme suit.

Data : U^{k-n}, α, h
Result : U^k

```

for  $i = 1, n$  do
   $U_p^{k-n+i} = \phi_1(U_p^{k-n+i-1}, \alpha, h);$ 
  if  $V_1 \neq -1$  then
     $\text{buffN} = ((p_x)_{i,m_p+1}^{k-n+i}, (p_z)_{i,m_p+1}^{k-n+i}), \quad i = k_p, K_p;$ 
    call envoi( $\text{buffN}$ ,  $V_1$ );
    call réception( $\text{buffS}$ ,  $V_1$ );
  end
  else
     $(p_x)_{i,m_p}^{k-n+i} + (p_z)_{i,m_p}^{k-n+i} = h_i^n, \quad i = k_p, K_p$ 
  end
  if  $V_2 \neq -1$  then
     $\text{buffS} = ((p_x)_{i,M_p-1}^{k-n+i}, (p_z)_{i,M_p-1}^{k-n+i}) \quad i = k_p, K_p;$ 
    call envoi( $\text{buffS}$ ,  $V_2$ );
    call réception( $\text{buffN}$ ,  $V_2$ );
  end
  else
     $(p_x)_{i,M_p}^{k-n+i} + (p_z)_{i,M_p}^{k-n+i} = 0, \quad i = k_p, K_p$ 
  end
  if  $V_3 \neq -1$  then
     $\text{buffO} = ((p_x)_{k_p+1,j}^{k-n+i}, (p_z)_{k_p+1,j}^{k-n+i}) \quad j = m_p, M_p;$ 
    call envoi( $\text{buffO}$ ,  $V_3$ );
    call réception( $\text{buffE}$ ,  $V_3$ );
  end
  else
     $(p_x)_{k_p,j}^{k-n+i} + (p_z)_{k_p,j}^{k-n+i} = 0, \quad j = m_p, M_p$ 
  end
  if  $V_4 \neq -1$  then
     $\text{buffE} = ((p_x)_{K_p-1,j}^{k-n+i}, (p_z)_{K_p-1,j}^{k-n+i}) \quad j = m_p, M_p;$ 
    call envoi( $\text{buffE}$ ,  $V_4$ );
    call réception( $\text{buffO}$ ,  $V_4$ );
  end
  else
     $(p_x)_{K_p,j}^{k-n+i} + (p_z)_{K_p,j}^{k-n+i} = 0, \quad j = m_p, M_p$ 
  end
end

```

Algorithme 7 : Algortihme parallèle ϕ_k^{par}

Il convient de remarquer que lorsqu'un processeur est dépourvu de voisin dans une direction donnée, la frontière du sous-domaine auquel il est rattaché constitue

une portion de la frontière du domaine initial Ω_d , dans cette direction. Dans ce cas, les conditions aux limites de Dirichlet originelles sont appliquées : CP au sommet du domaine, Dirichlet homogènes sur la frontière inférieure et les frontières latérales.

Lorsqu'un processeur possède un voisin, les routines *envoi* et *réception* permettent de communiquer avec lui. Pour plus de concision, il est ici supposé que la routine *réception*, comprend en plus la mise à jour des variables p_x et p_z après réception du message dans la mémoire tampon `buffN`, `buffS`, `buffO`, ou `buffE`, suivant le cas.

La même méthode est appliquée pour le problème adjoint, permettant de définir une version parallèle $\psi_{-n}^{par}(\lambda^{n-k+1}, U^{n-k}, \alpha, h)$ de la fonction $\psi_{-n}(\lambda^{n-k+1}, U^{n-k}, \alpha, h)$.

L'utilisation d'un schéma de Yee centré d'ordre 2, explicite, permet donc de procéder à une parallélisation par décomposition de domaines extrêmement simple. La taille des données échangées est limitée, car seule la donnée de la pression sur les bords internes des sous-domaines est nécessaire. De plus, comme ceci est expliqué dans la partie précédente et en annexe 3.7, cette décomposition de domaines permet de réduire la taille des données stockées par processeur, ce qui permet de mettre en place les algorithmes très efficaces proposés dans la partie précédente. Cette parallélisation permet de mettre en oeuvre la méthode numérique d'inversion en réduisant les temps de calcul à quelques dizaines d'heures. Il est cependant important d'étudier attentivement la scalabilité de cette parallélisation, pour voir quels en sont les éléments limitants. Ceci est effectué dans le chapitre 4, consacré notamment aux performances numériques de la méthode d'inversion mise en place. Avant de conclure, quelques indications sur l'utilisation de l'interface de programmation parallèle MPI (Message Passing Interface) pour l'implémentation des algorithmes parallèles présentés ici sont données dans la partie suivante.

3.6.3 Utilisation de l'interface MPI

Loin d'une présentation exhaustive de l'interface de programmation parallèle MPI, ce paragraphe est dédié à quelques rappels et commentaires sur son utilisation. Une documentation exhaustive sur les routines disponibles et une description globale de l'interface ont été mises en ligne par l'IDRIS [Requena et al., 2004].

Le principe adopté par MPI est simple : le même code est lu par tous les processeurs recrutés. Lors de l'initialisation du calcul, chacun d'entre eux se voit attribué un numéro identifiant, ou "rang". C'est au sein du code que sont réparties les tâches à effectuer, suivant le rang du processeur. Cet identifiant permet également aux processeurs de communiquer entre eux. Les communications peuvent se faire d'un processeur à un autre (échange point à point) ou de manière collective. Dans ce cas un ou plusieurs processeurs communiquent avec plusieurs processeurs en distribuant

les données à échanger à un ensemble de processeurs. Ce type d'échange collectif permet d'induire une hiérarchie entre les processeurs. Un processeur maître est défini, chargé de lire les fichiers d'initialisation et de construire les principales structures de données, qu'il distribue ensuite aux processus esclaves, en charge de l'exécution de noyau de calcul. De même, à l'issue du calcul, ce processeur maître est en charge de la collecte des résultats en provenance des processus esclaves, qu'il peut ensuite écrire sur le disque.

Comme expliqué dans 3.6.1, dans l'implémentation de la méthode d'inversion de données sismiques de puits 2D, le processus maître est également en charge de la boucle d'optimisation principale (Quasi-Newton l-BFGS). A chaque itération de cette boucle, le gradient de la fonctionnelle et sa valeur au point courant lui sont demandés. Le processus maître se charge alors de relayer la demande auprès des processus esclaves qui fournissent le gradient par l'utilisation de l'algorithme 7.

Pour implémenter cet algorithme, il s'agit d'abord de déterminer une grille de découpage permettant de faire correspondre à chaque processeur un sous-domaine. Cette étape consiste en une spatialisation de la répartition des processeurs. Une fois cette grille établie, les processeurs peuvent prendre connaissance des identifiants de leur éventuels voisins (construction du vecteur V). On a vu dans la partie précédente qu'à chaque itération en temps de l'algorithme de rétropropagation, les processeurs échangent les données au bord de leur domaine avec leurs voisins. Ces communications sont effectuées par des échanges point à point, en se servant des identifiants des processeurs voisins. Il est important de noter que l'implémentation de cet algorithme requiert que ces échanges soient faits de manière non bloquante.

En effet, si chaque processeur attend, à l'issue de l'envoi de ses données, que le destinataire de son message acquitte réception de celui-ci, avant de poursuivre la séquence d'instructions (envoi bloquant), la situation est figée, chaque processeur attendant la réponse de son homologue. Il est donc important que chaque processeur poursuive sa séquence d'instructions après l'envoi de son message. Pour cela, des échanges point à point non bloquants sont utilisés. Ceci permet, de plus, de recouvrir les temps de communication. Une fois le message envoyé, il est pris en charge par l'architecture de communication du cluster, et avant la réception du message en provenance de leurs voisins, les processeurs peuvent effectuer des calculs propres au sous-domaine dont ils ont la charge. C'est une optimisation de temps de calcul courante pour la parallélisation par décomposition de domaines.

Une importante difficulté rencontrée a trait à la robustesse de la décomposition de domaines. En effet, il est impératif que les résultats fournis en bout de chaîne par l'algorithme d'optimisation soient les mêmes, quelque soit le nombre de sous-domaines et la configuration du découpage utilisés. Ceci peut sembler évident, néanmoins, dans le cas de l'application développée, le respect de ce principe demande des efforts parti-

culiers. En effet, d'une part, l'algorithme d'optimisation est très sensible aux valeurs du gradients. Une faible perturbation sur ses valeurs peut entraîner un cheminement dans l'espace des solutions très différent. D'autre part, les valeurs du gradient suivant α sont obtenues par accumulation de termes dont les ordres de grandeur sont très différents, certaines de ces valeurs pouvant être extrêmement faibles. Ainsi, si l'ordre d'accumulation n'est pas toujours strictement le même, suivant la décomposition du domaine utilisée, les erreurs d'arrondis produisent des perturbations sur les valeurs du gradient suffisantes pour modifier le comportement de l'algorithme d'optimisation. Un soin particulier doit donc être apporté pour veiller à ce que cet ordre d'accumulation soit identique, quelle que soit la décomposition du domaine.

Conclusion

Cette partie, consacrée au calcul numérique et l’algorithmie nécessaires à la mise en place de la méthode d’inversion 2D de données sismiques de puits, représente une part importante du travail fourni au cours de cette thèse. Chaque étape de la conception de la méthode a été réalisée dans un souci constant de compromis entre le temps de calcul et l’espace mémoire requis pour le stockage des données. L’enjeu est de déterminer une méthode robuste, bien adaptée au problème, et pouvant être étendue à un cadre tri-dimensionnel. Les différentes étapes de sa conception sont retracées dans ce qui suit.

Tout d’abord, la modélisation de la propagation des ondes acoustiques dans le sous-sol nécessite l’utilisation de conditions absorbantes. L’utilisation des équations PML adaptées au système d’Euler à partir des équations de Bérenger pour les ondes électromagnétiques [Bérenger, 1994] est choisie à cet effet. Ceci permet de définir le problème direct de modélisation, puis le problème inverse correspondant.

Ensuite, la conception d’une méthode d’optimisation adaptée à ce problème inverse permet de s’affranchir de problèmes d’indétermination liés au couplage des inconnues, en améliorant de manière radicale la convergence de l’algorithme d’optimisation et en proposant des solutions cohérentes sur le plan de la géophysique.

Cette méthode d’optimisation requiert le calcul du gradient de la fonctionnelle minimisée, qui est effectué par la méthode de l’état adjoint. Cette méthode permet d’obtenir une expression du gradient en fonction de l’état direct du problème, c’est-à-dire la pression et les vitesses de déplacement, et de leur variables adjointes, solutions d’une équation d’onde couplée au système PML, se rétropropageant à partir d’un état final nul.

La discréétisation du problème permet ensuite de définir une expression discrète du gradient adaptée au schéma numérique utilisé pour implémenter la résolution des équations directes. Celle-ci repose sur la transposition au cadre discret du problème d’optimisation, qui permet la détermination de la version discrète du système adjoint.

Des algorithmes de rétropropagation appropriés sont alors mis en place pour parvenir à un calcul efficace de l’expression du gradient fournie par la méthode de l’état adjoint. Ces algorithmes ont été implantés dans leur version la plus simple, ceci étant permis par l’utilisation du code de calcul sur une machine parallèle possédant une mémoire suffisante. Néanmoins des versions alternatives applicables à un contexte pour lequel le stockage mémoire doit être limité, soit en raison de l’importance du cas traité, soit en raison de limitations inhérentes aux moyens de calcul disponibles, ont été testées avec succès, et sont présentées en annexe,

Finalement, les calculs de propagation de l'état direct et de rétropropagation de l'état adjoint sont parallélisés par une méthode de décomposition de domaine, afin de pouvoir disposer de la puissance de calcul d'une machine à architecture parallèle et réduire efficacement les temps de calcul.

La conjonction de ces efforts permet de définir une méthode d'optimisation parallèle originale, adaptée à l'inversion non linéaire 2D de données sismiques de puits, et à son extension 3D. Il convient maintenant d'analyser son comportement tant du point de vue numérique (temps de calcul, convergence), que du point de vue géophysique (résolution de la méthode, sensibilité). Cette analyse est menée dans le chapitre suivant.

3.7 Annexes

Annexe 1. Introduction d'un terme de régularisation sur les CP après interpolation

L'interpolation permet de prendre en compte la majeure partie de l'information *a priori* sur la régularité des CP, néanmoins, dans certains cas, par exemple si le nombre de traces sélectionnées est faible, il peut être intéressant d'ajouter un terme de régularisation, qui va porter directement sur les traces sélectionnées, et non sur la CP entière. Ce terme est noté $\widehat{R}_3(q_1, \dots, q_s)$.

On s'intéresse à une CP particulière, notée $h(x, t)$ à laquelle correspond un vecteur de traces

$$q(t) = (q_1(t), \dots, q_p(t)) = (h(x_{i_1}, t), \dots, h(x_{i_p}, t)), \quad (i_1, \dots, i_p) \in \mathbb{N}^p$$

Les points x_{i_k} sont indexés dans l'ordre croissant,

$$x_{i_k} < x_{i_{k+1}}, \quad k = 1, \dots, p.$$

La CP $h(x, t)$ présente une régularité impliquant de faibles variations suivant la direction $\gamma \in \mathbb{R}^2$. Le principe employé est de minimiser l'écart des points trace à trace, suivant cette direction γ . Considérons les deux premières traces sélectionnées,

$$h(x_{i_1}, t), \quad h(x_{i_2}, t).$$

La régularité suivant la direction γ implique que les points $h(x_{i_1}, t)$ doivent être proches de $h(x_{i_2}, t + \tau_{12})$ avec

$$\tau_{12} = \frac{\gamma_2}{\gamma_1}(x_{i_2} - x_{i_1})$$

Si le coefficient directeur $\frac{\gamma_2}{\gamma_1}$ est positif, les décalages en temps $\tau_{k,k+1}$ sont positifs, et le terme de régularisation à appliquer à $q(t)$ est

$$r(q) = \sum_{k=1}^{p-1} \int_0^{T-\tau_{k,k+1}} (q_k(t) - q_{k+1}(t + \tau_{k,k+1}))^2 dt.$$

Dans le cas contraire,

$$r(q) = \sum_{k=1}^{p-1} \int_{\tau_{k,k+1}}^T (q_k(t) - q_{k+1}(t + \tau_{k,k+1}))^2 dt.$$

De plus, le gradient de $r(q)$ est donné par

$$\partial_q r(q) = \begin{bmatrix} \partial_{q_1} r(q) \\ \vdots \\ \partial_{q_p} r(q) \end{bmatrix},$$

avec

$$\partial_{q_k} r(q) = 2(q_k(t) - q_{k+1}(t + \tau_{k,k+1})) - 2(q_{k-1}(t) - q_k(t + \tau_{k-1,k})), \text{ pour } k = 2, \dots, p$$

$$\partial_{q_1} r(q) = 2(q_1(t) - q_2(t + \tau_{12})), \quad \partial_{q_p} r(q) = -2(q_{p-1}(t) - q_p(t + \tau_{p-1,p})),$$

Ainsi, le terme de régularisation $\widehat{R}_3(q_1, \dots, q_S)$ s'exprime

$$\widehat{R}_3(q_1, \dots, q_S) = \sum_{s=1}^S r(q_s),$$

et son gradient est

$$\partial_{q_s} \widehat{R}_3(q_1, \dots, q_S) = \partial_{q_s} r(q_s).$$

Annexe 2. Calcul des matrices d'interpolation B_s

La construction des matrices B_s repose sur un schéma d'interpolation à 4 points. Le nombre p de traces en temps $tr_i(t)$, $i = 1, \dots, p$ sélectionnées ne peut être inférieur à 3. Les trois traces sélectionnées par défaut sont les traces extrêmales $h(-a - l, t)$, $h(a + l, t)$, et la trace milieu $h(0, t)$. Cette configuration minimale découle du problème d'indétermination décrit dans le chapitre 2. La composante bien déterminée de l'inconnue $h(x, t)$ se situe en effet au voisinage du puits en $x = 0$. L'indétermination augmente à mesure que l'on s'en éloigne. Cette configuration minimale peut ensuite être raffinée en ajoutant des traces intermédiaires. Le choix du nombre de traces à utiliser est un compromis entre l'importance accordée à l'information *a priori* introduite par l'interpolation linéaire suivant la direction γ_s , et l'information sismique. Plus le nombres de traces est important, plus le nombre de degrés de liberté est important, et plus l'information sismique devient importante. Dans le cas contraire, c'est l'information de régularité qui prime.

Dans la suite, afin de simplifier les notations, l'indice s est abandonné. Le nombre de pas de discrétisation en x est $K+1$, et $N+1$ est le nombre de pas de discrétisation en temps. Soient

$$\Delta x = \frac{2(a + l)}{K + 1}, \quad \Delta t = \frac{T}{N + 1},$$

et la grille discrète en espace temps

$$G = \{x_i = -a - l + i\Delta x, t^n = n\Delta t, \quad i = 0, \dots, K + 1, \quad n = 0, \dots, N + 1\}.$$

Soit une BC $h(x, t)$. Sur la grille, on utilise la notation $h(x_i, t^n) = h_i^n$, et on définit le vecteur

$$H \in \mathbb{R}^{(K+1)(N+1)}, \quad H = (h_1^1, \dots, h_1^{N+1}, \dots, h_{K+1}^1, \dots, h_{K+1}^{N+1}).$$

Une sélection de p traces en temps est effectuée sur la BC $h(x, t)$, notée

$$q(t) = (h(x_{i_1}, t), \dots, h(x_{i_p}, t)), \quad (i_1, \dots, i_p) \in I^p, \quad \text{avec } I = \{i \in \mathbb{N}, \quad i = 0, \dots, K + 1\}.$$

Le vecteur correspondant est

$$Q \in \mathbb{R}^{p(N+1)}, \quad Q = (h_{i_1}^1, \dots, h_{i_1}^{N+1}, \dots, h_{i_p}^1, \dots, h_{i_p}^{N+1}).$$

Pour une direction dans le plan $\gamma \in \mathbb{R}^2$ donnée, une matrice d'interpolation $B \in \mathbb{M}_{NK, Np}$ est donc recherchée, telle que $\tilde{H} = BQ$ soit l'interpolation linéaire suivant la direction γ des traces sélectionnées Q . L'objet de cette annexe est de présenter la construction de cette matrice. La figure 3.2 représente la situation pour deux traces sélectionnées x_{i_1} et x_{i_2} . Le but est de déterminer la valeur de la BC au point (x_i, t^n) pour

$$i = i_1 + 1, \dots, i_2 - 1, \quad n = 0, \dots, N + 1.$$

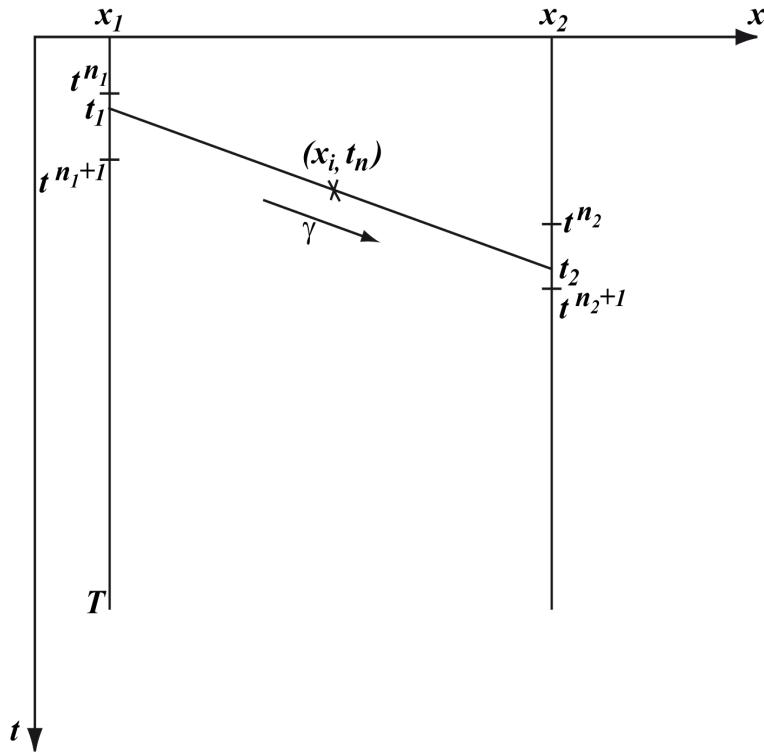


FIG. 3.2 – Schéma d'interpolation à quatre points.

Dans un premier temps, il convient de calculer l'intersection de la droite de direction γ passant par (x_i, t^n) avec les traces sélectionnées. Ces points d'intersection sont notés (x_{i1}, t_1) et (x_{i2}, t_2) , avec

$$t_1 = \frac{\gamma_2}{\gamma_1}(x_{i1} - x_i) + t^n, \quad t_2 = \frac{\gamma_2}{\gamma_1}(x_{i2} - x_i) + t^n.$$

Ceci permet de déterminer les 4 noeuds de la grille qui vont servir à l'interpolation, c'est-à-dire

$$(x_{i1}, t^{n1-1}), \quad (x_{i1}, t^{n1}), \quad (x_{i2}, t^{n2-1}), \quad (x_{i2}, t^{n2}),$$

avec

$$t_1 \in [t^{n1-1}, t^{n1}], \quad t_2 \in [t^{n2-1}, t^{n2}].$$

On procède par deux interpolations successives. Dans un premier temps, on effectue une interpolation entre les points

$$(x_{i1}, t^{n1-1}), \quad (x_{i1}, t^{n1}),$$

puis entre les points

$$(x_{i2}, t^{n2-1}), \quad (x_{i2}, t^{n2}),$$

qui donnent les quantités

$$\begin{cases} f_1 = \frac{1}{\Delta t}(t_1 - t^{n_1-1})h(x_{i_1}, t^{n_1}) + \frac{1}{\Delta t}(t_1 - t^{n_1})h(x_{i_1}, t^{n_1-1}) \\ f_2 = \frac{1}{\Delta t}(t_2 - t^{n_2-1})h(x_{i_2}, t^{n_2}) + \frac{1}{\Delta t}(t_2 - t^{n_2})h(x_{i_2}, t^{n_2-1}). \end{cases}$$

Dans les cas tels que

$$t_1 > T, \quad \text{ou} \quad t_2 > T, \quad \text{ou} \quad t_1 < 0, \quad \text{ou} \quad t_2 < 0,$$

l'interpolation mène hors du domaine de définition de $h(x, t)$. La valeur f_i , $i = 1, 2$ est alors fixée arbitrairement à 0, selon que le "débordement" a lieu pour la variable t_1 ou t_2 .

Dans un second temps, on procède à l'interpolation de ces deux valeurs entre les points x_{i_1}, t_1 et x_{i_2}, t_2 . Soient

$$\begin{cases} d_{12} = \sqrt{(x_{i_2} - x_{i_1})^2 + (t_2 - t_1)^2} \\ d_{1i} = \sqrt{(x_i - x_{i_1})^2 + (t^n - t_1)^2} \\ d_{i2} = \sqrt{(x_{i_2} - x_i)^2 + (t_2 - t^n)^2}. \end{cases}$$

Alors

$$\begin{aligned} \tilde{H}_{n+(i-1)(N+1)} &= \frac{1}{d_{12}} (d_{1i}f_2 + d_{i2}f_1) = \\ &\frac{d_{1i}}{d_{12}} \frac{(t_2 - t^{n_2-1})}{\Delta t} h(x_{i_2}, t^{n_2}) + \frac{d_{1i}}{d_{12}} \frac{(t_2 - t^{n_2})}{\Delta t} h(x_{i_2}, t^{n_2-1}) + \\ &\frac{d_{i2}}{d_{12}} \frac{(t_1 - t^{n_1-1})}{\Delta t} h(x_{i_1}, t^{n_1}) + \frac{d_{i2}}{d_{12}} \frac{(t_1 - t^{n_1})}{\Delta t} h(x_{i_1}, t^{n_1-1}). \end{aligned}$$

Soit,

$$\begin{aligned} \tilde{H}_{n+(i-1)(N+1)} &= \frac{d_{1i}}{d_{12}} \frac{(t_2 - t^{n_2-1})}{\Delta t} Q_{n_2+(i_2-1)(N+1)} + \frac{d_{1i}}{d_{12}} \frac{(t_2 - t^{n_2})}{\Delta t} Q_{n_2-1+(i_2-1)(N+1)} + \\ &\frac{d_{i2}}{d_{12}} \frac{(t_1 - t^{n_1-1})}{\Delta t} Q_{n_1+(i_1-1)(N+1)} + \frac{d_{i2}}{d_{12}} \frac{(t_1 - t^{n_1})}{\Delta t} Q_{n_1-1+(i_1-1)(N+1)}. \end{aligned}$$

Ainsi, la matrice B est une matrice creuse dont tous les coefficients sont nuls, exceptés, pour $n = 0, \dots, N+1$, $i = i_1 + 1, \dots, i_2 - 1$

$$\begin{aligned} B_{n+(i-1)(N+1), n_1-1} &= \frac{d_{i2}}{d_{12}} \frac{(t_1 - t^{n_1})}{\Delta t} \\ B_{n+(i-1)(N+1), n_1} &= \frac{d_{i2}}{d_{12}} \frac{(t_1 - t^{n_1-1})}{\Delta t} \\ B_{n+(i-1)(N+1), n_2-1+N+1} &= \frac{d_{1i}}{d_{12}} \frac{(t_2 - t^{n_2})}{\Delta t} \\ B_{n+(i-1)(N+1), n_2+N+1} &= \frac{d_{1i}}{d_{12}} \frac{(t_2 - t^{n_2-1})}{\Delta t}. \end{aligned}$$

De plus, pour $i = i_1$ et $i = i_2$, il n'y a pas d'interpolation, puisqu'on restitue la valeur des traces en temps sélectionnées dans le vecteur Q . Ainsi,

$$B_{n+(i_1-1)(N+1), n} = 1, \quad n = 0, \dots, N+1$$

$$B_{n+(i_2-1)(N+1), n+N+1} = 1, \quad n = 0, \dots, N+1.$$

La constitution de la matrice B en entier se fait en itérant ce procédé pour chaque couple de traces voisines sélectionnées pour l'interpolation.

Annexe 3. Calcul de $\partial_\alpha \mathcal{J}(\alpha, h)$ pour une mémoire limitée

Annexe 3.1 Intérêt

Suivant les moyens de calcul disponibles, il peut être impossible de stocker l'état direct U^n à tous les pas de temps. En effet, pour n fixé, U^n représente une matrice de points définis sur la grille discrète en espace. La discréétisation spatiale étant fine, cette matrice est de grande taille. De plus, pour assurer la stabilité du schéma numérique, le pas de temps choisi doit respecter la condition CFL,

$$\|c\|_\infty \Delta t \leq \frac{1}{\sqrt{\frac{1}{\Delta x^2} + \frac{1}{\Delta z^2}}}.$$

Ainsi, le pas de temps Δt est d'autant plus petit que les pas de discréétisation spatiale Δx et Δz le sont également. Pour fixer les idées, pour le problème test du chapitre 4, les pas de discréétisation Δx et Δz sont pris égaux respectivement à 12 m et 8 m. Le domaine Ω est un rectangle d'environ 800 m de large pour 2500 m de profondeur. Le nombre de point en espace résultant est de l'ordre de 22000. La vitesse maximale utilisée est de l'ordre de 6000 m.s⁻¹ et le pas de temps correspondant par CFL est de 1 ms. Le temps d'enregistrement utilisé étant fixé à 2 s, le nombre de points en temps s'élève à 2000. La variable U_{ij}^n , qui possède 4 composantes, nécessite donc pour son stockage un espace mémoire de l'ordre de 1.5 Go. Cette quantité reste dans les limites admissibles de l'espace mémoire disponible sur un ordinateur dédié au calcul.

Cependant, la discréétisation en espace à échelle plus fine, par exemple $\Delta x = 8$ m et $\Delta z = 4$ m, a pour effet de multiplier le nombre de points en espace (62500), mais également le nombre de points en temps. En effet, le pas de temps choisi devant respecter la CFL, il s'abaisse aux alentours de 0.5 ms, ce qui produit 4000 points en temps pour le même temps d'enregistrement. Au final, l'espace mémoire requis pour la variable U_{ij}^n est de 8 Go, ce qui commence à dépasser les tailles standards de mémoire vive disponible.

De plus, dans l'optique d'une application de ces résultats au problème 3D, le nombre de points à stocker devient alors véritablement problématique. Dans ce cas pour lequel une discréétisation fine est utilisée, la définition d'une extension suivant la dimension y de la même taille que l'extension latérale x , pareillement discréétisée, mène à une taille mémoire qui dépasse le To pour le stockage de U_{ij}^n .

La parallélisation par décomposition de domaines procure le double avantage de limiter la taille des données à stocker par processeur, tout en augmentant la taille totale de mémoire vive disponible (sans compter le gain en temps de calcul). La parallélisation sur n processeurs pour des sous-domaines de même taille permet de diviser l'espace mémoire requis par processeur par n . Les moyens de calcul

disponibles à l'IFP ont permis d'utiliser simultanément jusqu'à 128 processeurs, disposant chacun de 2 Go de mémoire. Dans le cadre de l'application 2D traitée dans le chapitre 4, ceci se révèle amplement suffisant, la taille des données à stocker par processeur atteignant environ 80 Mo par processeur. En revanche, dans le cadre d'une application 3D, ce nombre de processeurs serait notoirement insuffisant, la taille des données par processeur avoisinant alors les 8 Go.

La conclusion de cette analyse est la suivante. Dans le cadre du cas expérimental sur lequel l'application a été testée, les moyens de calcul disponibles à l'IFP, et dans les centres de calcul en général, sont suffisamment puissants pour permettre d'utiliser les algorithmes procédant au stockage massif de l'état direct U_{ij}^n présentés dans la partie 5. Cependant, la taille des données à stocker dépend de nombreux paramètres : finesse de la discrétisation, vitesse maximale de propagation des ondes (CFL), temps d'enregistrement total, et bien sûr taille du domaine imaginé. Ceci peut conduire à la définition de cas tests pour lesquels ces algorithmes pourraient être mis en difficulté malgré l'utilisation de la méthode de décomposition de domaines. Ceci est d'autant plus vraisemblable dans le cadre d'une application 3D de la méthode. Ainsi, des algorithmes spécifiques doivent être mis en place, permettant de calculer $\partial_\alpha \mathcal{J}(\alpha, h)$ sans stocker intégralement les valeurs de l'état direct U_{ij}^n en temps et en espace, et sans non plus pénaliser le temps de calcul par le recours à un nombre d'opérations trop important. C'est ce qui est présenté dans cette annexe.

Annexe 3.2 Principe de l'algorithme

Le principe de la solution retenue se décompose en deux étapes. Dans un premier temps, l'état direct à l'instant final U^N est calculé, par résolution du problème direct. Il est en effet nécessaire de connaître l'état direct au temps final pour amorcer la résolution du problème adjoint par rétropropagation. En revanche, seul l'état final est stocké en mémoire. Dans un second temps, l'état adjoint et l'état direct sont calculés conjointement dans une boucle de rétropropagation, à partir du temps final. Chaque itération de cette boucle consiste en le calcul de l'état direct et de l'état adjoint au temps $n - 1$ à partir du temps n , et permet le calcul successif des termes $I^n(U^n, \lambda^n)$, qui sont accumulés dans une variable initialisée à 0. L'algorithme correspondant est

```

Data :  $\alpha, h$ 
Result :  $\nabla \mathcal{J}(\alpha, h)$ 
 $U^0 = 0;$ 
 $U^N = \phi_N(U^0, \alpha, h);$ 
 $\lambda^{N+1} = 0;$ 
 $\partial_\alpha \mathcal{J}(\alpha, h) = 0;$ 
for  $n = N, 0$  do
     $U^n = \phi_{-1}(U^{n+1}, \alpha, h);$ 
     $\lambda^n = \psi_{-1}(\lambda^{n+1}, U^n, \alpha, h);$ 
    calculer  $I^n(U^n, \lambda^n)$  par (3.25);
     $\partial_\alpha \mathcal{J}(\alpha, h) = \partial_\alpha \mathcal{J}(\alpha, h) + I^n$ 
end

```

Algorithme 8 : Algorithme de rétropropagation classique

La différence avec la version pour laquelle le stockage mémoire est possible est donc le recalcul de U^n à chaque itération de la boucle de rétropropagation. Le nombre d'opérations en est donc augmenté d'autant. Cependant, ceci repose sur la possibilité de pouvoir effectivement calculer l'état direct par rétropropagation. Ceci est normalement possible pour les équations d'ondes, car ces équations sont réversibles en temps.

Néanmoins, lorsque les ondes se propagent dans des milieux absorbants, le calcul rétrograde de la propagation montre des instabilités. La modélisation acoustique de la propagation dans le sous-sol ne prenant pas en compte d'éventuels phénomènes d'absorption, cette hypothèse est, de prime abord, valide. Cependant, l'introduction de couches PML, qui sont des milieux absorbants, change la donne. *De facto*, la rétropropagation de l'état direct devient instable dans les couches PML, ce qui interdit l'utilisation de l'algorithme classique de rétropropagation présenté ci-dessus. Des aménagements de cet algorithme doivent donc être trouvés afin de pouvoir calculer les termes de gradient liés à la PML.

Ceci soulève de plus une question légitime : pourquoi est-il nécessaire de calculer les valeurs du gradient par rapport à α dans la couche PML ? En effet, cette couche est artificielle, les géophysiciens ne sont donc pas intéressés par le calcul des valeurs d'impédance dans cette zone. En réalité, l'impédance dans cette zone est fixée par continuité normale avec le bord, comme cela a été présenté dans la première partie. C'est cette prolongation par continuité qui oblige à pouvoir calculer les termes du gradient par rapport à α dans la couche PML. Ceci est présenté dans le paragraphe suivant. La solution implémentée pour surmonter cette difficulté est donnée ensuite.

Annexe 3.3 Nécessité du calcul de $\partial_\alpha \mathcal{J}(\alpha, h)$ dans la PML

Comme cela a été mentionné dans la première partie, afin d'optimiser le comportement absorbant de la couche PML, les modèles d'impédance et de vitesse sont prolongés par continuité en son sein. Définissons $\mathring{\alpha} = \alpha|_{\mathcal{O}_d}$ la restriction de α aux valeurs prises dans le domaine d'intérêt, et $\widehat{\alpha} = \alpha|_{\Omega_d \setminus \mathcal{O}_d}$ restriction de α aux valeurs prises dans la couche PML, telles que

$$\left\{ \begin{array}{ll} \widehat{\alpha}_{ij} = \mathring{\alpha}_{K'j}, & 1 \leq i \leq K' - 1, \quad 1 \leq j \leq M - M' \\ \widehat{\alpha}_{ij} = \mathring{\alpha}_{K-K'j}, & K' + 1 \leq i \leq K, \quad 1 \leq j \leq M - M' \\ \widehat{\alpha}_{ij} = \mathring{\alpha}_{i,M-M'}, & K' \leq i \leq K - K', \quad M' + 1 \leq j \leq M \end{array} \right. \quad (3.28)$$

Ceci permet de définir formellement la fonction f de prolongation par continuité de α telle que

$$\alpha = f(\mathring{\alpha}), \quad \alpha \text{ définie par les équations (3.28).}$$

Ainsi, la fonctionnelle coût sismique dont le minimum est recherché est en fait la fonctionnelle

$$\mathcal{H}(\mathring{\alpha}, h_1, \dots, h_S) = \mathcal{J}(f(\mathring{\alpha}), h_1, \dots, h_S) \quad (3.29)$$

L'utilisation de l'algorithme d'optimisation à la fonctionnelle $\mathcal{H}(\mathring{\alpha}, h_1, \dots, h_S)$ (augmentée des termes de régularisation $\mathcal{R}_1(\alpha)$ et $\mathcal{R}_2(\alpha)$) conduit donc à s'intéresser au calcul de

$$\partial_{\mathring{\alpha}} \mathcal{H}(\mathring{\alpha}, h_1, \dots, h_S),$$

qui est donné à l'intérieur du domaine d'intérêt \mathcal{O}_d par

$$\partial_{\mathring{\alpha}} \mathcal{H}(\mathring{\alpha}, h_1, \dots, h_S)_{ij} = \partial_\alpha \mathcal{J}(\alpha, h_1, \dots, h_S)_{ij}, \quad i = K' + 1, \dots, K - K' - 1, \quad j = 0, \dots, M - M' - 1.$$

De plus, sur le bord $\mathring{\Gamma}_d$, on a, en utilisant la notation $\tilde{h} = (h_1, \dots, h_S)$

$$\begin{aligned} \partial_{\mathring{\alpha}} \mathcal{H}(\mathring{\alpha}, \tilde{h})_{K'j} &= \partial_\alpha \mathcal{J}(\alpha, \tilde{h})_{K'j} + \sum_{i=0}^{K'} \partial_\alpha \mathcal{J}(\alpha, \tilde{h})_{ij}, \quad j = 0, \dots, M - M' - 1. \\ \partial_{\mathring{\alpha}} \mathcal{H}(\mathring{\alpha}, \tilde{h})_{K-K',j} &= \partial_\alpha \mathcal{J}(\alpha, \tilde{h})_{K-K',j} + \sum_{i=K-K'+1}^K \partial_\alpha \mathcal{J}(\alpha, \tilde{h})_{ij}, \quad j = 0, \dots, M - M' - 1. \\ \partial_{\mathring{\alpha}} \mathcal{H}(\mathring{\alpha}, \tilde{h})_{i,M-M'} &= \partial_\alpha \mathcal{J}(\alpha, \tilde{h})_{i,M-M'} + \sum_{j=M-M'+1}^M \partial_\alpha \mathcal{J}(\alpha, \tilde{h})_{ij}, \quad i = K' + 1, \dots, K - K' - 1 \end{aligned}$$

Et finalement pour les coins,

$$\begin{aligned} \partial_{\mathring{\alpha}} \mathcal{H}(\mathring{\alpha}, \tilde{h})_{K',M-M'} &= \partial_\alpha \mathcal{J}(\alpha, \tilde{h})_{K',M-M'} + \sum_{i=1}^{K'} \sum_{j=M-M'+1}^M \partial_\alpha \mathcal{J}(\alpha, \tilde{h})_{ij}, \\ \partial_{\mathring{\alpha}} \mathcal{H}(\mathring{\alpha}, \tilde{h})_{K-K',M-M'} &= \partial_\alpha \mathcal{J}(\alpha, \tilde{h})_{K-K',M-M'} + \sum_{i=K-K'+1}^{K'} \sum_{j=M-M'+1}^M \partial_\alpha \mathcal{J}(\alpha, \tilde{h})_{ij}, \end{aligned}$$

Ainsi dans la pratique, la minimisation s'effectue sur la fonctionnelle $\mathcal{H}(\alpha, \tilde{h})$ par rapport aux variables d'impédance contenues dans le domaine d'intérêt \mathcal{O}_d . Cependant, la prolongation par continuité du modèle d'impédance dans les couches PML impose de connaître les valeurs de $\partial_\alpha \mathcal{J}(\alpha, h_1, \dots, h_S)$ dans les couches PML pour déterminer $\partial_\alpha \mathcal{H}(\alpha, \tilde{h})$ sur le bord du domaine d'intérêt $\mathring{\Gamma}_d$. Ainsi, la nécessité de pouvoir calculer ces valeurs apparaît. Ceci est à l'origine de la difficulté particulière qui intervient dans l'algorithme de calcul de $\partial_\alpha \mathcal{J}(\alpha, h_1, \dots, h_S)$ lorsque les ressources de calcul ne permettent pas de stocker en mémoire l'intégralité de l'état direct, pour toutes les itérations en temps.

Annexe 3.4 Algorithme de calcul de $\partial_\alpha \mathcal{J}(\alpha, h)$ dans la PML

L'algorithme de rétropropagation classique présenté dans l'annexe 3.2 n'est pas adapté à l'emploi de couches PML. Néanmoins, il est possible de l'utiliser dans le domaine d'intérêt \mathcal{O}_d , dans lequel l'absorption est nul. Ceci consiste à considérer le problème direct posé sur \mathcal{O}_d seulement, dans la deuxième phase de l'algorithme. Cependant, pour pouvoir résoudre le problème direct dans ce domaine, il est nécessaire de posséder l'information sur la pression sur le bord $\mathring{\Gamma}_d$ notée

$$U_{\mathring{\Gamma}_d}^n = \left(U_{\mathring{\Gamma}_d,1}^n, U_{\mathring{\Gamma}_d,2}^n, U_{\mathring{\Gamma}_d,3}^n \right),$$

avec

$$\begin{cases} U_{\mathring{\Gamma}_d,1}^n = (U_3 + U_4)_{K'j}^n, & U_{\mathring{\Gamma}_d,2}^n = (U_3 + U_4)_{K-K',j}^n, & 0 \leq j \leq M - M' \\ U_{\mathring{\Gamma}_d,3}^n = (U_3 + U_4)_{i,M-M'}^n, & K' \leq i \leq K - K'. \end{cases}$$

La condition au sommet du modèle en $j = 0$ est connue

$$(U_3 + U_4)_{i0}^n = h_i^n, \quad K' \leq i \leq K - K'.$$

En revanche, les conditions de pression sur les frontières internes sont *a priori* inconnues. Ainsi, durant la première phase de l'algorithme classique, qui consiste en le calcul de l'état direct à l'instant final U^N , les conditions de pression au bord $U_{\mathring{\Gamma}_d}^n$ sont enregistrées, pour $0 \leq n \leq N$. Dans une seconde phase, l'état direct et l'état adjoint sont rétropropagés conjointement, l'état direct étant solution du problème direct (3.20) dans le domaine \mathcal{O}_d avec les conditions au bord

$$\begin{cases} (U_3 + U_4)_{K'j}^n = U_{\mathring{\Gamma}_d,1}^n, & (U_3 + U_4)_{K-K',j}^n = U_{\mathring{\Gamma}_d,2}^n, & 0 \leq j \leq M - M' \\ (U_3 + U_4)_{i,M-M'}^n = U_{\mathring{\Gamma}_d,3}^n, & (U_3 + U_4)_{i0}^n = h_i^n, & K' \leq i \leq K - K'. \end{cases} \quad (3.30)$$

Bien que, pendant la phase de rétropropagation, le calcul de l'état direct soit limité à \mathcal{O}_d , il est à noter que l'état adjoint est calculé dans le domaine entier Ω_d . En effet, la rétropropagation ne présente pas d'instabilité pour l'état adjoint. De plus, il est couplé à la résolution du problème direct par les valeurs de l'état direct prises au niveau des récepteurs, qui sont dans le domaine d'intérêt \mathcal{O}_d , et donc accessibles au calcul par rétropropagation. Une troisième phase doit être ensuite ajoutée à l'algorithme initial, qui consiste en le calcul de l'état direct dans la zone PML ($\Omega_d \setminus \mathcal{O}_d$) au temps n , noté \widehat{U}^n .

Soit \mathring{U}^n la solution du problème direct (3.20) avec les conditions aux limites (3.30) au temps n , restreinte à \mathcal{O}_d . Soit $\mathring{\phi}_n$ la fonction de propagation de n pas de temps, $0 \leq n \leq N$, dans \mathcal{O}_d , telle que

$$\forall k \in \mathbb{N}, \quad 0 \leq k \leq N + 1, \quad \mathring{U}^k = \mathring{\phi}_n(\mathring{U}^{k-n}, \alpha, h).$$

La structure de l'algorithme de rétropropagation adapté à l'utilisation de couches PML est

```

Data :  $\alpha, h$ 
Result :  $\nabla \mathcal{J}(\alpha, h)$ 
 $U^0 = 0;$ 
 $U^N = \phi_N(U^0, \alpha, h)$ , enregistrement de  $U_{\mathring{\Gamma}_d}^n$ ,  $0 \leq n \leq N$  ;
extraire  $\mathring{U}^N$  de  $U^N$ ;
 $\lambda^{N+1} = 0$ ;
 $\partial_\alpha \mathcal{J}(\alpha, h) = 0$ ;
for  $n = N, 0$  do
     $\mathring{U}^n = \mathring{\phi}_{-1}(\mathring{U}^{n+1}, \alpha, h);$ 
     $\lambda^n = \psi_{-1}(\lambda^{n+1}, U^n, \alpha, h);$ 
     $\partial_h \mathcal{J}(\alpha, h)_i^n = \alpha_{i,1/2} c_{i,1/2} (\lambda_2)_i^n (1 - \delta_{i,0}) (1 - \delta_{n,0});$ 
    calculer  $\widehat{U}^n$ ;
     $U^n = (\mathring{U}^n, \widehat{U}^n);$ 
    calculer  $I^n(U^n, \lambda^n)$  par (3.25);
     $\partial_\alpha \mathcal{J}(\alpha, h) = \partial_\alpha \mathcal{J}(\alpha, h) + I^n$ ;
end

```

Algorithme 9 : Algorithme de rétropropagation adapté aux couches PML

La question qui reste à régler est maintenant celle du calcul de l'état direct dans la couche PML \widehat{U}^n . Celui-ci peut être calculé par la résolution du problème direct (3.20) avec les conditions aux limites

$$\left\{ \begin{array}{l} (U_3 + U_4)_{i0}^n = h_i^n, \quad 0 \leq i \leq K' - 1 \text{ et } K - K' \leq i \leq K, \\ (U_3 + U_4)_{iM'}^n = 0, \quad 0 \leq i \leq K + 1 \\ (U_3 + U_4)_{i,M-M'}^n = U_{\Gamma_d,3}^n, \quad K' \leq i \leq K - K' \\ (U_3 + U_4)_{0j}^n = 0, \quad (U_3 + U_4)_{K+1,j}^n = 0, \quad 0 \leq i \leq M + 1 \\ (U_3 + U_4)_{K',j}^n = U_{\Gamma_d,1}^n, \quad (U_3 + U_4)_{K-K',j}^n = U_{\Gamma_d,2}^n, \quad 0 \leq j \leq M - M' \\ \end{array} \right. \quad (3.31)$$

Soit ϕ_n^{PML} la fonction de propagation de n pas de temps, $0 \leq n \leq N$, dans $\Omega_d \setminus \mathcal{O}_d$, telle que

$$\forall k \in \mathbb{N}, \quad 0 \leq k \leq N + 1, \quad \widehat{U}^k = \widehat{\phi}_n(\widehat{U}^{k-n}, \alpha, h).$$

Une première solution, naïve, pour le calcul de \widehat{U}^n au sein de la boucle de rétropropagation, consiste en la propagation de l'état direct dans la PML depuis l'état initial à chaque itération de la boucle de rétropropagation. Cependant cette solution est très coûteuse en temps de calcul. En effet, par récurrence, pour $n \in \mathbb{N}$, un appel à la fonction $\widehat{\phi}_n$ représente n appels à la fonction $\widehat{\phi}_1$, qui correspond à la résolution du problème direct pour un pas de temps, dans la zone PML. Ainsi, pour cette version de l'algorithme, le nombre d'appels à la fonction $\widehat{\phi}_1$ au cours de la rétropropagation s'élève à

$$\sum_{n=1}^N n = \frac{N(N+1)}{2}.$$

Comme cela a déjà été mentionné, le nombre de pas de temps utilisé est très important, ceci découlant du choix de pas d'espace fins, et du respect de la condition CFL. Même si la résolution du problème direct est limitée à la zone PML, qui comporte beaucoup moins de points d'espace que le domaine entier, cette complexité en $O(N^2)$ rend cet algorithme extrêmement coûteux en temps de calcul. Une version plus efficace doit donc être implémentée.

Le surcoût de calcul occasionné par cette version de l'algorithme de rétropropagation vient de la propagation à partir de l'état initial de l'état direct dans la PML. L'approche inverse consisterait à stocker tous les itérés en temps de l'état direct, dans la PML, ce qui éviterait de les recalculer à chaque itération de la boucle de rétropropagation. Cette solution doit cependant être évitée dans le contexte précis pour lequel l'espace mémoire disponible est limité.

L'idée retenue repose sur la définition d'une solution intermédiaire entre ces deux solutions extrêmes. A chaque propagation de l'état direct dans la PML, celui-ci est stocké pour un pas de temps situé à mi-chemin entre l'état initial considéré et l'état final. A l'itération suivante, au lieu de propager l'état direct dans la PML depuis l'état initial, celui-ci est propagé depuis l'état stocké à l'itération précédente. Ceci est ensuite répété jusqu'à ce que le pas de temps auquel l'état direct doit être propagé coïncide avec le dernier état stocké. A cette étape, aucun calcul n'est nécessaire, et à l'étape suivante, on repart de l'état stocké antérieurement. Ceci est traduit par l'algorithme suivant

Data : α, h
Result : $\partial_\alpha \mathcal{J}(\alpha, h)$

$U^0 = 0;$
 $U^N = \phi_N(U^0, \alpha, h)$, enregistrement de $U_{\Gamma_d}^n$, $0 \leq n \leq N$;
extraire \dot{U}^N de U^N ;
 $\lambda^{N+1} = 0;$
 $\partial_\alpha \mathcal{J}(\alpha, h) = 0$;
 $q = 1$;
 $p(q) = 0$;
extraire \widehat{U}^0 de U^0 ;
 $U(q) = \widehat{U}^0$;

for $n = N, 0$ **do**

$\dot{U}^n = \dot{\phi}_{-1}(\dot{U}^{n+1}, \alpha, h);$
 $\lambda^n = \psi_{-1}(\lambda^{n+1}, \dot{U}^n, \alpha, h);$

if $n - p(q) > 3$ **then**

$\widehat{U}^n = \widehat{\phi}_{n-p(q)}(U(q), \alpha, h);$
 $p(q+1) = E\left(\frac{n+p(q)}{2}\right);$
 $U(q+1) = \widehat{U}^{p(q+1)}$;
 $q = q + 1$;

else

if $1 \leq n - p(q) \leq 3$ **then**
 $\widehat{U}^n = \widehat{\phi}_{n-p(q)}(U(q), \alpha, h)$

end

if $n - p(q) = 0$ **then**
 $\widehat{U}^n = U(q)$;
 $q = q - 1$;

end

end

$U^n = (\dot{U}^n, \widehat{U}^n);$
calculer $I^n(U^n, \lambda^n)$ par (3.25);
 $\partial_\alpha \mathcal{J}(\alpha, h) = \partial_\alpha \mathcal{J}(\alpha, h) + I^n$;

end

Algorithme 10 : Algorithme de rétropropagation adapté aux couches PML

Evaluons la complexité de cet algorithme. Soit $f(N)$ le nombre d'itérations nécessaires pour obtenir l'état direct dans la PML dans une boucle rétrograde en temps. Pour ce nouvel algorithme, il est facile d'établir la relation de récurrence suivante

$$f(N) = N + 2f\left(\frac{N-1}{2}\right), \quad f(0) = 0.$$

En effet, N opérations mènent de l'état initial à l'état final, et l'état intermédiaire

à $N/2$ est stocké en mémoire. Pour calculer le reste des états en temps, il reste à appliquer la même méthode sur la portion allant de $N/2$ à $N - 1$, et sur celle allant de 0 à $N/2 - 1$. Pour chacune de ces portions, le nombre d'opération à effectuer est égal à $f\left(\frac{N-1}{2}\right)$.

A partir de cette relation de récurrence, il est facile de donner une borne sur la complexité de l'algorithme. La fonction f est par définition croissante. Plus le nombre de points en temps est grand, plus le nombre d'itérations à réaliser est important. Ainsi,

$$f(N) \leq N + 2f\left(\frac{N}{2}\right).$$

De plus, il existe k entier tel que

$$2^{k-1} \leq N \leq 2^k.$$

Alors,

$$f(N) \leq f(2^k) = 2^k + 2f(2^{k-1})$$

Comme $f(0) = 0$,

$$f(2^k) = k2^k$$

D'autre part,

$$k - 1 \leq \log_2(N) \leq k, \quad \text{et} \quad 2^k \leq 2N,$$

ce qui permet d'établir

$$f(N) \leq 2N(\log_2(N) + 1)$$

Ainsi, la complexité de l'algorithme modifié est en $O(N \log_2(N))$, ce qui est bien meilleur que la complexité initiale en $O(N^2)$. De plus, le nombre de matrices à stocker en mémoire est de l'ordre de $\log_2(N)$. C'est donc un compromis idéal entre temps de calcul et espace mémoire, dans le cadre d'une implémentation pour laquelle le stockage mémoire doit être limité.

Chapitre 4

Evaluation de la méthode d'inversion

Introduction

La description du problème inverse non linéaire correspondant à l'inversion 2D de données sismiques de puits et celle des moyens employés pour sa résolution ont été réalisées dans les deux chapitres précédents. Il convient maintenant d'évaluer le potentiel de cette méthode d'imagerie sismique. Ce chapitre a donc pour objectif de fournir un éclairage sur les avantages et les limitations de la méthode développée tant sur le plan purement numérique (convergence de la méthode) que sur le plan de l'application géophysique (amélioration de la résolution, sensibilité).

Dans une première partie, le modèle test qui sert à l'essentiel des expériences numériques est décrit. Le premier aspect qui est ensuite évalué, au sein de la deuxième partie, est la scalabilité du code. Cette partie propose une étude du comportement des algorithmes parallèles de rétropropagation, servant à évaluer le gradient, et du comportement global de la méthode d'inversion, en fonction du nombre de processeurs employés pour la résolution du problème test.

Les résultats qui sont présentés dans les parties suivantes proposent quant à eux une évaluation de l'intérêt que représente la méthode d'inversion 2D de données sismiques de puits en vue de l'amélioration de la résolution des méthodes d'imagerie sismique actuelles. L'objet est en effet de proposer une méthode d'inversion devant permettre, par la prise en compte du phénomène non linéaire que sont les réflexions multiples, de fournir une image du sous-sol à plus fine échelle que les méthodes conventionnelles. Plus précisément, nous nous attachons à la reconstruction d'un modèle d'impédance acoustique, connaissant la distribution de vitesse de propagation, dans un milieu 2D à faibles variations latérales, à partir d'enregistrements de données sismiques de puits.

Nous commençons par analyser la résolution verticale de l'image obtenue. A titre de comparaison, les résultats obtenus sont confrontés à un résultat issu d'une méthode d'inversion linéaire classique, afin de mieux appréhender le gain obtenu.

Nous analysons ensuite la résolution latérale de l'image. En effet, de par la configuration de l'acquisition, l'information sur les variations latérales du milieu est moins importante, et pour lutter contre l'indétermination, des termes de régularisation doivent être introduits. Il est donc important de s'interroger sur la pertinence de la reconstruction des variations latérales du milieu, et en particulier, d'évaluer le niveau de détail qui peut être atteint.

Ces deux questions ont trait à la résolution de la méthode d'inversion, et sont d'importance primordiale. Elles sont toutes deux traitées d'abord dans un cas dit "idéal" pour lequel les CP associés à chaque jeu de données sont connues. Cette expérience permet de donner une estimation de la meilleure résolution pouvant être obtenue par la méthode. En effet, si les CP sont connues, seul le modèle d'impédance doit être déterminé, et les problèmes d'indétermination de couplage sont levés. Dans un second temps, ces mêmes questions sont traitées dans le cadre concret d'utilisation de la méthode, c'est-à-dire lorsqu'à la fois les CP et le modèle d'impédance sont inconnus. Il est alors intéressant d'évaluer la dégradation du résultat en terme de résolution par rapport à la situation idéale précédente.

