HAL
open science

Quelle confiance accordée au modèle solution de la tomographie de réflexion 3D ?

Carole Duffet

To cite this version:

Carole Duffet. Quelle confiance accordée au modèle solution de la tomographie de réflexion 3D ?. Optimisation et contrôle [math.OC]. Université de Montpellier 2, 2004. Français. NNT : . tel-02494284

HAL Id: tel-02494284

https://hal.science/tel-02494284v1

Submitted on 28 Feb 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

UNIVERSITÉ MONTPELLIER II
SCIENCES ET TECHNIQUES DU LANGUEDOC

THÈSE
pour obtenir le grade de

DOCTEUR DE L'UNIVERSITÉ MONTPELLIER II

Discipline : *Mathématiques Appliquées*
Formation doctorale : *Mathématiques*
École doctorale : *Informations Structures et Systèmes*

présentée et soutenue publiquement

par

Carole DUFFET
le Lundi 13 Décembre

Sujet de la thèse :

QUELLE CONFIANCE ACCORDER AU MODÈLE SOLUTION DE LA TOMOGRAPHIE DE RÉFLEXION 3D ?

JURY

Jacques Blum	Rapporteur
Michel Cuer	Directeur de thèse
Marc Daignières	Président du jury
Jean-Robert Grasso	Rapporteur
Patrice Guillaume	Examinateur
Delphine Sinoquet	Encadrant IFP

À mes formidables parents, Michèle et Jean,

Remerciements

J'adresse mes sincères remerciements à

Delphine Sinoquet, ma responsable à l'Institut Français du Pétrole, qui m'a accueillie chaleureusement au sein de la division géophysique, m'a fait découvrir la «tomo» et les joies de l'informatique et m'a fait partager ses grandes compétences et son professionnalisme ! Un GRAND merci pour ton aide et tes précieux conseils et pour la confiance que tu m'as témoignée durant ces trois années de thèse.

Michel Cuer, mon directeur de thèse, pour ses encouragements et son soutien durant ces trois années.

Jacques Blum et Jean-Robert Grasso pour m'avoir fait le grand honneur d'accepter de juger ce travail de thèse.

Marc Daignières et Patrice Guillaume qui ont accepté d'être membres de mon jury de thèse.

Michel Gaven, mon instituteur du CM1 et ami, le premier témoin de mes débuts dans l'univers des mathématiques et de ses applications, merci pour ton soutien : « C'est farimineux, n'est-ce pas, Michel !».

Maud Cavalca, ma p'tite Choupette, avec qui j'ai partagé cette belle aventure de la thèse, pour les moments inoubliables passés ensemble et nos éclats de rire. A nos falaises !

Timothée Perdrizet, mon collègue de bureau, pour m'avoir «supportée» dans le bureau les deux premières années de ma thèse et pour son aide technique précieuse. Finalement, ..., «That's all folks !».

Caroline Joseph, pour m'avoir fait partager ses connaissances géostats et pour les bons moments passés ensemble. Alors, et tes petits patins définitivement oubliés ? !

Judith Pirot, un coeur gros comme ça et une vraie mère pour moi, pour avoir rendu plus facile ma vie à l'IFP, sans oublier, bien sûr, les fameux tupperwares ! ! !

Evelyne Maffiolo, pour sa douceur et sa sensibilité et pour nos longues conversations qui m'ont apporté beaucoup de réconfort dans les moments où on y croit un peu moins fort.

Jean-philippe Cattin, capitaine Nemo au coeur tendre, pour son écoute et ses précieux conseils.

Bruno Léty, pour sa générosité, son humour et sa joie de vivre qui font démarrer la journée d'un bon pied !

Jean-Marc Richard, mon informaticien préféré, une oreille attentive, un homme de bon conseil et un fin mélomane à ses heures : «Mange des tomates mon amour ...», pour m'avoir redonné très souvent le sourire.

Merci à mes collègues et amis stagiaires, thésards ou post-doctorants avec lesquels j'ai partagé des moments de doute et d'enthousiasme durant ces années de recherche : Alexandre, Arnaud, Aymeric, Benoit, Diane, Gilles (Allez l'OM ! !), François, Frédéric A., Frédéric D., Frédéric J., Hervé, Laurent (Lapin !), Laure, Martin, Olivier, Xavier, Yannick (Out of control !) et Yohan (Yoyo, vive les limaces féroces !). Certains ont déjà fini, d'autres y travaillent encore ... Bon courage, donc !

Andreas Ehinger, pour m'avoir accueillie dans sa division à l'Institut Français du Pétrole et pour les discusions qui ont pu enrichir ce travail.

Merci à toutes les personnes des Primevères et des Giroflées avec qui j'ai passé trois années de ma thèse formidables et qui m'ont encouragée dans mon travail et m'ont aidée à m'acclimater à la vie parisienne : Martine Ancel, Axelle Baroni (ma fidèle co-équipière !), Josette Bruneau, Karine Broto, Florence Delprat-Jannaud (LA référence tomo !), Bertrand Duquet, Annne Jardin, Fabrice Jurado (l'expert jerry), Karine Labat, Patrick Lailly (ton fluide informatique est légendaire : Patrick.csh un véritable shell déclencheur de fous rires !), Michel Léger, Danièle Macé-Prieur, Thomas Mensch, Philippe Nivelet, Jacques Quiblier, et bien sûr tous les autres. Un merci tout particulier à nos secrétaires, Sylvie Louis et Christine Le Cleac'h, mes rayons de soleil.

Merci à mes amies (et à leur moitié, bien sûr !) pour m'avoir soutenue et souvent remonté le moral dans les moments difficiles au cours de cette longue épreuve : Caro et Bernard (les aveyronnais de mon coeur sans oublier mes p'tits chouchous Lilou et Clem), Élo et Arnaud (qui m'ont «bichonnée» et qui ont rendu mes trois années dans la capitale heureuses et agréables et «RDV au H&M ! !»), Jojo et Marco (toujours là pour m'aider à RE-LA-TI-VI-SER !), Caro et Laurent (et Loïc ..., merci pour votre soutien).

Et puis, un bisou particulier à Arnaud pour son soutien de chaque instant (et c'est vrai que je n'ai pas toujours été facile ...).

Un grand merci à mes parents, mes supporters inconditionnels, qui, à distance, m'ont toujours soutenue et encouragée quelle que soit la situation. Ce travail, c'est aussi un peu le vôtre...

Et enfin, j'ai une pensée émue pour mes «papis et mamies» : Albert & Yvonne, André & Noële, Marcel & Georgette.

Table des matières

1 INTRODUCTION **11**

2 CONTEXTE **15**
2.1 Introduction générale 15
2.2 Principaux concepts géophysiques pour «imager» le réservoir 17
2.3 Panorama sur les méthodes de détermination d'un modèle de vitesse 18
2.3.1 Analyse de vitesse dans le domaine temps : les corrections NMO et DMO . . . 18
2.3.2 Analyse de vitesse de migration profondeur 20
2.3.3 Inversion par forme d'ondes 21
2.3.4 Inversion de temps de trajet : la tomographie de réflexion 21

3 LA TOMOGRAPHIE DE RÉFLEXION **23**
3.1 Description du modèle de sous-sol 23
3.1.1 Modèles de type blocky 24
3.1.2 Paramétrisation du modèle par des fonctions B-splines 25
3.2 Le problème direct 26
3.2.1 Le tracé de rayons 26
3.2.2 La méthode de tracé de rais utilisée : le bending 27
3.3 Le problème inverse 29
3.3.1 Prise en compte des incertitudes sur les données 31
3.3.1.1 Les incertitudes sur les données 31
3.3.1.2 Comment prendre en compte les incertitudes sur les données ? 32
3.3.2 Régularisation et information a priori 33
3.3.2.1 Régularisation : mise en oeuvre 34
3.3.2.1.1 Choix de poids de régularisation : méthode de continuation . 34
3.3.2.1.2 Critère de la L-curve 35
3.3.2.2 Introduction d'information a priori : l'approche bayesienne 37
3.3.2.3 Résolution du problème inverse 41
3.3.2.3.1 Le problème inverse continu 41
3.3.2.3.2 La méthode de Gauss-Newton 43
3.3.2.3.3 Le problème inverse linéarisé 44
3.3.3 Conclusions 45

4 QUANTIFICATION DES INCERTITUDES SUR LE MODÈLE SOLUTION **47**
4.1 Indétermination du problème inverse 47
4.1.1 Indétermination de la vitesse 47
4.1.1.1 Présentation du modèle 48

4.1.1.2 Inversion 48
4.1.1.3 Exploration de l'ensemble des modèles 48
4.2 Comment estimer les incertitudes : les méthodes classiques 51
4.2.1 Incertitudes sur les paramètres du modèle 54
4.2.2 Décomposition en vecteurs propres de la matrice de covariance a posteriori : détermination des directions de modèle les mieux/les moins bien déterminées . 55
4.3 Deux méthodes proposées 61
4.3.1 Les simulations de modèles admissibles 61
4.3.1.1 Principe 61
4.3.1.2 Applications 61
4.3.1.2.1 Premier exemple (voir Annexe A) 62
4.3.1.2.2 Second exemple (voir Annexe B) 62
4.3.1.3 Conclusions 62
4.3.2 Utilisation de macro-paramètres géologiques pour le calcul d'incertitude 64
4.3.2.1 Principe 64
4.3.2.2 Applications 65
4.3.2.2.1 Premier exemple (voir Annexe A) 65
4.3.2.2.2 Second exemple (voir Annexe B) 66
4.3.2.3 Illustration du macro-paramètre volume 70
4.3.2.3.1 Principe 70
4.3.2.3.2 Application 71
4.3.2.4 Conclusions 72
4.3.3 Premières discussions sur ces deux méthodes 76

5 LIMITES DES MÉTHODES DE CALCUL D'INCERTITUDE : RÉALISME DES INCERTITUDES CALCULÉES 77
5.1 Influence de la paramétrisation et de la discrétisation du modèle sur les incertitudes calculées 77
5.1.1 Importance du choix de paramétrisation du modèle de Terre sur les incertitudes calculées 77
5.1.2 Influence de la discrétisation sur l'analyse d'incertitude 78
5.1.2.1 Choix du nombre de paramètres B-splines 78
5.1.2.2 Résultat théorique et démonstration 79
5.1.2.3 Application : résultats numériques 82
5.2 Influence de l'échantillonnage des données 84
5.2.1 Influence du choix de la norme dans l'espace des données 84
5.2.1.1 Résultats numériques 85
5.2.2 Influence de l'échantillonnage des données 86
5.2.2.1 Résultats numériques 87
5.3 Influence de l'introduction de régularisation et d'information a priori 88
5.3.1 Influence de l'introduction de régularisation 88
5.3.2 Influence de l'introduction d'information a priori (cas anisotrope) 90
5.3.3 Conclusions 94

6 APPLICATIONS SUR DONNEES REELLES **99**
6.1 Première application : HOD 99
6.2 Deuxième application : KIMASI 118
6.2.1 Présentation de l'application 118
6.2.2 Résultats numériques 120

7 CONCLUSIONS **123**

A MODÈLE SYNTHÉTIQUE 1 **127**
A.1 Présentation du modèle 127

B MODÈLE SYNTHÉTIQUE 2 **133**
B.1 Présentation du modèle 133
B.2 La méthode d'inversion utilisée 133

C FONCTIONS B-SPLINES **137**

D CONSTRUCTION DE LA NORME INTERVENANT DANS LE TERME GÉOLOGIQUE : CAS ANISOTROPE **139**
D.1 Calcul de l'expression de la norme en 2D 141
D.2 Calcul de l'expression de la norme en 3D 144

E CONTINUITÉ DE $J(m_n)$ **151**

F UN PROBLÈME BIEN POSÉ **153**

G L'APPROCHE BAYESIENNE **157**

H PUBLICATIONS/COMMUNICATIONS À CONGRÈS **161**

Chapitre 1

INTRODUCTION

L'exploration pétrolière vise à détecter et à estimer avec le plus de précision possible les ressources en hydrocarbure du sous-sol. Ces hydrocarbures sont piégés dans des structures géologiques par nature tri-dimensionnelles et d'une complexité plus ou moins grande. Pour leurs opérations d'exploration, les compagnies pétrolières consacrent un effort important de recherche dans le développement de nouveaux outils afin de mieux déterminer la géométrie et la nature des réservoirs d'hydrocarbure. La méthode la plus utilisée est la sismique réflexion. Cette méthode consiste à engendrer des ébranlements dans le sol (grâce à des explosifs, camions vibrateurs en sismique terrestre, canons à air ou à eau en sismique marine) et à enregistrer en surface les ondes réfléchies par les discontinuités d'impédance acoustique.

Les structures géologiques qui nous intéressent sont de nature complexe et nécessitent des techniques d'imagerie particulières. Dans ce contexte, la tomographie de réflexion est l'une de ces méthodes. Le but de la tomographie est la reconstruction du champ de vitesse de propagation et de la géométrie des interfaces ou réflecteurs, à partir des temps de trajet des ondes sismiques associés à des réflexions sur des interfaces géologiques. On parle alors d'inversion des temps d'arrivée. Nous choisissons une paramétrisation pour les modèles de type «blocky» qui autorise des discontinuités de part et d'autre des réflecteurs (par opposition aux modèles lisses). Le problème direct consiste à calculer les temps de trajet des ondes se propageant dans un milieu donné pour des réflecteurs donnés. Les données du problème direct sont donc les interfaces et la distribution de vitesse ainsi que le dispositif d'acquisition des données, les inconnues sont les temps de trajet. Le problème inverse consiste à comparer les temps associés aux réflexions majeures pointés sur les sections sismiques et les temps calculés par le problème direct dans le modèle courant. Le modèle ajustant «au mieux» les temps pointés est la solution du problème inverse.

Le modèle solution de ce problème inverse est le modèle qui minimise une fonctionnelle de type moindres carrés mesurant les écarts entre les temps de trajet mesurés et les temps de trajet calculés par tracé de rayons dans ce modèle (problème direct). Ce problème non linéaire est résolu par la méthode de Gauss-Newton, par linéarisations successives. Ce modèle obtenu est un modèle parmi les modèles admissibles. En effet, l'accès au temps de trajet peut être très difficile dans le cas de structures géologiques complexes. Il est souvent difficile de distinguer les évènements qui nous intéressent : repérer les réflexions, savoir à quelles interfaces géologiques elles correspondent, etc. Ces temps de trajet observés, données d'entrée de la tomographie, entachés d'erreur et plus généralement la sous-détermination du problème inverse entraînent des incertitudes sur le modèle solution de la tomographie de réflexion. Il est alors essentiel de quantifier ces incertitudes.

Classiquement l'analyse de la solution du problème inverse consiste à comparer les écarts entre les temps de trajet observés et les temps de trajet calculés en calculant la valeur du résidu moyen et la valeur du résidu maximum, et en étudiant la répartition spatiale de ces résidus (zones locales avec des résidus élevés). Cette analyse est souvent insuffisante : ce modèle est un des modèles qui vérifie les données, une analyse plus fine devra être menée pour quantifier l'ensemble des modèles admissibles que nous pouvons obtenir à partir de ces données.

Le travail de la thèse a consisté à étudier et à développer différentes méthodes d'estimation des incertitudes a posteriori.

Après un état des lieux du contexte de notre étude (chapitre 2) et une présentation détaillée du problème de tomographie de réflexion (chapitre 3), le travail de thèse a concerné les méthodes étudiées pour accéder aux incertitudes sur le modèle solution. A ce niveau nous faisons un point sur les méthodes classiques d'analyse des incertitudes a posteriori et proposons deux approches appropriées à notre problème (chapitre 4).

Pour cette étude des différentes méthodes d'estimation des incertitudes a posteriori nous nous plaçons dans le cadre linéarisé, c'est à dire au voisinage de la solution. Comme cela a été fait par Tarantola, 1987, et Gouveia et Scales, 1997, nous étudierons la matrice de covariance a posteriori (voir aussi der Sluis and der Vost, 1987, Nolet et al., 1999). L'analyse de cette matrice (inverse de la matrice Hessienne) donne les incertitudes sur le modèle solution mais son calcul est généralement très coûteux (la matrice est de grande taille pour des problèmes 3D) et l'interprétation physique des résultats est difficile. Nous mènerons ensuite une analyse d'incertitude analogue à celle faite par Delprat-Jannaud et Lailly, 1992, par exemple. Elle repose sur l'analyse des valeurs propres et vecteurs propres de la matrice hessienne et permet d'accéder aux directions de modèle les mieux (respectivement moins bien) déterminées, à condition d'utiliser la norme appropriée.

Après avoir établi les limites de chacune de ces méthodes, nous proposerons deux approches adaptées à notre problème pour quantifier les incertitudes sur le modèle solution de la tomographie de réflexion : la méthode de simulation de modèles et l'approche macro-paramètre.

La première permet d'échantillonner la fonction de densité de probabilité a posteriori et donc de simuler des modèles équiprobables. La seconde permet de calculer les variances de quantités géologiques intéressantes, décrites par des combinaisons linéaires de paramètres (ce que nous appelons des «macro-paramètres»), avec un temps de calcul réduit (les matrices qui sont manipulées sont réduites à l'espace des macro-paramètres). Ce formalisme général permet une grande flexibilité pour la construction des macro-paramètres (moyenne de la vitesse dans une région, épaisseur d'une couche, volume d'un piège pétrolier). Cependant, nous devons garder en tête que cette approche (comme les approches classiques) est valide uniquement dans un voisinage du modèle solution (cadre linéarisé) et que les cas complexes peuvent nécessiter une approche non linéaire.

Une deuxième étape de la thèse est consacrée à l' étude des limites de ces deux méthodes de calcul d'incertitude et du réalisme de ces incertitudes (chapitre 5). En effet, l'analyse via ces deux approches a été réalisée jusqu'alors, sous certaines hypothèses : cadre linéaire (au voisinage de la solution), à paramétrisation fixée, poids de régularisation fixés, échantillonnage des données fixé, etc. Le but est donc de relaxer ces hypothèses et de voir l'influence sur les incertitudes calculées. Chaque étude est illustrée par une application sur un exemple synthétique.

L'analyse d'incertitude linéarisée permet de délimiter l'ensemble des modèles solutions possibles qui vérifient, avec la précision espérée, les données temps de trajet et l'information a priori. Néanmoins cette approche est uniquement valide au voisinage du modèle solution et les cas complexes peuvent nécessiter une approche non-linéaire qui consiste à rechercher les minima locaux de la fonction coût (chapitre 6). Une approche expérimentale consiste à résoudre le problème inverse sous contraintes (par exemple sur la position d'un réflecteur, sur un gradient vertical de vitesse, etc.) en vue de tester diverses hypothèses ou scenarii géologiques. On réalise donc plusieurs inversions permettant de tester différents scenarii géologiques en vue de délimiter l'espace des solutions admissibles. Ces limites de l'approche linéarisée seront illustrées par deux applications sur données réelles 2D et 3D.

Chapitre 2

CONTEXTE

2.1 Introduction générale

La géophysique est la discipline scientifique qui s'attache à l'étude des phénomènes physiques affectant le globe terrestre et son atmosphère (Dictionnaire Universel Francophone, Hachette). Elle permet la description du sous-sol, à partir de mesures télémétriques.

L'exploration pétrolière vise à détecter et à estimer, avec autant de précision que possible, les ressources en hydrocarbures du sous-sol. Ces hydrocarbures sont piégés dans des structures géologiques par nature tridimensionnelles et d'une complexité plus ou moins grande. Pour leurs opérations d'exploration, les compagnies pétrolières consacrent un effort important de recherche dans le développement de nouveaux outils afin de mieux déterminer la géométrie et la nature de leurs réservoirs d'hydrocarbures.

Parmi les techniques utilisées à grande échelle pour caractériser le sous-sol, on trouve des mesures gravimétriques, magnétiques, électriques, électromagnétiques et sismiques. Les premières s'intéressent aux variations locales du champ de pesanteur, significatives d'une variation de la densité des roches sous-jacentes ; les secondes travaillent sur la variation du champ magnétique terrestre, ou du paramètre de conductivité du sous-sol pour les deux suivantes. Enfin les mesures sismiques dépendent des caractéristiques de propagation des ondes des roches traversées. On trouve ici deux grandes familles de disciplines, travaillant soit au moyen de signaux naturels créés par les tremblements de terre, soit de signaux générés par des ébranlements artificiels produits par explosion ou par vibrations du sol.

À plus petite échelle, de multiples techniques existent, pour caractériser différentes propriétés des roches. Ainsi, le grand nombre d'outils de sondage de puits (logging) existant actuellement illustre la quantité importante de méthodes concevables.

Ces différentes techniques sont appliquées en exploration pétrolière, en fonction de l'état d'avancement de la recherche en cours. Les techniques gravimétriques et magnétiques sont utilisées à l'échelle du bassin sédimentaire, pour une première définition des prospects envisageables. Les techniques sismiques (avec signal généré) permettent une image précise de la géométrie des couches sédimentaires, et une localisation des pièges pétroliers potentiels. Enfin, les sondages de puits conduisent à un ajustement fin des modèles de remplissage sédimentaire, et au calibrage des modèles de propagation des fluides.

Dans ce contexte, la tomographie* qui contribue à retrouver la structure du sous-sol revêt un intérêt stratégique.

Bois et al., 1971, ont introduit la tomographie pour la sismique de transmission, méthode très proche de la tomographie médicale. Plus tard, Bording et al., 1987, et Bishop et al., 1985, ont développé la tomographie de réflexion. Au lieu de rayons X, les sismologues† utilisent, pour imager la surface proche de la Terre, les ondes sismiques, résultant de tremblements de terre importants. Ces ondes traversent la Terre et ralentissent ou s'accélèrent en profondeur. Les données de la tomographie sismique sont des temps de propagation de l'onde entre l'épicentre du séisme et diverses stations sur le globe terrestre. Ces mesures permettent une estimation des variations de vitesses sismiques associées. Cette tomographie permet de reconstituer au moins d'une manière approchée la structure des roches traversées par les ondes sismiques.

Nous allons nous intéresser dans tout ce qui va suivre au problème de sismique réflexion (Lavergne, 1986). La prospection par réflexion sismique est une méthode de la géophysique appliquée‡ utilisée dans la recherche du pétrole et qui permet d'obtenir des informations structurales et/ou quantitatives sur les couches géologiques du sous-sol jusqu'à plusieurs kilomètres de profondeur.

Dans une campagne de réflexion sismique, on déplace, sur la surface de la région de la Terre à explorer, un dispositif constitué d'une source sismique, camion vibrateur ou explosif enterré en sismique terrestre, canon à eau ou autre dispositif immergé dans la mer à quelques dizaines de mètres en sismique marine, l'autre partie du dispositif étant constituée d'une nappe de capteurs. En sismique terrestre, ces capteurs sont des géophones et mesurent, en fonction du temps, le mouvement du sol aux points où ils sont placés. En sismique marine, les capteurs sont des hydrophones placés à quelques mètres de profondeur et mesurent, toujours en fonction du temps, la pression dans l'eau. Actuellement, on développe des campagnes, essentiellement pour le «monitoring» des exploitations des gisements de pétrole, où en plus des hydrophones on ajoute des capteurs de la même nature que ceux utilisés en sismique terrestre et posés au fond de la mer.

Pour chaque «tir», c'est à dire chaque position de la source et pour chaque capteur, on dispose donc, en général d'une fonction réelle du temps. C'est à partir de cette famille de fonctions d'une variable, le temps, famille indicée par les positions des sources et des capteurs qu'il faut extraire des informations sur la constitution du sous-sol. On notera que ces fonctions ne sont que les enregistrements, observés en surface, des échos et réverbérations des ondes acoustiques et élastiques qui se propagent à la suite de chaque tir. Compte-tenu de la complexité des phénomènes de propagation des vibrations dans le sous-sol et du fait que les signaux importants sont de petites amplitudes donc noyés dans du «bruit», l'extraction de l'information n'est pas possible sans traitements numériques. Le traitement a donc pour but l'amélioration du rapport signal sur bruit et la mise en forme de l'information pour faciliter l'interprétation.

*Le terme «tomographie» (du grec *tomos* : coupe) provient du fait que les données du problème sont les intégrales, le long de lignes appelées rayons, de l'inverse de la vitesse des ondes sismiques.

†A noter la différence entre le terme de «sismologie» qui s'attache à l'étude des séismes, et plus généralement des mouvements du sous-sol, et celui de «sismique» qui associé à la prospection correspond à une méthode géophysique qui va être détaillée par la suite.

‡La plupart des méthodes de la géophysique, les mesures du champ de gravité de la Terre, puis du champ magnétique, les sondages électriques et électromagnétiques, et bien sûr exploitation des données de la sismologie, conduisent d'ailleurs à des problèmes inverses.

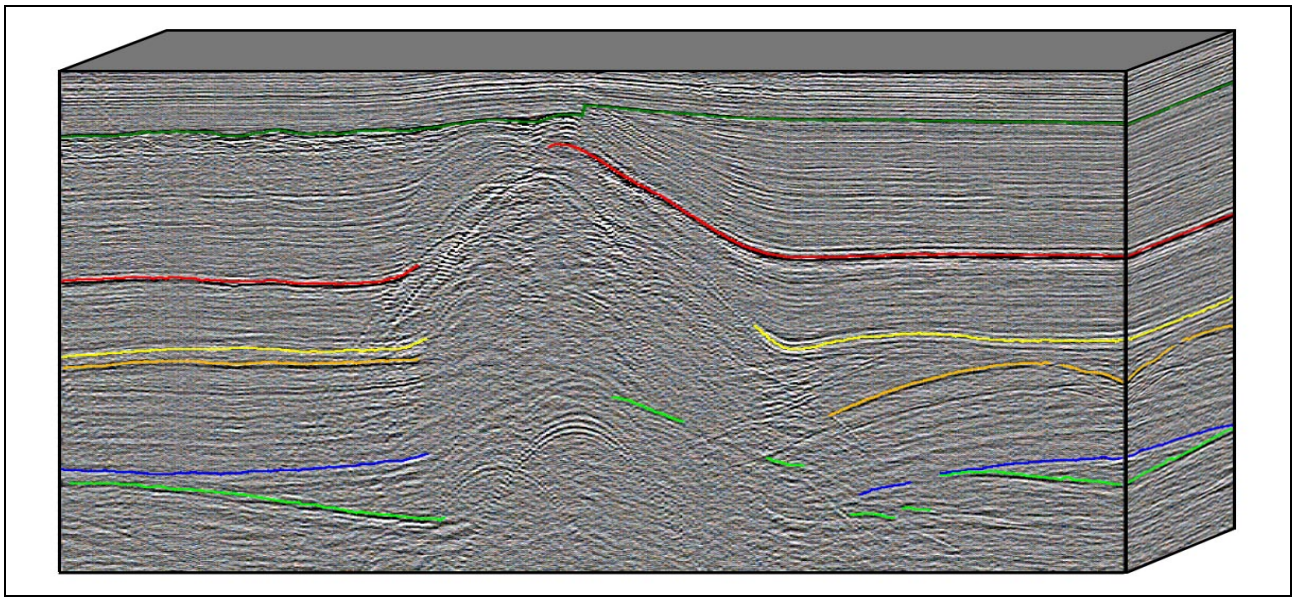

FIG. 2.1 – **Illustration de l'interprétation d'une coupe sismique :** les évènements sismiques identifiés par des marques de couleur correspondent à des contrastes de vitesse majeur. (D'après Broto, 1996).

Après traitement des données, on obtient une première image de la coupe plane verticale appelée section sismique (voir Figure 2.1). Le but de la tomographie est la reconstruction du champ de vitesse de propagation et des interfaces ou réflecteurs, à partir des temps de trajet des ondes sismiques associés à des réflexions sur des interfaces géologiques. On parle alors d'inversion des temps d'arrivée et le modèle de vitesse obtenu sera alors utilisé pour l'imagerie des réservoirs.

2.2 Principaux concepts géophysiques pour «imager» le réservoir

L'imagerie sismique des premiers kilomètres du sous-sol constitue l'une des étapes initiales à mettre en oeuvre avant toute phase d'exploitation pétrolière. Elle permet d'obtenir une première représentation des structures géologiques et d'avoir une idée de la localisation des réservoirs potentiels, pour autant que l'image soit de bonne qualité. Une image mal focalisée peut conduire à sur ou sous estimer la capacité potentielle d'un réservoir et donc modifier son intérêt économique, ou encore à implanter des puits dans des zones sans pétrole.

La construction d'une image du sous-sol de bonne qualité, directement utilisable par le géologue, s'obtient de manière non triviale par l'opération dite de migration qui permet de repositionner en profondeur les données enregistrées en temps à la surface après leur avoir fait subir un premier traitement (corrections d'amplitudes, filtrage du bruit, déconvolution, élimination des multiples (Pellé, 2003), corrections statiques, etc.).

Cette étape nécessite la disponibilité d'une modélisation de la loi physique liant le domaine temps et le domaine profondeur et d'un modèle de vitesse du sous-sol. Actuellement, différentes méthodes permettent de résoudre la conversion temps-profondeur de façon fiable, même dans le cas de propaga-

tion des ondes les plus complexes. En particulier, dans le cas de milieux présentant des réflecteurs à géométrie complexe et/ou des vitesses variant latéralement, c'est la migration profondeur avant sommation qui est utilisée pour replacer les images temporelles des réflecteurs à leur position latérale et verticale réelle dans le sous-sol.

Diverses méthodes existent pour réaliser la migration : par tracé de rayons (migration de Kirchhoff, voir French, 1975, Schneider, 1978, Berkhout, 1980), en domaine de Fourier (migration FK, phase-shift, voir Stolt, 1978), ou par continuation (Claerbout, 1976). Notons que, indépendamment de la méthode adoptée, l'étape de migration demande un temps de calcul considérable, compte-tenu de la quantité «gigantesque» de données sismiques.

Malgré de nombreux progrès, la difficulté principale de l'imagerie sismique dans le cas de structures complexes réside aujourd'hui encore dans la détermination du modèle de vitesse du sous-sol. En effet, l'opération d'imagerie nécessite de connaître les vitesses de propagation des ondes dans le sous-sol pour passer du domaine temps (celui des données) au domaine profondeur (celui de l'image). Or, ces vitesses sont précisément a priori inconnues puisque le processus d'imagerie a lieu pendant la phase d'exploration.

2.3 Panorama sur les méthodes de détermination d'un modèle de vitesse

Il existe plusieurs méthodes de détermination du modèle de vitesse du sous-sol. Dans ce qui suit, nous allons en donner une brève description avant de décrire *in fine* celle qui nous intéresse : la tomographie de réflexion.

2.3.1 Analyse de vitesse dans le domaine temps : les corrections NMO et DMO

Ce qui suit est illustré par la Figure 2.2.

Dans le cas d'un seul réflecteur plan horizontal délimitant une couche homogène, les données d'un tir particulier sont des arrivées qui s'alignent dans le plan «distance horizontale - temps», le long d'une hyperbole : en effet, si h est la profondeur du réflecteur, si x est l'offset c'est à dire la distance horizontale entre la source et le capteur et si v est la vitesse des ondes dans la couche, le temps d'arrivée au capteur, $t(x)$, d'un signal émis par la source vérifie

$$t(x) = \frac{2}{v}\sqrt{\left(\frac{x}{2}\right)^2 + h^2} \tag{2.1}$$

et donc

$$t^2(x) = \frac{4h^2}{v^2} + \frac{x^2}{v^2}. \tag{2.2}$$

Dans le cas de plusieurs couches homogènes horizontales superposées, cette formule n'est plus exacte mais est une première approximation. Cela conduit au concept d'analyse de vitesse : pour déterminer la vitesse des ondes sismiques à la verticale d'un tir particulier, on pourrait au moins en première approximation si les réflecteurs sont presque horizontaux, considérer à chaque profondeur h, les hyperboles d'équations $t^2(x) - \frac{x^2}{v^2} = \frac{4h^2}{v^2}$ et ce pour toutes les vitesses v raisonnables.

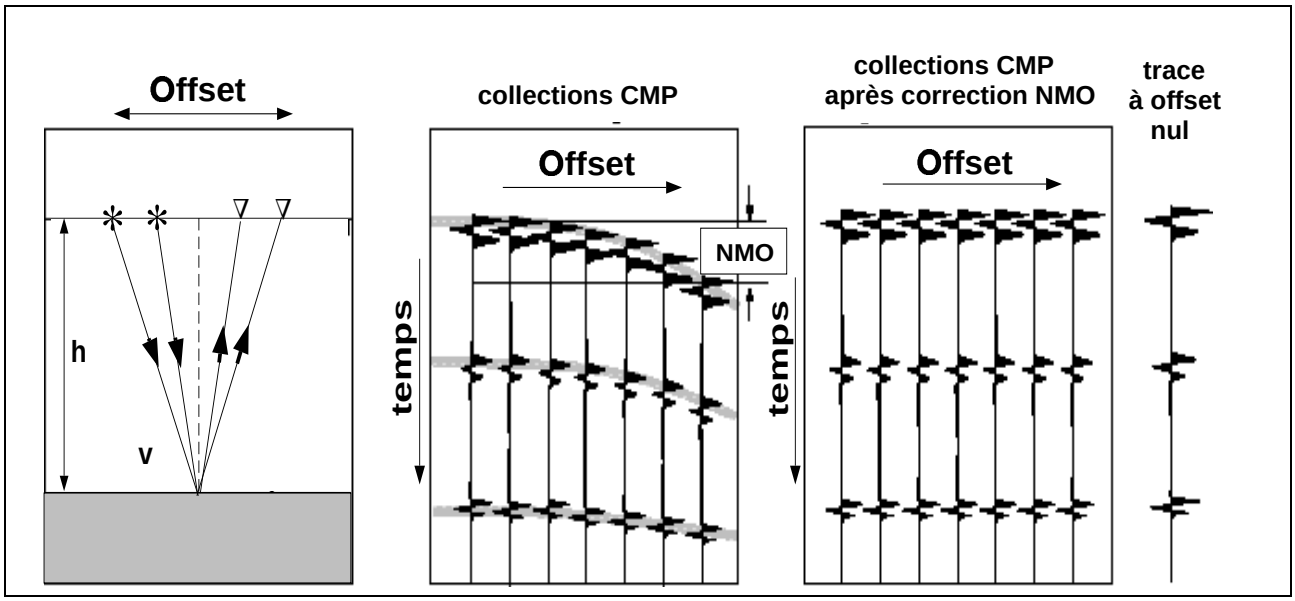

FIG. 2.2 – **La correction NMO - Réflexion sur la base d'une couche plane horizontale homogène et isotrope.** Les traces sismiques réorganisées en collections point milieu commun présentent un évènement sismique qui varie hyperboliquement avec l'offset. A chaque trace sismique d'une collection point milieu commun est appliquée une correction temporelle $-\Delta t_{NMO}$. Puis les traces sismiques sont sommées de façon à obtenir la trace sismique à offset nul associée au point milieu commun considéré.

Pour choisir la «bonne» vitesse, on somme toutes les traces du tir particulier le long de ces hyperboles et on retient la vitesse qui donne la plus grande somme. En pratique, on ne fait pas cette opération sur les traces provenant d'un tir particulier, mais sur toutes les traces telles que le milieu du segment joignant le point source au capteur soit fixe : on dit qu'on met en commun les points milieux (CMP pour «Common Mid Point» ou CDP «Common Depth Point»). Cela ne change pas grand chose en théorie mais cela moyenne les différences dues au fait que les sources sismiques ne sont pas parfaitement répétitives et il peut y avoir plus de points milieux que de tirs.

Ensuite on reporte la somme obtenue au temps $t = \frac{2h}{v}$ à la verticale du point milieu. A noter que sommer les signaux le long de l'hyperbole d'équation $t^2(x) - \frac{x^2}{v^2} = \frac{4h^2}{v^2}$ revient à décaler le signal en temps de $\delta t(x) = \sqrt{\frac{4h^2}{v^2} + \frac{x^2}{v^2}} - t$. On parle de «correction NMO» (*Normal Move Out*, i.e. Courbure Normale) (voir Figure 2.2).

Il existe une autre correction, dite «correction DMO» (*Dip Move Out*, Déplacement «amont-pendage») comme moyen de détermination du modèle de vitesse mais qui reste inapplicable dès que l'on ne se trouve plus dans le cas de réflecteurs plans pentés et d'un milieu de propagation homogène. En bref, cette correction DMO correspond à une correction NMO dans le cas de réflecteurs pentés et de conflits de vitesse (cas de la migration temps). Dans la pratique, la correction DMO est généralement appliquée en supposant que le milieu est à vitesse constante, c'est à dire en utilisant un opérateur DMO elliptique. Il s'avère que les résultats obtenus rapportés à leur coût, sont satisfaisants dans de nombreux cas (Deregowski, 1986).

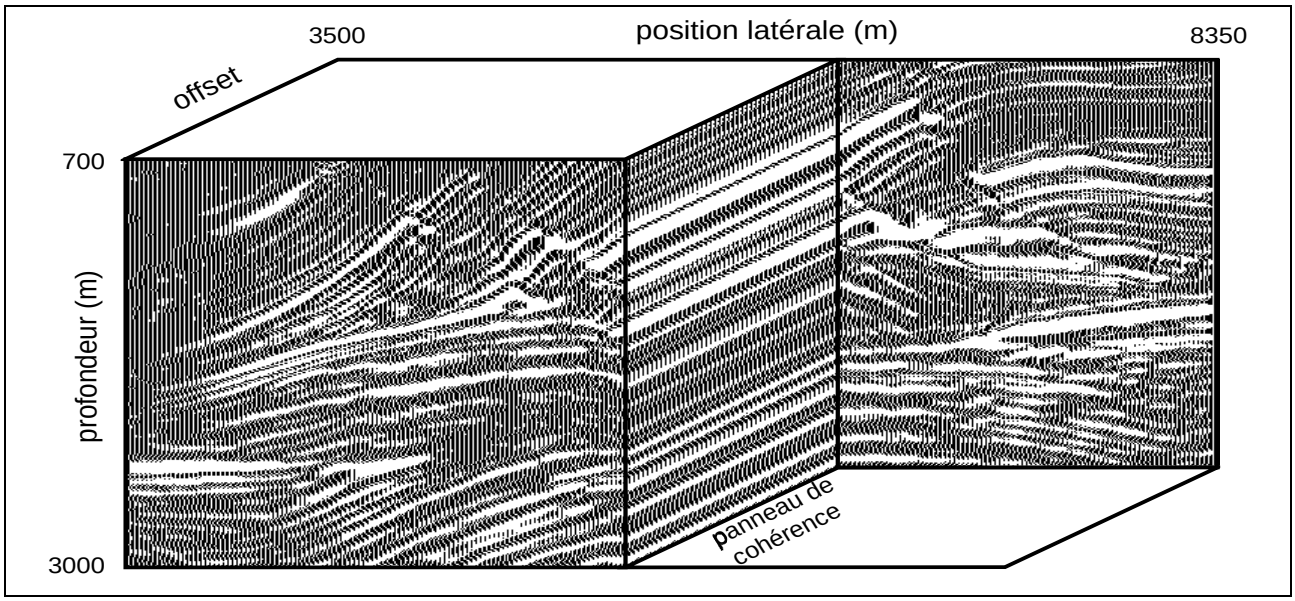

FIG. 2.3 – **Principe des méthodes d'analyse de vitesse de migration.** La migration profondeur avant sommation avec un modèle de vitesse cinématiquement correct produit des images qui sont identiques d'un offset à un autre. Par conséquent, les évènements sismiques dans les sections verticales pour une position latérale donnée (panneaux de cohérence) sont plats et horizontaux, indépendamment de la complexité de la structure imagée. (D'après Duquet, 1996).

2.3.2 Analyse de vitesse de migration profondeur

Les méthodes de type analyse de vitesse de migration sont toutes basées sur l'idée que l'opération de migration profondeur est très sensible à la qualité du modèle de vitesse utilisé (Gardner et al., 1974). Après une «migration profondeur avant sommation», si on utilise le modèle de vitesse exact, les images profondeurs individuelles devraient montrer la même image structurale du sous-sol. Par contre, si le modèle de vitesse utilisé est erroné, les images profondeurs individuelles représentent un même réflecteur à des positions différentes. De plus, la forme de ce réflecteur change d'une image à l'autre. L'examen des différences entre ces images permet alors de juger le modèle de vitesse et, le cas échéant, de l'améliorer pour tendre vers le modèle exact. Pour cela, on peut par exemple utiliser les panneaux de cohérence qui sont constitués des traces du cube d'images profondeur associées à la même position latérale (Figure 2.3). Cet examen est en effet réalisé en analysant la déformation des évènements contenus dans ces derniers (qui devraient être plats et continus lorsque l'on utilise le modèle de vitesse exact).

D'autres méthodes utilisant ce principe, telles que :
- les méthodes DFA («Depth Focusing Analysis»),
- les méthodes RCA («Residual Curvature Analysis»),
- les tomographies MVA («Migration velocity Analysis»),

ont été développées (voir Broto, 1996, pour le détail de ces différentes méthodes). Elles diffèrent soit par le critère utilisé pour mesurer l'horizontalité des panneaux de cohérence (analyse de focalisation (méthodes DFA) ou de courbe résiduelle (méthodes RCA)), soit par le moyen utilisé pour la mise à jour du modèle de vitesse (mise à jour locale (méthodes DFA, RCA) ou inversion globale (tomographies MVA)).

2.3.3 Inversion par forme d'ondes

Nous avons vu dans la section 2.2 que la détermination du modèle de vitesse de propagation des ondes sismiques dans le sous-sol suivie d'une migration avant ou après sommation permettait de transformer des sections sismiques dans le domaine temps en sections sismiques dans le domaine profondeur. L'imagerie sismique utilise donc uniquement l'information cinématique contenue dans les traces. Notons à ce niveau, que la tomographie utilisant essentiellement, pour l'instant, cette information cinématique, l'inversion des temps de trajet des ondes est bien une technique appropriée. L'information dynamique, c'est à dire l'amplitude des réflexions, n'est donc pas prise en compte en imagerie sismique.

Cette information est cependant très utile, elle peut permettre d'avoir accès à des propriétés physiques du sous-sol (autres que la vitesse de propagation des ondes sismiques), comme par exemple la densité des roches, ou encore l'impédance (produit de la densité des roches par la vitesse de propagation des ondes sismiques). L'inversion par forme d'ondes (Noble et Tarantola, 1991), qui vise à rechercher le modèle de vitesse dont la réponse sismique (constituée cette fois-ci à la fois des temps de trajet des ondes réfléchies et des amplitudes associées) s'ajuste au mieux sur les traces sismiques, permet d'utiliser cette information. Notons que cette technique d'inversion par forme d'ondes peut être aussi bien utilisée après sommation (Bamberger et al., 1982, Brac et al., 1992) qu'avant sommation (Bourgeois et al., 1989, Kolb et al., 1986, Lambaré et al., 1992).

Dans l'inversion par formes d'ondes, tous les types d'ondes (c'est à dire y compris les arrivées directes, les arrivées multiples, les ondes converties, etc) étant pris en compte, la redondance d'information devrait permettre de résoudre ce problème de façon satisfaisante. Le caractère fortement non linéaire de l'équation caractéristique de l'inversion par formes d'ondes, ainsi que le nombre très important de données prises en compte dans l'inversion rendent cette méthode difficilement applicable. En pratique, on est confronté à des problèmes de convergence vers la «vraie» solution et à un coût de calcul très élevé. C'est pourquoi il semble préférable d'utiliser la tomographie de réflexion, qui, elle, est beaucoup moins coûteuse en terme de temps de calcul.

2.3.4 Inversion de temps de trajet : la tomographie de réflexion

Les méthodes de détermination du modèle de vitesse dites temporelles décrites précédemment (correction NMO et correction DMO, voir section 2.3.1) supposent une variation hyperbolique des temps de trajet des ondes réfléchies en fonction de l'offset (l'offset correspondant à la distance source-récepteur). Dès que l'on ne se trouve plus dans le cas mono-couche plane horizontale, cette hypothèse n'est au mieux vérifiée que pour les faibles offsets. Plus précisément, cette hypothèse est en fait quasiment toujours fausse mais les erreurs commises (i.e. incertitudes provenant de cette hypothèse) restent acceptables jusqu'à un certain degré de complexité de la propagation des ondes lié aux variations latérales de vitesse qui conduisent à la non-hyperbolicité des réflexions. Or la considération d'une faible gamme d'offsets ne permet pas de lever l'ambiguïté vitesse-profondeur : plusieurs hyperboles caractérisées par des vitesses et profondeurs différentes peuvent alors correctement ajuster la courbure d'un évènement sismique. D'autre part, pour que cette hypothèse soit vérifiée, le milieu ne doit pas présenter de variations latérales de vitesses à l'intérieur des couches et des pendages structuraux trop importants. Ces méthodes sont donc impropres à la détermination du modèle de vitesse dans le cas de structures complexes, qui sont précisément caractérisées par des variations latérales de vitesse et/ou des réflecteurs qui peuvent être non seulement pentés mais aussi à géométrie complexe.

La tomographie de réflexion apparaît alors comme un outil «puissant», puisqu'elle ne demande

aucune hypothèse quant aux variations des temps de trajet en fonction de l'offset. Cette méthode a très rapidement connu un grand succès en géophysique pétrolière (voir par exemple Woodward et al., 1998, Guillaume et Zhang, 2001). En effet, elle est particulièrement appropriée lorsqu'il s'agit de déterminer un modèle de vitesse composé de structures complexes puisqu'elle permet de s'affranchir des hypothèses 1D interdisant toute variation latérale de vitesse et toute interface non horizontale.

En tomographie de réflexion il est possible d'inverser toutes sortes de données dont par exemple :
- les réflexions multiples (Cavalca et lailly, 2004),
- les arrivées multiples (Clarke, 1997).

Dans notre étude, seuls les temps d'arrivée des réflexions primaires sont considérés. Pour obtenir ces temps d'arrivée, une interprétation des données sismiques est nécessaire. Cette interprétaton est délicate, notamment pour les structures complexes, puisque les données peuvent être entachées d'erreurs de mesure mais aussi d'erreurs d'interprétation (voir section 3.3.1.1). Il est évident à ce stade, que la confiance que l'on aura en le modèle solution de la tomographie de réflexion est directement liée à la qualité des données interprétées et que l'on aura à inverser ; données qui dépendent, elles, de la qualité des données sismiques enregistrées.

Par ailleurs (et nous approfondirons cela dans le chapitre suivant), la formulation du problème de tomographie de réflexion, correspondant à la minimisation dans le processus d'inversion d'une fonctionnelle moindres carrés, permet une intégration facile d'information de différents types comme par exemple, des contraintes de calage en profondeur au puits (Le Stunff et Grenié, 1998), en contenant les vitesses et les profondeurs d'interfaces dans certaines marges (idéalement selon les connaissances géologiques) (voir Lailly et Sinoquet, 1996), ou encore en orientant des régularisations selon certaines directions ou pentes motivées par des contraintes géologiques structurales (Clapp et Claerbout, 1998).

Toutefois, la principale difficulté de la tomographie de réflexion est «le pointé» des temps de trajet (Apostoiu-Marin et Delprat-Jannaud, 1995). Cette étape d'interprétation des données sismiques qui consiste à identifier dans les enregistrements sismiques les évènements correspondant à des réflexions primaires est délicate : elle peut conduire à des données d'inversion erronées et/ou incomplètes. Billette et Lambaré, 1998, ont proposé la stéréotomographie, qui est une méthode de tomographie de pente (et non plus seulement de temps de trajet) qui permet, en théorie au moins, de s'affranchir de cette difficulté.

Avant d'appliquer la tomographie, soit à des modèles synthétiques simples, soit à des structures géologiques complexes (applications sur données réelles), nous allons étudier (d'un point de vue théorique) dans le chapitre suivant cette méthode qui semble bien appropriée pour calculer la vitesse de propagation des ondes.

Chapitre 3

LA TOMOGRAPHIE DE RÉFLEXION

Nous avons vu dans le chapitre précédent que l'imagerie sismique nécessite la détermination aussi précise que possible de la distribution des vitesses de propagation du sous-sol pour effectuer des migrations. C'est pourquoi les géophysiciens de l'exploration pétrolière se sont intéressés de très près à la tomographie de réflexion pour essayer de déterminer cette distribution des vitesses à partir de la cinématique des données sismiques. Ainsi dans le cas de structures géologiques complexes, la tomographie de réflexion remplacerait l'analyse de vitesse.

Dans ce chapitre, nous allons voir que la tomographie est un problème inverse pour lequel nous avons choisi une formulation de moindres carrés qui consiste à chercher le modèle qui minimise les résidus entre les temps observés et les temps calculés par le problème direct et nous expliciterons les motivations de ce choix. Nous utilisons pour résoudre le problème direct associé, la technique du tracé de rayons entre deux points (i.e. «two-point raytracing») : nous calculons ainsi pour un modèle donné, le temps de trajet le long du rayon reliant une source et un capteur.

L'accès au temps de trajet peut être difficile dans le cas de structures géologiques complexes. Il est souvent difficile de distinguer les évènements qui nous intéressent : repérer les réflexions pertinentes, savoir à quelles interfaces géologiques elles correspondent, etc. Ces temps de trajet observés, données d'entrée de la tomographie, sont donc entachés d'erreur. Il est alors essentiel de quantifier ces incertitudes et de les intégrer dans le formalisme du problème inverse pour une meilleure résolution. Nous verrons comment prendre en compte ces erreurs sur les données et comment introduire dans la formulation du problème inverse des informations a priori sur le modèle.

Notons que le consortium KIM a développé un outil de tomographie de réflexion 3D, «*jerry*» (Jurado et al., 1996a, et Jurado et al., 1996b). De nombreuses applications ont été réalisées et ont montré l'efficacité de cette méthode (Broto et Ehinger, 1998, Ehinger et al., 2001, Stopin et Ehinger, 2001 et Broto et Ehinger, 2001).

3.1 Description du modèle de sous-sol

Considérons des milieux tri-dimensionnels et notons $\Omega = I_x \times I_y \times I_z \subset \mathbb{R}^3$, $I_x = [x_{min}, x_{max}]$, $I_y = [y_{min}, y_{max}]$ et $I_z = [z_{min}, z_{max}]$.

Définissons, avant toute chose, la représentation mathématique du modèle de sous-sol étudié : un

modèle m est consitué d'une distribution de vitesses $V(x, y, z) \in \Omega$ et de N^Z réflecteurs $Z_i(x, y)$ ($x \in I_x$, $y \in I_y$ et $i = 1, ..., N^Z$).

Actuellement deux types de représentation du milieu sont généralement utilisés :

- les modèles **lisses** (Lailly et Sinoquet, 1996, Billette et Lambaré, 1998) qui se composent d'une distribution unique de la vitesse définie globalement sur Ω et des interfaces. La vitesse et les réflecteurs sont supposés être au moins $\mathcal{C}^1$ sur Ω.

- les modèles dits **blocky** (défini par bloc) qui repose sur une partition du milieu en couches. Ces modèles se composent des distributions de vitesse associées aux couches et des interfaces qui délimitent les différentes couches. La vitesse peut présenter des discontinuités aux interfaces et est supposé au moins $\mathcal{C}^1$ dans les couches.

Dans notre approche, nous représentons le modèle de Terre à l'aide de blocs (représentation **Blocky**, voir Figure 3.1) : les interfaces et les vitesses dans chaque couche sont paramétrées par des fonctions B-splines cubiques (cf. Annexe C ou de Boor, 1978, pour plus de détails concernant la théorie des B-Splines cubiques).

3.1.1 Modèles de type blocky

L'utilisation des modèles *blocky* (Lines et Treitel, 1985, Chiu et Stewart, 1987), modèles présentant des discontinuités de vitesse de part et d'autre d'une interface (voir Figure 3.1), permet de travailler avec des distributions de vitesse très régulières à l'intérieur des couches (les variations de grande amplitude ont été absorbées par les discontinuités) ce qui nous assure une très bonne stabilité même avec des données clairsemées et de limiter le nombre de coefficients B-splines pour décrire ces vitesses. Ce choix nous permet d'autre part de simplifier le tracé de rayons comme cela sera expliqué dans la section 3.2.

Les interfaces délimitent des blocs dans le modèle de vitesse et à chaque bloc est associée une loi de vitesse définie par :

$$v_0(x, y) + k(x, y).(z - z_0) \quad \text{ou} \quad v(x, y, z), \tag{3.1}$$

où $v_0(x, y)$ et $k(x, y)$ sont respectivement les variations latérales de vitesse à la profondeur de référence z_0 et k le gradient vertical de vitesse.

Notons que l'on peut prendre en compte la vitesse des ondes P (ondes de compression) et des ondes S (ondes de cisaillement) et qu'il est également possible d'introduire des coefficients, η et δ, permettant de prendre en compte l'anisotropie des vitesses. Le paramètre η peut être vu comme une mesure de la non-ellipticité de la vitesse, alors que δ contrôle la propagation quasi-verticale de la vitesse des ondes P (voir Thomsen, 1986, pour une définition précise de ces paramètres d'anisotropie).

Chaque interface est représentée par une fonction explicite :

$$z = Z(x, y) \quad \text{ou} \quad x = X(y, z) \quad \text{ou} \quad y = Y(z, x). \tag{3.2}$$

Ce choix nous permet de repérer un bloc de vitesse par sa position relative aux interfaces (signe de

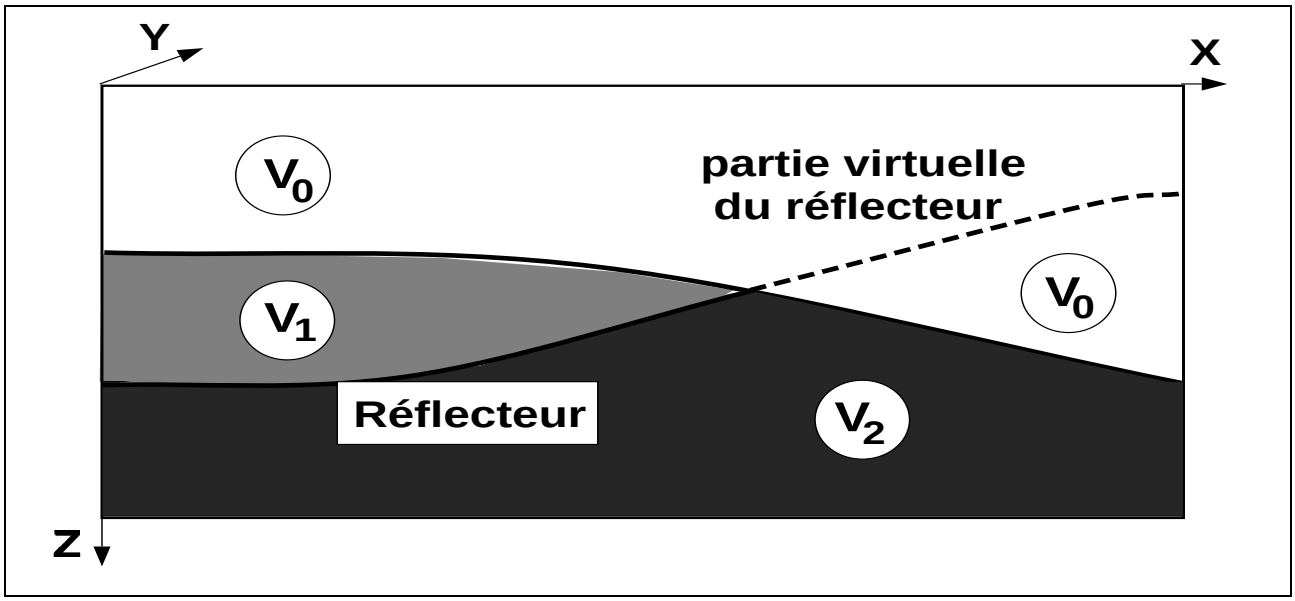

FIG. 3.1 – **Le modèle :** illustration d'un modèle blocky de sous-sol.