Enfin nous abordons la question de la sensibilité de la procédure d'inversion à des erreurs introduites sur les données. Deux types d'erreur sont considérés. Tout d'abord, la méthode d'inversion requérant l'établissement d'un modèle de vitesse *a priori*, il est nécessaire de se poser la question de la sensibilité de la méthode à une erreur sur ce modèle de vitesse. Les expériences précédentes étant réalisées à modèle de vitesse exact, ce dernier paragraphe montre l'évolution de la solution du cas test suivant l'introduction de diverses erreurs sur le modèle de vitesse. Dans un second temps, un test est effectué à modèle de vitesse exact mais pour des données bruitées. Dans la pratique, une partie du signal contenu dans les données enregistrées ne peut en effet être expliquée par l'opérateur de modélisation. Cette partie du signal est considérée comme un bruit. L'évaluation de la sensibilité de la méthode à ce type d'erreur est important pour des développements futurs et l'application à des cas réels.

4.1 Description du cas test

4.1.1 Modèle exact et acquisition

Le modèle de sous-sol servant de base à l'étude numérique est un modèle synthétique inspiré de relevés de terrain effectués en Algérie. Les nombreuses hétérogénéités verticales qu'il présente, à l'échelle de 8 m, sont à l'origine de très importantes réflexions multiples, dont la présence rend l'utilisation des méthodes d'inversion linéaires inappropriée [Delprat-Jannaud, Lailly, 2008]. Les variations latérales sont quant à elles assez faibles, conformément au contexte d'application qui a été défini dans le chapitre 2. L'utilisation d'un tel modèle permet donc d'illustrer l'intérêt de la mise en place de la procédure d'inversion non linéaire qui a été développée. Les profils d'impédance et de vitesse sont donnés en figure 4.1.

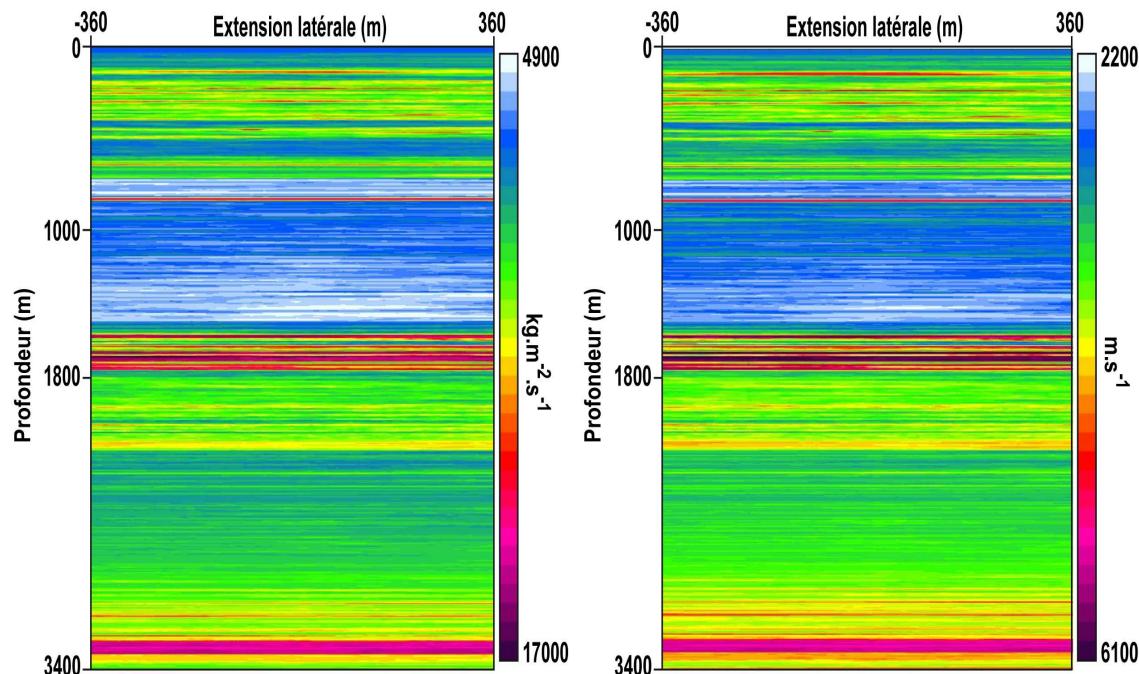


FIG. 4.1 – Modèles exacts d'impédance (gauche), et de vitesse de propagation (droite).

Ces modèles sont définis sur un domaine de 3400 m de profondeur et 720 m de largeur. Le puits dans lequel sont placés les récepteurs se situe à l'abscisse $x = 0$ m, au milieu du domaine. Ce puits possède une profondeur de 1800 m. Les récepteurs, au nombre de 100, sont disposés tous les 8 m à partir de 1000 m de profondeur.

Un ensemble de 30 sources, réparties symétriquement par rapport à la tête du puits, en surface, tous les 24 mètres, génèrent les données. Le signal émis par ces sources est un signal centré sur la fréquence $f = 25$ Hz, appelé signal de Ricker.

Ce signal, couramment utilisé en géophysique, est représenté figure 4.2, et noté $r(t)$. Son expression est

$$r(t) = C \left(t - \frac{1}{f} \right) \exp \left(-\Pi^2 f^2 \left(t - \frac{1}{f} \right)^2 \right),$$

où C est une constante multiplicative.

Le signal émis par ces sources se propage dans le sous-sol, et est enregistré au niveau des récepteurs pendant le temps d'enregistrement T , fixé à 2 s. Le domaine au sein duquel une estimation du modèle d'impédance est recherchée est le rectangle de 2400 m de profondeur et de 720 m de largeur, qui s'étend de la profondeur de 1000 m, correspondant à celle du premier récepteur, jusqu'à la profondeur de 3400 m, et couvrant toute l'extension latérale. Le temps d'enregistrement est suffisamment grand pour que ce domaine soit entièrement éclairé. Par là, il est entendu que les données sismiques générées sont sensibles en chaque valeur d'impédance de ce domaine.

Conformément à ce qui est expliqué dans le chapitre 2, afin de pouvoir introduire un terme de régularisation sur les CP, ces données sont recombinées, afin d'approcher une situation d'éclairage par ondes planes dont la direction de propagation forme différents angles avec la verticale du milieu. Ainsi, 13 jeux de données par angles sont constitués, de -6° à $+6^\circ$, avec des variations de 1° par jeu de données. Le choix d'un angle maximal de 6° est conforme à la géométrie de l'acquisition (offset maximal des sources et profondeur des récepteurs). De plus, il permet qu'un réflecteur horizontal à la profondeur de 2500 m soit éclairé de 150 m de part et d'autre du puits, suivant l'approximation de l'optique géométrique. Les jeux de données recombinées correspondant à la propagation d'une pseudo-onde plane se propageant avec un angle de 0° et $+6^\circ$ sont représentés figure 4.3.

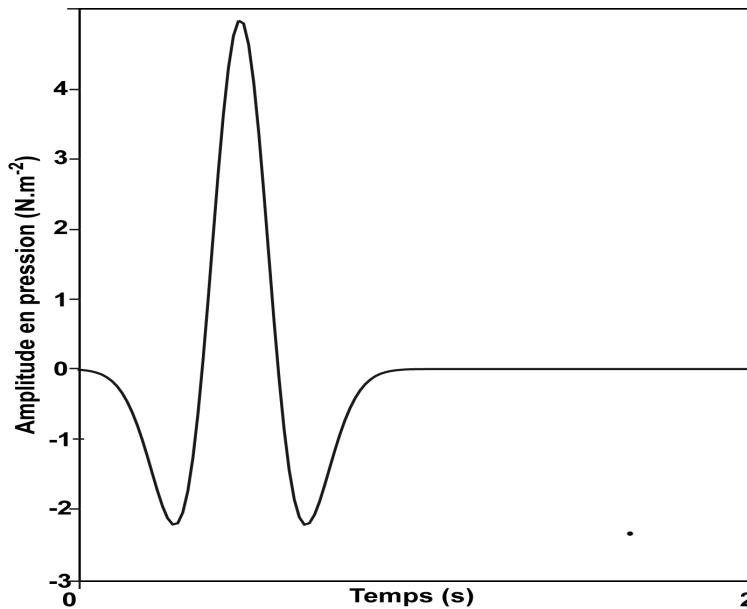


FIG. 4.2 – Signal de Ricker.

Cette représentation, comme dans le chapitre 1, utilise un axe vertical en temps et un axe horizontal correspondant à la profondeur des récepteurs. Les amplitudes positives sont soulignées en blanc, et les amplitudes négatives en noir. Le dégradé de gris correspond aux valeurs intermédiaires. La première arrivée correspond à la propagation de l'onde directe, visible sur les données. Celle-ci correspond au signal d'intensité maximale qui se propage du premier récepteur au récepteur le plus profond. En dessous de cette première arrivée, un signal complexe est enregistré, qui correspond aux réflexions multiples engendrées par les nombreuses hétérogénéités d'impédance du modèle exact (fig 4.1). Parmi ces réflexions peuvent notamment se distinguer des ondes montantes, qui se propagent du dernier récepteur au récepteur le moins profond.

Le modèle physique servant de cas test et le mode d'acquisition des données ayant été présenté, il convient maintenant d'apporter quelques précisions sur la manière dont ce modèle est discréétisé pour simuler la propagation des ondes acoustiques en son sein.

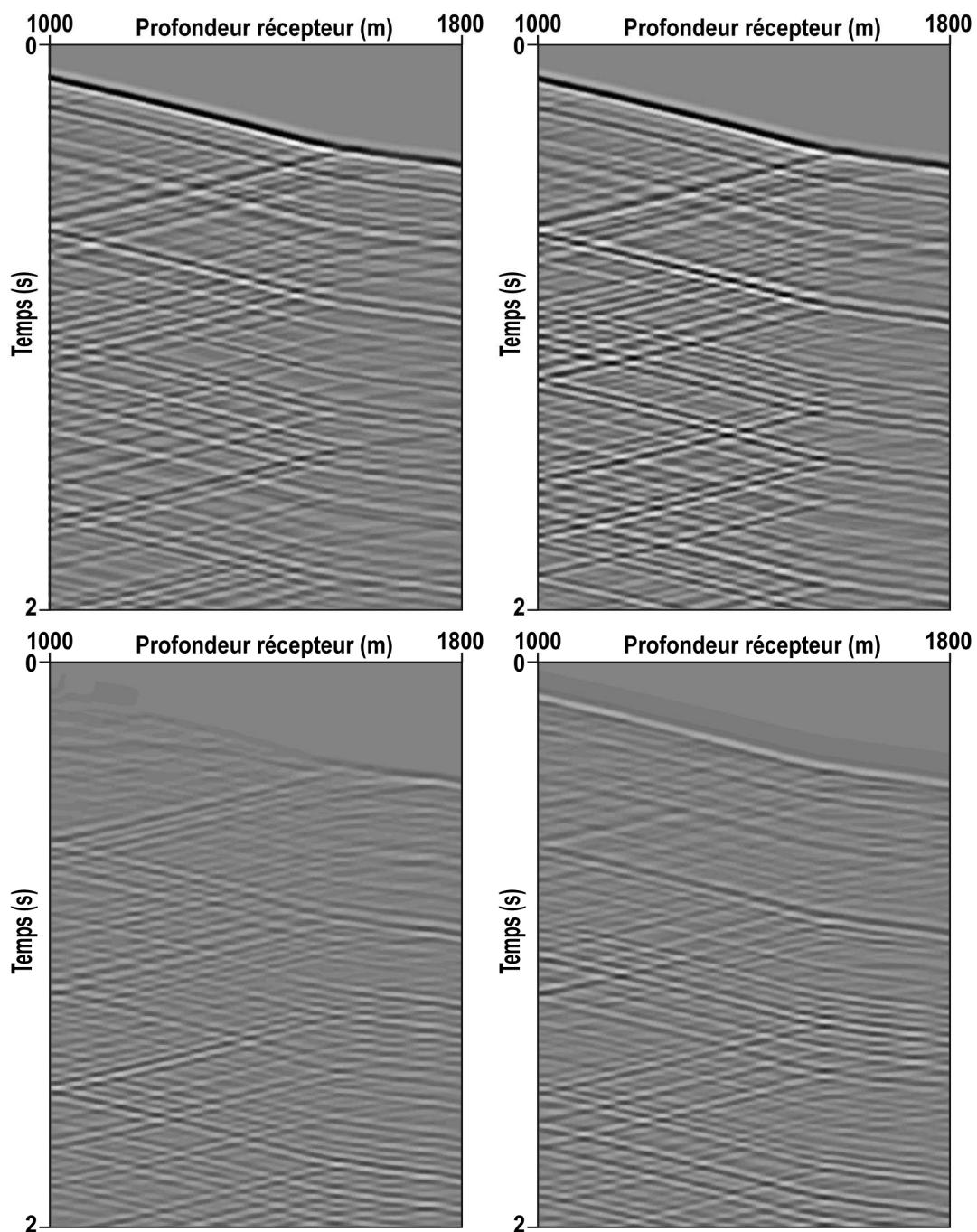


FIG. 4.3 – Données recombinées par angle. Vitesse verticale 0° (haut gauche), vitesse verticale $+6^\circ$ (haut droite), vitesse horizontale 0° (bas gauche), vitesse horizontale $+6^\circ$ (bas droite).

4.1.2 Discréétisation

La discréétisation spatiale utilisée doit répondre à plusieurs impératifs. Tout d'abord, elle doit être suffisamment fine pour promouvoir un gain dans la résolution

du résultat. Les pas de discréétisation doivent ainsi être choisis plus petits que la résolution moyenne des méthodes d'inversion linéaires, qui est de l'ordre de plusieurs dizaines de mètres. De plus, ils doivent permettre une bonne représentation des ondes dans le domaine de propagation. Pour cela, une règle empirique consiste à utiliser pour le moins 10 points de discréétisation par longueur d'onde moyenne. Celle-ci, notée λ_{moy} , est évaluée par

$$\lambda_{moy} = \frac{c_{moy}}{f} = \frac{3531}{25} \simeq 140 \text{ m},$$

où c_{moy} représente la moyenne de la vitesse de propagation en norme L^2 .

La longueur d'onde moyenne des ondes se propageant dans le sous-sol est donc de l'ordre de 140 m, ce qui conduit à choisir des pas de discréétisation inférieurs à 14 m. Les valeurs choisies au final sont $\Delta x = 12 \text{ m}$ et $\Delta z = 8 \text{ m}$. Le pas de discréétisation vertical est choisi plus petit que le pas de discréétisation horizontal, en raison du gain attendu en résolution verticale.

Le choix du pas de discréétisation temporel Δt est effectué conformément à la condition CFL du schéma de Yee centré d'ordre 2 [Yee, 1966]. Cette condition, qui assure la stabilité du schéma numérique, s'écrit

$$c_{max} \Delta t \leq \frac{1}{\sqrt{1/\Delta x^2 + 1/\Delta z^2}},$$

où c_{max} est la valeur maximale de vitesse.

Cette vitesse maximale c_{max} avoisinant les 6000 m.s^{-1} , le pas de discréétisation temporel Δt doit donc respecter la condition

$$\Delta t \leq 1.1 \times 10^{-3} \text{ s}.$$

Cependant cette condition est valable dans le contexte de milieux de propagation homogènes, ce qui n'est clairement pas le cas ici. Elle doit donc être légèrement abaissée, et les expériences numériques ont montré que le choix

$$\Delta t = 0.9 \times 10^{-3} \text{ s},$$

convenait.

Pour fixer les idées sur la taille du problème résolu, il est bon de donner le nombre d'inconnues et de données discrètes. La discréétisation mène à la définition de 13 jeux de données composés chacun de 100 signaux (un par récepteur) comportant 2000 points, ce qui donne au total 2.6×10^6 valeurs discrètes. Le modèle d'impédance est lui représenté par une matrice de 101×321 points, soit 32421 inconnues, tandis que chaque CP est représentée par 15 traces de 2000 points. Le nombre

total d'inconnues s'élève donc à $4,22421 \times 10^5$.

La discrétisation étant définie, il reste à déterminer l'épaisseur de la couche PML utilisée pour simuler la propagation des ondes dans un demi-espace. La détermination de cette épaisseur, et celle des coefficients d'absorption utilisés dans ces couches, est donnée au chapitre 5, qui traite de manière plus exhaustive du problème de la couche PML, en proposant une analyse mathématique de stabilité, et des expériences numériques illustrant leur comportement absorbant. On se contentera ici de donner la taille finalement retenue qui est de 20 points de discrétisation, la précision obtenue pour une telle taille se révélant suffisante pour l'application envisagée.

Le cas test étant défini, il est maintenant possible d'aborder les expériences numériques qui ont été effectuées, en commençant par les tests de scalabilité du code.

4.2 Scalabilité de la parallélisation par décomposition de domaine

4.2.1 Description du cluster de l'IFP

Les ressources informatiques de l'IFP ont été mises à disposition pour le développement du code et la réalisation des expériences numériques. Il convient d'apporter quelques précisions sur la description de ces ressources. L'IFP a fait l'acquisition depuis décembre 2005 d'un cluster de calcul dédié au calcul parallèle. Ce cluster comporte 114 noeuds de calcul Supermicro 1041M, chaque noeud comportant 4 processeurs quadri coeurs AMD Barcelona, cadencés à 2.3 Ghz, et possédant 1 Mo de cache L2, 32 Go de mémoire RAM (de type DDR2), 160 Go d'espace disque, et une carte Infiniband Mellanox ConnectX double port 20 Gbs. Chaque noeud comporte donc 16 coeurs de calcul, et le cluster dispose au total de 1824 coeurs. Cet équipement représente un outil de calcul performant et de construction récente, qui permet le développement efficace de codes parallèles. Le cluster a même fait l'objet d'une apparition dans le classement des 500 supercalculateurs les plus puissants de par le monde, à la 498^e place, dans le classement bi-annuel de juin 2008 [<http://www.top500.org>]. La connection Infiniband, notamment, permet une communication très rapide entre les coeurs de calcul. Cependant, les temps de communication diffèrent suivant l'éloignement des coeurs. Deux coeurs associés au même processeur communiquent extrêmement rapidement. Ce temps de communication est légèrement plus long si ces deux coeurs sont sur des processeurs différents mais sur le même noeud. Enfin, les temps de communication entre deux coeurs situés sur deux noeuds différents sont sensiblement plus longs. Il est donc important de veiller à concentrer les coeurs de calcul utilisés sur les mêmes noeuds. Les utilisateurs du cluster pouvant disposer d'un maximum de 128 coeurs simultanément, la scalabilité du code de calcul a été testée pour une décomposition allant jusqu'à 128 sous-domaines, en groupant les coeurs de calcul utilisés sur 8 noeuds de calcul, afin d'optimiser les temps de communication.

4.2.2 Tests de performance

Lorsqu'une méthode de parallélisation est utilisée pour résoudre un problème numérique, il est normal de s'interroger sur l'efficacité de cette parallélisation, en d'autres termes son rendement. Le rendement théorique optimal est atteint lorsqu'en utilisant n processeurs, le nombre d'opérations à effectuer par processeurs est exactement divisé par n . Le temps d'exécution correspondant est alors lui aussi divisé par n .

Dans le cas présenté ici, on s'intéresse d'abord à la scalabilité de l'algorithme parallèle 4 (chapitre 3, section 6) qui correspond à l'algorithme de rétropropagation

pour le calcul des dérivées partielles

$$\partial_{h_s} \mathcal{J}(\alpha, h_1, \dots, h_S).$$

Ces dérivées partielles servent à la résolution du problème de gradient conjugué (chapitre 3, section 2). De plus la parallélisation du calcul des dérivées partielles

$$\partial_\alpha \mathcal{J}(\alpha, h_1, \dots, h_S),$$

suit le même principe, puisque le choix a été fait de stocker l'intégralité des valeurs de l'état direct $U(x, z, t)$ en temps dans la couche PML (voir la section 5 du chapitre 3). L'algorithme 4 est donc représentatif de la parallélisation employée. L'idée directrice est de s'intéresser d'abord à la scalabilité de la partie parallèle du code. De celle-ci dépend la scalabilité du code d'inversion global, qui est étudiée par la suite.

Le temps de calcul requis par l'algorithme 4 est calculé pour un nombre de processeurs doublant à chaque mesure, de 1 à 128 processeurs. Les résultats obtenus sont présentés dans le tableau suivant.

Nombre de coeurs	1	2	4	8	16	32	64	128
Temps de calcul (s)	6.74	2.92	1.53	0.38	0.18	0.14	0.12	0.11

FIG. 4.4 – Temps de calcul suivant le nombre de processeurs pour l'algorithme 4.

Il est déjà possible de constater que si, pour les premières valeurs, le temps de calcul est approximativement divisé par deux à mesure que le nombre de processeurs double, à partir de 16 processeurs, cette tendance ne se vérifie plus, et le gain en temps de calcul est moins important. Pour mieux appréhender celui-ci, il convient de le comparer avec le gain optimal théorique. Pour cela, il est utile de représenter ces résultats sur une échelle logarithmique en base 2. Sur cette échelle, le gain optimal théorique correspond à une droite de pente -1 , à laquelle il est aisément de confronter les résultats expérimentaux. La courbe de performance optimale est déduite du temps de calcul nécessaire à un unique cœur pour effectuer le calcul. Cette comparaison est effectuée sur la figure 4.5.

Le profil de la performance optimale théorique est représenté en noir tandis que les mesures effectuées sont représentées en gris. Les premières mesures montrent une convergence linéaire, et même superlinéaire pour l'utilisation de 1 à 16 processeurs. Ce dépassement de la performance optimale peut sembler étrange. Cependant, lorsque la performance optimale est définie, seul le nombre d'opérations est pris en compte, et la problématique du temps d'accès mémoire est escamotée. Celle-ci est pourtant d'importance majeure. Ainsi des données qui dépassent la taille de la mémoire cache des coeurs de calcul pour l'utilisation de 1 ou 2 sous-domaines, peuvent voir leur taille réduire suffisamment à partir de 4 sous-domaines

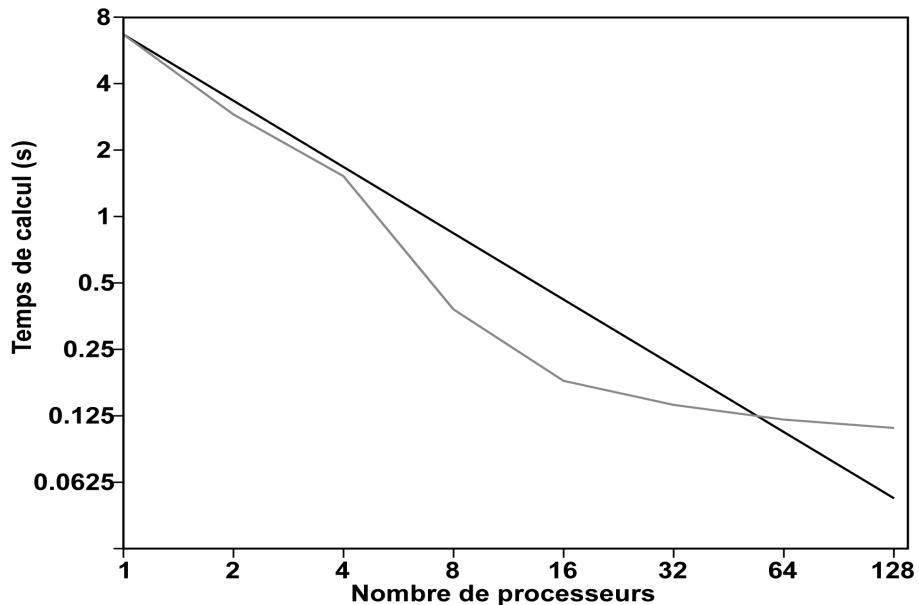


FIG. 4.5 – Performance suivant le nombre de coeurs utilisés pour l'algorithme 4.

pour être stockées dans cette mémoire d'accès plus rapide. La réduction du nombre d'opérations s'accompagne donc d'une accélération de l'accès aux données, ce qui résulte en cette convergence "superlinéaire".

Pour une utilisation plus massive de processeurs ($n > 16$), les performances sont nettement moins bonnes. Dans ce contexte, la réduction de la taille des domaines de calcul contribue toujours à diminuer le nombre d'opérations à effectuer, mais n'améliore pas les temps d'accès à la mémoire, la taille des données étant désormais suffisamment réduite pour que celles-ci puissent être stockées en mémoire cache. De plus, le gain en temps de calcul est compensé par le temps de communication entre processeurs requis par les échanges de données. Dans le cas présenté ici, les processeurs échangeant des données à chaque itération en temps, le nombre de communications à effectuer est rapidement très important.

Il est maintenant nécessaire de s'interroger sur le comportement de la méthode globale d'inversion. En particulier, il convient de vérifier si ce comportement correspond à celui du noyau de calcul parallèle, où s'il s'en éloigne. Ceci correspond en réalité à l'évaluation de la part du temps de calcul passé à réaliser des opérations de type séquentiel par rapport au temps de calcul parallèle. Ces opérations de type séquentiel sont réalisées par le processus maître, et correspondent aux étapes de l'algorithme d'optimisation de type Quasi-Newton (calcul de la direction de descente, calcul du pas par recherche linéaire, mise à jour de l'approximation de l'inverse de la matrice Hessienne). Ainsi, une comparaison de temps de calcul identique à la

précédente est effectuée, mais cette fois pour 10 itérations de la boucle d'optimisation Quasi-Newton (cf algorithme 3, section 2 du chapitre 3). Les résultats sont présentés dans le tableau suivant.

Nombre de processeurs	1	2	4	8	16	32	64	128
Temps de calcul (s)	58 997	26 132	14875	3998	2275	1834	1730	1482

FIG. 4.6 – Temps de calcul suivant le nombre de processeurs pour l'algorithme 3 limité à 10 itérations.

De la même manière, une représentation logarithmique de ces résultats et de la performance optimale théorique est utilisée (fig.4.7).

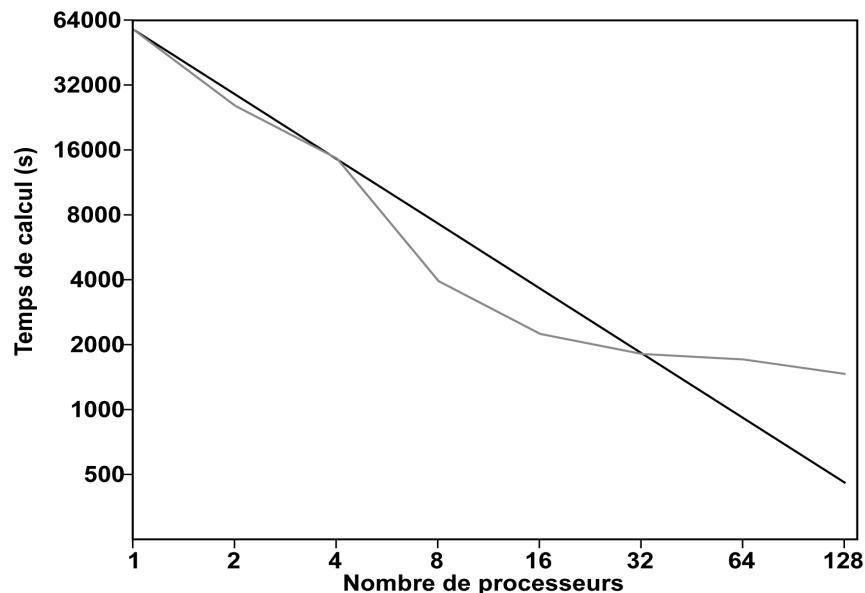


FIG. 4.7 – Performance suivant le nombre de processeurs pour l'algorithme 3 limité à 10 itérations.

Ces résultats montrent que la performance de la méthode globale suit à peu près celle du noyau parallèle. Une convergence superlinéaire est observée pour une décomposition de 1 à 16 coeurs. Pour une utilisation plus massive ($n > 16$) le gain en temps de calcul est à nouveau concurrencé par les temps de communication entre coeurs. La performance globale est cependant dégradée par rapport aux résultats précédents, la courbe de comportement observée dépassant la courbe de comportement optimal à partir de 32 coeurs de calcul, tandis que ce dépassement n'intervient

qu'à partir de 64 coeurs dans le cas précédent.

La méthode de décomposition de domaines utilisée permet donc une réduction conséquente des temps de calcul, même si la scalabilité n'est pas optimale. Pour 10 itérations de l'algorithme Quasi-Newton sur le cas test défini dans la partie précédente, l'utilisation de 128 processeurs permet de réduire le temps de calcul d'un facteur proche de 40. Ceci provient en partie du nombre de communications à effectuer, qui est extrêmement important, même si la taille des données échangées est réduite. En effet, à chaque itération en temps, chaque cœur de calcul doit communiquer avec 1 à 4 voisins, suivant le découpage utilisé. Ceci joue certainement un rôle pénalisant dans la scalabilité de la méthode numérique. De plus, les calculs liés à l'algorithme d'optimisation Quasi-Newton L-BFGS restent intrinsèquement séquentiels, seul le calcul du gradient étant parallélisé. Certaines améliorations peuvent être cependant apportées.

4.2.3 Améliorations possibles

Une amélioration simple consiste à introduire un niveau supérieur de parallélisation, en traitant chaque jeu de données séparément. En effet, il a été montré que la fonctionnelle minimisée se décompose comme une somme de fonctionnelles $\mathcal{J}_s(\alpha, h_s)$ telles que

$$\mathcal{J}_s(\alpha, h_s) = \|\mathcal{D}_s - \widetilde{\mathcal{M}}_d(\alpha, h_s)\|_t^2.$$

De plus le calcul du gradient de la fonctionnelle globale est donné par les équations

$$\begin{cases} \partial_\alpha \mathcal{J}(v) = \sum_{s=1}^S \partial_\alpha \mathcal{J}_s(\alpha, h_s), \\ \partial_{h_s} \mathcal{J}(v) = \partial_{h_s} \mathcal{J}_s(\alpha, h_s), \quad 1 \leq s \leq S. \end{cases} \quad (4.1)$$

Ainsi, il est possible de définir S groupes de processeurs, chaque groupe étant chargé d'évaluer les termes

$$\partial_\alpha J_s(\alpha, h_s), \quad \partial_{h_s} J_s(\alpha, h_s),$$

pour s entier tel que $1 \leq s \leq S$, en utilisant la méthode de décomposition de domaines qui a déjà été introduite.

Ceci permet de répartir le traitement du calcul du gradient à un niveau supérieur, et de fournir un code dont le temps de calcul ne dépend pas du nombre de jeux de données considéré. De plus, l'optimalité de la scalabilité de cette parallélisation supplémentaire est garantie, car les calculs par jeux de données sont totalement indépendants. La seule communication ajoutée consiste en l'envoi des composantes

du gradient calculées par chaque groupe de processeurs au processeur maître. Dans le cas test présenté, qui utilise 13 jeux de données, en gardant une décomposition en 128 sous-domaines, ceci nécessite cependant l'usage de 1664 coeurs de calcul, ce qui implique des moyens de calcul trop conséquents pour que cette méthode ait pu être testée dans le cadre de cette thèse.

D'autres pistes d'amélioration consistent en une gestion plus attentive de la mémoire et un contrôle des temps d'accès. Comme le montre en effet le phénomène de convergence superlinéaire survenant pour les premières décompositions, la problématique du temps d'accès mémoire joue un rôle fondamental pour la scalabilité. L'implémentation de calculs parallèles sur des architectures de type carte graphique (GPU computing), discipline suscitant un engouement prononcé à l'heure actuelle, permet d'effectuer ce contrôle sur la répartition des structures de données en mémoire, de manière à optimiser les temps d'accès. L'interface de programmation MPI n'offre pas cette possibilité. Aussi les codes parallèles actuels couplent différents niveaux de parallélisme, un premier niveau étant effectué à l'aide de MPI, le second par GPU.

Ceci dépasse cependant largement le cadre de cette thèse. L'objectif fixé était de proposer une première parallélisation, profitant des moyens de calcul offerts. L'algorithme de décomposition de domaine a permis de remplir cet objectif, et de fournir une méthode d'inversion non linéaire 2D convergeant en des temps raisonnables, typiquement de l'ordre de 36 heures pour 1000 itérations de l'algorithme Quasi-Newton l-BFGS sur le cas test présenté dans la partie précédente. Il convient maintenant de déterminer les apports de cette méthode sur le plan de la problématique de l'imagerie sismique.

4.3 Analyse de la résolution dans un cas idéal (CP connues)

Cette partie est consacrée à l'évaluation du potentiel de la méthode en termes de résolution verticale et horizontale de l'image calculée. Ce potentiel est évalué dans le cas idéal où la stimulation du milieu par les CP est connue, pour chaque angle d'éclairage. Ainsi, seul le modèle d'impédance optimal est recherché. L'indétermination de couplage entre les inconnues, évoquée dans le chapitre 2, n'apparaît pas. C'est une situation parfaitement fictive, éloignée de toute application concrète, néanmoins, elle permet de donner une idée de la meilleure résolution qu'il est possible d'atteindre, pour le cas test choisi, en utilisant notre méthode d'inversion. Cette expérience numérique est également réalisée sans information *a priori* : les poids de régularisation sont fixés à 0. C'est donc seulement l'information sismique qui est prise en compte. Ceci permet de mettre en évidence l'indétermination inhérente au problème en ce qui concerne l'extension latérale du modèle d'impédance. L'objectif de cette partie est donc de proposer une évaluation de la meilleure résolution qu'il est possible d'obtenir, tout en donnant une idée des composantes bien déterminées du modèle d'impédance lorsque seule l'information sismique est prise en compte.

D'un point de vue numérique, pour cette expérience, les calculs sont allégés : l'algorithme du gradient conjugué n'est pas utilisé au sein de la boucle d'optimisation, les CP optimales étant supposées connues. De manière pratique, celles-ci sont enregistrées lors de la constitution des données synthétiques, ce qui permet de les utiliser dans la suite du code sans avoir à les évaluer.

4.3.1 Présentation du résultat

La boucle d'optimisation prend fin à l'issue de 5000 itérations, pour 10 heures de calcul. Celui-ci est effectué sur 128 processeurs, conformément aux expériences menées sur la scalabilité des algorithmes de rétropropagation parallèles. Les critères d'arrêt de l'algorithme d'optimisation, qui portent sur la réduction de la fonction coût et de la norme du gradient, sont calibrés de manière à obtenir le résultat le plus précis possible (les seuils choisis sont extrêmement faibles, 10^{-25} exactement), aussi le véritable critère d'arrêt est le nombre d'itérations maximal autorisé, ici égal à 5000.

Dans un premier temps, il convient d'analyser les différences entre les données et le sismogramme synthétique correspondant au modèle d'impédance optimal (les résidus, par la suite). La courbe de décroissance de la fonction coût en fonction du temps est présentée figure 4.8.

Normalisée à 1 au commencement des itérations, la fonction coût atteint la valeur de 1×10^{-4} à l'issu de 500 itérations, en 1 heure de calcul. Cette décroissance cor-

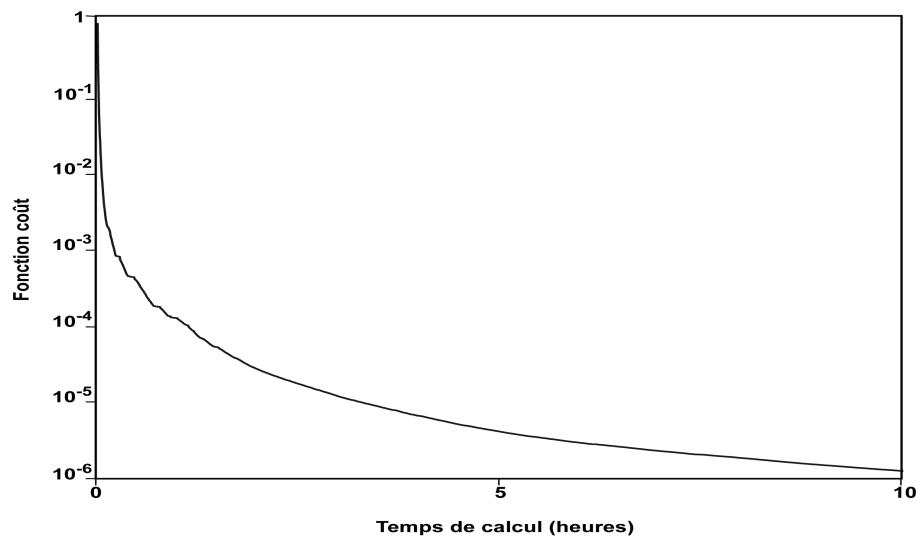


FIG. 4.8 – Décroissance de la fonction coût en fonction du temps (cas 2D, CP exactes).

respond à la division par 100 des résidus. A l'issue de 5000 itérations, elle atteint finalement 1.25×10^{-6} , soit près d'une réduction par 1000 des résidus, ce qui implique que le modèle d'impédance calculé explique très précisément les données. Ceci est confirmé par la représentation des résidus, donnée figure 4.9.

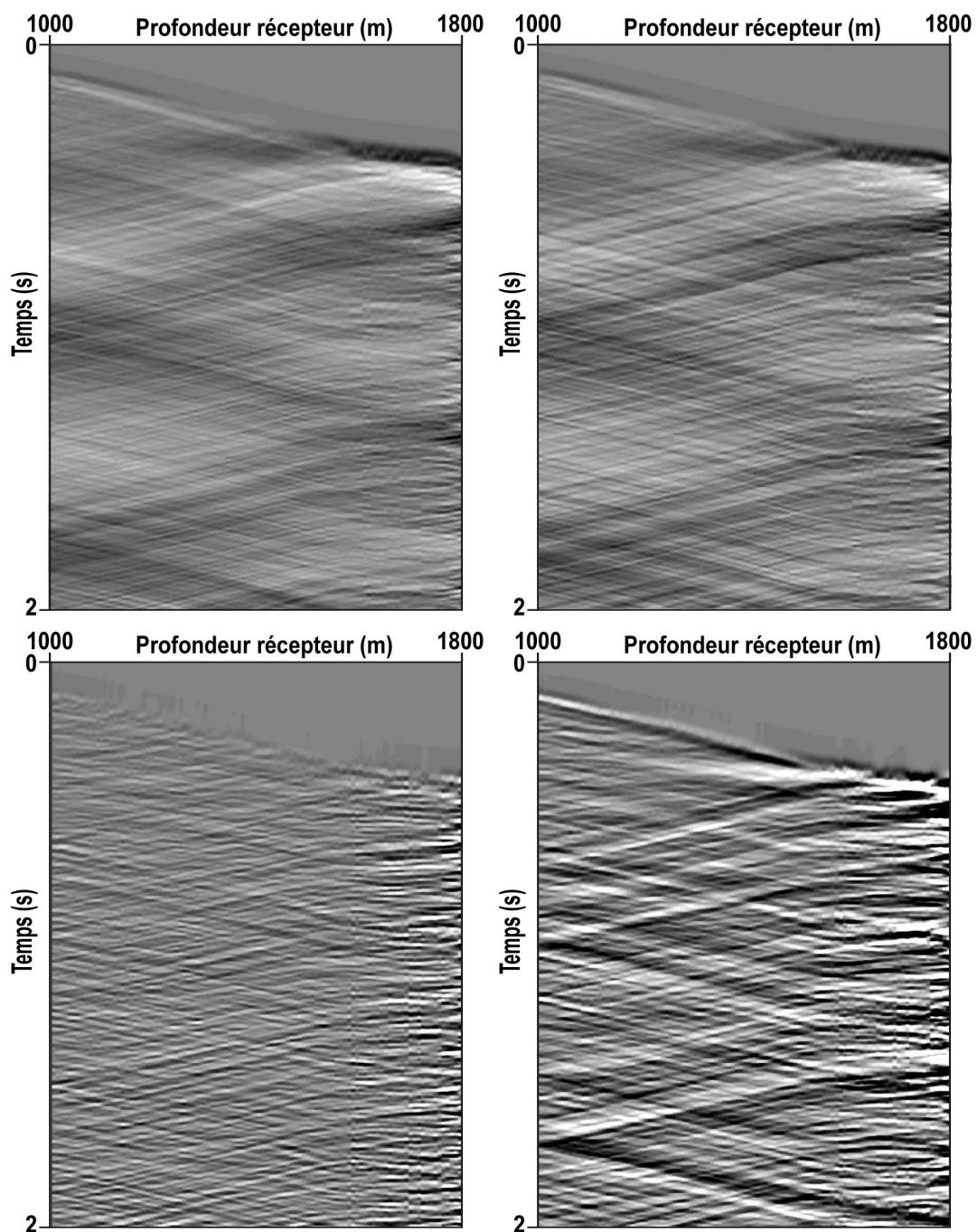


FIG. 4.9 – Résidus vitesse verticale 0° (haut gauche), vitesse verticale 6° (haut droite), vitesse horizontale 0° (bas gauche), vitesse horizontale 6° (bas droite).

Cette figure représente les résidus, avec une amplification d'un facteur 1000 par rapport à l'échelle utilisée pour la représentation des données exactes en figure 4.3.

Les résidus sont donc très faibles, et le modèle d'impédance calculé explique les données avec une grande précision. Ce modèle calculé et sa différence avec le modèle exact sont présentés figure 4.10.

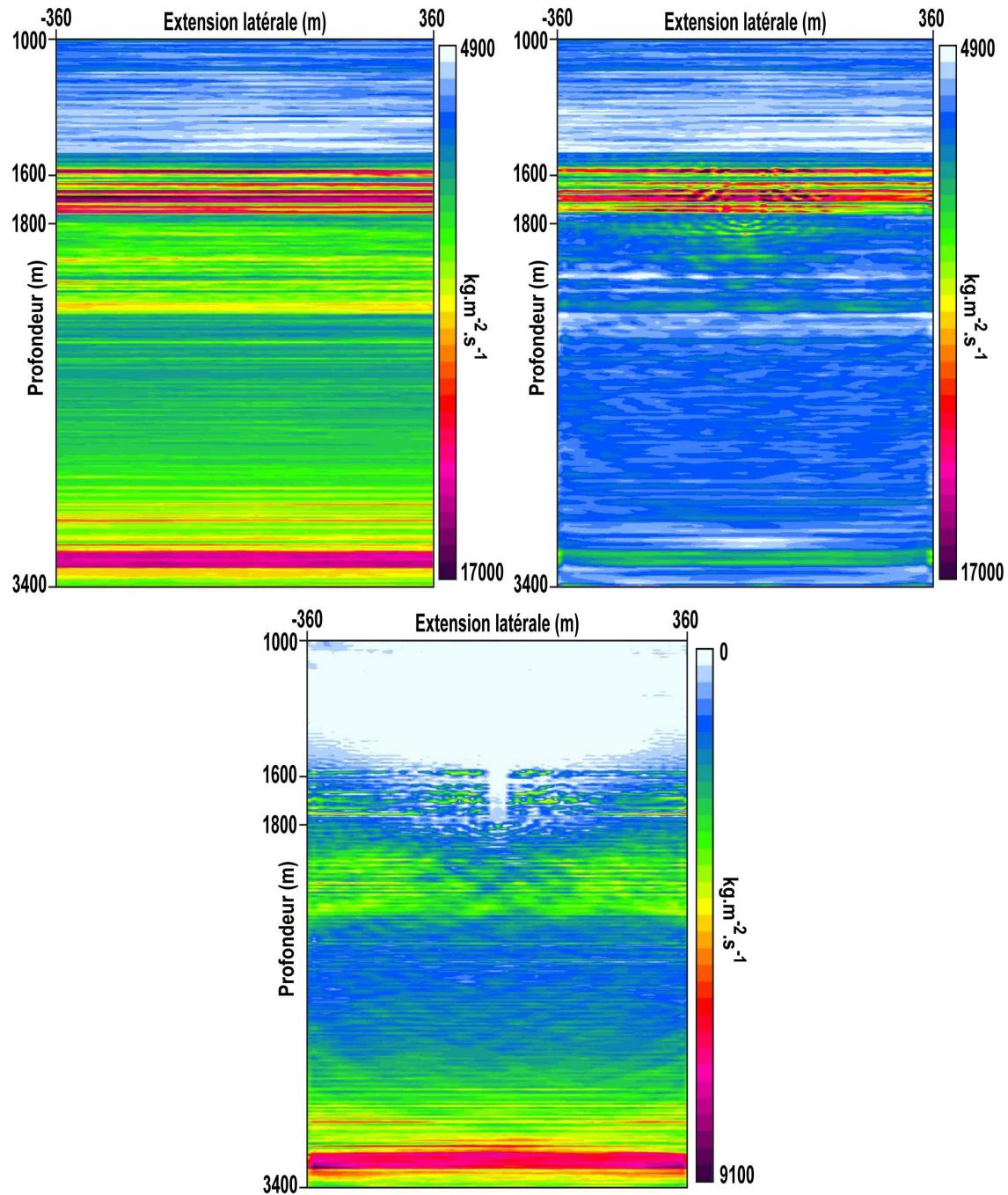


FIG. 4.10 – Modèle d'impédance exact (haut gauche), calculé (haut droite), différence (bas).

Beaucoup d'enseignements peuvent être tirés de cette vision 2D du modèle d'impédance calculé. Tout d'abord, il semble que la partie la mieux déterminée du modèle soit la partie supérieure (1000 m à 1800 m). Entre les profondeurs 1000 m et 1600 m elle est retrouvée de manière presque exacte sur toute la largeur (cf différence, fig. 4.10). Entre 1600 m et 1800 m de profondeur, ce résultat est perturbé. Le modèle est déterminé de manière exacte au niveau du puits, mais la régularité latérale n'est pas strictement conservée. Rappelons qu'ici aucun terme de régularisation n'est employé. Pour les profondeurs supérieures à 1800 m, c'est-à-dire sous le dernier récepteur, le résultat semble être sensiblement dégradé. Cependant, il convient de rappeler que l'information basse fréquence en dessous du puits n'est pas captée par les données sismiques, en raison notamment de la fréquence utilisée pour les sources ($f = 25$ Hz). Aussi, comme aucune information *a priori* sur cette basse fréquence n'est introduite dans l'inversion, les valeurs du modèle calculé peuvent s'éloigner considérablement de celles du modèle exact, sans qu'il soit impossible que ces deux modèles présentent le même profil de variations. Ici, la vue 2D du modèle calculé ne permet donc pas de conclure. Pour l'analyse de la résolution verticale dans la zone profonde, il convient donc de s'intéresser à une visualisation de profils sismiques verticaux, sélectionnés pour différentes positions autour du puits. C'est ce qui est réalisé dans la partie suivante.

4.3.2 Résolution verticale

4.3.2.1 Profils sismiques verticaux

La figure 4.11 présente les profils verticaux des modèles d'impédance exact (en noir) et calculé (en gris). Ces profils verticaux sont pris d'abord au niveau du puits ($x = 0$ m), puis pour différentes positions latérales ($x = 40$ m, $x = 100$ m, $x = 200$ m, $x = 300$ m).

Il est d'abord intéressant de noter que l'interprétation précédente concernant la partie supérieure du modèle est confirmée. Le modèle calculé correspond de très près au modèle exact dans la zone comprise entre 1000 m et 1600 m de profondeur. Dans la zone intermédiaire comprise entre 1600 m et 1800 m de profondeur, ces variations sont retrouvées de manière exacte au niveau du puits. Aux proches abords du puits ($|x| \leq 40$ m) des sauts d'impédance anormalement élevés sont visibles. Ces pics d'impédance évoquent une forme d'instabilité, qui peut être combattue à l'aide de termes de régularisation. Ensuite, à mesure que l'on s'éloigne du puits, ces contrastes semblent être au contraire atténusés.

En ce qui concerne le résultat dans la zone profonde ($z > 1800$ m), au niveau du puits, le résultat rappelle fortement celui obtenu pour l'inversion de données de puits 1D. La basse fréquence n'est pas retrouvée, mais les variations du modèle exact sont retrouvées avec une précision inférieure à la dizaine de mètres, et ce jusqu'à

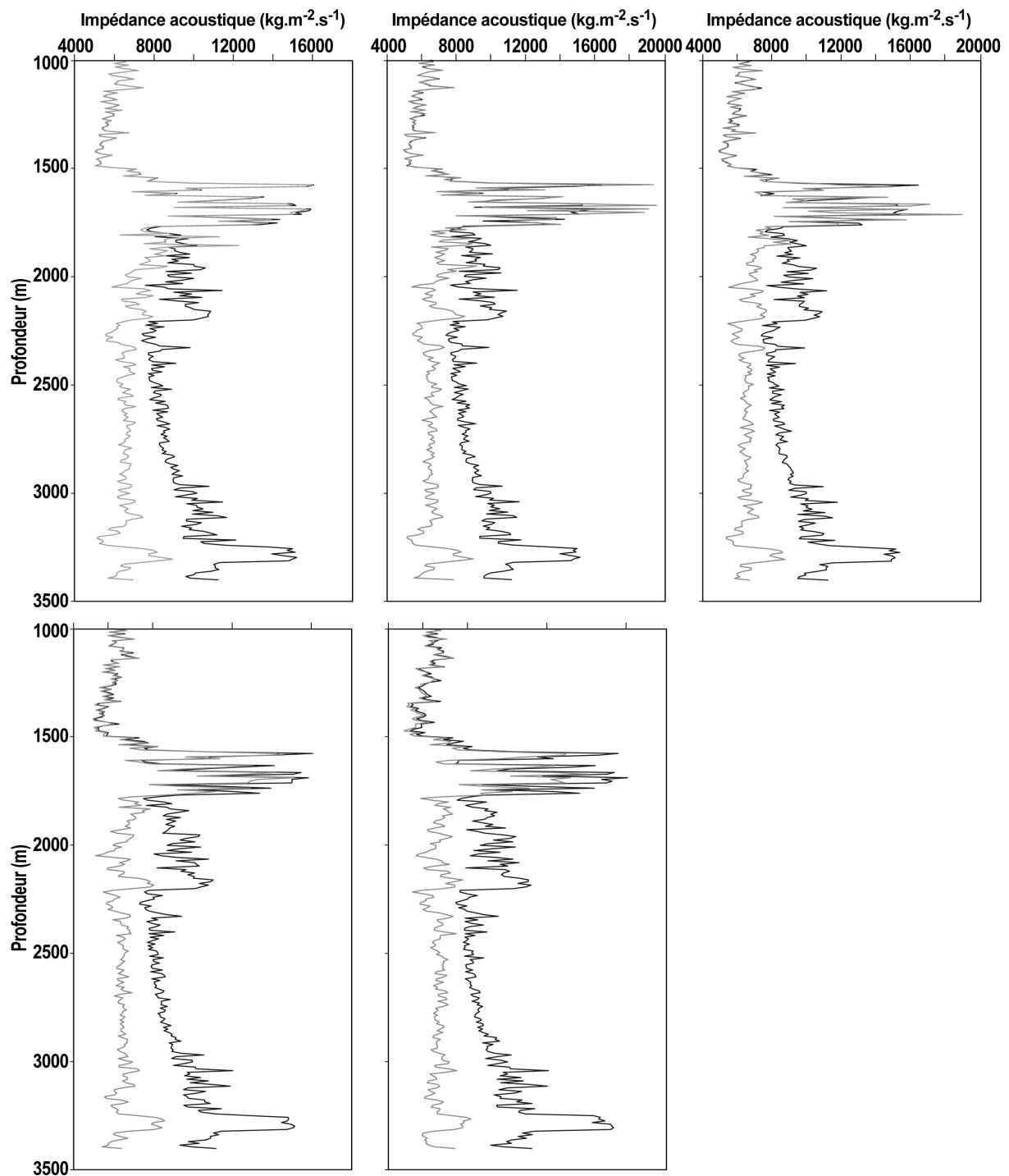


FIG. 4.11 – Profils sismiques verticaux exacts (noir), et calculés (gris). De gauche à droite, et de bas en haut : traces au puits ($x = 0 \text{ m}$), $x = 40 \text{ m}$, $x = 100 \text{ m}$, $x = 200 \text{ m}$, (e) $x = 300 \text{ m}$.

une profondeur de plus de 3000 m, soit plus de 1200 m sous le puits. De plus, ce constat reste valable, pour une grande part, pour les profils verticaux situés jusqu'à 300 m du puits. La précision des détails retrouvés diminue, néanmoins, même dans ces zones éloignées du puits, les variations verticales du modèle d'impédance sont retrouvées avec une grande acuité.

Afin de préciser encore ce résultat en termes de résolution verticale, il est possible d'effectuer une comparaison du contenu fréquentiel de la partie profonde des profils sismiques verticaux. Une comparaison du modèle exact, du modèle calculé, et d'un modèle obtenu à partir des mêmes données par un processus d'inversion linéaire classique permet de mettre en avant le gain en résolution résultant de l'inversion non linéaire de données sismiques.

4.3.2.2 Comparaison des spectres de fréquence

La densité spectrale des profils sismiques verticaux est obtenue en utilisant une conversion profondeur/temps à partir du modèle exact de vitesse correspondant à chacun des profils. Une fois le signal converti en temps, il est rééchantilloné avec un pas de discréétisation égal à 1×10^{-4} s, sur la fenêtre $[0, T]$. Ce sont les densités spectrales de ces signaux qui sont présentées ici, soit, pour un signal $s(t)$ et sa transformée de Fourier $\tilde{s}(w)$, la fonction

$$\frac{\tilde{s}(w)\tilde{s}^*(w)}{T},$$

où $\tilde{s}^*(w)$ désigne le conjugué de $\tilde{s}(w)$.

Comme la densité spectrale est symétrique par rapport à la fréquence 0, on ne s'intéresse qu'à la partie positive du spectre. Les résultats sont présentés sur la figure 4.12.

Le spectre des profils sismiques verticaux du modèle calculé par une méthode d'inversion linéaire classique est centré sur la fréquence de 40 Hz, et, au-delà de 60 Hz, l'amplitude du spectre est quasi-nulle. Au contraire, le modèle calculé par la méthode d'inversion de données sismiques de puits 2D permet de disposer de profils dont le spectre est beaucoup plus riche, et qui s'apparente à celui du modèle exact. Ainsi, la bande de fréquence pour laquelle le spectre est non-nul s'étend au delà de 80 Hz. De plus, la largeur de bande du signal est plus grande. Ceci atteste du gain potentiel en résolution verticale de la méthode d'inversion par rapport aux techniques d'inversion classique, et confirme les impressions obtenues après l'analyse des profils sismiques verticaux. Cette amélioration de la résolution résulte de la prise en compte au sein de l'inversion des réflexions multiples, qui sont ici nombreuses, en vertu de l'hétérogénéité du modèle imagé. Le modèle obtenu par cette méthode d'inversion non linéaire présente beaucoup plus de détails que celui obtenu par l'utilisation d'une méthode d'inversion linéaire classique.

Le gain en résolution vertical ayant été confirmé, il convient maintenant de s'intéresser à la précision avec laquelle le modèle est estimé dans son extension latérale. La question de l'indétermination mérite d'être soulevée : quelles sont les composantes latérales du modèle bien déterminées par l'information sismique ? De plus, quelle est la résolution latérale maximale qui peut être atteinte ? La partie suivante tente d'apporter des réponses à ces questions.