$z - Z(x, y)$ par exemple).

3.1.2 Paramétrisation du modèle par des fonctions B-splines

On appelle «paramétrisation» l'ensemble des variables indépendantes d'un problème. En général la paramétrisation d'un problème n'est pas unique et la complexité des calculs dépendra de la dimension de l'espace de paramétrisation. Cette étape de paramétrisation est donc importante car elle caractérise la taille du problème d'optimisation, mais aussi définit l'espace dans lequel on peut trouver des solutions optimales.

Dans notre approche, le modèle, m, est composé de fonctions B-splines 2D ou 3D décrivant les variations de vitesse dans une couche (3.1) et de fonctions B-splines 2D décrivant les interfaces (3.2) (pour une description détaillée de notre paramétrisation du modèle, voir aussi Clarke, 1996a et Clarke, 1996b).

Cette paramétrisation du modèle repose sur le choix de fonctions deux fois différentiables (i.e. de classe $\mathcal{C}^2$) pour décrire la géométrie des interfaces et des vitesses dans les couches. Cette paramétrisation permet, d'une part, de modéliser des structures complexes et, d'autre part, de calculer les dérivées des temps de trajet (jacobienne) exactement, ce qui est essentiel pour le problème inverse.

Ainsi, une interface explicite en z est définie par

$$z = Z(x, y) = \sum_{u=0}^{N_x^Z - 1} \sum_{v=0}^{N_y^Z - 1} Z^{uv} B_u(x) B_v(y), \tag{3.3}$$

où les Z^{uv} sont les $N_x^Z * N_y^Z$ paramètres de B-spline dans les directions indépendantes x et y, et où les

B_u, B_v sont les fonctions de base de la représentation B-spline. Les interfaces sont définies sur tout le domaine et peuvent se croiser.
Les vitesses peuvent être 2D ou 3D et sont alors définies par

$$v(x,y) = \sum_{a=0}^{N_x^V-1} \sum_{b=0}^{N_y^V-1} V^{ab} B_a(x) B_b(y), \tag{3.4}$$

ou par

$$v(x,y,z) = \sum_{a=0}^{N_x^V-1} \sum_{b=0}^{N_y^V-1} \sum_{c=0}^{N_z^V-1} V^{abc} B_a(x) B_b(y) B_c(z), \tag{3.5}$$

où les V^{ab} sont les $N_x^V * N_y^V$ paramètres de B-spline dans les directions x et y et les V^{abc} sont les $N_x^V * N_y^V * N_z^V$ paramètres de B-spline dans les directions x, y et z.

3.2 Le problème direct

3.2.1 Le tracé de rayons

La propagation des ondes dans un matériau élastique isotrope est décrite par l'équation des ondes. Cette équation découle directement des «grandes» lois de la physique. La solution analytique de cette équation est possible par l'introduction de fonctions de Green, mais l'obtention de solutions explicites n'est en général pas possible, dès que l'on se trouve en présence de milieu complexe.

Dans le cas d'un milieu quelconque, plusieurs approximations de l'équation des ondes permettent de décrire de façon approchée la propagation des ondes dans un milieu isotrope non homogène. Elles sont fondées sur l'équation de propagation des ondes dans un milieu élastique, et reposent sur différents types d'approximations. Nous allons nous intéresser ici à la propagation des ondes acoustiques et introduire l'approximation de l'équation des ondes choisie : l'approximation asymptotique haute fréquence.

L'objectif du problème direct de la tomographie de réflexion consiste à calculer les temps de trajet des réflexions primaires sur les interfaces d'un modèle de vitesse du sous-sol. Dans ce problème direct, nous cherchons à déterminer le rayon qui relie une source et un récepteur donnés et qui satisfasse au principe de Fermat* : on parle de «tracé de rayons point à point» (ou en anglais «two-point raytracing»). Ce tracé de rayons consiste à fournir une solution haute fréquence de l'équation des ondes. Nous nous contenterons ici de donner les équations (pour plus de détails voir Jannaud, 1992).

Etant donnés un milieu modélisé (géométrie des interfaces, champ de vitesse de propagation des ondes), un dispositif d'acquisition des données (sources S et récepteurs R), un ensemble de signatures de rais (une signature (S, R, i) est un ensemble (source, récepteur, liste des interfaces rencontrées (y compris le réflecteur)), voir Figure 3.3), on cherche à calculer les temps de trajet $T_k(m)$ des ondes se propageant dans la distribution de vitesse et se réfléchissant sur une ou plusieurs interfaces.

Le temps de trajet associé à un triplet (S, R, i) est donné par :

$$T_k(m) = \int_{\mathcal{R}_k} \frac{ds}{v}, \tag{3.6}$$

*Ce principe établit que le temps de trajet total associé à une trajectoire entre un point source S et un récepteur R doit être stationnaire. La trajectoire particulière qui vérifie le principe de Fermat est le rai.

où s est l'abscisse curviligne le long du rayon $\mathcal{R}_k$ et l'indice k correspond au triplet (S, R, i).

Une source sismique émet des ondes dans toutes les directions. Quand on observe le champ d'ondes associé (Figure 3.2), on peut isoler, entre la source et la station d'enregistrement, un rai sismique qui est perpendiculaire en tout point au front d'ondes qui se propage depuis la source. Un front d'ondes correspond à une surface de temps de propagation constant. Dans notre cas, les fronts d'ondes sont donc définis par les courbes ou surfaces d'équation $T(x) = constante$. On définit alors les rais comme des trajectoires orthogonales en tout point aux fronts d'ondes.

Dans un milieu uniforme, pour une source ponctuelle, le front d'ondes est une sphère, et le rai sismique une demi-droite. Mais dans un milieu hétérogène, la vitesse varie spatialement, le front d'ondes est plus compliqué et reflète la structure du milieu. La trajectoire du rai sismique n'est plus une droite, mais est toujours perpendiculaire en chaque point au front d'ondes. La théorie des rais étudie la trajectoire des rais d'ondes sismiques dans les milieux hétérogènes. Comme en optique, on doit supposer que la longueur d'onde est suffisamment petite devant la taille des domaines où les caractéristiques du sous-sol varient notablement.

L'équation de l'Eikonal

$$(\nabla T(x))^2 = \frac{1}{c^2(x)}, \tag{3.7}$$

permet le calcul des temps de trajet. C'est une équation aux dérivées partielles en général difficile à intégrer qui décrit la courbure du rayon.

La trajectoire du rayon est alors définie par son équation paramétrique $(x(s), p(s))$, donnée par :

$$\begin{cases} \frac{dx}{ds} = c(x)p(s) \\ \frac{dp}{ds} = \frac{c(x(s))}{2}\nabla\left(\frac{1}{c^2(x)}\right) \end{cases}. \tag{3.8}$$

Ces équations du tracé de rais (3.8) sont très utilisées en géophysique pétrolière pour calculer la propagation d'un champ d'onde dans un milieu hétérogène (Červený, 2001). Bien que les conditions d'application théoriques supposent que la longueur d'onde caractéristique de la source soit plus petite que les longueurs d'ondes spatiales caractéristiques du milieu de propagation, l'expérience montre qu'en général la théorie des rais peut être appliquée dans la majorité des cas réels traités en exploration pétrolière (même dans des milieux fortement hétérogènes).

3.2.2 La méthode de tracé de rais utilisée : le bending

Dans le logiciel de tomographie de réflexion, «*jerry*», la méthode de tracé de rais utilisée est une méthode de bending[†].

Le principe de la méthode de «bending», méthode de courbure de la trajectoire (Guiziou, 1993, Thisse, 1994, Jurado et al., 1996a, Jurado et al., 1997 et Jurado et al., 1998), consiste à fragmenter

[†]Dans les méthodes de tracé de rais point à point il existe aussi des méthodes dites de shooting (V. Červený, I. A. Molotkov et I. Pšenčík, 1977, R. T. Langan, I. Lerche et R. T. Cutler, 1985, Virieux et al., 1988, J. Virieus et D. Farra, 1991, Jannaud, 1992, Clarke, 1997) : elles consistent à calculer un rai à partir d'une direction intiale à la source, puis à perturber cette direction jusqu'à ce que le rayon aboutisse au récepteur visé. Le terme de «shooting» vient du fait que l'on «tire» un rayon à partir de la source.

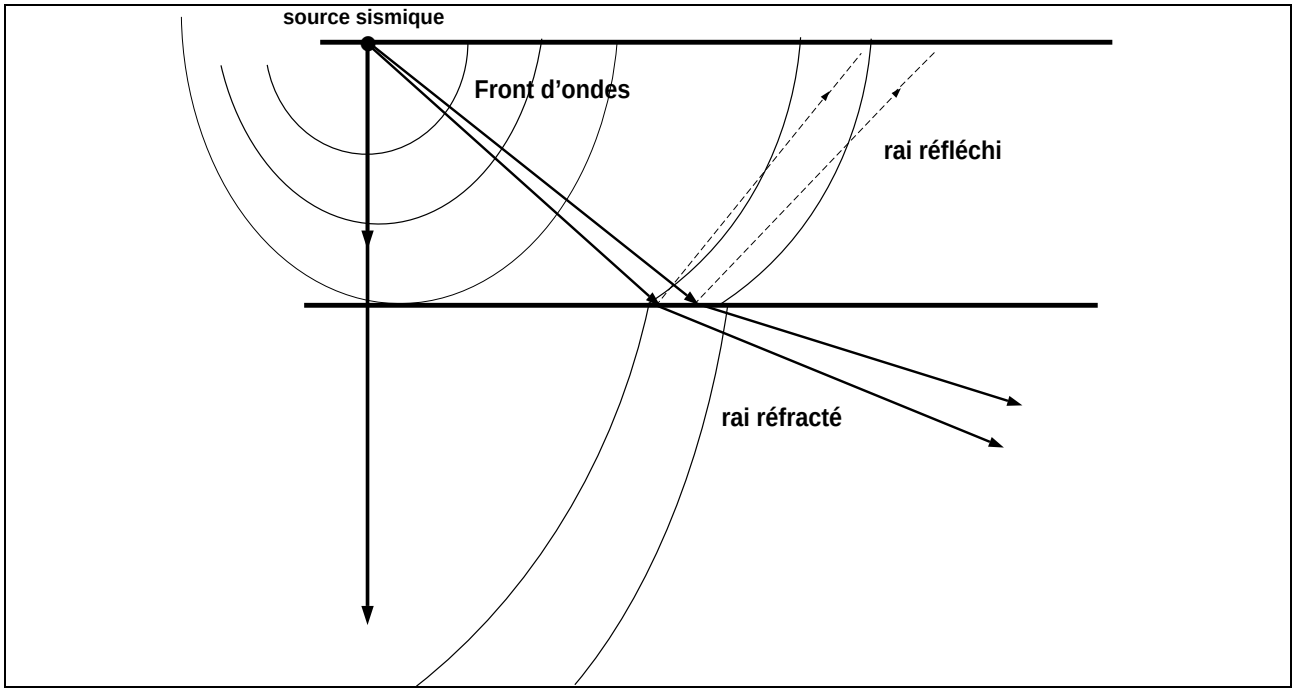

FIG. 3.2 – Rais sismiques associés à la propagation des ondes.

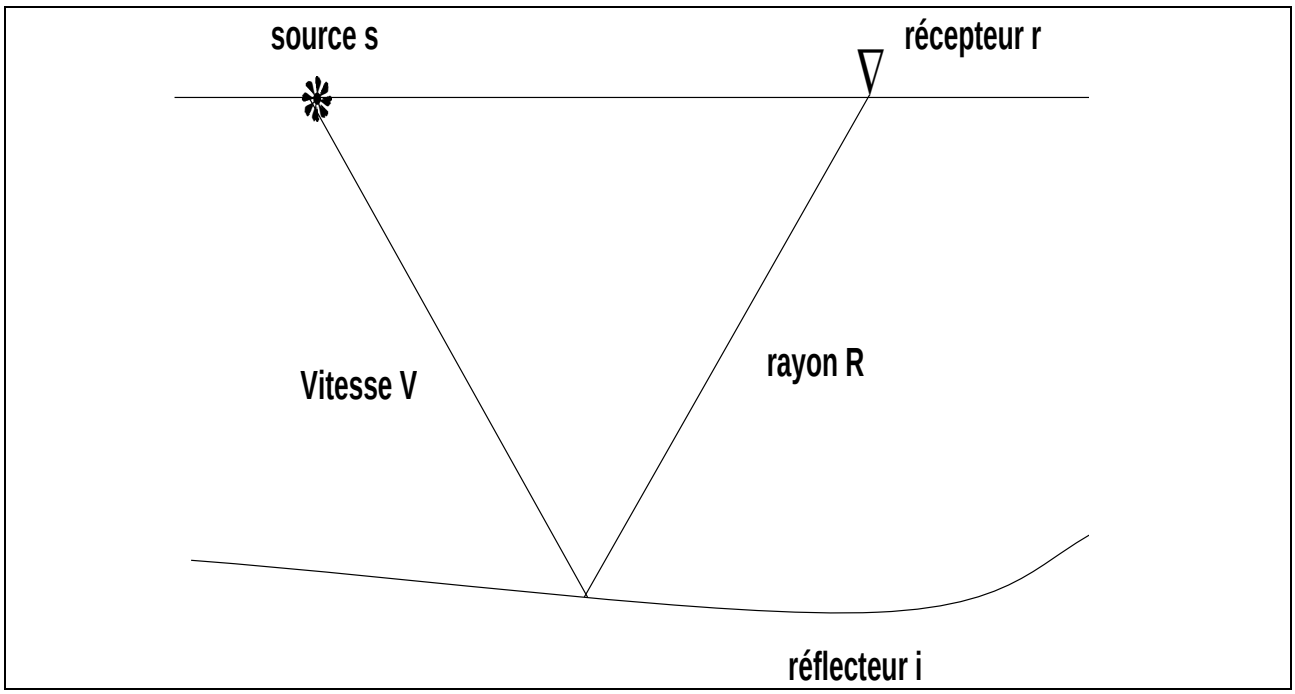

FIG. 3.3 – **Illustration d'un rayon de signature** (S, R, i) **:** l'onde est émise de la source S, se propage dans le champ de vitesse V, se réfléchit sur le réflecteur i et se propage jusqu'à la surface au récepteur r où le temps de trajet est enregistré.

la trajectoire d'un rayon en segments reliant les points d'impact. Chaque segment est défini par une courbe analytique (Figure 3.4). Le premier (respectivement le dernier) segment relie la source (respectivement le récepteur) à un point d'impact. Cette méthode de bending suppose que la trajectoire analytique reliant deux points d'impact est une droite. Cette hypothèse est valable car les variations de vitesse à l'intérieur d'un bloc sont supposées faibles, les principales hétérogénéités étant modélisées par les interfaces.

Le temps de trajet le long de la trajectoire $(\mathcal{P}) = (P_0, ..., P_N)$ d'un rayon est donné par

$$t_{P_0 \to P_N}(\mathcal{P}) = \sum_{n=1}^{N} t_n(\mathcal{P}), \tag{3.9}$$

où P_0 est la source et P_N le récepteur, t_n est le temps de trajet entre les points d'impact successifs P_{n-1} et P_n et N le nombre de segment définissant la trajectoire du rayon.

La méthode de bending consiste tout d'abord à initialiser une première trajectoire (Figure 3.4). Puis les positions des points d'impact sur les différentes interfaces sont perturbées jusqu'à ce que la trajectoire vérifie le principe de Fermat.

Avec les notations introduites précédemment, cela revient à rechercher la trajectoire $(\mathcal{P}) = (P_0, \ldots, P_N)$ solution du problème

$$\frac{\partial t_{P_0 \to P_N}}{\partial P_n}(\mathcal{P}) = 0, \qquad 1 \leq n \leq (N-1). \tag{3.10}$$

Notons que $\forall n \in [1; N-1]$, le point d'impact P_n n'influe que sur les temps de trajet partiels t_n et t_{n+1}. Nous recherchons donc la trajectoire $(\mathcal{P})$ qui vérifie

$$\frac{\partial (t_n + t_{n+1})}{\partial P_n}(\tilde{\mathcal{P}}) = 0, \qquad 1 \leq n \leq (N-1). \tag{3.11}$$

Le système (3.11) est résolu par une méthode de quasi-Newton[‡] qui fournira les temps de trajet localement minimals et non pas les points selles. Les dérivées nécessaires à ce calcul peuvent être trouvées dans Jurado et al., 1998.

3.3 Le problème inverse

Un problème inverse consiste à déterminer les paramètres d'un système physique à partir de la connaissance de la réponse de ce système à une excitation donnée, et de la connaissance d'un modèle théorique qui permet de générer des réponses synthétiques. Dans le cas du problème inverse de la tomographie, il s'agit de retrouver les paramètres de vitesse du sous-sol (modèle) à partir des temps de trajet des ondes sismiques : le problème inverse en tomographie de réflexion consiste donc à déterminer le ou un modèle qui satisfait les temps de trajet observés.

Le pointé de ces temps de trajet sur les sections sismiques n'étant pas exact et les données pouvant être bruitées, le modèle solution peut ne pas satisfaire exactement les données. Le problème inverse en

[‡]On peut aussi utiliser une méthode classique de Newton. Cependant les dérivées secondes des temps de trajet par rapport aux coordonnées spatiales des points d'impacts sont difficilement calculables.

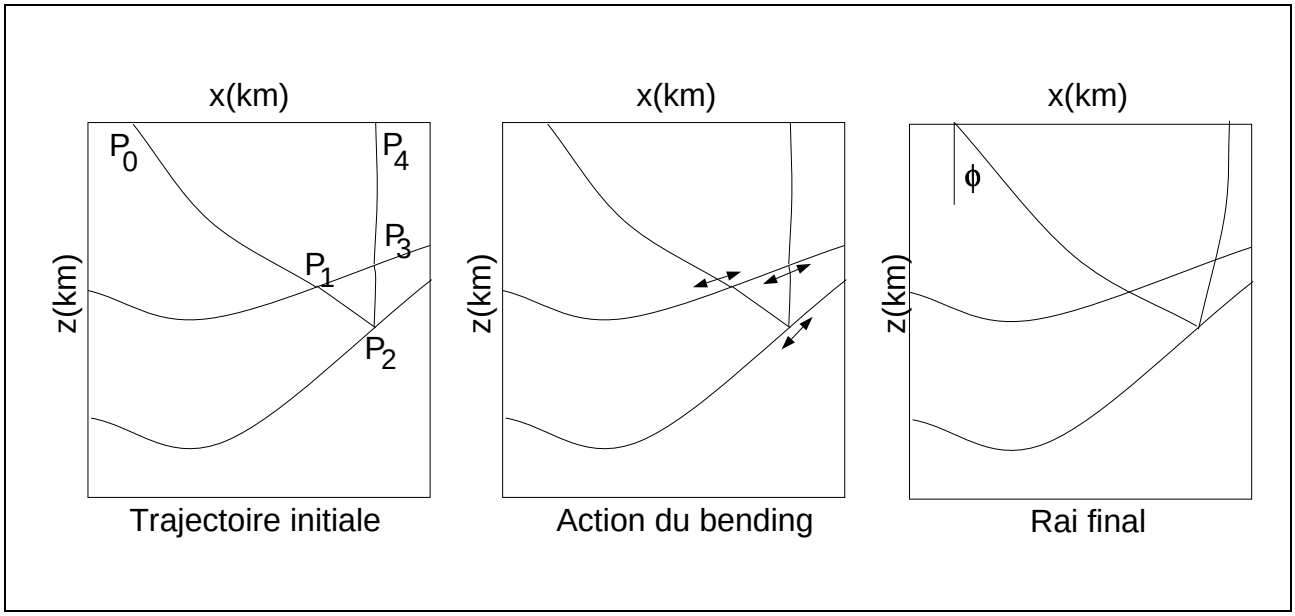

FIG. 3.4 – **Le tracé de rais :** illustration de la méthode de bending.

tomographie de réflexion est posé sous la forme d'un problème de moindres carrés.

Il s'agit donc de minimiser une fonction coût qui mesure l'écart entre les données observées et les données synthétiques associées à un modèle par le biais de la norme L^2 définie sur l'espace des données. Notons que les principales difficultés proviennent tout d'abord de la taille de l'espace des données (de l'ordre de 500000 données) et de l'espace des modèles (de l'ordre de 50000 inconnues).

Soit T l'application de modélisation qui permet de générer des données synthétiques T^{cal} associées à un modèle m. Cette application opère de l'espace des modèles $\mathcal{M}$ vers celui des données $\mathcal{D}$:

$$T^{cal} = T(m). \tag{3.12}$$

On notera $T^{cal}(m)$, les temps de trajet calculés par tracé de rayon sur le modèle m.

La plupart des problèmes inverses sont mal posés (Backus et Gilbert, 1970, Lines et Treitel, 1984, et Delprat-Jannaud, 1991) au sens où une des trois conditions suivantes n'est pas respectée :

- existence d'une solution,
- unicité de la solution,
- stabilité de la solution par rapport aux données ; en d'autres termes, une petite erreur sur les données entraîne une petite erreur sur la solution.

En effet, l'inverse de l'opérateur de modélisation, T, s'il existe, n'est la plupart du temps pas continu. Quant aux données observées, outre les bruits incohérents et cohérents (mais indésirables), elles sont enregistrées sur un temps limité, avec des dispositifs qui induisent des incertitudes sur les

mesures.

De part leur aspect mal posé, les problèmes inverses sont très souvent sensibles aux incertitudes. Nous verrons dans le paragraphe 3.3.1.1 que ces incertitudes sont nombreuses et proviennent de diverses origines. Et nous verrons dans le paragraphe 3.3.1.2 comment prendre en compte ces incertitudes sur les données dans la formulation du problème inverse.

D'autre part, le problème inverse peut être exprimé comme un problème d'optimisation : «trouver le modèle qui explique au mieux les données». Le terme «explique au mieux» doit être défini et requiert l'introduction d'une fonction coût résumant la «qualité» du modèle donné. Cette fonction coût doit tenir compte de la connaissance que l'on a a priori sur le modèle. A ce stade deux approches sont utilisées de façon classique : l'approche déterministe et l'approche bayesienne.

Le problème inverse consite donc, respectivement, à minimiser la fonction coût régularisée (Delprat-Jannaud and Lailly, 1993) :

$$C(m) = \frac{1}{2}||T^{cal}(m) - T^{obs}||^2_{C_d^{-1}} + \epsilon^2||D^2(m - m^{prior})||^2_{L^2}, \tag{3.13}$$

ou à minimiser (Gouveia and Scales, 1997, Mitsuhata, 2004) :

$$C(m) = \frac{1}{2}||T^{cal}(m) - T^{obs}||^2_{C_d^{-1}} + ||m - m^{prior}||^2_{C_m^{-1}}, \tag{3.14}$$

avec C_d l'opérateur de covariance a priori sur les données, m^{prior} le modèle a priori et C_m l'opérateur de covariance a priori sur les modèles qui permet de décrire les erreurs sur le modèle a priori.

Dans ces deux approches, régulariser le problème inverse consiste à introduire des informations additionnelles sur la solution dans le but de déterminer une solution parmi plusieurs solutions possibles. Il y a deux types d'informations : les informations a priori sur le modèle dues à des considérations physiques (3.14) et la régularisation (3.13) (on modifie la fonction coût en ajoutant un terme de stabilisation qui permet, en général, de pénaliser une solution improbable (oscillations par exemple)). Nous allons dans ce qui suit préciser ces deux formulations de la tomographie de réflexion (section 3.3.2.1 et 3.3.2.2), ou plus exactement la façon d'introduire cette information a priori sur le modèle.

3.3.1 Prise en compte des incertitudes sur les données

3.3.1.1 Les incertitudes sur les données

Il est intéressant ici d'être en mesure de déterminer les différentes sources d'erreurs ou approximations réalisées pour en connaître ensuite l'impact sur le modèle solution. De plus, Il faut souligner qu'une mauvaise prise en compte des incertitudes associées aux données et de leurs corrélations introduit un biais sur la loi de probabilité a posteriori et donc sur la localisation de son maximum.

Ces incertitudes sur les données peuvent être dues à des erreurs d'acquisition, d'interprétation et de modélisation.

- Les principales sources d'incertitude au niveau du processus d'acquisition sont dues à une mauvaise connaissance de la position des éléments du dispositif (en particulier le suivi de la position

des flûtes tirées par un bateau est difficile). Une imprécision des mesures provient aussi de la sensibilité des capteurs utilisés.

- Au niveau de l'interprétation, les principales sources d'incertitude sont liées au pointé des horizons et des failles : il est difficile de suivre un évènement continu sur les sections sismiques.

- Au cours des différentes étapes du traitement des données, on utilise de nombreuses approximations, qui permettent de modéliser les phénomènes réels au moyen d'équations simples. Chacune de ces modélisations est entachée d'erreur, en voici quelques exemples :

  - L'utilisation de la norme euclidienne dans la fonctionnelle de moindres carrés revient à considérer que les données ne sont pas corrélées en espace : en conséquence, l'échantillonnage des données influe sur la solution du problème. Renard, 2002, a étudié la prise en compte de ces corrélations entre les données.

  - Le bruit de modélisation est aussi une source d'erreur car en général il est fortement corrélé. Par exemple, la prise en compte de l'anisotropie de propagation dans la modélisation (Mensch et Farra, 1998, ou Stopin, 2000) peut être importante.

  - Une autre source d'erreur provient de la paramétrisation du modèle : comme nous l'avons vu dans la section 3.1 la représentation «blocky» du modèle autorise des discontinuités de vitesse au niveau des interfaces et contraint les variations verticales de vitesse à des formes très simples, une telle représentation introduit un biais sur les données modélisées.

3.3.1.2 Comment prendre en compte les incertitudes sur les données ?

Ce que l'on cherche à faire en tomographie, c'est résoudre d'une manière satisfaisante le problème inverse. Pour cela il faut prendre en compte explicitement les erreurs sur les données : non seulement celle de la discrétisation mais aussi celles dues aux mesures et celle dues à la modélisation. La *théorie de probabilités* est alors l'outil idéal. On utilise cet outil pour représenter l'incertitude sur les mesures et/ou le manque de connaissance parfaite sur une grandeur par une loi de probabilité. Dans ce cas, l'approche utilisée est l'approche bayesienne (voir Annexe G).

Dans cette approche, la matrice de covariance C_d des équations (3.13) et (3.14) définit les incertitudes sur les données. Ces erreurs sont modélisées par des distributions de probabilité gaussiennes.

De plus, en ce qui concerne cet opérateur, il a été montré par Tarantola, 1987, que, sous l'hypothèse gaussienne, la matrice de covariance C_d est en fait une somme de covariances individuelles représentant d'une part les erreurs d'observation et d'autre part les erreurs de modélisation. Ce résultat peut être utilisé pour incorporer différents types d'incertitudes sur les données comme celles décrites dans le paragraphe précédent (paragraphe 3.3.1.1).

D'autre part, toujours relativement à la prise en compte des incertitudes sur les données, Renard et Lailly, 2001, ont proposé un formalisme permettant d'intégrer explicitement notre connaissance a priori des corrélations spatiales des données. Ce qui permet de remédier au manque de réalisme du

choix d'une norme L^2 dans l'espace des données, choix qui revient à considérer que les erreurs sur les données sont spatialement indépendantes, ce qui semble peu réaliste dès lors que les données sont associées à des couples source-récepteur proches en surface.

3.3.2 Régularisation et information a priori

L'introduction de régularisation et/ou d'information a priori dans la formulation de moindres carrés est nécessaire pour bien poser le problème. En effet, l'inversion tomographique de données géophysiques (données sismiques, gravimétriques, électromagnétiques, etc.) peut être contrainte par l'information dont on dispose sur la géologie de la zone où la campagne géophysique a eu lieu. On introduit alors cette information, qui concerne les paramètres de l'espace des modèles, comme information a priori dans le problème inverse.

En pratique, construire un modèle a priori est difficile et le terme de la fonction coût contenant l'information a priori, est souvent remplacé par des termes de régularisation composés des dérivées secondes du modèle (ce qui correspond à une régularisation sur la courbure du modèle), ce qui donne $C_m^{-1} = \mathcal{Q}_{reg}$ et $m^{prior} = 0$ dans (3.14). Dans ce cas, trouver la solution du problème régularisé revient à trouver le modèle le plus lisse qui vérifie les données avec la précision voulue .

Delbos et al., 2004, introduisent des informations géologiques a priori en résolvant sous contraintes le problème inverse. Nous verrons par la suite que nous pouvons introduire différents types de contraintes selon les différents types d'information a priori. Ces contraintes peuvent être :

- des contraintes de natures physiques différentes : vitesse ou interface, trois ordres de dérivations (par exemple une interface peut être contrainte sur sa profondeur, sa pente et/ou sa courbure),

- des contraintes d'égalité ou d'inégalité (valeur fixé du gradient de vitesse, profondeur minimale d'une interface),

- des contraintes locales ou globales (information locale provenant d'un puits, courbure d'une interface dans une région particulière).

Sinoquet, 1995, est un autre exemple d'introduction d'un terme supplémentaire dans la fonction coût pour permettre l'incorporation d'informations a priori adaptées aux modèles lisses.

Concernant l'introduction d'information a priori, nous étudierons dans ce qui suit, une méthode qui consiste à introduire de l'information a priori via la construction d'une norme intervenant dans le terme géologique de la fonction coût (3.14) (Tarantola, 1987, Tonellot, 2000) et par ce biais on construit une matrice de covariance a priori sur le modèle. Ceci conduit à un problème de réglage des poids qui peut être délicat notamment quand il y a un conflit entre les différentes informations et les données sismiques. Une estimation de l'ordre de grandeur des incertitudes associées à chaque information permet de choisir les poids. La conclusion est que l'utilisation de la matrice de covariance a priori, C_m^{-1}, peut se comparer à une régularisation, mais est plus flexible quant aux différents types d'information que l'on peut introduire, le problème étant de la construire et de la manipuler pour des problèmes de grandes tailles.

Dans ce qui suit nous allons donc décrire ces deux techniques d'introduction d'information a priori.

3.3.2.1 Régularisation : mise en oeuvre

La régularisation des problèmes mal posés, due initialement à Tikhonov et Arsenin, 1977, cherche à re-définir les notions d'inversion et de solution (quasi-solution, solution approchée, ...) de manière à ce que la «solution régularisée» obtenue par «inversion régularisée» dépende continûment des données et soit proche de la solution exacte (en supposant que celle-ci existe pour des données proches des valeurs effectivement obtenues par la mesure). En d'autres termes, on remplace le problème initial mal posé par un autre «proche» du premier et bien posé.

Une régularisation convenable du problème de tomographie de réflexion originel, initialement proposée dans Constable et al., 1987, consiste à rajouter dans la fonction coût un terme contenant la norme des dérivées secondes des vitesses de couche ainsi que des interfaces : cette méthode dite de continuation revient à baisser les poids au fur et à mesure (Bube et Langan, 1994, Renard et Lailly, 1999).

Dans ce qui suit, nous allons donc considérer la version régularisée du problème inverse linéarisé et préciser la méthodologie d'inversion.

Nous cherchons donc à minimiser la fonction coût suivante :

$$C(m) = ||T^{cal}(m) - T^{obs}||^2_{C_d^{-1}} + \epsilon^2 ||D^2(m - m^{prior})||^2_{L^2}, \tag{3.15}$$

où m^{prior} est le modèle a priori et ϵ^2 est un poids qui permet de régler l'effet de la régularisation.

Nous allons ici expliciter la méthode utilisée pour déterminer le paramètre ϵ, puisque la régularisation du problème inverse dépend du choix du paramètre ϵ. Ce choix est assez délicat :

- Si on donne un poids trop important à la fonctionnelle «stabilisatrice» $||D^2(m - m^{prior})||^2_{L^2}$ la solution peut être altérée de façon très importante (des données différentes peuvent être vérifiées), c'est à dire que l'on ne va pas suffisamment minimiser les résidus temps de trajet.

- Si le poids est trop faible, on donne trop d'importance aux irrégularités du modèle, et par voie de conséquence on a une mauvaise convergence : l'effet régularisateur n'est plus actif et l'inversion devient instable.

Le choix de ϵ va donc correspondre à un compromis entre la stabilité et la vraisemblance de la solution. En pratique, on utilise le critère de la L-curve (voir ci-dessous pour une description détaillée de ce critère).

3.3.2.1.1 Choix de poids de régularisation : méthode de continuation

Pour le choix du paramètre ϵ, nous suivons la méthodologie décrite par Bube et Langan, 1994 (méthode de continuation) et Renard et Lailly, 1999, méthodologie similaire à l'approche Occam mentionnée dans Gouveia et Scales, 1997 : en partant d'un poids de régularisation élevé, nous le faisons décroître régulièrement jusqu'à obtenir un modèle solution lisse qui vérifie les données (voir Figures 3.5 et 3.6). Cette méthodologie nous donne le modèle le plus lisse qui vérifie les données avec la précision voulue (cette méthodologie satisfait le principe de la L-curve, voir Hansen, 2001). Plus on dispose de

données cinématiques et plus le poids de régularisation est grand, plus la solution est stable.

La méthodologie permet donc une inversion stable : sur le plan pratique, le problème apparaît d'autant plus stable que l'on pénalise fort, c'est à dire que l'on sait a priori que le modèle recherché est très régulier, ce qui est bien naturel (voir Renard et Lailly, 1999, pour une étude expérimentale visant à quantifier cette stabilité en fonction de la régularité du modèle recherché et du nombre de données disponibles).

Le critère de la L-curve (détaillé plus loin) permet de détecter le poids de régularisation limite au-dessous duquel aucune amélioration sensible du résidu ne sera obtenue et l'inversion devient instable. Cette méthode est très efficace en tomographie de réflexion (voir Comby, 2004) : la diminution régulière des poids de régularisation permet de trouver des modèles de moins en moins lisses pour lesquels les temps de trajet calculés se rapprochent de plus en plus des temps de trajet observés.

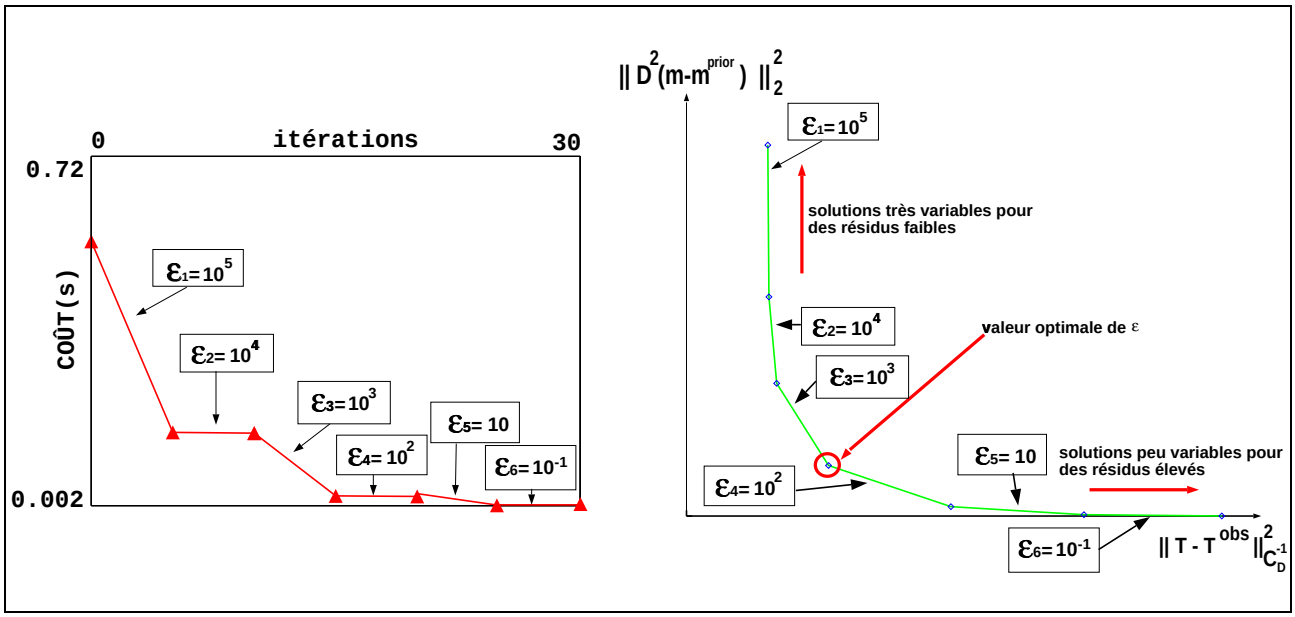

FIG. 3.5 – A gauche : représentation de l'évolution du coût au cours des itération et illustration de la méthodologie décrite dans le paragraphe 3.3.2.1.1. A droite : représentation de la courbe en «L» caractéristique (courbe paramétrée en fonction de ϵ).

3.3.2.1.2 Critère de la L-curve

L' utilisation du critère de la L-curve revient à Miller et Lawson, 1974, et c'est P. C. Hansen, 2001, qui l' a popularisée. C'est une courbe (Figure 3.5) représentant l' évolution de la solution régularisée $||D^2(m - m^{prior})||^2_{L^2}$ en fonction du résidu $||T^{cal}(m) - T^{obs}||^2_{C_d^{-1}}$ (plus précisément la moyenne quadratique des résidus) pour différentes valeurs de ϵ. Les points de cette courbe correspondent à des modèles solutions pour des ϵ différents. Cette courbe tracée en échelle log-log a la forme d'un «L» dans le cas où le problème non régularisé est mal posé, d'où le nom du critère.

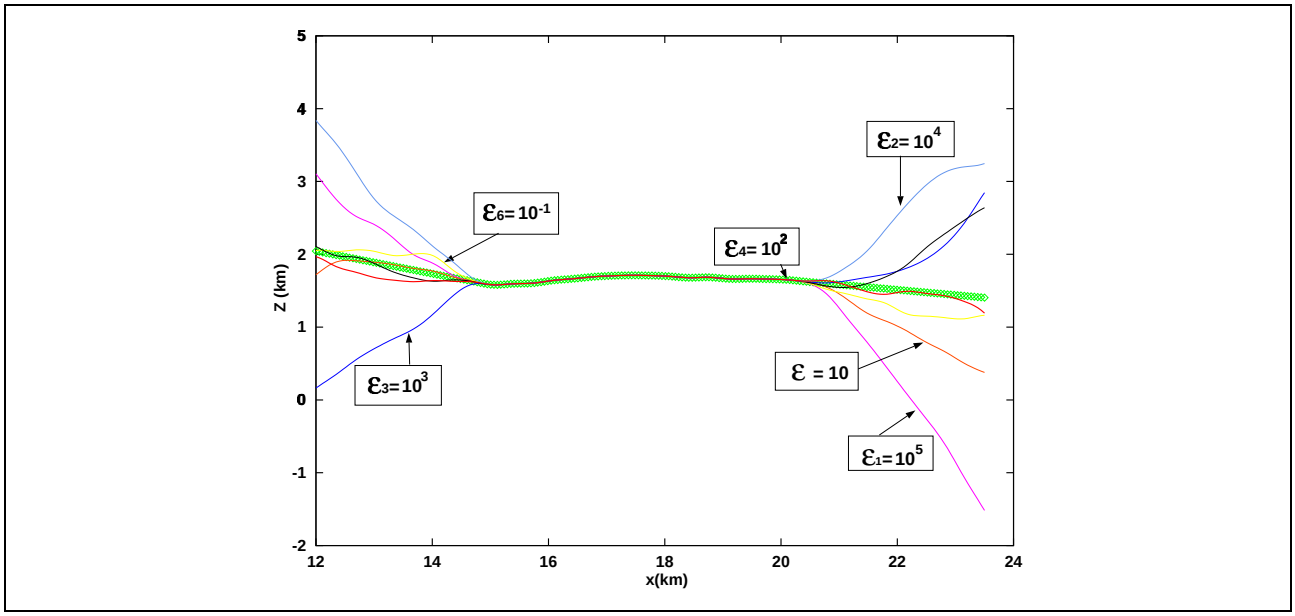

FIG. 3.6 – Illustration des différents modèles obtenus au cours des itérations pour différents poids ϵ. La figure représente une coupe de l'interface d'un modèle le long de l'axe des x (le modèle exact est représenté par le trait vert épais).

D'une manière générale, l'erreur sur le modèle solution commise en régularisant est la somme d'une erreur de perturbation qui est la différence avec la solution obtenue avec les données exactes et celle obtenue avec les données bruitées (elle correspond aux «imperfections» du terme $||T^{cal}(m) - T^{obs}||^2_{C_d^{-1}}$ de la fonction coût régularisée (3.15)) et d'une erreur de régularisation (due au choix de ϵ) qui se produit lors de la minimisation des effets des perturbations des données sur la solution, cette erreur est généralement la source d'une perte d'information pertinente.

Soient $\tilde{m}(\epsilon)$, les solutions du problème de minimisation de $C(m)$ de la fonction coût régularisée (3.15)) :

- La partie horizontale de la courbe en «L» correspond aux solutions où l'erreur de régularisation prédomine (grandes valeurs de ϵ). Dans cette zone, la solution est très filtrée et reste lisse et donc $||D^2(\tilde{m}(\epsilon) - m^{prior})||^2_{L^2}$ varie très peu avec ϵ, alors que le résidu $||T^{cal}(\tilde{m}(\epsilon)) - T^{obs}||^2_{C_d^{-1}}$ est très sensible aux variations de ϵ. Les écarts entre les temps de trajet observés et les temps de trajet calculés ne sont donc pas bien minimisés.

- La partie verticale de la courbe en «L» correspond aux solutions où l'erreur due à la perturbation du terme $||T^{cal}(m) - T^{obs}||^2_{C_d^{-1}}$ prédomine. Dans cette zone, $||D^2(m - m^{prior})||^2_{L^2}$ est très sensible aux variations de ϵ.

Il existe alors une valeur optimale de ϵ qui équilibre l'erreur de régularisation et l'erreur de perturbation. Cette valeur se situe à l'intersection des parties horizontales et verticales, le coin de la courbe en «L».

Toutefois, le critère de la L-curve comporte certains inconvénients. En effet, l'existence du «coin» de la courbe en L n'est pas garantie, d'une part, et d'autre part, la solution dépend de l'échelle choisie car la notion de «coin» est mal définie. Enfin, il est nécessaire de connaître de nombreux points de la courbe afin de pouvoir réaliser une interpolation efficace du coin. En effet, l'obtention d'un point de la courbe est une tâche difficile : par exemple dans le cas de la régularisation de Tikhonov cela nécessite la minimisation pour un ϵ fixé du problème (3.15)). Les travaux de Hanke, 1995, et Vogel, 1996, ont mis en évidence d'autres difficultés liées à l'utilisation de ce critère pour estimer le paramètre de régularisation.

En pratique, le critère de la L-curve consiste tout d'abord à obtenir une première solution de l'inversion en partant d'un fort poids de régularisation. Puis en initialisant l'inversion avec ce modèle, on obtient un nouveau modèle pour un poids de régularisation plus faible. On renouvelle cette opération. Le critère de la L-curve permet de détecter le poids de régularisation limite au-dessous duquel aucune amélioration sensible du résidu ne sera obtenue et l'inversion devient instable.

3.3.2.2 Introduction d'information a priori : l'approche bayesienne

L'approche décrite précédemment, c'est à dire l'introduction de régularisation, souffre de quelques limitations :

- on manque d'arguments pour le choix des poids de régularisation,
- le bruit est pris en compte mais de manière implicite ; la prise en compte explicite des caractéristiques du bruit n'est pas aisée,
- il n'est pas non plus aisé de prendre en compte des informations statistiques sur la solution.

Pour remédier à ces difficultés, il n'y a pas d'autres solutions que de considérer la résolution du problème inverse comme un problème d'estimation probabiliste (voir Annexe G).

Dans ce qui suit nous allons donc résoudre le problème inverse linéarisé en introduisant de l'information a priori et en tentant de donner une signification physique aux poids de régularisation (approche bayesienne). On espère alors améliorer l'inversion par l'intégration d'informations a priori sur les modèles de façon à restreindre l'espace des solutions à des modèles plus réalistes.

Nous allons chercher dans cette approche à minimiser la fonction coût suivante :

$$C(m) = \frac{1}{2}||T^{cal}(m) - T^{obs}||^2_{C_d^{-1}} + \epsilon^2||m - m^{prior}||^2_{\mathbf{C}_m^{-1}}, \tag{3.16}$$

où m^{prior} est le modèle a priori et $\mathbf{C}_m$ la matrice de covariance sur les modèles qui permet de décrire les erreurs sur ce modèle a priori.

Dans la suite nous allons voir comment construire la norme $||\,||_{\mathbf{C}_m^{-1}}$ (Tarantola, 1987, Tonellot, 2000).

Les calculs qui permettent de construire la norme intervenant dans le terme géologique à partir de l'opérateur de covariance exponentiel peuvent être retrouvés dans les travaux de Tarantola, 1987 et de Tonellot, 2000. En annexe D seuls les calculs de la norme en 3D dans le cas anisotrope sont développés.

Nous nous contenterons de donner ici les expressions de la norme intervenant dans le terme géologique dans les cas isotrope et anisotrope, en 2D et en 3D.

Par l'introduction d'information a priori, on cherche à apporter plus de réalisme au formalisme du problème inverse. Pour cela nous devons prendre en compte les corrélations des incertitudes sur le modèle a priori, c'est à dire utiliser un opérateur de covariance adéquate, et donc modifier la norme dans l'espace des modèles.

Le terme de la fonction coût de l'inversion stochastique correspondant à l'information a priori, c'est à dire mesurant l'écart du modèle courant au modèle a priori a pour expression (cf. Tarantola, 1987, et Tarantola et Valette, 1982) :

$$S_{geol}(m) = (m - m^{prior})^T \mathbf{C}_m^{-1} (m - m^{prior}). \tag{3.17}$$

Le modèle a priori m^{prior} intègre l'ensemble des informations quantitatives dont nous disposons a priori sur les modèles. Lors de la construction du modèle a priori, la connaissance du vrai modèle en certains points peut-être plus précise qu'ailleurs. Par exemple, si on dispose de mesures dans certaines zones du modèle, ces informations sont précises et par conséquent, la variance en ces points devrait être plus faible qu'ailleurs. De la même façon, on peut envisager que les corrélations des incertitudes sur le modèle a priori soient plus faibles en certains points du milieu, ou que ces corrélations soient plus importantes dans certaines directions. Ces considérations vont nous amener à introduire des variations des variances et des longueurs de corrélations dans la fonction de covariance.

La fonction de covariance $\mathcal{C}_m$ est une application de $\Omega \times \Omega$ dans $\mathbb{R}$ (où $\Omega = \Omega_x \times \Omega_y$ en 2D et $\Omega = \Omega_x \times \Omega_y \times \Omega_z$ en 3D est le domaine de définition) qui décrit les incertitudes sur le modèle a priori en tout point de Ω et la dépendance spatiale de ces incertitudes en deux points x_1 et x_2 de Ω selon la définition :

$$\mathcal{C}_m(x_1, x_2) = E[(m(x_1) - m^{prior}(x_1))(m(x_2) - m^{prior}(x_2))], \tag{3.18}$$

où E est l'espérance mathématique.

Quand $x_1 = x_2$, $\mathcal{C}_m(x_1, x_2)$ est égale à la variance de m en x_1, i.e., $\mathbf{C}_m$ quantifie la confiance que l'on a en m en x_1.

Si $x_1 \neq x_2$ alors $\mathcal{C}_m(x_1, x_2)$ mesure la corrélation des incertitudes sur le modèle a priori en x_1 et en x_2.

Pour modéliser l'opérateur de covariance $\mathcal{C}_m$, nous faisons deux hypothèses quant aux comportements des erreurs. Les incertitudes en deux points proches sont fortement corrélées, et cette corrélation diminue si la distance entre ces deux points augmente.

Pour modéliser ce genre de comportement des erreurs, dans le cas isotrope, on peut utiliser la fonction suivante (fonction de covariance exponentielle) :

$$\begin{aligned} \mathcal{C}_m : \Omega \times \Omega &\longrightarrow \mathbb{R} \\ (x_1, x_2) &\mapsto \mathcal{C}_m(x_1, x_2) = \sigma^2 \exp\left(-\frac{\| x_1 - x_2 \|}{\lambda}\right), \end{aligned} \tag{3.19}$$

où σ^2 est la variance et λ la longueur de corrélation, c'est à dire la distance au-delà de laquelle les corrélations deviennent négligeables (σ et λ sont des constantes réelles non nulles).

Les Figures 3.7 et 3.8 illustrent, respectivement, les effets d'une variation de λ pour un σ fixé et d'une variation de σ pour un λ fixé.

L'opérateur de covariance décrit donc les incertitudes sur le modèle a priori et va nous permettre de construire via son inverse la norme intervenant dans le terme géologique.

A noter que la formulation (3.19) où σ et λ sont fixes est trop rigide. C'est pourquoi on introduit le cas qu'il convient d'appeler anisotrope où nous construisons une nouvelle fonction de covariance dérivée de la fonction de covariance exponentielle. Son expression est :

$$\mathcal{C}_m(x_1, x_2) = \sigma^2 \exp\left(- \parallel \Lambda(x_1) - \Lambda(x_2) \parallel\right), \tag{3.20}$$

où σ est la variance et $\Lambda(x_1)$ et $\Lambda(x_2)$ sont les longueurs de corrélations selon les deux directions principales d'anisotropie.

Nous voulons donc inverser l'opérateur de covariance exponentiel $\mathbf{C}_m$ dont le noyau est la fonction $\mathcal{C}_m(\ ,\ )$ définie par (3.19) ou par (3.20) respectivement dans le cas isotrope et anisotrope. Comme nous l'avons précisé au début du paragraphe, ces calculs peuvent être trouvés dans Tarantola, 1987, et Tonellot, 2000, et en Annexe D.

Donc finalement, les expressions de la norme associée à l'inverse de l'opérateur de covariance exponentiel selon la dimension de Ω (en 2D, $\Omega = \Omega_x \times \Omega_y$ et en 3D, $\Omega = \Omega_x \times \Omega_y \times \Omega_z$) sont respectivement :

- <u>Dans les cas isotrope</u> :
  - dans le cas 2D :

$$\parallel m_1 \parallel^2_{\mathbf{C}_m^{-1}} = \frac{1}{2\pi\lambda^2\sigma^2}\int_\Omega m_1^2(x)dx + \frac{1}{\pi\sigma^2}\int_\Omega \parallel \text{grad } m_1 \parallel^2 dx + \frac{\lambda^2}{2\pi\sigma^2}\int_\Omega (\Delta m_1(x))^2 dx, \tag{3.21}$$

  - dans le cas 3D :

$$\parallel m_1 \parallel^2_{\mathbf{C}_m^{-1}} = \frac{1}{8\sigma^2\lambda^3\pi}\int_\Omega m_1^2(x)dx + \frac{1}{4\lambda\sigma^2\pi}\int_\Omega \parallel \text{grad } m_1 \parallel^2 dx + \frac{\lambda}{8\sigma^2\pi}\int_\Omega (\Delta m_1(x))^2 dx. \tag{3.22}$$

- <u>Dans les cas anisotrope</u> :
  - dans le cas 2D :

$$\begin{aligned} \parallel m_1 \parallel^2_{\mathbf{C}_m^{-1}} |_{2D} &= \frac{1}{2\pi\sigma^2}\int_\Omega \frac{1}{\lambda_x\lambda_y} m_1(u)^2 du \\ &+ \frac{1}{\sigma^2\pi}\int_\Omega \lambda_x\lambda_y \left(\frac{1}{\lambda_y^2}\left(\frac{\partial\ m_1}{\partial x}\right)^2 + \frac{1}{\lambda_x^2}\left(\frac{\partial m_1}{\partial y}\right)^2\right) du \\ &+ \frac{1}{2\pi\sigma^2}\int_\Omega (\lambda_x\lambda_y)^3 \left(\frac{1}{\lambda_y^2}\frac{\partial^2 m_1}{\partial x^2} + \frac{1}{\lambda_x^2}\frac{\partial^2 m_1}{\partial y^2}\right)^2 du \end{aligned} \tag{3.23}$$

  avec $\Omega = \Omega_x \times \Omega_y$ et λ_x et λ_y les longueurs de corrélations selon les deux directions principales d'anisotropie.

- dans le cas 3D :

$$
\begin{aligned}
\| m_1 \|^2_{\mathbf{C}_m^{-1}} |_{3D} &= \frac{1}{8\pi\sigma^2} \int_\Omega \frac{1}{\lambda_x \lambda_y \lambda_z} m_1(u)^2 du \\
&+ \frac{1}{4\pi\sigma^2} \int_\Omega \lambda_x \lambda_y \lambda_z \left(\frac{1}{(\lambda_y \lambda_z)^2} \left(\frac{\partial m_1}{\partial x} \right)^2 + \frac{1}{(\lambda_x \lambda_z)^2} \left(\frac{\partial m_1}{\partial y} \right)^2 + \frac{1}{(\lambda_x \lambda_y)^2} \left(\frac{\partial m_1}{\partial z} \right)^2 \right) du \\
&+ \frac{1}{8\pi\sigma^2} \int_\Omega (\lambda_x \lambda_y \lambda_z)^3 \left(\frac{1}{(\lambda_y \lambda_z)^2} \frac{\partial^2 m_1}{\partial x^2} + \frac{1}{(\lambda_x \lambda_z)^2} \frac{\partial^2 m_1}{\partial y^2} + \frac{1}{(\lambda_x \lambda_y)^2} \frac{\partial^2 m_1}{\partial z} \right)^2 du \qquad (3.24)
\end{aligned}
$$

avec $\Omega = \Omega_x \times \Omega_y \times \Omega_z$ et λ_x, λ_y et λ_z les longueurs de corrélations selon les trois directions principales d'anisotropie.

On reconnaît dans les deux cas la norme H^2 aux coefficients près.