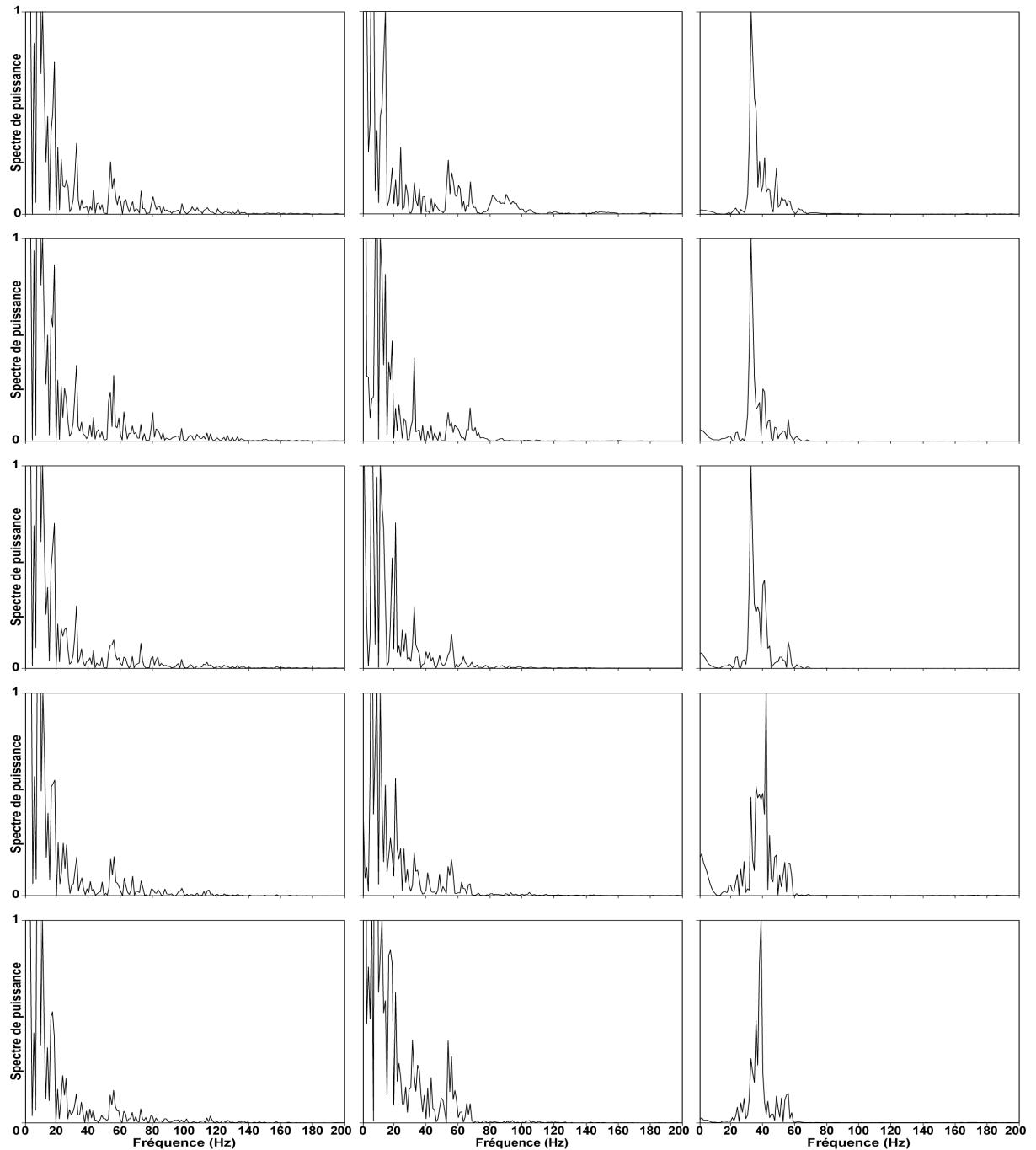


FIG. 4.12 – Spectres de puissance des profils sismiques verticaux. La première ligne correspond à l'offset 0 m, la seconde +40 m, la troisième +100 m, la quatrième +200 m, la cinquième +300 m. La première colonne correspond au modèle exact, la seconde au modèle calculé, la troisième au résultat d'inversion linéarisée.

4.3.3 Résolution latérale

L'information sismique disponible sur les variations latérales du milieu est moins importante que celle qui peut être obtenue sur les variations verticales. Ceci est dû principalement au mode d'acquisition. En effet, la gamme d'angle choisie pour la constitution des données est restreinte, et les champs d'ondes se propagent dans des directions proches de la verticale. Les signaux enregistrés par les récepteurs, alignés verticalement, prennent en compte principalement les réflexions issues de discontinuités verticales. Ainsi, un tel dispositif d'acquisition éclaire mieux les variations verticales que les variations latérales du sous-sol, ceci étant accentué par la profondeur, l'angle d'éclairage diminuant à mesure que celle-ci augmente. Ceci semble se vérifier sur le modèle d'impédance calculé. En effet, si la partie supérieure du modèle est parfaitement imagée, et avec elle sa structure particulière de stratification horizontale, il semble à l'inverse que dans la partie profonde, cette structure ne soit pas correctement retrouvée. Comme annoncé dans le chapitre 2, en l'absence de termes de régularisation latérale, l'information sismique semble ne pas être suffisante pour déterminer les paramètres du modèle éloignés latéralement avec précision. La possibilité de retrouver parfaitement la couche supérieure du modèle provient probablement, quant à elle, de l'hypothèse de connaissance des CP exactes, qui est une information *a priori* très forte. Il convient donc de pousser l'analyse plus avant afin de déterminer quelle est l'importance de la sous-détermination de l'extension latérale du modèle calculé. Pour cela, nous nous intéressons tout d'abord à la représentation de profils sismiques latéraux.

4.3.3.1 Profils sismiques latéraux

Les profils sismiques latéraux extraits des modèles d'impédance exact et calculé sont représentés sur les figures 4.13 et 4.14.

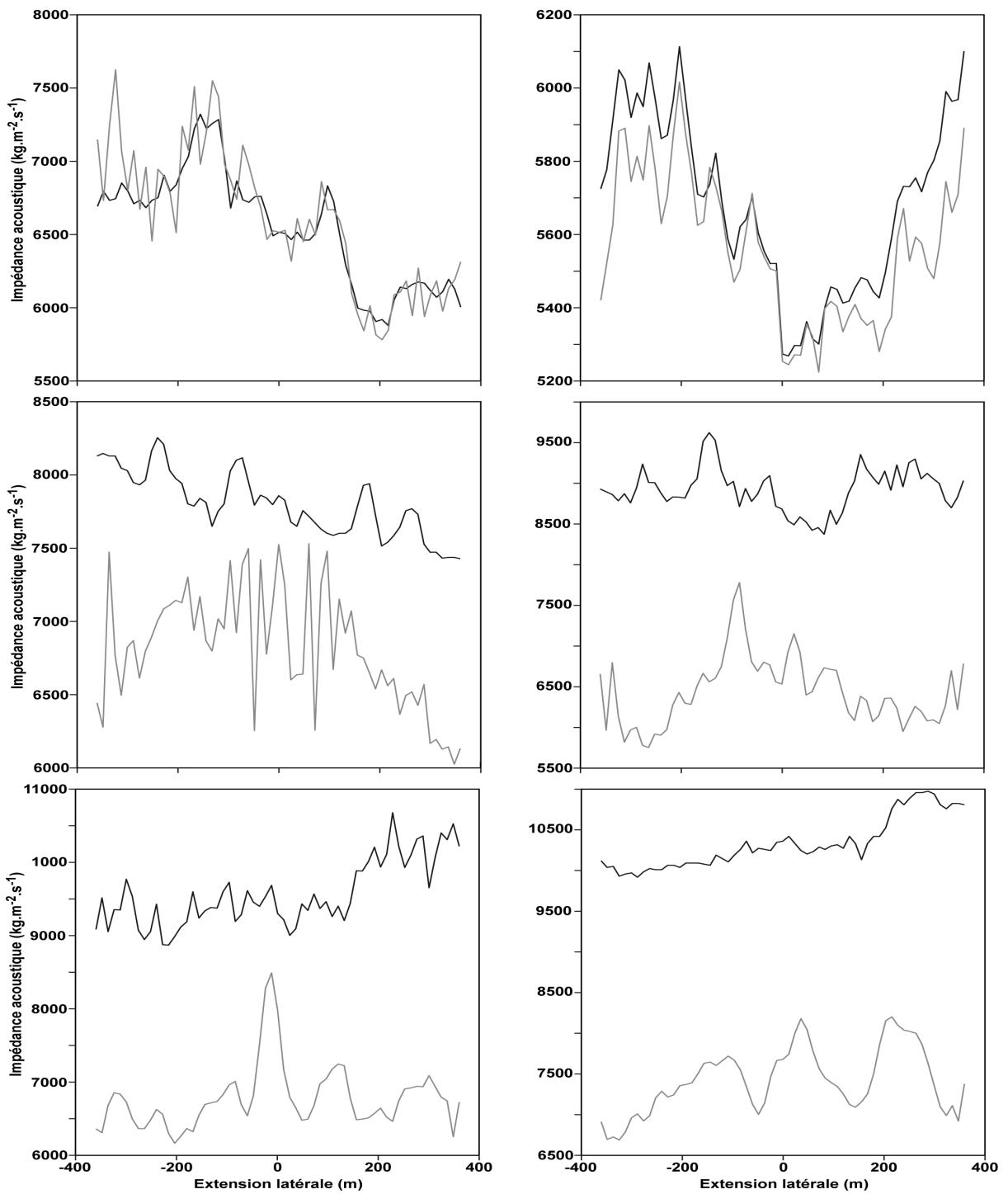


FIG. 4.13 – Profils sismiques latéraux exacts (noir) et calculés (gris). De gauche à droite, et de haut en bas, $z = 1000 \text{ m}$, $z = 1400 \text{ m}$, $z = 1800 \text{ m}$, $z = 2000 \text{ m}$, $z = 2100 \text{ m}$, $z = 2200 \text{ m}$.

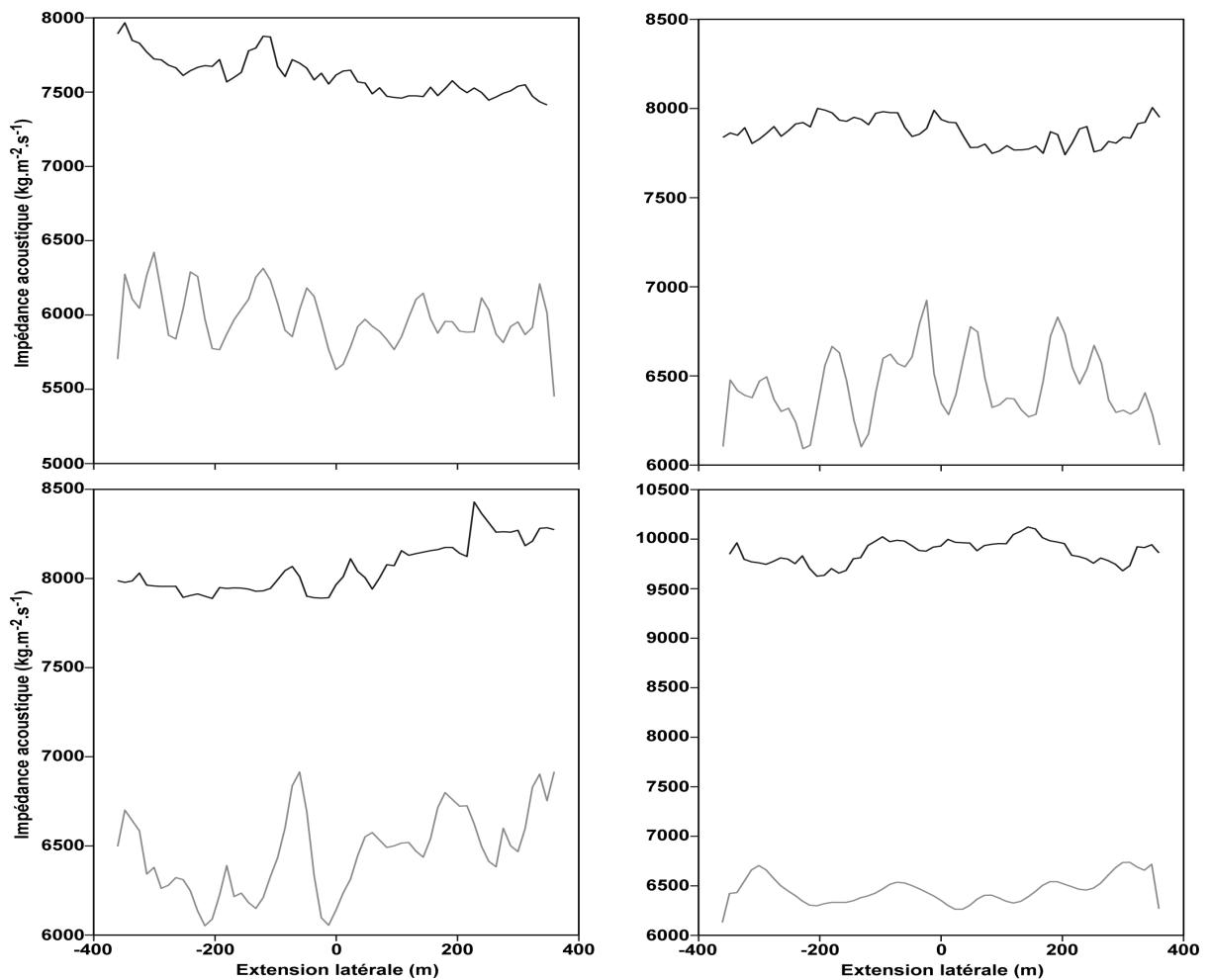


FIG. 4.14 – Profils sismiques latéraux exacts (noir) et calculés (gris). De gauche à droite, et de haut en bas, $z = 2300$ m, $z = 2400$ m, $z = 2500$ m, $z = 3000$ m.

Les profils latéraux extraits aux profondeurs correspondant à la zone supérieure du modèle corroborent les observations déjà effectuées jusqu'ici. La résolution latérale dans cette partie est excellente. Il est possible de noter le caractère instable du profil d'impédance à la profondeur du premier récepteur (fig. 4.13) : celui-ci présente des variations brutales non corrélées avec le modèle exact. Ceci est vraisemblablement dû à l'extrême sensibilité du modèle en ces premiers paramètres. On retiendra cependant l'excellente corrélation entre les profils latéraux du modèle exact et du modèle calculé à la profondeur de 1400 m. A 1800 m, on découvre une instabilité marquée : le profil sismique latéral du modèle calculé est très perturbé, même à proximité du puits. Ceci est probablement dû au statut particulier de cette zone qui marque la transition entre la zone supérieure où sont situés les récepteurs, et la zone profonde, d'où ils sont absents.

En zone profonde ($z > 1800$ m), les traces latérales semblent indiquer que l'information sismique n'est pas suffisante à la bonne détermination de l'extension latérale du modèle. La sous-détermination de l'extension latérale apparaît clairement. La tendance basse fréquence des variations semble être grossièrement retrouvée, mais les variations latérales à petite ou moyenne échelle ne sont globalement pas traduites dans le modèle d'impédance calculé. Au contraire, certains contrastes semblent être amplifiés, et l'établissement de corrélations entre le modèle exact et le modèle calculé est malaisé. De plus, ceci semble être vrai non seulement pour les variations lointaines du puits ($|x| > 200$ m), mais également pour les variations proches ($|x| < 200$ m). Cet exemple illustre donc les limites de l'information sismique pour la bonne détermination des variations latérales en zone profonde, et la nécessité d'introduire une information *a priori* sur la régularité du milieu pour s'affranchir de l'indétermination.

Une nouvelle expérience est réalisée dans le but d'approfondir ces résultats. Cette expérience consiste en l'analyse du comportement de la méthode sur des données obtenues à partir d'un modèle basé sur le modèle exact (fig. 4.1) et modifié par l'introduction d'une perturbation latérale très localisée. L'enjeu est d'obtenir une information sur la résolution latérale maximale de la méthode d'imagerie, par l'évaluation de sa capacité à reconstruire cette perturbation, et de la finesse de cette reconstruction. C'est ce qui est réalisé dans la partie suivante.

4.3.3.2 Evaluation de la résolution latérale

Nous appelons Point Spread Function la réponse de la méthode d'inversion à une perturbation ponctuelle du modèle exact d'impédance. L' "étalement" de l'image de cette perturbation, c'est-à-dire la précision avec laquelle elle est retrouvée, témoigne de la résolution maximale de la méthode. Comme cela a été vu, le modèle d'impédance dans la partie supérieure (de 1000 m à 1800 m) est reconstruit presque exactement. La résolution latérale de la méthode y est donc excellente. C'est en revanche dans la partie inférieure ($z > 1800$ m) que la sous-détermination latérale

de la solution est mise en évidence. C'est donc dans cette zone que l'évaluation de la résolution latérale va être effectuée. La recombinaison des données qui a été choisie simule un éclairage par ondes planes se propageant avec un angle maximal de 6° , ce qui doit permettre, en première approximation, d'imager correctement un réflecteur horizontal jusqu'à 150 m de chaque côté du puits, à la profondeur de 2500 m, comme cela a été dit dans la première partie de ce chapitre. Il semble donc naturel que la résolution latérale de la méthode soit mauvaise dans les parties très éloignées, et nous nous focalisons sur une zone d'extension latérale ne dépassant pas les 200 m autour du puits.

Aussi, nous nous intéressons à la réponse de la méthode à une perturbation de la taille d'une maille (12 m de large sur 8 m de hauteur), située à 144 m à droite du puits, et successivement à 2200 m et 2600 m de profondeur. L'amplitude de cette perturbation est fixée à 500% de la valeur du modèle initial. Cette amplitude n'est pas réaliste, néanmoins, elle permet de rendre l'effet de la perturbation sur les données suffisamment important pour que celle-ci soit prise en compte par la méthode d'inversion. L'intérêt de l'expérience est d'observer dans le modèle d'impédance résultat la présence éventuelle de cette perturbation et son étirement. Celui-ci permet d'évaluer la précision maximale avec laquelle il est possible d'imager une perturbation latérale. La méthode d'inversion est donc utilisée sur les données obtenues à partir de ce modèle perturbé, en utilisant le même dispositif de sources et la même recombinaison des données. Ces données sont qualifiées dans la suite de données perturbées. Le critère d'arrêt établi est le même que pour le cas précédent : la procédure d'inversion prend fin à l'issue de 5000 itérations de l'algorithme Quasi-Newton l-BFGS. La fonctionnelle coût et l'amplitude des résidus obtenus sont du même ordre de grandeur que pour le cas précédent (10^{-6}). L'acuité avec laquelle le modèle d'impédance calculé explique les données perturbées est donc identique à celle avec laquelle le modèle calculé dans l'exemple précédent explique les données non perturbées. Ceci rend possible la comparaison des résultats obtenus ici avec les précédents résultats.

La figure 4.15 présente le modèle d'impédance perturbé au point $x = 144$ m, $z = 2200$ m, et le résultat d'inversion correspondant.

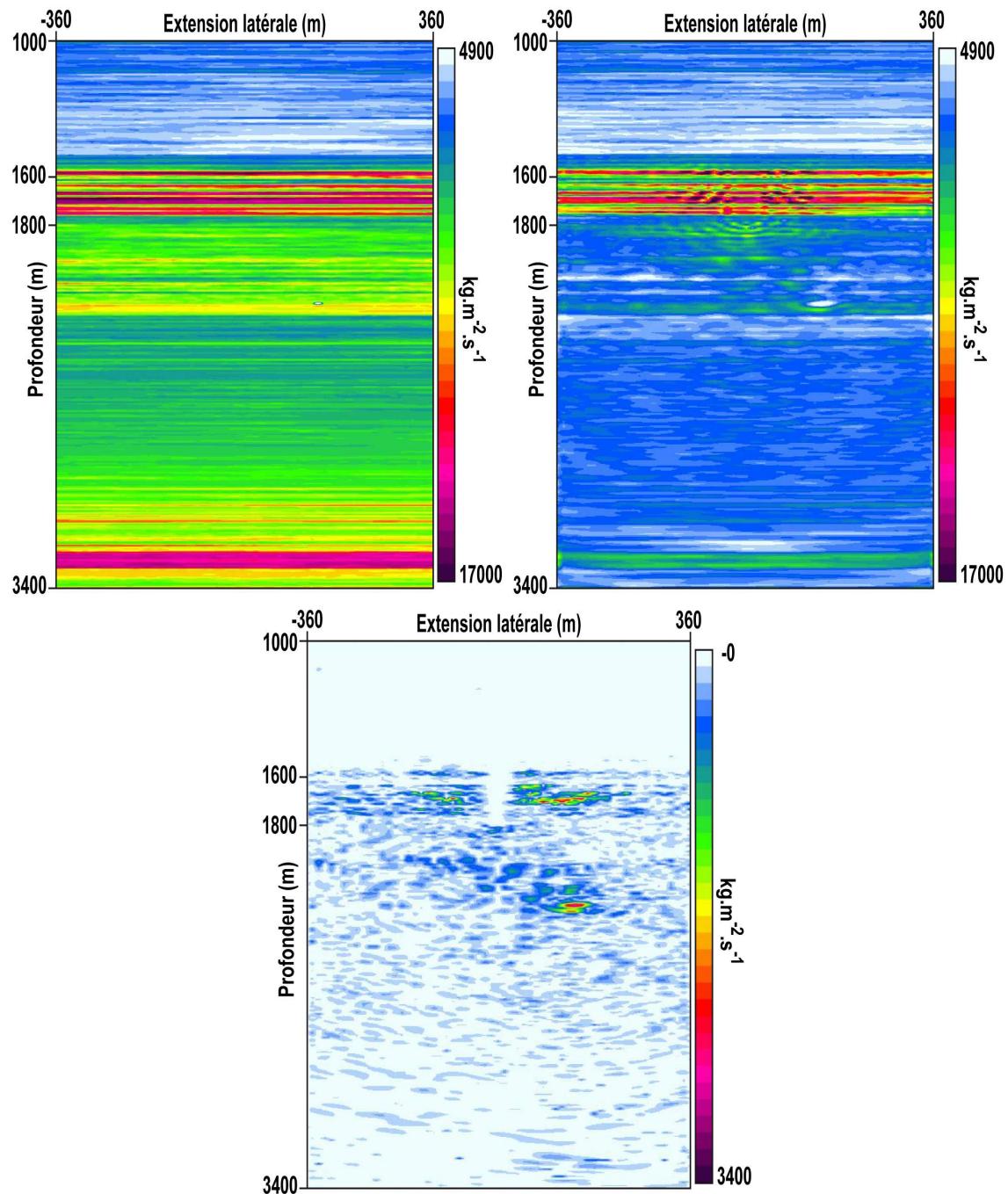


FIG. 4.15 – Modèle d’impédance après perturbation située en $x = 144$ m et $z = 2200$ m (haut gauche), calculé à partir des données perturbées (haut droite), différence entre les modèles d’impédance calculés avec et sans perturbation.

La perturbation du modèle d’impédance recherché provoque une discontinuité latérale blanche nettement visible dans le modèle calculé, au voisinage de l’emp-

lacement de la perturbation. Cette discontinuité est en revanche beaucoup plus étendue que la perturbation initiale. La taille de cette discontinuité avoisine en effet un carré de 30 m sur 30 m, pour une perturbation de 8 m de profondeur sur 12 m de largeur. Cet étalement témoigne bien d'une plus faible résolution latérale, de l'ordre de la trentaine de mètres, à comparer avec la résolution verticale inférieure à 10 m. Cette résolution latérale reste cependant de l'ordre de celle proposée par les méthodes d'inversion linéaires.

De plus, il est intéressant de comparer le résultat obtenu sur les données perturbées avec le résultat calculé sans perturbation (fig. 4.10, bas). Cette figure rend visible l'apparition de la discontinuité latérale à la profondeur de la perturbation ponctuelle qui a été introduite. Cependant, elle met en évidence le caractère instable de la bande située entre 1600 m et 1800 m de profondeur. En effet, le résultat obtenu pour les données perturbées a subi dans cette zone des modifications d'un ordre de grandeur proche de celui de celles qui correspondent à la localisation de la perturbation ponctuelle, bien que sur des zones plus restreintes. Le mode de visualisation par interpolation peut être ici trompeur, car il semble que ce soient des valeurs isolées qui prennent des valeurs anormales. Cependant, hors de cette zone, la différence entre le modèle calculé à partir des données sans perturbation et celui calculé avec les données perturbées fait apparaître une tâche correspondant à l'image de la perturbation par la méthode d'inversion. Celle-ci est donc susceptible de détecter une telle perturbation, et sa résolution latérale n'excède pas la trentaine de mètres. De plus, ceci tend à indiquer qu'en dehors de cette zone d'instabilité, les composantes de la solution soient assez bien déterminées, puisque des données sensiblement différentes mènent à un résultat similaire.

La même expérience, effectuée pour une perturbation située à 144 m du puits, mais à la profondeur de 2600 m cette fois, est présentée par la figure 4.16.

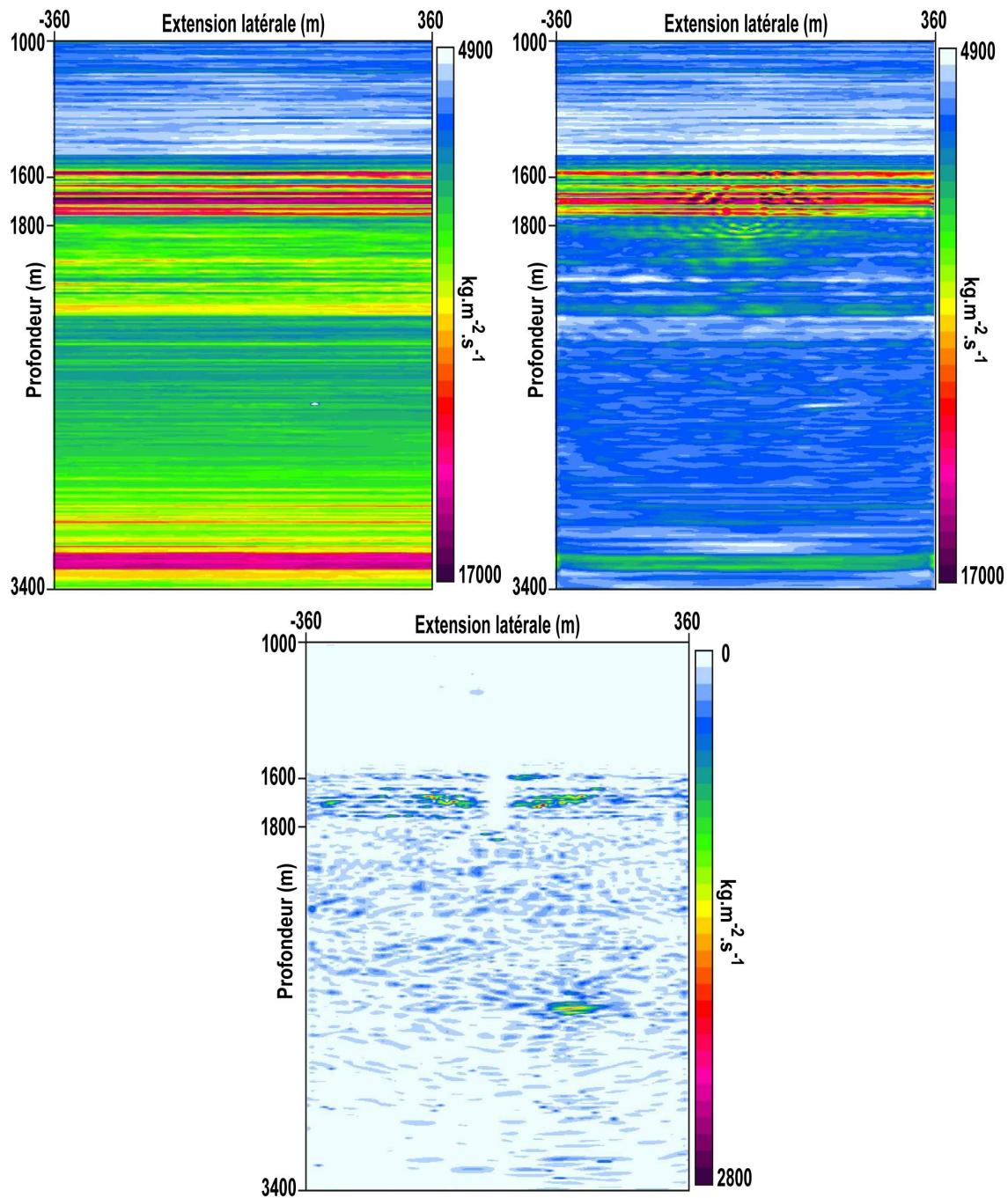


FIG. 4.16 – Modèles d’impédance, après perturbation située en $x = 144$ m et $z = 2600$ m (haut gauche), calculé à partir des données perturbées (haut droite), différence entre les modèles d’impédance calculés avec et sans perturbation.

Les résultats obtenus pour cette position sont similaires aux résultats obtenus précédemment. En dehors d’une zone dont l’instabilité a déjà été mise en avant,

la perturbation latérale est détectée efficacement. L' "étalement" constaté est légèrement plus important que pour la position précédente, proche de 40 m, attestant d'une dégradation de la résolution latérale en profondeur, ce qui est conforme à l'intuition : l'information sismique enregistrée par les récepteurs se fait plus faible à mesure que la profondeur augmente.

4.3.4 Premières conclusions

L'investigation du comportement de la méthode dans le cadre idéal pour lequel l'indétermination de couplage peut être évitée par la connaissance des CP est riche d'enseignements. Le gain en résolution vertical, pour un cas test faisant intervenir 30 sources réparties tous les 24 m en surface, émettant un signal dont le spectre est centré sur la fréquence de 25 Hz, est prometteur. L'analyse des spectres de fréquence des profils sismiques verticaux, comparés à ceux obtenus pour un résultat classique, montre clairement un enrichissement de la bande de fréquence, y compris vers des valeurs voisines de 80 Hz. Ceci atteste d'un important gain en résolution verticale, que ce soit dans la zone supérieure ($z < 1800$ m) ou dans la zone profonde ($z > 1800$ m).

En ce qui concerne la détermination des variations latérales, celles-ci sont retrouvées de manière presque exacte dans la zone supérieure. La résolution latérale dans cette zone est donc excellente. Ce résultat peut cependant être imputé au contexte de l'expérience qui a été mené, dans lequel les CP sont supposées connues, de manière exacte. Une attention particulière est donc requise sur ce point dans l'expérience suivante, qui est effectuée dans un contexte d'application plus réel pour lequel les CP ne sont pas connues.

L'analyse de la résolution latérale en zone profonde est plus problématique. En particulier, la méthode d'inversion semble être assujettie à une forte indétermination des variations latérales. Cette indétermination ne peut être levée que par l'introduction d'une information *a priori* sur la régularité latérale du modèle d'impédance. Cependant, il semble que la méthode d'inversion développée ne soit pas complètement insensible aux variations latérales du sous-sol. En effet, les profils sismiques verticaux montrent une très bonne corrélation entre le modèle calculé et le modèle exact, ceci même pour les profils éloignés du puits ($x = 300$ m). De plus, les profils latéraux montrent que le modèle calculé suit approximativement la basse fréquence du modèle exact, et la structure latérale des forts contrastes d'impédance semblent être reconstruite efficacement, en l'absence d'information *a priori* sur cette structure, et ce même à une profondeur de 3300 m (cf vue 2D du modèle calculé figure 4.10). Pour finir, la méthode a montré son aptitude à détecter des perturbations latérales ponctuelles en zone profonde ($z > 1800$ m). Il est également possible d'en déduire, par l'étalement de la perturbation reconstruite, une résolution latérale de l'ordre de la trentaine de mètres, ce qui rejoint sur ce point les méthodes d'inversion

linéaires classiques.

Ces résultats donnent donc une idée du potentiel d'imagerie de la méthode d'inversion de données sismiques de puits 2D. Il convient maintenant d'observer dans quelle mesure ces résultats sont dégradés pour une utilisation "réelle" de la méthode, lorsque les CP ne sont pas connues. L'indétermination de couplage entre l'impédance et les CP doit alors être compensée par l'introduction d'information *a priori* sur la régularité latérale de ces inconnues. Ces résultats font l'objet de la partie suivante.

4.4 Analyse de la résolution réelle (CP inconnues)

Dans cette partie, les résultats numériques obtenus pour l'inversion globale sont exposés. Après avoir donné quelques précisions sur les poids de régularisation, sur le choix du schéma d'interpolation des CP et le modèle d'impédance *a priori* utilisé dans le puits, le résultat 2D est présenté. Avant de passer à l'analyse de la résolution verticale et latérale de ce résultat, une comparaison est effectuée avec le résultat obtenu avec les mêmes données en utilisant une méthode d'optimisation globale, agissant sur toutes les inconnues, ceci afin d'illustrer la forte indétermination de couplage entre les inconnues d'impédance et celles liées aux CP, et la nécessité d'employer la méthode d'optimisation emboîtée définie au chapitre 3. Le fil directeur de cette partie est l'évaluation de la dégradation du résultat par rapport à la partie précédente, et celle de l'influence des termes de régularisation.

4.4.1 Poids de régularisation

L'utilisation de la méthode d'inversion dans le contexte où les CP sont inconnues requiert l'introduction de termes de régularisation, comme cela a été mentionné dans le chapitre 2, afin de limiter les effets de l'indétermination. Les poids affectés à ces termes de régularisation sont évalués par une méthode d'essai-erreur, et l'examen attentif sur des cas tests de l'apparition d'un conflit entre le coût sismique et le coût associé à la régularisation. En effet, les critères de régularisation étant par nature inexacts, leur minimisation rentre en contradiction avec celle du coût sismique. En réalité, le réglage des poids de régularisation se fait par la recherche du poids critique à partir duquel l'information *a priori* et l'information sismique deviennent antagonistes. Ce poids critique fournit un seuil en-deça duquel l'information *a priori* est prise en compte sans gêner la minimisation du coût sismique. Cette information permet aux composantes mal déterminées de la solution d'être stabilisées. Une prise en compte efficace de cette information est possible pour des valeurs de poids de régularisation proche de ce poids critique. Un critère complémentaire de calibration est l'analyse du résultat *a posteriori*. Ceci permet de juger sur pièce si le résultat obtenu n'est pas suffisamment régulier, ce qui implique une augmentation de la pondération, ou inversement. Dans la pratique, ces poids de régularisation sont également ajustés en fonction du niveau de bruit sur les données. Si celles-ci sont très bruitées, l'information sismique est moins fiable, et il convient donc de compenser cette perte d'information par l'information *a priori*, en affectant un poids plus élevé aux termes de régularisation.

Dans l'exemple traité, le poids de régularisation utilisé pour les CP interpolées, (cf Annexe 3.1) est fixé à 10^{-15} . Celui utilisé pour l'information *a priori* dans le puits est fixé à 10^{-6} . D'autre part, l'expérience numérique acquise sur l'exemple traité nous a conduit à définir un poids de régularisation variable avec la profondeur. En effet, la zone supérieure ($z < 1800$ m) est sujette à d'importantes instabilités, tandis que la zone profonde ($z > 1800$ m) l'est relativement moins. Ainsi, il est naturel d'u-

tiliser un poids de régularisation plus fort dans la partie supérieure, qui diminue progressivement pour atteindre une valeur constante dans la partie inférieure. Dans la pratique, le poids de régularisation est fixé à 10^{-4} à la profondeur du premier récepteur $z = 1000$ m, et à 10^{-6} à partir de la profondeur du dernier récepteur $z = 1800$ m. Les valeurs intermédiaires dans la partie supérieure sont calculées par interpolation linéaire.

Ces précisions ayant été données sur la détermination des poids de régularisation, il convient maintenant d'aborder le problème de l'interpolation des CP.

4.4.2 Interpolation des CP

L'interpolation des CP dans l'algorithme d'inversion de données sismiques de puits 2D requiert la définition d'une trame de traces temporelles du champ de pression à la profondeur du premier récepteur, sur laquelle l'interpolation doit se baser. Cette trame doit être suffisamment légère en nombre de points pour favoriser le gain en temps de calcul, tout en étant suffisamment étouffée pour fournir les degrés de liberté nécessaires à l'algorithme d'optimisation, en charge de retrouver les CP optimales. Dans le cas test présenté ici, le nombre de points de discréétisation dans la dimension latérale vaut 101 et le nombre de traces sélectionnées vaut 15, soit une réduction voisine d'un facteur 7 du nombre d'inconnues prises en compte dans l'optimisation grâce à l'interpolation. La répartition de ces traces temporales est présentée en figure 4.17.

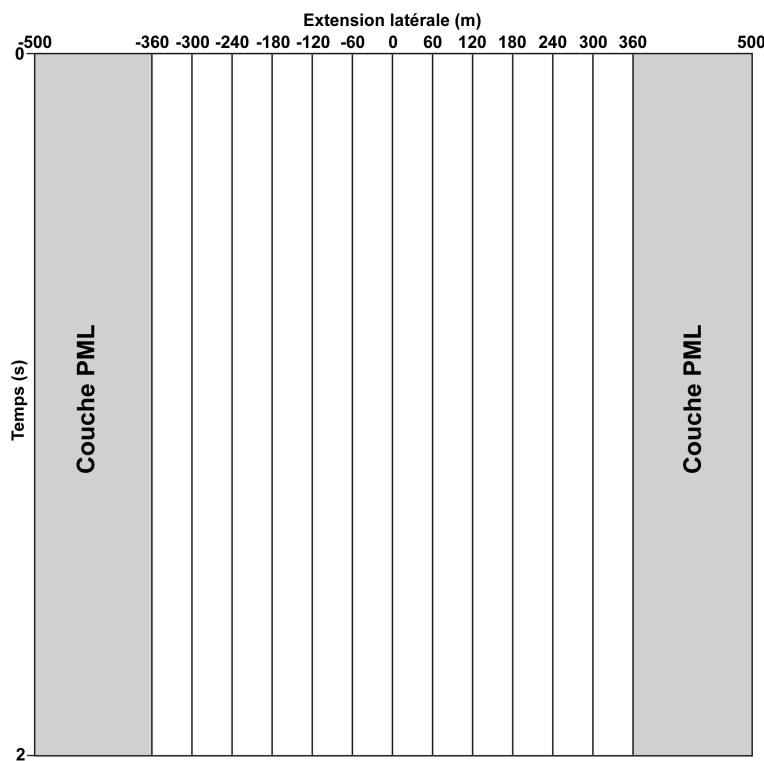


FIG. 4.17 – Trame utilisée pour l'interpolation des CP.

La trame choisie utilise un échantillonage régulier de part et d'autre du puits, une trace étant sélectionnée tous les 60 m, soit une trace tous les 5 points de discréttisation, à l'exception de la zone PML, où aucune trace n'est sélectionnée. Cette disposition vise à fournir les degrés de liberté à l'algorithme d'optimisation dans la zone où la solution est la mieux déterminée. En effet, la CP dans la PML n'a que très peu d'influence sur les signaux enregistrés aux récepteurs et, par conséquent, elle y est fortement indéterminée. Ainsi, les traces situées au bord de la PML sont fixées à 0, et les premières traces non nulle sélectionnées correspondent à la frontière entre le domaine d'intérêt et la couche PML latérale. La CP dans la PML est calculée exclusivement par interpolation, entre la trace au bord fixée à 0 et cette trace à l'interface entre la PML et le domaine d'intérêt.

Pour donner une idée de la précision de l'interpolation, les CP exactes et les CP calculées par interpolation à partir des traces exactes sont présentées figure 4.20.

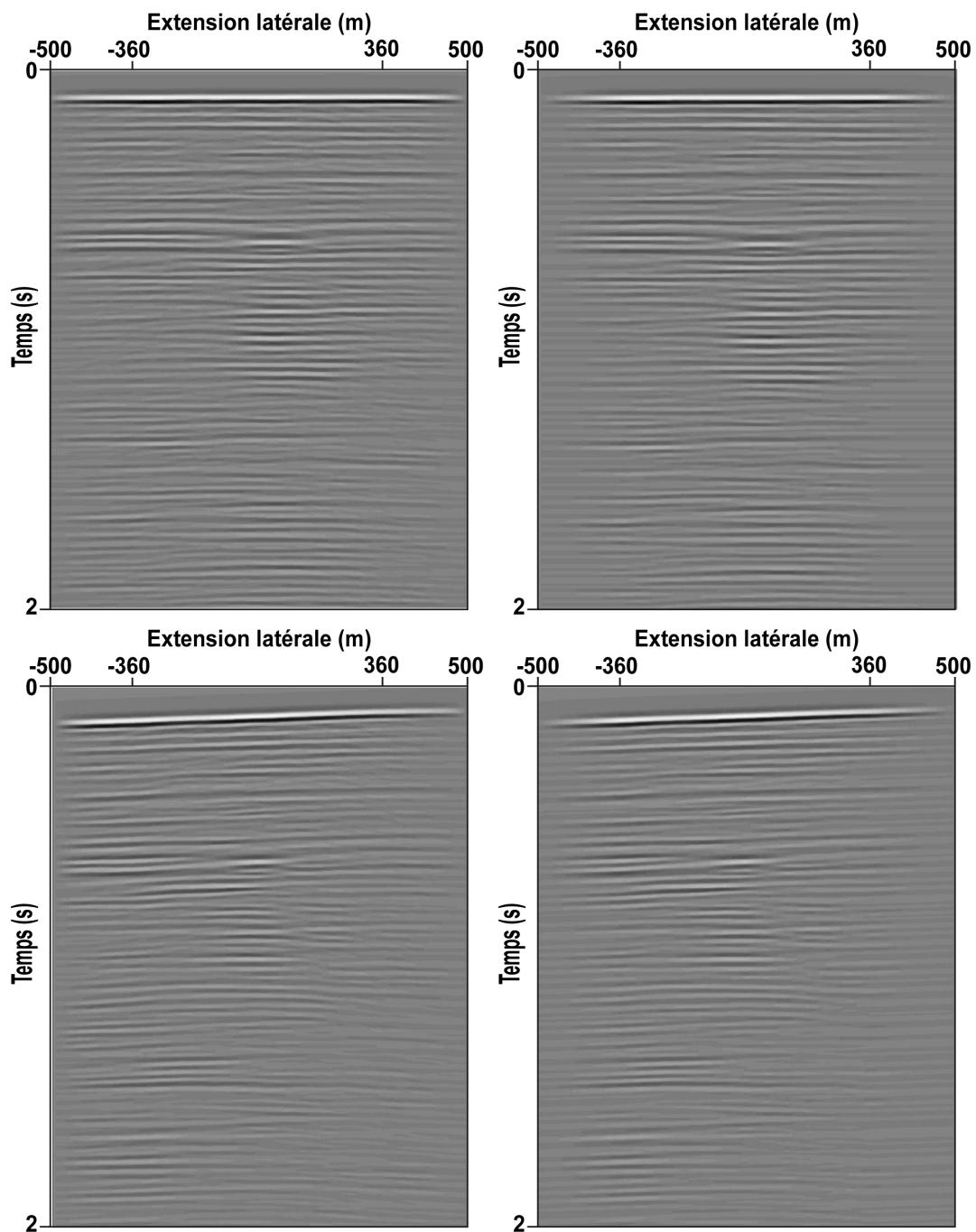


FIG. 4.18 – Comparaison entre les CP exactes (gauche) et les CP interpolées à partir des traces exactes (droite). Première ligne : données recombinées par angle d'éclairage 0° , deuxième ligne -6° .

Une perte d'information due à l'interpolation est donc observée dans les couches PML latérales, ce qui, comme cela a été mentionné, n'est pas gênant, en raison de

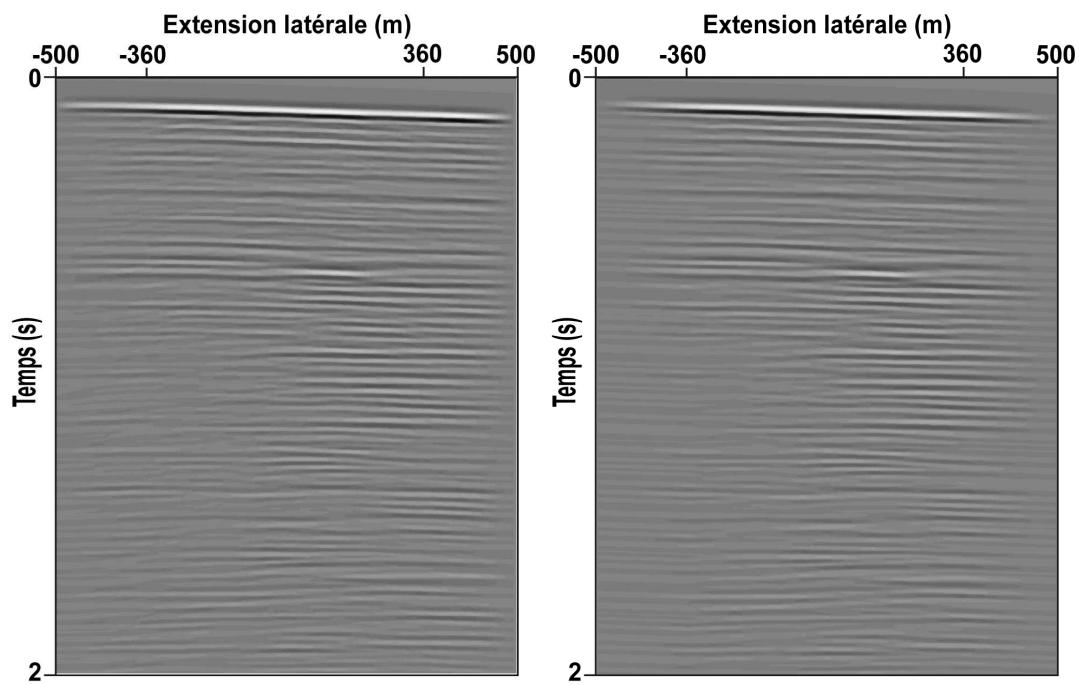


FIG. 4.19 – Comparaison entre les CP exactes (gauche) et les CP interpolées à partir des traces exactes (droite), pour les données recombinées par angle d'éclairage 6° .

l'indétermination inhérente à cette partie des CP. Il est possible d'observer également dans la partie interne de la CP la perte de l'information non conforme à l'information *a priori* sur la régularité. L'amplitude de cette information reste cependant faible relativement à l'amplitude maximale du signal.

La définition de la trame d'interpolation ayant été donnée, il reste, avant d'analyser les résultats, à présenter le modèle d'impédance *a priori* utilisé dans le puits.

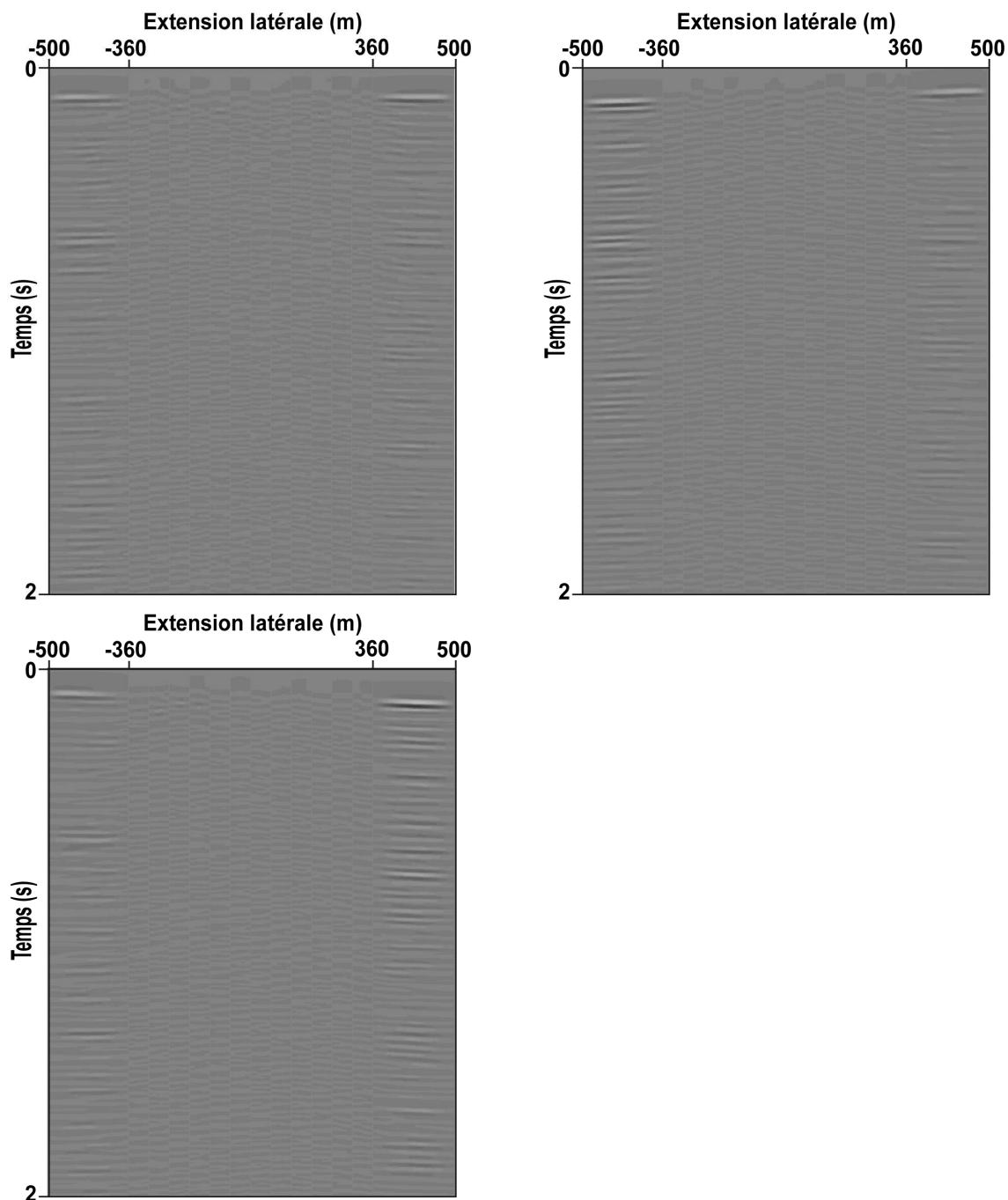


FIG. 4.20 – Différences entre les CP exactes et les CP interpolées, correspondant aux données recombinées par angle d'éclairage 0° (haut gauche), -6° (haut droite), 6° (bas).

4.4.3 Modèle d'impédance *a priori* dans le puits

Comme cela a été évoqué dans le chapitre 2, il est possible en pratique, à partir de relevés stratigraphiques, de déterminer un modèle d'impédance *a priori* 1D,

au niveau du puits ($x = 0$ m). Un tel modèle a été utilisé pour cette expérience. Néanmoins, pour ne pas fausser l'analyse de la résolution de la méthode d'inversion, et pour utiliser une hypothèse plus conforme à ce qui est pratiqué en réalité, le modèle d'impédance *a priori* choisi est une estimation locale basse fréquence du modèle exact, et non pas le modèle exact lui-même. Ainsi, aucune information haute résolution n'est introduite, et aucun artefact expérimental ne vient s'immiscer dans l'analyse de celle-ci. Cette information basse fréquence se révèle toutefois très utile pour réduire l'indétermination de couplage entre les CP et l'impédance, notamment en ce qui concerne l'amplitude des CP. Ce modèle local 1D basse fréquence est présenté par la figure 4.21.

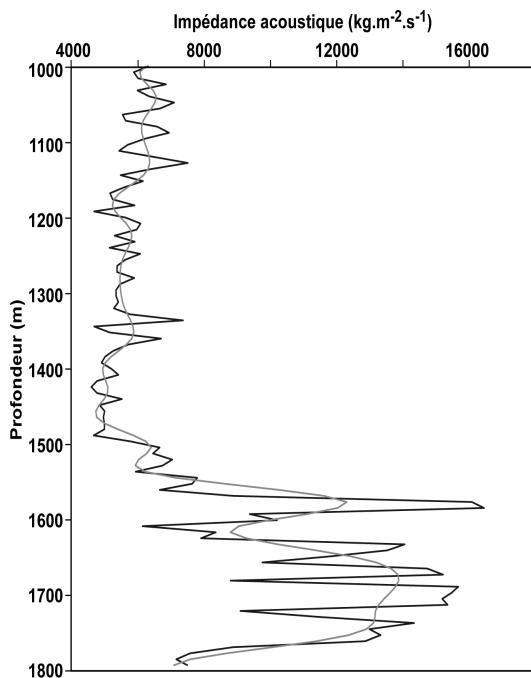


FIG. 4.21 – Comparaison entre le modèle exact (noir) et le modèle basse fréquence utilisé comme modèle *a priori* (gris) localement au puits ($x = 0$ m et $1000 \text{ m} < z < 1800 \text{ m}$).

Les outils nécessaires à la mise en place de l'inversion globale ayant été introduits, il est maintenant possible de présenter le résultat obtenu.

4.4.4 Présentation du résultat

Le modèle initial d'impédance est un modèle homogène, initialisé à la valeur du modèle d'impédance exact au niveau du premier récepteur ($x = 0$ m, $z = 1000$ m). Le modèle initial pour les traces représentant les CP est un modèle nul. Les CP initiales sont donc nulles.

Le résultat est obtenu à l'issue de 1000 itérations de l'algorithme Quasi-Newton, en autorisant à chaque fois au plus 60 itérations de l'algorithme du gradient conjugué. Le temps requis pour y parvenir est d'environ 36 heures, sur 128 coeurs de calcul du cluster de l'IFP.

La fonction coût J s'exprime formellement de la manière suivante

$$J = \frac{s + r}{s_0 + r_0}$$

où s est le coût sismique, s_0 le coût sismique initial, r le coût de régularisation, et r_0 le coût de régularisation initial (avant inversion). Cette fonction coût J vaut donc 1 avant l'inversion. Sa courbe de décroissance en fonction du temps est présentée figure 4.22.

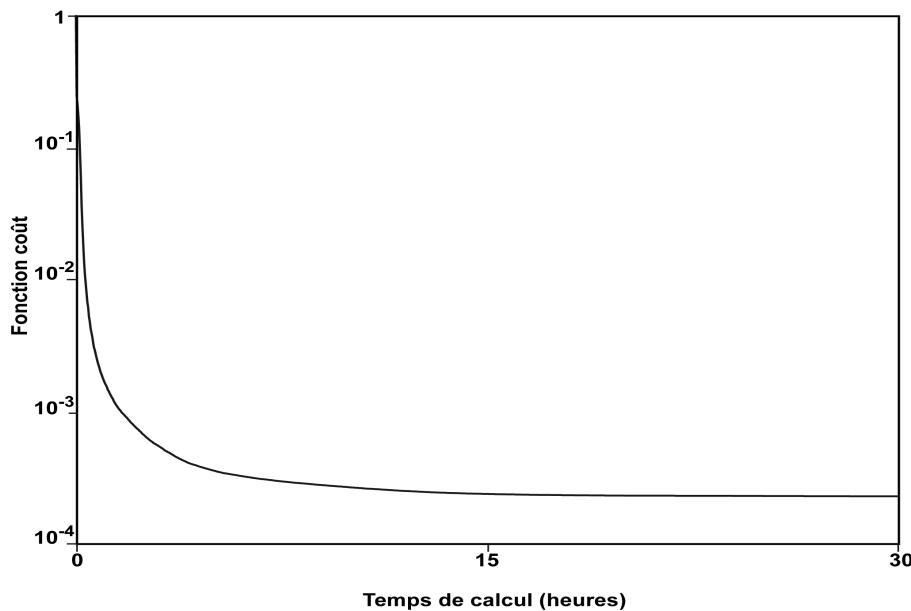


FIG. 4.22 – Courbe de décroissance de la fonction coût en fonction du temps pour l'inversion 2D à CP inconnues.