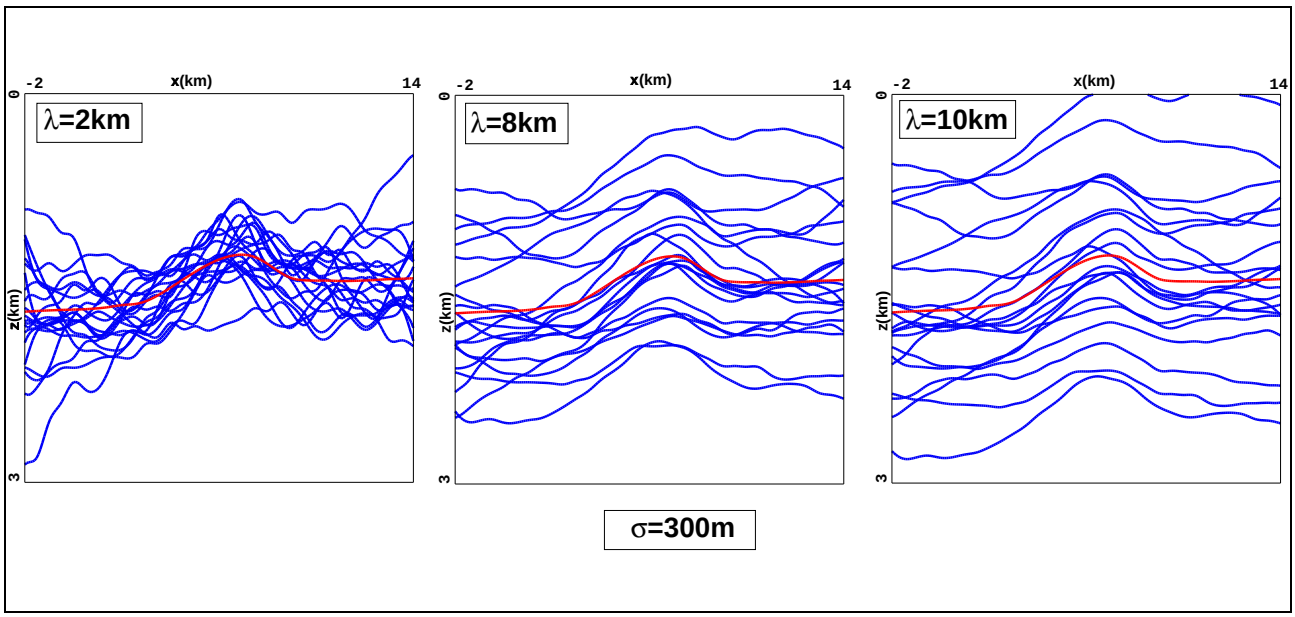

FIG. 3.7 – Illustration de l'effet de λ sur les modèles. Coupes suivant l'axe des x sur les modèles de vitesse 3D simulés a priori (représentation de la profondeur de la première interface du modèle solution de la Figure B.3) dans le cas isotrope. Les différentes longueurs de corrélation λ testées sont : $\lambda = 2km$, $\lambda = 8km$ et $\lambda = 10km$, $\lambda = 2km$. L'écart type est $\sigma = 300m$.

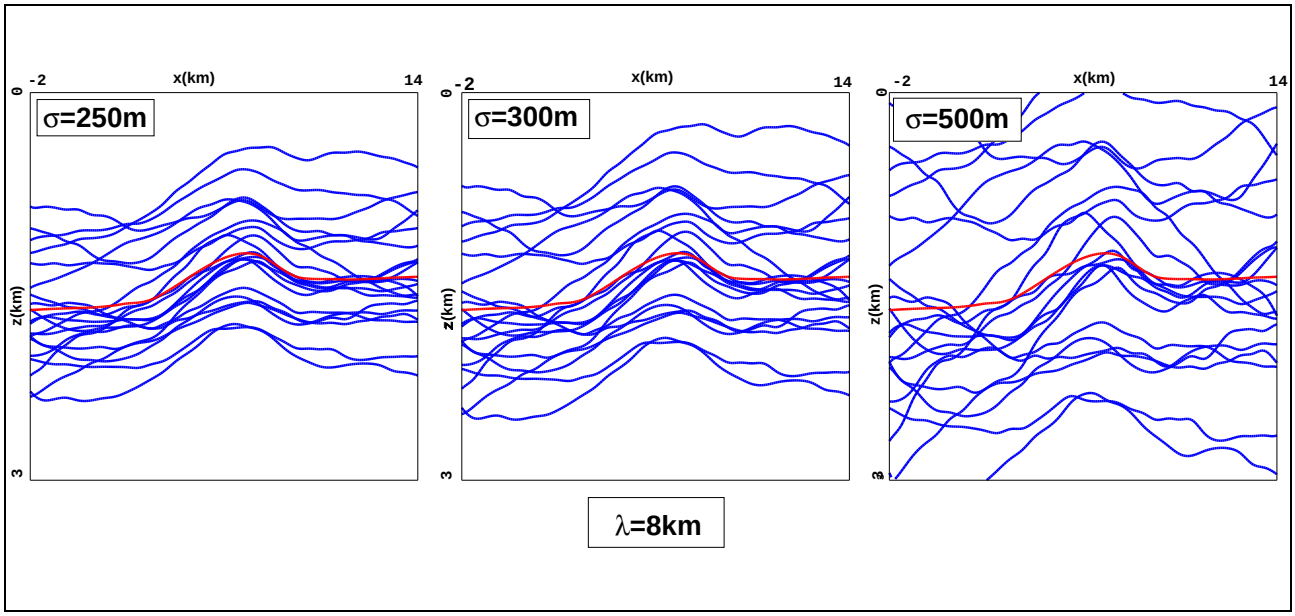

FIG. 3.8 – Illustration de l'effet de σ sur les modèles. Coupes suivant l'axe des x sur les modèles de vitesse 3D simulés a priori (représentation de la profondeur de la première interface du modèle solution de la Figure B.3) dans le cas isotrope. Les différents écarts types σ testés sont : $\sigma = 250m$, $\sigma = 300m$ et $\sigma = 500m$. La longueur de corrélation est $\lambda = 8km$.

3.3.2.3 Résolution du problème inverse

3.3.2.3.1 Le problème inverse continu

Pour résoudre notre problème de tomographie de réflexion, nous avons vu dans la section 3.1 que nous nous intéressons à des modèles particuliers.

Les données de notre problème sont les temps d'arrivée associés aux ondes réfléchies obtenus par pointé des évènements sur les sections sismiques (le nombre de données peut être de l'ordre de 50000-100000 couples (T^{obs}, T^{cal}) pour une interface).

Le problème inverse est formulé en termes de minimisation d'une fonction coût qui mesure l'écart entre les données observées et les données synthétiques associées à un modèle m par le biais d'une norme L^2 définie sur l'espace des données. On parle de formulation moindres carrés du problème inverse. Bien que ce choix puisse être critiquable, car il est intimement lié à une hypothèse d'incertitudes gaussiennes sur les données, il n'en reste pas moins qu'il s'agit d'un critère simple à mettre en oeuvre et pour lequel on dispose d'une large palette de méthodes de traitements. Si l'hypothèse d'incertitudes gaussiennes est gênante, car elle ne correspond pas forcément à la distribution réelle des erreurs, on pourra envisager d'autres critères tel que la minimisation d'une norme L^1, L^∞ (minimax), ou éventuellement les normes L^p (Tarantola, 1987, Crase et al., 1990, Renard, 2002).

On notera au préalable $\mathcal{M}$ l'espace des modèles.

Soit l'ensemble des temps de trajet $(T_k^{obs})_{k=1,2,\ldots,N^{obs}}$, où N^{obs} est le nombre de données observées.

le problème inverse consiste à :

Chercher le modèle m appartenant à $\mathcal{M}$ qui minimise :

$$C(m) = ||T(m) - T^{obs}||^2_{C_d^{-1}}, \tag{3.25}$$

avec T l'opérateur de modélisation défini par (3.12) qui donne, pour un réflecteur, les temps de trajet associés à chaque couple (source, récepteur)) et $|| - ||_{C_d^{-1}}$ la norme dans l'espace des données (norme qui prend en compte les incertitudes dans le pointé des temps.

Le problème inverse tel qu'il vient d'être formulé est mal-posé. Delprat-Jannaud et Lailly, 1993, ont défini une formulation du problème inverse de tomographie régularisée qui consiste à :

Chercher le modèle m appartenant à $\mathcal{M}$ minimisant la fonctionnelle :

$$C(m) = ||T(m) - T^{obs}||^2_{C_d^{-1}} + \epsilon^2 ||D^2(m - m^{prior})||^2_{L^2}, \tag{3.26}$$

où

$$||D^2(m)||^2_{\mathcal{M}} = \sum_{i=1}^{N_V} \int_{\Omega_x \times \Omega_y \times \Omega_z} [D^2(v^i(x,y,z))]^2 \, dxdydz + \sum_{j=1}^{N_I} \int_{\Omega_x \times \Omega_y} [D^2(Z^j(x,y))]^2 \, dxdy, \tag{3.27}$$

et D^2 est l'opérateur de dérivation jusqu'à l'ordre 2, par exemple dans le cas à trois variables :

$$\begin{aligned} [D^2(f(x,y,z))]^2 &= \left[\frac{\partial^2 f}{\partial x^2}(x,y,z)\right]^2 + \left[\frac{\partial^2 f}{\partial y^2}(x,y,z)\right]^2 + \left[\frac{\partial^2 f}{\partial z^2}(x,y,z)\right]^2 \\ &+ \left[\frac{\partial^2 f}{\partial x \partial y}(x,y,z)\right]^2 + \left[\frac{\partial^2 f}{\partial x \partial z}(x,y,z)\right]^2 + \left[\frac{\partial^2 f}{\partial y \partial z}(x,y,z)\right]^2 . \end{aligned} \tag{3.28}$$

où m est un modèle de Terre défini par N_I interfaces et N_V vitesses de couches et m^{prior} le modèle a priori.

$||.||_{C_d^{-1}}$ est la norme dans l'espace des données, C_d étant l'opérateur de covariance a priori sur les données. Ses éléments diagonaux représentent les incertitudes sur les données (erreurs sur les données elles-mêmes et erreurs de modélisation), et ses éléments hors-diagonaux représentent les corrélations entre ces incertitudes (Tarantola, 1987).

$||.||_{\mathcal{M}}$ est la norme dans l'espace des modèles $\mathcal{M}$ (ici la norme L^2).

ϵ^2 est un poids de régularisation qui permet de régler l'effet de la régularisation (On reviendra plus tard sur le choix de la valeur de ce paramètre). On a choisi de régulariser le problème de la tomographie par les dérivées secondes des variations de vitesse et des profondeurs d'interface, ces fonctions étant régulières (Delprat-Jannaud et Lailly, 1993, Bube, 1995).

La fonction coût (3.26) n'étant pas quadratique, l'opérateur de modélisation direct $T(m)$ étant non linéaire, nous résolvons ce problème de minimisation non linéaire par un algorithme de Gauss-Newton, qui consiste en des linéarisations successives de l'opérateur direct.

Les principaux arguments qui conduisent à utiliser la méthode de Gauss-Newton pour résoudre (3.26) sont les suivants :

- L'approximation du Hessien de C par $J(m_n)^T J(m_n)$ est intéressante car elle utilise uniquement le calcul de la jacobienne $J(m_n)$, le coût de calcul des dérivées secondes des résidus pouvant être important.

- Dans de nombreuses applications le terme $J(m_n)^T J(m_n)$ approchant le Hessien de C est dominant par rapport aux termes regroupant les dérivées secondes des résidus. Cela se produit plus particulièrement dans le cas d'applications où :

  * les résidus sont petits ($r_i \approx 0$),

  * les résidus sont quasi-linéaires ($\| \nabla^2 r_i \| \approx 0$).

3.3.2.3.2 La méthode de Gauss-Newton

Soit le problème non linéaire général suivant :

$$\begin{cases} \text{Nous cherchons } m \in \mathcal{M} \text{ minimisant} \\ \qquad C_F(m) = \|F(m)\|_D^2 \,. \end{cases} \tag{3.29}$$

où F est une application de l'espace des modèles $\mathcal{M}$ dans un espace de données $\mathcal{D}$ qu'on supposera muni d'un produit scalaire $<,>_D$. $\mathcal{D}$ est un espace de Hilbert dont la norme associée est notée $\| \ \|_D$. On suppose que F est différentiable (notons que dans notre cas $F = T$ l'est puisqu'on a fait le choix de fonctions B-splines).

On effectue alors un développement de Taylor d'ordre 1 de la fonction F,

$$F(m_n + \delta m) = F(m_n) + J(m_n)\delta m + ||\delta m||_{\mathcal{M}} o(||\delta m||_{\mathcal{M}}), \tag{3.30}$$

où

$$\lim_{||\delta m||_{\mathcal{M}} \to 0} o(||\delta m||_{\mathcal{M}}) = 0,$$

et $J(m_n) = \frac{\partial F}{\partial m}(m_n) = d_{m_n} F$ est la jacobienne de F en m_n.

On définit le problème inverse linéaire comme la minimisation de la fonctionnelle :

$$\begin{aligned} \tilde{C}_F(\delta m) &= ||J(m_n)\delta m + F(m_n)||_D^2 \\ &= ||F(m_n)||_D^2 + 2\,\langle J(m_n)\delta m, F(m_n)\rangle_D + \langle J(m_n)\delta m, J(m_n)\delta m\rangle_D \\ &= ||F(m_n)||_D^2 + \langle G_n, \delta m\rangle_{\mathcal{M}} + \langle H_n \delta m, \delta m\rangle_{\mathcal{M}} \end{aligned} \tag{3.31}$$

où $G_n = 2J^*(m_n)F(m_n)$ est le gradient de la fonctionnelle C_F en m_n, $H_n = J^*(m_n)J(m_n)$ est la hessienne de la fonctionnelle C_F en m_n, $J^*(m_n)$ est l'adjoint de $J(m_n)$.

Les étapes de la méthode de Gauss-Newton sont les suivantes (voir, par exemple, Nocedal et Wright, 1999, ou Kelley, 1999) :

(1) Choix d'un modèle initial $m^0 = m_0$ $(n = 0)$

(2) $n^{\text{ième}}$ itération : résolution du problème linéarisé autour de m_n

(i) Calcul de $F(m_n)$ et de $J(m_n)$

(ii) Détermination de δm_n solution du problème linéarisé :
Détermination de m_{n+1} qui est alors le modèle solution.

$$m_{n+1} = m_n + \delta m_n$$

(iii) Recherche linéaire éventuelle dans la direction δm_n :

$$m_{n+1} = m_n + \epsilon \delta m_n, 0 < \epsilon \leq 1$$

On applique cette recherche linéaire quand le coût non linéaire ne diminue pas suffisamment (règles d'Armijo). Cela signifie que le pas δm_n est trop grand (la linéarisation de la fonctionnelle n'est plus valide).

3.3.2.3.3 Le problème inverse linéarisé

La résolution du problème inverse linéarisé constitue une étape de Gauss-Newton.

La fonction coût quadratique résultante que l'on minimise à chaque itération de Gauss-Newton est :

$$\tilde{C}(\delta m) = \frac{1}{2}||J(m_n)\delta m - \delta T^{obs}||^2_{C_d^{-1}} + \frac{\epsilon^2}{2}||D^2(m_n + \delta m - m^{prior})||^2_{L^2}. \tag{3.32}$$

où m_n est le modèle courant, δm est la perturbation de modèle, $J(m_n) = \frac{\partial T^{cal}}{\partial m}(m_n)$ est la matrice Jacobienne évaluée en m_n, et $\delta T^{obs} = T^{obs} - T^{cal}(m_n)$.

Résoudre ce problème de minimisation est équivalent à résoudre le système linéaire suivant

$$(J(m_n)^T C_d^{-1} J(m_n) + \mathcal{Q}_\epsilon)\Delta m = J(m_n)^T C_d^{-1} \delta T^{obs} - \mathcal{Q}_\epsilon(m_n - m^{prior}), \tag{3.33}$$

où $\mathcal{Q}_\epsilon$ est la matrice constituée des termes de régularisation.

Ce problème inverse linéarisé $\tilde{C}$ est bien posé comme l'ont montré Delprat-Jannaud et Lailly, 1993 : si le nombre de données est suffisant, ce problème a une unique solution qui dépend continûment des données (voir Annexes E et F).

En général, on résout le problème inverse linéarisé par un algorithme de gradient conjugué (Ciarlet, 1988).

Dans notre cas, l'opérateur de covariance sur les données, C_d est diagonal : les composantes du vecteur résidu, c'est à dire le vecteur dont les éléments sont les différences entre les temps de trajet observés et les temps de trajet calculés par le problème direct (tracé de rayons), sont considérées comme des informations spatialement indépendantes. Donc, la fonction coût non linéaire discrétisée s'écrit :

$$\tilde{C}'(m) = \frac{1}{2}\sum_{i=1}^{N^{obs}} \frac{(T_i(m) - T_i^{obs})^2}{\sigma_i^2} + \frac{\epsilon^2}{2} \parallel D^2(m - m^{prior}) \parallel^2, \tag{3.34}$$

où σ_i^2 est l'incertitude associée à la donnée T_i^{obs}.

3.3.3 Conclusions

Il existe donc deux manières de concevoir le problème inverse en relation avec l'objectif que l'on se fixe. Lorsque le problème inverse a un objectif décisionnel, le concept de meilleur modèle à un sens et le problème peut se poser comme un problème d'optimisation (3.13). En revanche, si le but poursuivi est de représenter la connaissance que l'on a sur un système physique à travers des paramètres sensés caractériser ce système, alors l'approche probabiliste est mieux adaptée (3.14). Cela n'est pas en contradiction avec le fait que l'on puisse finalement décider au vu de la structure des incertitudes de résoudre le problème en utilisant des méthodes d'optimisation.

Les incertitudes liées à la connaissance dont on dispose sur le système physique ou encore les incertitudes liées à l'expérimentation physique elle-même doivent être prises en compte dans la manière de poser le problème. Par conséquent, l'approche probabiliste du problème inverse géophysique semble être incontournable, au moins pour poser conceptuellement le problème. De plus, cette approche permet d'aborder naturellement l'étude et l'interprétation des résultats d'une inversion (incertitudes, fiabilité, résolution) puisque l'on a accès à l'information a posteriori sur l'espace des modèles. Dans le cas où on utilise des méthodes d'optimisation pour résoudre un problème inverse, il est nécessaire de pouvoir décrire ou caractériser les incertitudes associées à la solution obtenue afin d'être capable d'interpréter correctement les résultats de l'inversion, c'est à dire fournir les barres d'erreur associées à la solution.

Chapitre 4

QUANTIFICATION DES INCERTITUDES SUR LE MODÈLE SOLUTION

Ce chapitre est consacré à l'estimation des incertitudes sur le modèle solution de la tomographie de réflexion. Nous allons faire le point sur les méthodes classiques d'analyse a posteriori et proposer deux approches adaptées à notre problème pour quantifier les incertitudes sur le modèle solution de la tomographie de réflexion.

Classiquement, l'analyse de la solution du problème inverse consiste à comparer les écarts entre les temps de trajet observés et les temps de trajet calculés en calculant la valeur du résidu moyen et celle du résidu maximum, et en étudiant la répartition spatiale de ces résidus (zones locales avec des résidus élevés et tendances). Cette analyse est souvent insuffisante : ce modèle est un des modèles admissibles que nous pouvons obtenir à partir de nos données. Une analyse d'incertitude est donc cruciale pour quantifier la confiance que l'on accorde au modèle solution déterminé.

4.1 Indétermination du problème inverse

Ce problème inverse de tomographie de réflexion est sous-déterminé : nous avons à faire face à une ambiguité entre la vitesse et la profondeur des réflecteurs qui ne peut pas être résolue par l'information des données temps de trajet (Bishop et al., 1985, Farra et Madariaga, 1988, Stork et Clayton, 1991, Bickel, 1990, Bube, 1995, Sinoquet, 1995). En effet, la profondeur des interfaces et la vitesse ne peuvent pas être déterminées de manière unique sans l'introduction d'information a priori sur la position des réflecteurs, c'est à dire qu'il est difficile de séparer l'information fournie par les temps de trajet en composantes de vitesse et profondeur d'interface. Par conséquent, pour un jeu de données fixé, plusieurs modèles pourraient correspondre au modèle le plus probable qui vérifie les données avec la précision espérée.

4.1.1 Indétermination de la vitesse

En tomographie, on peut trouver des situations telles qu'il y ait plusieurs ou une infinité de modèles expliquant les données.

Prenons par exemple, le cas d'un réflecteur plan, z, et d'une vitesse constante, v. Pour inverser cette interface, si nous ne possédons que des données zéro-offset, il y aura alors une infinité de couples (v, z) définissant notre modèle et satisfaisant les temps de trajet ($t = \frac{z}{v}$ peut prendre toutes les valeurs à z ou v fixes). Plus généralement, plus nous aurons de données, meilleure sera la détermination de notre modèle, même complexe. Notons, qu'une acquisition sismique classique enregistre jusqu'à plusieurs centaines d'offsets différents. Illustrons cette indétermination de la vitesse liée à l'offset sur un exemple synthétique multi-couches.

4.1.1.1 Présentation du modèle

Dans les expériences numériques qui vont suivre, on réalise l'inversion des données associées au modèle présenté sur la Figure 4.1. Ce modèle est un modèle constitué de plusieurs couches. Les interfaces sont planes, et les vitesses de couches constantes. Les données (temps observés) sont calculées par tracé de rayons sur le modèle exact (z, v). Seules la vitesse et la position de la dernière couche (Lias, voir Figure 4.3) sont inversées. L'interface considérée (Lias) est définie par 4×4 paramètres. La vitesse est décrite par une fonction B-spline $v(x, y)$ de 4×4 paramètres.

Le choix de ce modèle provient du fait que, la couche précédant le Lias, la craie, a une vitesse plus rapide. Ainsi, l'offset maximum apparent au niveau de la réflexion sur la base du Lias sera plus faible, ceci rendant forcément l'indétermination plus grande (le cône d'illumination de la couche étant réduit) (Figure 4.2). D'autre part, le fait que la couche du Lias soit plus étroite entraîne aussi une moins bonne résolution. En effet, une forte perturbation de vitesse dans cette couche a très peu d'influence sur le temps de trajet total (le temps de trajet étant l'intégrale de l'inverse de la vitesse le long des rayons).

Le dispositif est composé d'un seul CDP en x, éclairé par des couples source-récepteur disposés de part et d'autre, tous les 12.50 mètres et de trois CDP en y (Figure 4.2). L'offset maximal est de 4000 mètres, ce qui correspond à 161 offsets.

4.1.1.2 Inversion

Seules la vitesse et la position de la dernière couche (Lias, voir Figure 4.3) sont inversées. On a de forts poids de régularisation et on cherche à obtenir un modèle à vitesse constante et interface horizontale plane.

Lors de l'inversion des temps de trajet, le RMS ($RMS = \sqrt{\frac{1}{NT}\sum_{i=1}^{NT}(T_i^{cal} - T_i^{obs})^2}$) et le maximum des résidus donnent un critère de validité du modèle qui permet de savoir si on a obtenu un «bon» modèle solution. Comme on l'a déjà vu précédemment, lors de l'inversion des données, plusieurs modèles peuvent expliquer les temps de trajet et donc satisfaire à nos critères de validité. Le modèle obtenu (Figure 4.4) vérifiant les données est le modèle solution que nous retiendrons. Dans ce qui va suivre, nous allons tenter de quantifier l'erreur que l'on peut faire sur ce modèle à critère fixé, plus précisément à RMS fixé ($1ms$).

4.1.1.3 Exploration de l'ensemble des modèles

L'exploration de l'espace des modèles consiste à faire varier simultanément profondeur et vitesse, autour des valeurs exactes, et à effectuer sur cette série de modèle des tracés de rayons, ceci simulant

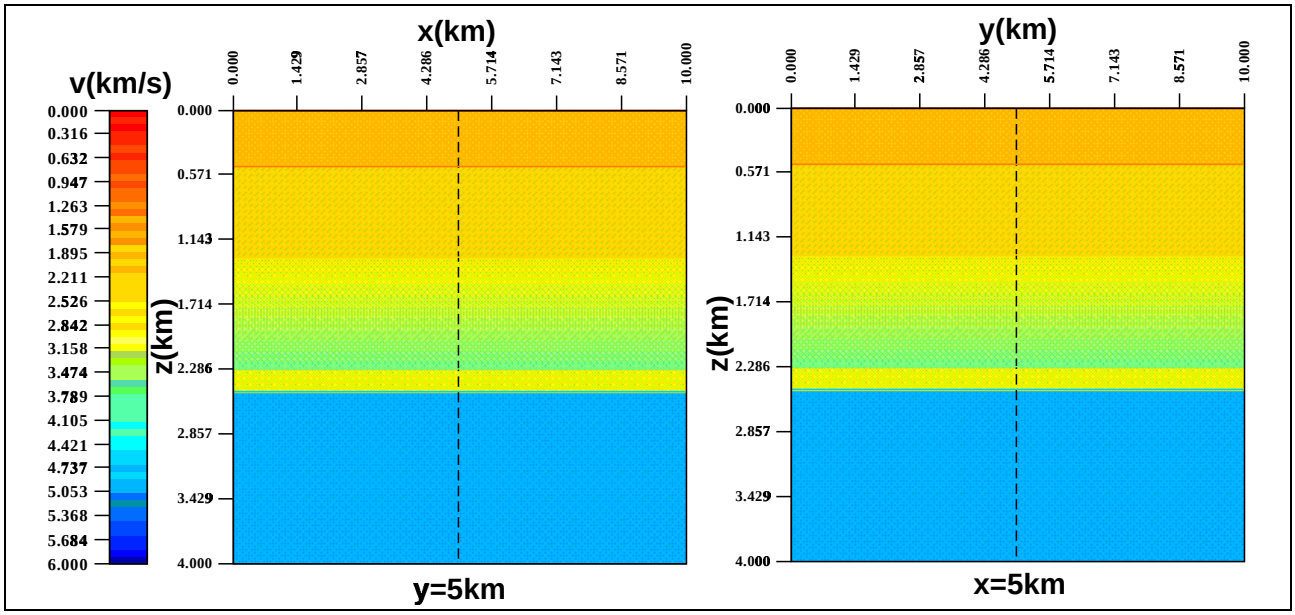

FIG. 4.1 – Coupes verticales du modèle de vitesse exact (Gauche : le long de l'axe des x. Droite : le long de l'axe des y.).

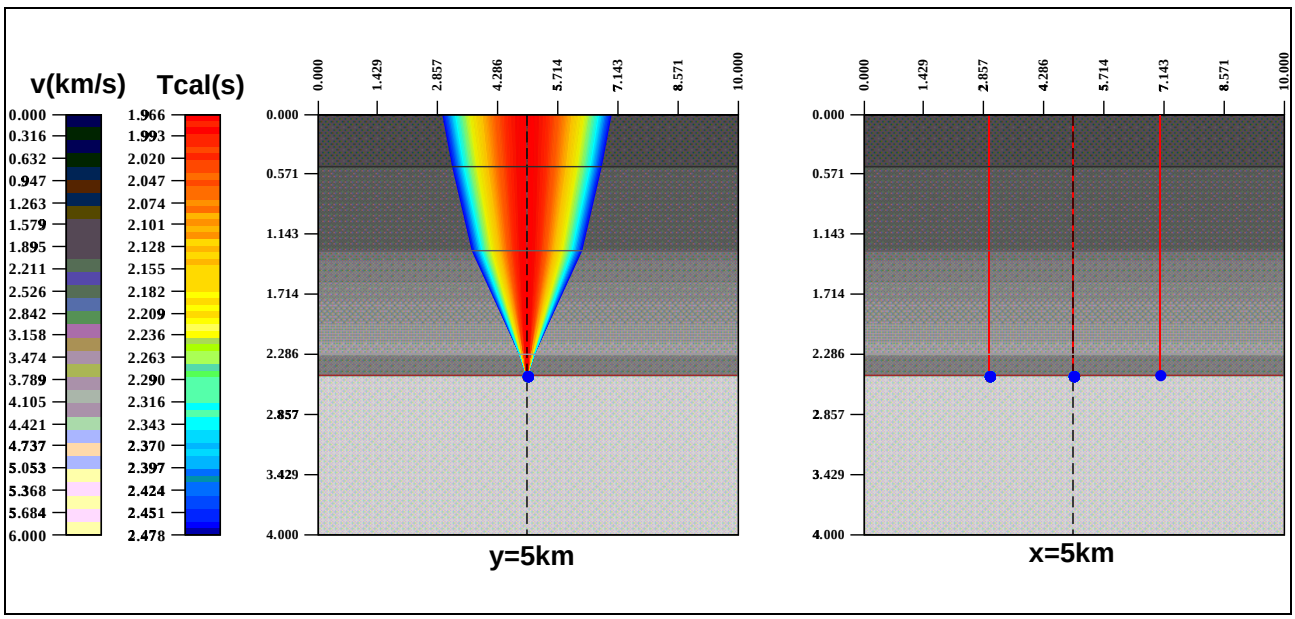

FIG. 4.2 – Système d'acquisition.

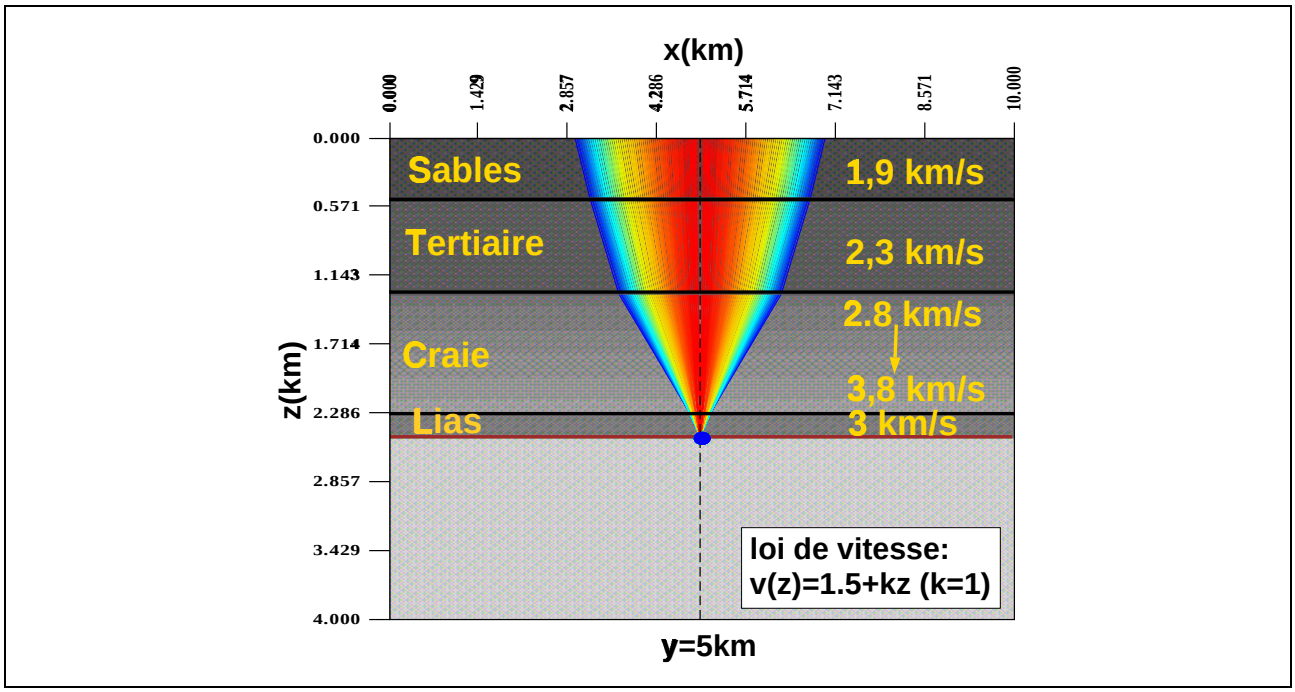

FIG. 4.3 – Représentation des différentes couches du modèle étudié.

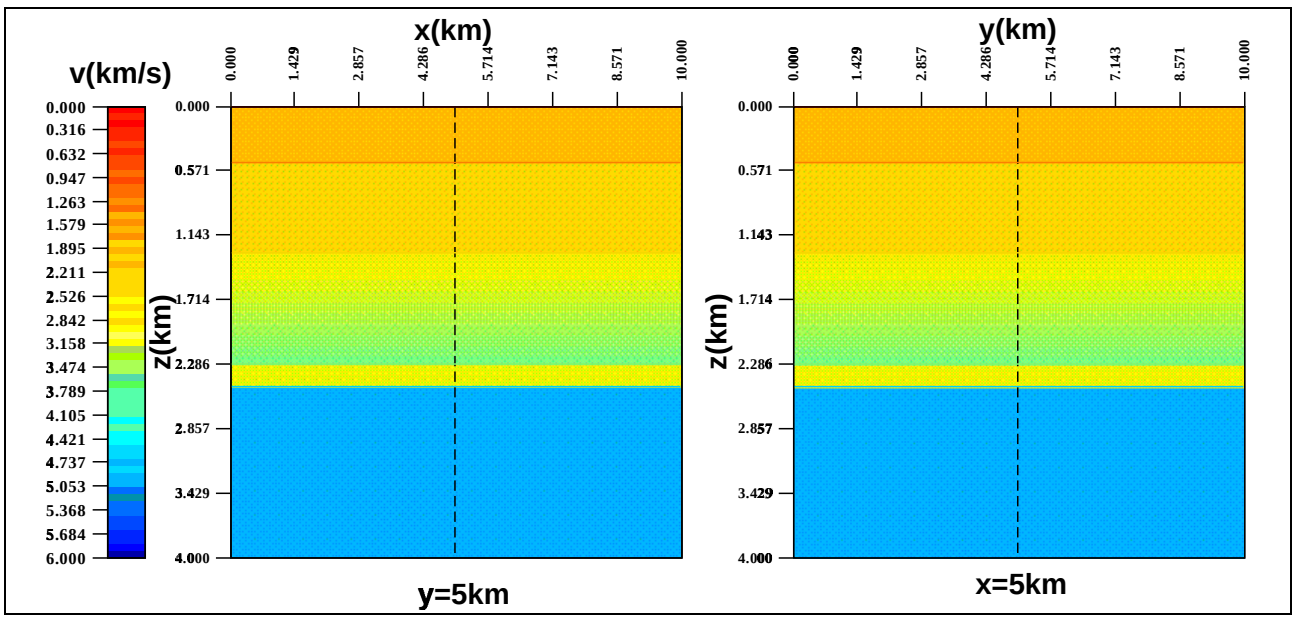

FIG. 4.4 – Coupes verticales du modèle de vitesse solution (Gauche : le long de l'axe des x. Droite : le long de l'axe des y.). RMS moyen des résidus temps de trajet $= 7,95E-02\mu s$ et Maximum des résidus $= 1\mu s$.

les temps calculés. Nous calculons alors le RMS et le maximum des résidus, en fonction du nombre d'offsets utilisés, et observons les courbes iso-RMS et iso-résidus maximum. Une iso-RMS (respectivement iso-résidus maximum), permet de voir l'ensemble des modèles (dans l'espace profondeur, vitesse, nombre d'offsets), dont le RMS (respectivement maximum des résidus) est inférieur à une certaine valeur, choisie par le tomographe.

La Figure 4.6 représente respectivement les cartes en RMS et maximum des résidus pour les offsetmax $4000m$, $3000m$, $2000m$ et $1000m$.

Pour donner un ordre de grandeur sur l'indétermination, avec une gamme d'offsets comprises entre 0 et 4000 mètres (première carte en haut à gauche de la Figure 4.6) un modèle avec un RMS inférieur à une milli-seconde peut se trouver à $105m/s$ de la vraie vitesse et à 10 mètres de la position exacte de l'interface.

De plus, l'incertitude autour du modèle solution (Figure 4.4) est liée au dispositif. En effet, lorsque nous réduisons l'offset de l'acquisition (Figure 4.2), l'incertitude sur la vitesse et la position augmentent, en constatant toutefois que l'incertitude sur la vitesse augmente plus rapidement que celle sur la position (Figure 4.6). Ceci est probablement dû à l'ensemble des couches recouvrant le Lias.

D'autre part, lorsqu'on prend en considération la pente (Figure 4.5), on observe un coefficient de corrélation de l'ordre de 0.1 par exemple pour la gamme d'offsets entre 0 et 4000 mètres. Ce coefficient correspond au facteur qui lie les paramètres de vitesses aux paramètres d'interfaces. Il nous indique de combien augmenter l'interface pour une certaine augmentation de la vitesse.

Pour conclure, les incertitudes sur les données de la tomographie liées à l'identification de la signature associée à un évènement, tâche souvent délicate qui incombe à l'interprétateur (section 3.3.1.1 du chapitre 3), ainsi que l'indétermination du problème inverse entraînent des incertitudes sur le modèle solution de la tomographie de réflexion. Nous allons voir dans ce qui suit comment estimer ces incertitudes.

4.2 Comment estimer les incertitudes : les méthodes classiques

Nous allons ici présenter deux approches classiques pour accéder aux incertitudes sur le modèle solution de la tomographie de réflexion. Elles consistent à analyser la matrice de covariance a posteriori au voisinage du modèle solution. Mais auparavant, nous allons préciser le cadre de notre analyse d'incertitude a posteriori.

Une approche classique pour quantifier les incertitudes consiste à analyser la matrice Hessienne (ou son inverse : la matrice de covariance a posteriori) associée au problème linéarisé (3.32) autour de la solution m_∞ (approche bayesienne, voir Annexe G). Cette approche est valide dans le voisinage du modèle solution (cadre linéarisé), la taille de ce voisinage dépendant de la non-linéarité de l'opérateur direct.

La forme bi-linéaire associée à la matrice Hessienne mesure l'influence d'une perturbation de modèle δm autour du modèle solution m_∞ sur la fonction coût quadratique :

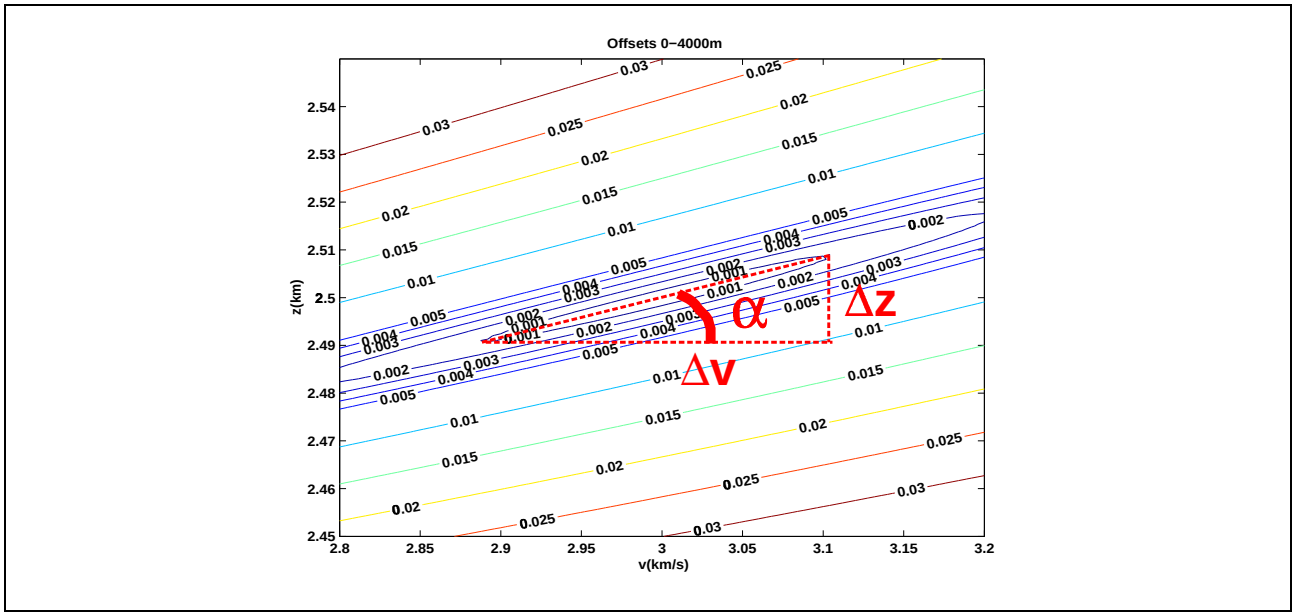

FIG. 4.5 – Calcul du coefficient de corrélation pour un RMS de $1ms$ pour une gamme d'offsets entre 0 et 4000 mètres.

$$
\begin{aligned}
\tilde{C}(\delta m) - \underbrace{\tilde{C}(0)}_{=C(m_\infty)} &= \frac{1}{2}(J(m_\infty)\delta m)^T C_d^{-1}(J(m_\infty)\delta m) + \frac{1}{2}\delta m^T C_m^{-1}\delta m \\
&\quad - \delta m^T \underbrace{(J(m_\infty)^T C_d^{-1}\delta T^{obs}(m_\infty) - C_m^{-1}m_\infty)}_{=g(m_\infty)=0,\ \text{gradient of } C(m) \text{ vanishes at the solution } m_\infty} \\
&= \frac{1}{2}\delta m^T (J(m_\infty)^T C_d^{-1} J(m_\infty) + C_m^{-1})\delta m \qquad (4.1) \\
&= \frac{1}{2}\delta m^T H(m_\infty)\delta m. \qquad (4.2)
\end{aligned}
$$

La matrice de covariance a posteriori est définie par :

$$C'_M = (J(m_\infty)^T C_d^{-1} J(m_\infty) + C_m^{-1})^{-1}. \qquad (4.3)$$

L'espace des modèles admissibles peut être caractérisé par :

$$(m - m_\infty)^T C_m'^{-1}(m - m_\infty) = \text{constant}, \qquad (4.4)$$

qui sont des ellipsoïdes de centre m_∞ et ce domaine est aussi défini par la fonction de densité de probabilité Gaussienne a posteriori

$$\exp\left(-\frac{1}{2}\delta m^T C_m'^{-1}\delta m\right). \qquad (4.5)$$

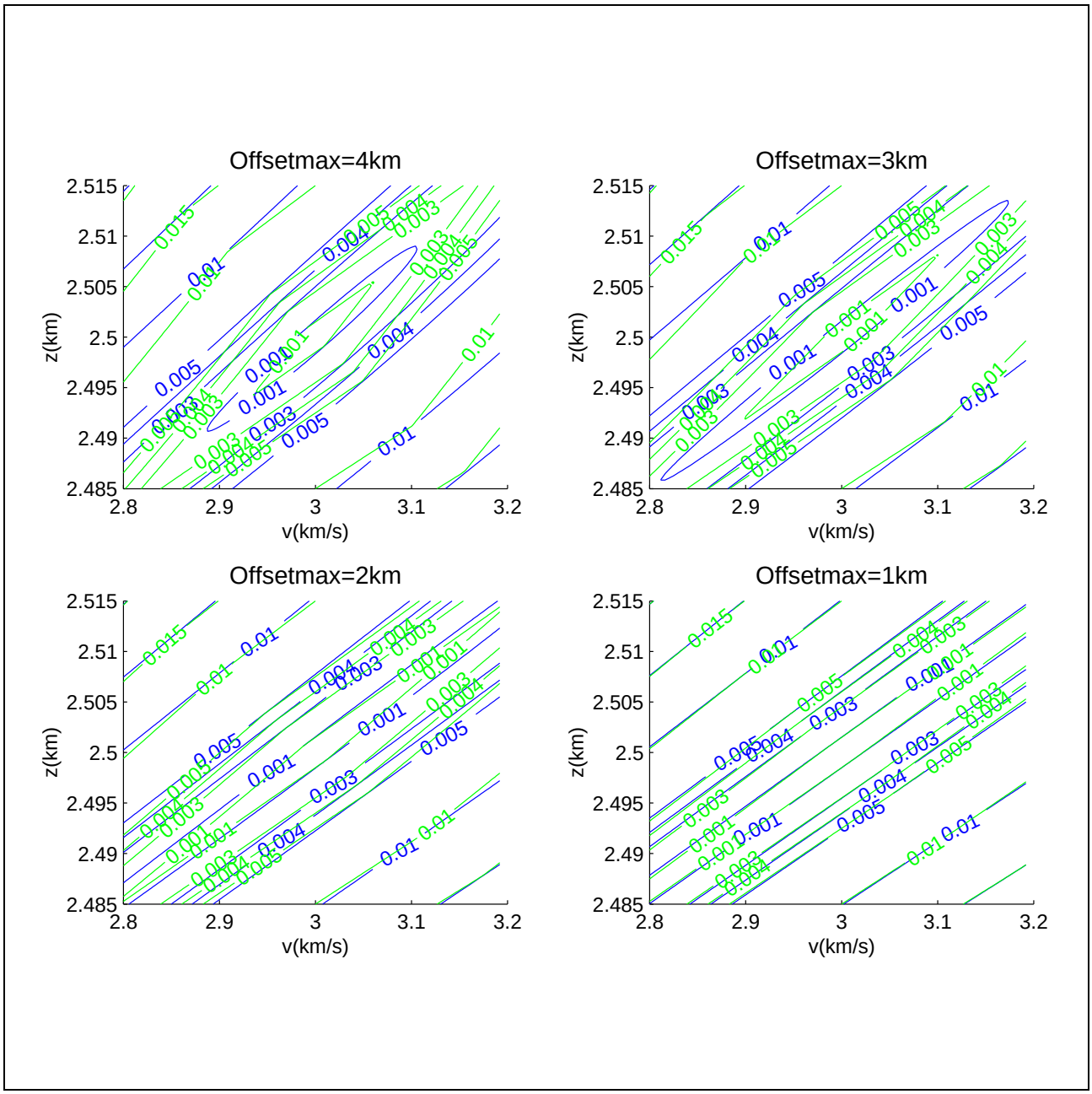

FIG. 4.6 – Représentation des iso-RMS et iso-résidus maximum pour les offsetmax $4000m$, $3000m$, $2000m$ et $1000m$.

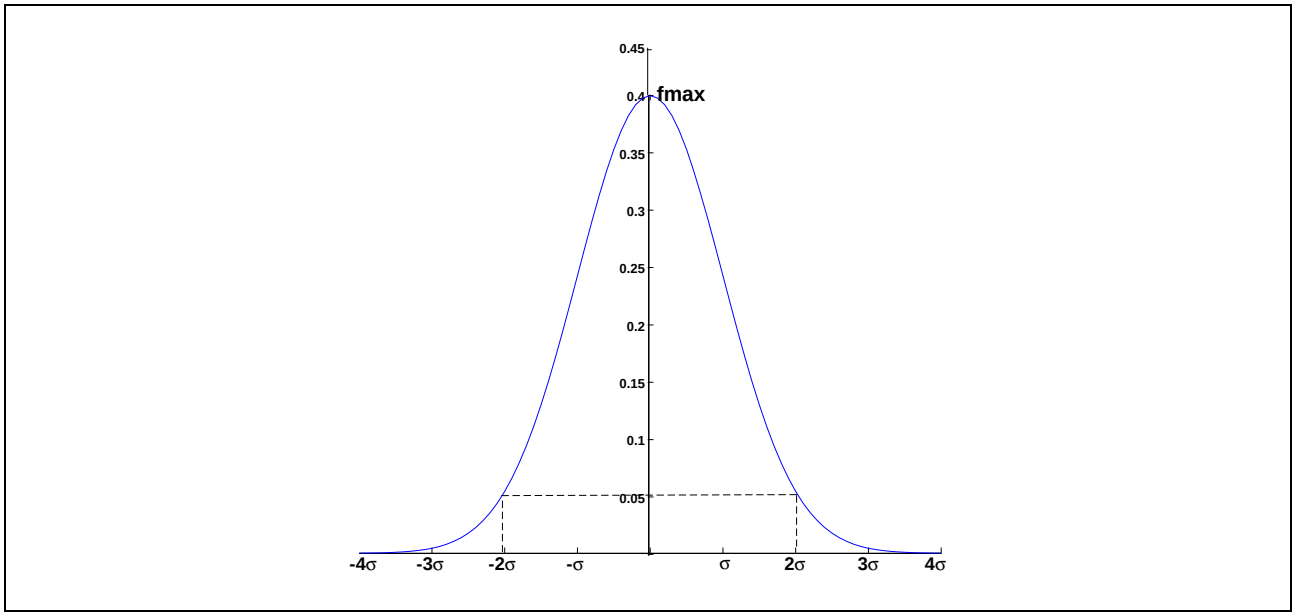

FIG. 4.7 – Fonction de densité de probabilité Gaussienne a posteriori.

Les termes diagonaux de C'_m sont les incertitudes sur les paramètres décrivant le modèle et les termes hors diagonaux sont les corrélations entre ces incertitudes.

Par exemple, la probabilité $P(-2\sigma \leq \delta m - m^{sol} \leq 2\sigma)$ que l'incertitude associée à la perturbation de modèle δm autour du modèle solution m^{sol} dans l'espace des paramètres soit comprise entre -2σ et 2σ est de 95%. Pour prendre en compte les corrélations entre les paramètres, on doit étudier l'ellipsoïde de confiance de 95%. Les axes des ellipsoïdes (4.4) sont définis par les vecteurs propres de C'_m, la racine carrée des valeurs propres donnant les incertitudes sur le vecteur propre associé.

Si l'on veut visualiser l'ellipsoïde de confiance associée à cette probabilité, on écrit tout d'abord :

$$\exp\left(-\frac{1}{2}\delta m^T \tilde{C}_m'^{-1} \delta m\right) = f(-2\sigma) = f(2\sigma). \tag{4.6}$$

Et en se reportant à la Figure 4.7 qui représente la fonction de densité de probabilité Gaussienne a posteriori, on observe que $f(-2\sigma) = f(2\sigma) = \frac{fmax}{8}$. Donc les ellipsoïdes que l'on cherche à visualiser ont pour équation :

$$\delta m^T \tilde{C}_m'^{-1} \delta m = 2\log(8). \tag{4.7}$$

Cette représentation permet de mettre en évidence les corrélations d'incertitudes entre les divers paramètres selon que l'ellipsoïde est de forme allongée pour des paramètres corrélés ou de forme proche d'un cercle pour des paramètres non-corrélés.

4.2.1 Incertitudes sur les paramètres du modèle

Une approche naturelle de l'analyse d'incertitude a posteriori est de calculer les incertitudes sur les paramètres à partir de la racine carrée des termes diagonaux de la matrice de covariance a posteriori

C'_m (voir par exemple le travail de D. Grenié et Y. Le Stunff, 2000).

La Figure 4.8 représente la matrice de covariance a posteriori associée au modèle solution de la Figure A.4 (Annexe A). Sur la Figure 4.9, nous représentons la racine carrée des termes diagonaux de C'_m, i.e. les incertitudes sur les paramètres B-splines du modèle.

Nous pouvons noter de très grandes valeurs pour les coefficients B-splines au bord. Ce n'est pas surprenant : la sensibilité de la fonction B-spline à ces coefficients est faible. Nous observons aussi de fortes corrélations entre les incertitudes sur ces coefficients : c'est un effet de la régularisation, ces coefficients étant principalement déterminés par cette information (voir section 4.2.2).

Les incertitudes sur les paramètres de B-spline ne sont pas très intéressantes en pratique : les incertitudes les plus grandes sur les coefficients de bord sont intrinsèques à la définition de la B-spline. Nous préfèrerions calculer des incertitudes sur des quantités physiques, par exemple, l'évaluation des fonctions B-spline sur un domaine physique (voir section 4.3.2).

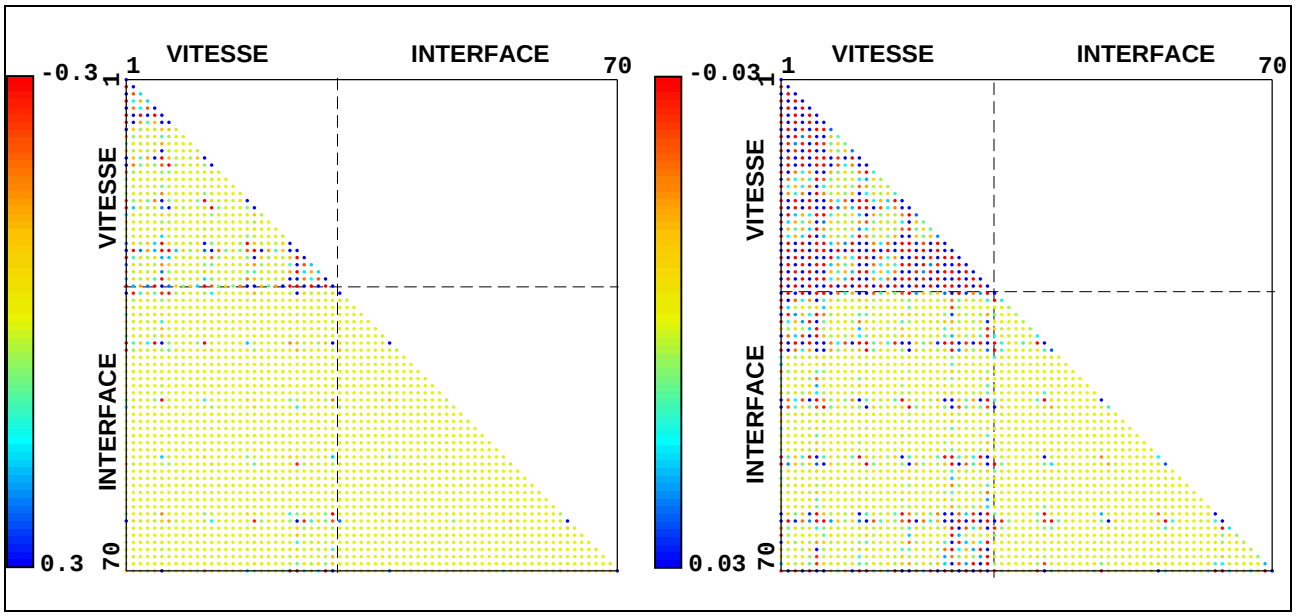

FIG. 4.8 – Matrice de covariance a posteriori du modèle solution de la Figure A.4 de l'Annexe A (triangulaire inférieure, la matrice est symétrique). La matrice est représentée avec deux échelles de couleur différentes pour mettre en évidence les termes de vitesse (gauche) ou les termes d'interface (droite).

4.2.2 Décomposition en vecteurs propres de la matrice de covariance a posteriori : détermination des directions de modèle les mieux/les moins bien déterminées

Une approche classique en vue d'estimer les incertitudes sur le modèle solution est de réaliser une décomposition en vecteurs propres de la Hessienne. L'étude du spectre de la matrice de covariance va

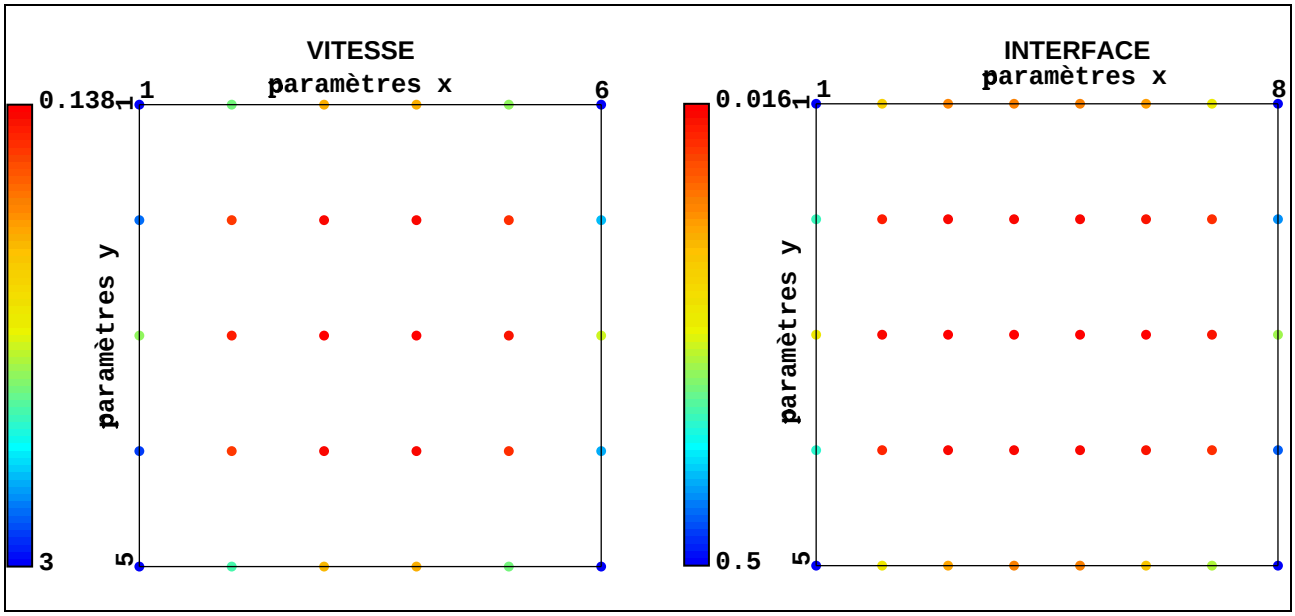

FIG. 4.9 – Incertitudes associées aux paramètres du modèle (racine carrée des termes diagonaux de la matrice de covariance a posteriori de la Figure A.4 (Annexe A). L'axe horizontal (resp. vertical) represente l'indice du paramètre B-spline dans la direction x (resp. y).

nous permettre d'apporter les premières conclusions quant à l'estimation des incertitudes. En effet, on vérifiera que les vecteurs propres associés aux plus grandes valeurs propres de la Hessienne sont les composantes du modèle les mieux déterminées par les données, l'inverse de la valeur propre donnant l'incertitude sur les vecteurs propres. L'analyse des composantes du modèle déterminé peut donc être réalisée en étudiant les valeurs propres et les vecteurs propres de la Hessienne.