Cette décroissance indique que l'algorithme d'optimisation a fonctionné de manière satisfaisante. La fonctionnelle atteint la valeur de 1×10^{-3} en moins de deux heures de calcul. A l'issue des 1000 itérations de la boucle externe Quasi-Newton, elle est réduite de près de 4 ordres de grandeur, et vaut 2.33×10^{-4} . Il convient de détailler plus avant cette décroissance en distinguant la progression du coût sismique et du

coût de régularisation. Les valeurs numériques sont

$$s_0 = 4.26 \times 10^{-8}, \quad r_0 = 2.38 \times 10^{-12}.$$

La première chose importante à noter est

$$\frac{r_0}{s_0} \ll 1.$$

Le modèle initial d'impédance étant homogène, et les CP initiales étant nulles, le coût initial de régularisation provient exclusivement du terme de régularisation au puits (minimisation de l'écart en norme L^2 du modèle d'impédance au puits et du modèle *a priori*). Ce coût est largement dominé par le coût sismique. Au commencement de la minimisation, l'optimiseur, pour réduire le coût total, doit donc privilégier les directions de descente qui minimisent le coût sismique.

A l'issue des 1000 itérations de Quasi-Newton, les valeurs finales du coût sismique et du coût de régularisation sont

$$s_f = 4.35 \times 10^{-12}, \quad r_f = 5.58 \times 10^{-12}.$$

Ainsi, à l'issue des itérations

$$\frac{r_f}{s_f} \simeq 1.$$

Le coût sismique et le coût de régularisation finaux sont du même ordre de grandeur. De plus, le coût final de régularisation r_f est plus élevé que le coût initial de régularisation r_0 . Cette information est très importante. En effet, c'est le signe qu'au cours de l'optimisation, les modifications apportées au modèle d'impédance et aux CP ont engendré un conflit entre l'information sismique et l'information de régularisation. Les modèles finaux résultent donc d'un compromis entre ces deux informations. D'autre part, les valeurs finales de ces coûts, relativement aux valeurs initiales, indiquent qu'un modèle expliquant les données avec une bonne précision (10^{-4}) a pu être déterminé. Les poids de régularisation ont donc été choisis suffisamment faibles pour laisser prédominer l'information sismique. Il reste à savoir maintenant si l'information de régularité a été prise en compte de manière satisfaisante sur les modèles résultats, ou si elle a été pondérée trop faiblement. Ceci est donné par l'analyse directe des résultats.

L'analyse des résidus, tout d'abord, permet de confirmer la diminution du coût sismique. La précision avec laquelle les données sont expliquées est, sans surprise, moins bonne que dans le cas idéal pour lequel les CP sont connues, néanmoins, elle reste excellente. La figure 4.23 présente ces résidus.

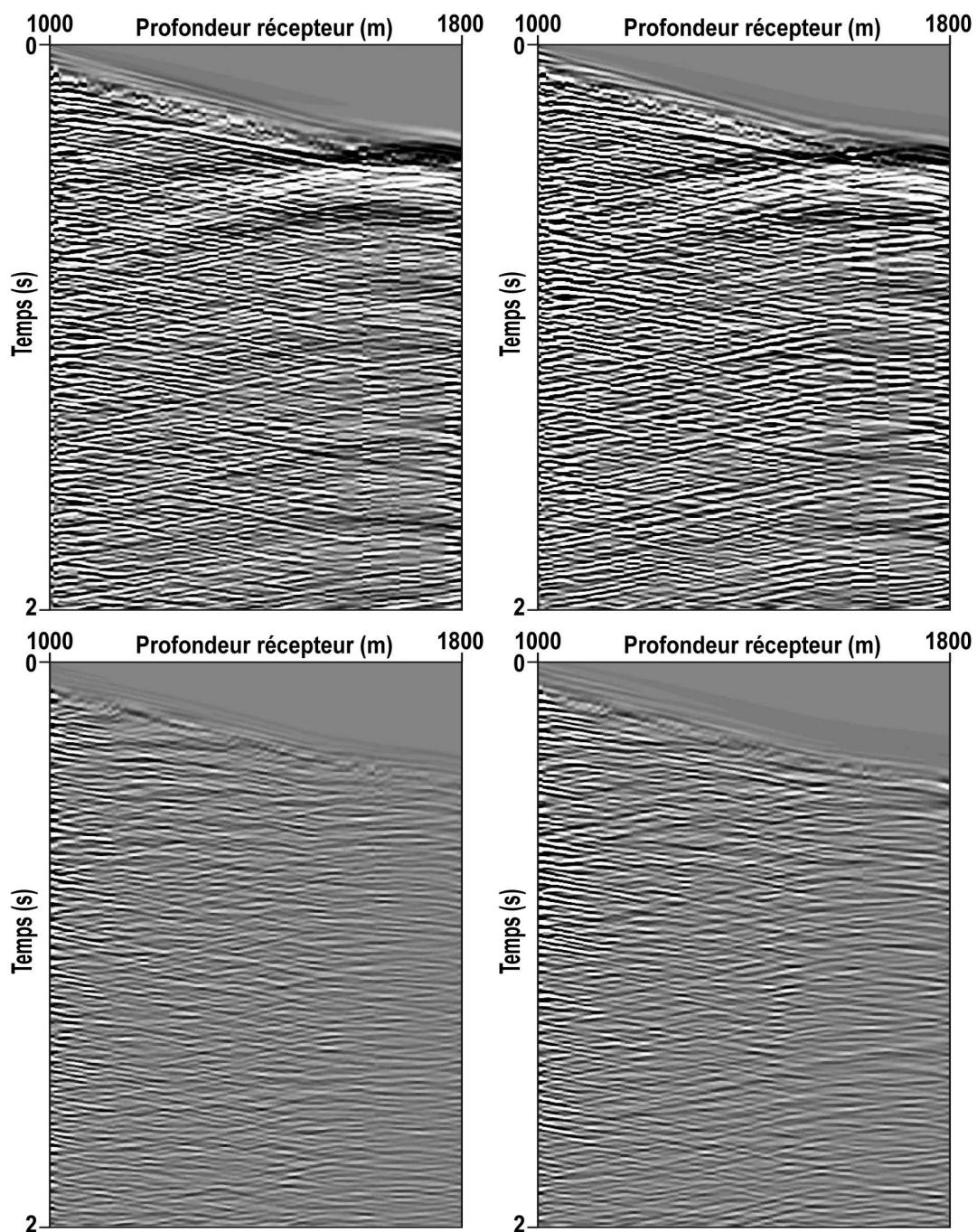


FIG. 4.23 – Résidus. Vitesse verticale 0° (haut gauche), vitesse verticale 6° (haut droite) , vitesse horizontale 0° (bas gauche), vitesse horizontale 6° (bas droite).

Cette figure représente les résidus correspondant aux données recombinées pour un éclairage par onde plane d'incidence 0° et 6° , avec une amplification d'un facteur 1000 par rapport à l'échelle utilisée pour représenter les données pour les données

de vitesse verticale, et d'un facteur 100 seulement pour les données de vitesse horizontale. Ces résidus sismiques, bien que plus élevés que dans le cas précédent, sont suffisamment faibles pour considérer que le modèle d'impédance et les CP calculées expliquent les données avec une grande précision.

Il convient maintenant de comparer les CP calculées par optimisation avec les CP exactes. Cette comparaison est présentée par les figures 4.24 et 4.25. Les différences entre les CP calculées et les CP exactes permettent de délimiter une zone de 300 m de part et d'autre du puits au sein de laquelle la détermination des CP est effectuée avec une très bonne précision. En dehors de cette zone, des différences plus marquées apparaissent, spécialement en ce qui concerne la zone PML. Cependant, ceci est conforme au choix de la trame de traces qui supporte l'interpolation. Les valeurs de la CP situées en dehors de cette zone ne jouent qu'un faible rôle dans la détermination du signal qui parvient aux récepteurs : elles sont en principe indéterminées. Elles sont donc fixées par un critère de régularité, qui, on le voit ici, n'est pas exactement conforme à la réalité. Ainsi, la partie principale des CP est reconstruite de manière satisfaisante et l'information de régularité supplée à l'information sismique pour la détermination de l'extension latérale des CP. Le poids de régularisation affecté aux CP d'une part, et la trame d'interpolation d'autre part, ont donc été choisis judicieusement.

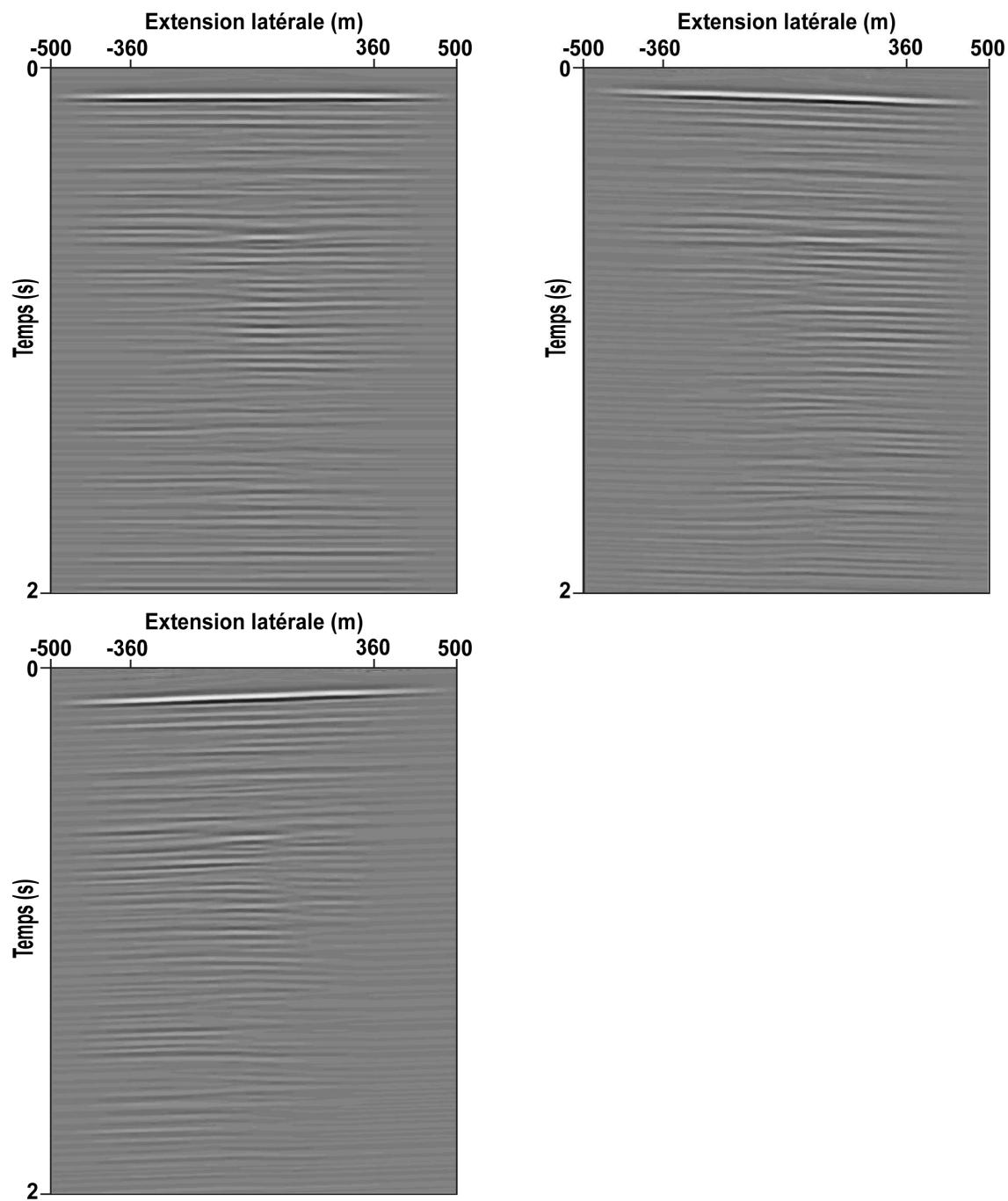


FIG. 4.24 – CP calculées, 0° (haut gauche), -6° (haut droite), 6° (bas).

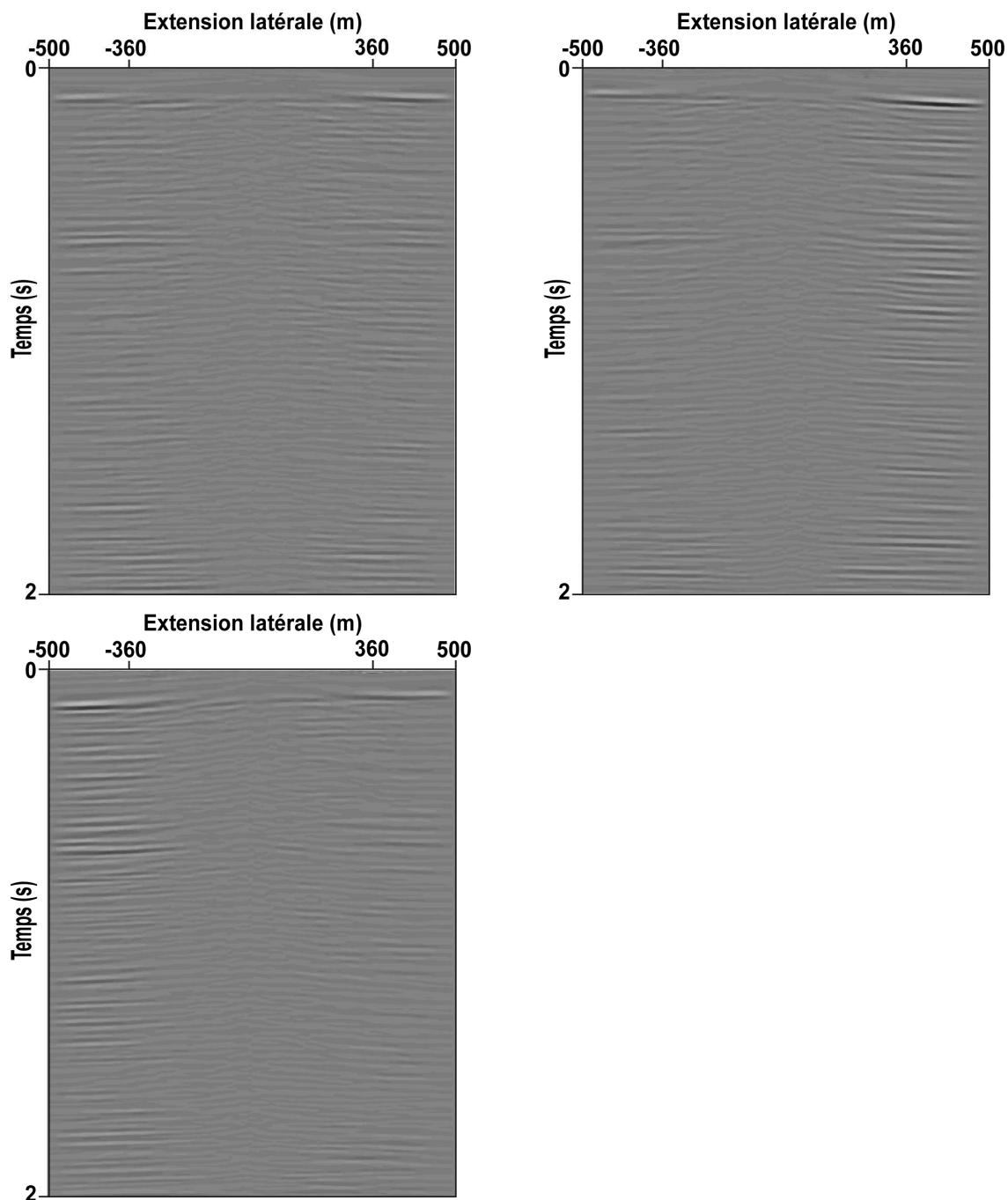


FIG. 4.25 – Différences entre CP calculées et CP exactes, 0° (haut gauche), -6° (haut droite), 6° (bas).

Il reste maintenant à analyser le modèle d'impédance obtenu. Celui-ci est présenté figure 4.26, ainsi que sa différence avec le modèle exact. La dégradation du résultat par rapport à celui obtenu pour des CP connues est surtout visible, de prime abord, dans la partie supérieure du modèle. En effet, cette partie est la plus sen-

sible à la bonne détermination des CP. Comme celles-ci ne sont plus maintenant qu'une approximation des CP exactes, le modèle d'impédance s'en retrouve dégradé. Néanmoins, cette dégradation reste assez limitée. En effet, la différence entre le modèle exact et le modèle calculé (fig. 4.26, bas) montre une zone de concordance presque parfaite s'étendant jusqu'à 200 m de part et d'autre du puits entre ces deux modèles. De plus les différences enregistrées pour des positions latérales plus lointaines dans la zone supérieure restent limitées.

En ce qui concerne la partie profonde, l'introduction de l'information *a priori* sur l'impédance permet de retrouver la structure latérale des évènements principaux. Le modèle final présente donc une régularité horizontale marquée, qui n'empêche pas pour autant l'expression de certaines hétérogénéités latérales. Le poids de régularisation affecté à la minimisation des variations latérales du modèle d'impédance a donc été lui aussi convenablement choisi.

Avant de passer à l'analyse détaillée de la résolution verticale par la visualisation de profils verticaux pris à différentes distances du puits, il est intéressant de mettre en évidence l'indétermination due au couplage des CP et du modèle d'impédance dans la procédure d'inversion, et l'importance d'utiliser la méthode d'optimisation emboîtée pour lutter contre cette indétermination. C'est ce qui est réalisé dans la partie suivante.

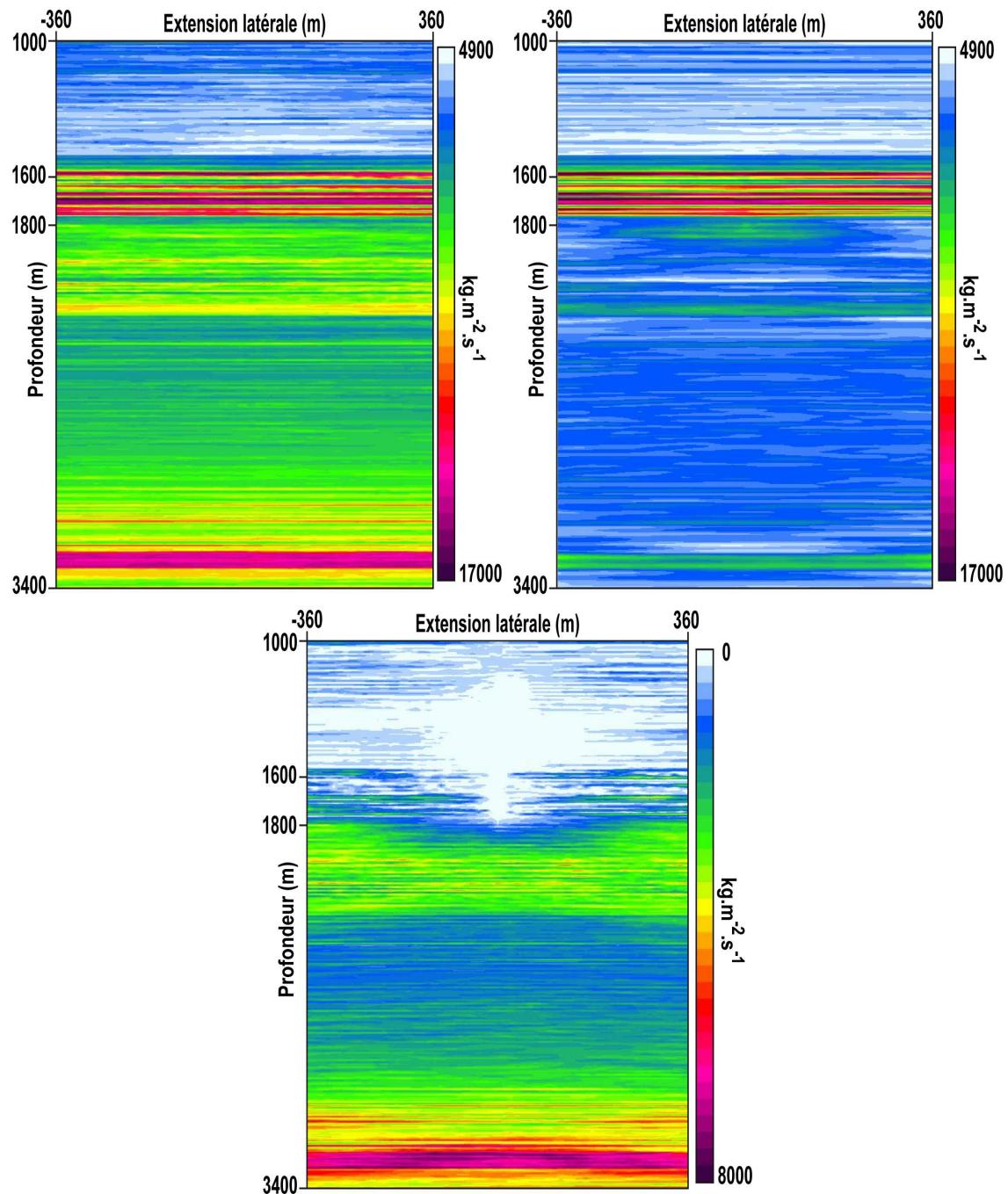


FIG. 4.26 – Modèle d’impédance exact (haut gauche), modèle d’impédance calculé par inversion avec CP inconnues (haut droite), différence (bas).

4.4.5 Indétermination et optimisation emboîtée

Ce paragraphe est consacré à une comparaison entre les résultats qui viennent d’être présentés, obtenus par la résolution du problème inverse par la méthode d’op-

timisation emboîtée décrite dans le chapitre 3, avec les résultats obtenus par l'utilisation d'une méthode d'optimisation qualifiée de globale, qui correspond à une résolution du problème par l'algorithme de Quasi-Newton appliqué à toutes les variables, sans introduire de distinction entre les CP et le modèle d'impédance.

L'application de l'algorithme de Quasi-Newton à toutes les inconnues implique la résolution d'un problème d'optimisation conséquent comportant un peu plus de 422000 inconnues. La méthode l-BFGS est spécifiquement dédiée à cette catégorie de problèmes de grande taille, celle-ci ne représente donc pas de limite intrinsèque à l'utilisation de cet algorithme. Afin de pouvoir comparer les résultats, le critère d'arrêt imposé à la méthode d'inversion globale est le temps maximal de calcul autorisé, qui est pris égal à 36 heures, le temps nécessaire à l'algorithme d'optimisation emboîtée pour effectuer 1000 itérations de la boucle principale de Quasi-Newton. Sur le même cas test, en ce temps imparti, la méthode globale ne parvient pas à réduire le coût global au delà de 6.82×10^{-2} , tandis que la méthode emboîtée parvient à un coût final égal à 2.33×10^{-4} . La figure 4.27 présente la comparaison des courbes de convergence en fonction du temps sur une échelle logarithmique en base 10.

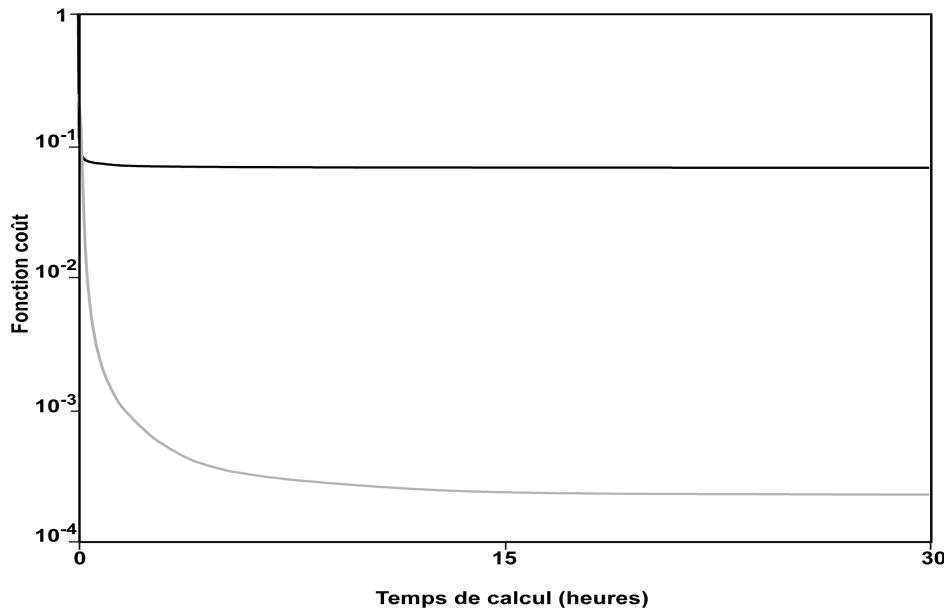


FIG. 4.27 – Convergence comparée de la méthode d'optimisation globale et la méthode d'optimisation emboîtée.

Le mauvais conditionnement du problème pris dans sa globalité est mis en évidence par la très mauvaise convergence de l'algorithme d'optimisation globale. Cette mauvaise convergence se traduit par des résultats très éloignés des modèles

recherchés. Les CP calculées par la méthode d'optimisation globale sont présentées par la figure 4.28.

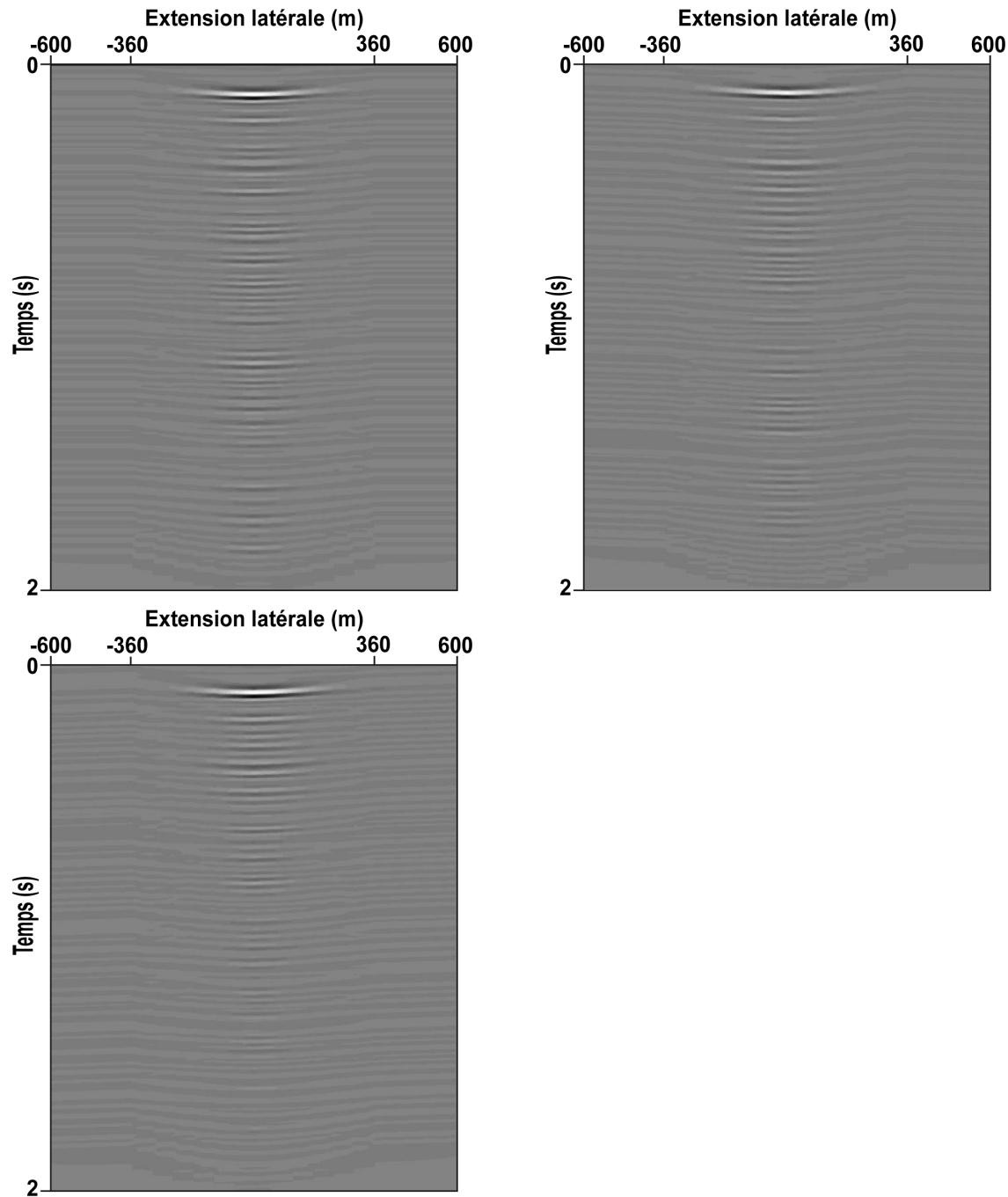


FIG. 4.28 – CP calculées par optimisation globale, 0° (haut gauche), -6° (haut droite), 6° (bas).

Le modèle d'impédance calculé par la méthode d'optimisation globale est présenté par la figure 4.29. Ces résultats mettent en évidence l'indétermination de cou-

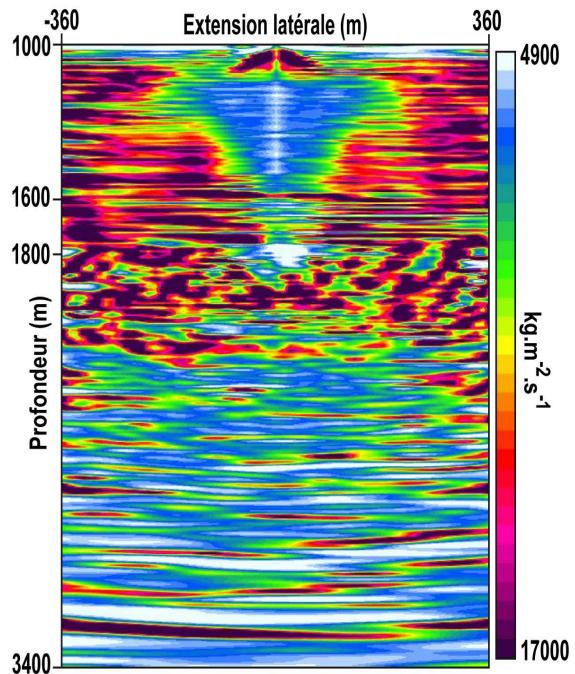


FIG. 4.29 – Modèle d’impédance calculée par optimisation globale.

plage inhérente à la reconstruction simultanée des CP et du modèle d’impédance. Les CP calculées sont déterminées dans une zone très proche du puits. Dans les zones éloignées latéralement, les CP ne sont pas reconstruites. Le modèle d’impédance présente quant à lui un aspect très irrégulier. Il est à peu près reconstruit au niveau du puits, mais de la même façon que pour les CP, l’extension latérale du modèle n’est pas imagée. Pourtant, les poids de régularisation appliqués, tant pour les CP que pour le modèle d’impédance, sont identiques à ceux qui sont utilisés dans le cas précédent. La procédure d’optimisation emboîtée, qui procède à la détermination des CP optimales correspondant au modèle d’impédance courant, avant chaque mise à jour de ce modèle par l’algorithme de Quasi-Newton, permet donc de s’affranchir d’une très forte indétermination. Plutôt que de tenter d’expliquer les données sismiques en utilisant toutes les inconnues simultanément, il explique d’abord la part du signal qu’il peut attribuer aux CP, puis celle due au modèle d’impédance, en affinant cette approximation à chaque itération. Son utilisation est cruciale pour la convergence de la méthode d’inversion développée.

Il convient maintenant de s’intéresser à l’analyse de la résolution de la méthode, en commençant par la résolution verticale, qui est menée d’abord par la visualisation de profils sismiques verticaux.

4.4.6 Résolution verticale

4.4.6.1 Profils sismiques verticaux

Les profils sismiques verticaux extraits du modèle calculé (fig. 4.26) et comparés à ceux extraits du modèle exact (fig. 4.1) sont représentés figure 4.30

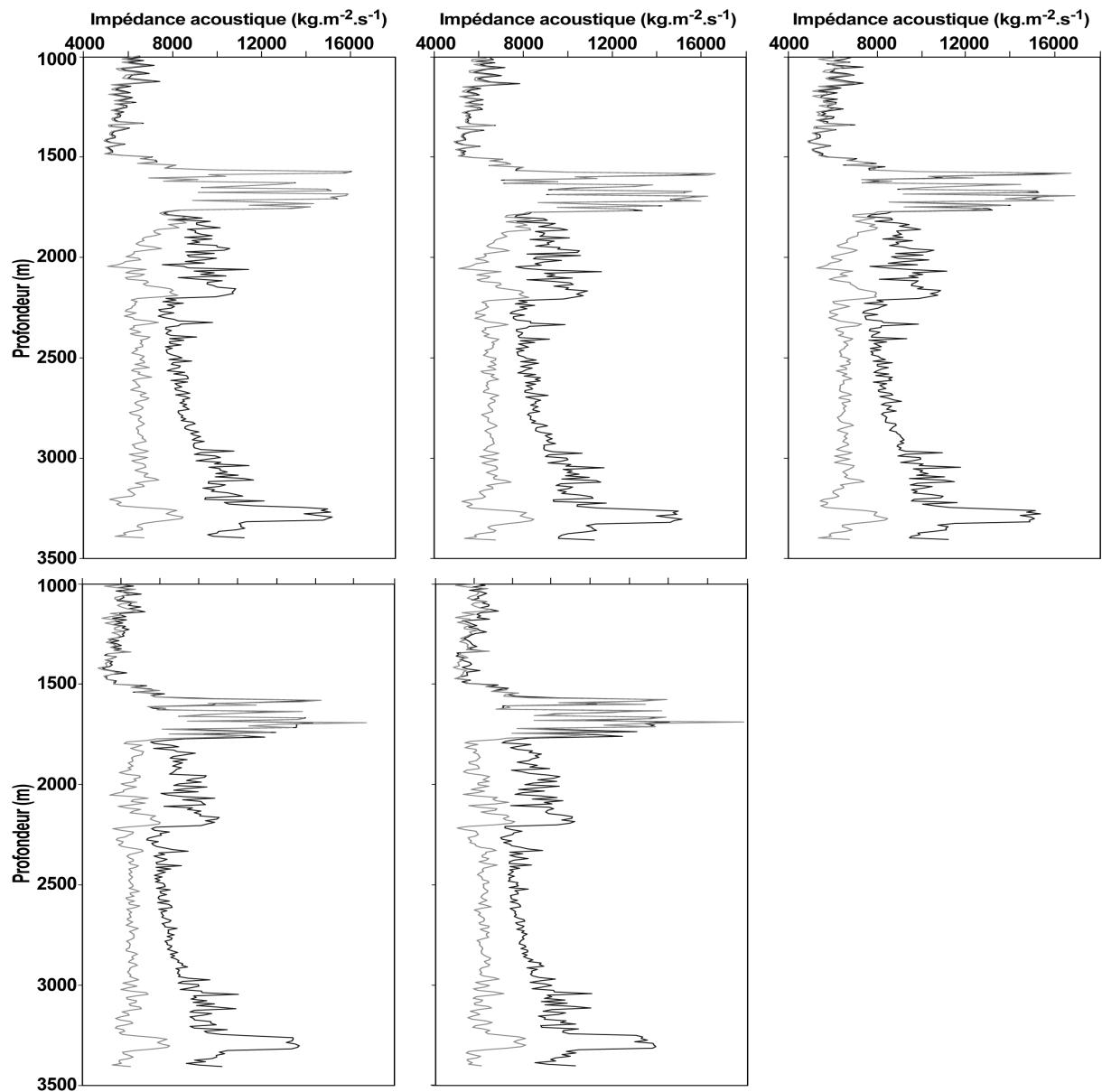


FIG. 4.30 – Profils sismiques verticaux exact(noir), et calculés (gris). De gauche à droite, et de haut en bas, traces au puits ($x = 0$ m), $x = 40$ m, $x = 100$ m, $x = 200$ m, $x = 300$ m.

L'analyse des profils sismiques verticaux confirme l'interprétation précédente concernant la reconstruction de la partie supérieure ($z < 1800$ m). Alors que précédemment, celle-ci était retrouvée de manière quasi-exacte, elle est ici sujette à quelques écarts. Néanmoins, jusqu'à 100 m du puits, il semble qu'il s'agisse simplement d'un décalage de la basse fréquence. A partir de 200 m, des erreurs plus conséquentes apparaissent mais la tendance principale est cependant très bien retrouvée.

La zone profonde ($z > 1800$ m), quant à elle, semble être retrouvée, à première vue, avec le même niveau de détails que dans la partie précédente. Cependant, la comparaison entre les profils sismiques verticaux du modèle résultat de la partie précédente et ceux du modèle calculé ici montre une certaine forme de lissage de l'information, qui accorde la thèse d'une légère perte de résolution. En dépit de ce lissage, il convient cependant de noter l'excellente résolution intrinsèque du résultat, qui présente des corrélations entre le modèle calculé et le modèle résultat à une échelle inférieure à 10 m. Afin de confirmer la dégradation de la résolution par rapport au cas précédent, il est nécessaire de procéder à la comparaison des spectres en fréquence des profils sismiques verticaux des différents modèles d'impédance.

4.4.6.2 Analyse du spectre de fréquence

La comparaison des spectres de fréquence des profils sismiques verticaux calculés pour des CP exactes avec celui des profils sismiques verticaux calculés pour des CP inconnues est présentée figure 4.31.

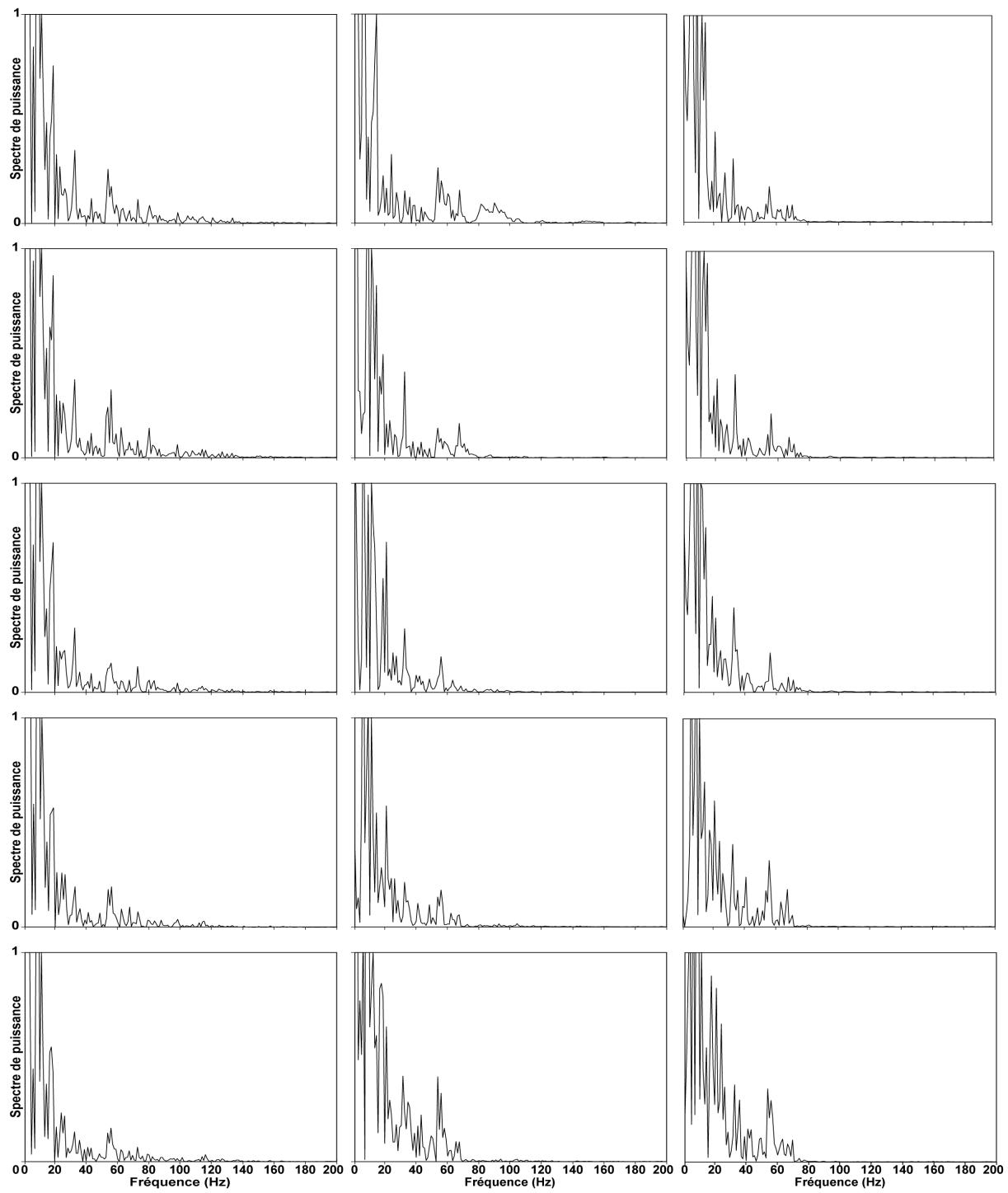


FIG. 4.31 – Spectres de puissance. Par lignes : $x = 0$ m, $x = 40$ m, $x = 100$ m, $x = 200$ m, $x = 300$ m. Par colonnes : modèle exact, modèle calculé avec CP inconnues, modèle calculé par migration.

Cette comparaison permet de mettre en évidence une perte de résolution au niveau du puits et de ses proches abords ($|x| < 40$ m). Alors que le spectre du modèle calculé pour des CP exactes prend des valeurs allant jusqu'à 100 Hz, celui du modèle calculé ici est limité à 80 Hz. En revanche, pour des profils sismiques plus lointains (à partir de 100 m) cette tendance s'estompe, et la bande passante associée à chacun des spectres est assez comparable.

Ainsi une perte de résolution au niveau du puits est enregistrée par rapport à l'exemple précédent. Toutefois, la résolution globale reste excellente, et le gain obtenu par rapport à une méthode d'inversion classique est sans équivoque. La prise en compte des non-linéarités dans l'inversion permet donc la construction d'un résultat haute fréquence dans sa dimension verticale, et ce dans le cas d'application concret où seules les données sismiques sont utilisées (les CP sont calculées en même temps que le modèle d'impédance).

Il convient maintenant de s'intéresser à la résolution latérale de la méthode, dans ce contexte d'application.

4.4.7 Résolution latérale

La partie précédente nous a montré que l'information sismique, dans le cas idéal où les CP sont exactes, ne suffit pas à la bonne détermination des variations latérales du modèle dans la zone profonde. La résolution latérale de la méthode dans cette zone est de l'ordre de la trentaine de mètres. L'apport ici d'une information *a priori* sur la structure latérale du milieu doit permettre de lutter contre cette indétermination. En revanche, cette information *a priori* peut entraîner une perte de résolution, car elle est par nature inexacte, et peut rentrer en conflit avec l'information sismique. Il convient donc d'analyser ici l'influence de la régularisation et de la recherche des CP optimales sur la résolution latérale.

4.4.7.1 Analyse de profils sismiques latéraux

Les profils sismiques latéraux extraits du modèle d'impédance calculé sont présentés par les figures 4.32 et 4.33.

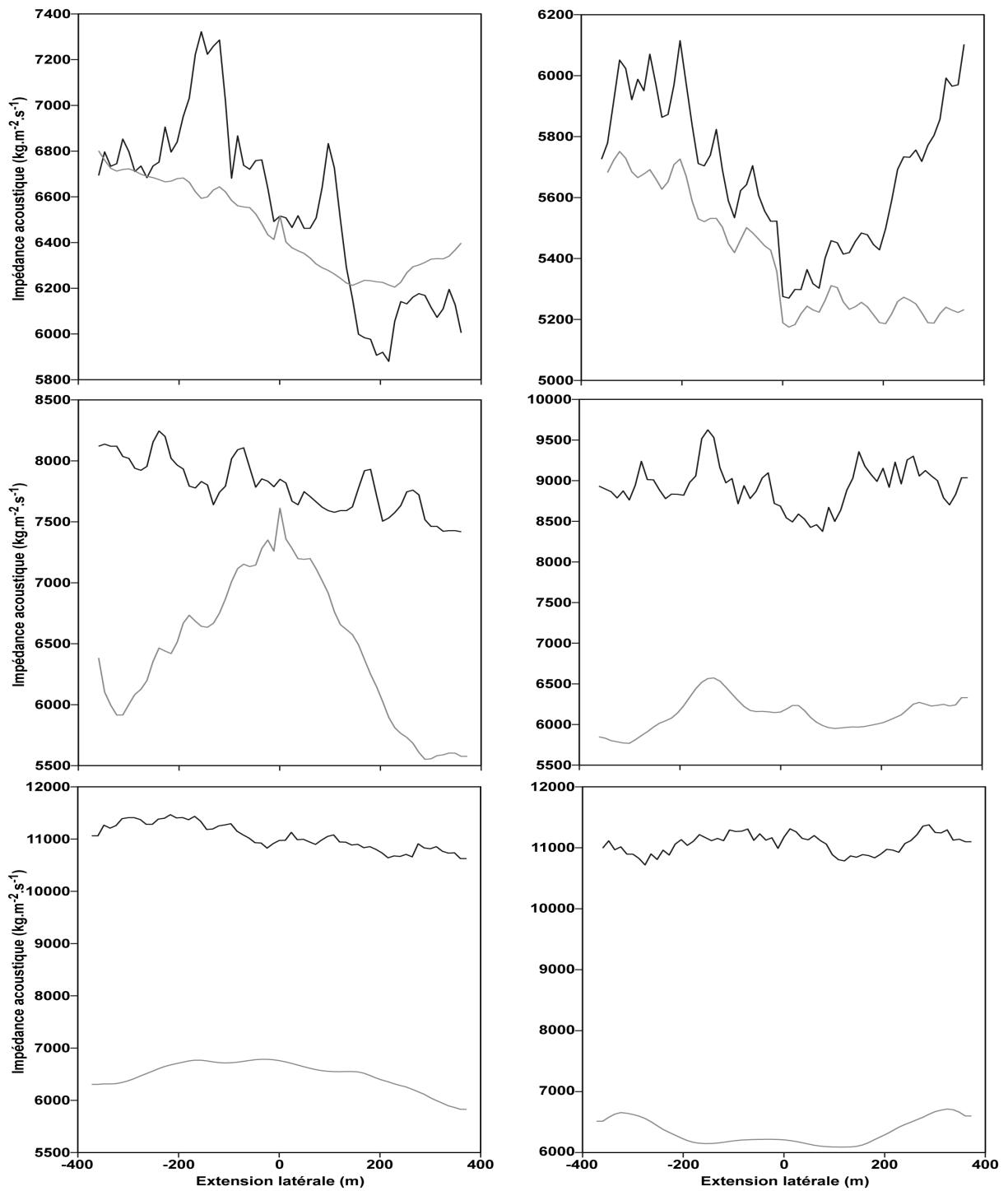


FIG. 4.32 – Profils sismiques latéraux exact (noir) et calculé (gris). De gauche à droite et de haut en bas, $z = 1000 \text{ m}$, $z = 1400 \text{ m}$, $z = 1800 \text{ m}$, $z = 2000 \text{ m}$, $z = 2100 \text{ m}$, $z = 2200 \text{ m}$.

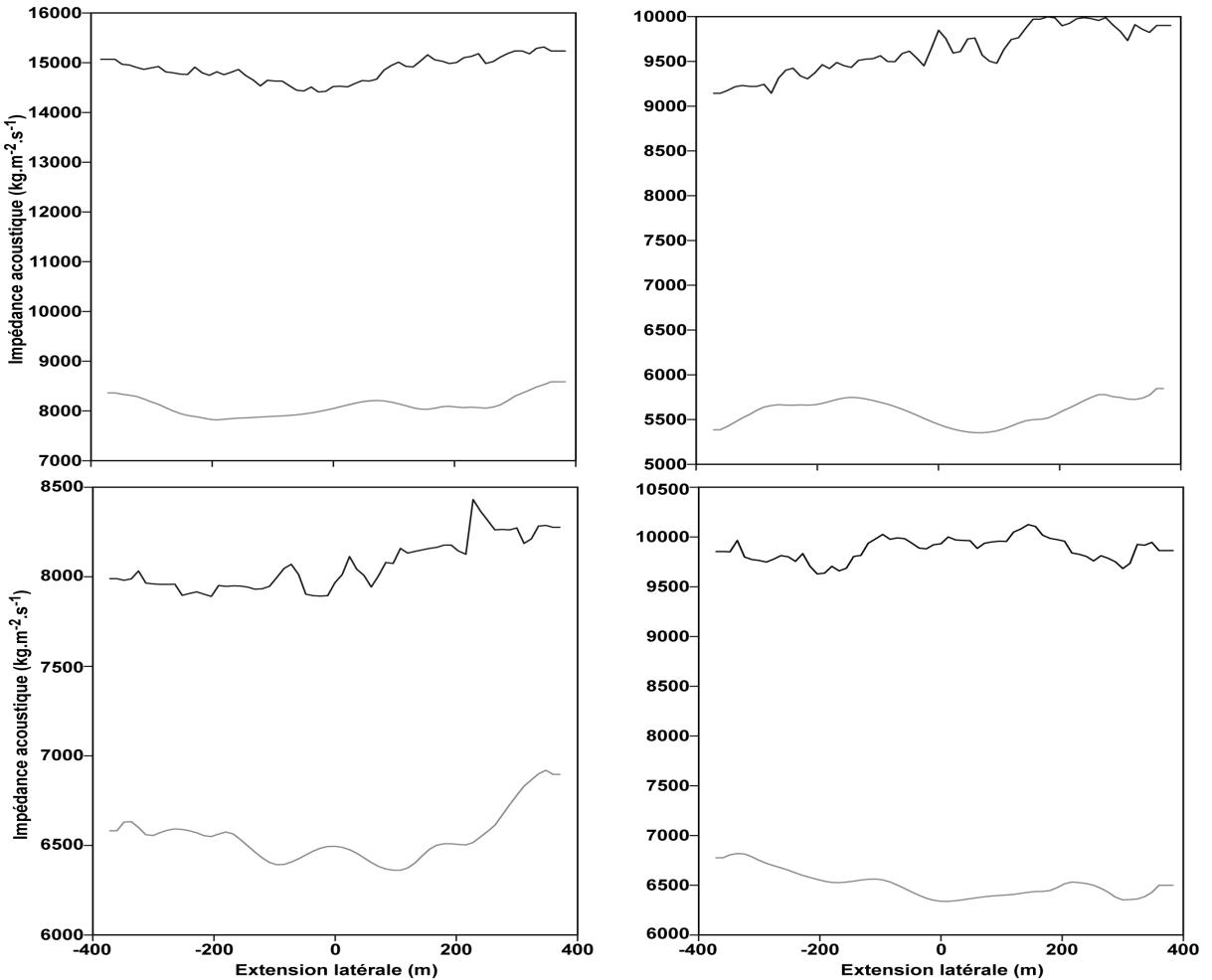


FIG. 4.33 – Profils sismiques latéraux exact (noir) et calculé (gris). De gauche à droite et de haut en bas, $z = 2300$ m, $z = 2400$ m, $z = 2500$ m, $z = 3000$ m.

Tout d'abord, nous nous concentrons sur la partie supérieure ($z < 1800$ m). L'effet de lissage induit par la régularisation forte est particulièrement visible tout en haut du modèle (fig 4.32). Contrairement au cas précédent où une forte instabilité était visible pour ces paramètres, cette instabilité est ici maîtrisée, et le modèle calculé suit la tendance basse fréquence du modèle exact. Plus bas ($z = 1400$ m), le poids de régularisation baisse et le modèle exact est très bien retrouvé. En revanche, au niveau du dernier récepteur ($z = 1800$ m), la zone d'instabilité demeure et le profil sismique latéral calculé semble totalement décorrélé du profil sismique latéral exact. Cependant, d'une manière générale, les résultats indiquant que le modèle d'impédance est bien reconstruit dans la zone supérieure sont confirmés.

Dans la zone profonde, la situation est plus complexe. Le poids de régularisation sur les variations latérales ne dépend plus de la profondeur, il est maintenant con-

stant. Il semble cependant que son effet soit très important. En effet, les profils sismiques latéraux du modèle calculé montrent un aspect plus lisse que pour l'inversion à CP connues. De plus, il semble que ces profils respectent la tendance basse fréquence des profils exactes, et ce jusqu'au moins la profondeur de 2500 m. Ainsi, l'introduction d'information *a priori* permet de juguler l'indétermination, tout en tirant profit de l'information sismique. En effet, même si les variations à fine échelle ne sont pas retrouvées, les variations principales tendent à l'être, si l'on excepte les parties très profondes du modèle ($z > 2500$ m) (fig 4.33).

Ces résultats ne donnent cependant aucune information sur la résolution maximale latérale. Pour essayer d'en obtenir une évaluation, nous reprenons l'expérience précédente consistant à analyser le comportement de la méthode face à une perturbation latérale localisée du modèle d'impédance.

4.4.7.2 Détection d'une perturbation latérale

Comme précédemment, l'objectif de cette partie est de donner une évaluation de la Point Spread Function de la méthode d'imagerie. Afin de pouvoir comparer les résultats, les perturbations ajoutées au modèle d'impédance exact sont les mêmes que celles qui ont été utilisées dans la partie précédente, c'est-à-dire des perturbations sur une maille, d'amplitude égale à 500% de l'amplitude locale, situées à $x = 144$ m du puits et à 2200 m et 2600 m de profondeur. L'étalement de la perturbation dans l'image reconstruite donne une évaluation de la résolution latérale maximale de la méthode. Celle-ci devrait être en principe dégradée par rapport au cas précédent, de par l'utilisation d'un terme de régularisation qui lisse les variations latérales.

Dans un premier temps, il faut s'assurer de la bonne convergence de la méthode. Les critères d'arrêt de l'algorithme de minimisation étant les mêmes, l'optimisation prend fin dans les deux cas à l'issue de 1000 itérations de l'algorithme Quasi-Newton, avec une valeur du coût total normalisé égale à 2.33×10^{-4} , identique au cas de l'inversion des données non perturbées. Les valeurs détaillées du coûts sismique et du coût de régularisation initiaux, sont, dans les deux cas,

$$s_0 = 4.27 \times 10^{-8}, \quad r_0 = 2.38 \times 10^{-12}.$$

Les coûts sismique et de régularisation finaux sont, pour la perturbation située en $x = 144$ m et $z = 2200$ m,

$$s_f = 4.38 \times 10^{-12}, \quad r_f = 5.61 \times 10^{-12},$$

et pour la perturbation située en $x = 144$ m et $z = 2600$ m,

$$s_f = 4.36 \times 10^{-12}, \quad r_f = 5.60 \times 10^{-12}.$$

Ces résultats sont très similaires à ceux obtenus pour l'inversion des données non perturbées. De fait, les valeurs des poids de régularisation sont les mêmes. Il permettent une prise en compte adéquate de l'information sismique et de l'information de régularisation. Ceci permet de plus de pouvoir comparer les résultats obtenus ici sur des modèles perturbés aux résultats obtenus sur le modèle non perturbé.

Les profils des CP calculées sont également très proches de ceux qui ont été établis par l'inversion des données non perturbées (fig.4.24). Leur exposition n'est donc pas nécessaire, de fait, elle n'apporte aucune information supplémentaire.

Il convient donc de s'intéresser au modèle d'impédance calculé. Les modèles perturbés et les modèles calculés à partir des données perturbées sont donnés figure 4.34.

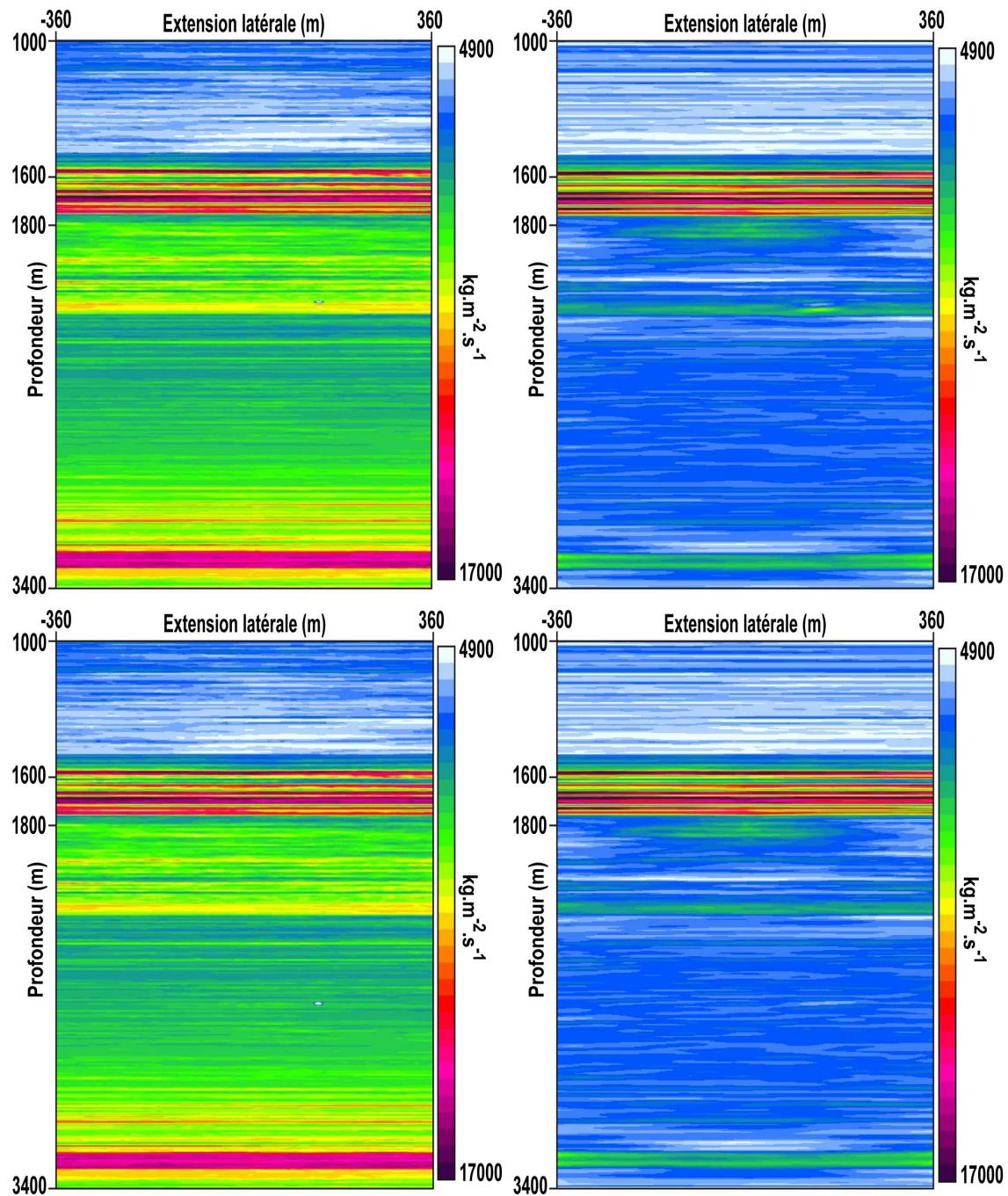


FIG. 4.34 – Modèle d'impédance après perturbation en $x = 144$ m et $z = 2200$ m (haut gauche), modèle calculé à partir des données perturbées correspondantes (haut droite). Modèle d'impédance après perturbation en $x = 144$ m et $z = 2600$ m (bas gauche), modèle calculé à partir des données perturbées correspondantes (bas droite)

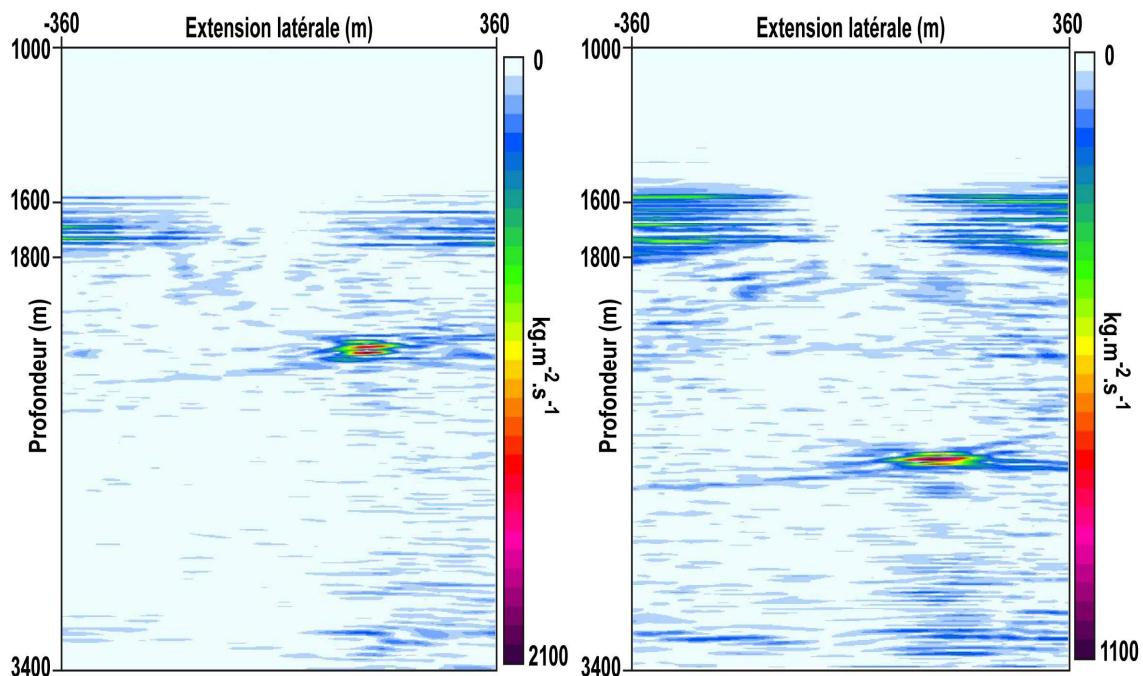


FIG. 4.35 – Différences entre modèles issus de données perturbées (fig.4.34 et modèle issu de données non perturbées (fig.4.26), perturbation en $x = 144$ m et $z = 2200$ m (gauche), perturbation en $x = 144$ m et $z = 2600$ m (droite).

La visualisation des résultats montre une bonne localisation de la perturbation d'impédance. Celle-ci est visible sur le modèle calculé (fig.4.34, et apparaît très bien sur la figure 4.35 représentant les différences. Son étalement est un peu plus prononcé que dans la partie précédente, la tache apparaissant au niveau de la perturbation couvrant une zone mesurant à peu près 50 m de large et 40 m de hauteur. Comme cela a été mentionné, cette perte de résolution provient de l'utilisation d'un terme de régularisation visant à lisser les variations latérales, et de la faiblesse de l'information sismique sur les variations latérales dans cette zone. Cependant, la possibilité de retrouver de telles perturbations tout en usant de régularisation prouve que le poids affecté à l'information *a priori* permet d'introduire un bon compromis entre cette information et l'information sismique.

De plus, les instabilités détectées précédemment dans la bande comprise entre 1600 m et 1800 m de profondeur ont été ici éradiquées par l'introduction de la régularisation latérale. Ainsi, les différences ne font apparaître qu'une seule zone

de forte amplitude, située au niveau de la perturbation. Aucune confusion n'est donc possible : l'effet de la perturbation est bien visible, et celle-ci est correctement détectée.

La méthode d'imagerie développée montre donc qu'elle n'est pas insensible à la présence d'hétérogénéités latérales. La finesse de la reconstruction, de l'ordre de plusieurs dizaines de mètres, est voisine de celle obtenue par les méthodes d'imagerie classique, même si elle est dégradée par rapport à l'expérience précédente, pour lesquelles les CP exactes étaient connues.

Pour terminer cette analyse de la résolution latérale, une dernière expérience est menée, pour une perturbation d'amplitude plus réaliste (30%), mais dont la taille est plus importante : 36 m de large pour 24 m de hauteur, située en $x = 144$ m et $z = 2200$ m. Cette taille se situe à la limite de la résolution latérale qui a été évaluée. Elle devrait donc en principe pouvoir être détectée.

Pour cette expérience, les coûts sismiques et de régularisation initiaux sont identiques aux précédents, et les coûts finaux sont tout à fait comparables,

$$s_f = 4.36 \times 10^{-12}, \quad r_f = 5.59 \times 10^{-12}.$$

Il en va de même pour les CP calculées, ce qui dispense à nouveau d'en donner une représentation. Nous nous focalisons donc sur le modèle d'impédance exact correspondant à cette perturbation et le modèle d'impédance calculé, présentés par la figure 4.36.

A première vue, la perturbation, d'amplitude plus faible, ne semble pas détectée, le modèle calculé ressemblant fortement à celui calculé à partir de données non perturbées (fig 4.26). Cependant, il est possible de noter une légère hausse d'amplitude (vert plus clair) à l'endroit où l'on s'attend à retrouver la perturbation. La visualisation de la différence entre le modèle calculé à partir des données perturbées et ceci calculé sans perturbation confirme cette intuition : la perturbation est effectivement détectée.

Ceci confirme l'analyse de la résolution latérale : une perturbation d'amplitude réaliste, de l'ordre de la trentaine de mètres, peut-être détectée par la méthode d'inversion qui a été développée.

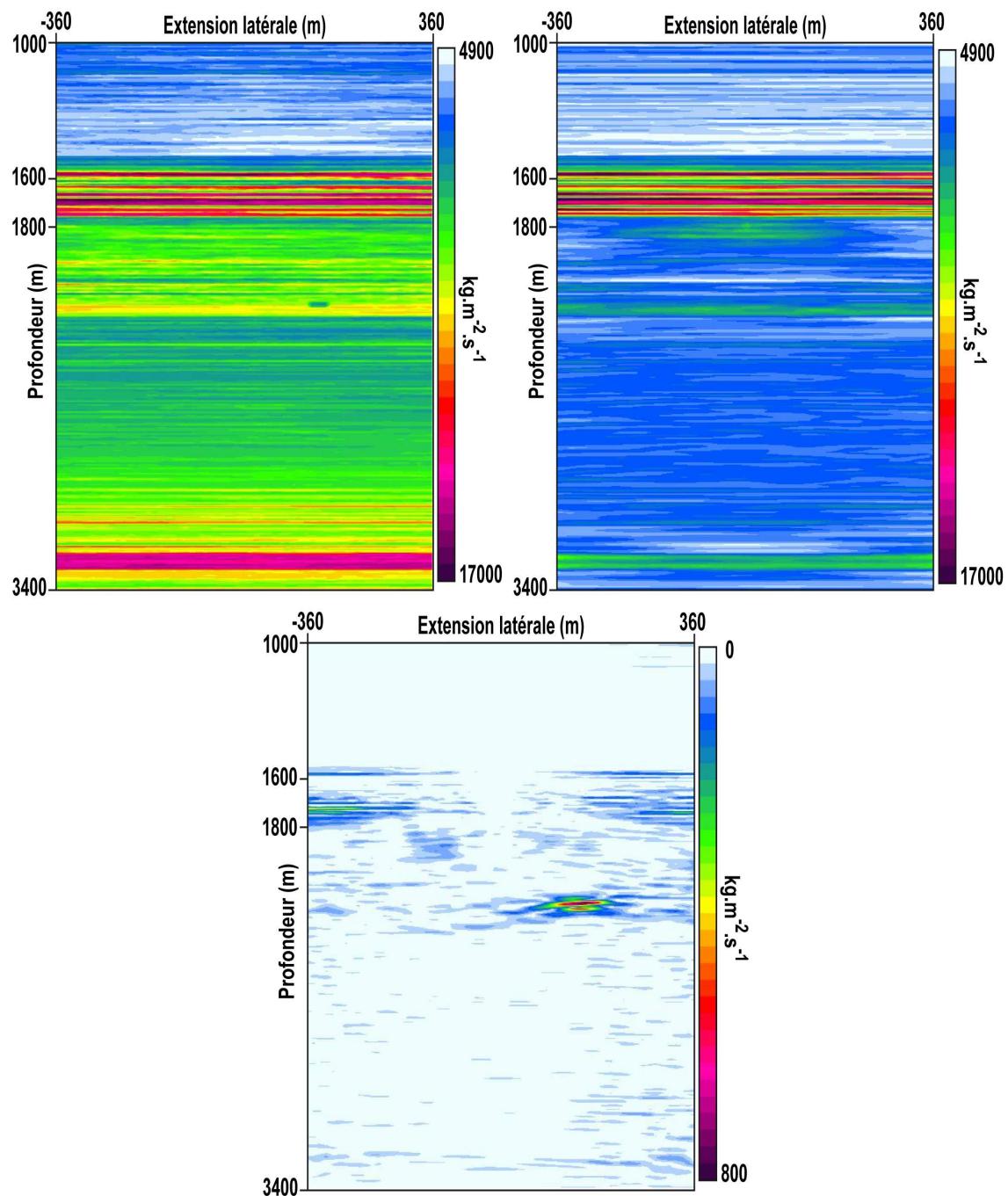


FIG. 4.36 – Modèle d’impédance après perturbation d’amplitude 30% en $x = 144$ m et $z = 2200$ m (gauche), modèle calculé à partir des données perturbées correspondantes (droite), différences avec modèle issu de données non perturbées fig.4.26 (bas).

4.4.8 Conclusion sur l'analyse de la résolution

L'intérêt de la mise en place d'une méthodologie d'inversion non linéaire de données sismiques est la prise en compte, dans les données, des phénomènes de réflexions multiples. Le contexte particulier de l'inversion de données sismiques de puits facilite cette prise en compte par le calcul au cours de l'inversion de la stimulation du milieu au niveau du premier récepteur (les CP), qui permet de s'affranchir de l'évaluation *a priori* des sources servant à générer les données. Dans le cadre de l'imagerie de milieux hétérogènes, pour lesquels ces réflexions constituent une part non négligeable du signal enregistré aux récepteurs, cette prise en compte doit conduire à une amélioration de la résolution de l'image établie. Les résultats présentés ici illustrent ce résultat issue de la théorie de l'inversion de données sismiques.

L'analyse de la résolution de la méthode sur le cas test présenté au début de ce chapitre conduit à définir deux zones du sous-sol pour lesquelles l'inversion montre un comportement très différent. La première d'entre elle est la zone supérieure, qui correspond à la bande comprise entre les profondeurs du premier et du dernier récepteur. La seconde est la zone profonde, qui correspond à la bande débutant à la profondeur du dernier récepteur et s'étendant jusqu'au fond du modèle.

La première zone est retrouvée avec une très grande précision, et ce sur une largeur de presque 300 m, de part et d'autre du puits. La résolution verticale et horizontale y est excellente. Il est probable que celle-ci soit gouvernée par le pas d'échantillonnage des récepteurs. Des expériences numériques dans le cas 1D ont en effet montré que l'emploi d'un grand nombre de récepteurs disposés de manière rapprochée permet d'augmenter la précision du résultat dans cette zone. De nouvelles expériences devraient être menées dans le cadre 2D pour vérifier cette hypothèse.

D'autre part, l'analyse de la résolution verticale montre une amélioration indéniable de la finesse des détails avec laquelle il est possible de calculer une estimation de la distribution d'impédance dans la zone profonde. Des variations verticales inférieures à 10 m peuvent ainsi être retrouvées, pour un éclairage du milieu par des ondes dont la longueur moyenne avoisine 140 m, dans l'exemple pris comme cas test et présenté dans ce chapitre. Cette observation est confirmée par l'analyse du spectre des profils sismiques verticaux, qui montre un enrichissement important par rapport aux spectres de profils obtenus par l'application d'une méthode d'imagerie linéaire classique. Cette amélioration de la résolution n'est plus liée ici à l'échantillonage des récepteurs, mais bien à la prise en compte des réflexions multiples.

La détermination des variations latérales dans la zone profonde est quant à elle soumise à une indétermination assez forte, qui est cependant contrôlée par l'ajout d'information *a priori*. Dans le cas test présenté, la résolution latérale est de l'ordre de 50 m, ce qui est à peu près équivalent à celle obtenue par les méthodes d'inversion linéaire pour l'éclairage du milieu utilisé. De plus, la résolution latérale peut être en

principe augmentée par l'utilisation d'un plus grand nombre de jeux données, et un élargissement de la gamme d'angle choisie pour éclairer le milieu. Dans le cas test présenté, les données sismiques correspondent en effet à la propagation de quasi-ondes planes se propageant dans des directions proches de la verticale du milieu, ce qui ne fournit qu'un faible éclairage des variations latérales du milieu.

L'investigation de la résolution de la méthode ayant été menée, il convient maintenant de s'intéresser à quelques tests supplémentaires évaluant la sensibilité de la méthode à une erreur introduite sur le modèle de vitesse, et la présence de bruit sur les données.

4.5 Analyses de sensibilité

La méthode a pour l'instant été testée dans un contexte idéal, pour lequel le modèle de vitesse est connu de manière exacte, et les données sont synthétiques, c'est-à-dire qu'elles correspondent (presque) parfaitement à l'opérateur de modélisation utilisé pour résoudre le problème direct.

Remarque. *Le seul biais introduit provient en réalité de l'interpolation des CP. Celle-ci induit que le sismogramme synthétique associé au modèle d'impédance exact et aux CP interpolées à partir des traces exactes, via la résolution du problème direct, ne correspond pas exactement au données utilisées. L'amplitude de ce biais est cependant largement en-deça des résidus obtenus à l'issue de l'inversion.*

Cette partie est consacrée à l'analyse du comportement de la méthode d'inversion lorsque ces deux hypothèses sont affaiblies. Dans un premier temps, trois expériences sont réalisées en introduisant des erreurs sur l'estimation du modèle de vitesse. Ensuite, deux expériences sont menées en introduisant un bruit coloré sur les données. D'une manière générale, ces tests sont loin d'être exhaustifs. Ils sont proposés comme une première approche du travail à effectuer dans une optique de validation de la méthode pour une utilisation sur des cas plus concrets, et permettent de mettre en lumière des paramètres auxquels la méthode est très sensible.

4.5.1 Analyse de sensibilité par rapport au modèle de vitesse

Les expériences précédentes ont été réalisées pour un champ de vitesse exact. Dans la pratique, on ne dispose que d'une estimation, plus ou moins précise, de celui-ci. Cette estimation est ensuite fixée au cours de l'inversion, elle n'est pas mise à jour en fonction de l'information sismique. Elle peut être établie au niveau des récepteurs de manière assez fiable, par l'analyse des temps d'arrivée de l'onde principale dans les données suivant la position du récepteur, et de manière générale dans l'axe du puits, en $x = 0$. Comme pour le modèle d'impédance, une estimation 1D du modèle de vitesse dans l'axe du puits est donc accessible, mais celle de ses variations latérales est beaucoup plus difficile. Trois expériences sont donc menées, avec des modèles de vitesse 1D. Le contexte d'application reste identique à celui qui a été présenté dans la partie précédente. L'inversion est menée pour des CP inconnues, sur les données issues du cas test présenté dans la première partie de ce chapitre. Les poids de régularisation utilisés sont également identiques à ceux qui ont été fixés dans l'exemple précédent. La seule différence introduite concerne donc le modèle de vitesse utilisé pour l'inversion.

La première expérience est réalisée, avec un modèle de vitesse 1D, calculé comme l'extension du modèle de vitesse exact pris à l'abscisse $x = 0$, dans l'axe du puits.

Ce modèle 1D présente l'avantage de posséder les valeurs de vitesse exactes au niveau des récepteurs, dont on peut supposer qu'elles jouent un rôle important dans l'évaluation du résultat. A l'issue de 1000 itérations de la boucle principale, la fonctionnelle normalisée est égale à 8.79×10^{-4} et le coût sismique s_f , et le coût de régularisation r_f , valent respectivement 2.03×10^{-11} et 1.71×10^{-11} .

Le modèle et les CP calculés expliquent les données avec moins de précision que dans le cas où le modèle de vitesse utilisé est exact. Ceci est conforme aux attentes : les données ne peuvent plus être expliquées avec la même précision, un bruit de modélisation étant introduit par l'utilisation dans l'inversion d'un modèle de vitesse inexact. Néanmoins, les résidus sont encore très faibles, ceux-ci étant divisés par un facteur 30. Une représentation du modèle d'impédance calculé ainsi que de sa différence avec le modèle exact est donnée figure 4.37.

Le résultat proposé est assez voisin des résultats obtenus en utilisant le modèle de vitesse exact. Il possède la régularité latérale désirée. Une légère perte de résolution verticale peut cependant être remarquée. La comparaison de profils sismiques verticaux confirme cette impression (fig. 4.38) : les variations d'impédance sont retrouvées avec une précision inférieure à 10 m, mais les amplitudes des variations du modèle calculé sont plus faibles. Cependant, l'estimation globale du modèle reste assez bonne. Ceci peut provenir de l'approximation d'un éclairage par ondes planes se propageant à incidence proche de la verticale, qui implique que les variations latérales du modèle de vitesse n'ont pas une importance primordiale dans le résultat.

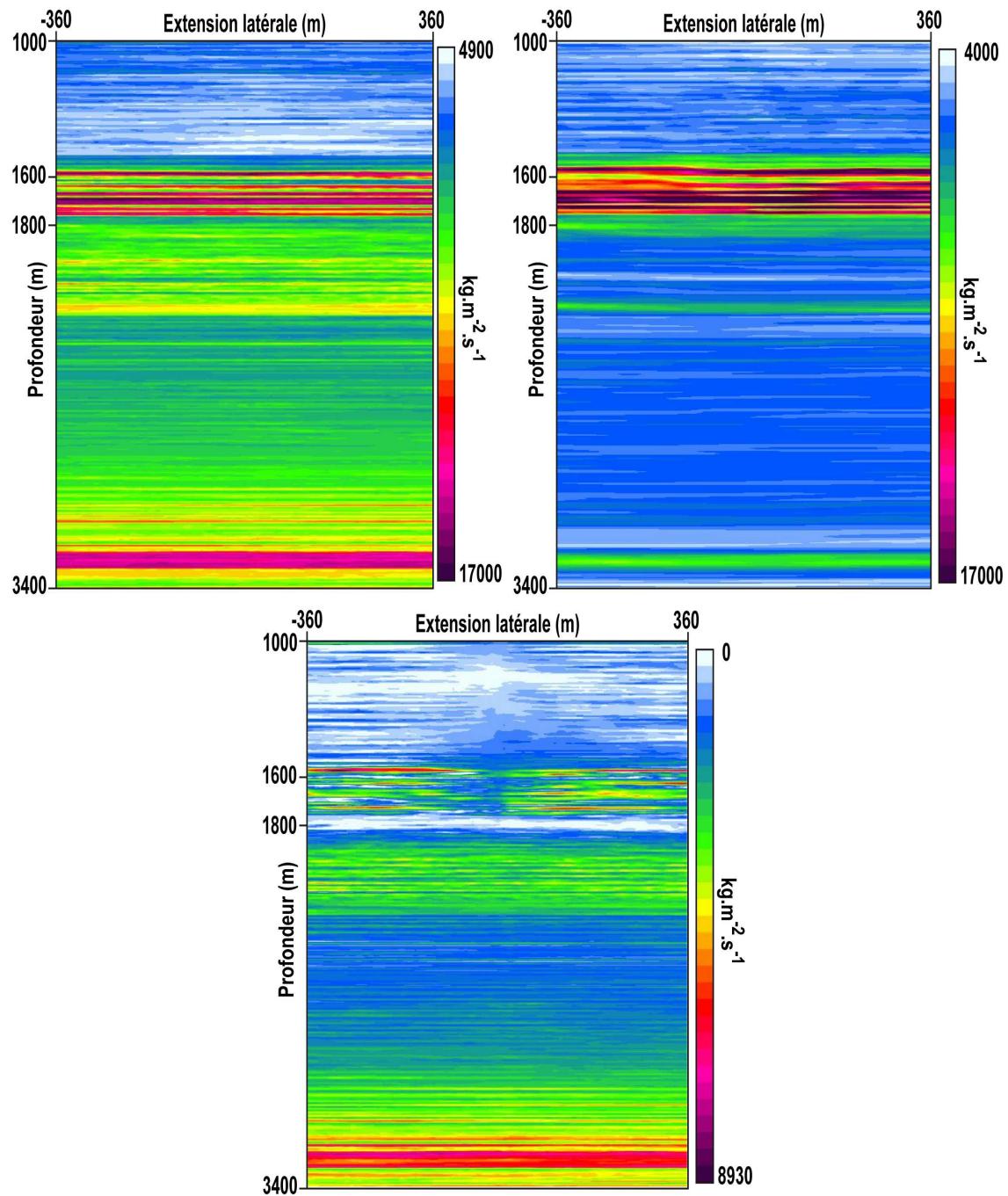


FIG. 4.37 – Impédance exacte (haut gauche), impédance calculée par inversion en utilisant un modèle de vitesse 1D extrapolé à partir du modèle de vitesse en $x = 0$ (haut droite), différence (bas).

Une deuxième expérience est menée. Pour celle-ci, le modèle de vitesse choisi est identique au précédent pour $z < 1800$ m (soit 1D construit par extension des valeurs

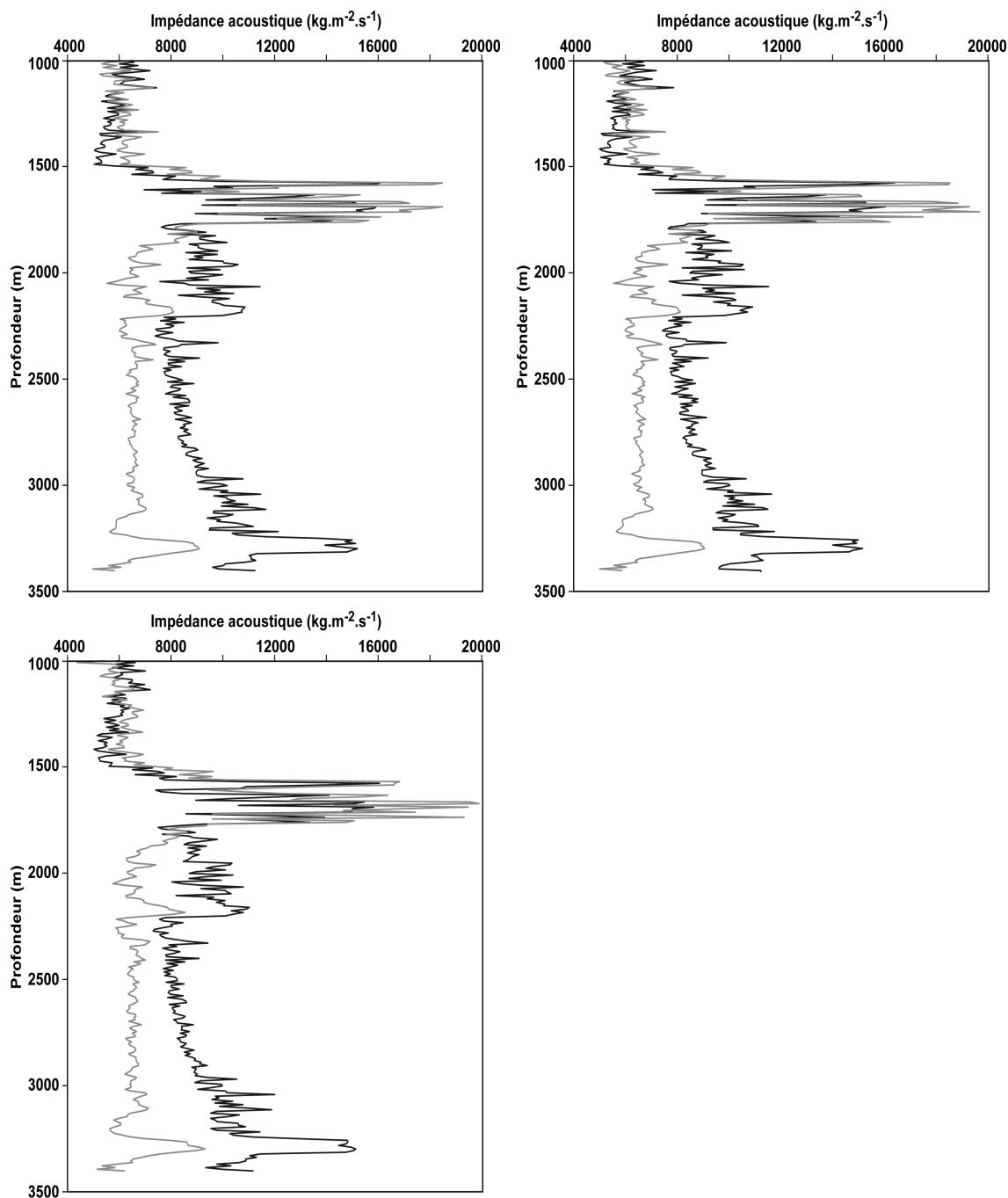


FIG. 4.38 – Profils sismiques verticaux du modèle d’impédance calculé en utilisant un modèle de vitesse 1D extrapolé (gris), et du modèle exact (noir). De droite à gauche, et de haut en bas, $x = 0 \text{ m}$, $x = 40 \text{ m}$, $x = 200 \text{ m}$.

prises au niveau des récepteurs dans la zone supérieure), et homogène pour $z > 1800 \text{ m}$, la valeur prise étant égale à la valeur du modèle de vitesse exact au niveau du récepteur le plus profond. Ce modèle de vitesse est représenté figure 4.39.

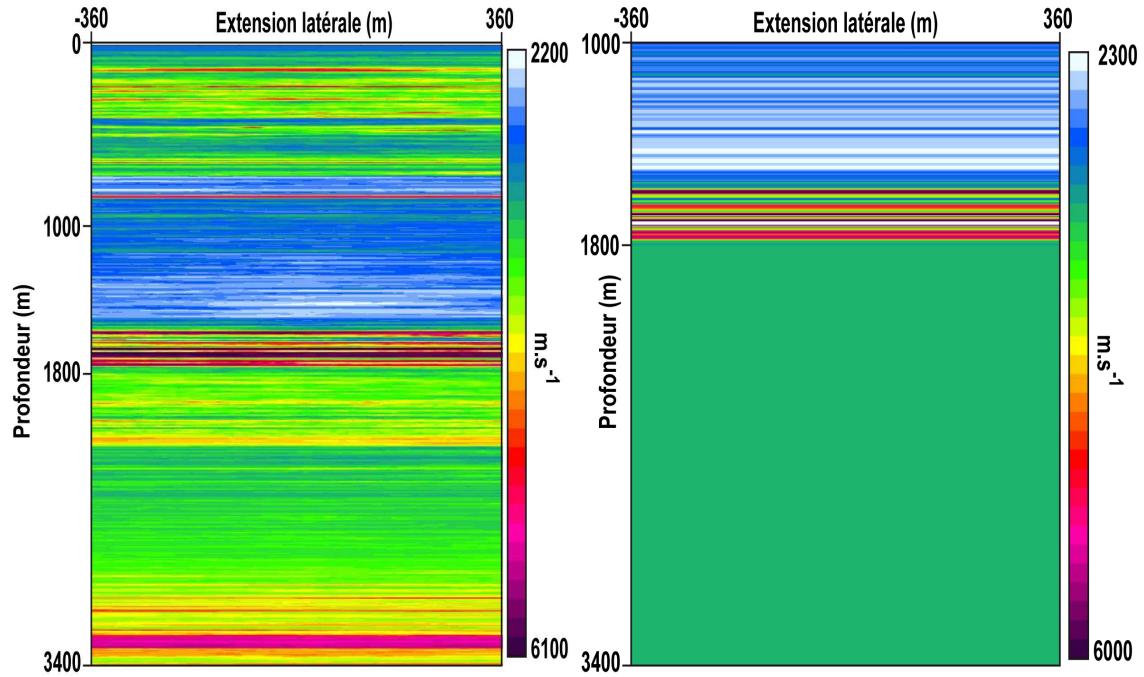


FIG. 4.39 – Modèle de vitesse exact (gauche) et modèle de vitesse 1D en zone supérieure et homogène en zone inférieure (droite).

Pour cette estimation du modèle de vitesse, la fonctionnelle normalisée atteint la valeur de 7.00×10^{-4} à l'issue de 1000 itérations de la boucle principale. Le coût sismique s_f vaut quant à lui 1.65×10^{-11} et le coût de régularisation r_f vaut 1.34×10^{-11} . Ces valeurs sont du même ordre de grandeur que pour le cas précédent, et sont même légèrement inférieures. Le résultat d'impédance correspondant, ainsi que sa différence avec le modèle d'impédance exact est présenté figure 4.40.

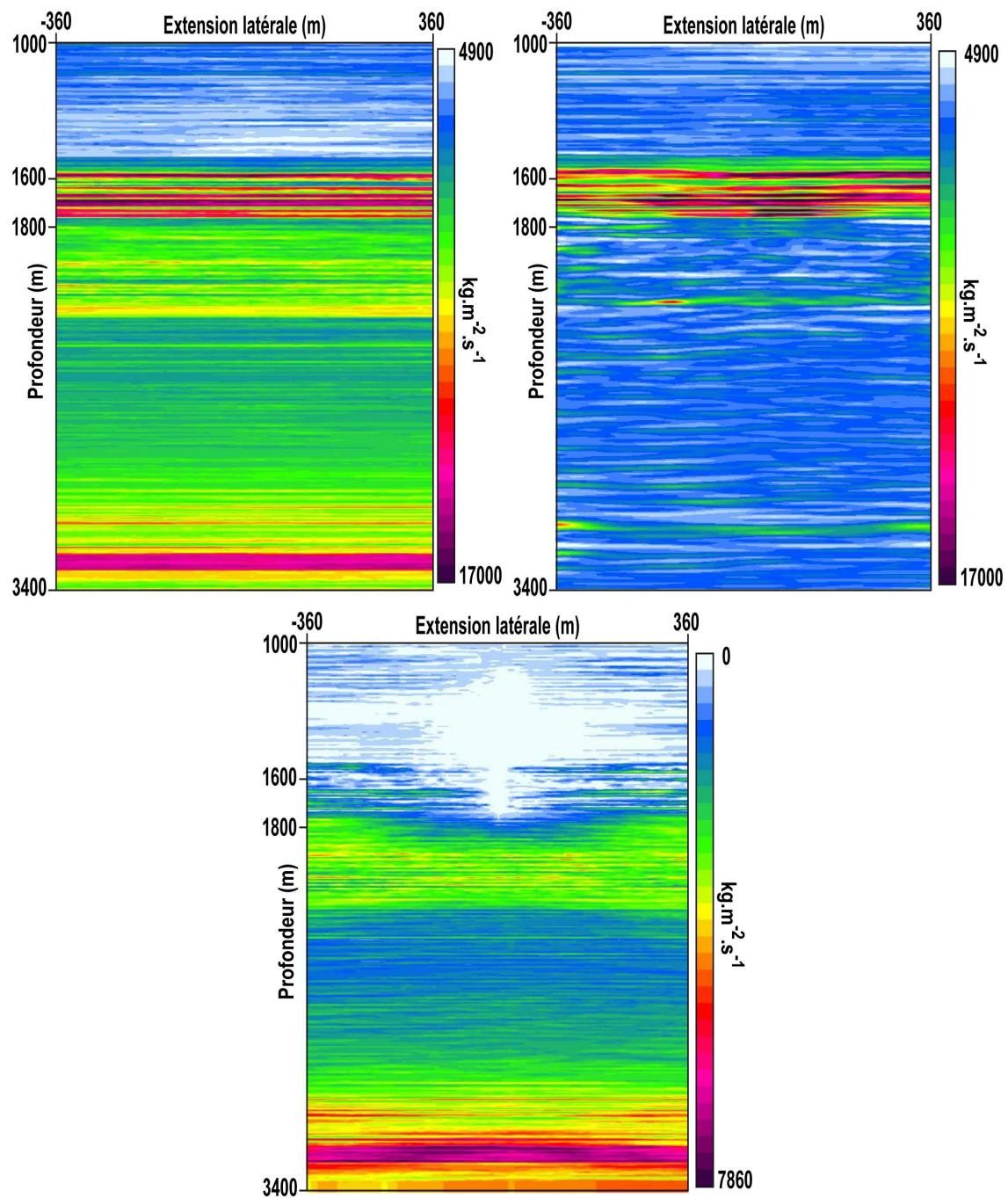


FIG. 4.40 – Impédance exacte (haut gauche), impédance calculée par inversion en utilisant un modèle de vitesse 1D extrapolé à partir des valeurs de vitesse au niveau des récepteurs dans la zone supérieure et homogène dans la zone inférieure (haut droite), différence (bas).

Le résultat obtenu est comparable au résultat précédent dans la zone supérieure, et paraît plus instable dans la zone inférieure. C'est là l'influence du modèle de vitesse qui est homogène dans cette zone pour le cas traité ici. La régularité en zone profonde pourrait être améliorée en renforçant le poids de l'information *a priori*. La comparaison de profils sismiques verticaux (fig.4.41) montre également l'influence de la partie homogène du modèle de vitesse : les variations sont retrouvées avec précision, mais subissent un décalage. En effet, les temps d'arrivée des ondes au niveau des récepteurs sont faussés par le modèle de vitesse, ainsi la distance des discontinuités calculées par inversion est elle-même faussée. En quelque sorte, une image en temps du modèle d'impédance est donc donnée, qui peut-être modifiée *a posteriori* par une information sur la vitesse en zone profonde. Si cette information est disponible, il est également possible de l'inclure directement dans l'inversion, ce qui correspond au test de sensibilité précédent.

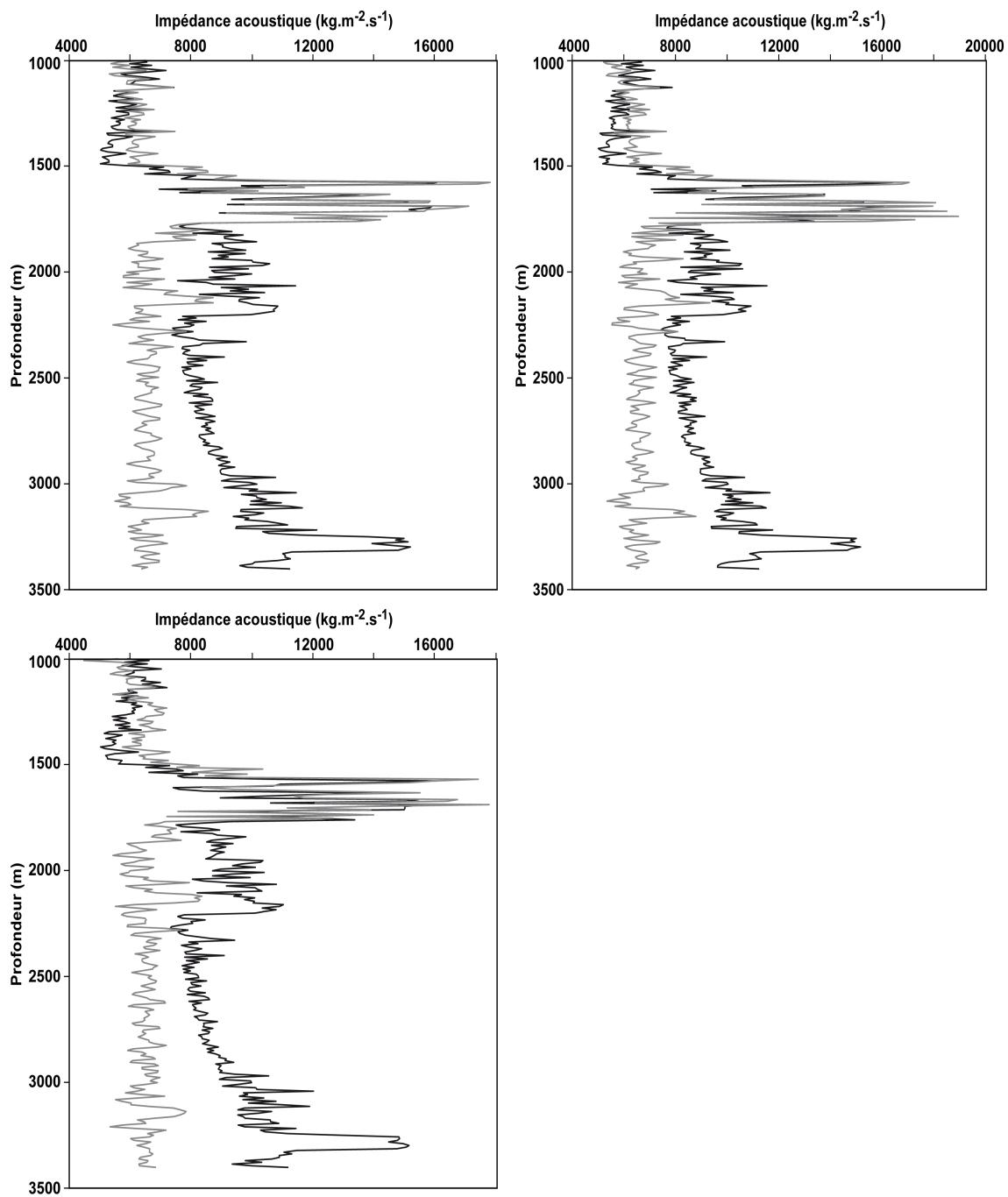


FIG. 4.41 – Profils sismiques verticaux du modèle d’impédance calculé en utilisant un modèle de vitesse 1D extrapolé dans la zone supérieure et homogène dans la zone inférieure (gris), et du modèle exact (noir). De gauche à droite et de haut en bas, $x = 0 \text{ m}$, $x = 40 \text{ m}$, $x = 200 \text{ m}$.

Ces deux premières expériences tendent à montrer que la méthode d’inversion développée ne montre pas une extrême sensibilité aux variations latérales du modèle

de vitesse. En revanche, il est possible de supposer que les valeurs importantes sont plutôt celles du champ de vitesse au niveau du puits. En effet, ces valeurs ont vraisemblablement une grande influence sur le signal enregistré par les récepteurs. Une troisième expérience est donc mise en place. Pour celle-ci, le modèle de vitesse utilisé est un modèle 1D, calculé en réalisant la moyenne arithmétique des valeurs discrètes de vitesse de chaque ligne. Ainsi, on construit un modèle de vitesse 1D c_{1D} tel que

$$(c_{1D})_j = \frac{1}{K - 2K' + 2} \sum_{i=K'}^{K-K'+1} c_{ij}$$

Ce modèle de vitesse est présenté figure 4.42.

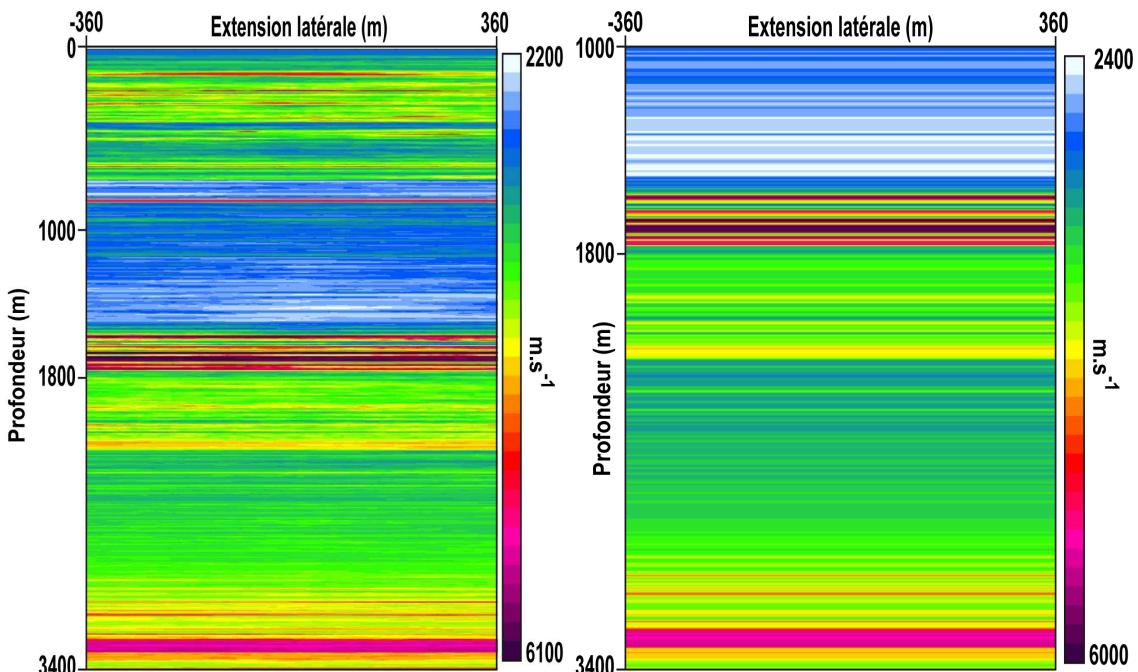


FIG. 4.42 – Modèle de vitesse exact (gauche) et modèle de vitesse 1D moyen (droite).

L'intérêt d'utiliser un modèle moyenné réside dans l'introduction d'une erreur sur les valeurs de vitesse au niveau des récepteurs. Le biais introduit sur ces valeurs est présenté figure 4.43.

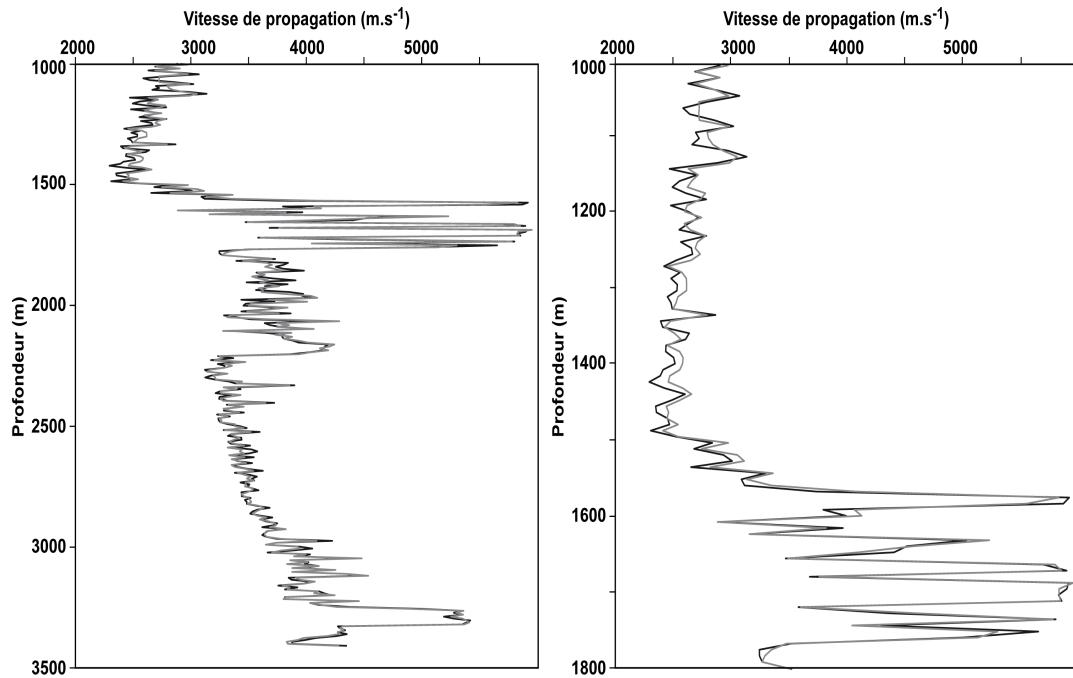


FIG. 4.43 – Profils du modèle de vitesse exact (noire) et du modèle de vitesse 1D moyen (gris) au niveau du puits (gauche), zoom sur la zone récepteur (droite).

A l'issue de 1000 itérations, la fonctionnelle normalisée vaut 3.05×10^{-3} . Le coût sismique s_f vaut pour sa part 8×10^{-11} et le coût de régularisation r_f vaut 4.93×10^{-11} . Un ordre de grandeur dans la précision de l'explication des données sismiques est donc perdu, par rapport au cas où le modèle de vitesse est exact. L'influence de l'erreur du modèle de vitesse sur le résultat semble ici largement plus importante que précédemment. Ceci est confirmé par l'analyse du modèle d'impédance calculé, ainsi que de la différence entre ce modèle et le modèle exact, présentée en figure 4.44.

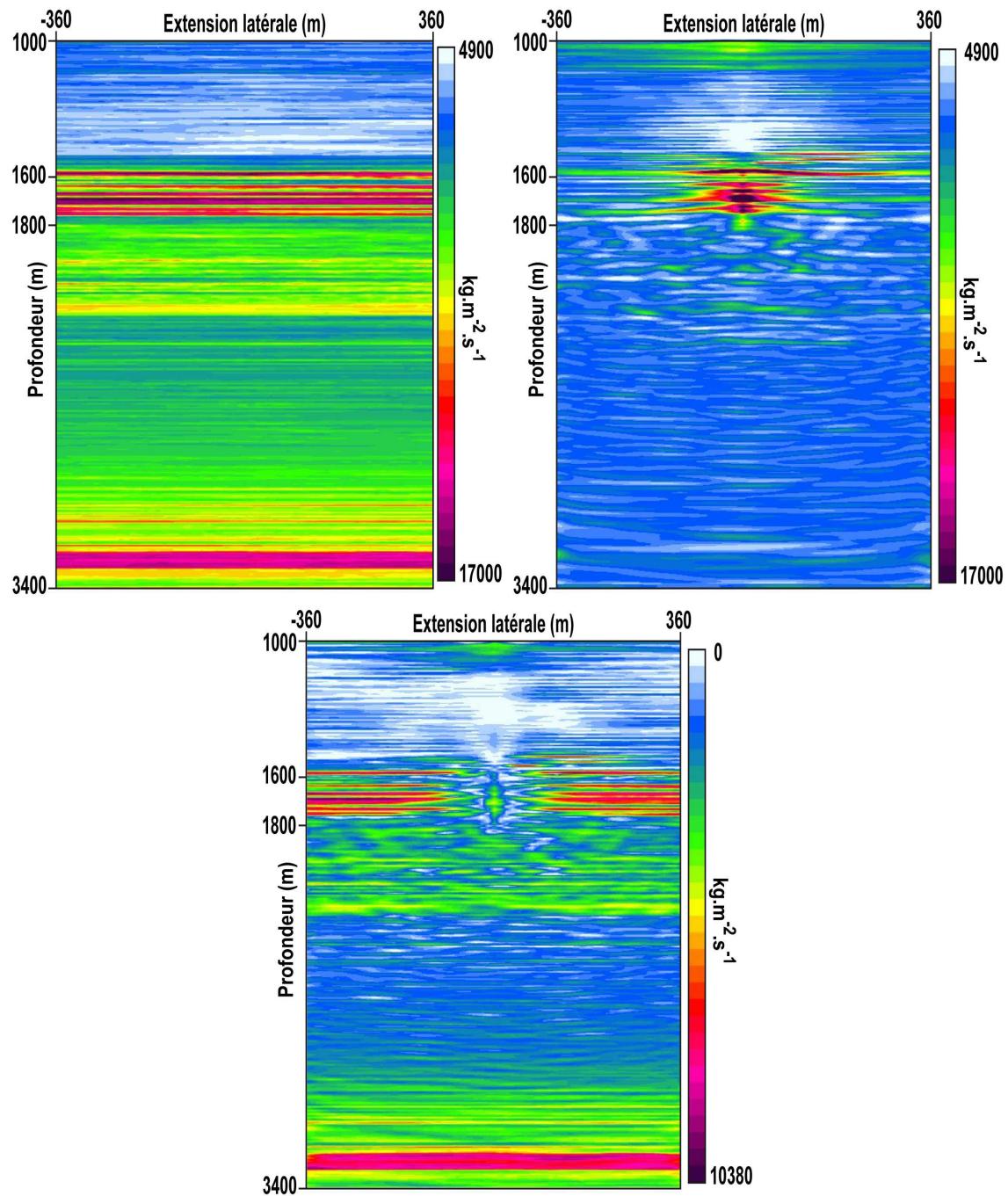


FIG. 4.44 – Impédance exacte (haut gauche), impédance calculée par inversion en utilisant un modèle de vitesse 1D moyen (haut droite), différence (bas).

La perturbation enregistrée sur le résultat d'impédance est en effet importante. Le modèle proposé est très peu régulier, et la zone reconstruite de manière satisfaisante est localisée au niveau du puits. La dégradation du résultat est donc très forte, ce qui indique une forte sensibilité dans l'estimation du modèle de vitesse au niveau des récepteurs.

Ces trois tests de sensibilité par rapport au modèle de vitesse indiquent donc l'importance primordiale d'une estimation précise de ce modèle au niveau des récepteurs. Des perturbations de l'ordre de 100 m.s^{-1} suffisent à fausser totalement le résultat. En revanche, la sensibilité de la méthode au modèle de vitesse semble être concentrée sur ces valeurs. La donnée d'un modèle 1D extrapolé à partir de ces valeurs dans la zone supérieure, et homogène dans la zone inférieure, mène à une bonne estimation de l'impédance dans la zone supérieure, et à une estimation juste mais décalée en profondeur dans la zone inférieure. La méthode ne semble donc pas très sensible à la perte d'information sur les variations latérales du modèle de vitesse. Cependant, ceci est vraisemblablement lié au cas test défini, pour lequel, d'une part, les variations latérales sont faibles et, d'autre part, une information *a priori* sur la régularité latérale est introduite. D'autre part, comme cela a été mentionné, les données sismiques correspondent à un éclairage du milieu par des quasi-ondes planes se propageant dans des directions proches de la verticale pour lesquelles les variations latérales du modèle de vitesse ont peu d'influence. Sur cette base, des tests plus précis pourraient donc être menés pour évaluer l'influence d'une estimation précise 2D du modèle de vitesse sur la résolution latérale du résultat d'impédance.