Delprat-Jannaud et Lailly, 1992, ont étudié la décomposition en vecteurs propres de la matrice de covariance a posteriori pour avoir accès aux directions de modèle les mieux ou les moins bien déterminées ainsi qu'aux incertitudes associées.

Nous appliquons la même technique sur le premier exemple présenté dans l'annexe A. Comme cela vient d'être expliqué en introduction, les vecteurs propres de la matrice de covariance (ou inverse de la Hessienne) sont les axes des ellipsoïdes d'incertitude. Une autre interprétation est la suivante : la matrice Hessienne mesure l'influence d'une perturbation de modèle δm autour du modèle solution m_∞ sur la fonction coût quadratique

$$\tilde{C}(\delta m) - \tilde{C}(0) = \langle H\delta m, \delta m\rangle = \sum_i \alpha_i^2 \lambda_i, \tag{4.8}$$

où $(1/\lambda_i, e_i)$ sont les valeurs propres et vecteurs propres de la matrice de covariance a posteriori, δm est la perturbation de modèle définie par $\delta m = \sum_i \alpha_i e_i$ (décomposition dans la base des vecteurs propres).

La direction de modèle la mieux (respectivement la moins bien) déterminée correspond à la plus petite (respectivement la plus grande) valeur propre (λ_{min}) (resp. λ_{max}) de la matrice Hessienne, i.e. la plus grande (resp. la plus petite) perturbation de la fonction coût quadratique

$$\delta m = e_{max} \Rightarrow \tilde{C}(e_{max}) - \tilde{C}(0) = \langle He_{max}, e_{max} \rangle = \lambda_{max},$$
$$(\text{resp. } \delta m = e_{min} \Rightarrow \tilde{C}(e_{min}) - \tilde{C}(0) = \langle He_{min}, e_{min} \rangle = \lambda_{min}),$$

$1/\sqrt{(\lambda_i)}$ sont donc les incertitudes sur les perturbations de modèle colinéaires au vecteur propre associé e_i.

On calcule les vecteurs propres de la matrice de covariance a posteriori (Annexe G) du modèle solution présenté sur la Figure A.4 de l'Annexe A (figure 4.10 et Figure 4.11). La direction de modèle la mieux déterminée correspond aux perturbations lisses de la vitesse et de l'interface alors que la direction de modèle la moins bien déterminée correspond aux grandes perturbations de la vitesse et de l'interface (ceci est dû à l'opérateur intégral $T(m)$ qui n'est pas très sensible aux perturbations de modèle haute fréquence de petite amplitude). De plus, le vecteur propre associé à la direction de modèle la moins bien déterminée présente de grandes valeurs sur les bords et particulièrement dans les coins qui sont des zones peu «éclairées» par les rayons.

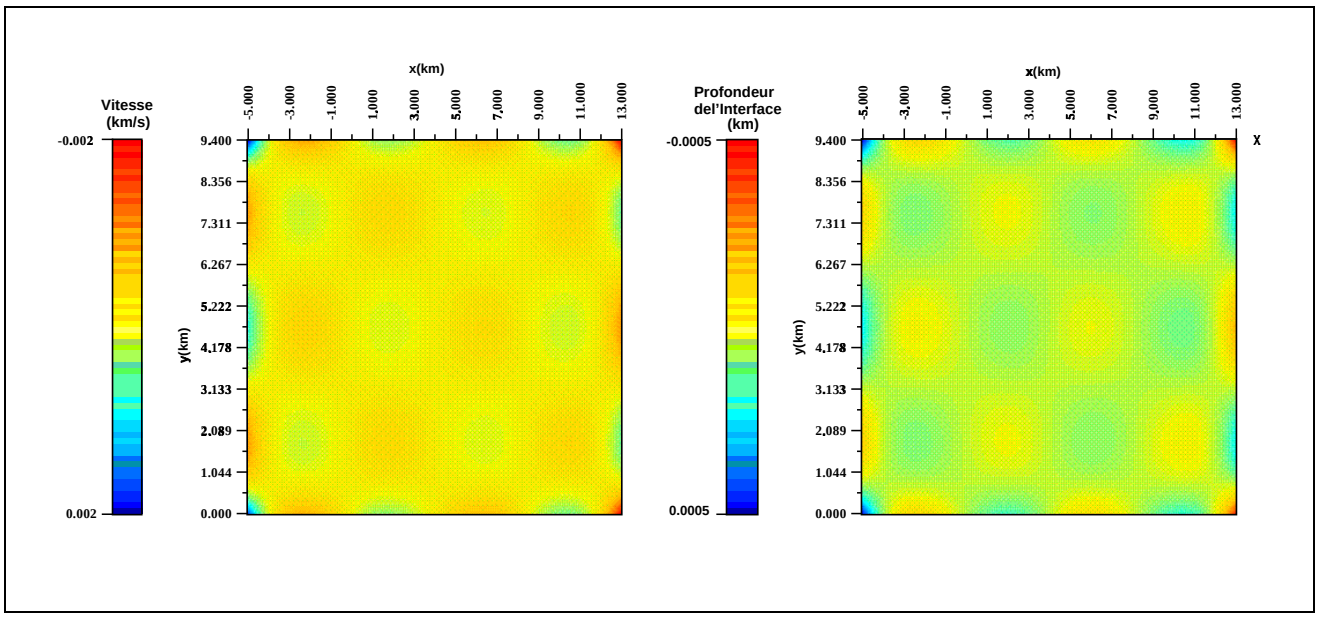

FIG. 4.10 – Vecteurs propres associés à la plus petite valeur propre de la matrice de covariance a posteriori du modèle solution de la Figure A.4 (voir Annexe A). Gauche : carte des variations latérales de vitesse. Droite : carte des variations de la profondeur de l'interface.

Les incertitudes asssociées à ces vecteurs propres, à savoir, la racine carrée des valeurs propres, ne sont pas significatives, les vecteurs propres étant composés à la fois des paramètres de vitesse et d'interface.

Comme cela a été montré par Tarantola, 1987, et par Delprat-Jannaud et Lailly, 1992, un choix judicieux de la norme dans l'espace des modèles destinée à être utilisée pour la décomposition de la matrice de covariance a posteriori doit être fait. Tarantola, 1987, s'interroge sur la signification de la matrice de covariance a posteriori associée à des paramètres d'unités physiques différentes, alors que

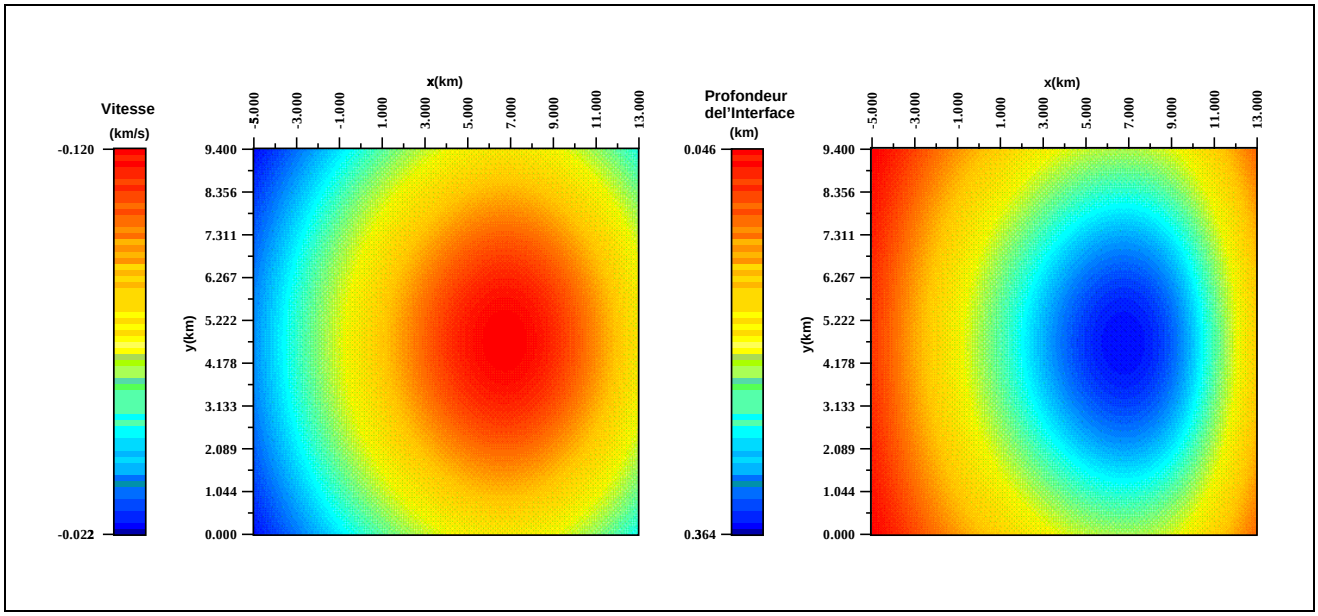

FIG. 4.11 – Vecteurs propres associés à la plus grande valeur propre de la matrice de covariance a posteriori du modèle solution de la Figure A.4 (voir Annexe A). Gauche : carte des variations latérales de vitesse. Droite : carte des variations de la profondeur de l'interface.

Delprat-Jannaud et Lailly, 1992, soulignent l'importance du choix approprié d'une norme pour obtenir la convergence du problème discret vers le problème continu (indépendance des incertitudes calculées par rapport à la discrétisation du modèle).

Nous devrons donc vérifier que la méthode fournit des résultats intrinsèques qui ne dépendent pas des unités physiques ou de la discrétisation choisie. Tarantola, 1987, propose d'utiliser la norme définie par la matrice de covariance a priori ce qui signifie que les vecteurs propres sont normalisés pour cette métrique. Delprat-Jannaud et Lailly, 1992, proposent aussi d'utiliser la norme associée à l'information a priori sur le modèle (régularisation par les moyennes des dérivées secondes comme pour le problème (3.13)), i.e. la norme H^2. Ils montrent, pour cette norme, que

(i) la convergence des valeurs propres et vecteurs propres du problème discret vers les valeurs propres et vecteurs propres du problème continu est assurée,

(ii) les valeurs propres de la matrice de covariance a posteriori associée au problème continu sont sur-estimées par les valeurs propres associées au problème discret,

(iii) calculer précisément l'incertitude associée à la direction de modèle la moins bien déterminée requiert une discrétisation plus fine que celle requise par le calcul de l'incertitude associée aux directions les mieux déterminées.

Les résultats présentés sur les Figures 4.12, 4.13 et 4.14 confirment ces résultats théoriques. Malgré ces résultats de convergence intéressants, nous ne recommandons pas cette méthode dans des applications pratiques. La principale raison est l'interprétation physique difficile des vecteurs propres. Les

vecteurs propres sont une combinaison des paramètres de vitesse et d'interface, ce que nous calculons est donc l'incertitude de ces combinaisons qu'il est souvent difficile de relier à des quantités physiques. De plus, cette approche est coûteuse : la diagonalisation de la matrice de covariance a posteriori complète doit être réalisée.

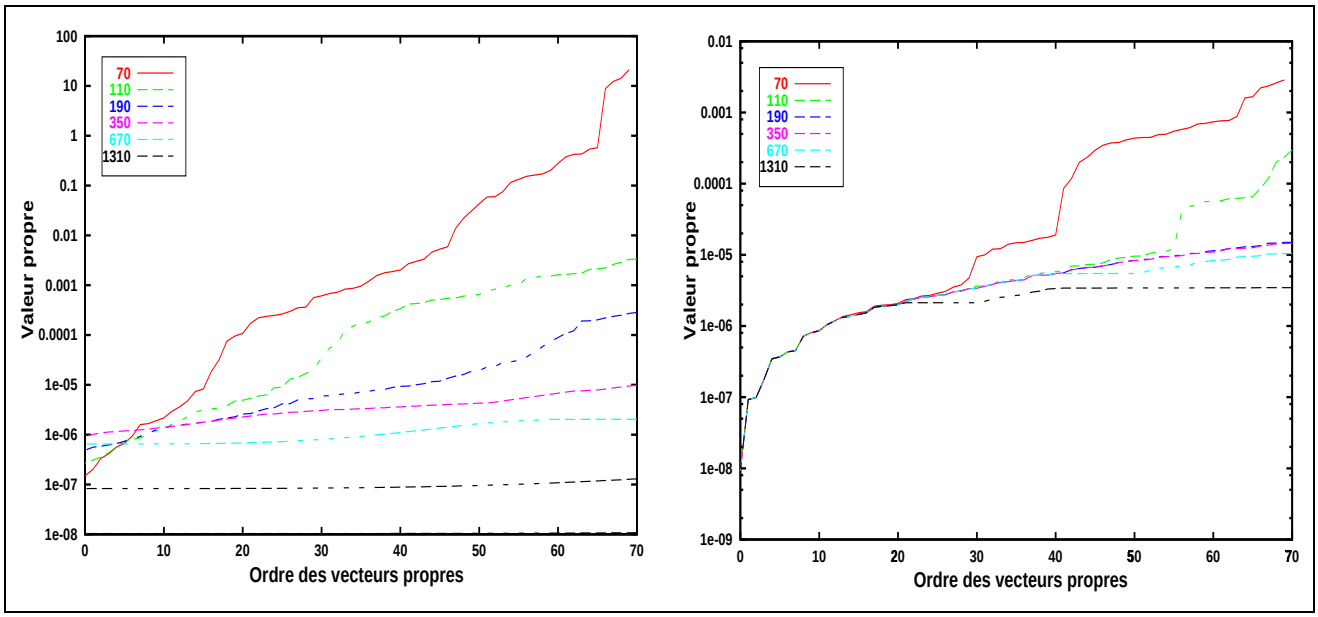

FIG. 4.12 – Valeurs propres de la matrice de covariance a posteriori du modèle solution de la Figure A.4 pour différentes discrétisations (Nombres de paramètres différents suivant la direction). Gauche : la norme L^2 a été utilisée pour réaliser la diagonalisation. Droite : la norme H^1 a été utilisée pour réaliser la diagonalisation.

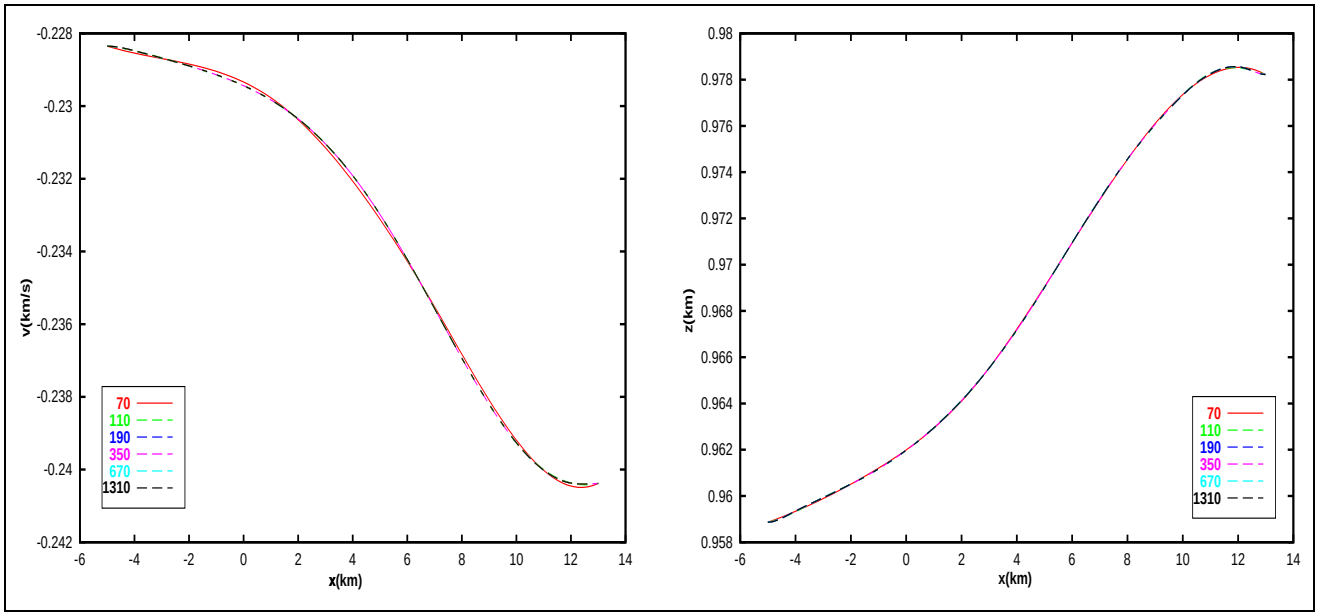

FIG. 4.13 – H^1-vecteur propre le mieux déterminé (associé à la plus faible valeur propre de la matrice de covariance a posteriori du modèle solution de la Figure A.4) pour différentes discrétisations (nous faisons varier le nombre de paramètres suivant la direction x). Gauche : coupe du modèle de vitesse suivant la direction x. Droite : coupe de l'interface suivant la direction x.

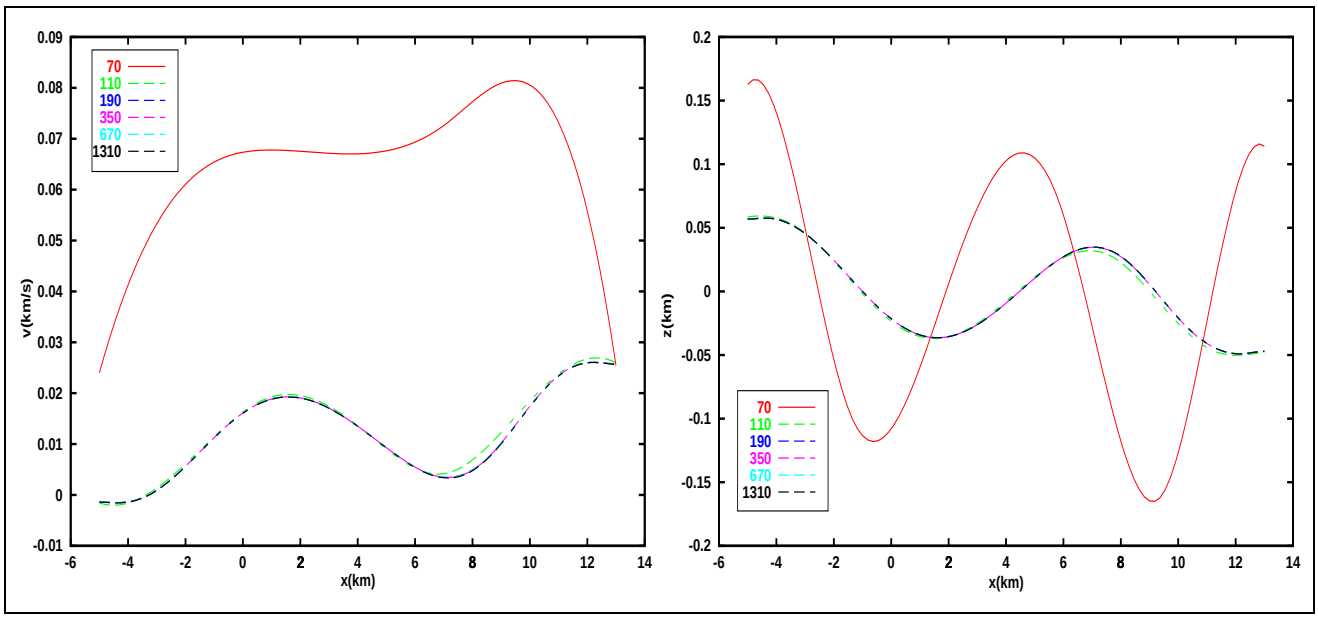

FIG. 4.14 – H^1-vecteur propre moins bien déterminé (associé à la plus faible valeur propre de la matrice de covariance a posteriori du modèle solution de la Figure A.4) pour différentes discrétisations (nous faisons varier le nombre de paramètres suivant la direction x). Gauche : coupe du modèle de vitesse suivant la direction x. Droite : coupe de l'interface suivant la direction x.

4.3 Deux méthodes proposées

Nous allons présenter ici deux méthodes qui nous paraissent intéressantes en pratique pour accéder aux incertitudes sur le modèle solution de la tomographie de réflexion. La première (section 4.3.1) permet d'échantillonner la fonction de densité de probabilité a posteriori et donc de simuler des modèles équiprobables. La seconde (section 4.3.2) permet de calculer les variances de quantités physiques du modèle, décrites par des combinaisons linéaires de paramètres (ce que nous appelons des «macro-paramètres»).

4.3.1 Les simulations de modèles admissibles

4.3.1.1 Principe

Contrairement à l'analyse des vecteurs propres (section 4.2.2), la simulation de modèles admissibles à partir de la fonction de densité de probabilité a posteriori permet une analyse plus complète que la seule analyse des termes diagonaux de la matrice de covariance a posteriori. En effet, les corrélations entre les incertitudes sont prises en compte. De plus, les simulations fournissent directement des résultats interprétables, i.e. des modèles physiques. Cette méthode consiste en des simulations de perturbation de modèle suivant la fonction de densité de probabilité a posteriori qui s'écrit :

$$\exp\{-\frac{1}{2}\delta m^T C_m'^{-1} \delta m\}. \tag{4.9}$$

La méthode (voir, par exemple, Parker, 1994) consiste en un changement de variable :

$$\delta m' = U \delta m, \tag{4.10}$$

où U est la matrice triangulaire supérieure obtenue par une décomposition de Cholesky (voir Ciarlet, 1988, pour le détail de la méthode) de l'inverse de la matrice de covariance a posteriori $C_m'^{-1} = H = U^T U$ (matrice symétrique définie positive).

Cela donne la fonction de densité de probabilité gaussienne non corrélée avec une variance unité

$$\exp\{-\frac{1}{2}\delta m'^T \delta m'\}. \tag{4.11}$$

Nous réalisons des simulations d'un vecteur aléatoire $\delta m'$ suivant (4.11) et construisons des simulations gaussiennes corrélées du vecteur δm en résolvant

$$\delta m = U^{-1} \delta m'. \tag{4.12}$$

La méthode des simulations nous permet donc d'avoir accès à plusieurs modèles admissibles.

4.3.1.2 Applications

On présente ici deux applications de la méthode des simulations sur des exemples synthétiques et nous verrons une application sur données réelles dans le chapitre 6.

4.3.1.2.1 Premier exemple (voir Annexe A)

La Figure 4.15 montre 20 modèles simulés obtenus par cette méthode appliquée au modèle solution décrit dans l'Annexe A. Les incertitudes les plus grandes sur la profondeur de l'interface ($\approx 10m$) et sur les variations de vitesse ($\approx 50m$) sont localisées sur les bords du modèle, ce qui correspond aux zones qui ne sont pas bien illuminées. Dans les autres zones, les incertitudes sur la profondeur sont de l'ordre de quelques mètres. On avait déjà observé cela sur les incertitudes sur les paramètres du modèle, mais ici, on obtient une estimation fiable des incertitudes sur la profondeur de l'interface et non pas sur le paramètre B-spline.

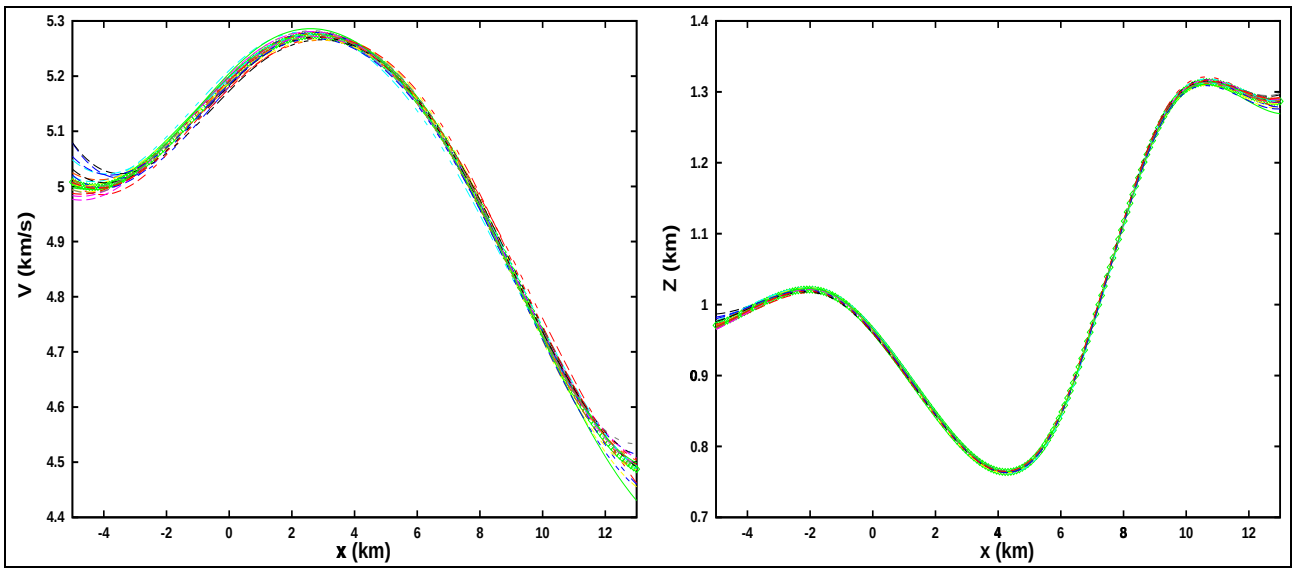

FIG. 4.15 – Coupes des 20 modèles simulés suivant l'axe des x sur les modèles de vitesse 3D simulés. Gauche : variations de vitesse. Droite : variations de la profondeur de l'interface.

4.3.1.2.2 Second exemple (voir Annexe B)

Nous avons réalisé des simulations sur la seconde interface (en bleu sur la partie droite de la Figure 4.16) du modèle solution décrit dans l'Annexe B. Les deux figures sur la gauche des Figures 4.16 et 4.17 représentent 20 modèles simulés obtenus par cette méthode. On constate ici aussi que les plus grandes incertitudes sont localisées aux bords du modèle.

4.3.1.3 Conclusions

Cette méthode est assez intéressante pour l'interprétation directe des résultats en dépit du coût de la factorisation de Choleski de la matrice de covariance a posteriori. Elle permet de visualiser différents modèles admissibles.

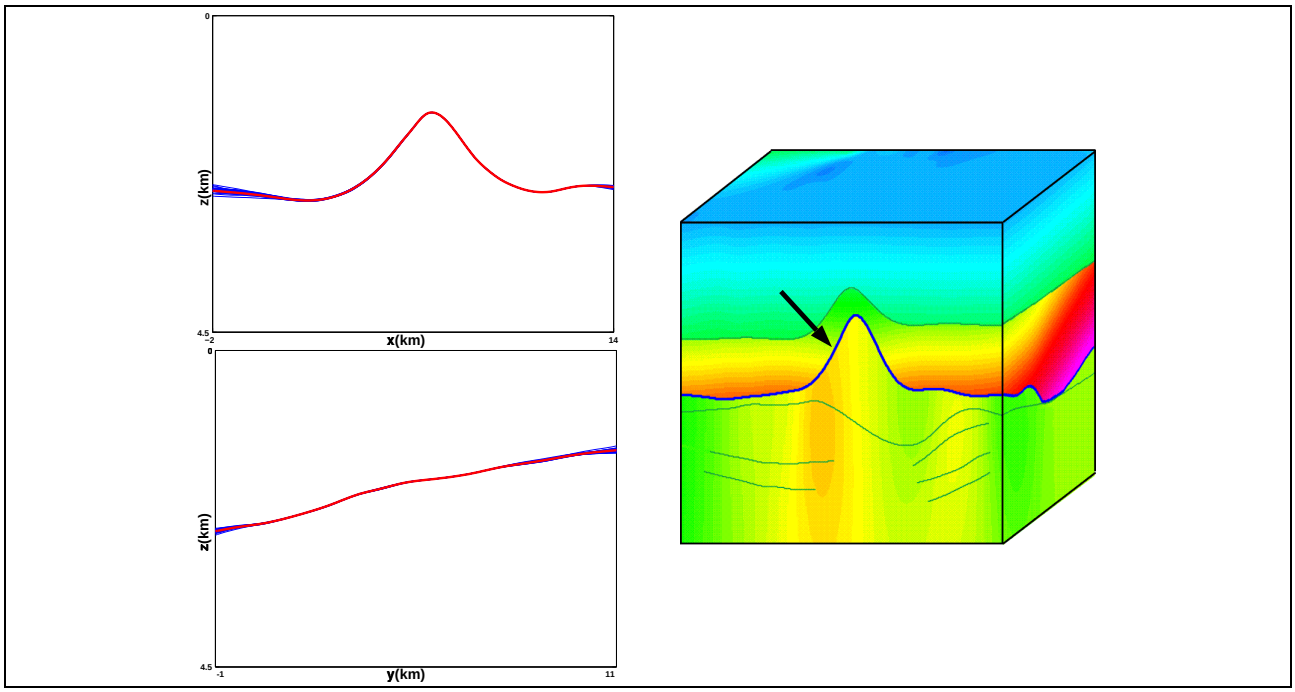

FIG. 4.16 – Coupes suivant l'axe des x (en haut à gauche) et des y (en bas à gauche) de la profondeur de la 2ième interface sur les 20 modèles simulés.

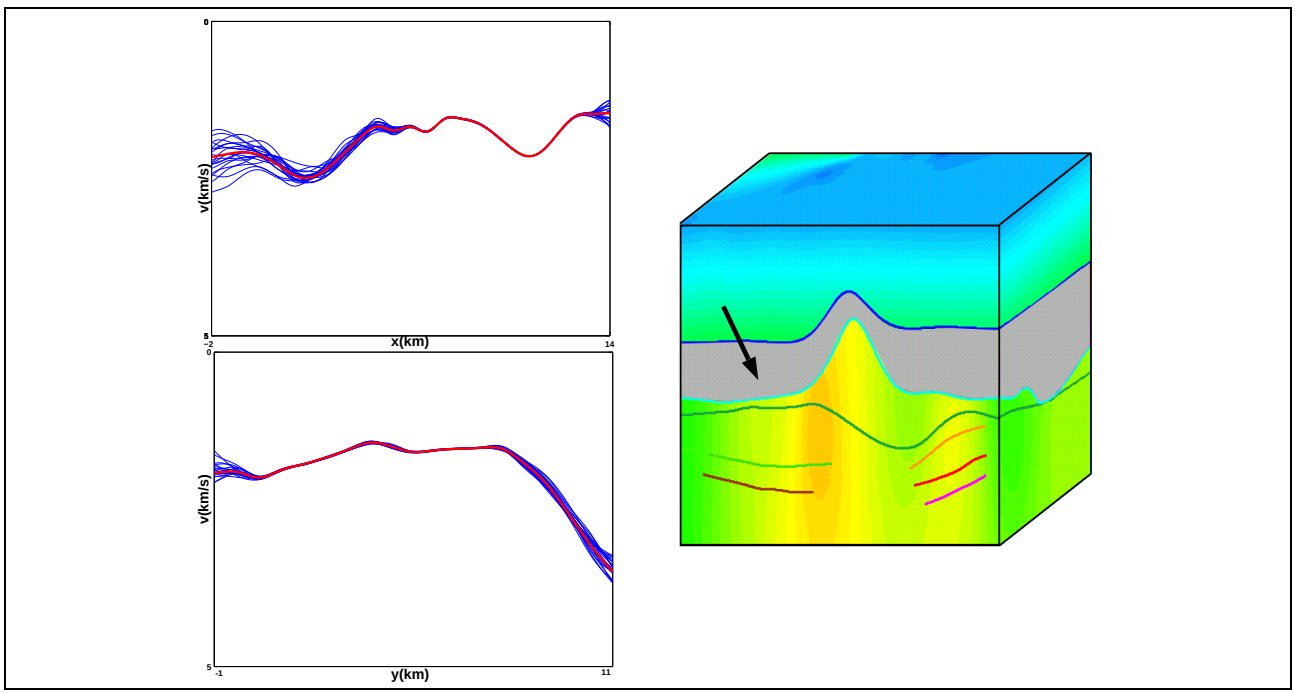

FIG. 4.17 – Coupes suivant l'axe des x (en haut à gauche) et des y (en bas à gauche) de la vitesse dans la 2ième couche sur les 20 modèles simulés.

4.3.2 Utilisation de macro-paramètres géologiques pour le calcul d'incertitude

4.3.2.1 Principe

Dans cette section, nous proposons une méthode qui permet de traiter des modèles de grande taille avec un coût de calcul raisonnable. De plus, nous voulons donner au géophysicien des incertitudes sur des quantités physiques, et non sur tous les paramètres numériques que nous inversons dans le processus de tomographie.

L'approche proposée consiste à construire quelques «macro-paramètres» ayant un intérêt géophysique. Ces macro-paramètres sont des combinaisons linéaires des paramètres à inverser : par exemple, la moyenne des variations verticales de vitesse dans une zone, la courbure d'une interface, la profondeur moyenne d'une couche ... Grenié, 2001, a introduit la notion de macro-paramètres (sa principale préoccupation étant d'éviter les problèmes numériques dans l'inversion du Hessien complet). Nous proposons ici une généralisation de son travail (définition générale des macro-paramètres) qui permet le calcul des incertitudes pour des problèmes 3D de grande taille.

Nous définissons un macro-paramètre comme une combinaison linéaire de paramètres :

$$P = Bp, \tag{4.13}$$

où p est le vecteur des n_p paramètres de modèle, P est le vecteur des n_{MP} macro-paramètres et B la matrice de passage des paramètres aux macro-paramètres (C. Duffet et D. Sinoquet, 2002).

Nous calculons la matrice de covariance a posteriori dans l'espace des macro-paramètres, $\tilde{C'_m}$, comme suit :

$$\tilde{C'_m} = BC'_mB^T, \tag{4.14}$$

où C'_m la matrice de covariance a posteriori dans l'espace des paramètres. Notons que la taille de $\tilde{C'_m}$, $n_{MP} \times n_{MP}$, est par définition inférieure à la taille de C'_m, $n_p \times n_p$.

Pour obtenir cette matrice, nous n'avons pas besoin de calculer l'inverse de la Hessienne complète, $H^{-1} = C'_m$ dans l'espace des paramètres.

En effet, pour obtenir

$$\tilde{C'_m} = \tilde{H}^{-1} = BH^{-1}B^T \tag{4.15}$$

dans l'espace des macro-paramètres, il suffit de disposer de

$$H^{-1}B_j^T, \tag{4.16}$$

où B_j^T sont les différentes colonnes de B^T (calculées par B^Te_j où les e_j sont des vecteurs tels que : $e_{ij} = 0 \ \forall i \neq j$ et $e_{jj} = 1$), et de résoudre n_{MP} systèmes linéaires (i.e., autant que de macro-paramètres).

Ces systèmes linéaires

$$Hx = B_j^T \tag{4.17}$$

sont similaires au problème linéarisé que nous résolvons à chaque itération de Gauss-Newton (3.33).

Par conséquent, le coût de calcul pour un macro-paramètre est comparable au coût d'une itération du processus d'inversion, sans prendre en compte la résolution du problème direct.

Voici différents types de macro-paramètres simples (notés P) :

- moyenne des paramètres B-splines p_i ($i = 1, ..., np$) pour chaque vitesse inversée et chaque interface inversée

$$P = \frac{1}{np} \sum_{i=1}^{np} p_i, \tag{4.18}$$

  avec np est le nombre de paramètres.

- moyenne de la fonction B-spline $u \in L^2(\Omega)$ ou de ses dérivées, pour chaque vitesse inversée et chaque interface inversée

$$P = \frac{1}{\Delta x \Delta y} \iint_{\Omega_x \times \Omega_y} u(x, y) dx dy, \tag{4.19}$$

  avec $\Omega = \Omega_x \times \Omega_y$ ($\Omega_x = [x_{min}; x_{max}]$ et $\Omega_y = [y_{min}; y_{max}]$) et $\Delta x = x_{max} - x_{min}$ et $\Delta y = y_{max} - y_{min}$.

- valeur en chaque point d'une grille de la fonction B-spline pour chaque vitesse inversée et chaque interface inversée

$$P = v_i \text{ ou } P = z_i, \quad i \in \{1, ..., n\} \tag{4.20}$$

  avec n le nombre de points de la grille.

Nous étudierons un macro-paramètre plus complexe et plus réaliste qui nous permettra d'évaluer l'incertitude sur le volume d'une couche ou d'un piège pétrolier dans le paragraphe **??** et nous l'illustrerons sur données réelles dans le chapitre 6.

4.3.2.2 Applications

On présente ici deux applications de cette méthode sur les exemples synthétiques présentés en Annexes A et B, et nous verrons une application sur données réelles dans le chapitre 6.

4.3.2.2.1 Premier exemple (voir Annexe A)

Ici nous illustrons l'approche macro-paramètre sur l'exemple de la Figure A.1 (Annexe A). Nous choisissons des macro-paramètres simples : la moyenne des variations de vitesse dans la couche, la moyenne de ses dérivées, la moyenne de la profondeur d'une interface et de ses dérivées. Les incertitudes sur ces macro-paramètres sont calculées sur le modèle solution de la Figure A.4. Les résultats sont récapitulés dans le tableau 4.1.

On observe des incertitudes faibles : cela s'explique par le fait que l'on considère des macro-paramètres associés à des moyennes donc les incertitudes fortes aux bords sont gommées. D'autre part, nous pouvons noter de fortes corrélations entre la vitesse et la profondeur de l'interface.

	Macro-paramètres		
	B-spline	**DX**	**DY**
VITESSE	3.74m/s	0.001/s	0.001/s
INTERFACE	0.74m	0.0003	0.0003
CORRELATIONS	0.97	0.98	0.98

TAB. 4.1 – Écarts types (racines carrées des termes diagonaux de $\tilde{C}'_m$) associées à différents macro-paramètres : moyenne de la fonction B-spline, de ses dérivées premières en x, y et z sur le domaine de définition et les corrélations normalisées entre vitesse et interface. Cette analyse de l'incertitude est réalisée pour le modèle solution de la Figure A.4. La moyenne de la vitesse est de $4.54km/s$ et la moyenne de la profondeur de l'interface est de $1.07km$.

4.3.2.2.2 Second exemple (voir Annexe B)

Ici nous illustrons l'approche macro-paramètre sur l'exemple de la Figure B.1 (Annexe B) et plus exactement sur l'interface représentée en bleu (partie droite de la Figure 4.18) et nous allons comparer l'analyse d'incertitude sur l'inversion «layer stripping» (inversion «couche par couche» en français) d'une part puis, sur l'inversion «layer stripping + inversion globale» (voir Annexe B pour le détail de ces techniques d'inversion).

Les macro-paramètres choisis sont : la moyenne de la profondeur de l'interface dans différentes régions du modèle et nous nous plaçons tout d'abord dans le cas de l'approche «layer stripping». Le domaine de définition est séparé en 5 régions (partie gauche des Figures 4.18 et 4.19 dans le cas de l'approche «layer stripping»). Notons que nous n'incluons pas les bords du modèle dans la définition des régions : les bords n'étant pas éclairés par les rayons, il est évident que les incertitudes associées sont fortes.

Ainsi, sur la Figure 4.18 on observe des différences importantes sur les incertitudes associées aux macro-paramètres des différentes régions. Par exemple, la moyenne de la profondeur de l'interface solution (en bleu sur la partie droite de la Figure 4.18) est $2.35km$. Si nous considérons la région représentée en vert sur la partie gauche de la Figure 4.18, nous pouvons dire que nous avons une incertitude sur la profondeur de l'interface de $\pm 2.76m$ (deux fois l'écart type).

De la même manière, considérons la vitesse de la couche grisée (partie droite de la Figure 4.19). Le macro-paramètre étudié est la moyenne de la vitesse dans différentes régions. Nous conservons les mêmes régions que précédemment et nous obtenons les mêmes conclusions. Par exemple, la moyenne de la vitesse (en gris sur la partie droite de la Figure 4.19) est $1.92km/s$. Si nous considérons la région représentée en jaune sur la partie gauche de la Figure 4.19, nous pouvons dire que nous avons une incertitude sur la variation de la vitesse dans la couche de $\pm 10.46m/s$ (deux fois l'écart type).

D'autre part, il est évident que la région la moins bien déterminée est la région la moins bien illuminée. Si nous comparons maintenant ces résultats avec ceux de l'approche «layer stripping + inversion globale» (résultats récapitulés dans le tableau 4.2), nous retrouvons la même «distribution» des incertitudes sur le domaine mais la valeur de ces incertitudes est plus faible que pour l'approche «layer-stripping». Par conséquent, la prise en compte des temps de trajet associés à toutes les couches simultanément (cas de l'inversion globale) dans l'analyse a posteriori permet de réduire les incertitudes

sur le modèle solution.

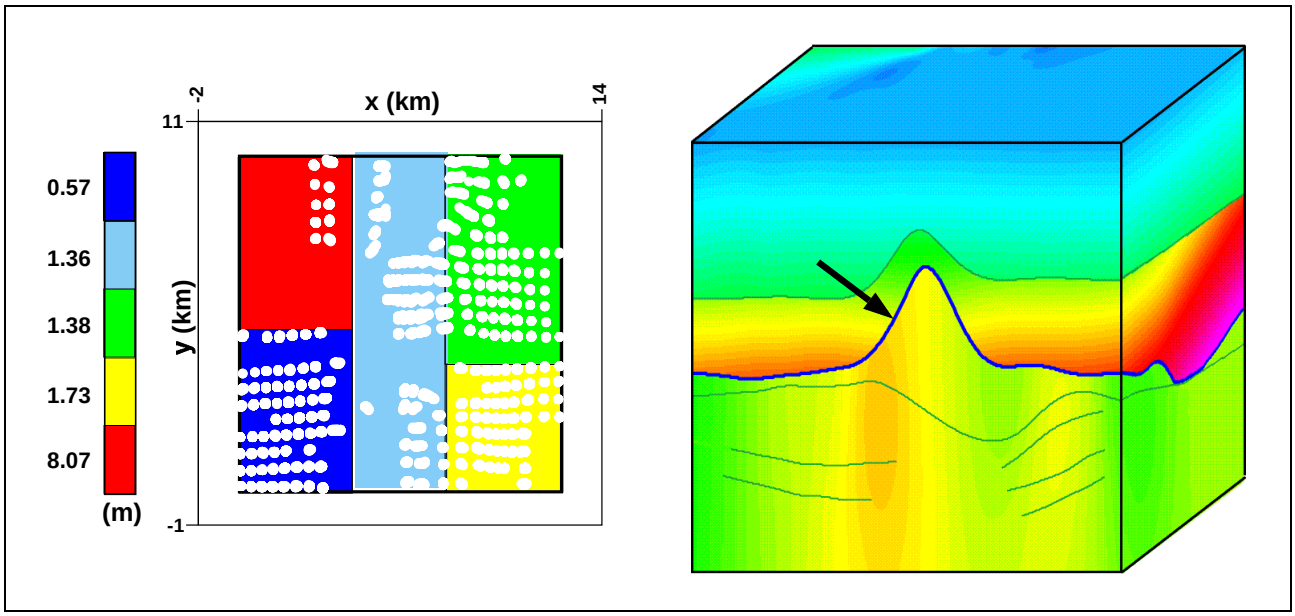

FIG. 4.18 – Écarts types associés à des macro-paramètres définis par la moyenne de l'interface (en bleu sur la figure de droite) dans 5 régions différentes (décrites par les rectangles sur la figure de gauche, la couleur indique l'écart-type du macro-paramètre). Les points sur la figure de gauche représentent les points d'impact des rayons sur l'interface. On est dans le cas de l'approche **«layer stripping»**.

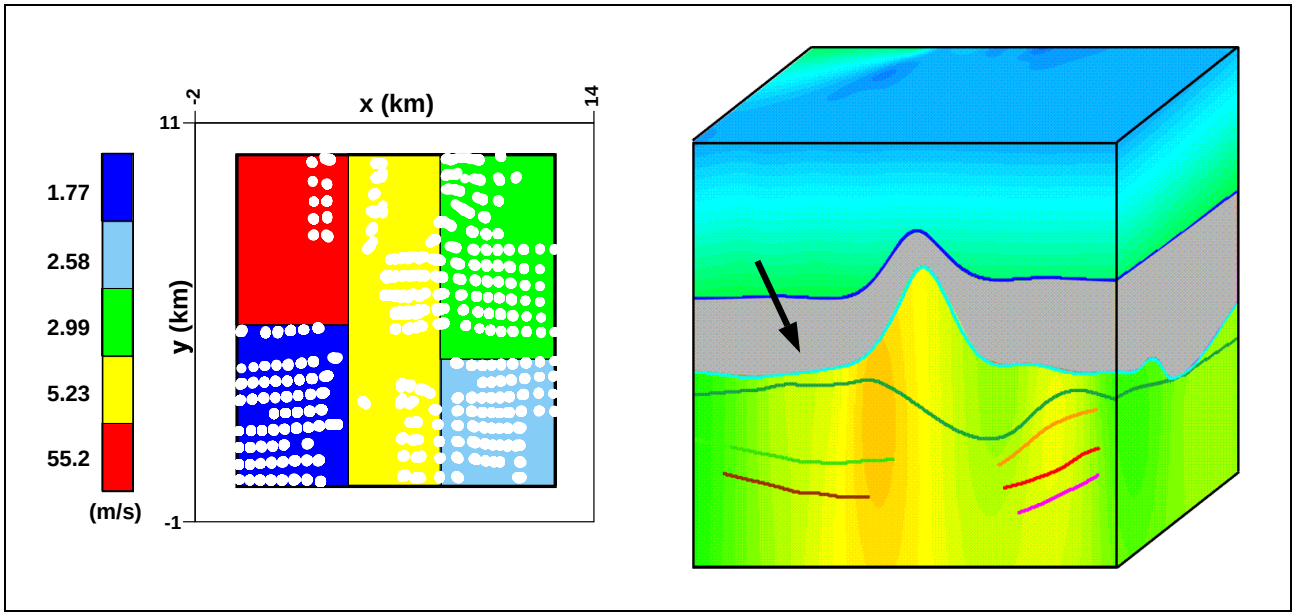

FIG. 4.19 – Incertitudes sur des macro-paramètres définis par la moyenne de la vitesse (en gris sur la figure de droite) dans 5 régions différentes (décrites par les rectangles sur la figure de gauche, la couleur indique l'écart-type du macro-paramètre). Les points sur la figure de gauche représentent les points d'impact des rayons sur l'interface. On est dans le cas de l'approche **«layer stripping»**.

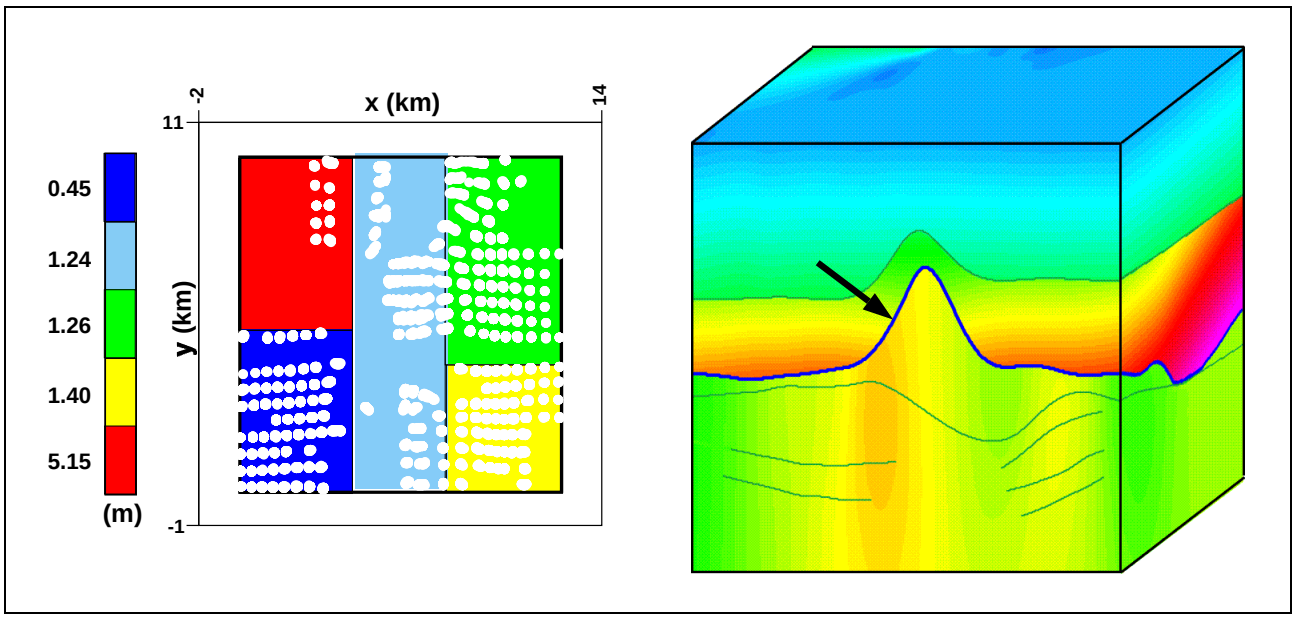

FIG. 4.20 – Écarts types associés à des macro-paramètres définis par la moyenne de l'interface (en bleu sur la figure de droite) dans 5 régions différentes (décrites par les rectangles sur la figure de gauche, la couleur indique l'écart-type du macro-paramètre). Les points sur la figure de gauche représentent les points d'impact des rayons sur l'interface. On est dans le cas de l'approche **«layer stripping + inversion globale»**.

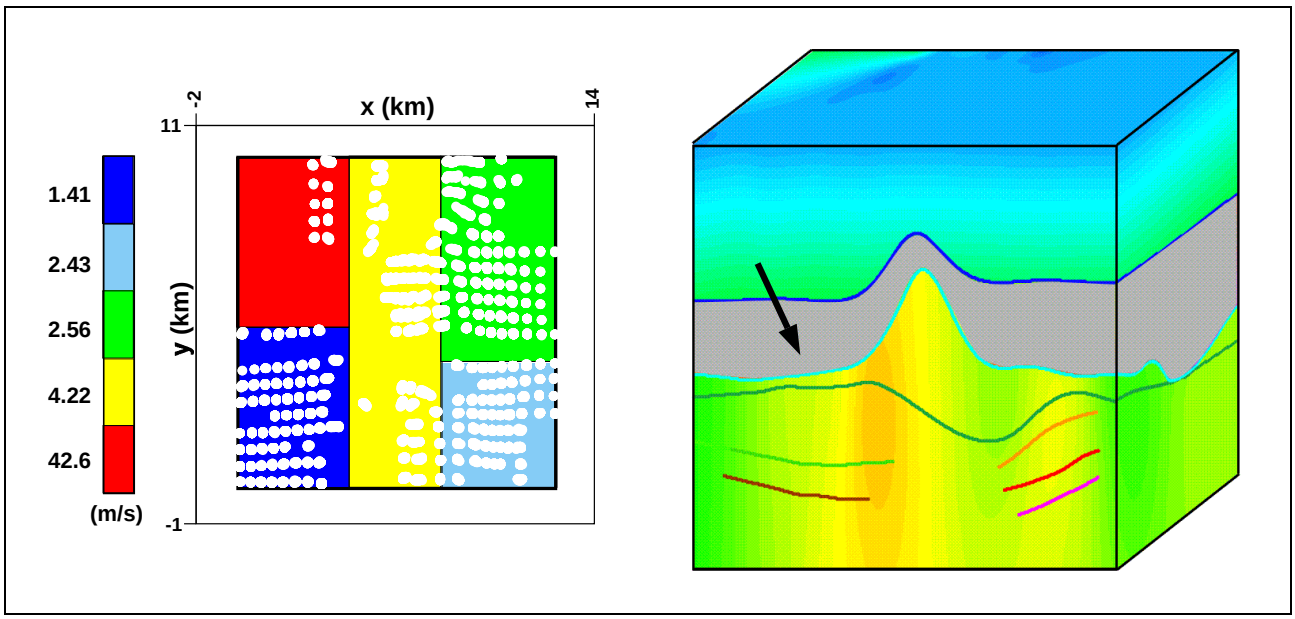

FIG. 4.21 – Incertitudes sur des macro-paramètres définis par la moyenne de la vitesse (en gris sur la figure de droite) dans 5 régions différentes (décrites par les rectangles sur la figure de gauche, la couleur indique l'écart-type du macro-paramètre). Les points sur la figure de gauche représentent les points d'impact des rayons sur l'interface. On est dans le cas de l'approche **«layer stripping + inversion globale»**.

Régions		**Layer stripping**	**Layer stripping + inversion globale**
région 1	INTERFACE	8.07m	5.15m
	VITESSE	55.2m/s	42.6m/s
région 2	INTERFACE	0.57m	0.45m
	VITESSE	1.77m/s	1.41m/s
région 3	INTERFACE	1.36m	1.24m
	VITESSE	5.23m/s	4.22m/s
région 4	INTERFACE	1.38m	1.26m
	VITESSE	2.99m/s	2.56m/s
région 5	INTERFACE	1.73m	1.40m
	VITESSE	2.58m/s	2.43m/s

TAB. 4.2 – Écarts types associés aux différents macro-paramètres sur les 5 régions du domaine de définition dans le cas de l'approche **«layer stripping»** et dans le cas de l'approche **«layer stripping + inversion globale»**. Noter que les régions sont numérotées de bas en haut et de gauche à droite (voir par exemple la Figure 4.18).

4.3.2.3 Illustration du macro-paramètre volume

4.3.2.3.1 Principe

Le formalisme général et la flexibilité de l'approche macro-paramètre vont nous permettre de construire un macro-paramètre (P) correspondant à des quantités qui sont intéressantes pour le géophysicien : un volume de couche ou de piège pétrolier.

Prenons l 'exemple de la Figure 4.22 : la couche qui nous intéresse est délimitée par les interfaces int1, int2, F1 et F2. Ici, l'exemple est 2D pour simplifier mais l'implémentation a été faite pour des modèles 3D. Nous nous sommes limités aux interfaces explicites en z ($z(x, y)$).

On veut donc calculer l'incertitude que l'on a sur le volume délimité par les $z_i(x)$ ($i = 1, ..., 4$). La couche est repérée par sa position relative aux interfaces la délimitant (cf Clarke, 1997) : ainsi la couche grisée de la Figure 4.22 est repérée par (int1- ;int2+ ;F1+ ;F2+), le signe + signifiant que les points de la couche sont au-dessus de l'interface considérée ($z_{int2} > z$) et le signe - l'inverse.

Le volume de la couche s'écrit donc :

$$P = -\int_{x_1}^{x_2} z_{int1-}(x)dx + \int_{x_2}^{x_3} z_{F2+}(x)dx + \int_{x_3}^{x_4} z_{int2+}(x)dx + \int_{x_4}^{x_1} z_{F1+}(x)dx, \tag{4.21}$$

avec x_i, $i = 1, ..., 4$ les points d'intersection entre les différentes interfaces. Le signe devant l'intégrale correspond au signe de la définition de la couche par (int1- ;int2+ ;F1+ ;F2+).

On peut aussi l'écrire sous la forme linéaire $P = Bp$:

$$P = Bp = -(x_2 - x_1)B_{int1-}p_{int1-} + (x_3 - x_2)B_{F2+}p_{F2+} + (x_4 - x_3)B_{int2+}p_{int2+} + (x_1 - x_4)B_{F1+}p_{F1+}, \tag{4.22}$$

avec x_i, $i = 1, ..., 4$ les points d'intersection entre les différentes interfaces, p_i, $i = int1-, F2+, int2+, F1+$, les paramètres des interfaces et B_i, $i = int1-, F2+, int2+, F1+$, l'opérateur moyenne de la fonction B-spline de la profondeur des interfaces.

Chaque $P_i = B_i p_i$, $i = int1-, F2+, int2+, F1+$, correspond au macro-paramètre moyenne de la profondeur de l'interface.

La difficulté dans la construction de ce macro-paramètre est de repérer les points d'intersection entre les différentes interfaces : pour chaque interface nous n'avons pas la même grille de B-spline. En pratique, on parcourt l'interface afin de repérer deux noeuds de B-spline situés de part et d'autre du point d'intersection, C'est-à-dire des noeuds pour lesquels la combinaison de signe servant à se repérer par rapport aux interfaces a changé. Ensuite on teste la combinaison de signes des points situés dans cet intervalle jusqu'à se rapprocher le plus possible du point d'intersection.

Notons que ce formalisme est généralisable à tous les types d'interface (ici on a réduit l'étude aux interfaces explicites en z). Le point clé réside dans le choix de considérer des interfaces explicites dans l'une des directions et qu'elles soient définies sur tout le domaine.

Nous allons dans le paragraphe suivant illustrer ce macro-paramètre volume sur un exemple synthétique.

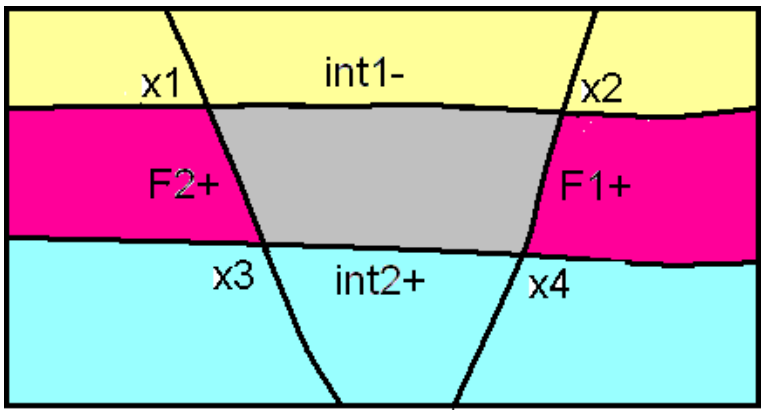

FIG. 4.22 – Illustration d'un volume.

4.3.2.3.2 Application

Le modèle étudié est un modèle synthétique 3D composé de trois interfaces régulières (Figure 4.23), s'étendant sur $10km$ dans la direction x et sur $10km$ dans la direction y et définies grâce à :

- 10 (respectivement 6) paramètres de B-spline dans la direction des x (respectivement dans la direction des y) pour l'interface int1,

- 8 (respectivement 4) paramètres de B-spline dans la direction des x (respectivement dans la direction des y) pour l'interface int2,

- 10 (respectivement 10) paramètres de B-spline dans la direction des x (respectivement dans la direction des y) pour l'interface int3.

Ici nous choissisons de ne pas inverser les vitesses, puisque ce qui nous intéresse est l'estimation du volume délimité par les trois interfaces.

On modélise les données synthétiques par un tracé de rayons sur le modèle exact (Figure 4.23) avec une géométrie d'acquisition décrite par la Figure 4.24.

On ajoute à ces temps de trajet exacts un bruit blanc (décrit par une loi de probabilité Gaussienne centrée) d'écart type $3ms$ (Figure 4.27).

L'inversion de ces données bruitées grâce à notre code de tomographie de réflexion *jerry* (voir chapitre 3 pour une description de l'outil de tomographie et la méthodologie d'utilisation) donne un modèle solution proche du modèle exact (Figure 4.28) avec un résidu moyen RMS de $4.7ms$ et un résidu maximum de $69.6ms$ pour l'interface int1, un résidu moyen RMS de $12.8ms$ et un résidu maximum de

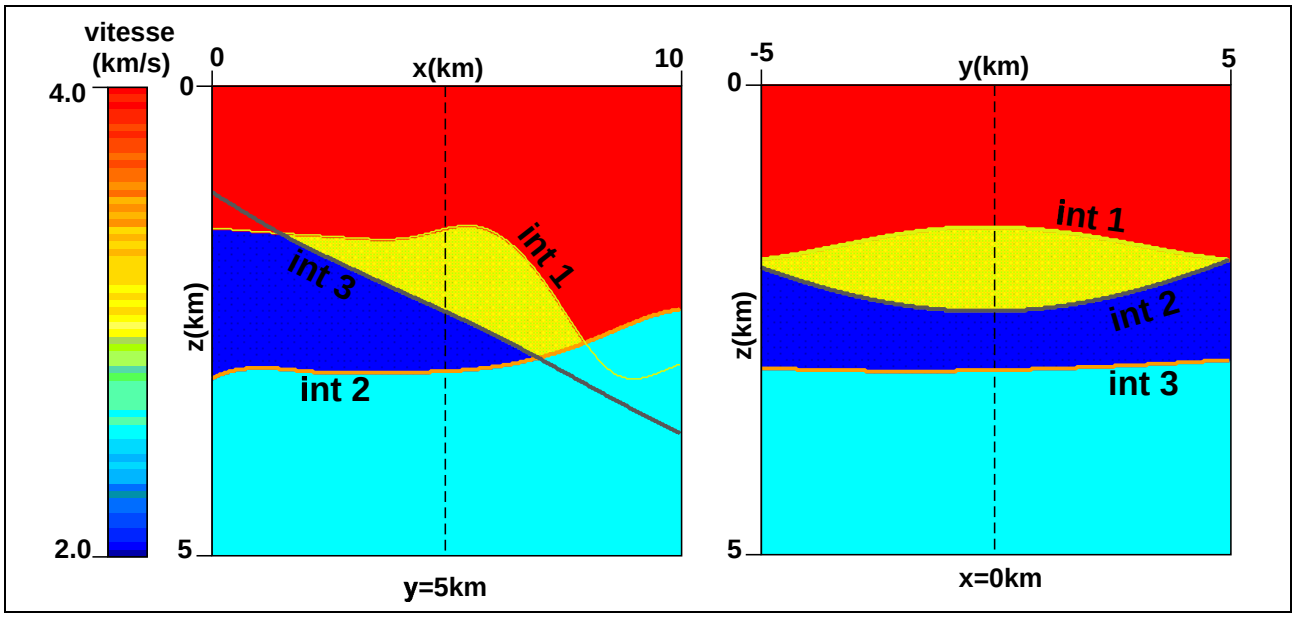

FIG. 4.23 – Coupes verticales du modèle de vitesse 3D exact (Gauche : le long de l'axe des x, droite : le long de l'axe des y).

$393ms$ pour l'interface int2 et un résidu moyen RMS de $4.15ms$ et un résidu maximum de $48.15ms$ pour l'interface int3. Ces résultats sont en accord avec le bruit introduit sur les données si on ne prend pas en compte les points d'impact des rayons au niveau des deux biseaux qui correspondent à des résidus élevés (voir Figures 4.25 et 4.26). Ces résidus élevés proches des biseaux sont liés à des problèmes de signature : la signature du rayon change de signe d'un côté à l'autre du biseau. Cavalca et Lailly, 2004, pour pallier à cette difficulté, ont proposé d'introduire des contraintes sur les points d'impact des rayons.

Ici nous illustrons l'approche macro-paramètre sur l'exemple de la Figure 4.23. Nous choisissons le macro-paramètre moyenne de la profondeur d'une interface et le macro-paramètre volume que nous venons d'expliciter dans le paragraphe 4.3.2.3. Les incertitudes sur ces macro-paramètres sont calculées sur le modèle solution de la Figure 4.28. Les résultats sont récapitulés dans le tableau 4.3.

Pour le premier macro-paramètre considéré (moyenne de la profondeur d'une interface), on observe des incertitudes faibles : cela s'explique par le fait que l'on considère des macro-paramètres associés à des moyennes donc les incertitudes fortes aux bords sont gommées. D'autre part, nous pouvons noter pour le macro-paramètre volume une incertitude de $25m^3$ sur le volume du piège délimité par int1, int2 et int3 qui est de $3.14km^3$.

4.3.2.4 Conclusions

Cette méthode avec son formalisme général permet de calculer les incertitudes sur des quantités géologiques intéressantes avec un coût de calcul raisonnable (voir tableau 4.4). Les simulations Gaussiennes (voir la section précédente) de macro-paramètres (utilisant la matrice de covariance a posteriori réduite (4.14)) peuvent être aussi réalisées : cela fournit une information intéressante sur les corrélations d'incertitudes des macro-paramètres et a un coût de calcul inférieur à celui des simulations sur les paramètres du modèle avec la matrice de covariance a posteriori complète.

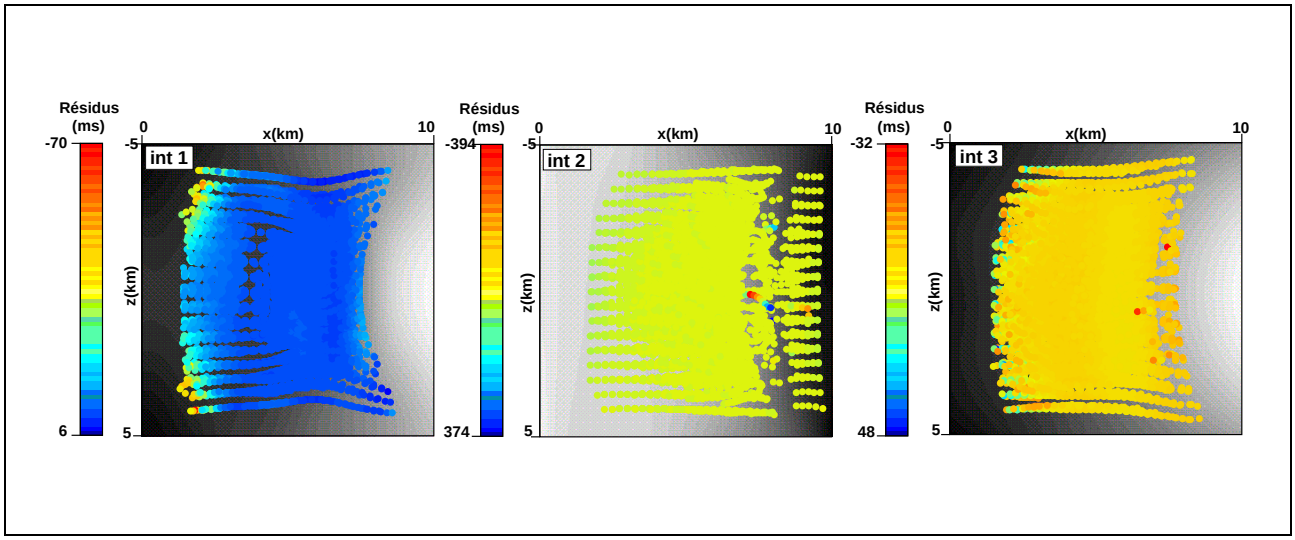

FIG. 4.24 – Illumination du domaine. Les points représentent les points d'impact des rayons sur l'interface ($x_min = 0km$ et $x_{max} = 10km$, $y_{min} = -5km$ et $y_{max} = 5km$, $z_{min} = 0km$ et $z_{max} = 5km$), les trois interfaces int1, int2 et int3 sont représentées.

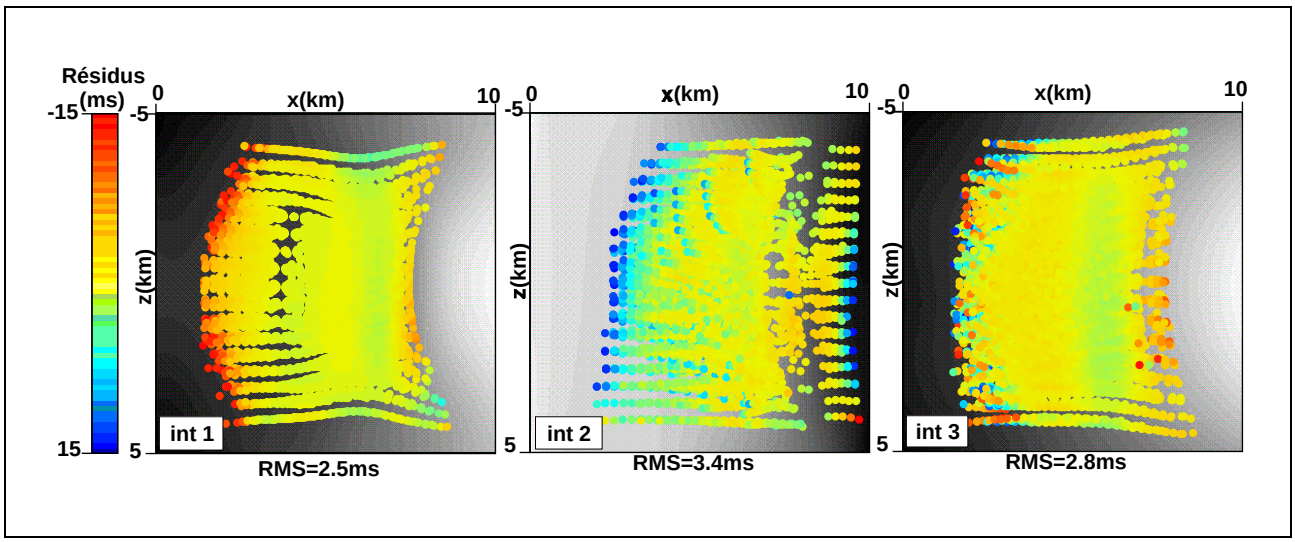

FIG. 4.25 – Illumination du domaine pour les résidus en accord avec le bruit introduit sur les données. Les points représentent les points d'impact des rayons sur l'interface ($x_{min} = 0km$ et $x_{max} = 10km$, $y_{min} = -5km$ et $y_{max} = 5km$, $z_{min} = 0km$ et $z_{max} = 5km$), les trois interfaces int1, int2 et int3 sont représentées.

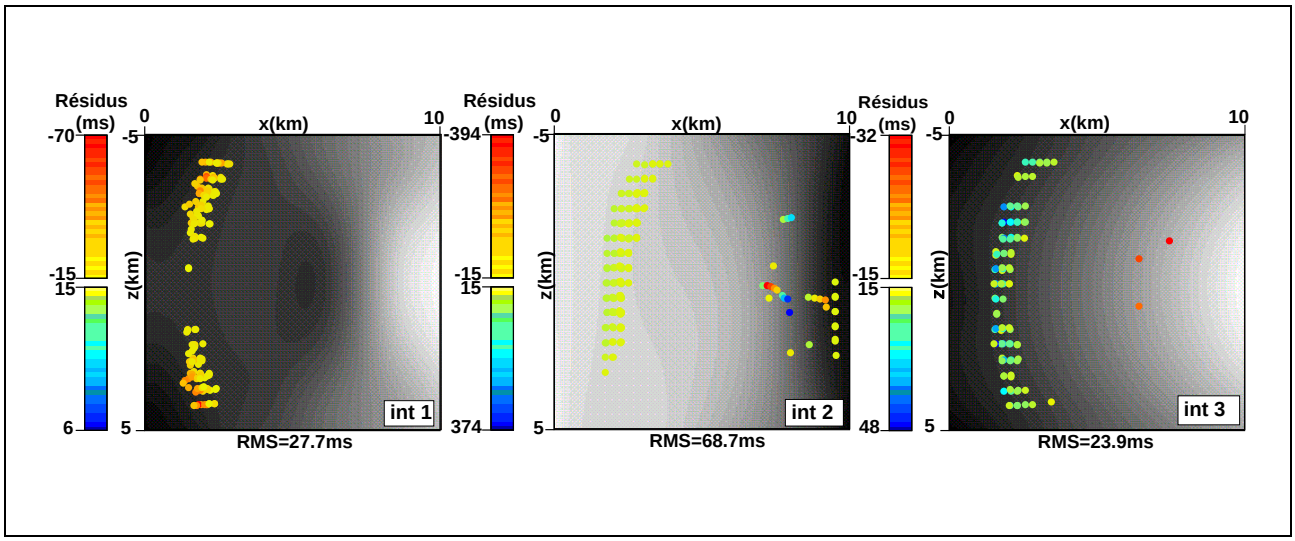

FIG. 4.26 – Illumination du domaine pour les résidus élevés. Les points représentent les points d'impact des rayons sur l'interface ($x_{min} = 0km$ et $x_{max} = 10km$, $y_{min} = -5km$ et $y_{max} = 5km$, $z_{min} = 0km$ et $z_{max} = 5km$), les trois interfaces int1, int2 et int3 sont représentées.

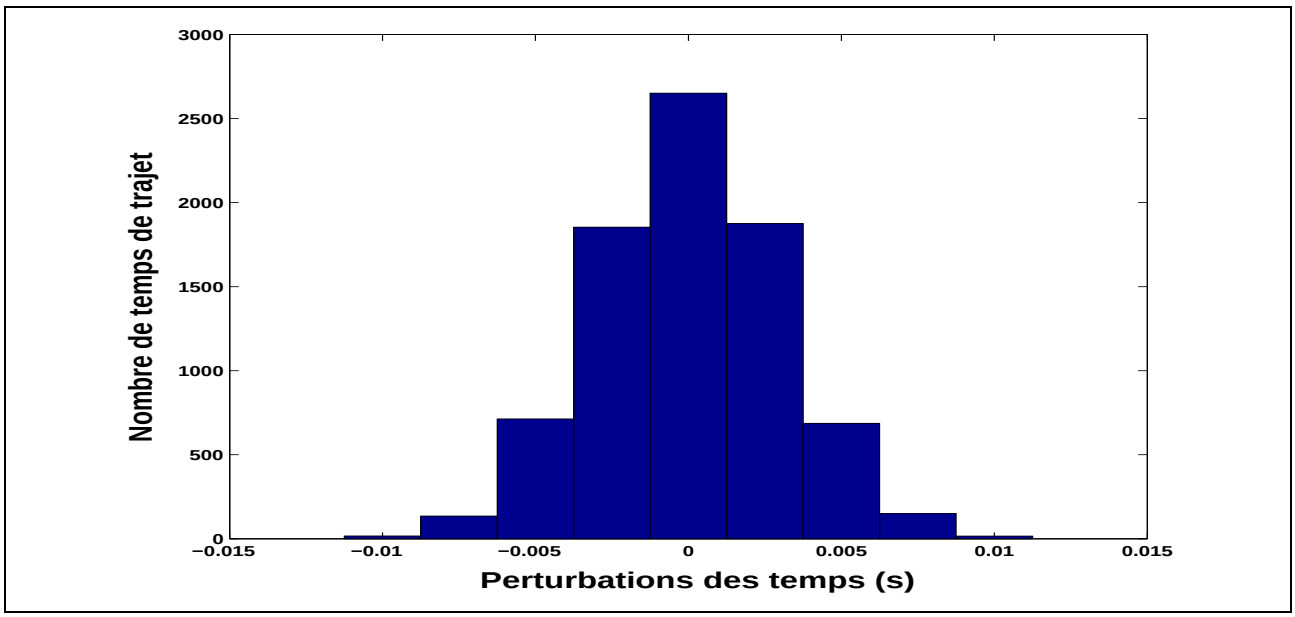

FIG. 4.27 – Histogramme des perturbations Gaussiennes non corrélées ajoutées aux données de temps de trajet de la Figure 4.24. L'écart type associé est de $3ms$.

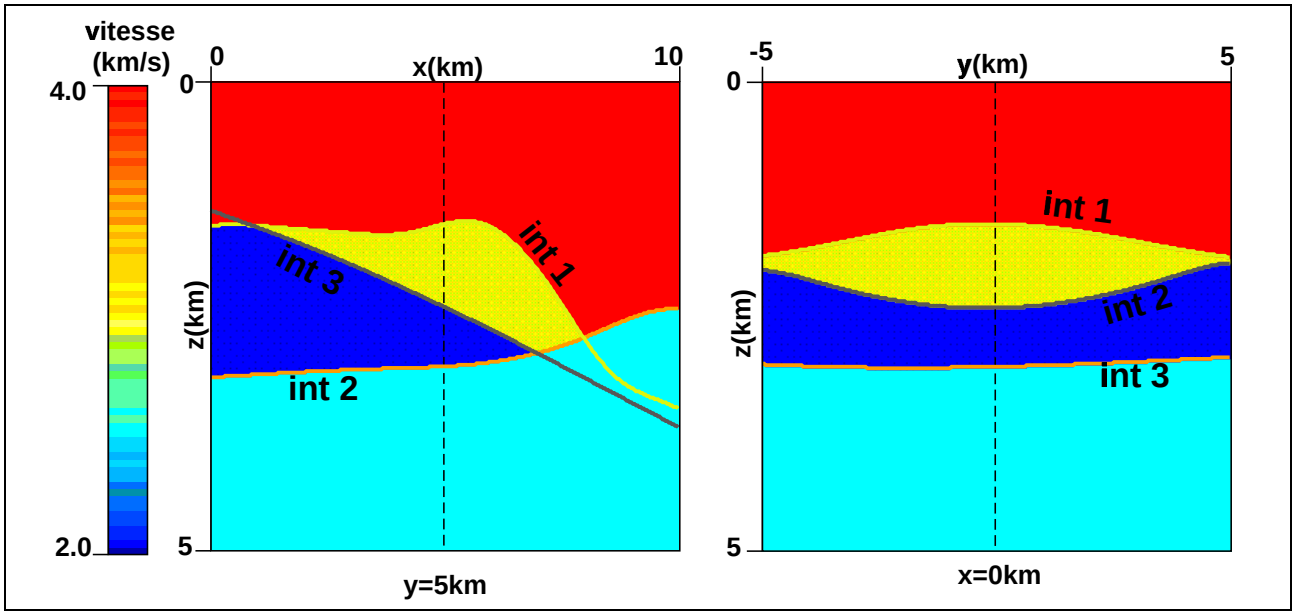

FIG. 4.28 – Coupes verticales du modèle solution obtenu par *jerry* pour le premier exemple (Gauche : le long de l'axe des x, droite : le long de l'axe des y). RMS des résidus temps de trajet $= 4.7ms$ et valeur maximum des résidus temps de trajet $= 69.6ms$ pour l'interface int1, RMS des résidus temps de trajet $=12.8ms$ et valeur maximum des résidus temps de trajet $= 393ms$ pour l'interface int2 et RMS des résidus temps de trajet $= 4.15ms$ et valeur maximum des résidus temps de trajet $= 48.15ms$ pour l'interface int3.

	Macro-paramètres	
	B-spline	**Volume**
INT1	1.612m	-
INT2	0.175m	-
INT3	0.419m	-
VOLUME	-	$25m^3$

TAB. 4.3 – Écarts types (racines carrées des termes diagonaux de $\tilde{C}'_m$) associées à différents macro-paramètres : moyenne de la fonction B-spline et volume du piège. Cette analyse de l'incertitude est réalisée pour le modèle solution de la Figure 4.28.

	Temps CPU
1 macro-Paramètre	28 min 4 s
50 simulations	3 h 3 min 50 s 52

TAB. 4.4 – Tableaux des temps CPU

4.3.3 Premières discussions sur ces deux méthodes

Ce que l'on peut déjà retenir à ce niveau, est que l'approche macro-paramètre fournit une information intéressante en ce qui concerne les corrélations d'incertitude des différents macro-paramètres et a un coût de calcul plus faible que les simulations de paramètres du modèle qui nécessite l'utilisation de la matrice de covariance a posteriori complète.

En effet, la méthode de simulation est plus coûteuse en temps CPU et en stockage (tableau 4.4) en raison de la décomposition de Choleski de la matrice de covariance a posteriori. Par exemple, pour le second modèle (Figure B.1) qui comporte 6869 paramètres, la méthode des simulations est 1.4 fois plus lente que la méthode des macro-paramètres (pour 50 simulations et 5 macro-paramètres).

D'autre part, les simulations fournissent directement des résultats interprétables, des modèles physiques.

Enfin, l'approche macro-paramètre peut fournir des incertitudes sur des quantités géologiques : la généralité du formalisme permet une large flexibilité en ce qui concerne la construction des macro-paramètres. Notons, également, que l'approche macro-paramètre permet d'avoir accès à une information intéressante concernant les corrélations entre les incertitudes associées aux divers macro-paramètres considérés : elle met en évidence facilement, les forts couplages entre les macro-paramètres choisis.

Chapitre 5

LIMITES DES MÉTHODES DE CALCUL D'INCERTITUDE : RÉALISME DES INCERTITUDES CALCULÉES

Ce chapitre est consacré à une étude concernant les limites des méthodes de calcul d'incertitudes décrites dans le chapitre précédent. Les différents thèmes abordés sont l'influence sur les incertitudes calculées de la paramétrisation et de la discrétisation du modèle, l'influence de l'échantillonnage des données, l'influence de l'introduction de régularisation et d'information a priori. Chaque cas est illustré par une application sur les exemples synthétiques décrits dans les Annexes A et B.

5.1 Influence de la paramétrisation et de la discrétisation du modèle sur les incertitudes calculées

5.1.1 Importance du choix de paramétrisation du modèle de Terre sur les incertitudes calculées

La paramétrisation de modèle choisie est un modèle décrit par blocs, dit *«blocky»* (voir chapitre 3, Clarke et Jannaud, 1996). Cette représentation permet d'introduire des discontinuités de vitesse aux interfaces. Cependant, cette paramétrisation du modèle peut être une source d'erreur de modélisation : le fait d'autoriser des discontinuités de vitesse au niveau de telle ou telle interface et de contraindre les variations de vitesse entre les interfaces à être faibles est un choix qui est loin d'être neutre et qui biaise fortement les données modélisées.

D'autre part, imposer le nombre de couches contraint fortement le modèle et peut engendrer des incertitudes artificiellement faibles. Nous illustrons ce point par l'application suivante : application de l'approche macro-paramètre (voir Chapitre 4, section 4.3.2) sur l'exemple de l'Annexe A. Nous choisissons des macro-paramètres simples : la moyenne de la vitesse dans la couche, la moyenne de ses dérivées, la moyenne de la profondeur d'une interface et de ses dérivées. Les incertitudes sur ces macro-paramètres sont calculées sur le modèle solution de la Figure A.4 et sur la solution de la Figure 5.1, modèle obtenu avec une paramétrisation de la vitesse par une B-spline 3D $v(x, y, z)$ à la place d'une B-spline 2D $v(x, y)$ (où l'on suppose aucune variation dans la direction z). Les résultats sont

récapitulés dans le tableau 5.1.

Pour l'inversion avec une paramétrisation de la vitesse en $v(x, y, z)$, nous notons la mauvaise détermination de la vitesse et tout particulièrement de ses variations verticales (incertitude de $0.17/s$ sur le gradient vertical (deux fois l'écart type)). Ce résultat n'est pas surprenant : les rayons se propagent près de la direction verticale, les temps de trajet ne sont donc pas très sensibles aux variations verticales de vitesse (uniquement sensibles à sa moyenne verticale). Par contre, les incertitudes obtenues sur le modèle obtenu avec une paramétrisation $v(x, y)$ sont beaucoup plus faibles : $3.74m/s$ au lieu de $126m/s$.

Ce résultat met bien en évidence le choix d'une paramétrisation adéquate : le choix d'une représentation de la vitesse $v(x, y)$ est une information a priori qui réduit énormément les incertitudes sur le modèle solution mais est-ce une information fiable ? Les incertitudes pour la paramétrisation $v(x, y, z)$ montrent qu'il existe d'autres solutions peut-être tout aussi réalistes.

Paramétrisation de la vitesse utilisée dans l'inversion		**Macro-paramètres**			
		B-spline	**DX**	**DY**	**DZ**
$v(x, y)$	VITESSE	3.74m/s	0.001/s	0.001/s	–
$v(x, y, z)$		126m/s	0.002/s	0.002/s	0.085/s
$v(x, y)$	INTERFACE	0.76m	0.0003	0.0003	–
$v(x, y, z)$		0.74m	0.0002	0.0003	–

TAB. 5.1 – Écarts types (racines carrées des termes diagonaux de $\tilde{C}'_m$) associés à différents macro-paramètres : moyenne de la fonction B-spline, de ses dérivées en x, y et z sur le domaine de définition. Cette analyse de l'incertitude est réalisée pour le modèle solution de la Figure A.4 obtenu avec la paramétrisation exacte $v(x, y)$ et pour le modèle solution de la Figure 5.1 obtenu avec une paramétrisation de la vitesse en $v(x, y, z)$.

5.1.2 Influence de la discrétisation sur l'analyse d'incertitude

5.1.2.1 Choix du nombre de paramètres B-splines

Comme nous l'avons vu dans la section 3.1.2 du chapitre 3, les vitesses et les interfaces sont paramétrées par des fonctions B-splines 2D ou 3D. Deux questions se posent : Comment choisir le nombre de paramètres B-splines ? Quelle est l'influence de ce choix sur les incertitudes sur le modèle solution ?

Classiquement l'analyse de la solution du problème inverse consiste à comparer les écarts entre les temps de trajet observés et les temps de trajet calculés en calculant la valeur du résidu moyen RMS (*Root Mean Square*) et la valeur du résidu maximum. Des résidus trop élevés peuvent être dus au fait que la paramétrisation B-spline choisie ne permet pas de modéliser les données complexes. En utilisant plus de paramètres on peut résoudre ce problème.

Dans le paragraphe suivant, nous allons montrer qu'à partir d'un certain pas de discrétisation suffisamment fin, la valeur de l'incertitude associée à un macro-paramètre tend à rester constante.

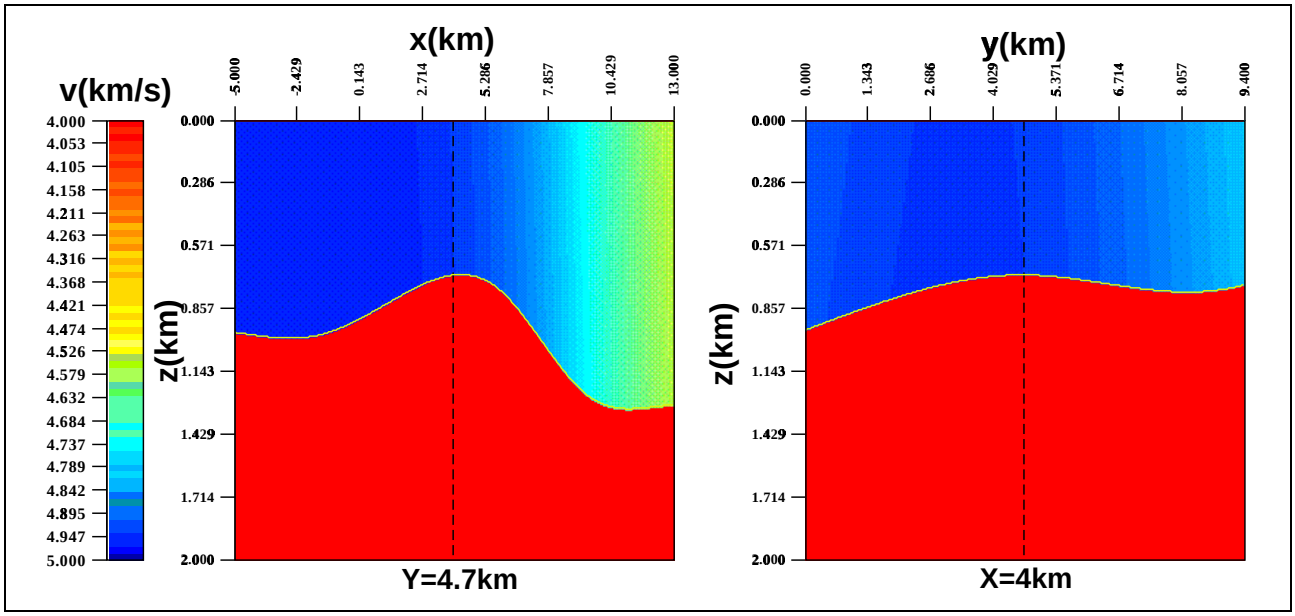

FIG. 5.1 – Coupes verticales du modèle solution obtenu par *jerry* avec une paramétrisation erronée de la vitesse : $v(x, y, z)$ à la place de $v(x, y)$ et une contrainte qui garde le gradient de vitesse vertical positif (Gauche : suivant l'axe x, droite : suivant l'axe y). RMS des résidus des temps de trajet = $5.5ms$, valeur maximale des résidus des temps de trajet = $22ms$.

5.1.2.2 Résultat théorique et démonstration

On considère un sous-espace de dimension finie $\mathcal{M}^h$ de $\mathcal{M}$, espace des modèles discrétisés (dans notre cas par des fonctions B-splines). h représente le pas de discrétisation. La norme considérée dans $\mathcal{M}^h$ est la norme H^2. On introduit l'hypothèse d'approximation suivante (Delprat-Jannaud et Lailly, 1992) :

$$\forall m \in \mathcal{M} \lim_{h \to 0} \inf_{m_h \in \mathcal{M}^h} \| m_h - m \|_{H^2} = 0, \tag{5.1}$$

qui est valide pour la plupart des procédures de discrétisation (en l'occurence pour une discrétisation par des B-splines). Cela signifie qu'en choisissant un h assez petit, on peut approcher avec la précision voulue (mesurée avec la norme $\| \|_{H^2}$) tout modèle m dans $\mathcal{M}$ par un modèle discret m_h dans $\mathcal{M}^h$.

Sous cette hypothèse, Delprat-Jannaud et Lailly, 1991, ont démontré la convergence au sens de la norme H^2 de la solution du problème discret vers la solution unique du problème continu.

Ce que nous voulons montrer ici c'est que l'incertitude associée à un macro-paramètre (voir paragraphe 4.3.2.1 du chapitre 4) converge à partir d'une certaine discrétisation, c'est à dire que la valeur de l'incertitude pour un macro-paramètre est indépendante de la discrétisation à partir d'un pas de discrétisation suffisamment fin.

Dans notre formalisme des macro-paramètres (voir chapitre 4 section 4.3.2), la matrice B, représente la matrice de passage des paramètres aux macro-paramètres. C'est à dire qu'elle vérifie la relation $P = Bm$ où m représente les paramètres du modèle et P le macro-paramètre. En d'autres termes et dans l'approche qui nous intéresse, B permet de réaliser une combinaison linéaire de paramètres afin

d'obtenir ce que nous appelons un macro-paramètre.

Pour mieux comprendre ce qui caractérise B, considérons le cas où B est l'opérateur **«moyenne de la fonction B-spline»** (pour la profondeur de l'interface et la vitesse). Notons que dans tout ce qui va suivre nous allons considérer ce macro-paramètre pour démontrer le résultat qui nous préoccupe mais que pour les autres macro-paramètres décrits dans la thèse le même principe s'applique.

Notons $a = Bu$ et $a_h = Bu_h$ où h représente le pas de discrétisation. Pour le macro-paramètre qui nous intéresse, B est défini par :

$$\begin{aligned} B : u \in H^2(\Omega) \subset L^2(\Omega) \quad &\longrightarrow \quad a \in \mathbb{R} \\ & \qquad a = \frac{1}{\text{mes}(\Omega_x \times \Omega_y)} \iint_{\Omega_x \times \Omega_y} u(x,y)dxdy, \end{aligned} \tag{5.2}$$

avec $\Omega = \Omega_x \times \Omega_y$ borné ($\Omega_x = [x_{min}; x_{max}]$ et $\Omega_y = [y_{min}; y_{max}]$) et $\text{mes}(\Omega_x \times \Omega_y) = (x_{max} - x_{min}) \times (y_{max} - y_{min})$.

De plus, $u(x,y) \in \mathcal{C}^0(\Omega)$, car d'après Brezis, 1987, on sait que que pour un ouvert $\Omega \subset \mathbb{R}^n$ on a $H^m(\Omega) \subset \mathcal{C}^k(\bar{\Omega})$ avec $0 \leq k < m - \frac{n}{2}$, sachant que dans le cas qui nous intéresse, $n = 2$ et $m = 2$.

Voyons maintenant les propriétés mathématiques de B pour ce macro-paramètre :

(1) $B : H^2(\Omega) \longrightarrow \mathbb{R}$ est un opérateur linéaire continu : en d'autres termes et dans le cas présent :

$$\mid a \mid = \mid Bu \mid = \mid \frac{1}{\text{mes}(\Omega_x \times \Omega_y)} \iint_{\Omega_x \times \Omega_y} u(x,y)dxdy \mid \leq K \parallel u \parallel_{L^2} \leq K \parallel u \parallel_{H^2}, \tag{5.3}$$

avec $K = \frac{1}{\sqrt{\text{mes}(\Omega_x \times \Omega_y)}}$.

(2) On définit B^T l'adjoint de B :

$$\begin{aligned} B^T : a \in \text{Im}_B(H^2(\Omega)) \subset \mathbb{R} \quad &\longrightarrow \quad u \in H^2(\Omega) \\ & \qquad u(x,y) = a, \forall x, y \in \Omega_x \times \Omega_y. \end{aligned} \tag{5.4}$$

Cet opérateur est aussi borné. En effet, comme :

$$\forall a \in \mathbb{R},\ \exists u \in H^2(\Omega)\ B^T a = u \text{ tel que } u(x,y) = a\ \forall x, y \in \Omega_x \times \Omega_y, \tag{5.5}$$

on a :

$$\parallel B^T a \parallel_{H^2(\Omega)} \leq \parallel u \parallel_{H^2(\Omega)} = \left(\iint_{\Omega_x \times \Omega_y} a^2 dxdy \right)^{\frac{1}{2}} = K' \mid a \mid, \tag{5.6}$$

avec $K' = \sqrt{\text{mes}(\Omega_x \times \Omega_y)}$.

On va maintenant utiliser ces propriétés de B pour démontrer le résultat qui nous intéresse :

$$\lim_{h \to 0} \left\langle BH^{-1}B^T a_h, a_h \right\rangle_{\mathbb{R}} - \left\langle BH^{-1}B^T a, a \right\rangle_{\mathbb{R}} = 0. \tag{5.7}$$

Preuve

Afin d'y parvenir nous allons majorer la quantité $\left\langle BH^{-1}B^T a_h, a_h \right\rangle_{\mathbb{R}} - \left\langle BH^{-1}B^T a, a \right\rangle_{\mathbb{R}}$ par une quantité dont nous serons capables de déterminer la limite lorsque le pas de discrétisation tendra vers 0.

On a :

$$\begin{aligned} \left\langle BH^{-1}B^T a_h, a_h \right\rangle_{\mathbb{R}} - \left\langle BH^{-1}B^T a, a \right\rangle_{\mathbb{R}} &= \left\langle BH^{-1}B^T u_h, Bu_h \right\rangle_{H^2} - \left\langle BH^{-1}B^T u, Bu \right\rangle_{H^2} \\ &= \left\langle A(u_h - u), u_h \right\rangle_{H^2} + \left\langle Au, u_h \right\rangle_{H^2} \\ &- \left\langle A(u - u_h), u \right\rangle_{H^2} - \left\langle Au_h, u \right\rangle_{H^2}, \end{aligned} \tag{5.8}$$

avec $A = B^T BH^{-1}B^T B$. Mais comme $A^T = A$, on peut avoir :

$$\left\langle Au, u_h \right\rangle_{H^2} = \left\langle u, A^T u_h \right\rangle_{H^2} = \left\langle A^T u_h, u \right\rangle_{H^2}. \tag{5.9}$$

Donc :

$$\begin{aligned} \left\langle BH^{-1}B^T a_h, a_h \right\rangle_{\mathbb{R}} - \left\langle BH^{-1}B^T a, a \right\rangle_{\mathbb{R}} &= \left\langle A(u_h - u), u_h \right\rangle_{H^2} - \left\langle A(u - u_h), u \right\rangle_{H^2}, \\ &= \left\langle A(u_h - u), u_h + u \right\rangle_{H^2} \end{aligned} \tag{5.10}$$

D'où :

$$| \left\langle BH^{-1}B^T a_h, a_h \right\rangle_{\mathbb{R}} - \left\langle BH^{-1}B^T a, a \right\rangle_{\mathbb{R}} | \leq \| A(u_h - u) \|_{H^2}^{\frac{1}{2}} \| u_h + u \|_{H^2}^{\frac{1}{2}} . \tag{5.11}$$

grâce à des opérations classiques sur le produit scalaire dans H^2 (Dautray et Lions, 1984).

À ce stade, nous allons majorer l'expression ci-dessus en utilisant le fait que $A = B^T BH^{-1}B^T B$ est un opérateur borné de L^2 dans H^2* †

On a alors :

$$\begin{aligned} | \left\langle BH^{-1}B^T a_h, a_h \right\rangle_{\mathbb{R}} - \left\langle BH^{-1}B^T a, a \right\rangle_{\mathbb{R}} | &\leq \| A \| \| u_h - u \|_{H^2}^{\frac{1}{2}} \| u_h + u \|_{H^2}^{\frac{1}{2}} \\ &\leq K 2a \sqrt{\text{mes}(\Omega x \times \Omega y)} \| u_h - u \|_{H^2}^{\frac{1}{2}} \end{aligned} \tag{5.12}$$

où K est une constante strictement positive.

*$\| A \| \leq \| H^{-1} \|_{L^2, H^2} \| B \|_{L^2, L^2}^4$.

†$H : H^2(\Omega) \longrightarrow L^2(\Omega)$ est continu et coercif sous l'hypothèse de complétude des données, cette hypothèse sera toujours vérifiée si le nombre de temps de trajet est suffisant (voir Annexe F). Par conséquent, comme tout opérateur linéaire continu et coercif sur un espace de Hilbert, H, est inversible, et $H^{-1} : L^2(\Omega) \longrightarrow H^2(\Omega)$ existe et est continu (H^{-1} est borné pour la norme H^2).

On peut maintenant passer à la limite dans (5.12) en faisant tendre le pas de discrétisation vers 0 et en utilisant l'hypothèse d'approximation (5.1) :

$$\forall u \in H^2 \lim_{h \to 0} \inf_{u_h \in H^2} \| u_h - u \|_{H^2} = 0, \tag{5.13}$$

On a alors bien démontré le résultat qui nous intéresse :

$$\lim_{h \to 0} \left\langle BH^{-1}B^T a_h, a_h \right\rangle_{\mathbb{R}} = \left\langle BH^{-1}B^T a, a \right\rangle_{\mathbb{R}}, \tag{5.14}$$

où h est le pas de discrétisation.

C'est à dire que les incertitudes sur les macro-paramètres convergent quand le pas de discrétisation tend vers 0 vers les valeurs obtenus pour le problème continu.

5.1.2.3 Application : résultats numériques

Les tests numériques dont les résultats sont récapitulés dans le tableau 5.2 illustrent le résultat théorique démontré dans la section 5.1.2.2.

L'expérience menée ici consiste à prendre une paramétrisation erronée de notre modèle et de réaliser l'analyse d'incertitude dans ce cas là et ensuite pour différentes paramétrisations emboîtées. On constate que la valeur de l'incertitude pour les différents macro-paramètres converge à partir d'une certaine discrétisation, c'est à dire que la valeur de l'incertitude pour les différents macro-paramètres est indépendante de la discrétisation à partir d'un certain pas de discrétisation.

D' autre part, on observe que ce formalisme des macro-paramètres nous permet d'obtenir un majorant de l'incertitude. En pratique on prend une discrétisation grossière, on se place donc dans une situation «au pire» et on obtient donc un majorant de l'incertitude (dans le cas bien sûr où B est un opérateur borné). Ceci se vérifie sur le tableau 5.2 et par :

$$0 \leq \| a_h - a \| \leq \| Bu_h - Bu \| \leq \| B \| \| \delta u \|, \tag{5.15}$$

avec $\delta u = u_h - u$.

Il existe une perturbation $\delta\tilde{u}$ maximale (c'est à dire pour un pas de discrétisation grossier), telle que $\| \delta u \| \leq \| \delta\tilde{u} \|$ et par conséquent on obtient :

$$\| B\delta u \| \leq \| B \| \| \delta u \| \leq \| B \| \| \delta\tilde{u} \| . \tag{5.16}$$

On observe ce même résultat de convergence dans le cas de macro-paramètres grilles, c'est à dire dans le cas où le macro-paramètre correspond à la valeur de la vitesse ou de la profondeur de l'interface en un point de la grille. Les résultats sont représentés sur la figure 5.2.

Paramétrisation de la vitesse utilisée dans l'inversion			**Macro-paramètres**			
Paramètres $(x \times y)$	Δ_x, Δ_y		**paramètres**	**B-spline**	**DX**	**DY**
6×5	$\Delta_x = 3.13km/s$, $\Delta_y = 4.7km/s$		64.44m/s	4.76m/s	0.0015/s	0.0016/s
9×5	$\Delta_x = 1.57km/s$, $\Delta_y = 4.7km/s$	V	35.65m/s	3.8m/s	0.0012/s	0.0013/s
15×5	$\Delta_x = 784m/s$, $\Delta_y = 4.7km/s$	I	15.45m/s	3.5m/s	0.0011/s	0.0018/s
27×5	$\Delta_x = 392m/s$, $\Delta_y = 4.7km/s$	T	10.3m/s	3.3m/s	0.0005/s	0.0013/s
51×5	$\Delta_x = 196m/s$, $\Delta_y = 4.7km/s$	E	9.6m/s	3.2m/s	0.0003/s	0.0011/s
99×5	$\Delta_x = 98m/s$, $\Delta_y = 4.7km/s$	S	9.4m/s	3.2m/s	0.0002/s	0.0011/s
195×5	$\Delta_x = 49m/s$, $\Delta_y = 4.7km/s$	S	9.4m/s	3.2m/s	0.0002/s	0.0011/s
387×5	$\Delta_x = 24.5m/s$, $\Delta_y = 4.7km/s$	E	9.4m/s	3.2m/s	0.0002/s	0.0011/s
771×5	$\Delta_x = 12.25m/s$, $\Delta_y = 4.7km/s$		9.4m/s	3.2m/s	0.0002/s	0.0011/s
8×5	$\Delta_x = 3.6km$, $\Delta_y = 9km$	I	23.1m	0.98m	0.00032	0.00035
10×5	$\Delta_x = 2.57km$, $\Delta_y = 9km$	N	19.9m	0.91m	0.0003	0.00033
17×5	$\Delta_x = 1.285km$, $\Delta_y = 9km$	T	8.7m	0.84m	0.00028	0.0003
31×5	$\Delta_x = 642.5m$, $\Delta_y = 9km$	E	3.9m	0.81m	0.00027	0.0029
59×5	$\Delta_x = 321.25m$, $\Delta_y = 9km$	R	3.2m	0.8m	0.00027	0.00029
115×5	$\Delta_x = 160.63m$, $\Delta_y = 9km$	F	3.05m	0.8m	0.0026	0.00028
227×5	$\Delta_x = 80.31m$, $\Delta_y = 9km$	A	3.02m	0.8m	0.0026	0.0028
451×5	$\Delta_x = 40.16m$, $\Delta_y = 9km$	C	3.01m	0.8m	0.0026	0.0028
899×5	$\Delta_x = 20.08m$, $\Delta_y = 9km$	E	3.01m	0.8m	0.0026	0.0028

TAB. 5.2 – Écarts types (racines carrées des termes diagonaux de $\tilde{C}'_m$) associés à différents macro-paramètres : moyenne des paramètres, moyenne de la fonction B-spline, de ses dérivées en x et en y sur le domaine de définition (voir 4.3.2.1). Cette analyse de l'incertitude est réalisée pour le modèle solution de la Figure A.4 obtenu avec la paramétrisation exacte et différentes paramétrisations emboîtées. La première ligne pour la vitesse et la première ligne pour l'interface correspondent à la valeur de l'écart type pour la paramétrisation exacte. Δ_x et Δ_y représentent respectivement le pas de discrétisation en x et le pas de discrétisation en y.

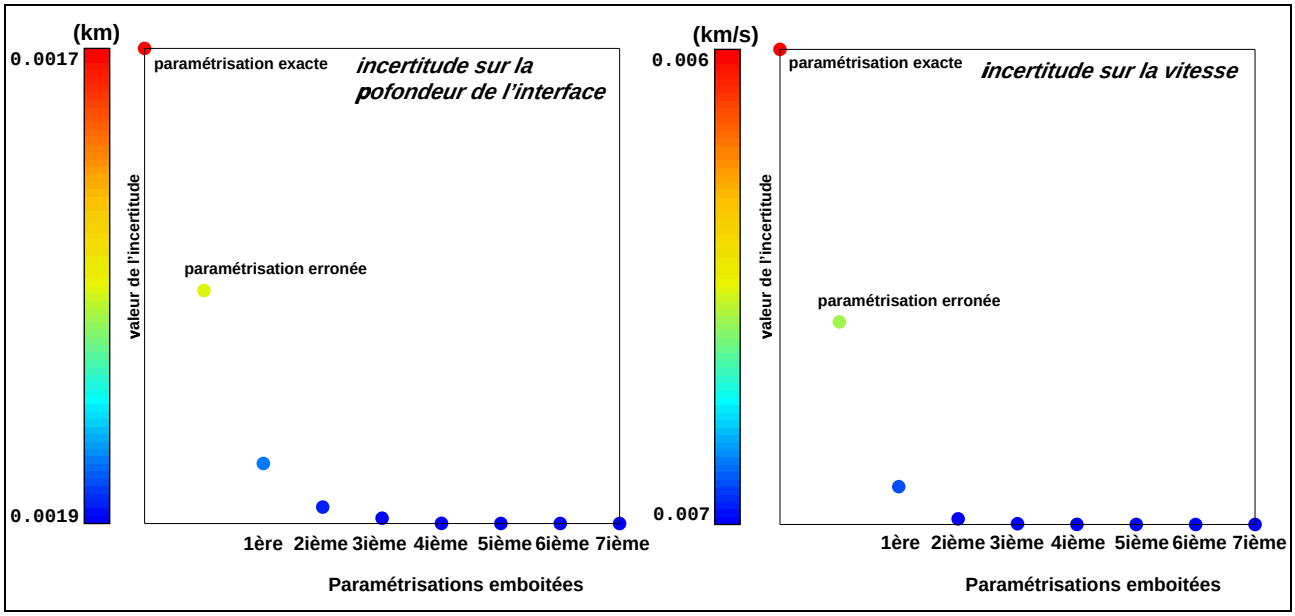

FIG. 5.2 – Écarts types (racines carrées des termes diagonaux de $\tilde{C}'_m$) associés aux macro-paramètres grilles. À gauche : valeur de la profondeur de l'interface en un point de la grille. À droite : valeur de la vitesse en un point de la grille. Cette analyse de l'incertitude est réalisée pour le modèle solution de la Figure A.4 obtenu avec la paramétrisation exacte et différentes paramétrisations emboîtées.

5.2 Influence de l'échantillonnage des données

5.2.1 Influence du choix de la norme dans l'espace des données

Renard, 2002, a montré dans ses travaux de thèse le manque de réalisme du choix classique de la norme dans l'espace des données. En effet, il souligne que le fait de choisir une norme L^2 dans l'espace des données (formulation classique de la tomographie de réflexion), revient à considérer les erreurs sur les données du problème inverse comme spatialement indépendantes voir la signification de la matrice de covariance dans la section 4.2 du chapitre 4 ou dans l'Annexe G), ce qui semble peu réaliste dès lors que des données sont associées à des couples source-récepteur proches en surface. C'est pour remédier à ce manque de réalisme qu'il a proposé un formalisme permettant de prendre en compte explicitement notre connaissance a priori des corrélations spatiales du bruit sur les données et pour obtenir des résultats indépendants de l'échantillonage des données.

L'idée qui a été la plus simple à mettre en oeuvre pour prendre en compte explicitement la nature corrélée du bruit entachant l'information cinématique en tomographie de réflexion est de définir le problème inverse continu en s'appuyant sur la norme H^1 dans l'espace des données. Ainsi, il inverse non seulement les temps de trajet observés mais aussi les pentes de ces temps dans les trois directions de l'espace (les coordonnées CMP (*Common Middle Point*) et l'offset), la confiance accordée à chacune de ces données étant quantifiée par des longueurs de corrélation agissant dans chacune des directions d'espace.

Les expériences numériques ont montré que les résultats de la tomographie de réflexion avec cette nouvelle norme sont similaires à ceux obtenus avec la norme L^2. François Renard explique ce résultat

par le fait que, sous certaines hypothèses très fréquemment vérifiées en tomographie de réflexion (notamment le caractère basse fréquence des résidus), les normes L^2 et H^1 sont sensiblement équivalentes.

Que l'on utilise une norme L^2 ou H^1 dans l'espace des données, les modèles solutions étant très proches, on peut cependant se demander si les incertitudes sur ces deux solutions voisines sont également similaires.

Une étude succinte des incertitudes associées aux solutions obtenues à l'aide des deux normes nous montre que les valeurs de ces incertitudes sont aussi très proches. En effet, considérons les deux solutions issues de la tomographie $\tilde{m}_{L^2}$ et $\tilde{m}_{H^1}$ comme quasiment identiques au même modèle $\tilde{m}$.

On sait qu'en la solution le gradient de la fonction coût est nul. On peut donc écrire, pour tout modèle m peu éloigné de $\tilde{m}$:

$$C(m) - C(\tilde{m}) = \frac{1}{2}\,(H(m-\tilde{m}),(m-\tilde{m})) = \frac{1}{2} \parallel J(m-\tilde{m}) \parallel^2, \tag{5.17}$$

avec H la matrice Hessienne et J la Jacobienne des temps de trajet.