Nous nous intéressons maintenant à la sensibilité de la méthode à la présence de bruit sur les données sismiques.

4.5.2 Analyse de sensibilité au bruit sur les données

Les données bruitées utilisées dans cette partie sont constituées par l'ajout d'un bruit coloré aux données synthétiques construites dans la première partie de ce chapitre. Ce bruit coloré est déterminé par la convolution d'un bruit blanc gaussien à moyenne nulle, avec le signal de Ricker utilisé pour les sources. Son amplitude est fixée à 50% de la norme L^2 en temps des données synthétiques. La figure 4.45 présente deux exemples de ces données.

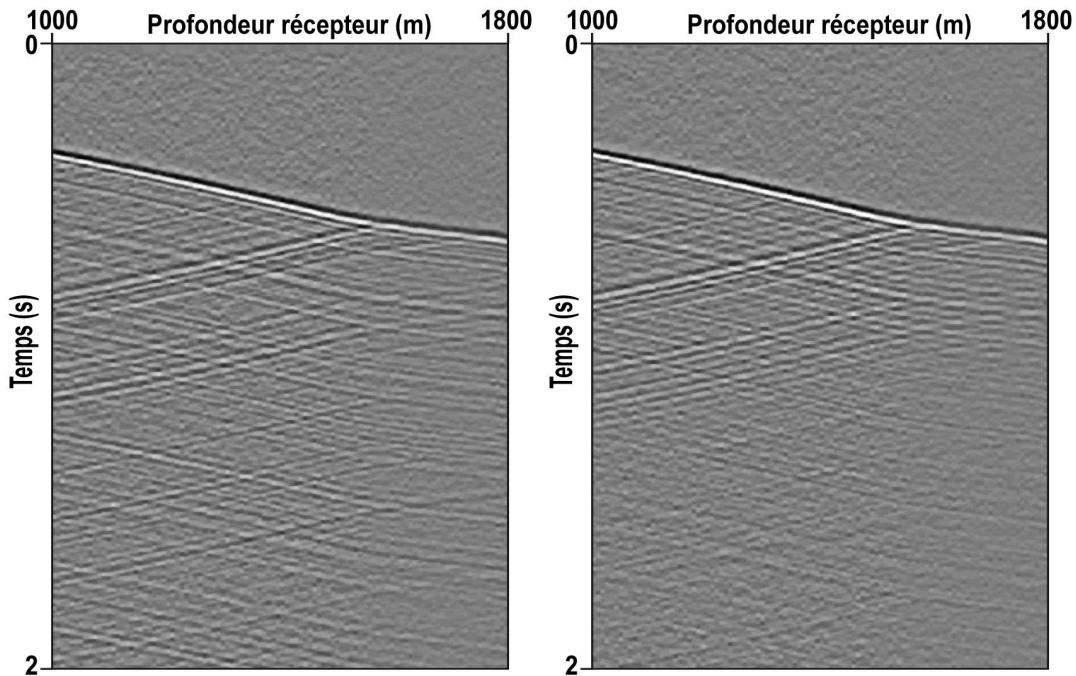


FIG. 4.45 – Données de vitesse verticale bruitées, recombinées par angle déclairage. A gauche, 0° , à droite 6° .

Un premier test est réalisé dans la configuration utilisée dans la partie précédente, soit une inversion pour des CP inconnues et modèle de vitesse exact, avec les mêmes poids de régularisation. A l'issue de 1000 itérations de la boucle d'optimisation principale, la valeur de la fonctionnelle normalisée est égale à 1.84×10^{-1} , ce qui est beaucoup plus élevé que dans les cas précédents. Néanmoins ceci s'explique par le fait que la modélisation de la propagation des ondes est incapable d'expliquer le bruit qui a été introduit dans les données. Le coût sismique reste donc important. Avant l'optimisation, les coûts sismique et de régularisation sont

$$s_0 = 5.34 \times 10^{-8}, \quad r_0 = 2.38 \times 10^{-12}.$$

Au terme des 1000 itérations, ils prennent les valeurs

$$s_f = 9.76 \times 10^{-9}, \quad r_f = 6.23 \times 10^{-11}.$$

La visualisation des résidus fournit une information complémentaire importante (fig. 4.46).

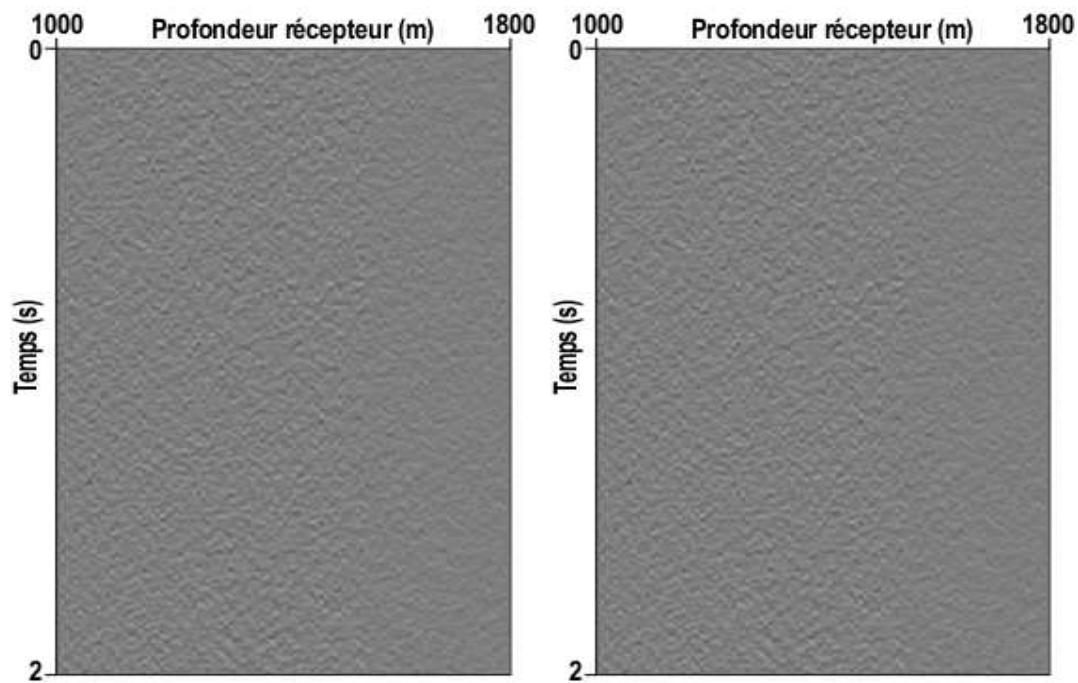


FIG. 4.46 – Résidus obtenues à l’issue de 1000 itérations sur les données bruitée données correspondant à l’onde plane 0° (gauche), et 6° (droite).

En effet, la représentation des résidus montre que l’information sismique contenue dans les données est expliquée par le modèle d’impédance et les CP calculées. La partie des données demeurant non expliquée est le bruit qui a été introduit. Le comportement de l’algorithme d’inversion est donc satisfaisant. Le résultat d’impédance correspondant est présenté figure 4.47

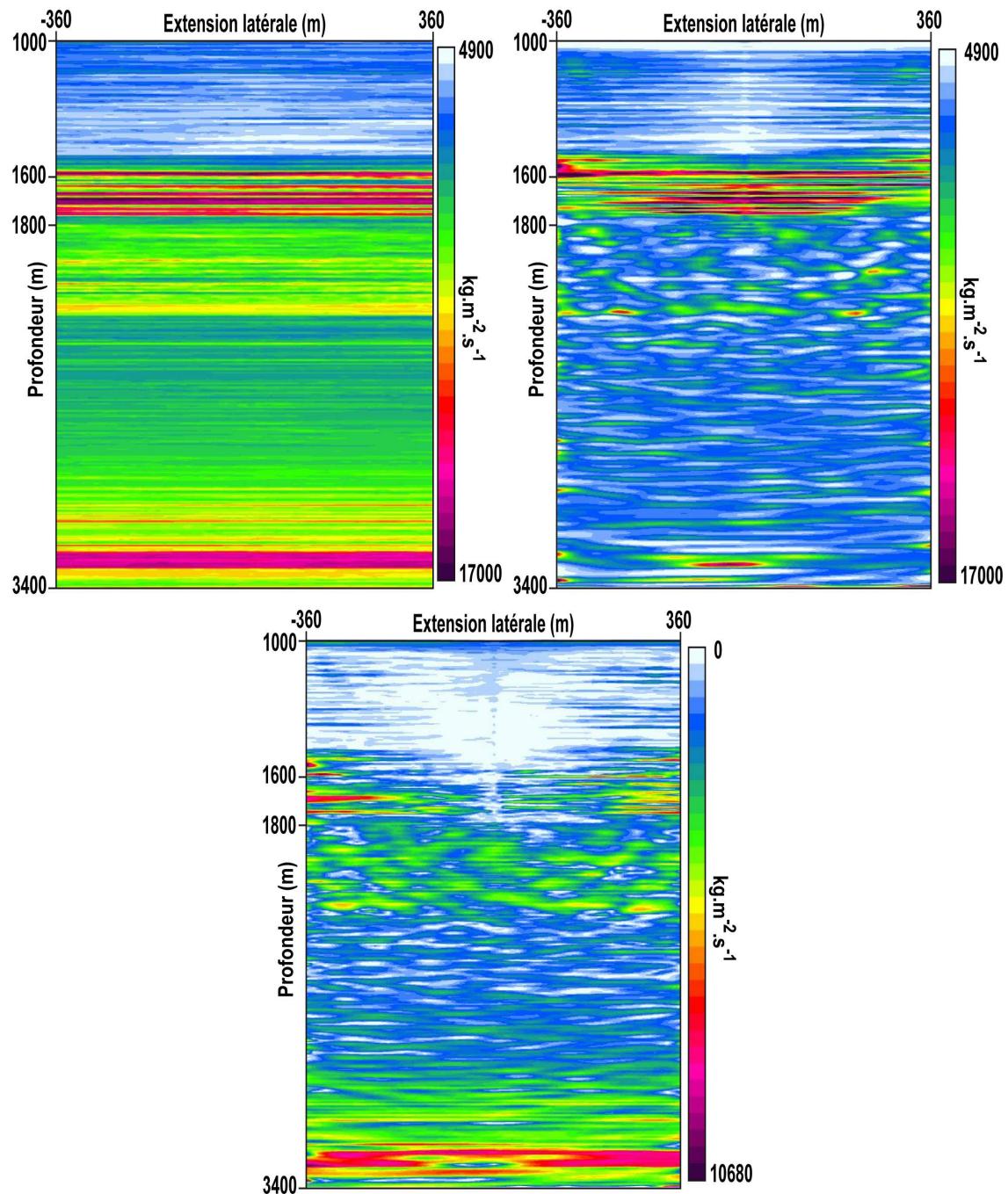


FIG. 4.47 – Impédance exacte (haut gauche), impédance calculée par inversion en utilisant des données bruitées (haut droite), différence (bas).

Une fois encore, la partie supérieure du modèle d'impédance est assez bien retrouvée. Le résultat se dégrade en profondeur, et l'indétermination des variations latérales de la zone profonde est à nouveau mise en avant, la régularité du modèle

dans cette zone étant très mauvaise. Ce résultat peut néanmoins être amélioré par l'utilisation d'un poids de régularisation plus fort pour l'impédance dans la zone profonde. Le résultat correspondant à cette configuration est présenté figure 4.48

L'introduction de bruit dans les données correspond à une perte de fiabilité de l'information sismique, dans ce contexte il est donc naturel d'accorder un rôle plus important à l'information *a priori*. Le résultat obtenu est conforme aux attentes, c'est-à-dire très régulier. Néanmoins, cette hausse du poids accordé à l'information *a priori* se traduit par une perte de résolution du résultat, qui est moins précis, notamment dans la zone supérieure.

Ces deux dernières expériences enseignent donc que la méthode demeure raisonnablement sensible au bruit sur les données. En présence de bruit, la zone supérieure reste bien imagée. La reconstruction de la zone profonde, plus instable, est en revanche dégradée, ce qui peut être compensé par une augmentation du poids associé à l'information *a priori*. La méthode d'inversion joue quant à elle son rôle, en expliquant l'information sismique contenue dans les données. Ces résultats préliminaires pourraient être complétés par une analyse précise de la perte de résolution engendrée par la présence de bruit sur les données, et l'évolution du résultat à mesure que l'amplitude du bruit augmente. Pour des données réelles, il n'est en effet pas rare de trouver un niveau de bruit proche de 100% du signal sismique, voire supérieur.

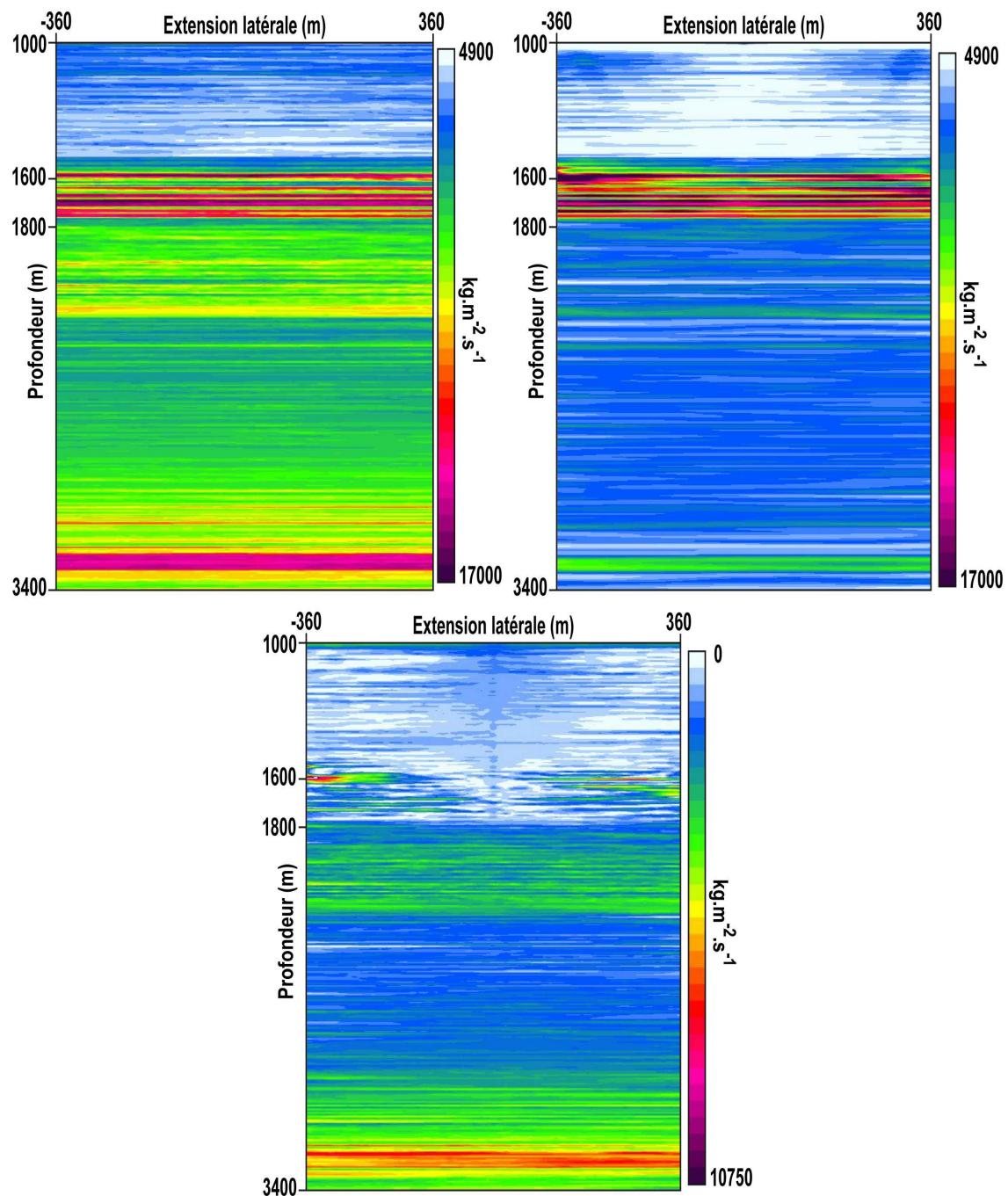


FIG. 4.48 – Impédance exacte (haut gauche), impédance calculée par inversion en utilisant des données bruitées et une régularisation plus forte (haut droite), différence (bas).

Conclusion

Les expériences décrites dans ce chapitre permettent d'acquérir une idée des performances de la méthode d'inversion développée. Le premier constat qui peut être fait concerne son bon comportement d'un point de vue numérique. L'algorithme d'optimisation emboîtée permet de lutter efficacement contre l'indétermination due au couplage des inconnues dans l'inversion. La parallélisation du calcul du gradient donne accès à une convergence rapide, permettant d'obtenir des temps de calcul de l'ordre de plusieurs dizaines d'heure sur 128 processeurs, pour un cas test faisant intervenir plusieurs millions de données discrètes et plusieurs centaines de milliers d'inconnues. Un niveau supplémentaire de parallélisation peut de plus être facilement introduit, ce qui laisse un potentiel facilement exploitable pour réduire encore les temps de calcul, si l'on dispose de l'architecture de calcul nécessaire.

Le deuxième point concerne l'évaluation de la résolution de la méthode. Deux zones distinctes se dégagent de cette analyse, la zone supérieure ($z < 1800$ m) et la zone profonde ($z > 1800$ m). Dans la première, la résolution latérale et verticale est excellente, le modèle étant retrouvé de manière presque exacte jusqu'à une distance proche de 300 m du puits. Comme cela a été mentionné, des expériences menées dans un contexte 1D laissent supposer que la résolution dépend dans cette zone du pas d'échantillonage des récepteurs, la résolution augmentant à mesure que celui-ci diminue. De nouvelles expériences devraient confirmer cette hypothèse.

Dans la zone profonde, la résolution est gouvernée par les données sismiques. Celles-ci, en raison de l'hétérogénéité du modèle de sous-sol utilisé, comportent une part très importante de réflexions multiples, qui sont prises en compte par la méthode d'inversion non linéaire développée, et permettent l'obtention d'une résolution verticale de l'ordre du pas de discréétisation utilisé ($\Delta z = 8$ m), pour un éclairage du milieu donc la longeur d'onde moyenne vaut 140 m. Ceci représente une amélioration très nette des méthodes d'inversion linéaire classiques, pour lesquelles, dans ce contexte, la résolution est de l'ordre de plusieurs dizaines de mètres.

La résolution latérale est, en revanche, nettement dégradée dans cette zone. Les expériences qui ont été menées montrent une assez forte indétermination des variations latérales, mise notamment en évidence par la représentation des profils sismiques latéraux. Cette indétermination peut être cependant compensée par l'information *a priori*. Le modèle calculé suit alors la basse fréquence des variations latérales, sans retrouver les détails de ces variations avec précision. Dans les cas présentés, l'évaluation de la Point Spread Function donne une estimation de la résolution latérale de l'ordre de la cinquantaine de mètres, ce qui est équivalent à la résolution des méthodes d'inversion linéaires. Cependant, l'utilisation d'un nombre plus important de jeux de données, avec une ouverture d'angle plus importante, devrait permettre d'améliorer la résolution latérale dans la zone profonde.

Pour conclure cette analyse de la résolution, il semble probable que, dans le contexte de l'expérience qui a été définie dans la première partie, la résolution verticale et latérale dans la zone supérieure, et la résolution verticale dans la zone profonde, dépendent des pas de discréétisation choisis ($\Delta x = 12$ m et $\Delta z = 8$ m). Des expériences menées avec des pas de discréétisation inférieurs devraient montrer une amélioration de la résolution. Ceci implique cependant la définition d'un cas test comportant encore plus d'inconnues.

Concernant la sensibilité de la méthode à l'estimation du champ de vitesse, il semble que les paramètres clés soient les valeurs de ce champ au niveau des récepteurs. Les expériences réalisées montrent en effet qu'une faible erreur commise sur ces paramètres engendre une très mauvaise détermination du modèle d'impédance. En revanche, l'utilisation d'un modèle de vitesse ne présentant pas de variations latérales, n'a pas engendré de perturbations majeures sur le résultat. Ceci est cependant dû au fait que la gamme d'angle utilisée pour la recombinaison des données reste faible (maximum de 6°). Ainsi, l'éclairage du milieu est fait par des ondes se propageant dans une direction proche de la verticale, sur lesquelles les variations latérales du champ de vitesse ont peu d'influence. Ceci est à prendre en considération si l'on décide d'ouvrir la gamme d'angle pour améliorer la résolution latérale de la zone profonde : dans ce cas, les variations latérales du modèle de vitesse gagnent probablement en importance.

Les expériences préliminaires menées sur la sensibilité de la méthode au bruit sur les données permettent de montrer que la méthode numérique est robuste, car l'information sismique contenue dans les données est efficacement dissociée du bruit. Le résultat obtenu montre également la stabilité de la méthode pour l'imagerie de la zone supérieure. Celle-ci est en effet bien retrouvée en présence de bruit. En revanche, la partie profonde est plus sensible, notamment en ce qui concerne les variations latérales. Une légère perte de résolution verticale est également enregistrée. Toutefois, la méthode semble globalement stable pour le niveau de bruit utilisé. Des investigations plus précises doivent être menées pour évaluer le comportement face à un niveau de bruit plus élevé. Si des instabilités sont mises à jour, l'utilisation de nouveaux termes de régularisation peut être envisagée, comme par exemple l'ajout d'un terme mesurant la norme L^1 des variations verticales du modèle d'impédance. Un tel terme est effectivement utilisé dans l'application 1D pour stabiliser la méthode utilisée sur des données bruitées [Macé, Lailly, 1986]. Ceci requiert cependant la prise en compte d'une mise à jour de l'algorithme d'optimisation, un tel terme n'étant pas différentiable. A cet effet, la reformulation du problème comme un problème d'optimisation sous contraintes linéaires est préférable à l'utilisation d'une méthode de sous-gradient.

La méthode d'inversion non linéaire 2D de données sismiques de puits développée

présente donc un comportement numérique satisfaisant. De plus, sur le plan de la géophysique, elle permet une très nette amélioration de la résolution verticale du modèle d'impédance, et l'estimation de celui-ci dans la zone supérieure est excellente. L'estimation des variations latérales dans la zone profonde est sujette à plus d'indétermination, qui doit être compensée par l'information *a priori* ou une hausse de l'information sismique. La sensibilité au modèle de vitesse montre l'importance de la détermination de celui-ci au niveau du puits. Finalement, le bruit sur les données influe principalement sur l'indétermination dans la zone profonde, qui doit être à nouveau compensée par une hausse de l'information *a priori*, ou de l'information sismique. Cette méthode semble donc pouvoir donner des renseignements à très fine échelle sur la composition du sous-sol dans la zone supérieure, et proposer une estimation présentant une haute résolution verticale et une résolution latérale de l'ordre de celle des méthodes d'inversion conventionnelles dans la zone profonde. D'autre d'investigations restent à mener pour poursuivre l'évaluation de son potentiel, notamment son application à des contextes où les variations latérales sont plus importantes et la régularisation plus difficile. Il est également important de tester la sensibilité de la résolution latérale en zone profonde par rapport à l'estimation du modèle de vitesse, et le comportement de la méthode sur des données fortement bruitées.

Chapitre 5

Analyse des équations PML

Introduction

La problématique de la simulation de la propagation des ondes dans des milieux d'extension infinie se pose depuis déjà plus d'une trentaine d'années. Elle a été à l'origine de très nombreux travaux sur la construction de conditions aux limites permettant de reproduire le mieux possible le comportement de la propagation d'ondes en milieu non borné. Les premiers efforts ont conduit à la définition de conditions aux limites absorbantes construites comme des approximations, à l'aide d'opérateurs différentiels, des conditions aux limites exacte. Cette démarche a été introduite par [Engquist, Majda, 1977] pour l'équation des ondes acoustiques

$$\partial_{tt}u - \Delta u = 0.$$

De très nombreux travaux ont été ensuite initiés pour l'amélioration et l'extension de cette technique à un ensemble d'équations plus large. Des conditions absorbantes d'ordre élevé ont notamment été proposées par [Collino, 1992], qui utilise des opérateurs différentiels d'ordre élevé pour réaliser l'approximation de la condition exacte au bord. Le principal défaut de ces méthodes réside dans leur difficulté d'implémentation, même si la précision qu'il est possible d'obtenir par leur emploi est assez bonne.

Parallèlement à ces travaux, d'autres auteurs se sont intéressés à la construction de couches qualifiées de "couches éponges", entourant le domaine d'intérêt, dont le rôle est d'absorber les ondes incidentes, sans les laisser reparaître dans le domaine où la propagation des ondes est étudiée. Différentes techniques ont été utilisées pour parvenir à cette fin : introduction d'un coefficient d'absorption [Colonius et al., 1993], accélération de l'onde [Ta'asan, Nark, 1995], décélération de l'onde [Mazaheri, Roe, 1997], ou une combinaison de ces techniques [Freund, 1997]. Bien que d'implémentation plus simple, l'efficacité de ces méthodes reste assez limitée.

Cette problématique a connu une véritable révolution lorsque le concept de PML (Perfectly Matched Layers) a été introduit par Bérenger, pour la modélisation de la propagation d'ondes électromagnétiques, d'abord dans un cadre 2D [Bérenger, 1994], puis 3D [Bérenger, 1996]. Cette couche parfaitement absorbante reprend le concept des couches éponges utilisant des coefficients d'absorption, mais en utilisant en plus une décomposition des variables suivant la direction de dérivation dans l'espace. Une analyse théorique de la propagation en ondes planes dans un milieu homogène bordé d'une couche PML montre que les coefficients de réflexion à l'interface avec la couche, dans un formalisme continu, sont nuls, quelque soit l'angle d'incidence. Bérenger a ainsi su définir une méthode simple d'implémentation, surpassant en efficacité celles qui avaient été imaginées auparavant. Ce concept a rapidement été étendu au cadre de l'acoustique [Hu, 1996]. Il est à noter cependant qu'aucun travail de comparaison n'a été mené, à notre connaissance, sur une comparaison exhaustive, utilisant des cas tests semblables, entre les conditions absorbantes et les couches PML. Celles-ci présentent en effet l'inconvénient de contraindre l'utilisateur à définir une couche d'épaisseur non nulle autour du domaine d'intérêt, ce qui induit un surcoût de calcul par rapport aux conditions aux limites absorbantes. Cette comparaison serait d'autant plus intéressante qu'il arrive parfois de voir, dans les expérimentations numériques, des auteurs utilisant des couches PML de taille de l'ordre de celle du domaine d'intérêt.

Cependant, les analyses mathématiques qui ont suivi le développement des couches PML ont montré leur caractère mal posé, d'abord pour les ondes électromagnétiques et les équations de Maxwell [Abarbanel, Gottlieb, 1997], puis pour les équations d'Euler linéarisées (modélisation acoustique) [Hesthaven, 1998]. Ce caractère mal posé peut se traduire numériquement par l'apparition d'instabilités. Une telle instabilité a notamment été mis en évidence par [Abarbanel et al., 2002], pour des temps longs : après le passage de l'onde dans le domaine d'intérêt, un nouveau front d'onde se crée dans la couche PML, qui finit par se propager, pour des temps suffisamment longs, dans le domaine d'intérêt. Un tel comportement peut-être gênant suivant l'application envisagée.

Certains auteurs ont tenté de reprendre la formulation PML introduite par Bérenger pour définir de nouvelles équations PML qui soient bien posés. Parmi ces travaux, se trouvent notamment ceux de [Hesthaven, 1998], [Rahmouni, 2004] et [Nataf, 2006] pour l'acoustique. D'autre part, une étude plus détaillée du caractère instable des équations de Bérenger a été effectué au cours de la thèse de [Petit-Bergez, 2006]. Cet auteur a montré que les équations de Bérenger pour les ondes électromagnétiques étaient en fait seulement faiblement bien posées, de défaut 1 (perte de régularité d'un ordre de dérivation dans les estimations d'énergie).

Tous les travaux qui ont été mentionnés jusqu'à maintenant étudient le comportement de la propagation des ondes pour des milieux homogènes, que ce soit dans le

cadre de la propagation d'ondes électromagnétiques ou acoustiques. Ainsi, l'ambition des résultats qui sont présentés dans la première partie de ce chapitre est de prolonger l'analyse de [Petit-Bergez, 2006] au cas d'un milieu de propagation hétérogène. La formulation PML utilisée dans la thèse dérive en effet des équations de Bérenger, adaptée aux équations d'Euler linéarisées, sans advection, pour la modélisation de la propagation d'ondes acoustiques en milieu hétérogène. Pour atteindre une portée plus générale, ces résultats sont obtenus dans un cadre de modélisation 3D.

Dans une seconde partie, une expérimentation numérique précise s'attache à montrer d'une part l'influence de la complexité du milieu sur l'efficacité de la couche PML comme couche absorbante, d'autre part que l'utilisation des équations PML est bien adaptée au problème d'inversion de données sismiques de puits 2D, le biais de modélisation introduit étant suffisamment faible pour être négligé.

5.1 Analyse mathématique

5.1.1 Contexte et résultats

Les équations Euler PML 3D sont dérivées du modèle de Bérenger pour l'électromagnétisme. Elles modélisent la propagation d'ondes acoustiques dans le milieu borné hétérogène $\Omega \subset \mathbb{R}^3$ de frontière Γ .

$$\left\{ \begin{array}{l} \partial_t u_x - \alpha c \partial_x (p_x + p_y + p_z) + c \sigma_x u_x = 0 \\ \partial_t u_y - \alpha c \partial_y (p_x + p_y + p_z) + c \sigma_y u_y = 0 \\ \partial_t u_z - \alpha c \partial_z (p_x + p_y + p_z) + c \sigma_z u_z = 0 \\ \partial_t p_x - \frac{c}{\alpha} \partial_x u_x + c \sigma_x p_x = 0 \\ \partial_t p_y - \frac{c}{\alpha} \partial_y u_y + c \sigma_y p_y = 0 \\ \partial_t p_z - \frac{c}{\alpha} \partial_z u_z + c \sigma_z p_z = 0, \end{array} \right. \quad (5.1)$$

avec les conditions initiales et les conditions aux limites

$$(p_x + p_y + p_z)(x, y, z, 0) = p_0(x, y, z) \quad (5.2)$$

$$(p_x + p_y + p_z)|_{\Gamma} = h(x, y, t). \quad (5.3)$$

Les variables u_x, u_y et u_z désignent les composantes spatiales de la vitesse de déplacement. Les variables p_x, p_y, p_z résultent de la décomposition de la variable de pression : leur somme donne la pression p dans le domaine Ω . Les coefficients $c \in C^\infty(\Omega)$ et $\alpha \in C_{pm}^\infty(\Omega)$, désignent respectivement la vitesse de propagation des ondes acoustiques, et l'inverse de l'impédance acoustique, où $C_{pm}^\infty(\Omega)$ est l'ensemble des fonctions de classe C^∞ par morceaux définies sur Ω . En effet, alors que la vitesse de propagation c est une fonction lisse, l'impédance est discontinue ; ses discontinuités donnent lieu aux réflexions observées lors de la modélisation de la propagation d'ondes sismiques acoustiques dans le sous-sol. La donnée de la pression à la frontière du domaine suffit à déterminer la solution.

Le résultat qui va être démontré est le suivant : alors que les équations d'Euler sans absorption, ni séparation des variables, sont bien posées pour un tel problème, les équations PML de Bérenger sont seulement faiblement bien posées, de défaut 1. Ceci a été déjà prouvé par Petit-Bergez dans ses travaux de thèse, pour les équations de Maxwell transverse électrique 2D, et les équations de Maxwell 3D, en milieu homogène et pour le problème de Cauchy. La nouveauté réside donc dans le caractère

borné et hétérogène du milieu de propagation.

La preuve de ce résultat s'effectue en deux étapes : un premier théorème établit le résultat pour le problème posé dans le demi-espace ($\Omega = [0, +\infty] \times \mathbb{R}^2$) pour des coefficients d'impédance lisses $\alpha \in C^\infty(\Omega)$ (modèle monocouche), le milieu étant stimulé à sa frontière $\Gamma = \{(x, y, z) \in \mathbb{R}^3, z = 0\}$. Ce problème est relié au caractère borné du domaine de propagation. Il impose notamment une certaine régularité sur la condition à la limite $h(x, y, t)$. Un deuxième théorème établit ensuite le résultat pour le problème de Cauchy ($\Omega = \mathbb{R}^3$) avec une discontinuité d'impédance en $z = 0$ (modèle bicouche). Une attention spéciale doit en effet être apportée aux discontinuités éventuelles du paramètre d'impédance. La preuve de ce deuxième résultat nécessite cependant une hypothèse forte restreignant quelque peu l'approche. Le problème d'une discontinuité d'impédance ne peut en effet être traité que dans le cas d'un modèle d'impédance 1D, c'est à dire parfaitement régulier dans les directions tangentes à l'interface. Ceci revient à considérer des milieux de propagation parfaitement stratifiés, ici suivant la direction z . La généralisation de ce résultat à un modèle d'impédance dépendant de toutes les directions de l'espace semble pour l'instant hors de portée.

Ces deux problèmes pourraient être traités simultanément, néanmoins, pour des raisons évidentes de clarté, ils sont exposés à l'aide de deux théorèmes distincts. La preuve de ceux-ci reposent sur la construction d'inégalités d'énergie pour la solution des équations PML. Celles-ci sont obtenues à partir de la définition d'un système augmenté, suivant l'idée présentée dans [Métral, Vacus, 1999], faisant intervenir les dérivées premières des inconnues initiales, et pour lequel un symétriseur peut être calculé. Il est possible de montrer que ce système est fortement bien posé ; il en découle que le système PML initial est seulement faiblement bien posé, de défaut 1.

5.1.2 Définition du système étendu

Les équations (5.1) se récrivent, en posant $v = c\sigma_x p_x + c\sigma_y p_y$

$$\left\{ \begin{array}{l} \partial_t u_x - \alpha c \partial_x p + c \sigma_x u_x = 0 \\ \partial_t u_y - \alpha c \partial_y p + c \sigma_y u_y = 0 \\ \partial_t u_z - \alpha c \partial_z p + c \sigma_z u_z = 0 \\ \partial_t p - \frac{c}{\alpha} \partial_x u_x - \frac{c}{\alpha} \partial_y u_y - \frac{c}{\alpha} \partial_z u_z + v + c \sigma_z p_z = 0 \\ \partial_t v - \sigma_x \frac{c^2}{\alpha} \partial_x u_x - \sigma_y \frac{c^2}{\alpha} \partial_y u_y - c^2 \sigma_x \sigma_y (p - p_z) + c(\sigma_x + \sigma_y)v = 0 \\ \partial_t p_z - \frac{c}{\alpha} \partial_z u_z + c \sigma_z p_z = 0. \end{array} \right. \quad (5.4)$$

Soit $U^1 = \begin{bmatrix} u_x \\ u_y \\ u_z \\ p \end{bmatrix}$, $U^2 = \partial_x U^1$, $U^3 = \partial_y U^1$, $U^4 = \partial_z U^1$, $U^5 = \begin{bmatrix} 0 \\ 0 \\ 0 \\ v \end{bmatrix}$, $U^6 = \begin{bmatrix} 0 \\ 0 \\ 0 \\ p_z \end{bmatrix}$,

$$D = \begin{pmatrix} c\sigma_x & 0 & 0 & 0 \\ 0 & c\sigma_y & 0 & 0 \\ 0 & 0 & c\sigma_z & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad E = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \quad S^1 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ \frac{c^2 \sigma_x}{\alpha} & 0 & 0 & 0 \end{pmatrix},$$

$$S^2 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & \frac{c^2 \sigma_y}{\alpha} & 0 & 0 \end{pmatrix}, \quad S^3 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & c^2 \sigma_x \sigma_y \end{pmatrix}, \quad S^4 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & c(\sigma_x + \sigma_y) \end{pmatrix},$$

$$S^5 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & \frac{c}{\alpha} & 0 \end{pmatrix}, \quad S^6 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & c \sigma_z \end{pmatrix},$$

$$A = \begin{pmatrix} 0 & 0 & 0 & \alpha c \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ \frac{c}{\alpha} & 0 & 0 & 0 \end{pmatrix}, \quad B = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \alpha c \\ 0 & 0 & 0 & 0 \\ 0 & \frac{c}{\alpha} & 0 & 0 \end{pmatrix}, \quad C = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \alpha c \\ 0 & 0 & \frac{c}{\alpha} & 0 \end{pmatrix}.$$

Le système augmenté considéré est le suivant

$$\left\{ \begin{array}{l} \partial_t U^1 - A \partial_x U^1 - B \partial_y U^1 - C \partial_z U^1 + D U^1 + E U^5 + S^6 U^6 = 0 \\ \partial_t U^2 - A \partial_x U^2 - B \partial_y U^2 - C \partial_z U^2 + D U^2 - \partial_x(A)U^2 - \partial_x(B)U^3 - \partial_x(C)U^4 + \partial_x(D)U^1 \\ + \partial_x(S^6)U^6 + E \partial_x U^5 + S^6 \partial_x U^6 = 0 \\ \partial_t U^3 - A \partial_x U^3 - B \partial_y U^3 - C \partial_z U^3 + D U^3 - \partial_y(A)U^2 - \partial_y(B)U^3 - \partial_y(C)U^4 + \partial_y(D)U^1 \\ + \partial_y(S^6)U^6 + E \partial_y U^5 + S^6 \partial_y U^6 = 0 \\ \partial_t U^4 - A \partial_x U^4 - B \partial_y U^4 - C \partial_z U^4 + D U^3 - \partial_z(A)U^2 - \partial_z(B)U^3 - \partial_z(C)U^4 + \partial_z(D)U^1 \\ + \partial_z(S^6)U^6 + E \partial_z U^5 + S^6 \partial_z U^6 = 0 \\ \partial_t U^5 - S^1 U^2 - S^2 U^3 - S^3(U^1 - U^6) + S^4 U^5 = 0 \\ \partial_t U^6 - S^5 U^4 + S^6 U^6 = 0, \end{array} \right. \quad (5.5)$$

ce qui, avec la définition de

$$\begin{aligned} \tilde{A} &= \begin{pmatrix} A & 0 & 0 & 0 & 0 & 0 \\ 0 & A & 0 & 0 & -E & -S^6 \\ 0 & 0 & A & 0 & 0 & 0 \\ 0 & 0 & 0 & A & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}, \quad \tilde{B} = \begin{pmatrix} B & 0 & 0 & 0 & 0 & 0 \\ 0 & B & 0 & 0 & 0 & 0 \\ 0 & 0 & B & 0 & -E & -S^6 \\ 0 & 0 & 0 & B & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}, \\ \tilde{C} &= \begin{pmatrix} C & 0 & 0 & 0 & 0 & 0 \\ 0 & C & 0 & 0 & 0 & 0 \\ 0 & 0 & C & 0 & 0 & 0 \\ 0 & 0 & 0 & C & -E & -S^6 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}, \quad \tilde{U} = \begin{bmatrix} U^1 \\ U^2 \\ U^3 \\ U^4 \\ U^5 \\ U^6 \end{bmatrix}, \\ \tilde{S} &= \begin{pmatrix} D & 0 & 0 & 0 & E & S^6 \\ \partial_x D & D - \partial_x A & -\partial_x B & -\partial_x C & 0 & \partial_x S^6 \\ \partial_y D & -\partial_y A & D - \partial_y B & -\partial_y C & 0 & \partial_y S^6 \\ \partial_z D & -\partial_z A & -\partial_z B & D - \partial_z C & 0 & \partial_z S^6 \\ -S^3 & -S^1 & -S^2 & 0 & S^4 & S^3 \\ 0 & 0 & 0 & -S^5 & 0 & S^6 \end{pmatrix}, \end{aligned}$$

est équivalent à l'équation

$$\partial_t \tilde{U} - \tilde{A} \partial_x U - \tilde{B} \partial_y U - \tilde{C} \partial_z U + \tilde{S} U = 0. \quad (5.6)$$

5.1.3 Définition d'un symétriseur pour le système étendu

Afin de pouvoir construire des inégalités d'énergie il convient dans un premier temps de définir un symétriseur du système étendu. Sa définition est donnée dans les lignes qui suivent.

Définition 1 Soit un opérateur

$$\begin{aligned} \mathcal{P}(x, t, \partial_x) &= \sum_{j=1}^J A_j(x, t) \partial_{x_j}, \quad x \in \mathbb{R}^J \\ \text{de symbole } P(x, t, i\omega) &= i \sum_{j=1}^J \omega_j A_j(x, t) \quad \omega \in \mathbb{R}^J. \end{aligned} \tag{5.7}$$

Un symétriseur de cet opérateur est une matrice hermitienne $H(x, t, \omega)$ définie positive pour tout $x \in \mathbb{R}^J$, $t \in [0, T]$, $\omega \in \mathbb{R}^J$, $|w| = 1$, C^∞ par rapport à tous ses arguments, et vérifiant

$$H(x, t, \omega) P(x, t, \omega) + P^*(x, t, \omega) H(x, t, \omega) = 0.$$

Lemme 1 Il existe un symétriseur réel $\tilde{H}(x, y, z)$ pour l'opérateur

$$\mathcal{P}_1 = \tilde{A} \partial_x + \tilde{B} \partial_y + \tilde{C} \partial_z.$$

De plus, son indépendance par rapport aux variables ω_1 , ω_2 et ω_3 lui confère la propriété de symétriser simultanément les matrices \tilde{A} , \tilde{B} et \tilde{C} .

5.1.3.1 Démonstration

Un symétriseur $H(x, y, z)$ pour les équations d'Euler sans absorption est

$$H(x, y, z) = \begin{pmatrix} c & 0 & 0 & 0 \\ \alpha & c & 0 & 0 \\ 0 & \alpha & c & 0 \\ 0 & 0 & 0 & \alpha c \end{pmatrix}.$$

On vérifie que

$$\tilde{H}(x, y, z) = \begin{pmatrix} H & 0 & 0 & 0 & 0 & 0 \\ 0 & H & 0 & 0 & X_1 & Y_1 \\ 0 & 0 & H & 0 & X_2 & Y_2 \\ 0 & 0 & 0 & H & X_3 & Y_3 \\ 0 & X_1^T & X_2^T & X_3^T & L & 0 \\ 0 & Y_1^T & Y_2^T & Y_3^T & 0 & L \end{pmatrix}, \tag{5.8}$$

est un symétriseur de \mathcal{P}_1 qui symétrise simultanément \tilde{A} , \tilde{B} , \tilde{C} avec

$$X_1 = \begin{pmatrix} 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad X_2 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix},$$

$$X_3 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad L = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & \frac{7\alpha}{c} \end{pmatrix},$$

$$Y_1 = c\sigma_z X_1, \quad Y_2 = c\sigma_z X_2, \quad Y_3 = c\sigma_z X_3.$$

Remarque. L'existence du symétriseur $\tilde{H}(x, y, z)$ prouve que le problème de Cauchy pour le système étendu et des coefficients $\alpha, c \in C^\infty(\Omega)$ est fortement hyperbolique et donc fortement bien posé [Kreiss, Lorenz, 2004].

5.1.4 Théorème 1 : cas du demi-espace pour des coefficients d'impédance lisses

5.1.4.1 Enoncé

Soit $\Omega = [0, +\infty] \times \mathbb{R}^2$ de frontière $\Gamma = \{(x, y, z) \in \mathbb{R}^3, z = 0\}$, $\mathcal{Q} = [0, T] \times \Omega$, $\Sigma = [0, T] \times \Gamma$.

Soit $h(x, t) \in H^{7/2, 7/2}(\Sigma)$ tel que

$$h(x, y, 0) = 0, \quad \partial_t h(x, y, 0) = 0, \quad \partial_{tt} h(x, y, 0) = 0.$$

Pour $(\sigma_x, \sigma_y, \sigma_z) \in (W^{1, \infty}(\Omega))^2$, $c \in C^\infty(\Omega)$, $\alpha \in C^\infty(\Omega)$ tels que

$$0 < c_{min} \leq c(x, y, z) \leq c_{max}, \tag{5.9}$$

$$0 < \alpha_{min} \leq \alpha(x, y, z) \leq \alpha_{max}, \quad \text{alors} \tag{5.10}$$

$\exists(C_1, C_2) \in \mathbb{R}^2$, tels que pour toute solution suffisamment régulière de (5.1), (5.2) on aie

$$\begin{aligned} & \| (u_x, u_y, u_z, p) \|_{L^\infty(0, T; H^1(\Omega))}^2 + \| (c(\sigma_x p_x + \sigma_y p_y)) \|_{L^\infty(0, T; L^2(\Omega))}^2 + \| p_z \|_{L^\infty(0, T; L^2(\Omega))}^2 \leq \\ & \frac{2N e^{\frac{(\gamma+1)T}{M}}}{M} \left(\| (u_x, u_y, u_z, p)(., ., ., 0) \|_{H^1(\Omega)}^2 + \| c(\sigma_x p_x + \sigma_y p_y)(., ., ., 0) \|_{L^2(\Omega)}^2 + \| p_z(., ., ., 0) \|_{L^2(\Omega)}^2 \right) \\ & + \frac{2C_1 e^{\frac{(\gamma+1)T}{M}}}{M} \| h \|_{H^{7/2, 7/2}(\Sigma)}^2 + 2C_2 \| h \|_{H^{7/2, 7/2}(\Sigma)}^2 \quad \text{avec} \end{aligned} \tag{5.11}$$

$$\gamma(\alpha, c) = \sup_{x,z} \|\partial_x(\tilde{H}\tilde{A}) + \partial_y(\tilde{H}\tilde{B}) + \partial_z(\tilde{H}\tilde{C}) + 2\tilde{H}\tilde{S}\|_2 \quad (5.12)$$

$$M(\alpha, c) = \min \left\{ 1, \min_{x,z} \left\{ \alpha(x, z)c(x, z), \frac{c(x, z)}{\alpha(x, z)}, \frac{7\alpha(x, z)}{c(x, z)}, \frac{7\alpha^2 + c^2 - \sqrt{49\alpha^4 + 10\alpha^2c^2 + c^4}}{2\alpha c}(x, z) \right\} \right\}, \quad (5.13)$$

$$N(\alpha, c) = \max \left\{ 1, \|\alpha c\|_\infty, \left\| \frac{c}{\alpha} \right\|_\infty, \left\| \frac{7\alpha}{c} \right\|_\infty, \left\| \frac{7\alpha^2 + c^2 + \sqrt{49\alpha^4 + 10\alpha^2c^2 + c^4}}{2\alpha c} \right\|_\infty \right\}. \quad (5.14)$$

5.1.4.2 Démonstration

Soit $(., .)_{L^2(\Omega)}$ le produit scalaire défini sur $L^2(\Omega)$, et $(., .)$ le produit scalaire défini sur \mathbb{R}^n , $n \in \mathbb{N}$. Pour toutes fonctions

$$U, V \in L^2(\Omega)^n \times L^2(\Omega)^n, \quad n \in \mathbb{N},$$

la notation suivante est adoptée

$$(U, V)_{L^2(\Omega)} = \sum_{i=1}^n (U_i, V_i)_{L^2(\Omega)} = \int_{\Omega} (U, V) d\Omega.$$

Le symétriseur \tilde{H} induit le produit scalaire $< ., . >$ tel que

$$\forall (U, V) \in L^2(\Omega)^{24} \times L^2(\Omega)^{24}, \quad < U, V > = (U, \tilde{H}V)_{L^2(\Omega)}.$$

La norme induite par ce produit scalaire, notée $|.|$, est équivalente à la norme L^2 , et

$$M\|\tilde{U}\|_{L^2}^2 \leq |\tilde{U}|^2 \leq N\|\tilde{U}\|_{L^2}^2,$$

où M et N sont définis par les équations (5.13) et (5.14). Les matrices suivantes, symétriques par construction, exceptée \tilde{S} , sont définies

$$\check{A} = \tilde{H}\tilde{A}, \quad \check{B} = \tilde{H}\tilde{B}, \quad \check{C} = \tilde{H}\tilde{C}, \quad \check{S} = \tilde{H}\tilde{S}.$$

Lemme 2 Il existe un relèvement W de \tilde{U} tel que $\overline{U} = \tilde{U} - W$ vérifie

$$\partial_t \overline{U} - \tilde{A} \partial_x \overline{U} - \tilde{B} \partial_y \overline{U} - \tilde{C} \partial_z \overline{U} + \tilde{S} \overline{U} = F.$$

où F est un second membre non nul dépendant de la condition de pression au bord $h(x, y, t)$. De plus,

$$\left\{ \begin{array}{l} \forall (x, y, z) \in \Omega, \quad \bar{U}(x, y, z, 0) = \tilde{U}(x, y, z, 0), \quad (P1) \\ \int_{\Gamma} (\check{C}\bar{U}(x, y, 0, t), \bar{U}(x, y, 0, t)) d\Gamma = 0, \quad (P2) \\ \exists C_1 \in \mathbb{R}^+ \quad \|F\|_{L^2(0,T;L^2(\Omega))}^2 \leq C_1 \|h\|_{H^{7/2,7/2}(\Sigma)}^2, \quad (P3) \\ \exists C_2 \in \mathbb{R}^+ \quad \|W(t)\|_{L^\infty(0,T;L^2(\Omega))}^2 \leq C_2 \|h\|_{H^{7/2,7/2}(\Sigma)}^2. \quad (P4) \end{array} \right.$$

La démonstration de ce lemme est donnée dans la partie suivante. Soit un tel relèvement W , et $\bar{U} = \tilde{U} - W$. \bar{U} vérifie

$$\begin{aligned} \frac{d}{dt} |\bar{U}(t)|^2 &= \left(\bar{U}, \tilde{H} \frac{d\bar{U}}{dt} \right)_{L^2(\Omega)} + \left(\tilde{H}\bar{U}, \frac{d\bar{U}}{dt} \right)_{L^2(\Omega)} = \int_{\Omega} (\bar{U}, \check{A}\partial_x \bar{U} + \check{B}\partial_y \bar{U} + \check{C}\partial_z \bar{U}) d\Omega \\ &+ \int_{\Omega} (\check{A}\bar{U}, \partial_x \bar{U}) + (\check{B}\bar{U}, \partial_y \bar{U}) + (\check{C}\bar{U}, \partial_z \bar{U}) d\Omega - 2 \int_{\Omega} (\bar{U}, \check{S}\bar{U}) d\Omega + 2 \int_{\Omega} (\bar{U}, F) d\Omega. \end{aligned}$$

Par intégration par parties des trois premiers termes de la dernière ligne, il vient

$$\begin{aligned} \frac{d}{dt} |\bar{U}(t)|^2 &= - \int_{\Omega} (\bar{U}, \partial_x(\check{A})\bar{U} + \partial_y(\check{B})\bar{U} + \partial_z(\check{C})\bar{U}) d\Omega - 2 \int_{\Omega} (\bar{U}, \check{S}\bar{U}) d\Omega \\ &\quad - \int_{\Gamma} (\check{C}\bar{U}(x, y, 0, t), \bar{U}(x, y, 0, t)) d\Gamma. \end{aligned}$$

D'après les propriétés de \bar{U} , le terme de bord est nul, et

$$\frac{d}{dt} |\bar{U}(t)|^2 \leq \left(\sup_{x,y,z} \|\partial_x \check{A} + \partial_y \check{B} + \partial_z \check{C} + 2\check{S}\|_2 + 1 \right) \|\bar{U}(t)\|_{L^2(\Omega)}^2 + \|F(t)\|_{L^2(\Omega)}^2.$$

Ainsi, en utilisant la première inégalité relative au produit scalaire induit par \tilde{H}

$$\frac{d}{dt} |\bar{U}(t)|^2 \leq \frac{\gamma + 1}{M} |\bar{U}(t)|^2 + \|\tilde{F}(t)\|_{L^2(\Omega)}^2,$$

où γ est défini par l'équation (5.23).

L'utilisation du lemme de Gronwall et les inégalités relatives au produit scalaire induit par \tilde{H} permettent d'établir

$$\|\bar{U}(t)\|_{L^2(\Omega)}^2 \leq \frac{1}{M} e^{\frac{(\gamma+1)t}{M}} \left(N \|\bar{U}(0)\|_{L^2(\Omega)}^2 + \int_0^t \|\tilde{F}(\tau)\|_{L^2(\Omega)}^2 d\tau \right).$$

D'après la propriété (P4) du lemme 2,

$$\|\overline{U}(t)\|_{L^\infty(0,T;L^2(\Omega))} \leq \frac{1}{\sqrt{M}} \exp^{\frac{(\gamma+1)T}{2M}} \left(N \|\widetilde{U}(0)\|_{L^2(\Omega)} + C_1 \|h\|_{H^{7/2,7/2}(\Sigma)}^2 \right)^{1/2}.$$

Comme $\widetilde{U} = \overline{U} - W$, en utilisant les résultats du lemme 2 et l'inégalité triangulaire, on obtient

$$\|\widetilde{U}(t)\|_{L^\infty(0,T;L^2(\Omega))} \leq \frac{1}{\sqrt{M}} \exp^{\frac{(\gamma+1)T}{2M}} \left(N \|\widetilde{U}(0)\|_{L^2(\Omega)} + C_1 \|h\|_{H^{7/2,7/2}(\Sigma)}^2 \right)^{1/2} + \sqrt{C_2} \|h\|_{H^{7/2,7/2}(\Sigma)}.$$

Finalement, l'inégalité $(a+b)^2 \leq 2(a^2 + b^2)$, pour (a, b) dans \mathbb{R}^2 , permet d'établir

$$\|\widetilde{U}(t)\|_{L^\infty(0,T;L^2(\Omega))}^2 \leq \frac{2}{M} \exp^{\frac{(\gamma+1)T}{M}} \left(N \|\widetilde{U}(0)\|_{L^2(\Omega)} + C_1 \|h\|_{H^{7/2,7/2}(\Sigma)}^2 \right) + 2C_2 \|h\|_{H^{7/2,7/2}(\Sigma)}^2.$$

Preuve du lemme 2

Le terme de bord apparaissant dans l'intégration par partie est

$$\begin{aligned} \int_\Gamma (\check{C}\overline{U}(x, y, 0, t), \overline{U}(x, y, 0, t)) d\Gamma &= 2 \int_\Gamma c^2 (pu_z + \partial_x p \partial_x u_z + \partial_y p \partial_y u_z + \partial_z p \partial_z u_z) (x, y, 0, t) d\Gamma \\ &\quad - 2 \int_\Gamma \alpha c \partial_z p v (x, y, 0, t) d\Gamma, \end{aligned}$$

en se servant du fait que $\sigma_z(0) = 0$.

L'idée directrice est de définir un relèvement tel que la pression p s'annule sur la surface $z = 0$ (ce qui entraîne l'annulation de $\partial_x p$ et $\partial_y p$ sur $z = 0$), et tel que les variables v et $\partial_z u_z$ s'annulent sur cette même surface. On considère une réécriture simple des équations (5.1) faisant intervenir directement la variable de pression p .

$$\left\{ \begin{array}{l} \partial_t u_x - \alpha c \partial_x p + c \sigma_x u_x = 0 \\ \partial_t u_y - \alpha c \partial_y p + c \sigma_y u_y = 0 \\ \partial_t u_z - \alpha c \partial_z p + c \sigma_z u_z = 0 \\ \partial_t p_x - \frac{c}{\alpha} \partial_x u_x + c \sigma_x p_x = 0 \\ \partial_t p_y - \frac{c}{\alpha} \partial_y u_y + c \sigma_y p_y = 0 \\ \partial_t (p - p_x - p_y) - \frac{c}{\alpha} \partial_z u_z + c \sigma_z (p - p_x - p_y) = 0. \end{array} \right.$$

Soient $h_0(x, t) = h(x, t) \in H^{7/2, 7/2}(\Sigma)$, $h_1(x, t) = 0 \in H^{5/2, 5/2}(\Sigma)$, $h_2(x, t) = 0 \in H^{3/2, 3/2}(\Sigma)$, $h_3(x, t) \in H^{1/2, 1/2}(\Sigma)$. Soient

$$\begin{cases} \psi_0(x, y, z) = 0 \in H^{7/2}(\Omega) \\ \psi_1(x, y, z) = 0 \in H^{5/2}(\Omega) \\ \psi_2(x, y, z) = 0 \in H^{3/2}(\Omega) \\ \psi_3(x, y, z) = 0 \in H^{1/2}(\Omega). \end{cases}$$

Alors $\psi_k, h_j, 0 \leq k, j \leq 3$ vérifient les conditions de compatibilité 2.47, 2.48 de Lions Magenes [Lions, Magenes, 1968] et il existe un relèvement $P(x, y, z, t)$ dans $H^{4,4}(\mathcal{Q})$, tel que

$$\begin{aligned} \frac{\partial^k P}{\partial t^k}(x, y, z, 0) &= \psi_k(x, y, z), & 0 \leq j, k \leq 3. \\ \frac{\partial^j P}{\partial z^j}(x, y, 0, t) &= h_j(x, y, t), \end{aligned}$$

De plus

$$\exists C_P \in \mathbb{R}^+, \quad \|P\|_{H^{4,4}} \leq C_1 \|h\|_{H^{7/2, 7/2}(\Sigma)}. \quad (5.15)$$

D'autre part, $u_x(x, y, 0, t)$ vérifie

$$\partial_t u_x(x, y, 0, t) + c\sigma_x u_x(x, y, 0, t) = -\alpha c \partial_x h(x, y, t).$$

Ainsi il existe une fonction g dans $H^{5/2, 5/2}(\Sigma)$ telle que $u_x(x, y, 0, t) = g(x, y, t)$ et $g(x, y, 0) = 0$, $\partial_t g(x, y, 0) = 0$.

Soient alors $g_1 = 0 \in H^{3/2, 3/2}(\Sigma)$, $g_2 = 0 \in H^{1/2, 1/2}(\Sigma)$, et

$$\begin{cases} \eta_0(x, y, z) = 0 \in H^{5/2}(\Omega) \\ \eta_1(x, y, z) = 0 \in H^{3/2}(\Omega) \\ \eta_2(x, y, z) = 0 \in H^{1/2}(\Omega). \end{cases}$$

Alors $\eta_k, g_j, 0 \leq k, j \leq 2$ vérifient les conditions de compatibilité 2.47, 2.48 Lions Magenes II et il existe un relèvement $U_x(x, y, z, t)$ dans $H^{3,3}(\mathcal{Q})$, tel que

$$\begin{aligned} \frac{\partial^k U_x}{\partial t^k}(x, y, z, 0) &= \eta_k(x, y, z), & 0 \leq j, k \leq 2. \\ \frac{\partial^j U_x}{\partial z^j}(x, y, 0, t) &= g_j(x, y, t), \end{aligned}$$

De plus,

$$\exists(C_{U_x}) \in \mathbb{R}^+, \quad \|U_x\|_{H^{3,3}} \leq C_{U_x} \|h\|_{H^{7/2,7/2}(\Sigma)}. \quad (5.16)$$

De la même façon, il existe une fonction f dépendant de h telle que u_y est égale à f sur le bord. Il est donc possible de définir $U_y(x, y, z, t)$ tel que

$$\begin{cases} \frac{\partial^k U_y}{\partial t^k}(x, y, z, 0) = 0, & 0 \leq k \leq 2. \\ \frac{\partial^j U_y}{\partial z^j}(x, y, 0, t) = 0, & 1 \leq j \leq 2. \\ U_y(x, y, 0, t) = f(x, y, t), \end{cases}$$

et telle que

$$\exists C_{U_y} \in \mathbb{R}^2, \quad \|U_y\|_{H^{3,3}} \leq C_{U_y} \|h\|_{H^{7/2,7/2}(\Sigma)}. \quad (5.17)$$

Au bord, p_x vérifie également

$$\partial_t p_x(x, y, 0, t) + c\sigma_x(x, y, 0)p_x(x, y, 0, t) = -\frac{c(x, y, 0)}{\alpha(x, y, 0)}\partial_x g(x, y, t).$$

De cette équation se déduit l'existence de $q(x, y, t) \in H^{3/2,3/2}(\Sigma)$ telle que $p_x(x, y, 0, t) = q(x, y, t)$ et $q(x, y, 0) = 0$, $\partial_t q(x, y, 0) = 0$. Soit $q_1 = 0 \in H^{1/2,1/2}(\Sigma)$, alors il existe P_x dans $H^{2,2}(\mathcal{Q})$, tel que,

$$\begin{cases} \frac{\partial^k P_x}{\partial t^k}(x, y, z, 0) = 0 \in H^{k+1/2}(\Omega) & 0 \leq j, k \leq 1, \\ \frac{\partial^j P_x}{\partial z^j}(x, y, 0, t) = q_j(x, t), \end{cases}$$

et ,

$$\exists C_{P_x} \in \mathbb{R}, \quad \|P_x\|_{H^{2,2}} \leq C_{P_x} \|h\|_{H^{7/2,7/2}(\Sigma)}. \quad (5.18)$$

Comme précédemment, il existe également $m(x, y, t) \in H^{3/2,3/2}(\Sigma)$ telle que p_y vaut m sur Σ et il existe P_y dans $H^{2,2}(\mathcal{Q})$, tel que,

$$\begin{cases} \frac{\partial^k P_y}{\partial t^k}(x, y, z, 0) = 0 \in H^{k+1/2}(\Omega), & 0 \leq k \leq 1. \\ \frac{\partial P_y}{\partial z}(x, y, 0, t) = 0 \in H^{1/2,1/2}(\Sigma) \\ P_y(x, y, 0, t) = m(x, y, t), \end{cases}$$

avec

$$\exists C_{P_y} \in \mathbb{R}, \quad \|P_x\|_{H^{2,2}} \leq C_{P_y} \|h\|_{H^{7/2,7/2}(\Sigma)}. \quad (5.19)$$

Finalement, la dernière équation donne sur le bord

$$\begin{aligned} \partial_z u_z(x, y, 0, t) &= \frac{\alpha(x, y, 0)}{c(x, y, 0)} (\partial_t q(x, y, t) + \partial_t m(x, y, t) - \partial_t h(x, y, t)) \\ &\quad + \alpha(x, y, 0) \sigma_z(0) (q(x, y, t) + m(x, y, t) - h(x, y, t)) = s(x, y, t) \end{aligned}$$

Ainsi, $\exists U_z \in H^{2,2}(\mathcal{Q})$ tel que

$$\left\{ \begin{array}{l} \frac{\partial^k U_z}{\partial t^k}(x, y, z, 0) = 0 \in Hk+1/2(\Omega), \quad 0 \leq k \leq 1 \\ \frac{\partial U_z}{\partial z}(x, y, 0, t) = s(x, y, t) \in H^{1/2,1/2}, \\ U_z(x, y, 0, t) = 0 \in H^{1/2,1/2}, \end{array} \right.$$

et

$$\exists C_{U_z} \in \mathbb{R}, \quad \|U_z\|_{H^{2,2}} \leq C_{U_z} \|h\|_{H^{7/2,7/2}(\Sigma)}. \quad (5.20)$$

Les variables $\bar{u}_x = u_x - U_x, \bar{u}_y = u_y - U_y, \bar{u}_z = u_z - U_z, \bar{p}_x = p_x - P_x, \bar{p}_y = p_y - P_y, \bar{p} = p - P$ vérifient

$$\left\{ \begin{array}{l} \partial_t \bar{u}_x - \alpha c \partial_x \bar{p} + c \sigma_x \bar{u}_x = -\alpha c \partial_x P + \partial_t U_x + c \sigma_x U_x \\ \partial_t \bar{u}_y - \alpha c \partial_y \bar{p} + c \sigma_y \bar{u}_y = -\alpha c \partial_y P + \partial_t U_y + c \sigma_y U_y \\ \partial_t \bar{u}_z - \alpha c \partial_z \bar{p} + c \sigma_z \bar{u}_z = -\alpha c \partial_z P + \partial_t U_z + c \sigma_z U_z \\ \partial_t \bar{p}_x - \frac{c}{\alpha} \partial_x \bar{u}_x + c \sigma_x \bar{p}_x = -\frac{c}{\alpha} \partial_x U_x + \partial_t P_x + c \sigma_x P_x \\ \partial_t \bar{p}_y - \frac{c}{\alpha} \partial_y \bar{u}_y + c \sigma_y \bar{p}_y = -\frac{c}{\alpha} \partial_y U_y \partial_t P_y + c \sigma_y P_y \\ \partial_t (\bar{p} - \bar{p}_x - \bar{p}_y) - \frac{c}{\alpha} \partial_z \bar{u}_z + c \sigma_z (\bar{p} - \bar{p}_x - \bar{p}_y) = \partial_t (P - P_x - P_y) + c \sigma_z (P - P_x - P_y) - \frac{c}{\alpha} \partial_z U_z. \end{array} \right.$$

En réorganisant les trois dernières équations par l'introduction de

$$\bar{p}_z = (p - \bar{p}_x - \bar{p}_y) \quad \text{et} \quad \bar{v} = \sigma_x \bar{p}_x + \sigma_y \bar{p}_y,$$

et en définissant

$$P_z = P - P_x - P_y, \quad V = \sigma_x P_x + \sigma_y P_y,$$

on retrouve le système

$$\left\{ \begin{array}{l} \partial_t \bar{u}_x - \alpha c \partial_x \bar{p} + c \sigma_x \bar{u}_x = -\alpha c \partial_x P + \partial_t U_x + c \sigma_x U_x \\ \partial_t \bar{u}_y - \alpha c \partial_y \bar{p} + c \sigma_y \bar{u}_y = -\alpha c \partial_y P + \partial_t U_y + c \sigma_y U_y \\ \partial_t \bar{u}_z - \alpha c \partial_z \bar{p} + c \sigma_z \bar{u}_z = -\alpha c \partial_z P + \partial_t U_z + c \sigma_z U_z \\ \partial_t \bar{p} - \frac{c}{\alpha} \partial_x \bar{u}_x - \frac{c}{\alpha} \partial_y \bar{u}_y - \frac{c}{\alpha} \partial_z \bar{u}_z - c^2 \sigma_x \sigma_y (\bar{p} - \bar{v}) = \\ -\frac{c}{\alpha} \partial_x U_x - \frac{c}{\alpha} \partial_y U_y - \frac{c}{\alpha} \partial_z U_z + \partial_t P - c^2 \sigma_x \sigma_y (P - V) \\ \partial_t \bar{v} - \sigma_x \frac{c^2}{\alpha} \partial_x \bar{u}_x - \sigma_x \frac{c^2}{\alpha} \partial_y \bar{u}_y - c^2 (\sigma_x + \sigma_y) \bar{v} = \partial_t V - \sigma_x \frac{c^2}{\alpha} \partial_x U_x - \sigma_x \frac{c^2}{\alpha} U_y - c^2 (\sigma_x + \sigma_y) V \\ \partial_t \bar{p}_z - \frac{c}{\alpha} \partial_z \bar{u}_z + c \sigma_z (\bar{p}_z) = \partial_t P_z + c \sigma_z P_z - \frac{c}{\alpha} \partial_z U_z. \end{array} \right.$$

Remarque : comme \bar{p}, \bar{p}_x et \bar{p}_y s'annulent sur le bord, \bar{v} et \bar{p}_z s'annulent aussi sur le bord.

En posant

$$\bar{U}_1 = \begin{bmatrix} \bar{u}_x \\ \bar{u}_y \\ \bar{u}_z \\ \bar{p} \end{bmatrix}, \quad \bar{U}_2 = \partial_x U^1, \quad \bar{U}_3 = \partial_y U^1, \quad \bar{U}_4 = \partial_z U^1, \quad \bar{U}_5 = \begin{bmatrix} 0 \\ 0 \\ 0 \\ \bar{v} \end{bmatrix}, \quad \bar{U}_6 = \begin{bmatrix} 0 \\ 0 \\ 0 \\ \bar{p}_z \end{bmatrix},$$

$$F_1 = \begin{bmatrix} -\alpha c \partial_x P + \partial_t U_x + c \sigma_x U_x \\ -\alpha c \partial_y P + \partial_t U_y + c \sigma_y U_y \\ -\alpha c \partial_z P + \partial_t U_z + c \sigma_z U_z \\ -\frac{c}{\alpha} \partial_x U_x - \frac{c}{\alpha} \partial_y U_y - \frac{c}{\alpha} \partial_z U_z + \partial_t P - c^2 \sigma_x \sigma_y (P - V) \end{bmatrix},$$

$$F_2 = \begin{bmatrix} 0 \\ 0 \\ 0 \\ \partial_t V - \sigma_x \frac{c^2}{\alpha} \partial_x U_x - \sigma_x \frac{c^2}{\alpha} U_y - c^2 (\sigma_x + \sigma_y) V \end{bmatrix},$$

$$F_3 = \begin{bmatrix} 0 \\ 0 \\ 0 \\ \partial_t P_z + c \sigma_z P_z - \frac{c}{\alpha} \partial_z U_z \end{bmatrix},$$

il vient

$$\partial_t \overline{U} - \tilde{A} \partial_x \overline{U} - \tilde{B} \partial_y \overline{U} - \tilde{C} \partial_z \overline{U} + S \overline{U} = F, \quad \text{avec} \quad \overline{U} = \begin{bmatrix} \overline{U}_1 \\ \vdots \\ \overline{U}_6 \end{bmatrix}, \quad \overline{F} = \begin{bmatrix} \overline{F}_1 \\ \partial_x \overline{F}_1 \\ \partial_y \overline{F}_1 \\ \partial_z \overline{F}_1 \\ \overline{F}_2 \\ \overline{F}_3 \end{bmatrix}.$$

Soit $W = (W^1, W^2, W^3, W^4, W^5, W^6)$ avec

$$W^1 = (U_x, U_y, U_z, P), \quad W^2 = \partial_x W^1, \quad W^3 = \partial_y W^1, \quad W^4 = \partial_z W^1,$$

$$W^5 = (0, 0, 0, V), \quad W^6 = (0, 0, 0, P_z),$$

Alors $\overline{U} = \tilde{U} - W$ et la définition des relèvements garantit les propriétés (P1) et (P2). D'autre part, (P3) et (P4) se déduisent des propriétés des relèvements (5.15),(5.16),(5.17),(5.18),(5.19),(5.20).

5.1.5 Théorème 2 : Problème de transmission pour un modèle d'impédance discontinu 1D

5.1.5.1 Enoncé

Soit $\Omega_1 = [-\infty; 0] \times \mathbb{R}^2$, $\Omega_2 = [0, +\infty] \times \mathbb{R}^2$, $\Omega = \Omega_1 \cup \Omega_2$, et $\Gamma = \{(x, y, z) \in \mathbb{R}^3, z = 0\}$ l'interface entre les sous-domaines Ω_1 et Ω_2 .

Pour $(\sigma_x, \sigma_y, \sigma_z) \in (W^{1,\infty}(\Omega))^3$, $c \in C^\infty(\Omega)$,

$$0 < c_{min} \leq c(x, y, z) \leq c_{max}, \tag{5.21}$$

et pour un modèle d'impédance 1D tel que

$$\alpha(x, y, z) = \begin{cases} \alpha_1(z) & \text{pour } z < 0 \\ \alpha_2(z) & \text{pour } z > 0, \end{cases}$$

avec

$$0 < \alpha_{min} \leq \alpha_i(z) \leq \alpha_{max}, \quad i = 1, 2$$

et

$$|\alpha_2(0^+) - \alpha_1(0^-)| > 0,$$

alors pour toute solution suffisamment régulière de (5.1), (5.2)

$$\begin{aligned} & \| (u_x, u_y, u_z, p) \|_{L^\infty(0,T;H^1(\Omega))}^2 + \| (c(\sigma_x p_x + \sigma_y p_y)) \|_{L^\infty(0,T;L^2(\Omega))}^2 + \| p_z \|_{L^\infty(0,T;L^2(\Omega))}^2 \leq \\ & \frac{N'}{M'} e^{\frac{(\gamma'+1)T}{2M'^2}} \left(\| (u_x, u_y, u_z, p)(., ., ., 0) \|_{H^1(\Omega)}^2 + \| c(\sigma_x p_x + \sigma_y p_y)(., ., ., 0) \|_{L^2(\Omega)}^2 + \| p_z(., ., ., 0) \|_{L^2(\Omega)}^2 \right) \end{aligned} \quad (5.22)$$

avec

$$\begin{cases} \gamma' = \max \{ \gamma(\alpha_1, c), \gamma(\alpha_2, c) \}, \\ M' = \min \{ M(\alpha_1, c), M(\alpha_2, c) \}, \\ N' = \min \{ N(\alpha_1, c), N(\alpha_2, c) \} \end{cases} \quad (5.23)$$

où $\gamma(\alpha_i, c), M(\alpha_i, c), N(\alpha_i, c)$ sont définis dans la partie précédente.

5.1.5.2 Démonstration

A chacun des demi-espaces Ω_i est associé un symétriseur $H_i(\alpha_i)$, $i = 1, 2$. Chacun de ces symétriseurs engendre un produit scalaire sur le demi-espace auquel il est associé, noté $\langle ., . \rangle_i$, $i = 1, 2$, tel que

$$\forall (U, V) \in (L^2(\Omega_i))^{24} \times (L^2(\Omega_i))^{24}, \quad \langle U, V \rangle_i = (U, H_i V)_{L^2(\Omega_i)}, \quad i = 1, 2.$$

La norme associée à ces produits scalaires est notée $|.|_i$ et vérifie

$$\forall U \in L^2(\Omega_i), \quad M_i \|U\|_{L^2(\Omega_i)}^2 \leq |U|_i^2 \leq N_i \|U\|_{L^2(\Omega_i)}^2$$

La restriction de la solution \tilde{U} des équations PML (5.6) au domaine i est notée \tilde{U}_i . Les matrices suivantes, symétriques par construction, sont définies

$$\check{A}_i = H_i \tilde{A}_i, \quad \check{B}_i = H_i \tilde{B}_i, \quad \check{C}_i = H_i \tilde{C}_i.$$

La quantité $|\tilde{U}_1(t)|_1^2 + |\tilde{U}_2(t)|_2^2$ vérifie alors, par intégrations par parties,

$$\begin{aligned} \frac{d}{dt} \left(|\tilde{U}_1(t)|_1^2 + |\tilde{U}_2(t)|_2^2 \right) &= -2 \int_{\Omega_1} \left(\tilde{U}_1, \check{S}_1 \tilde{U}_1 \right) d\Omega_1 - 2 \int_{\Omega_2} \left(\tilde{U}_2, \check{S}_2 \tilde{U}_2 \right) d\Omega_2 \\ &- \int_{\Omega_1} \left(\tilde{U}_1, \partial_x(\check{A}_1) \tilde{U}_1 + \partial_y(\check{B}_1) \tilde{U}_1 + \partial_z(\check{C}_1) \tilde{U}_1 \right) d\Omega_1 - \int_{\Gamma} \left(\check{C}_1 \tilde{U}_1(x, y, 0^-, t), \tilde{U}_1(x, y, 0^-, t) \right) d\Gamma \\ &- \int_{\Omega_2} \left(\tilde{U}_2, \partial_x(\check{A}_2) \tilde{U}_2 + \partial_y(\check{B}_2) \tilde{U}_2 + \partial_z(\check{C}_2) \tilde{U}_2 \right) d\Omega_2 + \int_{\Gamma} \left(\check{C}_2 \tilde{U}_2(x, y, 0^+, t), \tilde{U}_2(x, y, 0^+, t) \right) d\Gamma. \end{aligned}$$

La preuve du théorème 2 repose sur l'analyse précise du terme de saut à l'interface. Celui-ci est réécrit de la manière suivante

$$\begin{aligned} S(t) &= \int_{\Gamma} \left(\check{C}_2 \tilde{U}_2(x, y, 0^+, t), \tilde{U}_2(x, y, 0^+, t) \right) d\Gamma - \int_{\Gamma} \left(\check{C}_1 \tilde{U}_1(x, y, 0^-, t), \tilde{U}_1(x, y, 0^-, t) \right) d\Gamma = \\ &\int_{\Gamma} \left([2c^2 u_z p] + [2c^2 \partial_x u_z \partial_x p] + [2c^2 \partial_y u_z \partial_y p] + [2c^2 \partial_z u_z \partial_z p] - [2\alpha c \partial_z p v] - [2\alpha c^2 \sigma_z p_z \partial_z p] \right) d\Gamma, \end{aligned}$$

en utilisant la notation

$$\forall u \in L^2(\Omega), \quad t \in [0, T], \quad [u] = u|_{\Omega_2}(x, y, 0^+, t) - u|_{\Omega_1}(x, y, 0^-, t).$$

La donnée fondamentale du problème est la continuité à l'interface des quantités suivantes : la vitesse de propagation des ondes $c(x, y, z)$, fonction C^∞ sur tout le domaine Ω , la vitesse normale à l'interface $u_z(x, y, z)$ et le champ de pression $p(x, y, z)$. Ceci implique que les trois premiers termes de saut sont nuls. En effet ces termes font intervenir le produit de ces quantités et de leurs dérivées tangentes par rapport à l'interface $z = 0$, qui sont toutes continues. De plus, au bord, $\sigma_z(0) = 0$. En effet, le modèle d'impédance est prolongé par continuité dans la couche PML. Ainsi, il n'y a pas de discontinuités suivant la direction z dans la couche PML inférieure pour laquelle $\sigma_z(z) > 0$. Puisqu'ici, une discontinuité suivant z est définie en $z = 0$, ceci implique qu'elle est définie hors de la couche PML inférieure. Ceci annule le dernier terme de saut. Il reste alors à maîtriser le saut de $\partial_z u_z \partial_z p$ et celui de $\alpha \partial_z p v$.

Remarquons d'abord que sur le bord, l'égalité suivante est vérifiée

$$\partial_t u_z = \alpha c \partial_z p.$$

Ainsi, $\alpha \partial_z p$ est continue, et le saut de cette quantité est nul. Le terme de saut à l'interface peut donc être simplifié de la manière suivante

$$S(t) = \int_{\Gamma} 2c(x, y, 0)(\alpha \partial_z p)(x, y, 0, t) \left[\frac{c}{\alpha} \partial_z u_z \right] - (\alpha \partial_z p)(x, y, 0, t)[\sigma_x p_x + \sigma_y p_y].$$

De plus, suivant les équations (5.4),

$$\partial_t p_z = \frac{c}{\alpha} \partial_z u_z, \quad \text{et} \quad \partial_t p_z = \partial_t(p - p_x - p_y).$$

Ainsi, le saut de $\partial_t p$ étant nul, le saut total peut s'exprimer en fonction de celui de p_x et p_y et de leurs dérivées temporelles

$$S(t) = - \int_{\Gamma} 2(\alpha c \partial_z p)(x, y, 0, t) [\partial_t p_x + c \sigma_x p_x + \partial_t p_y + c \sigma_y p_y] d\Gamma.$$

Or, p_x et p_y vérifient

$$\begin{cases} \partial_t p_x - \frac{c}{\alpha} \partial_x u_x + \sigma_x p_x = 0 \\ \partial_t p_y - \frac{c}{\alpha} \partial_y u_y + \sigma_y p_y = 0. \end{cases}$$

D'autre part,

$$\begin{cases} \partial_t \frac{u_x}{\alpha} + \sigma_x \frac{u_x}{\alpha} = c \partial_x p \\ \partial_t \frac{u_y}{\alpha} + \sigma_x \frac{u_y}{\alpha} = c \partial_y p. \end{cases}$$

La continuité de p impliquant la continuité de ses dérivées tangentielles à l'interface, les quantités

$$\frac{u_x}{\alpha}, \quad \frac{u_y}{\alpha},$$

sont-elles aussi continues à l'interface.

Ceci implique également la continuité de leurs dérivées tangentielles, qui s'expriment, en se servant de l'hypothèse que le modèle d'impédance ne dépend que de z ,

$$\partial_x \left(\frac{u_x}{\alpha} \right) = \frac{\partial_x u_x}{\alpha} \text{ et } \partial_y \left(\frac{u_y}{\alpha} \right) = \frac{\partial_y u_y}{\alpha}.$$

Aussi,

$$\begin{cases} \partial_t p_x + \sigma_x p_x = c \partial_x \left(\frac{u_x}{\alpha} \right) \\ \partial_t p_y + \sigma_y p_y = c \partial_y \left(\frac{u_y}{\alpha} \right). \end{cases}$$

Les quantités p_x et p_y sont donc continues à l'interface. Il en va donc de même pour leur dérivées temporelles, et le terme de saut apparaissant dans les intégrations par parties s'annule. La suite se déduit simplement.

Il vient, par Cauchy-Schwartz,

$$\frac{d}{dt} \left(|\tilde{U}_1(t)|_1^2 + |\tilde{U}_2(t)|_2^2 \right) \leq \frac{\gamma'}{M'} \left(|\tilde{U}_1(t)|_1^2 + |\tilde{U}_2(t)|_2^2 \right).$$

Il est donc possible d'appliquer le lemme de Gronwall, qui donne

$$|\tilde{U}_1(t)|_1^2 + |\tilde{U}_2(t)|_2^2 \leq e^{\frac{\gamma' t}{M'}} \left(|\tilde{U}_1(0)|_1^2 + |\tilde{U}_2(0)|_2^2 \right).$$

En utilisant les inégalités liés aux produits scalaires $\langle \cdot, \cdot \rangle_i$, il vient

$$\|\tilde{U}_1(t)\|_{L^2(\Omega_1)}^2 + \|\tilde{U}_2(t)\|_{L^2(\Omega_2)}^2 \leq \frac{N'}{M'} e^{\frac{\gamma' t}{M'}} \left(\|\tilde{U}_1(0)\|_{L^2(\Omega_1)}^2 + \|\tilde{U}_2(0)\|_{L^2(\Omega_2)}^2 \right),$$

c'est à dire, en prenant les bornes supérieures en temps sur l'intervalle $[0, T]$,

$$\|\tilde{U}\|_{L^\infty(0,T;L^2(\Omega))}^2 + \frac{N'}{M'} e^{\frac{\gamma' T}{M'}} \|\tilde{U}(0)\|_{L^2(\Omega)}^2.$$

5.1.6 Conclusion de l'analyse mathématique

Les équations PML de Bérenger adaptées au cas des équations d'Euler 3D en milieu borné hétérogène, mais ne présentant pas de discontinuités, sont donc faiblement bien posées de défaut 1. L'extension des travaux de [Petit-Bergez, 2006], dans ce contexte précis, a donc été réussie. En revanche, en ce qui concerne des milieux pouvant présenter des discontinuités d'impédance, comme c'est le cas dans l'application traitée, les résultats ne sont valables que pour un milieu parfaitement stratifié. Dans le cadre de l'extension à des milieux 3D, il devient difficile de borner le terme de saut apparaissant dans les estimations d'énergie, et il est impossible de conclure. Une autre méthode devra sans doute être mise en oeuvre pour la généralisation de ce résultat.

La partie suivante est maintenant consacrée aux expérimentations numériques qui ont été réalisées sur ces équations PML pour l'acoustique.

5.2 Résultats numériques

5.2.1 Description des expériences

L'objectif de ces expériences numériques est de proposer une analyse de l'efficacité de la couche absorbante dans le contexte particulier de l'inversion de données sismiques de puits 2D. Dans un premier temps, cette analyse consistera à mettre en évidence les difficultés soulevées par le caractère hétérogène du milieu, en comparant l'erreur de modélisation engendrée par l'utilisation de la couche PML pour des milieux de propagation de complexité croissante : modèle homogène, modèle à deux couches, et modèle hétérogène.

Le contexte utilisé est le même que celui du cas test présenté dans le chapitre 4. Ceci concerne la taille du domaine d'intérêt, le nombre (respectivement S et Q) de sources et de récepteurs, et leurs positions, les pas de discréttisation spatiaux et temporel, la forme de la source utilisée (signal de Ricker $r(t)$ centré sur la fréquence $f = 25$ Hz), et le temps total d'enregistrement T . En voici un bref rappel :

$$\left\{ \begin{array}{l} \mathcal{O} = [-360, 360] \times [0; 3400], \quad T = 2 \text{ s}, \\ S = 30, \text{ réparties tous les } 24 \text{ m} \\ Q = 100, \text{ répartis tous les } 8 \text{ m à partir de } 1000 \text{ m, en } x = 0 \text{ m.} \\ r(t) = C \left(t - \frac{1}{f} \right) \exp \left(-\Pi^2 f^2 \left(t - \frac{1}{f} \right)^2 \right), \text{ où } C \text{ est une constante} \end{array} \right.$$

La taille de la couche PML est fixée à 20 points de discréttisation. En ce qui concerne le choix opéré pour les coefficients d'absorption $\sigma_x(x)$ et $\sigma_z(z)$, qui n'a pour l'instant pas été précisé, la formulation mise au point par [Hu, 1996] est utilisée. Elle correspond à définir σ_x et σ_z tels que

$$\sigma_x(x) = \sigma_{0,x} \left(\frac{d_x(x)}{D_x} \right)^n, \quad \sigma_z(z) = \sigma_{0,z} \left(\frac{d_z(z)}{D_z} \right)^n$$

avec $\sigma_{0,x} = \frac{8(n+1)}{D_x}$, $\sigma_{0,z} = \frac{8(n+1)}{D_z}$, $n = 3$,

où D_x et D_z représentent l'épaisseur des PML respectivement suivant x et z , tandis que les fonctions $d_x(x)$ et $d_z(z)$ mesurent la distance séparant le point courant du début de la couche PML, respectivement suivant x et z .

Pour évaluer l'erreur induite par l'utilisation de la couche PML, l'expérience témoin consiste à calculer la propagation des ondes dans un domaine d'extension infinie (latéralement et en profondeur), c'est-à-dire suffisamment grand pour que les réflexions sur ses frontières n'aient pas le temps de revenir dans le domaine

d'intérêt pour un temps de propagation égal à T . Les modèles initiaux de vitesse et d'impédance définis sur le domaine d'intérêt \mathcal{O} sont prolongés par continuité dans les direction x et $z > 0$ de manière à définir un modèle "demi-espace" englobant le domaine dans lequel la propagation est étudiée.

Les comparaisons portent sur l'erreur maximale en norme L^∞ commises sur l'évaluation du champ de vitesse de déplacement horizontal $u_x(x, z, t)$ et vertical $u_z(x, z, t)$. Soient $u_{x,\infty}(x, z, t)$ et $u_{z,\infty}(x, z, t)$ les champs de vitesse de déplacement obtenus dans le domaine d'extension infini, et $u_{x,PML}(x, z, t)$ et $u_{z,PML}(x, z, t)$ ceux correspondant au domaine bordé de la couche PML. Soit de plus $e_x(x, z, t)$ et $e_z(x, z, t)$ l'erreur commise dans \mathcal{O} au temps t ,

$$e_x(x, z, t) = |u_{x,\infty}(x, z, t) - u_{x,PML}(x, z, t)|,$$

$$e_z(x, z, t) = |u_{z,\infty}(x, z, t) - u_{z,PML}(x, z, t)|.$$

L'erreur qui est mesurée est l'erreur en norme L^∞ induite par la modélisation PML,

$$\sup_{(x,z) \in \mathcal{O}} |e_x(x, z, t)|, \quad \sup_{(x,z) \in \mathcal{O}} |e_z(x, z, t)|.$$

Il est de plus intéressant d'étudier la répartition spatiale de cette erreur. Soit t_m le temps pour lequel l'erreur en norme L^∞ est maximale. Une carte spatiale de l'erreur $e_x(x, z, t_m)$ et $e_z(x, z, t_m)$ est également donnée. Ceci permet de donner à la fois une vision de l'évolution au cours du temps de l'erreur maximale commise, et de sa répartition dans le domaine \mathcal{O} .

Dans un deuxième temps, la différence en norme $L^2(0, T)$ des signaux enregistrés aux récepteurs (vitesses de déplacement horizontal et vertical) sera calculée, pour le milieu hétérogène correspondant au cas test du chapitre 4. Cette différence donne une indication sur le biais induit par la modélisation PML. Si elle est plus élevée que les résidus sismiques obtenus à l'issue de l'inversion des données, ceci signifie que les différences introduites dans les données par la modélisation PML sont prises en compte par l'algorithme d'inversion. Cette situation doit être évitée, et il est possible de montrer qu'avec une couche PML de 20 points de discrétisation, le biais de modélisation reste trop faible pour être pris en compte.

5.2.2 Figures résultats

On s'intéresse tout d'abord au cas d'un modèle totalement homogène (α et c sont constants). La figure 5.1 donne l'évolution temporelle de l'erreur maximale commise pour les vitesses de déplacement horizontal et vertical.

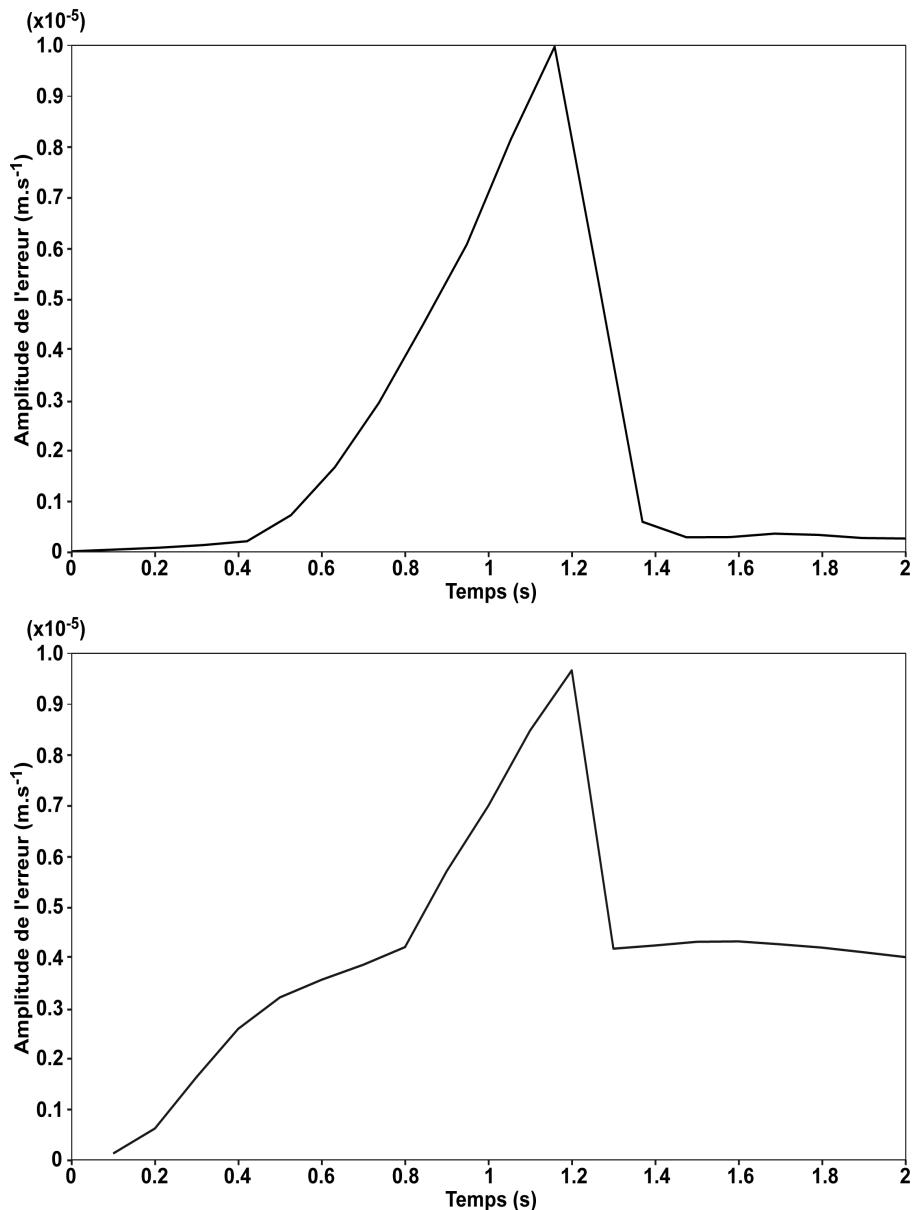


FIG. 5.1 – Erreur en norme L^∞ pour les vitesses de déplacement vertical (haut) et horizontal (bas) dans le cas d'un milieu homogène

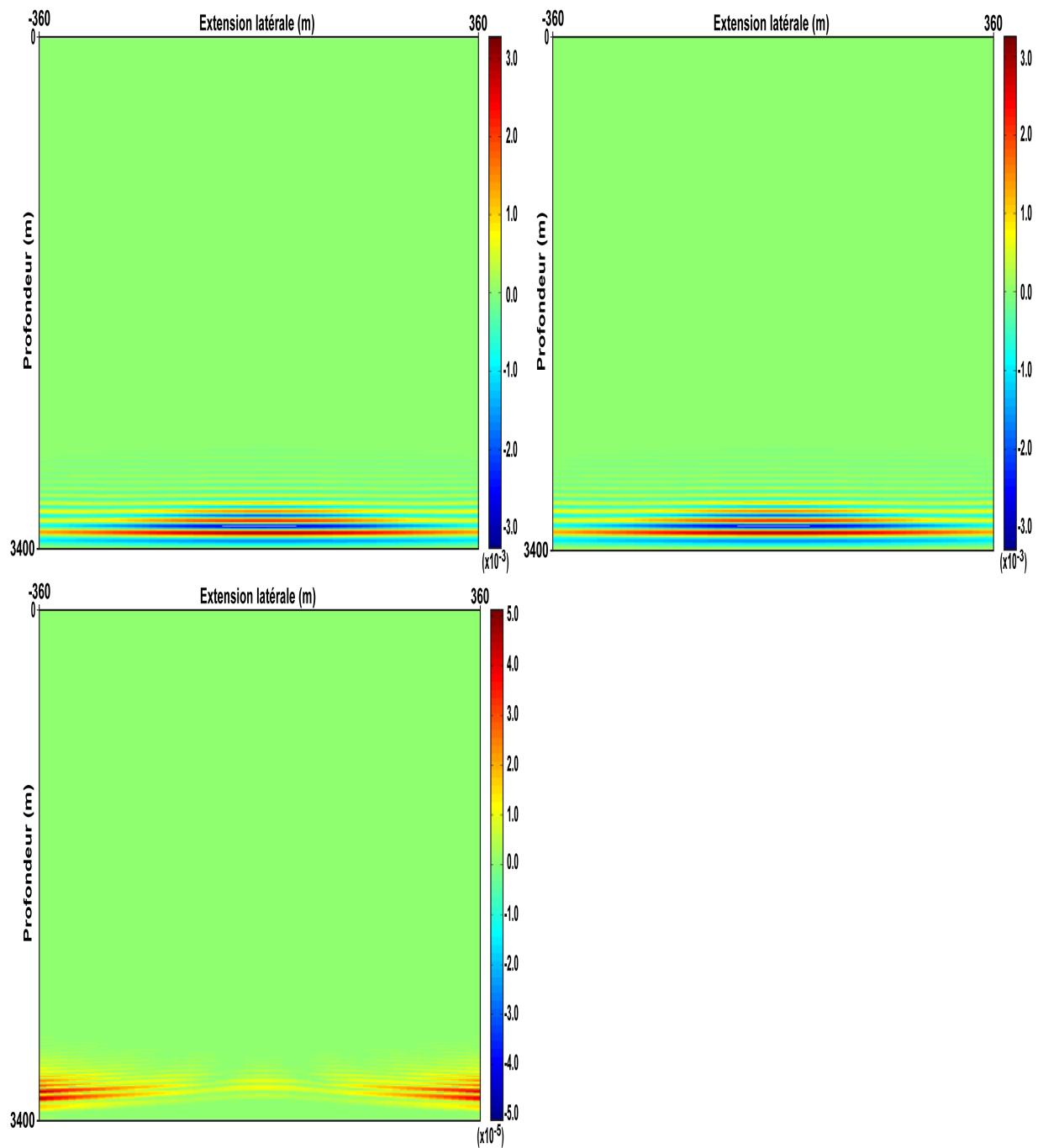


FIG. 5.2 – Champ de vitesse de déplacement vertical pour un modèle homogène à $t=1.2$ s. Milieu infini (haut gauche), milieu PML (haut droite), différence (bas).

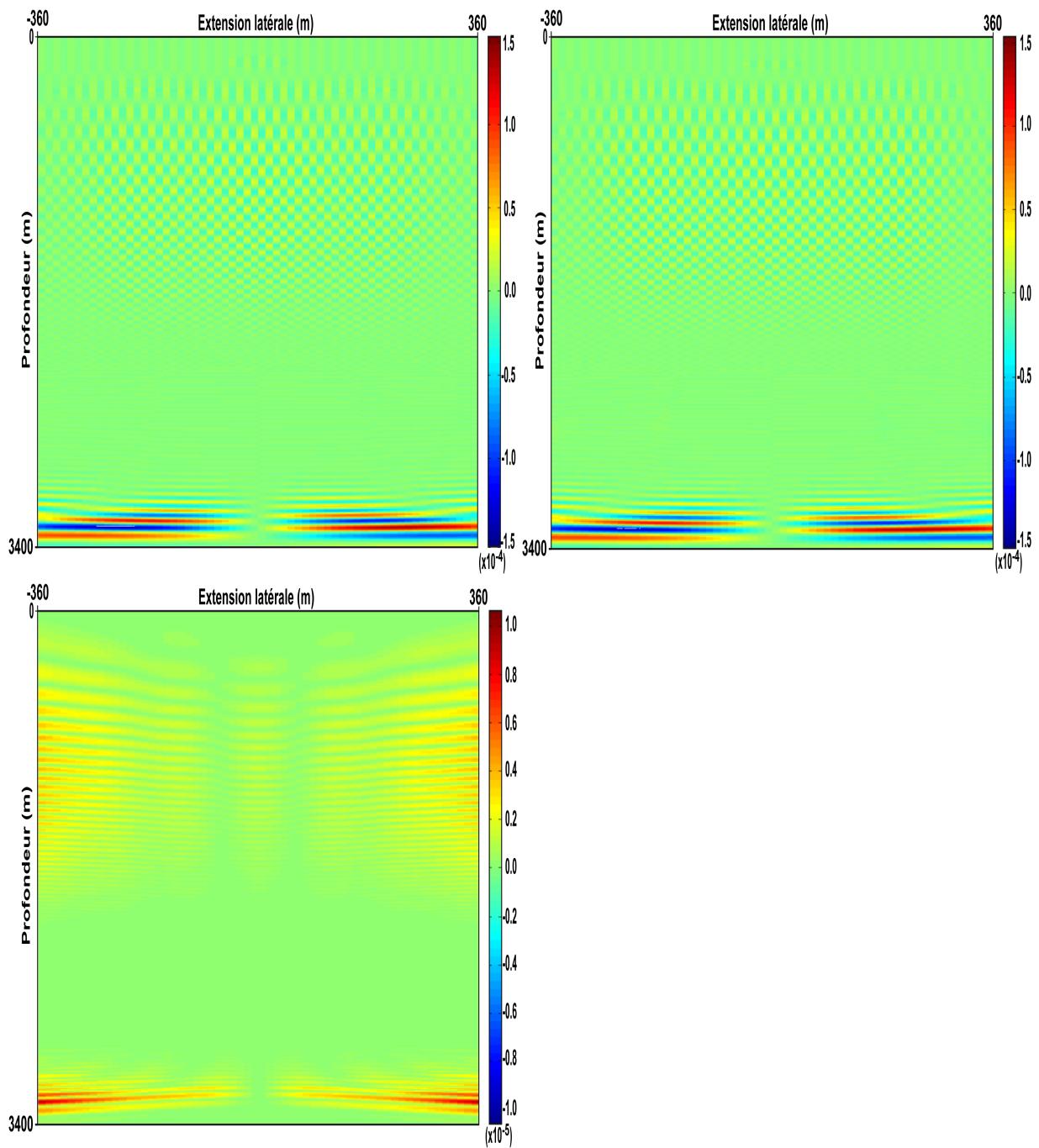


FIG. 5.3 – Champ de vitesse de déplacement horizontal pour un modèle homogène à $t=1.2$ s. Milieu infini (haut gauche), milieu PML (haut droite), différence (bas).

L'erreur maximale enregistrée s'élève à 4.5×10^{-5} pour la vitesse de déplacement vertical et 1×10^{-5} pour la vitesse de déplacement horizontal. Cette erreur maximale est atteinte aux alentours de $t = 1, 2$ s. La représentation spatiale de cette erreur pour $t = 1.2$ s est donnée par les figures 5.2 et 5.3

Deux enseignements sont à tirer de cette expérience. Tout d'abord, l'amplitude de l'erreur maximale est assez faible : de l'ordre de 1% de l'amplitude maximale du signal, pour la vitesse de déplacement vertical et, 7% pour la vitesse de déplacement horizontal. De plus, la localisation de cette erreur correspond aux zones où l'amplitude des ondes est la plus forte, et décroît significativement depuis les bords du domaine d'intérêt vers l'intérieur. Ceci représente un avantage pour l'application à l'inversion de données sismiques de puits, la partie la plus importante du signal étant enregistrée au niveau des récepteurs, placés au centre du domaine, zone où l'erreur est faible. Ceci témoigne cependant du caractère imparfaitement absorbant de la couche PML : une fraction du signal incident est réfléchie dans le domaine d'intérêt.

Cette expérience est réitérée dans le cadre d'un modèle à deux couches. Une discontinuité d'impédance est introduite au milieu du domaine, à la profondeur $z = 1700$ m. Les figures 5.4, 5.5, 5.6 et 5.7 présentent les résultats obtenus.

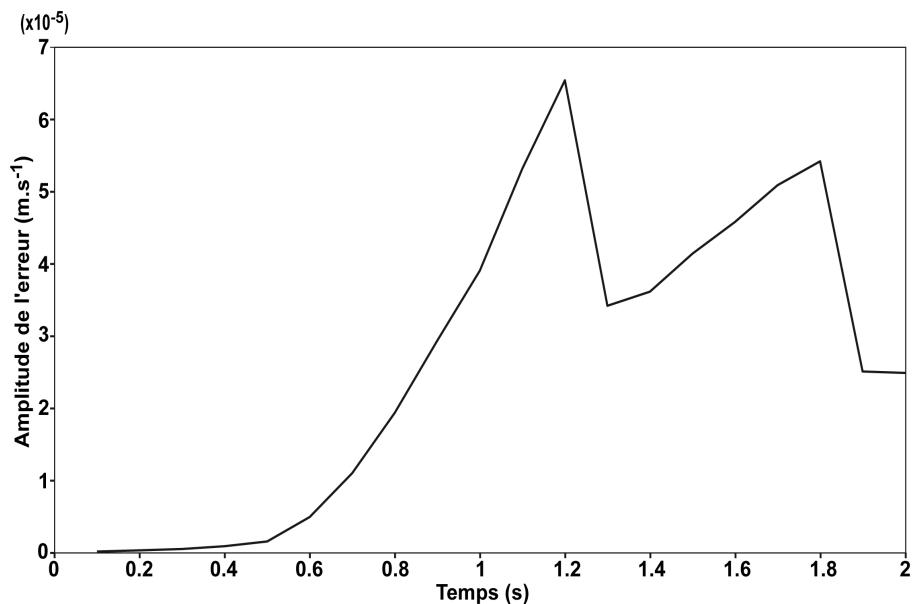


FIG. 5.4 – Erreur en norme L^∞ pour la vitesse de déplacement vertical dans le cas d'un milieu à deux couches.

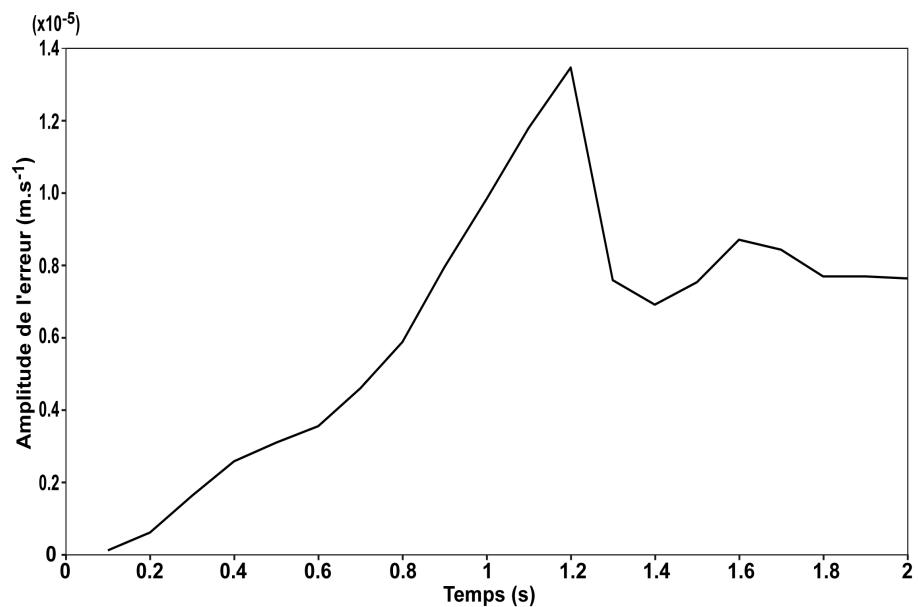


FIG. 5.5 – Erreur en norme L^∞ pour la vitesse de déplacement horizontal dans le cas d'un milieu à deux couches.

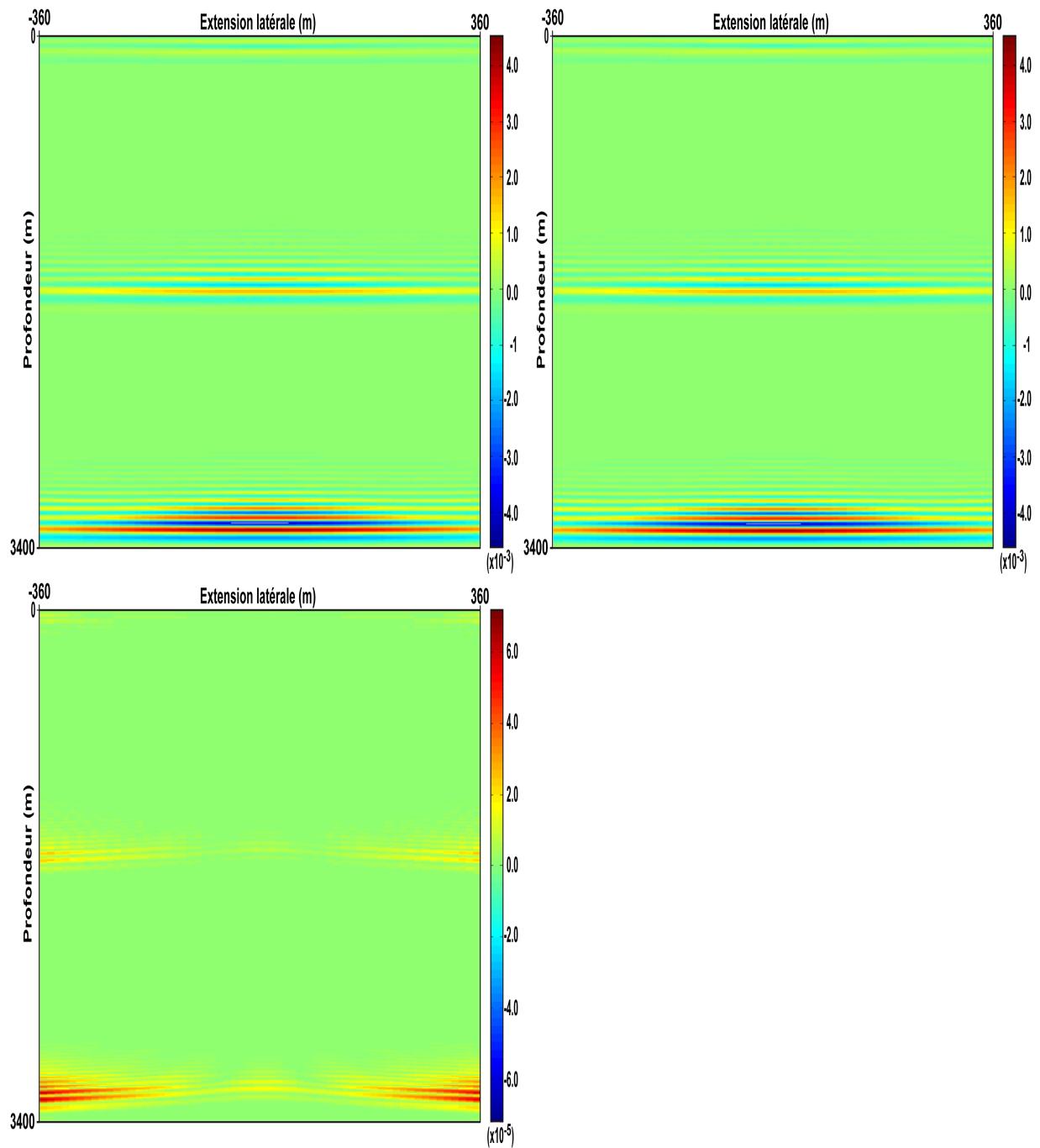


FIG. 5.6 – Champ de vitesse de déplacement vertical pour un modèle à deux couches à $t = 1, 2$ s. Milieu infini (haut gauche), milieu PML (haut droite), différence (bas).