De plus, on définit $\xi_{L^2,\tilde{m},r}$ et $\xi_{H^1,\tilde{m},r}$ les boules ouvertes de centre $\tilde{m}$ et de rayon r pour les normes L^2 et H^1, respectivement, par :

$$\xi_{L^2,\tilde{m},r} = \{m \in \mathcal{M} / \frac{1}{2} \parallel J(m-\tilde{m}) \parallel^2_{L^2} \leq r\}, \tag{5.18}$$

et

$$\xi_{H^1,\tilde{m},r} = \{m \in \mathcal{M} / \frac{1}{2} \parallel J(m-\tilde{m}) \parallel^2_{H^1} \leq r\}. \tag{5.19}$$

Puisque les normes L^2 et H^1 sont sensiblement équivalentes, on peut facilement remarquer que $\forall m \in \mathcal{M}$, $\parallel J(m-\tilde{m}) \parallel^2_{L^2} \simeq \parallel J(m-\tilde{m}) \parallel^2_{H^1}$: ces deux boules ouvertes sont donc quasiment identiques. Les incertitudes associées aux solutions obtenues à l'aide des deux normes sont donc très proches.

Ici, nous allons essayer de voir si une étude plus approfondie sur la quantification des incertitudes a posteriori associées aux modèles solutions obtenus avec la norme H^1 et la norme L^2 confirme la tendance.

5.2.1.1 Résultats numériques

Les tests numériques dont les résultats sont récapitulés dans le tableau 5.3 illustrent l'équivalence des deux normes sur les incertitudes calculées.

On applique l'approche macro-paramètre (voir Chapitre 4, section 4.3.2) sur l'exemple A.4 (voir annexe A) pour illustrer le choix d'une norme L^2 ou H^1 dans l'espace des données pour le calcul d'incertitude.

Les macro-paramètres que nous choisissons sont les suivants : la moyenne de la vitesse dans la couche, la moyenne de ses dérivées, la moyenne de la profondeur d'une interface et de ses dérivées. Les incertitudes sur ces macro-paramètres sont calculées sur le modèle solution de la Figure A.4 et sur la solution de la Figure 5.3, modèle obtenu avec une norme H^1 dans l'espace des données. Les résultats sont récapitulés dans le tableau 5.3 et confirment les résultats de l'étude succinte : les normes H^1 et

L^2 sont donc bien sensiblement équivalentes, les incertitudes associées aux solutions obtenues à l'aide des deux normes sont très proches.

Norme dans l'espace des données utilisée dans l'inversion		**Macro-paramètres**		
		B-spline	**DX**	**DY**
Norme L^2	VITESSE	3.74m/s	0.001/s	0.001/s
Norme H^1		3.69m/s	0.001/s	0.002/s
Norme L^2	INTERFACE	0.74m	0.0002	0.0003
Norme H^1		0.77m	0.0001	0.0001

TAB. 5.3 – Écarts types (racines carrées des termes diagonaux de $\tilde{C}'_m$) associés à différents macro-paramètres : moyenne de la fonction B-spline, de ses dérivées en x et y sur le domaine de définition. Cette analyse de l'incertitude est réalisée pour le modèle solution de la Figure A.4 obtenu avec la norme L^2 dans l'espace des données et pour le modèle solution de la Figure 5.3 obtenu avec une norme H^1 dans l'espace des données.

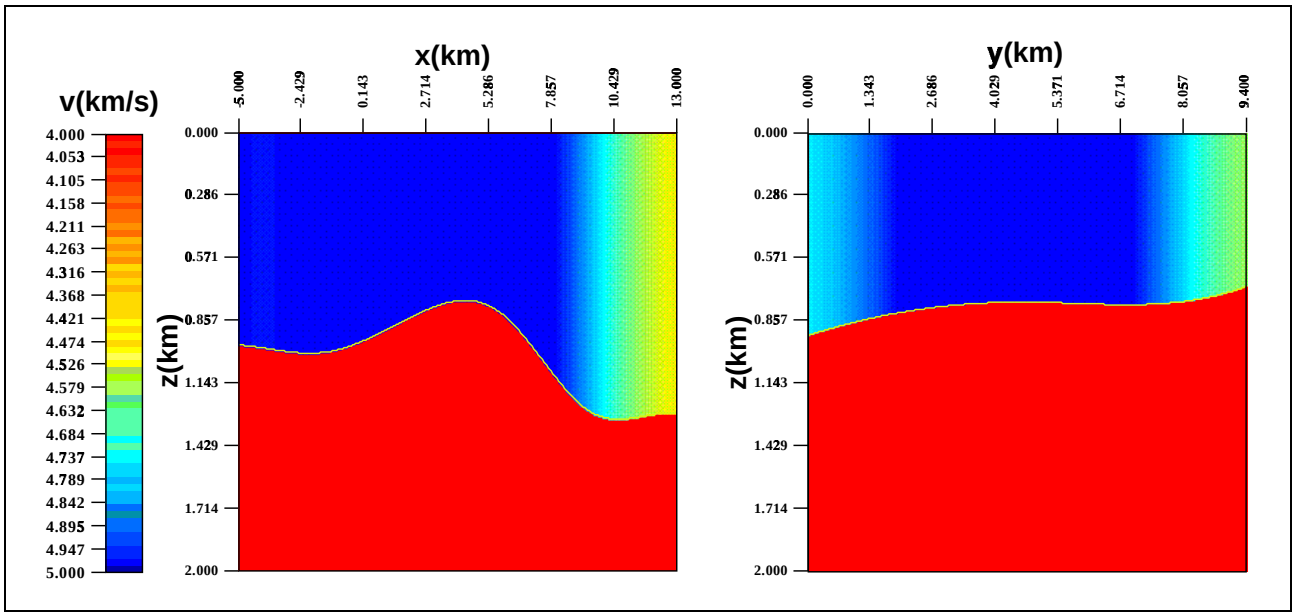

FIG. 5.3 – Coupes verticales du modèle solution obtenu par *jerry* avec la norme H^1 dans l'espace des données (Gauche : suivant l'axe x, droite : suivant l'axe y). RMS des résidus des temps de trajet = $5.08ms$, valeur maximum des résidus des temps de trajet = $25ms$.

5.2.2 Influence de l'échantillonnage des données

Nous venons donc de vérifier que le choix de la norme dans l'espace des données n'a pas d'incidence majeure sur les incertitudes calculées. Par conséquent, il n'est pas nécessaire d'introduire la norme H^1 dans l'espace des données pour déterminer une solution au problème de tomographie de réflexion plus satisfaisante que celle obtenue à l'aide d'une norme L^2. Nous allons maintenant essayer de voir dans

quelle mesure l'échantillonnage des données a une influence sur les incertitudes calculées, en utilisant la norme L^2 dans l'espace des données.

5.2.2.1 Résultats numériques

Les tests numériques dont les résultats sont récapitulés dans le tableau 5.4 illustrent l'influence de l'échantillonnage des données sur les incertitudes calculées.

On applique l'approche macro-paramètre (voir Chapitre 4, section 4.3.2) sur l'exemple A.4 (voir annexe A) pour illustrer l'importance du choix de l'échantillonnage des données pour le calcul d'incertitude.

Nous choisissons des macro-paramètres simples : la moyenne de la vitesse dans la couche, la moyenne de ses dérivées, la moyenne de la profondeur d'une interface et de ses dérivées. Les incertitudes sur ces macro-paramètres sont calculées sur le modèle solution de la Figure A.4 et sur la solution de la Figure 5.4, modèle obtenu avec une norme L^2 dans l'espace des données et un échantillonnage grossier (temps de trajets zéro-offset sur une grille $300m \times 300m$ et temps de trajet multi-offset pour des lignes espacées de $600m$ avec des CMP espacés de $300m$ et un pas d'offset de $100m$; on modélise ainsi 5969 temps de trajet.). Les résultats sont récapitulés dans le tableau 5.4 et montrent que dans le cas d'une norme L^2 dans l'espace des données, l'échantillonnage a très peu d'incidence sur les incertitudes calculées.

Echantillonnage dans l'espace des données utilisée dans l'inversion		**Macro-paramètres**		
		B-spline	**DX**	**DY**
Echantillonnage fin	VITESSE	3.74m/s	0.001/s	0.001/s
Echantillonnage grossier		3.83m/s	0.002/s	0.002/s
Echantillonnage fin	INTERFACE	0.74m	0.0002	0.0003
Echantillonnage grossier		0.78m	0.0004	0.0004

TAB. 5.4 – Écarts types (racines carrées des termes diagonaux de $\tilde{C}'_m$) associés à différents macro-paramètres : moyenne de la fonction B-spline, de ses dérivées en x et y sur le domaine de définition. Cette analyse de l'incertitude est réalisée pour le modèle solution de la Figure A.4 obtenu avec la norme L^2 dans l'espace des données pour un échantillonnage fin et pour le modèle solution de la Figure 5.4 obtenu avec une norme L^1 dans l'espace des données pour un échantillonnage grossier.

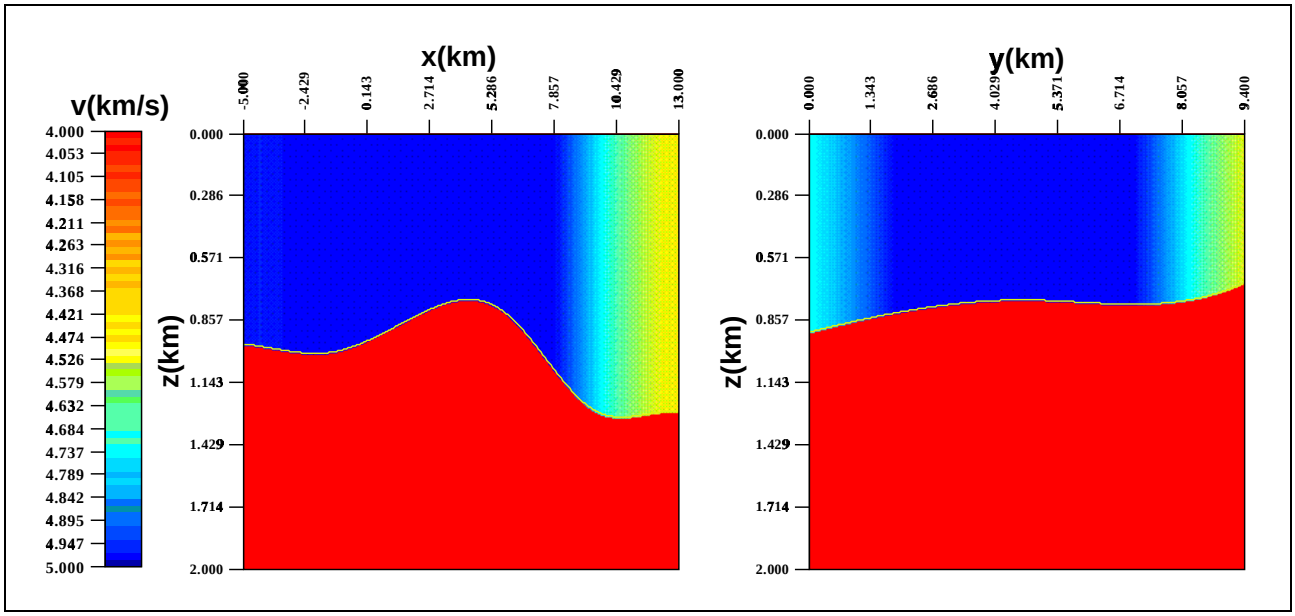

FIG. 5.4 – Coupes verticales du modèle solution obtenu par *jerry* avec la norme L^2 dans l'espace des données pour un échantillonnage des données grossier (Gauche : suivant l'axe x, droite : suivant l'axe y). RMS des résidus des temps de trajet $= 5.02ms$, valeur maximum des résidus des temps de trajet $= 23.6ms$.

5.3 Influence de l'introduction de régularisation et d'information a priori

5.3.1 Influence de l'introduction de régularisation

Pour illustrer l'influence des poids de régularisation sur les incertitudes, on a choisi ici d'étudier trois modèles (voir figure 5.5) : ils ont des résidus en accord avec le bruit introduit sur les données. Ces trois modèles correspondent à des modèles solutions obtenus avec des poids de régularisation différents ($\epsilon^2 = 1000$, $\epsilon^2 = 100$ et $\epsilon^2 = 10$).

On réalise des simulations de modèle (20 modèles simulés) autour de ces trois modèles solutions (voir le chapitre 4 section 4.3.1 pour la méthodologie) et on obtient les résultats de la Figure 5.6 pour la première interface.

De manière générale, pour ces trois modèles, les plus grandes incertitudes sont localisées sur les bords du modèle, ce qui correspond aux zones qui ne sont pas bien illuminées et dans la partie centrale qui correspond à la zone non illuminée (voir Figure B.2).

D'autre part, on constate que pour les deux premiers modèles de la Figure 5.6 (qui correspondent respectivement aux poids $\epsilon^2 = 1000$ et $\epsilon^2 = 100$), l'information fournie par la régularisation est trop régularisante, surtout dans la partie centrale (c'est à dire la zone non illuminée) : les incertitudes sur la partie centrale de l'interface, zone non-illuminée par les rayons, sont faibles, ce qui n'est pas réaliste. Cette zone est uniquement déterminée par l'information a priori introduite, c'est à dire la régularisation.

Notons que la valeur des incertitudes associées au troisième exemple de la Figure 5.6 qui correspond à un poids de régularisation plus faible, $\epsilon^2 = 10$, est élevée dans ces mêmes zones. Cet exemple semble fournir des incertitudes plus réalistes dans la zone centrale. Toutefois si on prend en compte le critère de la L-curve, ce modèle ne se situe pas dans le coin de la courbe en «L» (voir Figure 5.5). Ce modèle est un modèle vérifiant les données mais pas acceptable car trop irrégulier.

Par conséquent, le critère de la L-curve (section 3.3.2.1 du chapitre 4) nous fournit une méthode stable pour obtenir un modèle (modèle 1 de la Figure 5.5). Mais les incertitudes associées à ce modèle solution ne sont pas réalistes : en effet, on obtient de faibles incertitudes dans la zone centrale qui est non illuminée (Figure 5.6) ce qui n'est pas réaliste.

Ainsi cet exemple montre la sensibilité des incertitudes sur le modèle solution aux poids de régularisation introduits dans le problème inverse, ce qui n'a rien d'étonnant. La méthode L-curve est une méthode stable pour obtenir le modèle le plus régulier vérifiant les données. Cependant, les incertitudes sur ce modèle solution ne sont par réalistes, l'information a priori introduite étant trop régularisante.

Pour pallier à cette difficulté, il faut pouvoir construire un modèle a priori plus fiable, ce qui est loin d'être facile en pratique. Nous illustrons dans le paragraphe suivant la construction d'une norme dans l'espace des modèles (section 3.3.2 du chapitre 3) qui nous permet une interprétation plus physique des poids de régularisation et donc de faciliter le choix de ces poids.

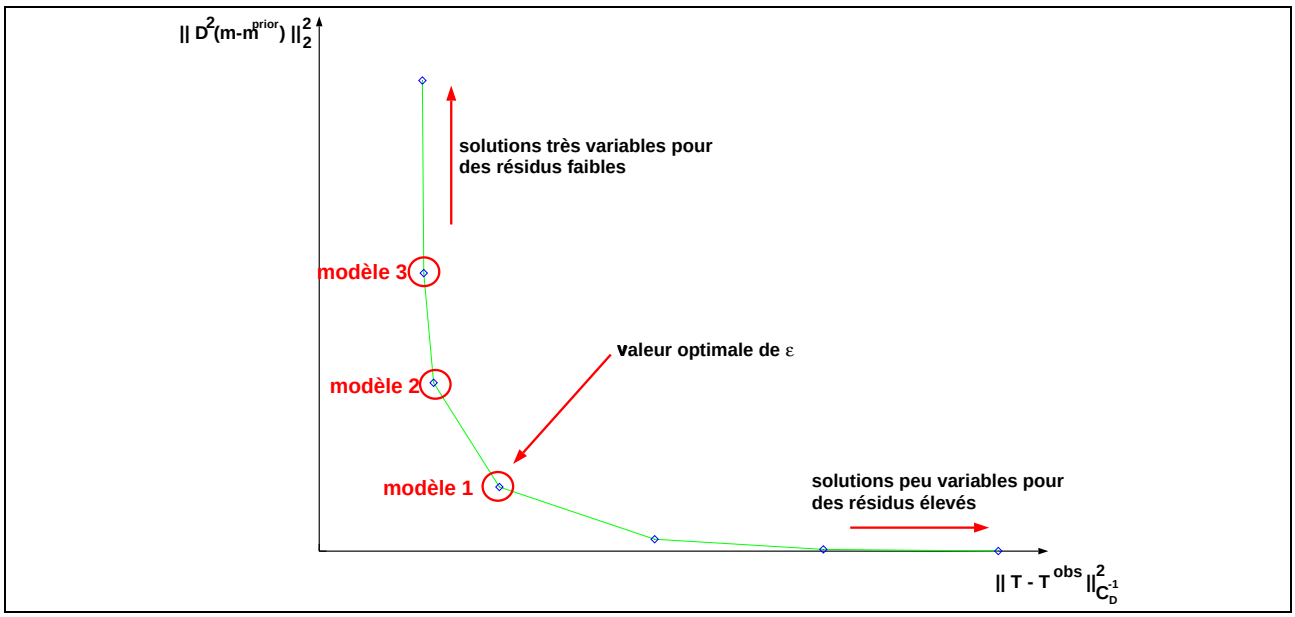

FIG. 5.5 – Représentation de la courbe en L caractéristique (courbe paramétrée en fonction de ϵ). Les trois modèles considérés pour mesurer l'influence des poids de régularisation sur les incertitudes sont placés sur la courbe.

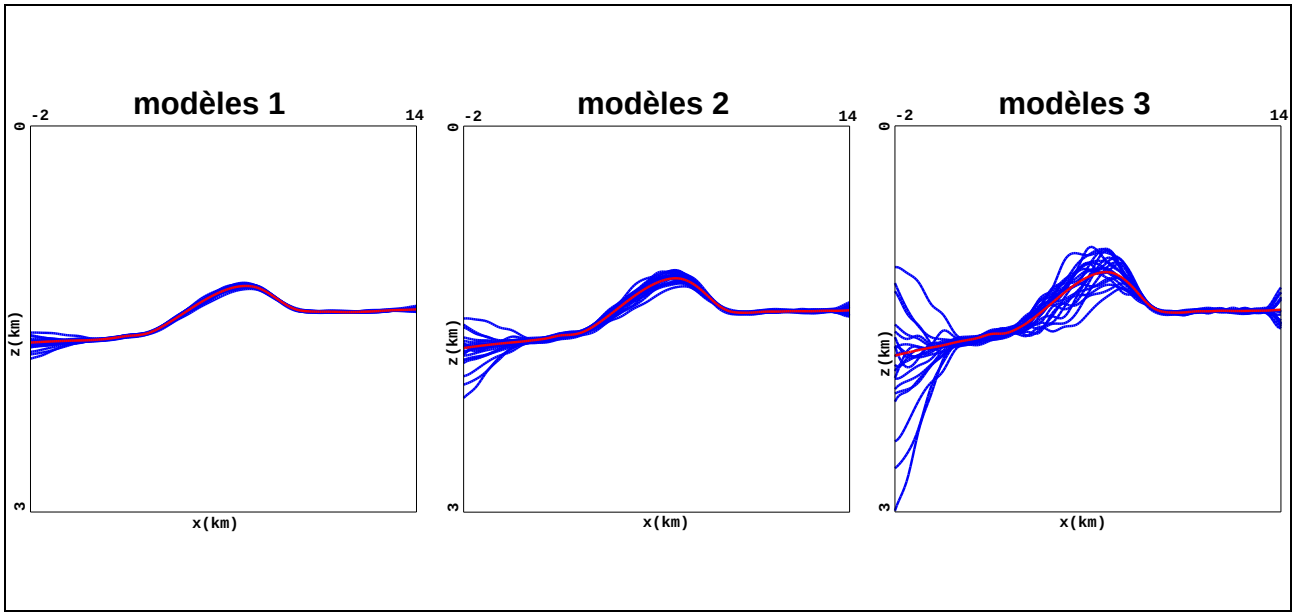

FIG. 5.6 – Coupes suivant l'axe des x sur les modèles de vitesse 3D simulés. De gauche à droite : $\epsilon^2 = 1000$, $\epsilon^2 = 100$ et $\epsilon^2 = 10$.

5.3.2 Influence de l'introduction d'information a priori (cas anisotrope)

Pour illustrer l'influence de l'introduction d'information a priori, on considère le modèle B.3 (voir Annexe B). On cherche à obtenir un modèle qui satisfait les données temps de trajet avec la précision voulue et l'information a priori. Si on considère 5.7 qui représente l'illumination du modèle, on veut ajouter de l'information a priori du type :

- Le modèle est régulier : aux bords et le long de l'axe y.

- On a peu (voire pas) d'informations a priori dans la zone non illuminée.

On modélise cette information a priori grâce à la formulation (3.24) introduite dans la section 3.3.2.2 du chapitre 3. Pour choisir les longueurs de corrélation λ et l' écart type σ qui vont nous permettre de définir les différents poids qui caractérisent le terme d'information a priori dans l'expression de la fonction coût (3.24), on réalise des simulations a priori en utilisant la méthode des simulations (se reporter à la section 4.3.1 du chapitre 4 pour un détail de la méthode), mais ici la matrice considérée est la matrice de covariance a priori, et on compare de manière visuelle les résultats obtenus. Dans notre étude, nous réaliserons l'analyse d'incertitude uniquement sur la première couche du modèle exact de la Figure B.1.

La Figure 5.8 représente 20 modèles simulés pour différentes longueurs de corrélation λ_x, λ_y et différents écarts types σ. Il suffit alors de choisir les familles de modèles qui correspondent le mieux à l'information a priori, par exemple, la deuxième famille ($\lambda = \lambda_x = \lambda_y = 8km$ et $\sigma = 300m$.) : la première famille qui correspond à $\lambda = \lambda_x = \lambda_y = 2km$ et $\sigma = 300m$ (on considère la profondeur de l'interface) est très oscillante et la troisième famille qui correspond à $\lambda = \lambda_x = \lambda_y = 10km$ et $\sigma = 300m$ est très lisse. On pourrait préciser un peu plus en mettant une longueur de corrélation plus faible dans la zone qui n'est pas illuminée (Figure 5.9) puisqu'a priori on ne sait rien sur cet exemple : ceci est un

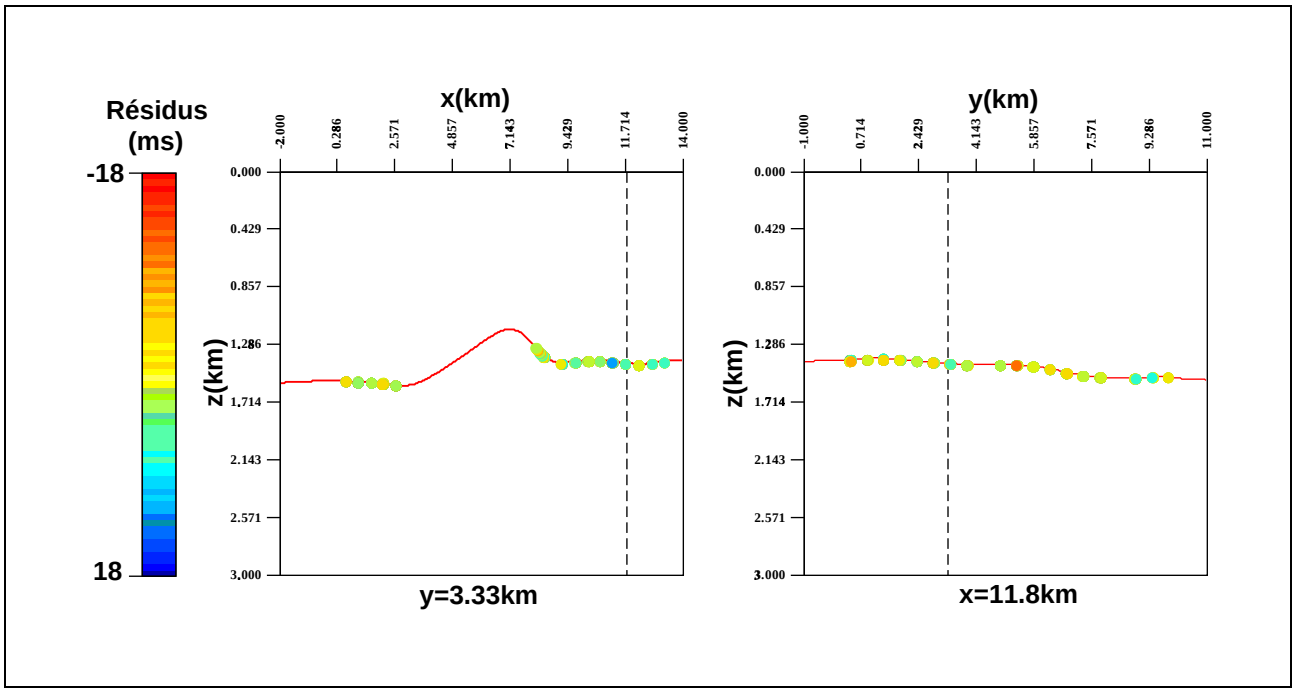

FIG. 5.7 – Illumination dans les deux directions principales d'anisotropie : coupes suivant l'axe des x (à gauche) et suivant l'axe des y (à droite) du modèle solution. Les points représentent les points d'impact des rayons sur l'interface.

exemple de prise en compte d'information a priori.

Dans tout ce qui va suivre notre a priori est associé à une longueur de corrélation de $\lambda = 2km$ (zone grisée sur la Figure 5.9) et partout ailleurs une longueur de corrélation de $\lambda = 8km$; l' écart type lui est de $\sigma = 300m$ sur tout le domaine.

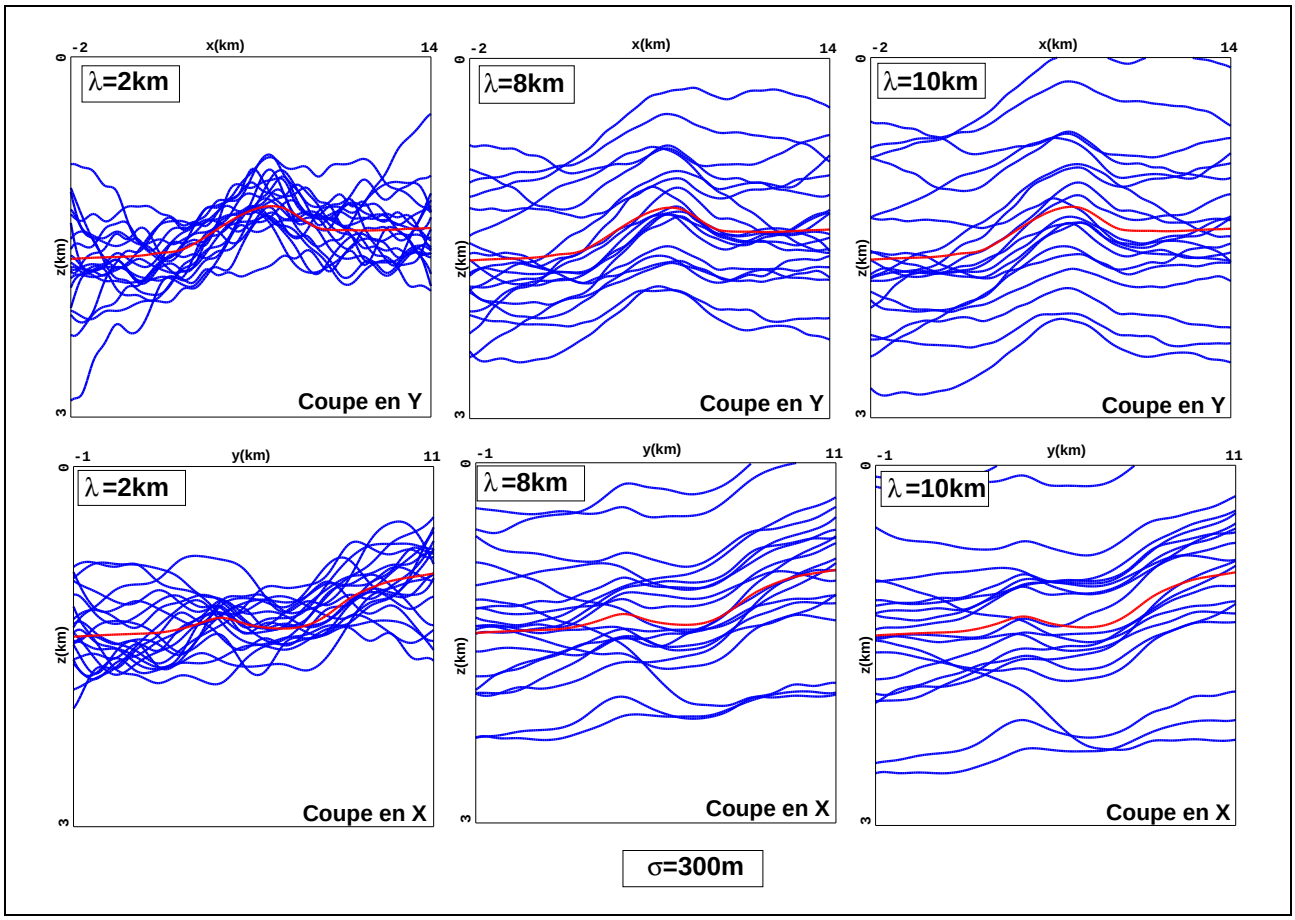

FIG. 5.8 – Coupes suivant l'axe des x (en haut) et suivant l'axe des y (en bas) sur les modèles de vitesse 3D simulés a priori (représentation de la profondeur de la première interface du modèle solution de la Figure B.3) dans le cas anisotrope. Les longueurs de corrélation λ pour les différentes familles sont : $\lambda = \lambda_x = \lambda_y = 2km$ pour la 1ère famille de modèles, $\lambda = \lambda_x = \lambda_y = 8km$ pour la deuxième famille de modèles, $\lambda = 10km$ pour la troisième famille de modèles. L'écart type est de $\sigma = 300m$.

Les résultats de l'inversion obtenus par la méthode de la L-curve (section 3.3.2.1 du chapitre 3) et par la méthode faisant intervenir les λ et σ, sont très similaires (à part peut-être légèrement moins dans la zone centrale) comme le montrent les Figures 5.10 et 5.11 (RMS moyen des résidus $\approx 5ms$ pour les deux modèles).

Les Figures 5.12 et 5.13 montrent 20 modèles simulés après inversion obtenus par la méthode décrite dans le chapitre 4 (section 4.3.1) pour accéder aux incertitudes sur le modèle solution selon les

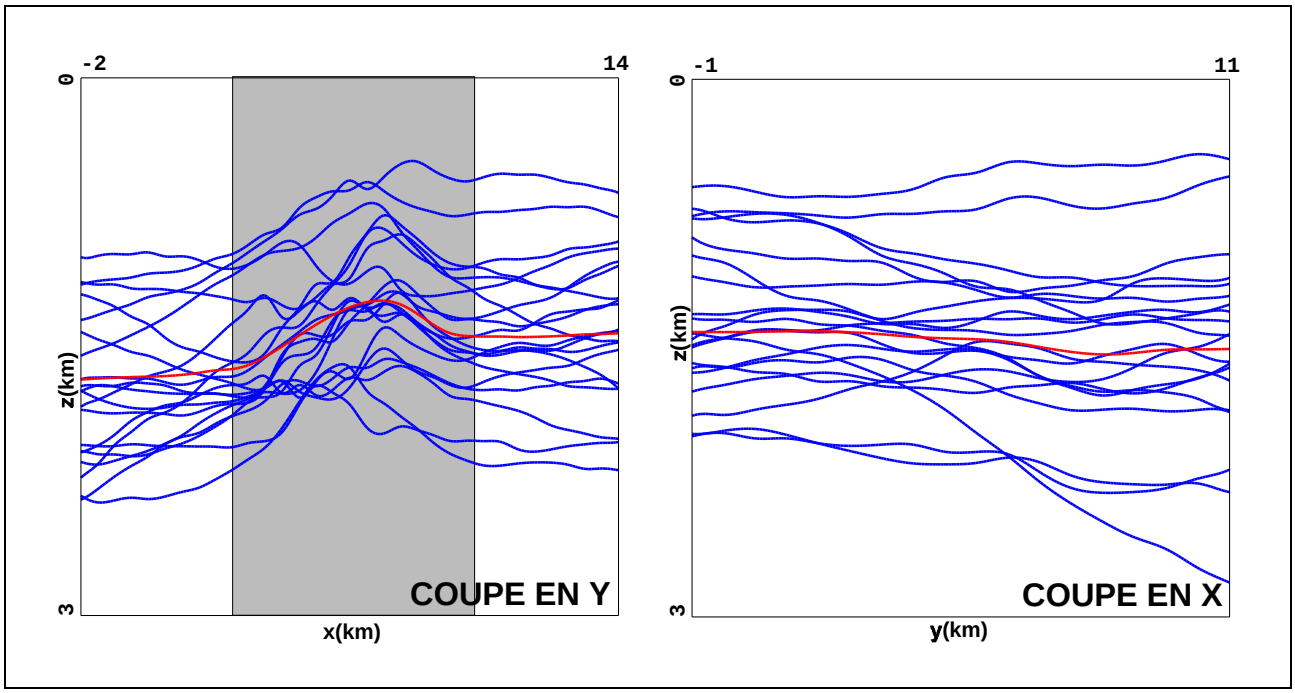

FIG. 5.9 – Coupes suivant l'axe des x (en haut) et suivant l'axe des y (en bas) sur les modèles de vitesse 3D simulés a priori (représentation de la profondeur de la première interface du modèle solution de la Figure B.3) dans le cas anisotrope. Les longueurs de corrélation λ choisies sont : $\lambda_x = 2km$ et $\lambda_y = 8km$ dans la zone grisée, $\lambda = \lambda_x = \lambda_y = 8km$ ailleurs. L'écart type est de $\sigma = 300m$.

deux directions d'anisotropie : coupes de l'interface et variations latérales de vitesse le long de l'axe des x et le long de l'axe des y.

Les incertitudes les plus grandes sur la profondeur de l'interface ($\approx 150m$) et sur les variations de vitesses ($\approx 60m/s$) sont localisées sur les bords du modèle, ce qui correspond aux zones qui ne sont pas bien illuminées. Dans la zone qui correspond à la partie grisée de la Figure 5.9, les incertitudes sur la profondeur de l'interface sont élevées et correspondent à $\approx 130m$, ce qui est une estimation réaliste des incertitudes.

Par contre, les résultats obtenus par la méthode de la L-curve (Figure 5.14) montrent des incertitudes faibles même dans la zone non illuminée, cette partie étant déterminée par la régularisation.

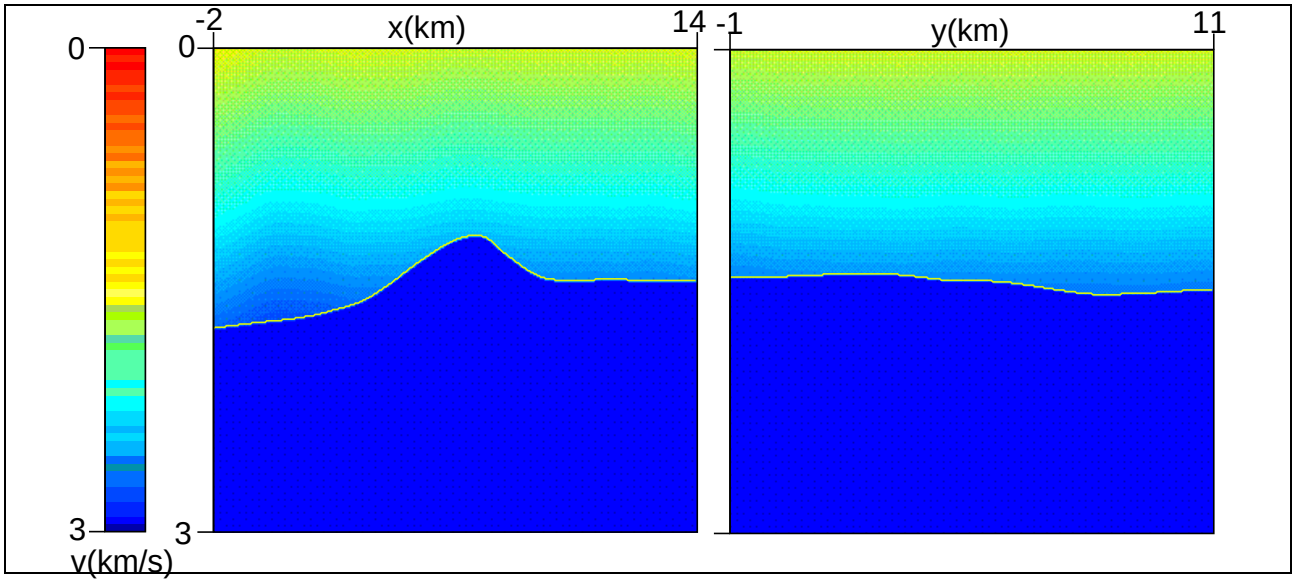

FIG. 5.10 – Coupes suivant l'axe des x et des y représentant le modèle solution associé à la première couche du modèle exact de la figure B.1 obtenu par une inversion avec introduction d'information a priori.

5.3.3 Conclusions

L'introduction d'informations a priori est cruciale pour réduire l'indétermination. Il est en pratique difficile de modéliser les informations géologiques disponibles et les incertitudes associées à ces informations. C'est pourquoi, en pratique, il est courant d'utiliser une méthode de continuation ou méthode L-curve qui consiste à rechercher le modèle le plus régulier vérifiant les données avec la précision voulue. Le problème inverse associé est un problème inverse régularisé où les poids de régularisation choisis forts en début d'inversion sont diminués au fur et à mesure jusqu'à l'obtention d'un modèle vérifiant les données. Cette méthode permet d'éviter les problèmes d'instabilité liés à une régularisation trop faible.

Cependant, nous avons montré que, bien que les modèles obtenus par cette méthode soient satisfaisants, les incertitudes associées peuvent ne pas être réalistes, les poids de régularisation obtenus par

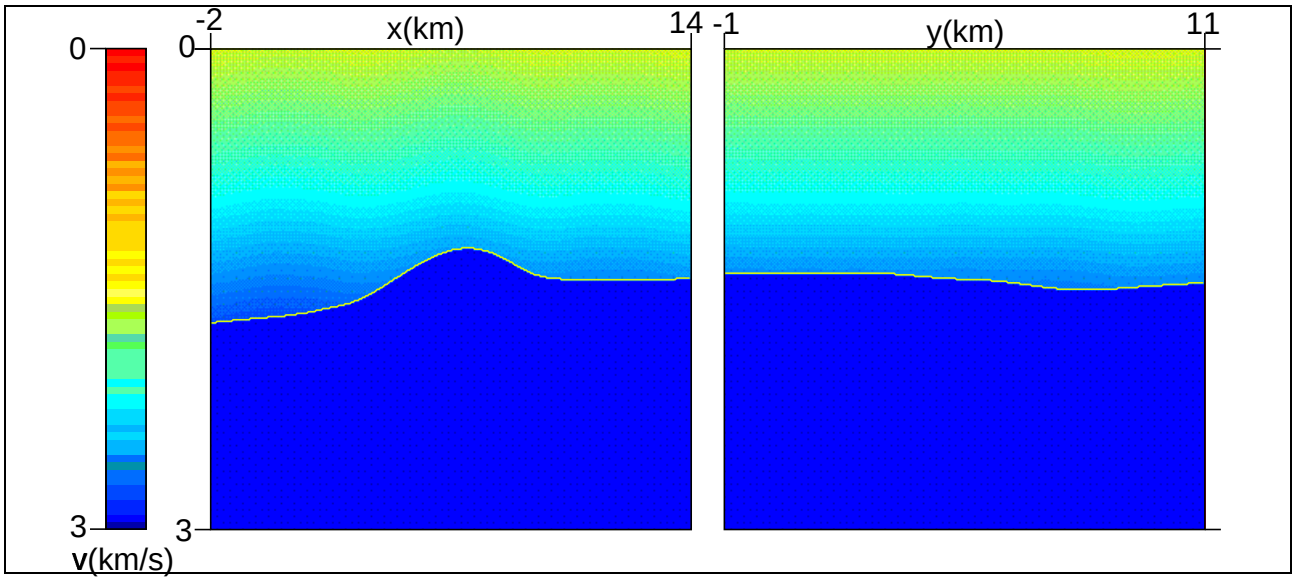

FIG. 5.11 – Coupes suivant l'axe des x et des y représentant le modèle solution associé à la première couche du modèle exact de la figure B.1 obtenu par une inversion utilisant la méthode de la l-curve.

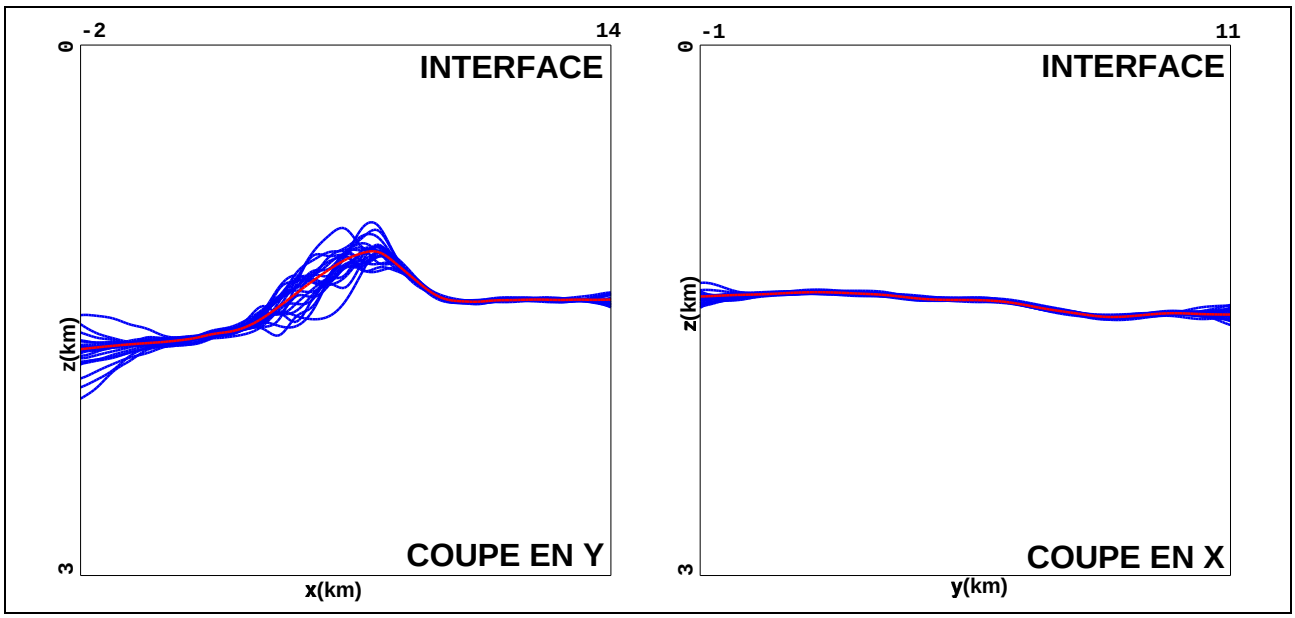

FIG. 5.12 – Coupes suivant l'axe des x (gauche) et suivant l'axe des y (droite) sur les modèles de vitesse 3D simulés. On a représenté les variations de la profondeur de la première interface du modèle solution de la Figure B.3.

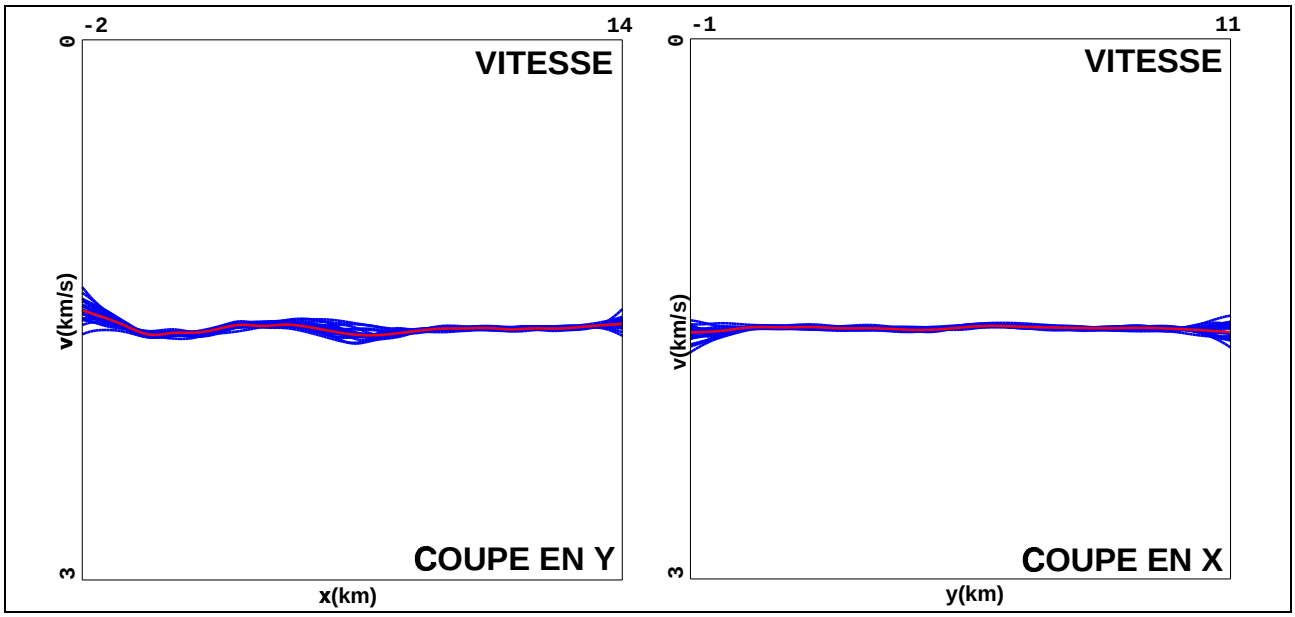

FIG. 5.13 – Coupes suivant l'axe des x (gauche) et suivant l'axe des y (droite) sur les modèles de vitesse 3D simulés. On a représenté les variations de vitesse du modèle solution de la Figure B.3.

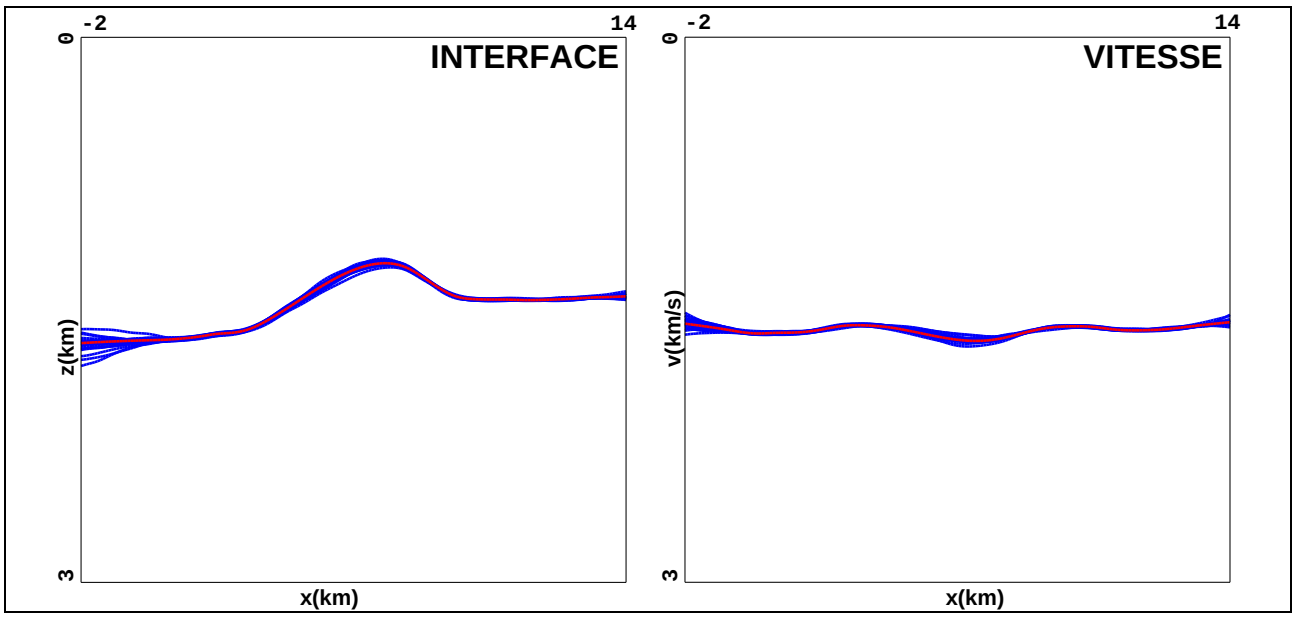

FIG. 5.14 – Coupes suivant l'axe des x sur les modèles de vitesse 3D simulés. Gauche : variations de la profondeur de la première interface du modèle solution de la Figure B.3. Droite : variations de vitesse associées à cette interface. Le modèle solution est obtenu dans ce cas en utilisant la méthode de la l-curve.

cette méthode ne correspondant pas à une information a apriori réaliste.

L'introduction d'une norme plus «géologique» avec des poids dépendant de quantités plus physiques (introduction d'une covariance exponentielle dans l'espace des modèles) permet l'obtention d'incertitudes plus réalistes. Cependant, dans le cas où nous disposons de très peu d'information a priori ou d'une information a priori peu fiable, les poids résultants peuvent conduire à une inversion instable. Dans ce cas, la méthode L-curve sera utilisée pour obtenir une solution et pour l'estimation des incertitudes sur ce modèle solution, on construira la norme «géologique» avec des poids plus conformes à l'information a priori disponible.

Chapitre 6

APPLICATIONS SUR DONNEES REELLES

6.1 Première application : HOD

Cette section comporte un article rédigé pour le congrès ECCOMAS 2004 qui a eu lieu en Finlande à Jßvaskßla du 24 au 28 juillet 2004. C'est pourquoi il est rédigé en anglais.

Cet article présente l'application sur données réelles des méthodes développées dans le chapitre 4 : la méthode des simulations de modèles et l'approche macro-paramètre. Ces deux méthodes sont ensuite comparées.

Nous avons vu dans les chapitres précédents que l'analyse d'incertitude linéarisée permet de délimiter l'ensemble des modèles solutions possibles qui vérifient, avec la précision espérée, les données temps de trajet et l'information a priori. Néanmoins cette approche est uniquement valide au voisinage du modèle solution et que les cas complexes peuvent nécessiter une approche non-linéaire. Cet article souligne ici l'utilité d'une approche non-linéaire.

L'approche non-linéaire consistera donc à étudier le comportement de la fonction coût et à en rechercher les minima locaux. Nous allons voir dans cet article qu'une approche expérimentale consiste à résoudre le problème inverse sous contraintes afin de tester la validité de certaines hypothèses géologiques.

European Congress on Computational Methods in Applied Sciences and Engineering
ECCOMAS 2004
P. Neittaanmäki, T. Rossi, K. Majava, and O. Pironneau (eds.)
H. W. Engl and F. Tröltzsch (assoc. eds.)
Jyväskylä, 24–28 July 2004

QUANTIFYING UNCERTAINTIES ON THE SOLUTION MODEL OF SEISMIC TOMOGRAPHY

Carole DUFFET⋆ and Delphine SINOQUET†

⋆Geophysics division
e-mail: carole.duffet@ifp.fr

† Technology, computer Science, and Applied Mathematics Division
1-4 avenue de Bois-Préau, 92500 Rueil-Malmaison, France
e-mail: delphine.sinoquet@ifp.fr

Key words: Tomography, uncertainties quantification, velocity model building

Abstract. *Reflection tomography allows the determination of a velocity model that fits the traveltime data associated with reflections of seismic waves in the subsurface. A least-square formulation is used to compare the observed traveltimes and the traveltimes computed by the forward operator based on a ray tracing. This non linear optimization problem is solved classicaly by a Gauss-Newton method based on successive linearizations of the forward operator. The obtained solution is only one among many possible models. Indeed, the uncertainties on the observed traveltimes (resulting from an interpretative event picking on seismic records) and more generally the underdetermination of the inverse problem lead to uncertainties on the solution. An a posteriori uncertainty analysis is then crucial to delimit the range of possible solutions that fit, with the expected accuracy, the data and the a priori information. A linearized a posteriori analysis is possible by an analysis of the a posteriori covariance matrix, inverse of the Gauss-Newton approximation of the matrix. The computation of this matrix is generally expensive (the matrix is huge for 3D problems) and the physical interpretation of the results is difficult. Then we propose a formalism which allows to compute uncertainties on relevant geological quantities for a reduced computational time. Nevertheless, this approach is only valid in the vicinity of the solution model (linearized framework) and complex cases may require a non linear approach. An application on a 2D real data set illustrates the linearized approach to quantify uncertainties on the solution of seismic tomography and the limitations of this approach are discussed.*

1 INTRODUCTION

Geophysical methods for imaging complex geological subsurface in petroleum exploration requires the determination of an accurate propagation velocity model. Seismic reflection tomography ([1]) turns out to be an efficient method for that: this method allows to determine a seismic velocity distribution from traveltimes data associated with the seismic waves reflecting on geological surfaces. This inverse problem is formulated as a least-square minimization problem which consists in the minimization of the mismatch between the observed traveltimes and the traveltimes computed by the forward problem (solved by a ray tracing method). The first section of this paper presents an overview of the reflection tomography method we have developed ([10], [11], [13], [9]).
Methods for the a posteriori analysis of the solution are then proposed and applied on a 2D real data set. Indeed, classically, the analysis of the solution consists only in checking the misfits between observed traveltimes and calculated traveltimes by computing characteristic values of the misfit distribution and by studying the spatial distribution of the misfits. This analysis is not sufficient: even if these quality control criteria are matched, the determined model is only one of many possible models that match the data. An uncertainty analysis should be performed to quantify the range of admissible models we can obtain from these data and the a priori information ([6]).
The different methods described here allow to access the uncertainties on the solution model thanks to the analysis of the a posteriori covariance matrix ([7], [14]) obtained in the linearized framework. The computation of this matrix is generally expensive for 3D problems and the physical interpretation of its terms may be cumbersome. The simulation of admissible models from the a posteriori probability density function allows a more complete analysis than the sole analysis of the diagonal terms of the a posteriori covariance matrix and furnishes directly physical models. This method requires a Cholesky decomposition of the Gauss-Newton matrix which may be unfeasible for 3D realistic models. To overcome this difficulty, we propose a formalism based on linear combinations of model parameters (macro-parameters) that allows to compute uncertainties on relevant geological quantities (the average thickness of a layer for example) for a reduced computational time (the a posteriori covariance matrix is reduced to the macro-parameter space). The application of these methods on a real data set shows their effectiveness and their limitations are discussed, particularly the limitations due to the linearization: a non linear approach is proposed in order to delimit the space of admissible solutions.