Le temps pour lequel l'amplitude de l'erreur maximale commise est la plus forte se situe à nouveau autour de $t = 1.2$ s. Le rapport de l'amplitude de l'erreur à l'amplitude maximale du signal augmente faiblement, 1,5% pour la vitesse de déplacement vertical, et 7% pour la vitesse de déplacement horizontal. La localisation de l'erreur

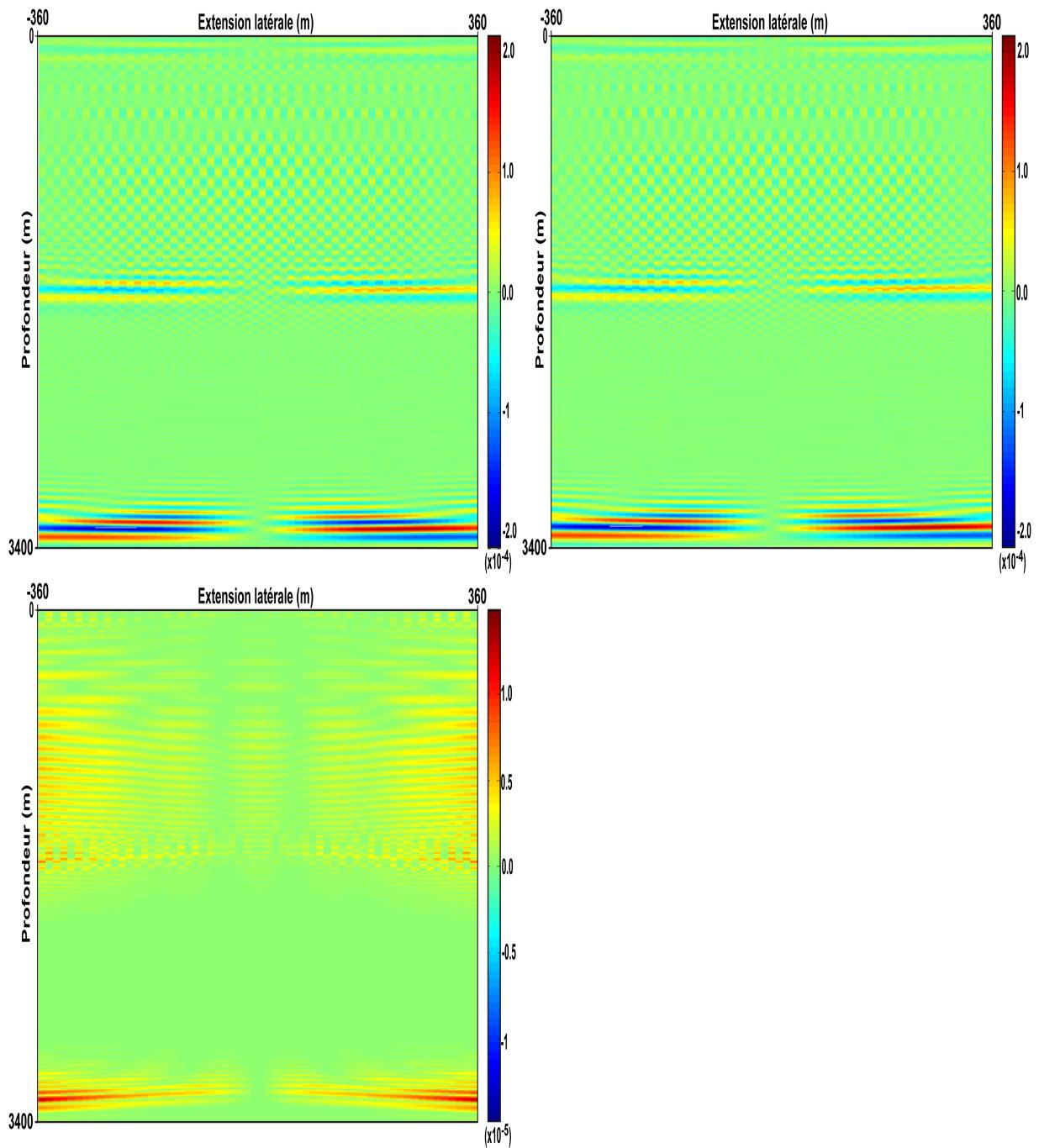


FIG. 5.7 – Champ de vitesse de déplacement horizontal pour un modèle à deux couches à $t = 1, 2$ s. Milieu infini (haut gauche), milieu PML (haut droite), différence (bas).

respecte également les règles esquissées pour l'exemple précédent : elle suit la localisation du signal et est surtout concentrée près des frontières du domaine d'intérêt \mathcal{O} . Les conclusions sont donc les mêmes que celles de l'exemple précédent.

On s'intéresse maintenant au cas du modèle très hétérogène utilisé pour réaliser les expériences numériques du chapitre 4 (fig.4.1). L'expérience précédente est réalisée à nouveau, avec ces modèles de vitesse et d'impédance. Les figures 5.8, 5.9 et 5.10 présentent les résultats obtenus.

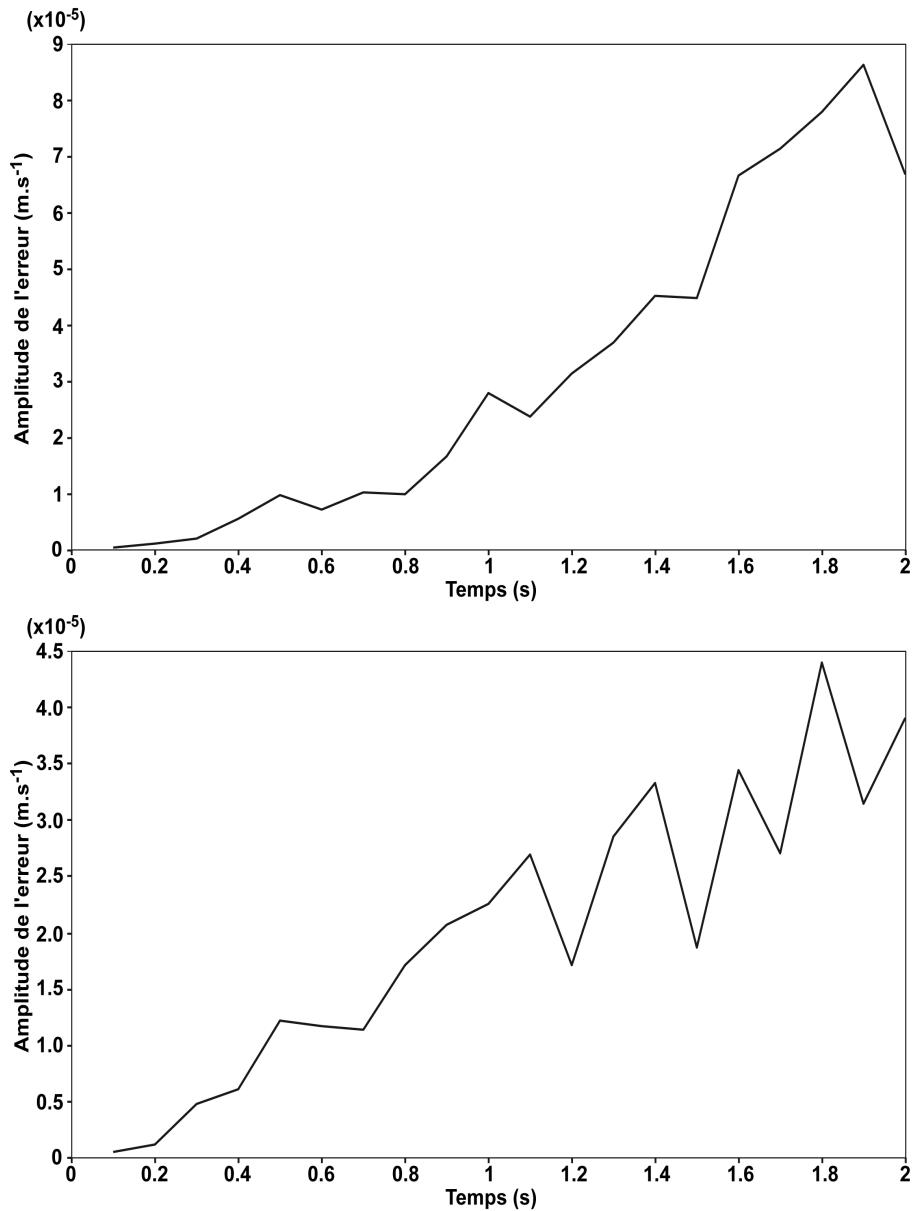


FIG. 5.8 – Erreur en norme L^∞ pour les vitesses de déplacement vertical (haut) et horizontal (bas) dans le cas du milieu hétérogène utilisé dans le chapitre 4 (fig.4.1).

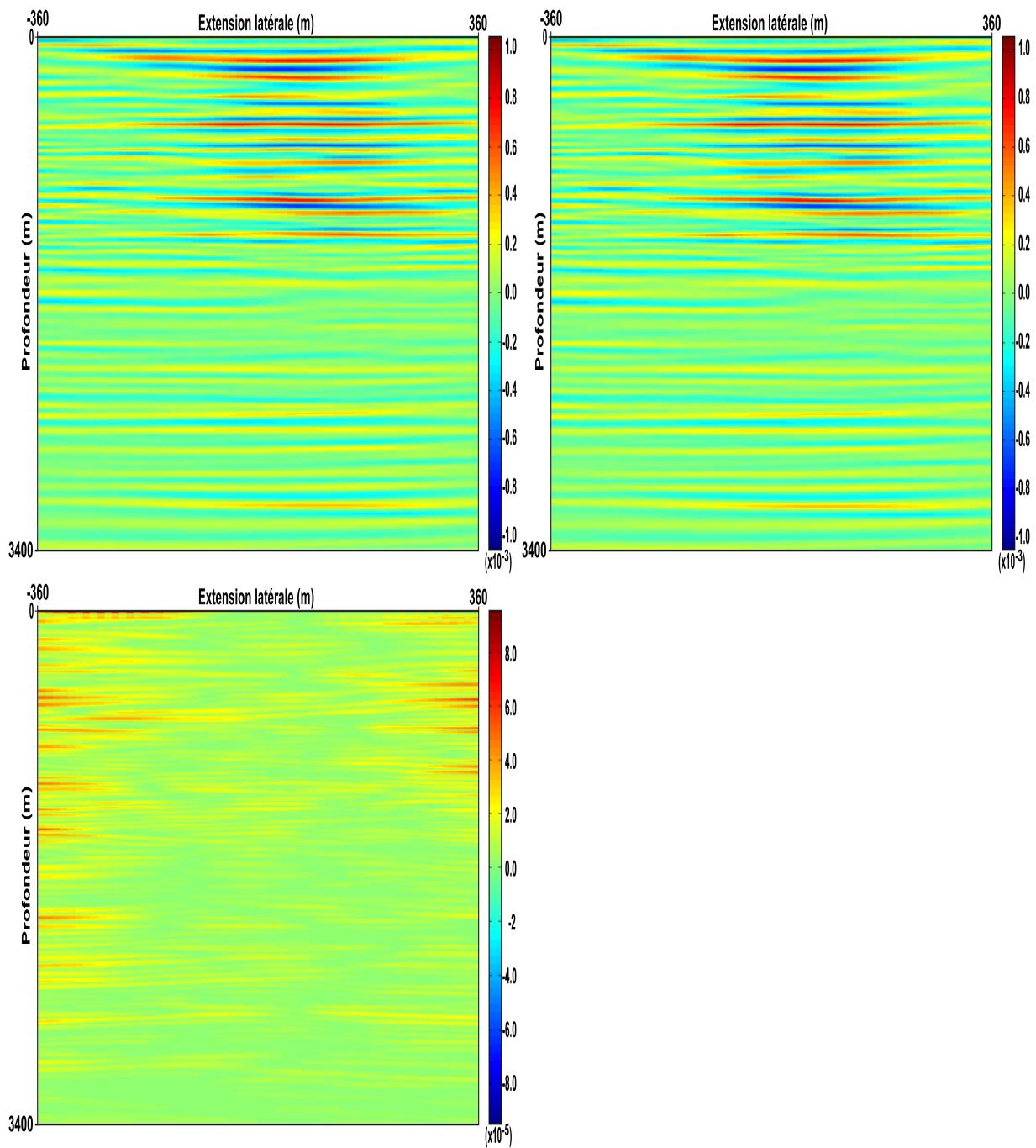


FIG. 5.9 – Champ de vitesse de déplacement vertical pour un modèle hétérogène à $t = 1,9$ s. Milieu infini (haut gauche), milieu PML (haut droite), différence (bas).

On constate dans un premier temps que l'amplitude la plus forte de l'erreur maximale commise n'est pas atteinte au même temps pour la vitesse de déplacement vertical ($t = 1,8$ s) et horizontal ($t = 1,9$ s). Les représentations spatiales de l'erreur sont donc effectuées pour ces temps, respectivement au cas traité. D'autre part, on

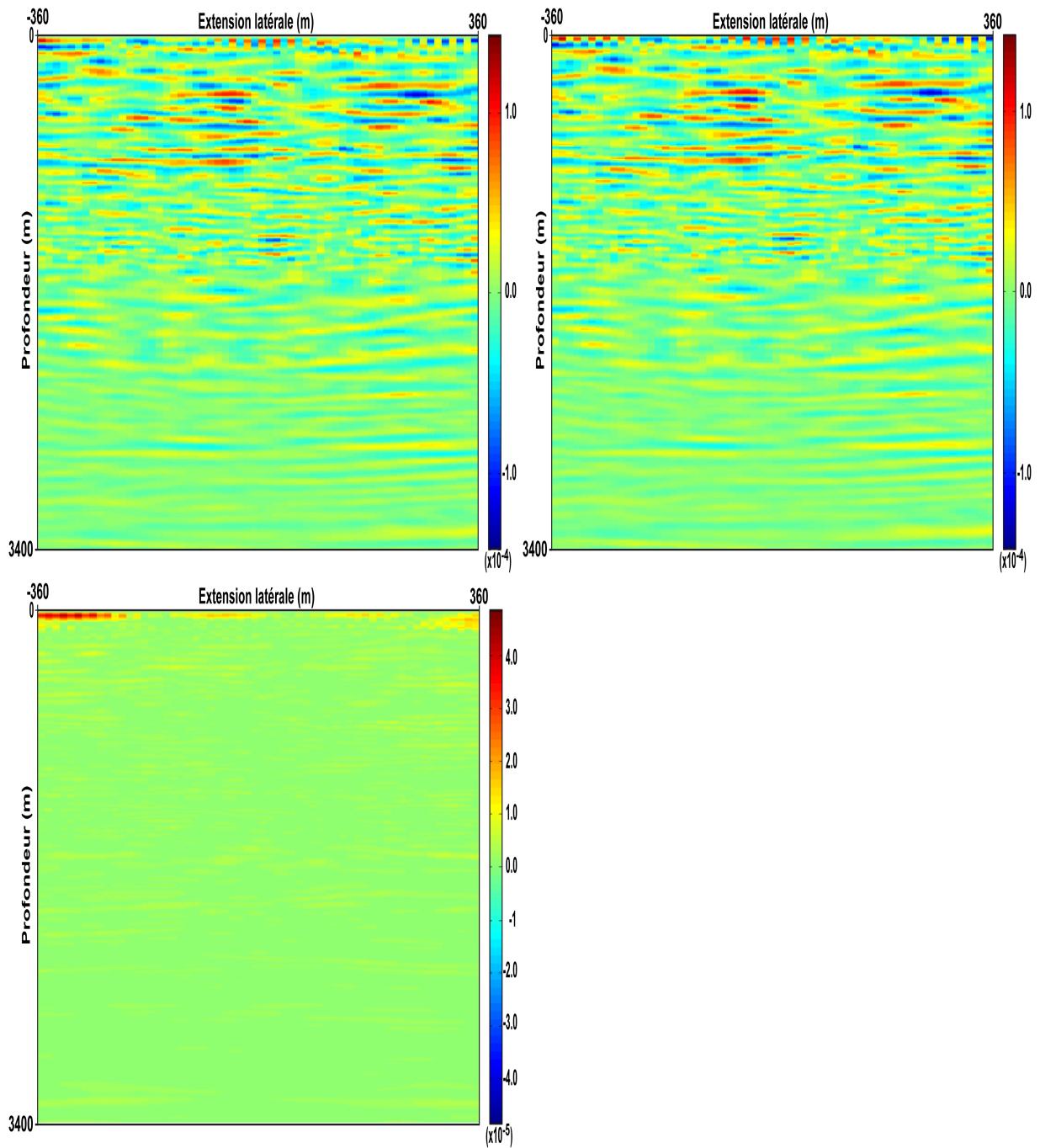


FIG. 5.10 – Champ de vitesse de déplacement horizontal pour un modèle hétérogène à $t = 1,8$ s. Milieu infini (haut gauche), milieu PML (haut droite), différence (bas).

constate une augmentation significative du rapport de l'erreur maximale à l'amplitude maximale du signal : 8% pour la vitesse de déplacement vertical, et près de 40% pour la vitesse de déplacement horizontal. Cependant, comme l'attestent les figures donnant la répartition spatiale de l'erreur, dans ce dernier cas, celle-ci est

extrêmement localisée près du coin gauche. De même pour la vitesse de déplacement vertical, l'erreur reste bien localisée sur les bords du domaine d'intérêt \mathcal{O} . Ainsi, l'utilisation d'un modèle d'impédance et d'un modèle de vitesse très hétérogènes, bien qu'augmentant la complexité du signal, ne rend pas caduque l'utilisation d'une couche PML. Des aberrations localisées, du type de celle enregistrée pour la vitesse de déplacement horizontal, peuvent cependant se produire. Dans le cas présent, cette aberration est peut-être due à la présence d'un coin tel que la partie supérieure n'est pas bornée par la couche PML (condition de Dirichlet homogène qui modélise la transition entre le sous-sol et la surface), tandis que la partie gauche est-elle bornée par la couche PML. L'erreur est "étalée" dans le domaine \mathcal{O} , car les réflexions sur les discontinuités d'impédance génèrent de multiples fronts d'ondes se propageant vers la couche PML. Comme celle-ci n'absorbe pas totalement le signal incident, une petite quantité de chacun de ces fronts est réfléchie vers le domaine d'intérêt, ce qui est à l'origine de cet étalement. Néanmoins, cette erreur reste confinée au bord du domaine d'intérêt.

Il est maintenant nécessaire de s'interroger sur le bruit de modélisation induit par l'utilisation de l'approximation PML au sein de la méthode d'inversion de données sismiques de puits. Pour cela, il convient de calculer la différence entre deux jeux de données obtenus par le dispositif sources/récepteur présenté au chapitre 4 et réutilisé dans cette partie, le premier étant calculé pour un modèle d'extension infini, et le second pour un modèle PML. La différence entre ces deux modèles au sens de la fonctionnelle mesurant le coût sismique permet de mesurer le biais de modélisation induit par l'utilisation de la couche PML dans le cas de l'inversion de données obtenues sur un domaine infini (par exemple des données réelles).

Soient D_{∞}^x les données de vitesse de déplacement horizontal (respectivement D_{∞}^z les données de vitesse de déplacement vertical) obtenues pour un modèle d'extension infini et D_{PML}^x les données de vitesse de déplacement horizontal (respectivement D_{PML}^z les données de vitesse de déplacement vertical) correspondant au modèle PML. L'expression du biais B est donnée par

$$B = \sum_{q=1}^Q \int_0^T (D_{\infty,q}^x(t) - D_{PML,q}^x(t))^2 + (D_{\infty,q}^z(t) - D_{PML,q}^z(t))^2 dt$$

Ce biais mesure donc le coût sismique associé au jeu de données issue de la modélisation PML. Afin de faciliter sa lecture, il est normalisé par la valeur du coût sismique associé à un jeu de données nul, B_0 , tel que

$$B_0 = \sum_{q=1}^Q \int_0^T (D_{\infty,q}^x(t))^2 + (D_{\infty,q}^z(t))^2 dt$$

ce qui correspond à la normalisation qui est opérée préalablement à la procédure

d'inversion. Le tableau suivant récapitule les valeurs du biais normalisé B/B_0 suivant la complexité des modèles utilisés

Modèle	Homogène	2 couches	Hétérogène
B/B_0	5.77×10^{-7}	5.73×10^{-6}	7.02×10^{-5}

FIG. 5.11 – Rapport B/B_0 suivant la complexité du modèle de sous-sol

Le biais de modélisation est donc d'autant plus fort que le modèle de propagation est complexe, ce à quoi il était possible de s'attendre, puisque les hétérogénéités du modèle exercent une sorte de "diffusion" sur l'erreur. Cependant, même dans le cas du modèle hétérogène, la valeur de ce biais est à rapporter à la valeur de la fonctionnelle à l'issue des itérations, voisine de 2×10^{-4} . Dans ce cas précis, le biais de modélisation serait donc négligeable. De plus, ce biais est diminué efficacement par l'utilisation d'une couche PML plus conséquente. Pour une couche PML d'une épaisseur de 30 points de discréttisation, les résultats obtenus sont les suivants

Modèle	Homogène	2 couches	Hétérogène
B/B_0	1.01×10^{-7}	7.63×10^{-7}	1.04×10^{-5}

FIG. 5.12 – Rapport B/B_0 suivant la complexité du modèle de sous-sol

Le biais de modélisation induit par l'utilisation de couches PML est donc faible, et peut être contrôlé par l'épaisseur de la couche PML mise en jeu.

Conclusion

Les expériences numériques menées dans ce chapitre montrent les imperfections de la couche PML (présence d'ondes réfléchies dans le domaine d'intérêt), tout en soulignant le fait que celles-ci sont suffisamment faibles pour être négligées dans le cadre de l'application à l'inversion de données sismiques de puits 2D. En effet, quelle que soit la complexité du milieu de propagation (homogène, deux couches, ou très hétérogène) une différence non nulle est enregistrée près de la frontière du domaine d'intérêt, entre les champs d'onde enregistrés pour un modèle d'extension infini, et le modèle PML. La présence d'hétérogénéités dans le milieu de propagation se traduit en fait par une diffusion dans tout le domaine du champ d'onde acoustique. Le nombre de zones où celui-ci pénètre dans la couche PML est donc considérablement augmenté, et à la diffusion du champ d'onde correspond la diffusion de l'erreur. L'amplitude maximale de celle-ci augmente, mais reste localisée près de la frontière. De plus, le calcul du biais de modélisation indique une différence, au sens de la fonction coût, négligeable devant la précision qui peut être atteinte pour la minimisation du coût sismique. La méthode PML qui a été définie est donc suffisamment précise pour être utilisée à profit dans le cadre de l'inversion de données sismiques de puits.

Les difficultés engendrées par l'utilisation de la couche PML résident plus dans l'aspect calcul haute performance du problème. En effet, d'une part, les équations PML imposent une séparation des inconnues qui mène de 3 à 4 équations en 2D, et de 6 à 9 en 3D. Ceci représente autant de données à stocker et de calculs à effectuer. D'autre part, comme cela est expliqué en Annexe 3.4, l'utilisation de la couche PML peut conduire, si les moyens de calcul disponibles ne sont pas suffisants, à l'utilisation d'algorithmes de rétropropagation spéciaux, qui réalisent un compromis entre le stockage et le coût de calcul, mais qui représentent cependant une charge supplémentaire non négligeable par rapport au contexte où toutes les informations peuvent être stockées. La taille de la couche PML doit ainsi être définie avec précaution, pour assurer un bias de modélisation faible sans toutefois surcharger les calculateurs.

Conclusion et perspectives

Le travail réalisé au cours de cette thèse a été mené dans l'objectif de définir une méthode d'inversion de données sismiques adaptée à l'imagerie de milieux très hétérogènes. Ces milieux, par leurs hétérogénéités, produisent des données sismiques comportant une part très importante de réflexions multiples. Celles-ci ne sont pas prises en compte par les méthodes d'inversion linéaires classiques, ce qui conduit à l'impossibilité, pour de telles méthodes, de fournir une image satisfaisante de ces milieux [Delprat-Jannaud, Lailly, 2004] [Delprat-Jannaud, Lailly, 2008]. L'imagerie de tels milieux nécessite donc la prise en compte de la totalité des données sismiques, ce qui revient à la résolution du problème d'inversion non linéaire. De plus, suivant une remarque de [Delprat-Jannaud, Lailly, 2008], inspirée d'une observation de [Fink, 1997], la résolution du problème inverse non linéaire peut mener à une nette amélioration de la précision du résultat. La résolution du problème inverse non linéaire est possible dans le cadre de l'inversion de données sismiques de puits, sans avoir à réaliser une estimation *a priori* des sources générant les données. Cette estimation est en effet un handicap pour la mise en place d'une méthode d'inversion de données sismiques non linéaire, car elle requiert une extrême précision. L'inversion de données sismiques de puits, étudiée dans les années 80 [Macé, Lailly, 1986], a donné naissance à une méthode d'inversion non linéaire utilisable uniquement dans le cadre d'hypothèses très restrictives de propagation 1D. L'enjeu du travail réalisé ici a donc été de proposer une nouvelle méthode d'inversion non linéaire de données sismiques de puits, multi-dimensionnelle, s'inspirant du problème 1D. Il s'attache au traitement du cas particulier bidimensionnel, pour une modélisation acoustique, dans l'optique d'une extension tridimensionnelle.

Le nouveau problème d'inversion consiste donc à déterminer une estimation du modèle d'impédance acoustique dans une zone définie autour d'un puits vertical dans lequel sont placés des récepteurs, enregistrant les vitesses de déplacement du milieu. Dans un premier temps, une méthode d'acquisition de données permettant d'éclairer les variations latérales du milieu est définie. Celle-ci consiste à utiliser plusieurs jeux de données, acquis pour différents départs de la source par rapport à la tête du puits. De plus, le contexte de modélisation acoustique nécessite d'utiliser une estimation du modèle de vitesse de propagation, ce qui conduit à s'intéresser à l'imagerie de milieux présentant de faibles variations latérales. Cette hypothèse est utilisée pour réduire l'indétermination du problème résultant de la forte augmenta-

tion du nombre d'inconnues, et de leur couplage dans la procédure d'inversion. Une recombinaison des données ("slant-stacking") est notamment utilisée pour se rapprocher d'un éclairage du milieu par des ondes planes se propageant avec différents angles d'incidence. Le problème d'inversion posé est donc mis sous la forme d'un problème d'optimisation d'une fonctionnelle régularisée.

La résolution numérique de ce problème requiert des efforts particuliers. Tout d'abord, la modélisation de la propagation des ondes acoustiques dans le sous-sol nécessite l'utilisation de conditions absorbantes, afin d'éviter la propagation de réflexions parasites. Une formulation utilisant des couches absorbantes de type PML est utilisée à cette fin. Une méthode d'optimisation adaptée au problème d'indétermination de couplage des inconnues est ensuite mise en place, baptisée "optimisation emboîtée". Cet algorithme nécessite le calcul du gradient de la fonctionnelle minimisée. Celui-ci est déterminé par la méthode de l'état adjoint, adaptée au cadre discret introduit pour résoudre le problème de modélisation de la propagation des ondes acoustiques en milieu PML. Des algorithmes rapides permettent un calcul efficace de ce gradient, basé sur les équations de l'état adjoint. De plus, pour optimiser les temps de calcul et bénéficier de la puissance de calcul des architectures distribuées, une parallélisation du calcul du gradient est effectuée à l'aide d'une méthode de décomposition de domaines.

Ceci permet de disposer d'une méthode de résolution numérique efficace, permettant de traiter des cas de taille conséquente (plusieurs centaines de milliers d'inconnues pour plusieurs millions de données discrètes) en des temps ne dépassant pas plusieurs dizaines d'heures de calcul pour l'utilisation de 128 processeurs. Des résultats sont obtenus sur un modèle de sous-sol hétérogène, pour lequel les réflexions multiples ne peuvent être négligées, éclairé par 30 sources disposées tous les 24 m jusqu'à 360 m du puits. Ces sources émettent un signal de Ricker centré sur 25 Hz, ce qui mène à un éclairage du milieu d'une longueur d'onde moyenne de 140 m. Les résultats numériques montrent la possibilité d'obtenir une image très précise de l'impédance dans la zone où se situent les récepteurs, jusqu'à 300 m de part et d'autre du puits, avec une précision verticale inférieure à 10 m. L'estimation du modèle d'impédance acquise dans la zone située sous les récepteurs bénéficie d'une résolution verticale du même ordre (détails inférieurs à 10 m), mais la résolution latérale est moins bonne (de l'ordre de la cinquantaine de mètres, soit la résolution d'une méthode d'inversion linéaire classique dans ce contexte), et le résultat souffre d'indétermination. Celle-ci peut-être cependant maîtrisée par l'information *a priori* sur la régularité du milieu. Les premiers tests de sensibilité à des erreurs sur l'estimation du modèle de vitesse montrent l'importance de la précision de cette estimation au niveau des récepteurs, quant à ceux effectués sur du bruit introduit sur les données, ils soulignent l'indétermination latérale de l'estimation du modèle d'impédance dans la zone située sous les récepteurs.

Ces premiers résultats sont encourageants. Les perspectives de travail à court terme consistent à poursuivre les tests liés à la résolution et à la sensibilité de la méthode, notamment l'évaluation de l'influence de l'estimation du modèle de vitesse sur la résolution latérale, couplée à la prise en compte de données couvrant une gamme d'angles d'éclairage plus grande. Des tests sur des données plus fortement bruitées sont également nécessaires, qui montreront peut-être la nécessité d'introduire de nouveaux termes de régularisation dans la fonctionnelle. Par exemple, il sera peut-être utile, en cas d'instabilité, d'utiliser une régularisation par la norme L^1 des dérivées verticales du modèle d'impédance, comme cela a été fait dans le cas 1D [Macé, Lailly, 1986]. D'autres tests peuvent aussi être effectués, concernant par exemple l'influence du nombre de récepteurs et de leur répartition sur la résolution de l'estimation du modèle d'impédance, ainsi que celle du nombre de sources utilisées, et de leur disposition en surface (distance entre les sources, et distance maximale par rapport à la tête de puits). Il faudrait également tester l'influence de la discrétisation sur la résolution du résultat, en procédant à des raffinements successifs, pour obtenir une évaluation de la résolution de la méthode qui soit détachée de toute limitation numérique liée au choix initial des pas de discrétisation. Pour finir, il est également envisageable de tester la méthode sur un domaine comportant plusieurs puits, afin de couvrir une zone plus large, et d'améliorer la résolution latérale de la zone située sous les récepteurs. D'un point de vue numérique, l'introduction d'un niveau de parallélisme supplémentaire, par rapport aux jeux de données, permettra de réduire encore les temps de calcul, et rendra celui-ci indépendant du nombre de jeux de données considérés.

Parallèlement à ces travaux, une étude mathématique plus spécifique a été menée, ayant trait à l'utilisation d'une couche PML pour la modélisation de la propagation d'ondes acoustiques en milieu 3D hétérogène borné, présentant des discontinuités d'impédance. Son ambition est de proposer une extension des travaux portant sur le caractère faiblement bien posé des équations PML, qui ont été réalisés dans le cadre de milieux homogènes non bornés. Des estimations d'énergie ont été établies, qui montrent que les équations PML pour la propagation d'ondes acoustiques en milieu hétérogène borné restent faiblement bien posées de défaut 1. Ce résultat est également prouvé pour un milieu hétérogène borné dans lequel l'impédance présente des discontinuités, mais ne dépend pas des directions tangentes aux surfaces de discontinuité. De plus, des expériences numériques ont été effectuées afin de montrer que cette modélisation peut être employée dans le cadre spécifique de l'application à l'inversion de données sismiques de puits. Le biais de modélisation introduit est suffisamment faible pour ne pas être pris en compte par la méthode d'inversion, et ce résultat est garanti pour une couche PML d'épaisseur réduite, n'induisant pas de surcoûts de calcul excessifs.

A plus long terme, il convient de procéder à l'extension des résultats acquis pour l'inversion de données sismiques de puits bi-dimensionnelle, à une modélisation

acoustique tri-dimensionnelle, ce qui fournira un outil d'imagerie sismique haute résolution, utilisable tant pour la caractérisation des réservoirs que pour le monitoring des sites de stockage de CO₂. Cette extension est immédiate du point de vue de la méthodologie, celle-ci ayant été développée en vue d'une application tri-dimensionnelle. Une adaptation de l'implémentation numérique est cependant nécessaire. La taille du problème d'inversion 3D est en effet considérable. Du fait du volume des données, il convient d'utiliser les algorithmes de calcul du gradient ne procédant pas à leurs stockage massif, qui sont proposés dans l'annexe 3 du chapitre 3. La répartition des charges entre processeurs doit également être étudiée avec soin. D'autre part, afin d'améliorer les performances numériques, il est possible d'envisager l'introduction d'une parallélisation en temps pour le calcul du gradient. L'utilisation d'une architecture de calcul sur carte graphique (GPU) peut également fournir la possibilité de mieux contrôler le stockage des données en mémoire et ainsi d'en optimiser les temps accès. Les évolutions de la méthode concernent ensuite l'intégration d'une modélisation plus réaliste de la propagation des ondes, comme par exemple le passage à une modélisation élastique, ou la prise en compte du phénomène d'atténuation. Ces évolutions nécessitent cependant une adaptation à la fois numérique et méthodologique. En effet, pour un modèle élastique, le nombre de paramètres augmente, et par là, l'indétermination du problème, ce qui implique d'adapter d'abord la définition du problème inverse. A cette fin, il convient de déterminer quels paramètres doivent être estimés dans l'inversion. De plus, les calculs à mener sont complexifiés par la définition d'un plus grand nombre d'équations, et la quantité croissante d'inconnues, ce qui demande une adaptation attentive de la méthode parallèle de résolution numérique.

Table des figures

1.1	Principe d'une expérience de sismique réflexion.	14
1.2	Modèle d'impédance exact.	27
1.3	Signal de Ricker.	28
1.4	Données PSV synthétiques.	29
1.5	Enregistrement sur le premier récepteur.	29
1.6	Enregistrement sur le 50 ^e récepteur.	30
1.7	Décroissance de la fonction coût (cas 1D).	31
1.8	Résidus obtenus sur l'ensemble des récepteurs.	32
1.9	Résidus obtenus pour le premier récepteur. En noir, données exactes et données synthétiques, en gris, différence entre ces deux signaux. . .	32
1.10	Résidus obtenus pour le 50 ^e récepteur (bas). En noir, données exactes et données synthétiques, en gris, différence entre ces deux signaux. . .	33
1.11	Résultat d'inversion 1D par rapport à la condition de pression au premier récepteur. En noir, condition exacte et condition calculée, en gris, différence.	33
1.12	Résultat d'inversion 1D par rapport à l'impédance.	34
1.13	Résultat d'inversion 1D par rapport à l'impédance. Zoom sur la zone des récepteurs (haut), zoom sur la partie profonde (bas).	35
2.1	Zone d'éclairage suivant l'approximation géométrique.	44
2.2	Prise en compte des multiples pour la zone d'éclairage.	45
2.3	Construction d'une onde plane à incidence oblique	54
2.4	CP exactes incidence normale (gauche), incidence oblique 6° (droite). . .	55
3.1	Découpage en 9 sous-domaines.	111
3.2	Schéma d'interpolation à quatre points.	122
4.1	Modèles exacts d'impédance (gauche), et de vitesse de propagation (droite).	137
4.2	Signal de Ricker.	139
4.3	Données recombinées par angle. Vitesse verticale 0°(haut gauche), vitesse verticale +6°(haut droite), vitesse horizontale 0°(bas gauche), vitesse horizontale +6°(bas droite).	140
4.4	Temps de calcul suivant le nombre de processeurs pour l'algorithme 4. 144	
4.5	Performance suivant le nombre de coeurs utilisés pour l'algorithme 4. 145	

4.6	Temps de calcul suivant le nombre de processeurs pour l'algorithme 3 limité à 10 itérations.	146
4.7	Performance suivant le nombre de processeurs pour l'algorithme 3 limité à 10 itérations.	146
4.8	Décroissance de la fonction coût en fonction du temps (cas 2D, CP exactes).	150
4.9	Résidus vitesse verticale 0 °(haut gauche), vitesse verticale 6 °(haut droite), vitesse horizontale 0 °(bas gauche), vitesse horizontale 6 °(bas droite).	151
4.10	Modèle d'impédance exact (haut gauche), calculé (haut droite), différence (bas).	152
4.11	Profils sismiques verticaux exacts (noir), et calculés (gris). De gauche à droite, et de bas en haut : traces au puits ($x = 0$ m), $x = 40$ m, $x = 100$ m, $x = 200$ m, (e) $x = 300$ m.	154
4.12	Spectres de puissance des profils sismiques verticaux. La première ligne correspond à l'offset 0 m, la seconde +40 m, la troisième +100 m, la quatrième +200 m, la cinquième +300 m. La première colonne correspond au modèle exact, la seconde au modèle calculé, la troisième au résultat d'inversion linéarisée.	157
4.13	Profils sismiques latéraux exacts (noir) et calculés (gris). De gauche à droite, et de haut en bas, $z = 1000$ m, $z = 1400$ m, $z = 1800$ m, $z = 2000$ m, $z = 2100$ m, $z = 2200$ m.	159
4.14	Profils sismiques latéraux exacts (noir) et calculés (gris). De gauche à droite, et de haut en bas, $z = 2300$ m, $z = 2400$ m, $z = 2500$ m, $z = 3000$ m.	160
4.15	Modèle d'impédance après perturbation située en $x = 144$ m et $z = 2200$ m (haut gauche), calculé à partir des données perturbées (haut droite), différence entre les modèles d'impédance calculés avec et sans perturbation.	163
4.16	Modèles d'impédance, après perturbation située en $x = 144$ m et $z = 2600$ m (haut gauche), calculé à partir des données perturbées (haut droite), différence entre les modèles d'impédance calculés avec et sans perturbation.	165
4.17	Trame utilisée pour l'interpolation des CP.	170
4.18	Comparaison entre les CP exactes (gauche) et les CP interpolées à partir des traces exactes (droite). Première ligne : données recombinées par angle d'éclairage 0°, deuxième ligne -6°.	171
4.19	Comparaison entre les CP exactes (gauche) et les CP interpolées à partir des traces exactes (droite), pour les données recombinées par angle d'éclairage 6°.	172
4.20	Définitions entre les CP exactes et les CP interpolées, correspondant aux données recombinées par angle d'éclairage 0°(haut gauche), -6°(haut droite), 6°(bas).	173

4.21	Comparaison entre le modèle exact (noir) et le modèle basse fréquence utilisé comme modèle <i>a priori</i> (gris) localement au puits ($x = 0$ m et 1000 m $< z < 1800$ m).	174
4.22	Courbe de décroissance de la fonction coût en fonction du temps pour l'inversion 2D à CP inconnues.	175
4.23	Résidus. Vitesse verticale 0° (haut gauche), vitesse verticale 6° (haut droite) , vitesse horizontale 0° (bas gauche), vitesse horizontale 6° (bas droite).	177
4.24	CP calculées, 0° (haut gauche), -6° (haut droite), 6° (bas).	179
4.25	Differences entre CP calculées et CP exactes, 0° (haut gauche), -6° (haut droite), 6° (bas).	180
4.26	Modèle d'impédance exact (haut gauche), modèle d'impédance calculé par inversion avec CP inconnues (haut droite), différence (bas).	182
4.27	Convergence comparée de la méthode d'optimisation globale et la méthode d'optimisation emboîtée.	183
4.28	CP calculées par optimisation globale, 0° (haut gauche), -6° (haut droite), 6° (bas).	184
4.29	Modèle d'impédance calculée par optimisation globale.	185
4.30	Profils sismiques verticaux exact(noir), et calculés (gris). De gauche à droite, et de haut en bas, traces au puits ($x = 0$ m), $x = 40$ m, $x = 100$ m, $x = 200$ m, $x = 300$ m.	186
4.31	Spectres de puissance. Par lignes : $x = 0$ m, $x = 40$ m, $x = 100$ m, $x = 200$ m, $x = 300$ m. Par colonnes : modèle exact, modèle calculé avec CP inconnues, modèle calculé par migration.	188
4.32	Profils sismiques latéraux exact (noir) et calculé (gris). De gauche à droite et de haut en bas, $z = 1000$ m, $z = 1400$ m, $z = 1800$ m, $z = 2000$ m, $z = 2100$ m, $z = 2200$	190
4.33	Profils sismiques latéraux exact (noir) et calculé (gris). De gauche à droite et de haut en bas, $z = 2300$ m, $z = 2400$ m, $z = 2500$ m, $z = 3000$ m.	191
4.34	Modèle d'impédance après perturbation en $x = 144$ m et $z = 2200$ m (haut gauche), modèle calculé à partir des données perturbées correspondantes (haut droite). Modèle d'impédance après perturbation en $x = 144$ m et $z = 2600$ m (bas gauche), modèle calculé à partir des données perturbées correspondantes (bas droite)	194
4.35	Differences entre modèles issus de données perturbées (fig.4.34 et modèle issu de données non perturbées (fig.4.26), perturbation en $x = 144$ m et $z = 2200$ m (gauche), perturbation en $x = 144$ m et $z = 2600$ m (droite).	195
4.36	Modèle d'impédance après perturbation d'amplitude 30% en $x = 144$ m et $z = 2200$ m (gauche), modèle calculé à partir des données perturbées correspondantes (droite), différences avec modèle issu de données non perturbées fig.4.26 (bas).	197

4.37	Impédance exacte (haut gauche), impédance calculée par inversion en utilisant un modèle de vitesse 1D extrapolé à partir du modèle de vitesse en $x = 0$ (haut droite), différence (bas).	202
4.38	Profils sismiques verticaux du modèle d'impédance calculé en utilisant un modèle de vitesse 1D extrapolé (gris), et du modèle exact (noir). De droite à gauche, et de haut en bas, $x = 0$ m, $x = 40$ m, $x = 200$ m.	203
4.39	Modèle de vitesse exact (gauche) et modèle de vitesse 1D en zone supérieure et homogène en zone inférieure (droite).	204
4.40	Impédance exacte (haut gauche), impédance calculée par inversion en utilisant un modèle de vitesse 1D extrapolé à partir des valeurs de vitesse au niveau des récepteurs dans la zone supérieure et homogène dans la zone inférieure (haut droite), différence (bas).	205
4.41	Profils sismiques verticaux du modèle d'impédance calculé en utilisant un modèle de vitesse 1D extrapolé dans la zone supérieure et homogène dans la zone inférieure (gris), et du modèle exact (noir). De gauche à droite et de haut en bas, $x = 0$ m, $x = 40$ m, $x = 200$ m.	207
4.42	Modèle de vitesse exact (gauche) et modèle de vitesse 1D moyen (droite).	208
4.43	Profils du modèle de vitesse exact (noire) et du modèle de vitesse 1D moyen (gris) au niveau du puits (gauche), zoom sur la zone récepteur (droite).	209
4.44	Impédance exacte (haut gauche), impédance calculée par inversion en utilisant un modèle de vitesse 1D moyen (haut droite), différence (bas).	210
4.45	Données de vitesse verticale bruitées, recombinées par angle déclairage. A gauche, 0° , à droite 6°	212
4.46	Résidus obtenues à l'issue de 1000 itérations sur les données bruitée données correspondant à l'onde plane 0° (gauche), et 6° (droite). . . .	213
4.47	Impédance exacte (haut gauche), impédance calculée par inversion en utilisant des données bruitées (haut droite), différence (bas).	214
4.48	Impédance exacte (haut gauche), impédance calculée par inversion en utilisant des données bruitées et une régularisation plus forte (haut droite), différence (bas).	216
5.1	Erreurs en norme L^∞ pour les vitesses de déplacement vertical (haut) et horizontal (bas) dans le cas d'un milieu homogène	244
5.2	Champ de vitesse de déplacement vertical pour un modèle homogène à $t=1.2$ s. Milieu infini (haut gauche), milieu PML (haut droite), différence (bas).	245
5.3	Champ de vitesse de déplacement horizontal pour un modèle homogène à $t=1.2$ s. Milieu infini (haut gauche), milieu PML (haut droite), différence (bas).	246
5.4	Erreurs en norme L^∞ pour la vitesse de déplacement vertical dans le cas d'un milieu à deux couches.	247
5.5	Erreurs en norme L^∞ pour la vitesse de déplacement horizontal dans le cas d'un milieu à deux couches.	248

5.6	Champ de vitesse de déplacement vertical pour un modèle à deux couches à $t = 1, 2$ s. Milieu infini (haut gauche), milieu PML (haut droite), différence (bas)	249
5.7	Champ de vitesse de déplacement horizontal pour un modèle à deux couches à $t = 1, 2$ s. Milieu infini (haut gauche), milieu PML (haut droite), différence (bas)	250
5.8	Erreur en norme L^∞ pour les vitesses de déplacement vertical (haut) et horizontal (bas) dans le cas du milieu hétérogène utilisé dans le chapitre 4 (fig.4.1)	251
5.9	Champ de vitesse de déplacement vertical pour un modèle hétérogène à $t = 1, 9$ s. Milieu infini (haut gauche), milieu PML (haut droite), différence (bas)	252
5.10	Champ de vitesse de déplacement horizontal pour un modèle hétérogène à $t = 1, 8$ s. Milieu infini (haut gauche), milieu PML (haut droite), différence (bas)	253
5.11	Rapport B/B_0 suivant la complexité du modèle de sous-sol	255
5.12	Rapport B/B_0 suivant la complexité du modèle de sous-sol	255

Références

- [Abarbanel, Gottlieb, 1997] S. Abarbanel, D.Gottlieb, 1997, A mathematical analysis of the PML method, *J.Comput. Phys.*, **134**, 357-363.
- [Abarbanel et al., 2002] S. Abarbanel, D.Gottlieb, J.S.Hesthaven, 2002, Long time behavior of the perfectly matched layer equations in computational electromagnetics, *J.Sci.Comput.*, **17**, 405-422.
- [Bamberger et al., 1977] A.Bamberger, G.Chavent, P.Lailly, 1977, Etude mathématique et numérique d'un problème inverse pour l'équation des ondes à une dimension, rapport interne 14, école Polytechnique, CMA.
- [Bamberger, 1978] A.Bamberger, 1978, Analyse, contrôle et identification de certains systèmes, Thèse de doctorat d'état es-sciences mathématiques soutenue à l'université Pierre et Marie-Curie Paris VI.
- [Benamou, 1994] J.D.Benamou, 1994, Domain decomposition methods with coupled transmission conditions for the control of systems governed by elliptic partial differential equations, *SIAM Journal of Numerical Analysis*, **33**, 2401-2416.
- [Benamou, Després, 1997] J.D.Benamou, B.Després, 1997, A domain decomposition methods for the Helmholtz equation and related optimal control problems, *J. Comput. Phys.*, **136**, 68-82.
- [Bérenger, 1994] J.P.Bérenger, 1994, A perfectly matched layer for the absorption of electromagnetic waves, *J. Comput. Phys.*, **114**, 185-200.
- [Bérenger, 1996] J.P.Bérenger, 1996, Three-dimensional perfectly matched layer for the absorption of electromagnetic waves, *J. Comput. Phys.*, **127**, 363-379.
- [Bleistein et al., 2000] N. Bleistein, J.K. Cohen, J.W.Stockwell, Jr., 2000, Mathematics of Multidimensional Seismic Imaging, Migration, and Inversion, Springer Verlag, New York, ISBN 0-387-95061-3.
- [Bensoussan et al., 1973] A.Bensoussan, R.Glowinski, J.L.Lions, 1973, Méthode de décomposition appliquée au contrôle optimal de systèmes distribués, Proceedings of the Fifth IFIP Conference on Optimization Techniques, Part 2, Lecture Notes in Computer Sciences, **3**, Springer Berlin, 141-153.
- [Bounaïm, 1999] A.Bounaïm, 1999, Méthodes de décomposition de domaine : application à la résolution de problèmes de contrôle optimal, thèse soutenue à l'université Joseph Fourier - Grenoble I, spécialité mathématiques appliquées.

- [Claerbout, 1976] J.F. Claerbout, 1976, Fundamental of geophysical data processing, Mc Graw-Hill Book Co.
- [Claerbout, Schultz, 1978] Philip S. Schultz and Jon F. Claerbout, 1978, Velocity estimation and downward continuation by wavefront synthesis, *Geophysics*, **43**, n°4, 691-714.
- [Collino, 1992] F.Collino, 1992, Conditions Absorbantes d'ordre élevé pour des modèles de propagation d'onde dans des domaines rectangulaires, Rapport de Recherche INRIA n°1790.
- [Colonius et al., 1993] T.Colonius, S.K.Lele, P.Moin, 1993, Boundary Conditions for Direct Computation of Aerodynamic Sound Generation, *AIAA Journal*, **31**, 1574-1582.
- [Delprat-Jannaud, Lailly, 2004] F.Delprat-Jannaud, P.Lailly, 2004, The insidious effects of fine-scale heterogeneity in reflection seismology, *Journal of Seismic Exploration*, **13**, 39-84.
- [Delprat-Jannaud, Lailly, 2005] F.Delprat-Jannaud, P.Lailly, 2005, A fundamental limitation for the reconstruction of impedance profiles from seismic data, *Geophysics*, **70**, n°1, R1-R14.
- [Delprat-Jannaud, Lailly, 2008] F.Delprat-Jannaud, P.Lailly, 2008, Waves propagation in heterogeneous media, *Geophysics*, **73**, n°3, T37-T49.
- [Duquet, 1996] B.Duquet, 1996, Amélioration de l'imagerie sismique de structures géologiques complexes, thèse soutenue à l'Institut Français du pétrole, Ecole doctorale de l'université Paris XIII en mathématiques appliquées.
- [Engquist, Majda, 1977] B.Engquist, A.Majda, 1977, Absorbing boundary conditions for the numerical simulation of waves, *Math. Comp.*, **31** 629-651.
- [Fink, 1997] M. Fink, 1997, Time-reversed acoustics, *Physics Today*, **50**, 34-40.
- [Freund, 1997] J.B.Freund, 1997, Proposed Inflow/Outflow Boundary Conditions for Direct Computation of Aerodynamic Sound, *AUAA Journal* **35**, 740-742.
- [Garnier, Papanicolaou, 2009] J.Garnier, G.Papanicolaou, Passive Sensor Imaging Using Cross Correlations of Noisy Signals in a Scattering Medium, *SIAM J. Imaging Science*, Volume 2, Issue 2, 386-437.
- [Gilbert and al., 2006] J.F.Bonnans, J.C.Gilbert, C.Lemaréchal, Claudia A.Safastizábal, 2006, Numerical Optimization, Theoretical and Practical Aspects, Second Edition, Springer, series Universitext.
- [Gilbert, 1999] J.Ch. Gilbert, 1999, Module N1CG1,
<http://www-rocq.inria.fr/estime/modulopt/modulopt.html>
- [Hadamard, 1902] J.Hadamard, 1902, Sur les problèmes aux dérivées partielles et leur signification physique, *Princeton University Bulletin*, 49-52.
- [Hesthaven, 1998] J.S.Hesthaven, 1998, On the Analysis and Construction of Perfectly Matched Layers for Linearized Euler Equations, *J.Comput. Phys.*, **142**, 129-147.

- [Hu, 1996] F.Q.Hu, 1996, On absorbing boundary conditions for linearized Euler equations by a perfectly matched layer, *J.Comput.Phys.* **129**(1996).
- [Kreiss, Lorenz, 2004] H.O.Kreiss, J.Lorenz, 2004, Initial-Boundary Value Problems and the Navier-Stokes Equations, SIAM, initial edition published by Academic Press, 1989, ISBN 0-89871-565-2.
- [Lions, Magenes, 1968] J.L. Lions, E.Magenes, Problèmes aux limites non homogènes, volume 2, Dunod, Paris, 1968.
- [Macé, Lailly, 1986] D.Macé, P.Lailly, 1986, Solution of the VSP one-dimensional inverse problem, *Geophysical Prospecting*, **34**, 1002-1021.
- [Mari et al., 2004] J.L.Mari, J.L.Lacoume, J.Mars, F.Glangeaud, 2004, Traitement du signal pour géologues et géophysiciens, Publications de l'Institut Français du Pétrole, Editions Technip.
- [Mazaheri, Roe, 1997] K.Mazaheri, P.L.Roe, 1997, Numerical Wave Propagation and Steady-State Solutions : Soft Wall and Outer Boundary Conditions, *AIAA Journal* **35**, 965-975.
- [Métivier, 2009] L.Métivier, 2009, Utilisation des équations Euler-PML en milieu hétérogène borné pour la résolution d'un problème inverse en géophysique, *ESAIM : Proceedings*, **27**, 156-170.
- [Métral, Vacus, 1999] J.Métral, O.Vacus, 1999, Caractère bien posé du problème de Cauchy pour le système de Bérenger, *C.R. Acad. Sci. Paris Sér.I Math.*, **10**, 847-852.
- [Mintrop, 1931] On the history of the seismic method for the investigation of underground formations and mineral deposits : Seismos, Hanovre.
- [Nataf, 2006] F.Nataf, 2006, A new construction of perfectly matched layers for the linearized Euler equations, *J. Comput. Phys.*, **214** n°2, 757-772.
- [Nocedal and al., 1995] R.H.Byrd, P.Lu, J.Nocedal, 1995, A Limited Memory Algorithm for Bound Constrained Optimization, *SIAM Journal on Scientific and Statistical Computing*, 16,5,1190-1208.
- [Nocedal and al., 1997] C.Zhu, R.H.Byrd, J.Nocedal, 1997, L-BFGS-B : Algorithm 778 : L-BFGS-B, FORTRAN routines for large scale bound constrained optimization, *ACM Transactions on Mathematical Software*, **23**, Num.4, 550-560.
- [Petit-Bergez, 2006] S.Petit-Bergez, 2006, Problèmes faiblement bien posés : discréétisation et applications, thèse soutenue à l'université de Paris XIII, discipline Mathématiques appliquées.
- [Pion, 2007] Y. Pions, 2007, Problème inverse de la sismique réflexion : analyse de la Hessienne du problème linéarisé et applications à l'imagerie 3D, thèse soutenue à l'Institut Français du pétrole, Ecole doctorale en Mathématique et Informatique de Marseille.
- [Radon, 1917] J.Radon, 1917, Ueber die Bestimmung von Funktionen durch ihre Integralwerte längs gewisser Mannigfaltigkeiten, *Ber. Verh. Sächs. Akad.*, **69**, 262-277.

- [Rahmouni, 2004] A.Rahmouni, 2000, Un modèle PML bien posé pour les équation d'Euler linéarisées, C.R. Acad. Sci. Paris Sér. I Math., **331**, 159-164.
- [Requena et al., 2004] J.Chergui, I.Dupays, D.Girou, N.Grima, S.Requena, 2004, Message Passing Interface (MPI-1), Version 2.12, Idris,
http://www.idris.fr/data/cours/parallel/mpi/choix_doc.html
- [Shanno, 1970] D.F.Shanno, 1970, Conditionning of Quasi-Newton Methods for Function Minimization, Mathematics of Computation, **24**, 647-656.
- [Sondhi, Gopinath, 1971] M.M.Sondhi, B.Gopinath, 1971, Determination of vocal-tract shape from impulse response at the lips, J.Acoust.Amer.,**49**, 6, 1867-1873.
- [Tarantola, 2005] A.Tarantola, 2005, Inverse Problem Theory and Model Parameter Estimation, SIAM book, ISBN 0-89871-572-5.
- [Ta'asan, Nark, 1995] S.Taásan, D.M.Nark, 1995, An absorbing Buffer Zone Technique for Acoustic Wave Propagation, AIAA Paper 95-0146.
- [Tikhonov, 1943] A.N. Tikhonov, 1943, On the stability of inverse problems, Dokl. Akad. Nauk SSSR, **39**, No. 5,195-198.
- [Yee, 1966] K.S.Yee, 1966, Numerical solution of initial boundary value problems involving Maxwell's equations in isotropic media, IEEE Trans. Antennas and Propagation, 302-307.
- [Yilmaz, 1987] O.Yilmaz, 1987, Seismic data processing, Society of Exploration Geophysicists.