2 THE REFLECTION TOMOGRAPHY PROBLEM

Let us first present the reflection tomography method which was developed in KIM consortium ([9]). The chosen model representation is a *blocky* velocity model where the velocity distribution is described by slowly varying layer velocities (also called velocity blocks) delimited by interfaces (see Figure 1 for an example). The model is thus composed of two kinds of parameters: parameters describing the velocity variations within the layers

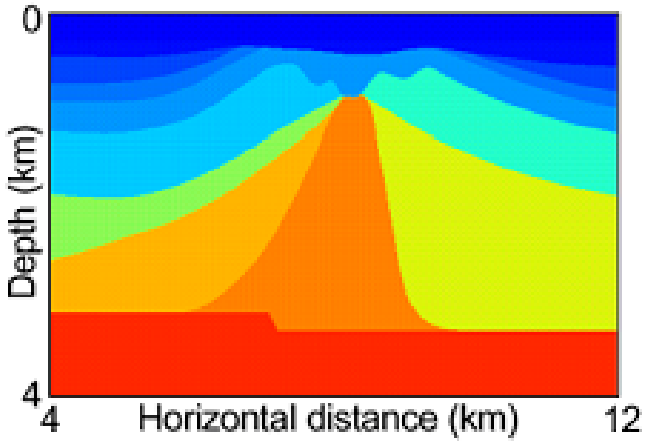

Figure 1: An exemple of a *blocky* subsurface model.

and parameters describing the geometry of the interfaces delimiting the velocity blocks[1]. Moreover, anisotropy of the velocity may be modelled by two parameters η and δ (see [13] for more details). This *blocky* representation allows the representation of complex models (e.g. presenting rapid lateral velocity variations and complex geological structures) with a reasonable number of unknowns. A strong a priori information is introduced, the number of layer being fixed: this reduces the underdetermination of the inverse problem.

The forward problem of reflection tomography consists in computing traveltimes of reflected seismic waves, given a model, an acquisition survey (locations of the sources and the receivers) and signatures (reflectors where the waves reflect). It is solved by a ray tracing method which is a high frequency approximation of the wave equation ([11]). We denote by $T(m)$ the forward modeling operator, that is the operator that gives, for a specified reflector, the traveltimes associated with all source-receiver pairs.

Reflection traveltime tomography is the corresponding inverse problem: its purpose is to adjust m such that $T(m)$ best matches a vector of traveltimes T^{obs} picked on seismic data. A natural formulation of this problem is the least squares formulation

$$\| T(m) - T^{obs} \|^2_{C_D^{-1}} + \| m - m^{prior} \|^2_{C_M^{-1}}, \tag{1}$$

where

- C_D^{-1} is the a priori covariance operator in the data space that allows to describe errors on the data,
- m^{prior} is an a priori model (coming, for instance, from geological knowledge and/or additionnal information resulting from well measurements),

[1]In our approach ([10]), the subsurface model m is composed of 2D or 3D B-spline functions describing velocity variations in a layer ($v(x,y)+k.z$ or $v(x,y,z)$) and 2D B-spline functions describing the interfaces ($Z(x,y)$, $Y(z,x)$ or $X(y,z)$).

- C_M^{-1} is the a priori covariance matrix in the model space that allows to describe errors on the a priori model.

In practice, building an a priori model is difficult and the a priori term, which is necessary to well pose the inverse problem, is often replaced by regularization terms made up of second derivatives of the model ([5]) which leads to $C_M^{-1} = \mathcal{Q}_{reg}$ and $m^{prior} = 0$

$$\| T(m) - T^{obs} \|^2_{C_D^{-1}} + \epsilon^2 \| D^2 m \|^2 \tag{2}$$

where a continuation technique ([3]) is applied for tuning the regularization weight ϵ: the weight is decreased until the model matches the data with the expected accuracy. The obtained model is then the smoothest model which fits the data. This methodology allows to stabilize the inversion.

This large size non linear least-square problem (the objective function is not quadratic since the forward modeling operator $T(m)$ is non linear) is solved classically by a Gauss-Newton method based on successive linearizations of the forward operator which needs the computation of the Jacobian matrix of T ($J(m) = \frac{\partial T}{\partial m}(m)$) [2]. The resulting approximation of the Hessian is noted $H(m) = J(m)^t C_D^{-1} J(m)$.

The resulting quadratic approximation of the objective function (2)

$$\tilde{C_n}(\delta m) = ||J_n \delta m - \delta T_n^{obs}||^2_{C_d^{-1}} + \epsilon^2 ||D^2(m_n + \delta m)||^2 \tag{3}$$

is minimized at each Gauss-Newton iteration n by a preconditionned conjugate gradient (m_n is the current model, δm is the model perturbation, $J_n = \frac{\partial T}{\partial m}(m_n)$ is the Jacobian matrix evaluated at m_n, and $\delta T_n^{obs} = T^{obs} - T^{cal}(m_n)$) ([4]).

3 PRESENTATION OF THE APPLICATION

Let us consider a 2D PP/PS[3] real data set studied by [2]. 45338 traveltime data were interpreted and an uncertainty of $5ms$ (resp. $8ms$) are associated with PP data (resp. PS data). A layer-stripping approach (separate inversion of each velocity layer from the shallower layer to the deeper one) furnished the velocity model of Figure 2. The model is described by four interfaces, corresponding to the interpreted events ($h1$, $h3$, $h4$ et $h5$) which define only three layers with lateral velocity variations for the two first upper layers ($v(x)$, there is no vertical variations) and the last layer stretching from $h3$ to $h5$ with a 2D velocity $vp5(x,z)$ and $vs5(x,z)$ (Figure 2). This model is composed of 4588 parameters, 592 for the interfaces and 3936 for the velocities. Anisotropy in the deepest layer is parameterized by η and δ ([2]) which are assumed constant per layer (η can

[2]The computation of the derivatives of T with respect to the model parameters is cheap, thanks to the Fermat principle ([1]).

[3]PS-wave results from the conversion at the reflector of a down-going P-wave (compressional wave) into an up-going S-wave (shear wave). In opposition to the PS-wave, the pure P mode is often called the PP-wave.

be seen as a measure of the velocity an-ellipticity, whereas δ controls the near vertical velocity propagation of the P-waves). For the two first layers, the inversion results are very satisfactory, with a traveltime misfit RMS (root mean square) of $3.5ms$ for the PP-data and of $7ms$ for the PS-data. For the last layer, the traveltime misfit RMS is of $3.7ms$ and of $5.6ms$ for, respectively, $h4$ and $h5$ PP-data and of $8.5ms$ for $h5$ PS-data, results which are consistent with the data uncertainties (Figure 3).

As already shown by [13], it turns out that the anisotropy parameter δ is strongly undetermined from seismic data. The value of δ parameter was obtained by a trial and error approach in order to match approximatly the depth of $h5$ horizon given at well location. We propose then to carry on an a posteriori uncertainty analysis in order to quantify the uncertainties on this solution model: we focus on the anisotropic layer delimited by $h3$ and $h5$ (velocities $vp5$ and $vs5$).

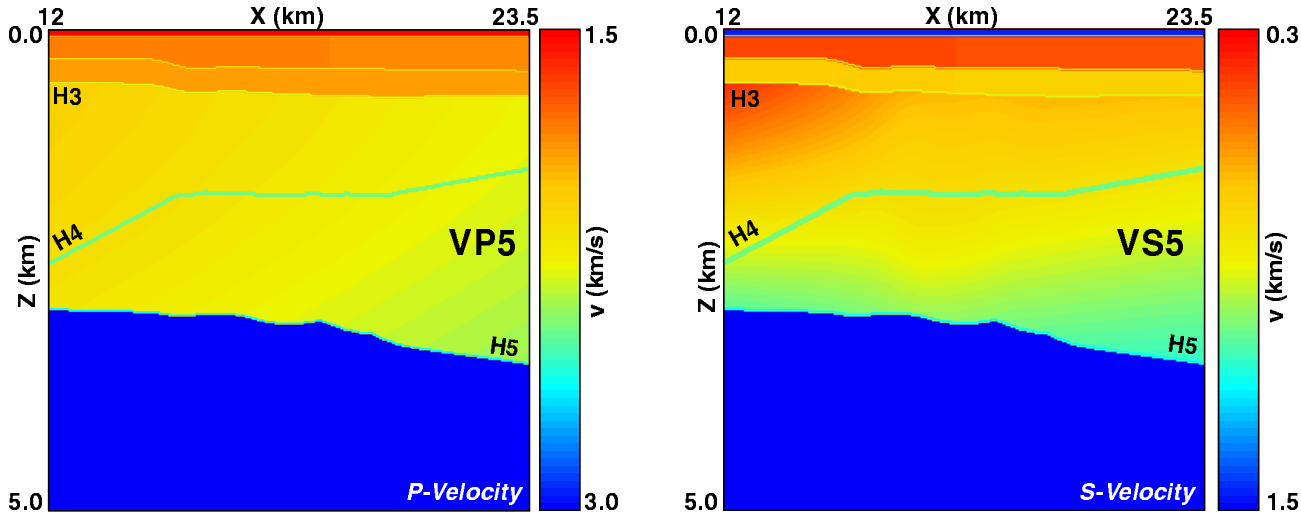

Figure 2: Solution velocity model obtained by tomography. Left: P velocity. Right: S velocity. The RMS value of the traveltime misfits is $6.2ms$. The anisotropy parameters values are: $\eta = 6.29\%$ and $\delta = -4.43\%$.

4 QUANTIFYING UNCERTAINTIES ON THE SOLUTION MODEL: A LINEARIZED APPROACH

A classical approach to quantify uncertainties consists in the analysis of the Hessian matrix (or its inverse: the a posteriori covariance matrix) associated with the linearized problem (3) around the solution m_∞. This approach is valid in the vicinity of the solution model, the size of the vicinity depending on the non-linearity of the forward map.

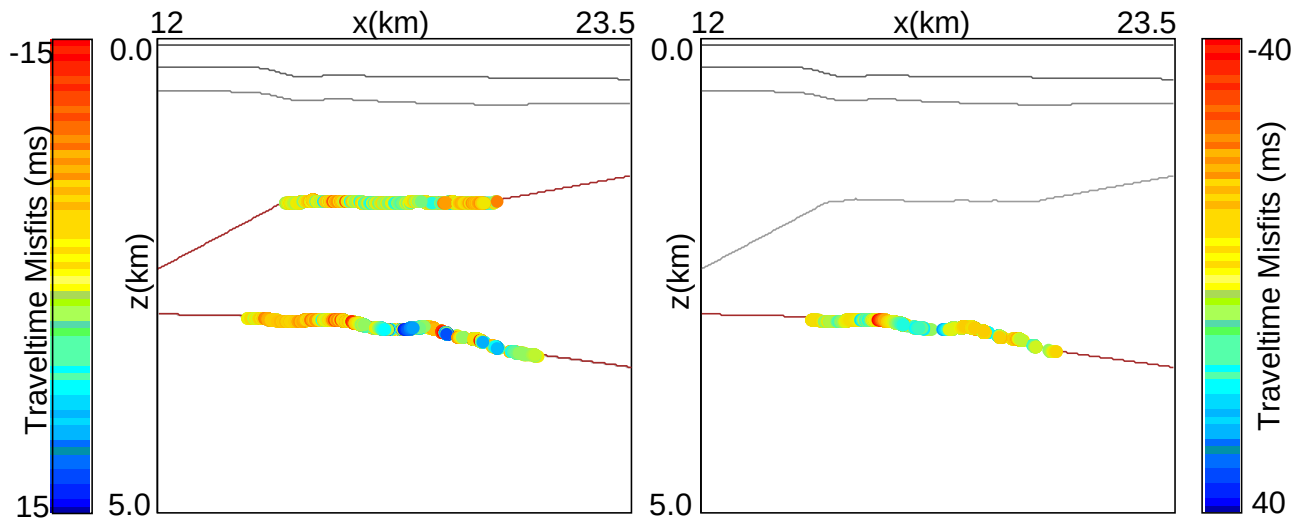

Figure 3: Interfaces of the solution velocity model obtained by tomography and their associated traveltime misfits: impact points of rays reflected on $h4$ and $h5$ (Left: PP rays. Right: PS rays), colors represent traveltime misfits. Notice that the illuminated zone extends from $14km$ to $20.5km$.

The bi-linear form associated with the Hessian matrix measures the influence of a model perturbation δm on the quadratic cost function defined around the solution m_∞:

$$\begin{aligned}
\tilde{C}(\delta m) - \underbrace{\tilde{C}(0)}_{=C(m_\infty)} &= \frac{1}{2}(J(m_\infty)\delta m)^T C_D^{-1}(J(m_\infty)\delta m) + \frac{1}{2}\delta m^T C_M^{-1}\delta m \\
&\quad - \delta m^T \underbrace{(J(m_\infty)^T C_D^{-1}\delta T^{obs}(m_\infty) - C_M^{-1}m_\infty)}_{=g(m_\infty)=0,\ \text{gradient of } C(m) \text{ vanishes at the solution } m_\infty} \\
&= \frac{1}{2}\delta m^T (J(m_\infty)^T C_D^{-1} J(m_\infty) + C_M^{-1})\delta m \qquad (4) \\
&= \frac{1}{2}\delta m^T H(m_\infty)\delta m. \qquad (5)
\end{aligned}$$

The a posteriori covariance matrix is defined by

$$C'_M = (J(m_\infty)^T C_D^{-1} J(m_\infty) + C_M^{-1})^{-1}. \qquad (6)$$

The space of admissible models can be characterized by the contour lines

$$(m - m_\infty)^T C_M'^{-1}(m - m_\infty) = constant, \qquad (7)$$

which are ellipsoïds of center m_∞ and correspond also to contour lines of the a posteriori Gaussian probability density function

$$\exp\left(-\frac{1}{2}\delta m^T C_M'^{-1}\delta m\right). \qquad (8)$$

The diagonal terms of C'_M are the uncertainties on the parameters describing the model and the off-diagonal terms are the correlations between these uncertainties. For instance, the probability that the true model parameter p_i verifies $-2C'_{Mi,i} \leq p_i - p_\infty \leq 2C'_{Mi,i}$, independently of the values of the other model parameters, is about 95%. To take into account the correlations between the parameters, we should study the 95% confidence ellipsoid. The axes of the ellipsoïds (7) are defined by the eigenvectors of C'_M, the square root of the eigenvalues giving the uncertainties on the associated eigenvector.

We propose in the next sections two methods to quantify geological uncertainties on the solution model which avoid the expensive computation and the cumbersome analysis of the generally huge a posteriori covariance matrix C'_M.

4.1 Simulations of admissible models

The first proposed method to quantify uncertainties is the simulation of admissible models from the a posteriori probability density function (8) (pdf). The simulations furnish directly interpretable results, i.e. physical models. The method (see for instance [12]) consists in random simulations of model perturbations following the pdf (8).

First, we apply the variable transformation:

$$\delta m' = U \delta m, \tag{9}$$

where U is the lower triangular matrix obtained by a Cholesky decomposition of the inverse of the a posteriori covariance matrix $C'^{-1}_M = H = U^T U$ (symmetric semi positive definite matrix). The method consists then in simulations of the uncorrelated Gaussian pdf with unit variance

$$\exp\left(-\frac{1}{2}\delta m'^T \delta m'\right). \tag{10}$$

Simulations $\delta m'$ are transformed into correlated Gaussian simulations of vector δm by the inverse transformation of (9).

Figures 4, 5, 6, 7 and 8 show 100 simulated models obtained by this method from the solution model of Figure 2: interfaces $h4$ and $h5$, variations of velocities $vp5$ and $vs5$ along x and z directions and histograms of the anisotropy parameters η and δ. From these simulations, we observe that the highest uncertainties on the lateral velocity variations are located at the boundaries of the model, areas that are not well illuminated by the rays. Indeed, for a slice at constant $z = 2.7km$, we observe uncertainties of $\approx 690m/s$ for $vp5$ and of $\approx 260m/s$ for $vs5$ at the boundaries of the model and uncertainties of $\approx 450m/s$ for $vp5$ and of $\approx 140m/s$ for $vs5$ in the illuminated parts of the model. Concerning the interface depths, we observe uncertainties of $\approx 180m$ for reflector $h4$ and of $\approx 220m$ for reflector $h5$ at the boundaries and of $\approx 90m$ for reflector $h4$ and of $\approx 100m$ for reflector

$h5$ elsewhere. For the anisotropy parameters η and δ, we notice uncertainties around 0.3% on η and 2% on δ. This method is quite attractive for the straightforward interpretation of the results despite its cost (cost of the Choleski decomposition): it furnishes physical models leading to small perturbations of the quadratic cost functions (3) (see Table 4.3).

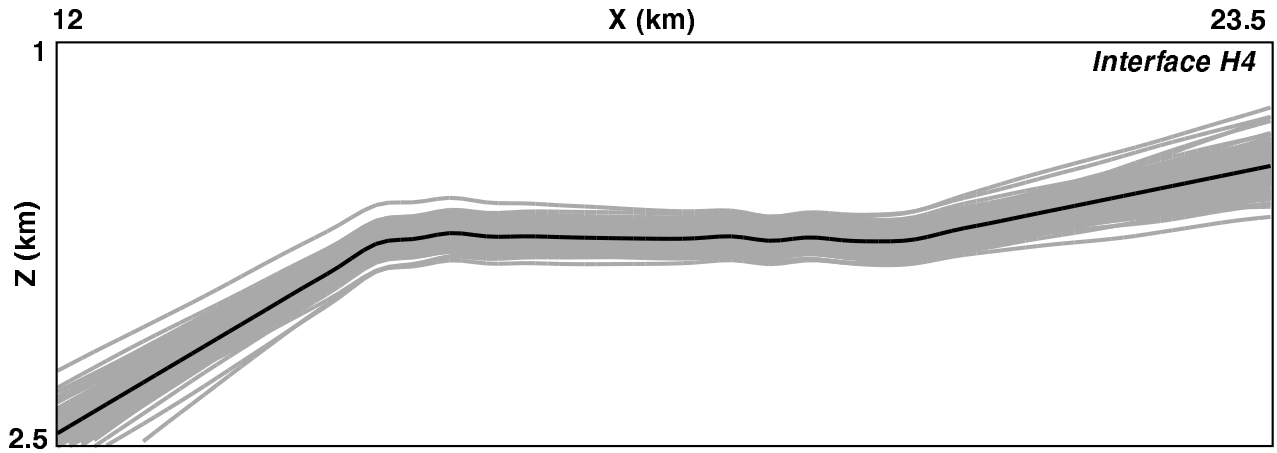

Figure 4: Reflector h4 extracted from the 100 simulated velocity models. The bold line represents the solution interface.

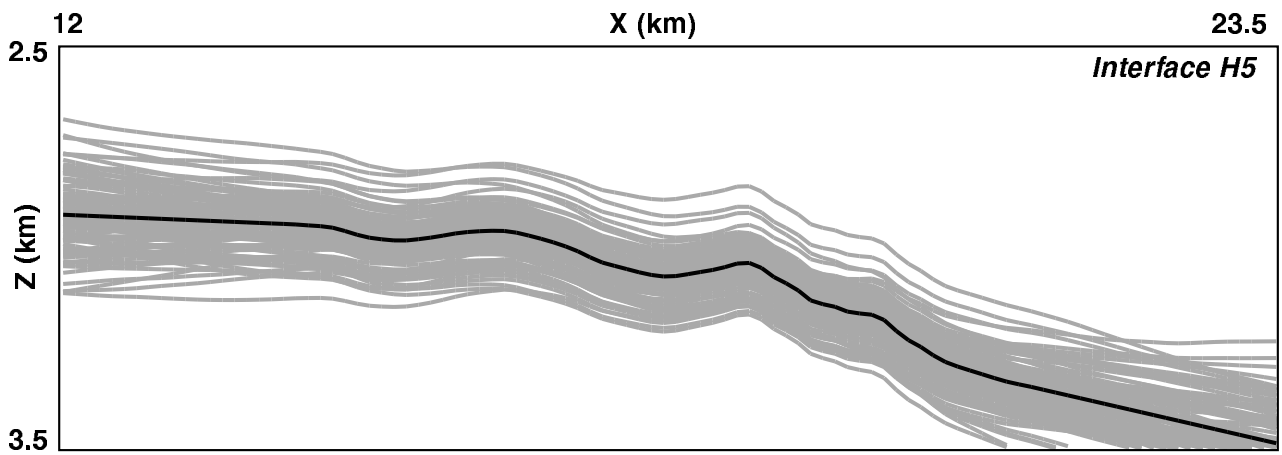

Figure 5: Reflector h5 extracted from the 100 simulated velocity models. The bold line represents the solution interface.

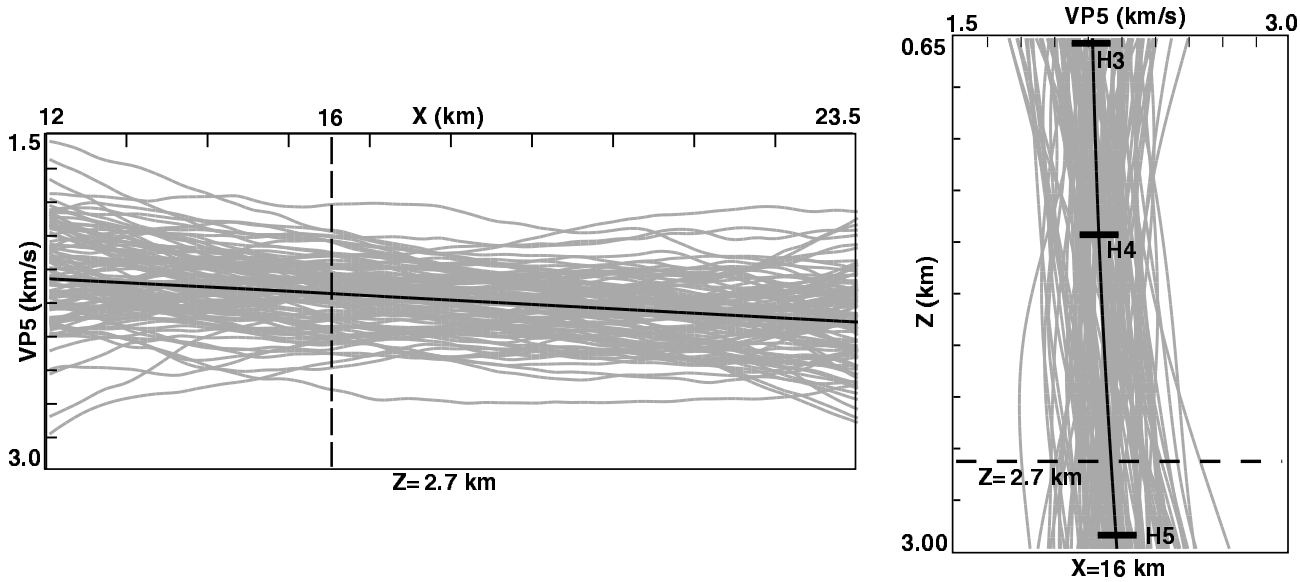

Figure 6: P-velocity variations along x axis at $z = 2.7km$ (left) and along z axis at $x = 16km$ (right) extracted from the 100 simulated velocity models.

4.2 Uncertainties on geological macro-parameters

We propose here a method that allows to deal with large size models at a reasonable cost and furnishes uncertainties on chosen physical quantities. The proposed approach consists in building macro-parameters (MP) with a geophysical interest. These macro-parameters are linear combinations of the inverted parameters such as the mean of the velocity variations in a zone, the slope of an interface, the average thickness of a layer, etc. [8] has introduced the notion of macro-parameter (his main motivation being to avoid numerical problems in the inversion of the complete Hessian). We propose here a generalization of his work (general definition of macro-parameters) which allows the computation of uncertainties for huge 3D problems.

We define a macro-parameter as:

$$P = Bp, \tag{11}$$

where p is the np model parameters vector, P is the n_{MP} macro-parameters vector and B is the condensation matrix. We compute the a posteriori covariance matrix $\tilde{C}'_M$ in the macro-parameter space:

$$\tilde{C}'_M = BC'_M B^T. \tag{12}$$

Note that $\tilde{C}'_M$ (a $n_{MP} \times n_{MP}$ matrix) is small compared to C'_M (a $n_p \times n_p$ matrix) since $n_{MP} \ll n_p$. To obtain this matrix, we do not need to compute the whole inverse of the Hessian, $C'_M = H^{-1}$, in the parameter space. Indeed, to obtain $\tilde{C}'_M = \tilde{H}^{-1} = BH^{-1}B^T$ in

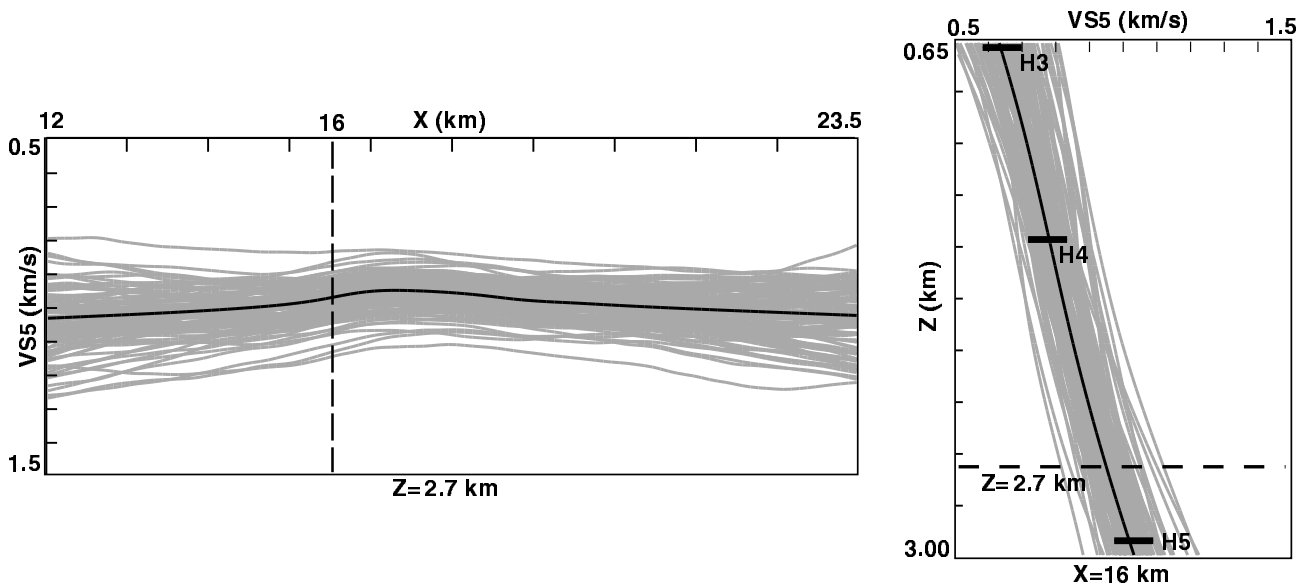

Figure 7: S-velocity variations along x axis at $z = 2.7km$ (left) and along z axis at $x = 16km$ (right) extracted from the 100 simulated velocity models.

the MP space, we just need $H^{-1}B_j^T$, where B_j^T are the different columns of B^T. We thus solve n_{MP} linear systems, $\tilde{H}_j^{-1} = BH^{-1}B_j^T$, that are similar to the linearized problem we solve at each Gauss-Newton iteration. Thus the computational cost for one MP is comparable to one iteration of the inversion process.

We applied this method on the model of Figure 2. We choose simple MP: the mean of the velocity variations and the mean of the interface depth in the illuminated part of the layer (region covered by rays)[4]. The results are listed in the Table 1. The uncertainty on anisotropy parameter δ is high: $\pm 3.2\%$ (if we consider twice standard deviation) and we observe high correlations between η and δ (0.93) and also correlations between anisotropy parameters and $H5$ depth (-0.33 and -0.36). We notice also the bad determination of the velocities: $\pm 22\%$ (relative value). All these results are consistent with the simulations results (section 4.1).

This method with its general formalism allows to compute uncertainties on relevant geological quantities with a reasonable computation cost and to highlight easily strong coupling between the chosen MP.

4.3 Discussions on the two methods

As already mentioned, the two methods described in sections 4.1 and 4.2 rely on the quadratic approximation (3) of the non linear cost function. In the table 4.3 we have listed the RMS of traveltime misfits for 20 simulated models. Note that the RMS traveltime

[4]From figure 3, we define the illuminated region by $x \in [14, 20.5km]$.

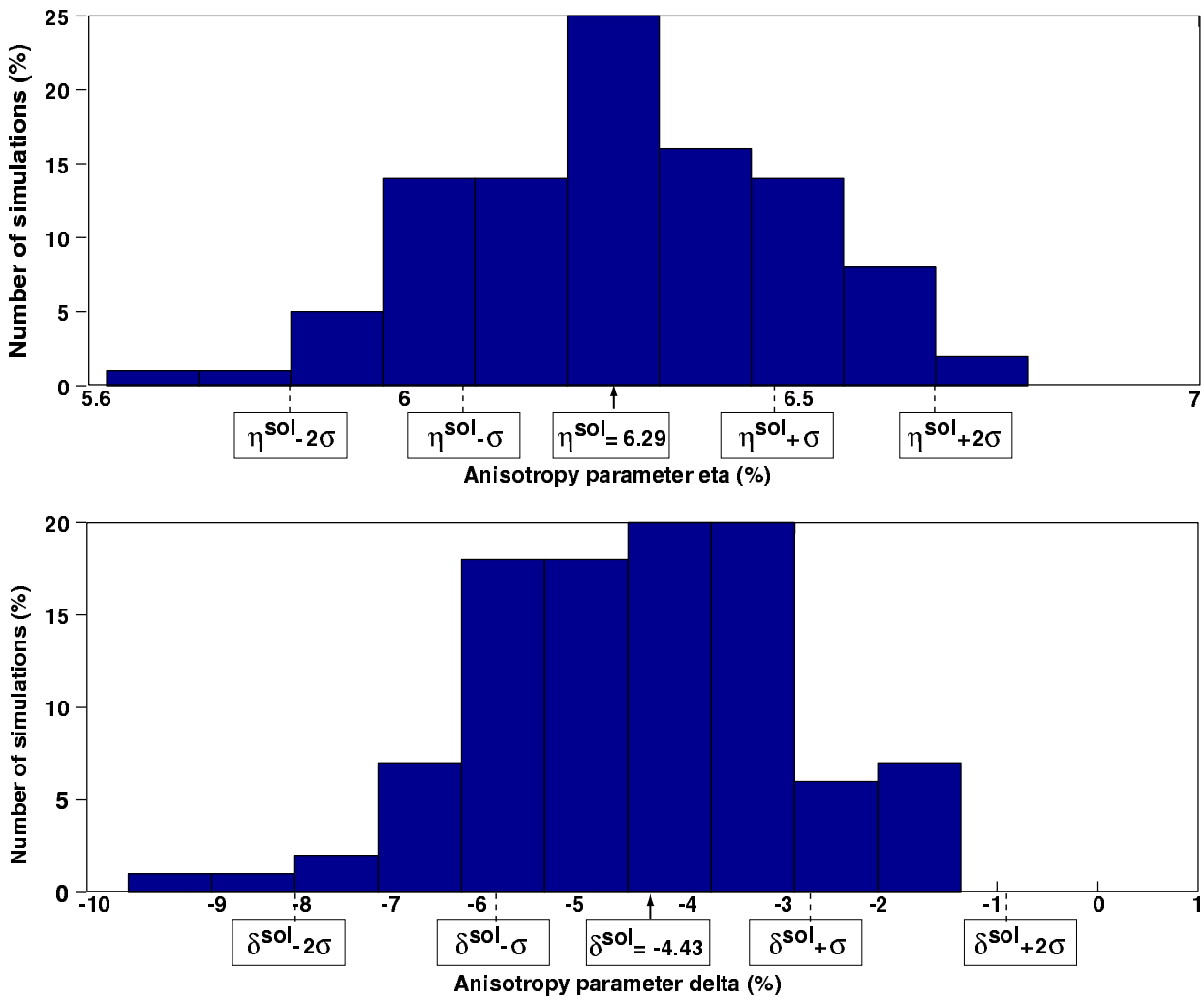

Figure 8: Histograms of the 100 simulated velocity models. Top: anisotropy parameter η. Bottom: anisotropy parameter δ. σ corresponds to the standard deviation on the anisotropy parameter η (resp. δ) given by the macro-parameter method.(see section 4.2).

misfit for the solution model is $6.2ms$. Thus, there is only one model which is not acceptable ($11.6ms$ for the model 15), the others are admissible models with a RMS traveltime misfit bounded by $7ms$. The simulation method has then furnished model perturbations that correspond to small perturbations of the quadratic cost function but also to small perturbations of the non linear cost function (RMS of traveltime misfits remain small). It shows for this example, the good concordance between the quadratic cost function and the non linear one around the solution model.

Figure 9 illustrates isovalues of the probability density function (8) in the MP space, i.e. ellipsoïds defined by $\delta m_{MP}^T \tilde{C}_M'^{-1} \delta m_{MP}$: we visualize the 2D-marginal probability density functions for several relevant MP couples, the 68% confidence ellipsoïd and the 95% confidence ellipsoïd for η and δ, $vp5$ and $vs5$, δ and $h5$. This representation allows to

	Vp5	***Vs5***	η	δ	***H4***	***H5***
Vp5	**475.1 m/s** *(22%)*	0.002	-0.03	-0.02	0.005	0.01
Vs5	0.002	**168.9 m/s** *(22%)*	-0.04	-0.03	0.005	0.01
η	-0.03	-0.04	**0.22%** *(4%)*	0.93	-0.16	-0.33
δ	-0.02	-0.03	0.93	**1.6%** *(37%)*	-0.17	-0.36
H4	0.005	0.005	-0.16	-0.17	**77.1 m** *(5%)*	0.06
H5	0.01	0.01	-0.33	-0.36	0.06	**80.3 m** *(3%)*

Table 1: Standard deviations in bold (square roots of the diagonal terms of the a posteriori covariance matrix) associated with the MP "mean in the illuminated part of the model" and the normalized correlations between the uncertainties on the different defined MP. This uncertainty analysis is performed for the solution model of Figure 2. The normalized values $\frac{\sigma}{|value^{sol}|} \times 100$ of the standard deviations (σ) are also indicated in italic.

underline the correlations between the anisotropy parameters η and δ which are quite strong in this example: the 68% and the 95% confidence ellipsoïds are elongated (left part of the figure 9). On the opposite, the P and S velocities are uncorrelated: we obtain almost a circle (middle part of the figure 9). We can also observe correlations between the anisotropy parameter δ and the mean of the last reflector depth $h5$ (right part of the figure 9). On this Figure, we have also superimposed the anisotropy parameters η and δ and the MP $vp5$ and $vs5$, δ and MP $h5$ built from the 100 simulated models. This figure allows to point out the equivalence of the simulation method and the macro-parameter approach.

Note that the macro-parameter approach furnishes interesting information on the correlations of macro-parameter uncertainties and has a lower computation cost than simulations of the model parameters with the complete a posteriori covariance matrix. We could also perform gaussian simulations of the macro-parameters using the reduced a posteriori covariance matrix.

5 NON-LINEAR A POSTERIORI ANALYSIS

We have shown in the previous sections that a linearized uncertainty analysis allows to delimit the range of possible solution models that fit, with the expected accuracy, the data

Models	*RMS(ms)*	*vp5(km/s)*	*vs5(km/s)*	*h4(km)*	*h5(km)*	*η(%)*	*δ(%)*
1	6.7	3.736	1.154	1.783	3.131	6.5	-2.6
2	6.6	3.121	1.222	1.925	3.222	6.2	-4.9
3	6.4	2.746	1.102	1.790	3.060	6.0	-5.9
4	6.9	2.701	1.315	1.857	3.259	6.2	-3.7
5	6.2	3.300	1.416	1.817	3.142	5.7	-8.3
6	6.5	2.542	1.245	1.926	3.181	6.1	-4.6
7	6.6	2.536	1.137	1.925	3.143	6.1	-5.5
8	6.2	1.919	1.087	1.844	3.102	6.3	-4.8
9	6.2	3.182	1.078	1.875	3.267	6.5	-3.5
10	6.4	2.960	1.388	1.736	3.198	6.0	-6.2
11	6.2	3.247	1.452	1.815	3.173	6.0	-6.4
12	6.2	2.592	1.073	1.664	3.147	6.6	-1.4
13	6.5	2.510	1.239	1.817	3.239	6.2	-4.9
14	6.6	3.595	1.130	1.717	3.208	6.5	-1.9
15	11.6	2.344	1.172	1.770	3.144	6.0	-6.1
16	6.5	2.154	1.323	1.853	3.2322	6.3	-4.5
17	6.3	2.855	1.307	1.797	3.296	6.3	-3.4
18	6.2	2.125	1.049	1.921	3.118	5.9	-6.8
19	6.6	1.431	1.048	1.772	3.144	6.4	-3.7
20	6.3	2.371	1.225	1.864	3.166	6.0	-6.6

Table 2: RMS values of the traveltime misfits, mean value of the velocities $vp5$ and $vs5$, of the interfaces $h4$ and $h5$ in illuminated region (MP defined in section 4.2 and anisotropy parameters η and δ for 20 simulated models (chosen among the 100 simulations presented in section 4.1).

and the a priori information. Nevertheless, we should keep in mind that this approach is only valid in the vicinity of the solution model (linearized framework) and complex cases may require a non linear approach.

Model 15 of table 4.3 indicates that some of the simulated models may produce unacceptable traveltimes misfits ($RMS \approx 11ms$ and MAX $= 25ms$ for PP data and $= 41ms$ for PS data). It shows the limitation of the quadratic approximation of the non linear cost function and thus the limitation of the linearized methods of uncertainty quantification. Moreover, [4] has shown that a tomographic inversion with constraints on the location of the interfaces at well locations furnished a very different model (model 2 of Figure 10) than the solution model of Figure 2. The two models verify the traveltime data with the expected accuracy. This second model does not belong to the range of possible models detected by the two methods we proposed: indeed, δ is equal to 15.94% whereas the expected range of values was between -8% and -1% (see for instance Figure 8). It shows

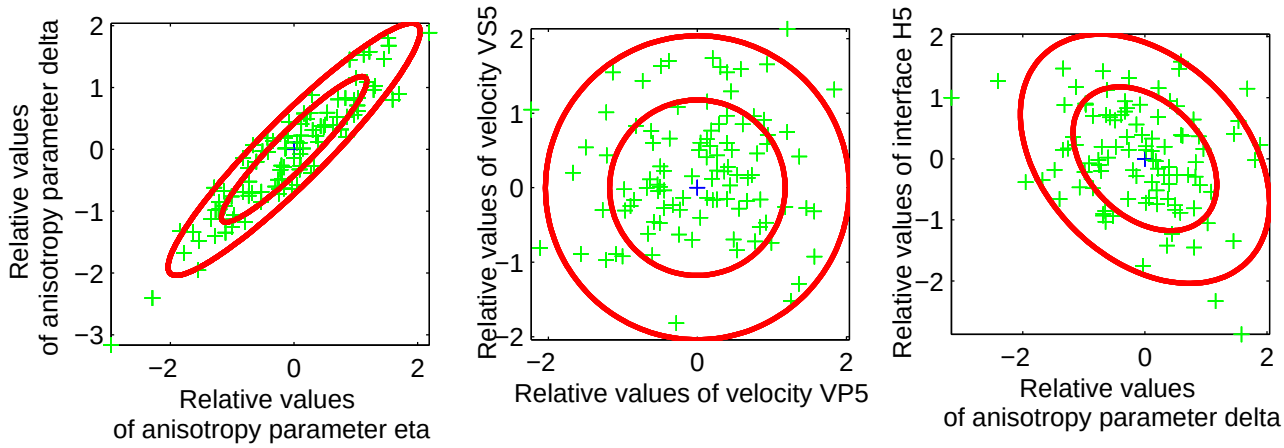

Figure 9: 68% and 95% isovalues of the marginal probability density function for the anisotropy parameters η and δ (left), P-velocity and S-velocity (middle), δ and the interface depth of reflector h5 (right). The represented values are the relative ones (for instance: $\frac{\eta-\eta^{sol}}{|\sigma^{\eta}|}$). The crosses represent the MP values of the 100 simulated models.

again the limitations of the linearized approach.

To perform a non linear analysis we have chosen an experimental approach which consists in performing several inversions allowing to test different geological scenarii to try to delimit the space of admissible solutions. For instance, we could test different hypothesis on the values of the anisotropy parameters for which a strong uncertainty has been detected by the linearized approach. If we introduce a constraint on the anisotropy parameter δ, such as $\delta > 0$, we find the model 3 displayed in Figure 11: this result expands the range of admissible models (variability of δ detected by MP approach was $-4.43 \pm 3.2\%$). In the same way, we can test the variability of the anisotropy parameter η. We notice (see table 5) that the value of the anisotropy parameter η of the model 2 (Figure 10) is 8.27%, but all the simulated values of this parameter are around 6% and the variablity of η furnished by the MP is 0.44%. Nevertheless, tomographic inversion under the constraint $\eta \leq 4\%$ or $\eta \geq 8\%$ does not furnish a model that satisfies both these constraints and the data with the expected accuracy. We have shown that, by constrained tomography, we can test different geological hypothesis and delimit the space of admissible models and go further than the linearized approach which explores only the vicinity of the solution model which may be one local minimum around others.

6 CONCLUSIONS

Reflection tomography furnishes the velocity model that best fits the traveltime data: however, this solution is only one among many admissible models. An a posteriori un-

	Model 1	***Model 2***	***Model 3***
mean of the P-velocity	2.849 km/s	3.623 km/s	2.895 km/s
mean of the S-velocity	1.229 km/s	1.621 km/s	1.213 km/s
η	6.29%	8.27%	6.2%
δ	-4.4%	15.94%	2%
mean of the depth of the reflector H4	1.838 km	1.755 km	1.659 km
mean of the depth of the reflector H5	3.167 km	3.021 km	2.882 km

Table 3: 3 different solution models. Model 1 is the model of Figure 2, Model 2 is the model of Figure 10 and Model 3 is the model of Figure 11.

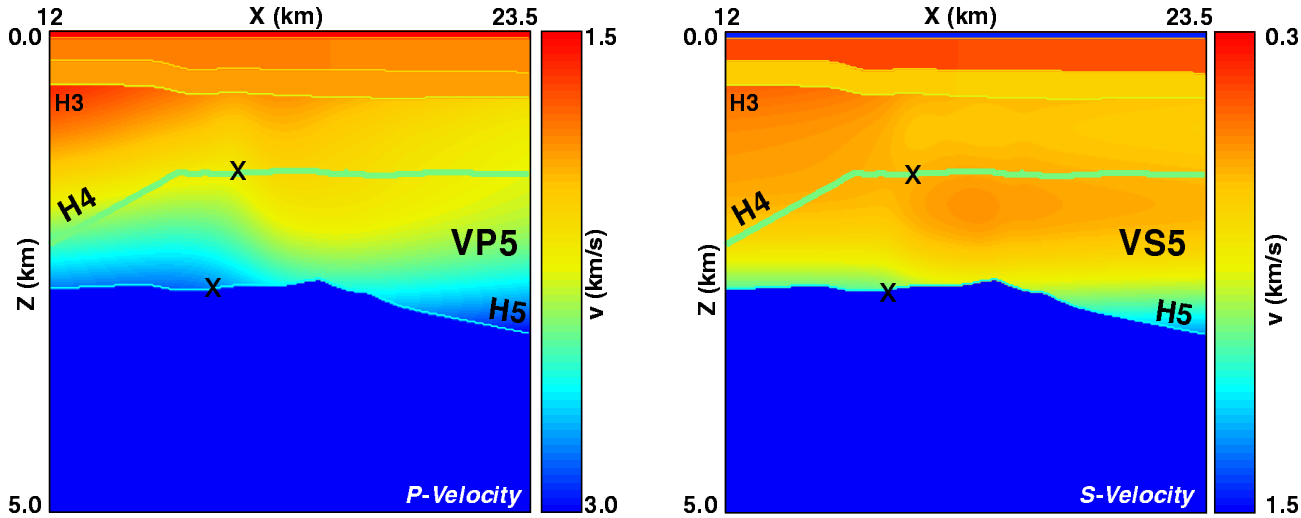

Figure 10: Solution model 2 obtained by tomography with additional constraints on the interface depth at well locations. The RMS value of the traveltime misfits is $6.0ms$. The anisotropy parameters values are: $\eta = 8.27\%$ and $\delta = 15.94\%$. The crosses represent the wells.

certainty analysis is crucial to delimit the range of possible solution models that would fit the data and the a priori information with the expected accuracy. In this paper, we describe two methods to perform a linearized a posteriori analysis, approach valid only in the vicinity of the solution model. The simulation method based on the analysis of the a posteriori covariance matrix (huge matrix for 3D models) is quite attractive for the straightforward interpretation of the results despite its cost. We propose a general formalism to reduce the a posteriori analysis to geological quantities of the model: we evaluate uncertainties on macro-parameters (linear combinations of model parameters) that have a geological interest. This method allows the manipulation of reduced matrices and thus becomes feasible in 3D. Nevertheless, we have confirmed on this real data set that this approach is only valid in the vicinity of the solution model (linearized framework) and that some cases require a non linear approach to search for local minima of the cost function. An experimental approach has consisted in solving the inverse problem

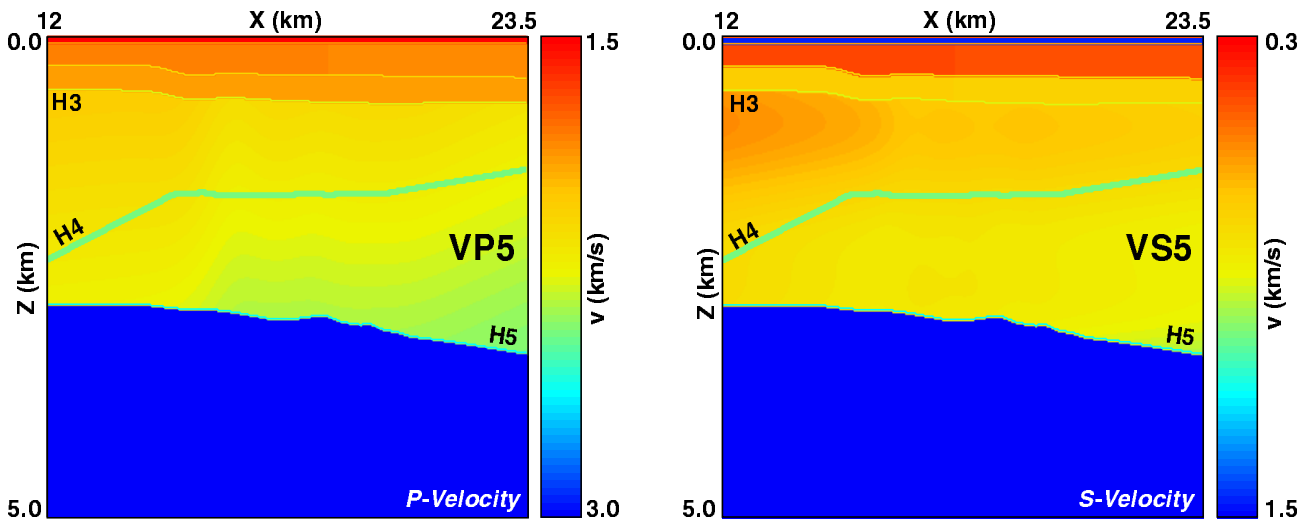

Figure 11: Solution model 3 obtained by tomography with a constrain on δ parameter to remain positive. The RMS value of the traveltime misfits is $6.371ms$. The anisotropy parameters values are: $\eta = 6.2\%$ and $\delta = 2\%$.

under constraints (for instance constraints on anisotropy parameters) to test the validity of some geological hypothesis.

7 ACKNOWLEDGMENTS

The authors would like to thank bp to have provided the data set, Karine Broto and Frédéric Delbos for many fruitful discussions.

REFERENCES

[1] T. N. Bishop, K. P. Bube, R. T. Cutler, R. T. Langan, P. L. Love, J. R. Resnick, R. T. Shuey, D. A. Spindler and H. W. Wyld. Tomographic determination of velocity and depth in laterally varying media, *Geophysics*, **50**, no 6, 903–923, 1985.

[2] K. Broto, A. Ehinger and J. Kommedal. Anisotropic traveltime tomography for depth consistent imaging of PP and PS data, The leading Edge, 114–119, 2003.

[3] K. P. Bube and R. T. Langan. A continuation approach to regularization for traveltime tomography. In *Expanded Abstracts*, 980–983, 64th Ann. Internat. Mtg., Soc. Expl. Geophys., 1994.

[4] D. Sinoquet, F. Delbos and J.C. Gilbert. A dedicated constrained optimization method for 3D reflection tomography. In *Expanded Abstracts*. EAGE 66th conference and exhibition, 2004.

[5] F. Delprat-Jannaud and P. Lailly. Ill-posed and well-posed formulations of the reflection traveltime tomography problem. *Journal of Geophysical Research*, **98**, 6589–6605, 1993.

[6] C. Duffet and D. Sinoquet. Quantifying geological uncertainties in velocity model building. In *Expanded Abstracts*, 926–929, 72nd Ann. Internat. Mtg., Soc. Expl. Geophys., 2002.

[7] W. P. Gouveia and J.A. Scales. Resolution of seismic waveform inversion: bayes versus occam. *Inverse problems*, 323–349, 1997.

[8] D. Grenié. *Tomographie de temps de trajet adaptée au problème de l'imagerie 3D en exploration pétrolière: théorie, mise en oeuvre logicielle et applications.* Ph.D. thesis, Université Paris VII, 2001.

[9] http;//consortium.ifp.fr/KIM/, 2001, *Web pages of former KIM consortium.*

[10] F. Jurado, D. Sinoquet and A. Ehinger. 3D reflection tomography designed for complex structures. In *Expanded Abstracts*, 711–714, 66th Ann. Internat. Mtg., Soc. Expl. Geophys., 1996.

[11] F. Jurado, P. Lailly and A. Ehinger. Fast 3D two-point raytracing for traveltime tomography. Proceedings of SPIE, *Mathematical Methods in geophysical Imaging.* Vol. V, **3453**, 70–81, 1998.

[12] R. Parker. *Geophysical inverse theory.* Springer, Berlin, 1994.

[13] A. Stopin. Reflection tomography of compressional and shear modes for determination of anisotropic velocity models: Phd Thesis of Université Louis Pasteur, Strasbourg I, 2001.

[14] A. Tarantola. *Inverse problem theory.* Elsevier, 1987.

6.2 Deuxième application : KIMASI

6.2.1 Présentation de l'application

Au cours du projet KIMASI, la tomographie de réflexion a été appliquée à un jeu de données réelles 3D provenant de la Mer du Nord fourni par BP (Ehinger et al., 2001). Un modèle de vitesse a été obtenu grâce à une approche «layer-stripping» (inversion couche par couche de haut en bas) où les variations latérales et verticales de vitesse à travers les unités Tertiaire, Paléocène et Crétacée (divisés en deux couches) ont été déterminées de manière séquentielle (voir Figure 6.1).

Une forte sous-détermination de la vitesse dans la partie supérieure du Tertiaire a été détectée au cours du processus d'inversion : elle est due à la large épaisseur de la couche ($\approx 2.5km$) et au nombre réduit de rayons. Plusieurs inversions avec différentes paramétrisations de la vitesse ont donné les mêmes résidus des temps de trajet. Ces différents tests sont coûteux en temps (chaque test correspond à une inversion complète) et les profondeurs des réflecteurs fournies par les données de puits ne sont pas très bien retrouvées, cette information n'étant pas explicitement introduite dans le processus d'inversion.

Pour obtenir un modèle plus consistant avec les données de puits, Delbos et al., 2004, ont proposé d'appliquer une inversion tomographique sous contraintes. Les profondeurs des interfaces sont contraintes aux 5 positions de puits et la variabilité du gradient vertical de vitesse dans la couche du Tertiaire est contrainte par les mesures de puits (tableau 6.1).

Pour limiter le mauvais calage des données pour les couches profondes, on réalise une inversion globale. Cette expérience consiste alors en une inversion globale de 127569 données temps de trajet pour 5960 inconnues décrivant 4 couches de vitesse et 5 interfaces, soumises à 2300 contraintes. Les résultats sont présentés sur la Figure 6.2 : le modèle obtenu vérifie les données avec la même précision que le modèle obtenu par Ehinger et al., 2001, et vérifie toutes les contraintes (tableaux 6.1 et 6.2).

Nous allons utiliser l'approche macro-paramètre (voir chapitre 4) pour illustrer l'indétermination des vitesses dans le tertiaire (en évitant de faire plusieurs inversions) et quantifier les incertitudes quand on ajoute les contraintes aux puits. Nous utiliserons également un nouveau macro-paramètre plus réaliste et plus complexe que ceux décrit dans la section 4.3.2.1 du chapitre 4 : le macro-paramètre qui correspond au volume d'une couche délimitée par des interfaces.

	Modèle obtenu avec l'inversion sans contraintes		**Modèle obtenu avec l'inversion avec contraintes**	
	Résidu moyen RMS	**Résidu Max.**	**Résidu moyen RMS**	**Résidu Max.**
layer1	4.3ms	24ms	3.3ms	26.6ms
tpal	4.2ms	18.4ms	4.4ms	20.4ms
tchalk	3.9ms	19.4ms	4.1ms	24.5ms
ichalk	8.5ms	59.6ms	7.6ms	52.1ms
bchlak	7.6ms	67.8ms	8.6ms	70.6ms

TAB. 6.1 – Tableau récapitulatif des résidus moyen RMS et résidus maximum pour les modèles solutions des Figures 6.1 et 6.2.

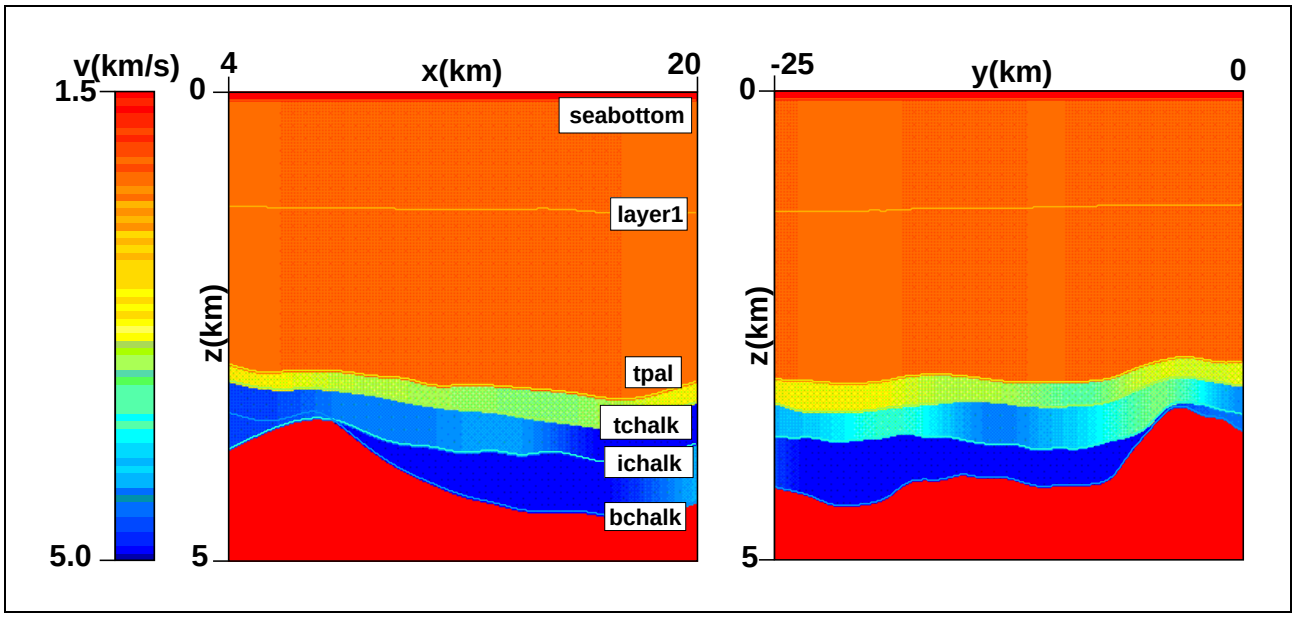

FIG. 6.1 – Coupes verticales du modèle de vitesse (Gauche : le long de l'axe des x, droite : le long de l'axe des y) obtenu grâ ce à une approche «layer-stripping». Le RMS des résidus temps de trajet = 6.1ms.

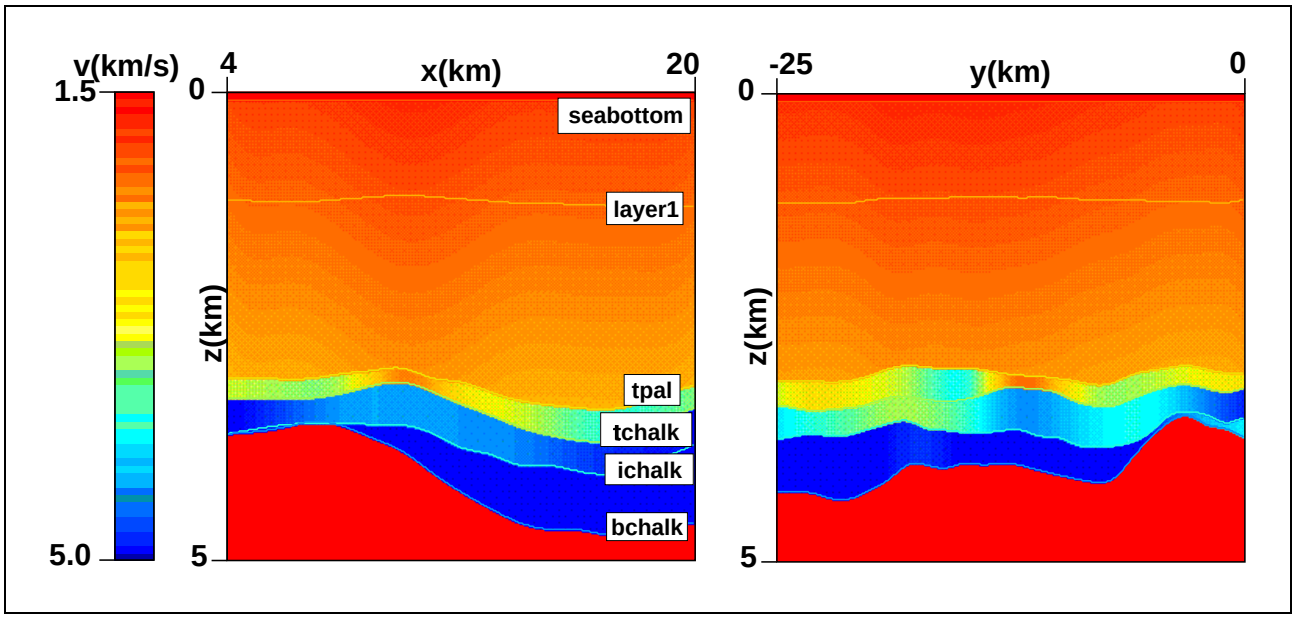

FIG. 6.2 – Coupes verticales du modèle de vitesse (Gauche : le long de l'axe des x, droite : le long de l'axe des y) obtenu grâ ce à une inversion globale sous contraintes (les contraintes sont décrites dans le tableau 6.2). Le RMS des résidus temps de trajet = 6.5ms.

Contraintes		**Modèle obtenu avec l'inversion sans contraintes**	**Modèle obtenu avec l'inversion avec contraintes**
Ecart moyen de la profondeur aux 5 puits	tpal	96m	0m
	tchalk	132m	0m
	bchalk	140m	0m
gradient vertical de vitesse du Tertiaire	0.1<k<0.3/s	k=0/s	k 0.18/s
intervalle de de vitesse	2.5<vpal<4km/s	ok	ok
	3.5<vichalk<5.7km/s	ok	ok
	4.2<vchalk<5.8km/s	ok	ok

TAB. 6.2 – Description des contraintes d'égalité et d'inégalités introduites dans l'inversion. On essaie d'obtenir un modèle qui vérifie les données et les 2300 contraintes.

6.2.2 Résultats numériques

Ici nous réalisons une analyse d'incertitude linéarisée et illustrons l'approche macro-paramètre sur l'exemple de la Figure 6.1 (pour les caractéristiques du modèle se reporter aux tableau 6.3).

Nous choisissons, d'une part, des macro-paramètres simples : la moyenne de la profondeur de l'interface, la moyenne des variations de vitesse et du gradient vertical de vitesse dans le Tertiaire. D'autre part, nous considèrerons le macro-paramètre volume d'une couche défini dans le paragraphe 4.3.2.3. Les résultats sont récapitulés dans le tableau 6.4.

Pour la couche qui nous intéresse, celle du Tertiaire, et pour laquelle l'étude Ehinger et al., 2001, a montré une forte sous-détermination de la vitesse dans la partie supérieure du Tertiaire, nous obtenons une incertitude de $1.3m$ sur l'interface tpal (voir Figure 6.1) qui est en moyenne à une profondeur de $3.04km$ (tableau 6.3), et une incertitude de $0.0042/s$ sur le gradient vertical de vitesse dans le Tertiaire, la moyenne de la vitesse dans le Tertiaire étant de $2.47km/s \pm 16.8m/s$ ($16.8m/s$, i.e. deux fois l'écart type, représente l'incertitude sur la moyenne de la vitesse dans la couche). D'autre part, nous pouvons noter de fortes corrélations (≈ 0.93) entre le gradient vertical de vitesse dans le Tertiaire et la profondeur de l'interface layer1 (tableau 6.6). Et en ce qui concerne le volume de cette couche Tertiaire, on note une incertitude de $206.4km^3$ pour un volume de $1414.4km^3$.

Cette méthode avec son formalisme général nous a permis de calculer des incertitudes sur des quantités géologiques intéressantes (par exemple sur le volume de la couche du tertiaire ou sur le gadient vertical de vitesse dans le Tertiaire) avec un coût de calcul raisonnable car sans inversion.

On peut s'étonner des incertitude trouvées : en effet, elles sont faibles alors que l'on sait qu'il y a une forte indétermination dans la couche Tertiaire. On peut remettre en cause l'approche linéarisée, c'est pourquoi nous testons l'adéquation de la fonction coût et de son approximation quadratique pour une perturbation de modèle correspondant à la différence entre les deux modèles obtenus par l'inversion avec (modèle 2, Figure 6.2) et sans contraintes (modèle 1, Figure 6.1). L'approximation quadratique étant réalisée autour du modèle 1.

Maintenant, si on note m_1 le modèle obtenu avec une inversion sans contraintes, m_2 le modèle

obtenu avec une inversion avec contraintes et $\delta m = m_2 - m_1$, et qu'on compare $\tilde{C}(m_2) - \tilde{C}(m_1)$ (perturbation de la fonction coût quadratique entre ces deux modèles) à $\delta m^T H \delta m$ (approximation quadratique pour $\delta m = m_2 - m_1$) on obtient deux résultats nettement différents :

$$\delta m^T H \delta m = 0.807 \quad \text{et} \quad \tilde{C}(m_2) - \tilde{C}(m_1) = 0.221 - 0.212 = 0.009. \tag{6.1}$$

Ce résultat met en évidence les limites de l'approche linéarisée et montre les limitations de l'approximation quadratique de la fonction coût non linéaire et donc les limitations des méthodes de quantification des incertitudes dans le cadre linéarisé.

L'approche linéarisée ne permet pas de délimiter l'ensemble des modèles possibles et nous mettons en relief ici la nécessité pour les cas complexes de réaliser une analyse non linéaire. L'optimisation sous contraintes peut être un moyen de tester des hypothèses géologiques et de délimiter l'espace des modèles admissibles : Delbos et al., 2004, ont montré que l'inversion tomographique avec contraintes sur la position des interfaces aux puits fournissait un modèle solution différent (Figure 6.2) du modèle solution de la Figure 6.1, ces deux modèles vérifiant les données temps de trajet avec la précision voulue. Elle permet aussi d'aller plus loin que l'approche linéarisée qui explore uniquement le voisinage du modèle solution qui peut être un minimum local parmi d'autres. Ce second modèle (Figure 6.2 n'appartient pas à la gamme de modèles possibles détectés par l'approche macro-paramètre d'après 6.1, pourtant une analyse linéarisée autour de ce modèle nous donne des résultats sur les incertitudes calculées équivalents à ceux du modèle de la figure 6.1.

	Volume	**Moyenne de la vitesse**
seabottom - tpal	$1414.4km^3$	2.47km/s
tpal - tchalk	$76.6km^3$	3.15km/s
tchalk - ichalk	$219.7km^3$	3.77km/s
ichalk - bchalk	$248.9km^3$	4.34km/s

TAB. 6.3 – Caractéristiques du modèle solution de la Figure 6.1. Les profondeurs moyennes des interfaces tpal, tchalk, ichalk, bchalk sont respectivement : $3.04km$, $3.29km$, $3.71km$ et $4.04km$.

		Macro-paramètres	
		B-spline	**DZ**
INTERFACE	layer1	1.8m	-
	tpal	1.3m	-
	tchalk	1.1m	-
	ichalk	2.2m	-
	bchlak	2.7m	-
VITESSE	vtert	8.4m/s	0.0042/s
	vpal	14.3m/s	-
	vichalk	12.4m/s	-
	vchalk	22.7m/s	-

TAB. 6.4 – Écarts types (racines carrées des termes diagonaux de $\tilde{C}'_m$) associés à différents macro-paramètres : moyenne de la fonction B-spline et moyenne du gradient vertical de vitesse. Cette analyse de l'incertitude est réalisée pour le modèle solution de la Figure 6.1.

	Macro-paramètre volume
seabottom - tpal	$206.4km^3$
tpal - tchalk	$10.1km^3$
tchalk - ichalk	$11.9km^3$
ichalk - bchalk	$4.3km^3$

TAB. 6.5 – Écarts types (racines carrées des termes diagonaux de $\tilde{C}'_m$) associés à différents macro-paramètres : épaisseur de la couche et volume de la couche. Cette analyse de l'incertitude est réalisée pour le modèle solution de la Figure 6.1 (voir tableau 6.3).

	$\frac{\partial vtert}{\partial Z}$	***vpal***	***vichalk***	***vchalk***	***layer1***	***tpal***	***tchalk***	***ichalk***	***bchalk***
$\frac{\partial vtert}{\partial Z}$	**0.0042/s**	-0.13	-0.01	0.0003	-0.93	0.17	0.09	0.04	0.03
vpal	-0.13	**14.3 m/s**	-0.10	0.003	-0.07	-0.69	0.02	-0.05	-0.03
vichalk	-0.01	-0.10	**12.4 m/s**	-0.24	-0.006	-0.05	-0.16	0.31	0.13
vchalk	-0.0003	-0.003	-0.24	**22.7 m/s**	-0.004	-0.01	-0.009	-0.11	0.33
layer1	-0.93	-0.07	-0.006	-0.004	**1.83 m**	0.11	0.08	0.03	0.02
tpal	0.17	-0.69	-0.05	-0.01	0.11	**1.3 m**	0.6	0.25	0.20
tchalk	0.09	0.02	-0.16	-0.009	0.08	0.6	**1.08 m**	0.25	0.20
ichalk	0.04	-0.05	0.31	-0.11	0.03	0.25	0.25	**2.24 m**	0.20
bchalk	0.03	-0.03	0.13	0.33	0.02	0.20	0.20	0.20	**2.69 m**

TAB. 6.6 – Écarts types (racines carrées des termes diagonaux de $\tilde{C}'_m$) associées à deux macro-paramètres : moyenne de la profondeur des interfaces layer1, tpal, tchalk, ichalk et bchalk, moyenne des vitesses vpal, vichalk et vchalk et du gradient vertical de vitesse dans le Tertiaire sur le domaine de définition et les corrélations normalisées entre vitesse et interface. Cette analyse de l'incertitude est réalisée pour le modèle solution de la Figure 6.1.

Chapitre 7

CONCLUSIONS

L'imagerie sismique de structures géologiques complexes nécessite la détermination précise de la distribution des vitesses de propagation dans le sous-sol. C'est pourquoi les géophysiciens de l'exploration pétrolière se sont intéressés de très près à la tomographie de réflexion de données sismiques qui permet la détermination d'un modèle de sous-sol à partir des temps de trajet des ondes sismiques se réfléchissant sur les interfaces géologiques. Le modèle solution de ce problème inverse est le modèle qui minimise une fonctionnelle de type moindre-carrés mesurant les écarts entre les temps de trajet mesurés et les temps de trajet calculés par tracé de rayons dans ce modèle (problème direct). Ce problème non linéaire est résolu par la méthode de Gauss-Newton, par linéarisations successives.

Dans le cas de structures géologiques complexes, l'accès au temps de trajet sur les sections sismiques peut être très difficile. En effet, le sous-sol est constitué d'une succession de dépôts aux caractéristiques lithologiques différentes, générant des réflexions des ondes sismiques. Il est souvent difficile de distinguer les évènements qui intéressent le géophysicien : repérer les réflexions pertinentes, savoir à quelles interfaces géologiques elles correspondent, repérer les discontinuités entre évènements sismiques, de même que les interférences entre évènements sismiques différents qui faussent l'interprétation. Ces temps de trajet observés, données d'entrée de la tomographie, sont donc entachés d'erreur. Par conséquent, les incertitudes sur les données et plus généralement la sous détermination du problème inverse (liée à l'ambiguïté vitesse/profondeur) entraînent des incertitudes sur le modèle solution de la tomographie de réflexion.

L'objet de cette thèse consiste en l'estimation de la confiance et plus précisément des incertitudes a posteriori sur le modèle solution obtenu par tomographie. Ces incertitudes dépendent évidemment fortement de la bonne estimation des incertitudes sur les données et de leur prise en compte dans la résolution du problème inverse, mais aussi de la densité et de la qualité des informations a priori sur le modèle introduites dans le problème inverse.

Une première partie du travail de thèse a été consacrée à une étude des différentes méthodes d'estimation des incertitudes a posteriori dans les problèmes inverses basées sur l'analyse de la matrice de covariance a posteriori (matrice de grande taille pour les modèles 3D) dans le cadre linéarisé, c'est à dire au voisinage de la solution. Nous avons pu constater que l'analyse classique de la matrice de covariance a posteriori (der Sluis et der Vost, 1987, Nolet et al., 1999) peut s'avérer délicate pour les applications de tomographie 3D qui nous intéressent ici : le calcul et l'interprétation de cette matrice restent difficiles pour les problèmes de grande taille. Pour des exemples de taille raisonnable, on a pu étudier le spectre de cette matrice via lequel on a évalué les directions de modèles les mieux (ou moins

bien) déterminées. Toutefois, la difficulté réside dans l'interprétation physique des vecteurs propres. Grâce à des expériences sur des modèles synthétiques, nous avons pu constater que pour des exemples de grande taille pour lesquels le calcul de l'inverse de la matrice hessienne est trop coûteux, on peut se restreindre à l'étude des composantes du modèle (ou à des combinaisons de paramètres décrivant le modèle) qui ont un intérêt pour l'explorateur (et non aux paramètres individuellement), par exemple, la position ou la pente d'une interface dans une région cible (Delprat-Jannaud, 1991). La difficulté est alors de définir les combinaisons de paramètres pertinents pour le problème géophysique posé.

Nous avons donc proposé deux approches adaptées à notre problème pour quantifier les incertitudes sur le modèle solution de la tomographie de réflexion : la méthode de simulation de modèles et l'approche macro-paramètre.

La première méthode permet d'échantillonner la fonction de densité de probabilité a posteriori et donc de simuler des modèles équiprobables. Les simulations fournissent directement des résultats interprétables, des modèles physiques. Toutefois, cette méthode est coûteuse en temps CPU et en stockage en raison de la décomposition de Choleski de la matrice de covariance a posteriori.

La deuxième méthode par son formalisme général permet de réduire l'analyse a posteriori aux quantités géologiques du modèle : nous évaluons les incertitudes sur des macro-paramètres (combinaisons linéaires des paramètres du modèle) qui ont un intérêt géologique. Cette méthode permet la manipulation de matrices réduites et donc devient faisable en 3D.

Une première application sur un exemple synthétique 3D de grande taille avec des macro-paramètres basiques a montré le potentiel de la méthode. La généralité du formalisme permet une large flexibilité pour la construction des macro-paramètres. Néanmoins, nous devrons garder à l'esprit que cette approche est seulement valide dans un voisinage du modèle solution (cadre linéarisé) et les cas complexes peuvent nécessiter une approche non-linéaire.

On a donc développé un formalisme permettant de calculer des incertitudes sur des quantités géologiques (et non pas sur les paramètres de b-splines) en sortie de tomographie. On peut se demander quelles sont les quantités qui intéresseraient les géologues ? On a proposé dans un premier temps, le calcul d'incertitude sur la profondeur des interfaces, sur leur pente, sur les valeurs de vitesse, les gradients, etc. Dans un deuxième temps, nous nous sommes intéressés à des macro-paramètres plus proches des préoccupations des géologues pétroliers tels que : le volume d'une couche (ou d'un piège pétrolier).

Nous avons également étudié les limites de ces méthodes de quantification des incertitudes et discuté du réalisme accordé aux incertitudes calculées. En effet, ces méthodes ont été testées dans le cadre linéarisé sous certaines hypothèses : paramétrisation fixée, poids de régularisation fixés, échantillonnage des données fixé, etc. L'idée a été ici de relaxer ces hypothèses et d'évaluer l'incidence sur les incertitudes calculées. Chaque étude est illustrée par une application sur un exemple synthétique 2D ou réaliste 3D.

Concernant l'influence de la paramétrisation et de la discrétisation du modèle, nous avons pu établir un résultat de convergence qui établit que les incertitudes sur les macro-paramètres convergent quand le pas de discrétisation tend vers 0 vers les valeurs obtenues pour le problème continu.

Concernant l'influence de l'échantillonnage des données, nous avons pu vérifier les résultats de Re-

nard, 2002, c'est à dire vérifier que le choix de la norme dans l'espace des données n'a pas d'incidence majeure sur les incertitudes calculées sous l'hypothèse que les résidus sont basse fréquence. Par conséquent, qu'il n'est pas nécessaire d'introduire la norme H^1 dans l'espace des données pour déterminer une solution au problème de tomographie de réflexion plus satisfaisante que celle obtenue à l'aide d'une norme L^2. Et finalement nous avons mis en évidence le fait que l'échantillonnage des données, pour la normeL^2, a peu d'influence sur les incertitudes calculées sous certaines hypothèses (on suppose un nombre suffisant de données et des résidus basses fréquences).

L'introduction de régularisation et/ou d'information a priori dans la formulation de moindres carrés est nécessaire pour bien poser le problème. En effet, l'inversion tomographique de données géophysiques (données sismiques, gravimétriques, électromagnétiques, etc.) peut être contrainte par l'information dont on dispose sur la géologie de la zone où la campagne géophysique a eu lieu. On introduit alors cette information, qui concerne les paramètres de l'espace des modèles, comme information a priori dans le problème inverse. En pratique, construire un modèle a priori est difficile et le terme de la fonction coût contenant l'information a priori, est souvent remplacé par des termes de régularisation composés des dérivées secondes du modèle. Dans ce cas, trouver la solution du problème régularisé revient à trouver le modèle le plus lisse qui vérifie les données avec la précision voulue.

L'introduction d'informations a priori est cruciale pour réduire l'indétermination. Il est en pratique difficile de modéliser les informations géologiques disponibles et les incertitudes associées à ces informations. C'est pourquoi, en pratique, il est courant d'utiliser une méthode de continuation ou méthode L-curve qui consiste à rechercher le modèle le plus régulier vérifiant les données avec la précision voulue. Le problème inverse associé est un problème inverse régularisé où les poids de régularisation choisis forts en début d'inversion sont diminués au fur et à mesure jusqu'à l'obtention d'un modèle vérifiant les données. Cette méthode permet d'éviter les problèmes d'instabilité liés à une régularisation trop faible.

Cependant, nous avons montré que, bien que les modèles obtenus par cette méthode soient satisfaisants, les incertitudes associées peuvent ne pas être réalistes, les poids de régularisation obtenus par cette méthode ne correspondant pas à une information a apriori réaliste.

L'introduction d'une norme plus «géologique» avec des poids dépendant de quantités plus physiques (introduction d'une covariance exponentielle dans l'espace des modèles) permet l'obtention d'incertitudes plus réalistes. Cependant, dans le cas où nous disposons de très peu d'information a priori ou d'une information a priori peu fiable, les poids résultants peuvent conduire à une inversion instable. Dans ce cas, la méthode L-curve sera utilisée pour obtenir une solution et pour l'estimation des incertitudes sur ce modèle solution, on construira la norme «géologique» avec des poids plus conformes à l'information a priori disponible.

Enfin, nous avons réalisé deux applications sur données réelles 2D et 3D. Ces deux applications nous ont permis de mettre en évidence les limites de l'approche linéarisée : cette approche est uniquement valide au voisinage du modèle solution et les cas complexes peuvent nécessiter une approche non-linéaire qui consiste à rechercher les minima locaux de la fonction coût. Une des pistes pour résoudre le problème non linéaire a été l'exploration expérimentale de l'ensemble des modèles admissibles par l'introduction de contraintes sur le modèle dans la résolution du problème inverse non linéaire afin de tester différentes hypothèses géologiques.

- Pour l'application 2D, nous avons testé différentes hypothèses sur les valeurs des paramètres

d'anisotropie pour lesquels l'approche linéarisée avait détecté de fortes incertitudes. En introduisant une contrainte sur la valeur de ces paramètres nous avons obtenu un modèle différent du modèle solution et qui vérifie aussi bien les données. De plus, ce modèle n'est pas un modèle détecté par l'approche linéarisée. On a donc pu remettre en cause l'approche linéarisée qui explore uniquement le voisinage du modèle solution qui peut être un minimum local parmi d'autres.

- Pour l'application 3D, on a introduit des données supplémentaires par le biais de l'optimisation sous contraintes. L'inversion tomographique avec contraintes sur la position des interfaces aux puits nous a fournit un modèle solution différent du modèle solution obtenu et qui vérifie les données temps de trajet avec la précision voulue. Ce modèle n'appartient pas à la gamme de modèles possibles détectés par l'approche linéarisée : les incertitude trouvées étaient faibles. Là encore nous avons constaté les limites de l'approche linéarisée.

Plusieurs pistes restent ouvertes pour de possibles développements des méthodes présentes dans cette thèse. Les travaux de cette thèse avaient pour but d'étudier la confiance que l'on peut avoir en le modèle solution de la tomographie de réflexion. Pour ce faire nous avons développé un formalisme permettant de calculer des incertitudes sur des quantités géologiques, à savoir l'approche macro-paramètre. On peut se demander quelles sont les quantités qui intéresseraient les géologues et envisager de nous intéresser à d'autres macro-paramètres proches des préoccupations des géologues pétroliers comme par exemple la fermeture du piège.

D'autre part, nous pourrions envisager d'autres méthodes pour développer l'approche non linéaire. Ainsi, l'approche non linéaire consistant à explorer l'espace des modèles admissibles en recherchant les minima locaux de la fonction coût, la méthode des plans d'expériences (Zabalza-Mezghani, 2000) pourrait par exemple permettre la résolution de ce type de problème de recherche de minimum. En effet, la construction d'un modèle analytique simplifié peut permettre d'éviter des simulations coûteuses : réaliser des simulations correspondant à différentes perturbations de modèle nécessite pour chaque simulation un tracé de rayons, ce qui est évidemment coûteux. Ces techniques de plans d'expériences suscitent un grand intérêt dans le domaine pétrolier, notamment pour la quantification des incertitudes sur les prévisions de production. Elles restent donc aussi applicables dans notre cas mais avec un bémol toutefois puisqu'on est très vite limité par le nombre élevé de paramètres des modèles utilisés en tomographie, il faudrait donc là aussi se limiter aux paramètres pertinents comme pour les macro-paramètres.

Annexe A

MODÈLE SYNTHÉTIQUE 1

Dans cette annexe est présenté un modèle synthétique mono-couche qui est utilisé au cours des chapitres pour valider les méthodes présentées et les expériences proposées.

A.1 Présentation du modèle

Le modèle étudié est un modèle 3D composé d'une interface régulière, s'étendant sur $18km$ dans la direction x et sur $9km$ dans la direction y et définie grâce à 8 (respectivement 5) paramètres de B-spline dans la direction des x (respectivement dans la direction des y) et d'une vitesse présentant des variations latérales ($v(x, y)$, il n'y a pas de variations verticales) définie par 6 (respectivement 5) paramètres de B-spline dans la direction x (respectivement dans la direction des y).

On modélise les données synthétiques par un tracé de rayons sur le modèle exact avec une géométrie d'acquisition décrite par la Figure A.2.

On ajoute à ces temps de trajet exacts un bruit blanc (décrit par une loi de probabilité Gaussienne centrée) d'écart type $5ms$ (Figure A.3).

L'inversion de ces données bruitées grâce à notre code de tomographie de réflexion *jerry* (voir la section 3.3.2.3 du chapitre 3 pour une description de l'outil de tomographie et la méthodologie d'utilisation) donne un modèle solution proche du modèle exact (Figure A.4) avec un résidu moyen RMS de $5ms$ et un résidu maximum de $20.47ms$, résultats qui sont en accord avec le bruit introduit sur les données.

Mais deux autres modèles solutions pourraient être obtenus avec un résidu moyen RMS équivalent (Figures A.5 et A.6). Sans connaître le modèle exact ou sans information a priori discriminante sur le modèle, on ne pourrait pas choisir entre ces trois modèles solutions : ils vérifient tous les données avec la précision espérée. Cet exemple illustre la nécessité d'être en mesure d'estimer l'ensemble des modèles possibles. Si les incertitudes obtenues sur le modèle solution sont jugées trop élevées, nous devrons introduire d'autres données ou information a priori sur le modèle pour les réduire.

Notons que nous avons aussi réalisé une inversion en ôtant des données (notamment au centre du modèle : partie qui correspond au dôme) et que nous avons obtenu un modèle solution quasi-identique (Figures A.7) avec des résidus équivalents.

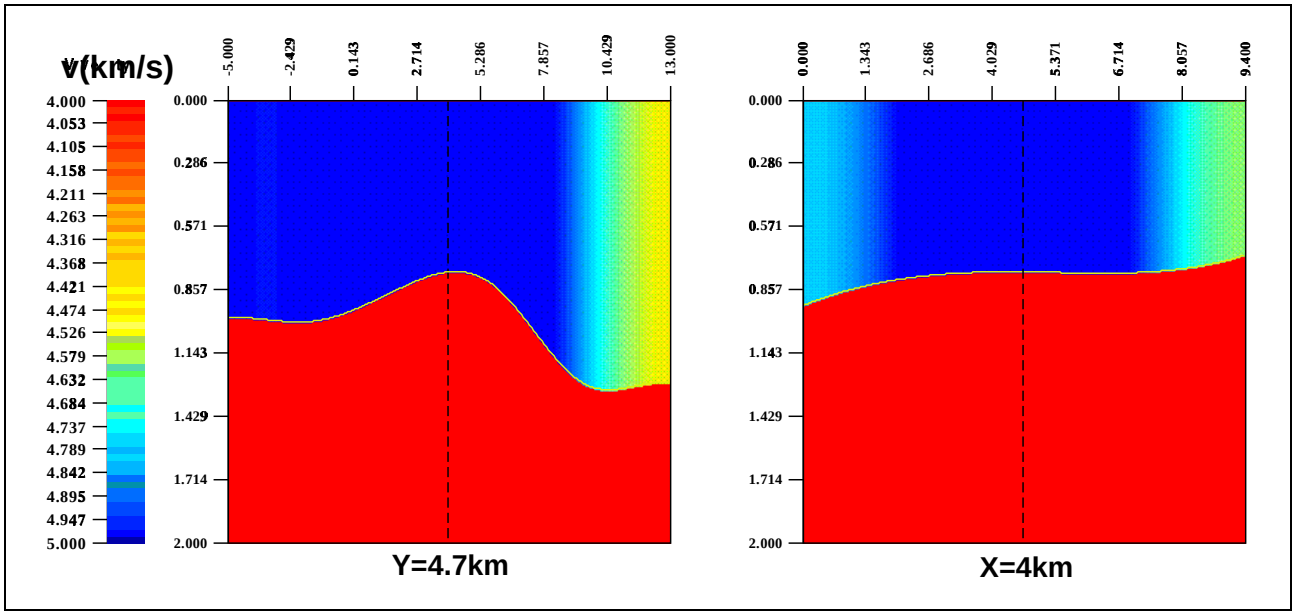

FIG. A.1 – Coupes verticales du premier modèle de vitesse 3D exact (Gauche : le long de l'axe des x, droite : le long de l'axe des y).

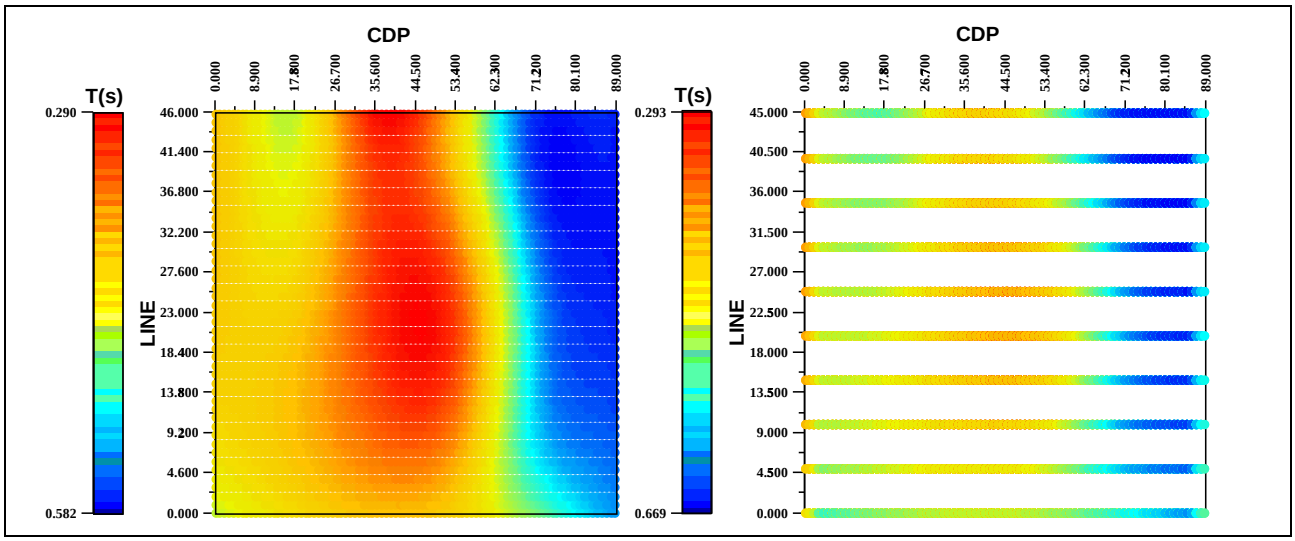

FIG. A.2 – Données temps de trajet générées par tracé de rayons sur le modèle exact de la figure A.1. Gauche : temps de trajet zéro-offset sur une grille $200m \times 200m$. Droite : temps de trajet multi-offset pour des lignes espacées de $1km$ avec des CMP (de l'anglais *Common Middle Point* : point milieu commun) espacés de $200m$ et un pas d'offset de $100m$.

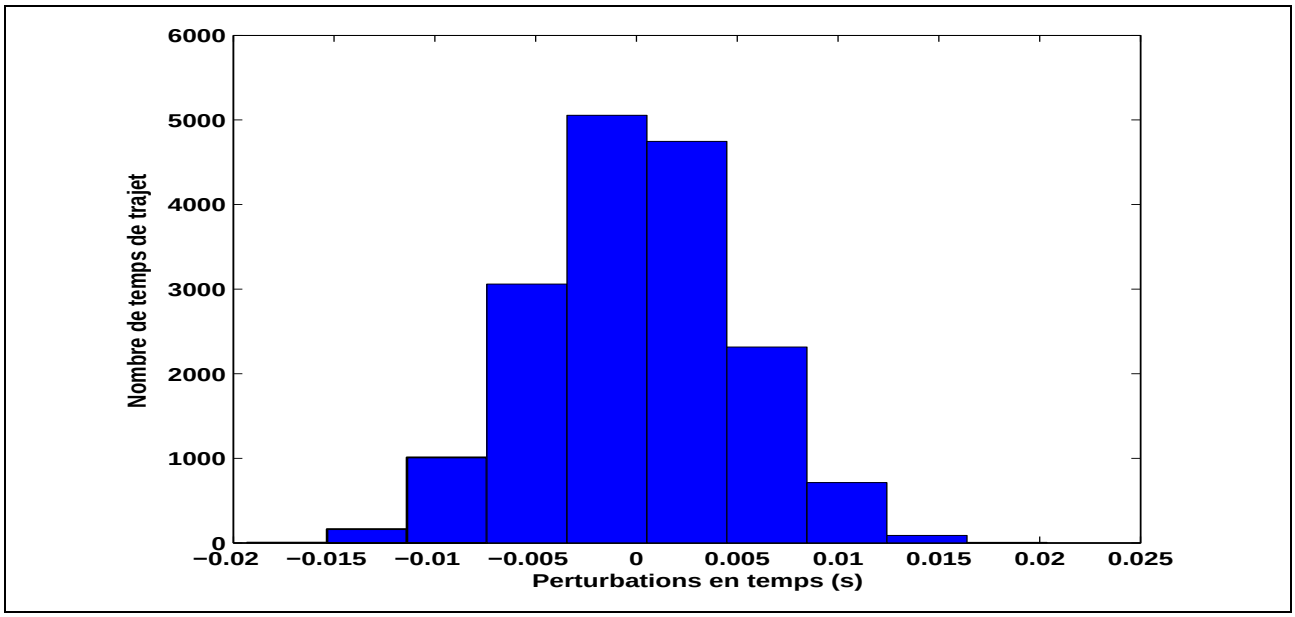

FIG. A.3 – Histogramme des perturbations Gaussiennes non corrélées ajoutées aux données de temps de trajet de la figure A.2. L'écart type associé est de $5ms$.

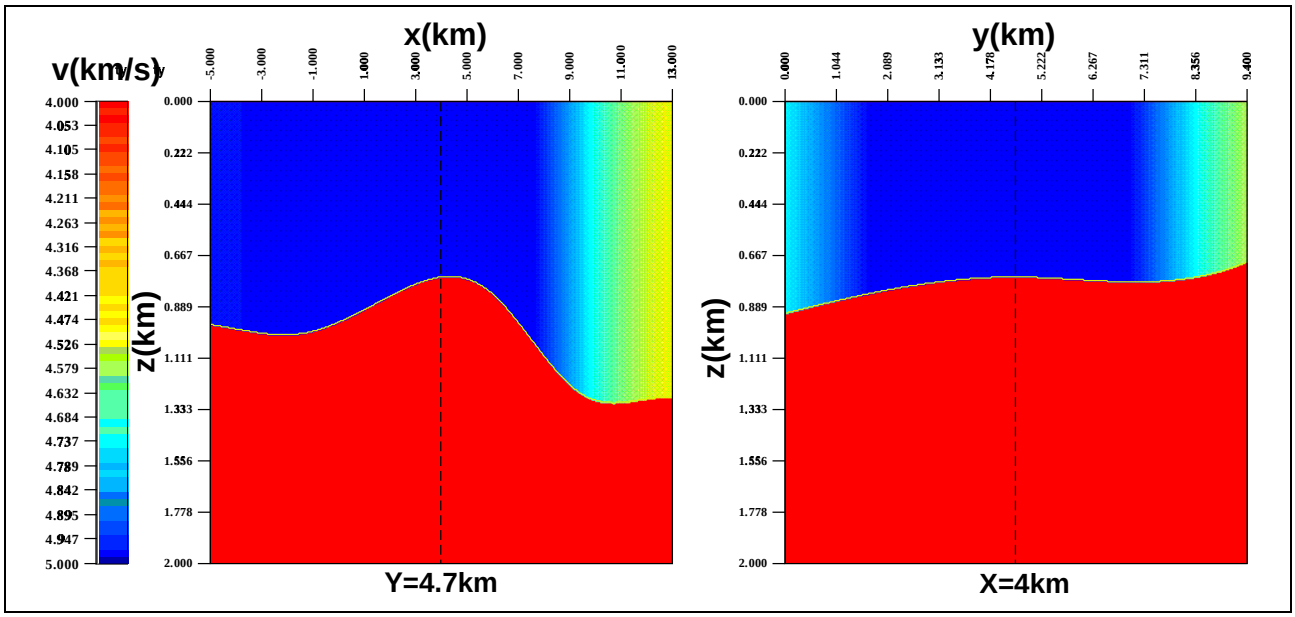

FIG. A.4 – Coupes verticales du modèle solution obtenu par *jerry* pour le premier exemple (Gauche : le long de l'axe des x, droite : le long de l'axe des y). RMS des résidus des temps de trajet $= 5ms$, valeur maximum des résidus des temps de trajet $= 20.47ms$.

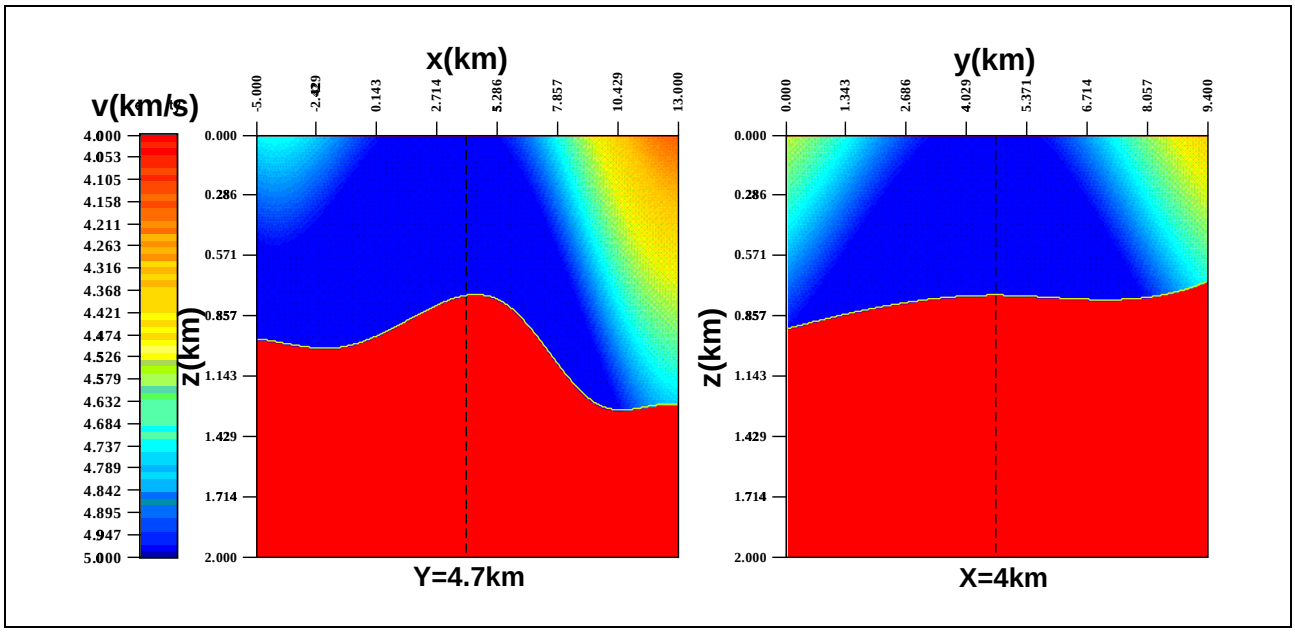

FIG. A.5 – Coupes verticales dans un autre modèle solution obtenu par *jerry* pour le premier exemple avec un gradient vertical fixé égal à $0.5/s$ (alors que la valeur exacte est 0) (Gauche : le long de l'axe des x, droite : le long de l'axe des y). RMS des résidus des temps de trajet $= 5ms$, valeur maximum des résidus des temps de trajet $= 20.46ms$.

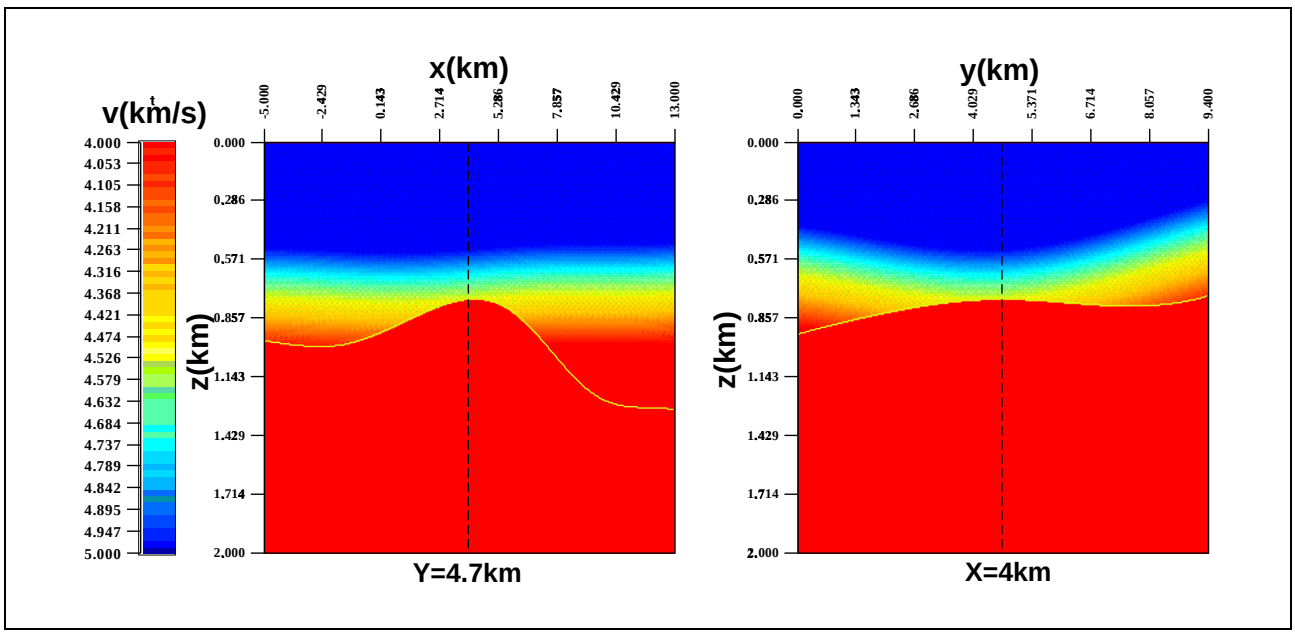

FIG. A.6 – Coupes verticales dans un autre modèle solution obtenu par *jerry* pour le premier exemple avec une paramétrisation erronée pour la vitesse : $v(x, y, z)$ à la place de $v(x, y)$. (Gauche : le long de l'axe des x, droite : le long de l'axe des y). RMS des résidus des temps de trajet $= 5.5ms$, valeur maximum des résidus des temps de trajet $= 22ms$.

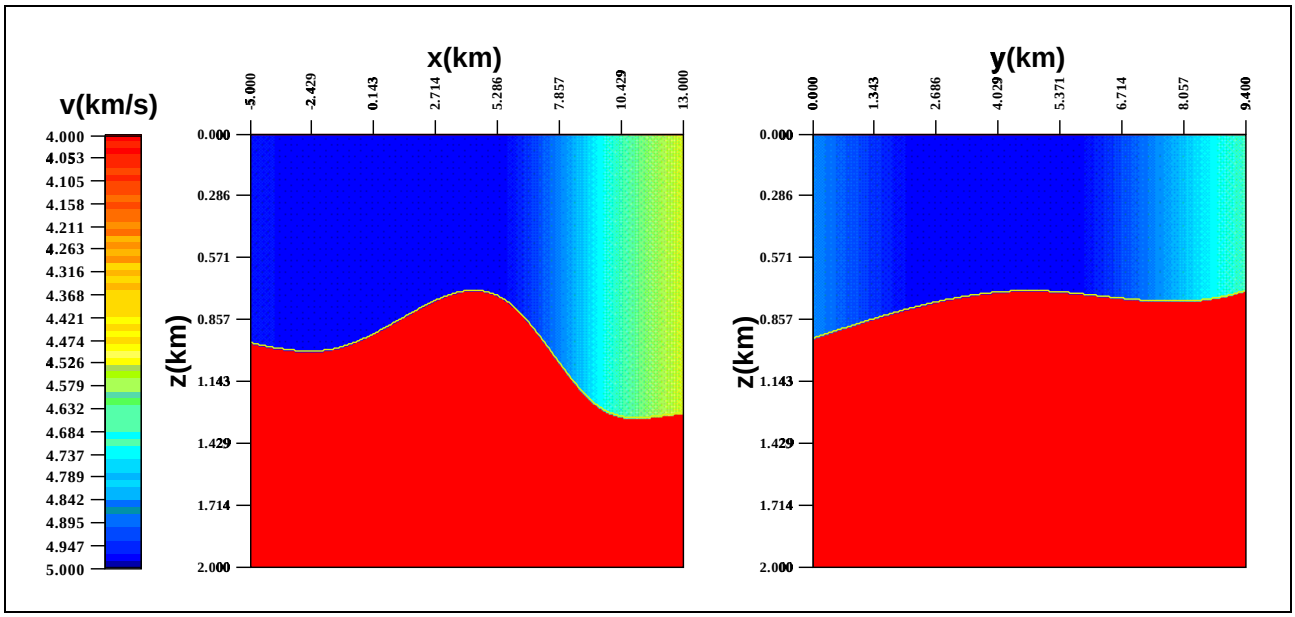

FIG. A.7 – Coupes verticales du modèle solution obtenu par *jerry* pour le premier exemple (Gauche : le long de l'axe des x, droite : le long de l'axe des y). RMS des résidus des temps de trajet $= 5.15ms$, valeur maximum des résidus des temps de trajet $= 21.04ms$.

Annexe B

MODÈLE SYNTHÉTIQUE 2

Dans cette annexe est présenté un modèle synthétique multi-couches 3D qui est utilisé au cours des chapitres pour valider les méthodes présentées et les expériences proposées. Ce modèle est un modèle 3D réaliste construit à partir d'un jeu de données réelles (voir Broto et Ehinger, 1998).

B.1 Présentation du modèle

Le modèle étudié ici est un modèle 3D réaliste composé de huit interfaces avec des vitesses de couche à gradient vertical constant et fixé pour les deux premières couches superficielles et une vitesse 3D $v(x, y, z)$ pour la couche plus profonde. Ce modèle comporte 6869 paramètres dont 4802 pour les interfaces et 2067 pour les vitesses.

On a modélisé par tracé de rayons 50844 temps de trajet multi-offsets ($0.197km <$ offset $< 4km$). L'illumination du modèle (c'est à dire la distribution des points d'impact) est décrite par la figure B.2. On ajoute à ces temps de trajet exacts un bruit blanc (décrit par une loi de probabilité Gaussienne centrée) d'écart type $5ms$.

Ces données sont inversées par le biais d'une approche layer-stripping (méthodologie décrite plus bas) suivie d'une inversion globale. on obtient un modèle solution proche du modèle exact, uniquement dans les zones illuminées (Figure B.3) avec un résidu moyen RMS proche de $5ms$ pour chaque interface et un résidu maximum autour de $20ms$ (les résultats sont récapitulés dans le tableau (B.1)), résultats qui sont en accord avec les incertitudes sur les données.

B.2 La méthode d'inversion utilisée

Pour cet exemple, au lieu d'utiliser directement une approche globale, on réalise une approche «layer stripping» («couche par couche» en français), c'est à dire qu'on utilise la méthode L-curve pour chaque couche.

Pour ce type d'approche, on détermine par inversion les paramètres associés à une couche (plus exactement à un bloc de vitesse), puis ce résultat est fixé et on passe à l'inversion des temps de trajet associés à la couche suivante.

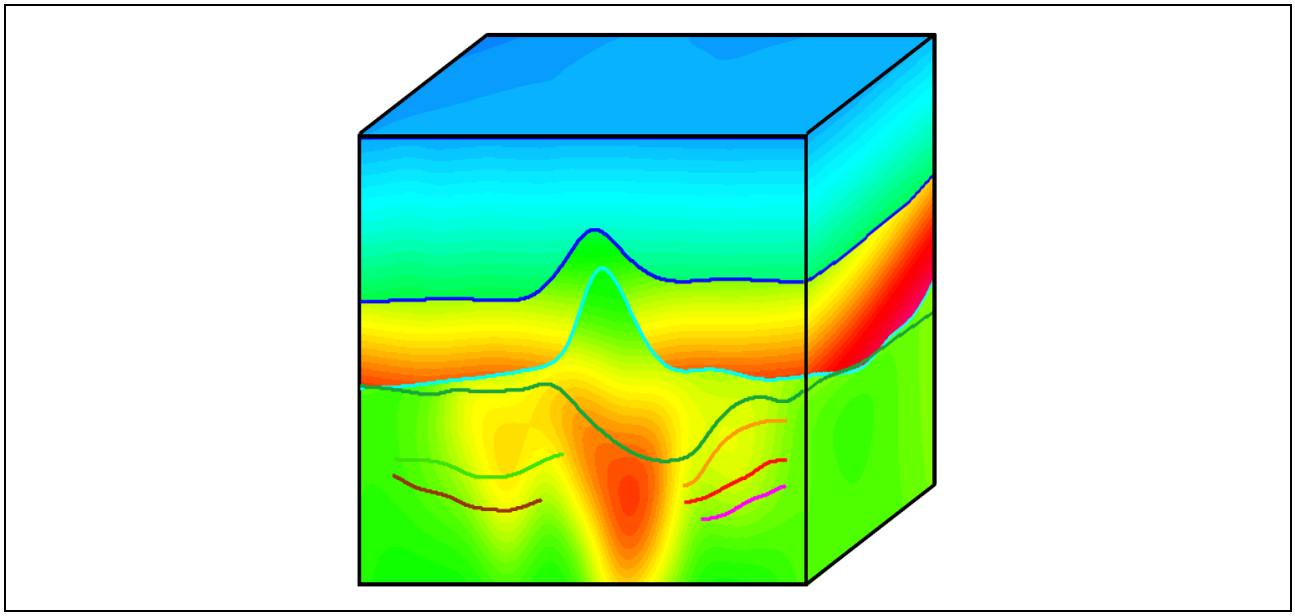

FIG. B.1 – Cube 3D du modèle 2 de vitesse 3D exact ($x_{min} = -2km$ et $x_{max} = 14km$, $y_{min} = -1km$ et $y_{max} = 11km$, $z_{min} = 0km$ et $z_{max} = 6km$).

Cette approche permet une inversion plus rapide (chaque inversion est moins lourde : on a moins de paramètres). Toutefois, elle peut présenter un inconvénient. En effet, s'il existe par exemple une zone du sous-sol pour laquelle l'information cinématique est inexistante, le modèle de vitesse obtenu par tomographie de réflexion ne sera déterminé dans cette zone que par l'information a priori introduite ; par conséquent, si aucune information a priori suffisamment précise n'a été spécifée pour cette zone, il y a un risque que le modèle de vitesse soit erroné dans cette région, et que cette erreur se propage sur les couches sous-jacentes.

C'est pourquoi, une fois toutes les couches déterminées par l'approche «layer stripping», nous lançons une inversion globale avec comme modèle initial celui obtenu par l'approche «layer stripping», de sorte que les rais qui traversent les zones de certaines couches non éclairées en réflexion puissent contribuer à leur détermination. Le résultat final est représenté sur la Figure B.3.

	Résidu moyen RMS	**Résidu maximum Max**
interface 1	5ms	20ms
interface 2	5.2ms	18.8ms
interface 3	5.5ms	19.9ms
interface 4	5.6ms	20ms
interface 5	5.4ms	19ms
interface 6	5.2ms	19ms
interface 7	5.3ms	19.2ms
interface 8	5.8ms	20ms

TAB. B.1 – Tableau récapitulatif des résidus moyen RMS et résidus maximum pour le modèle solution (B.3)

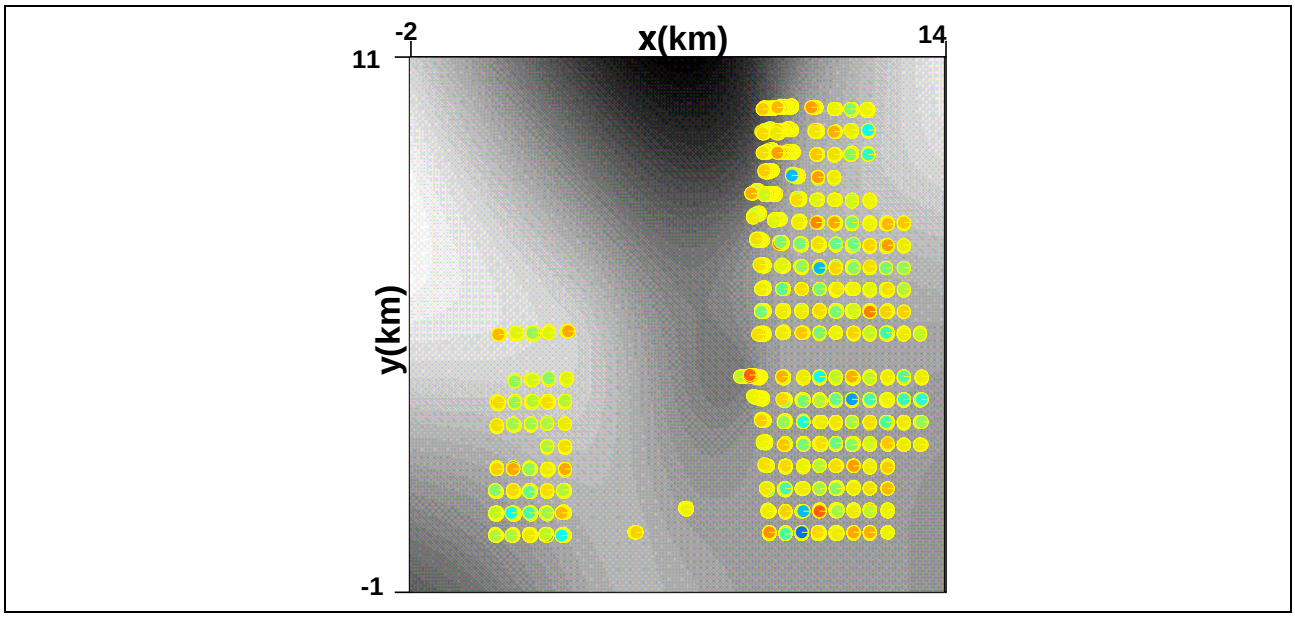

FIG. B.2 – Illumination du domaine. Les points représentent les points d'impact des rayons sur l'interface ($x_{min} = -2km$ et $x_{max} = 14km$, $y_{min} = -1km$ et $y_{max} = 11km$, $z_{min} = 0km$ et $z_{max} = 6km$).

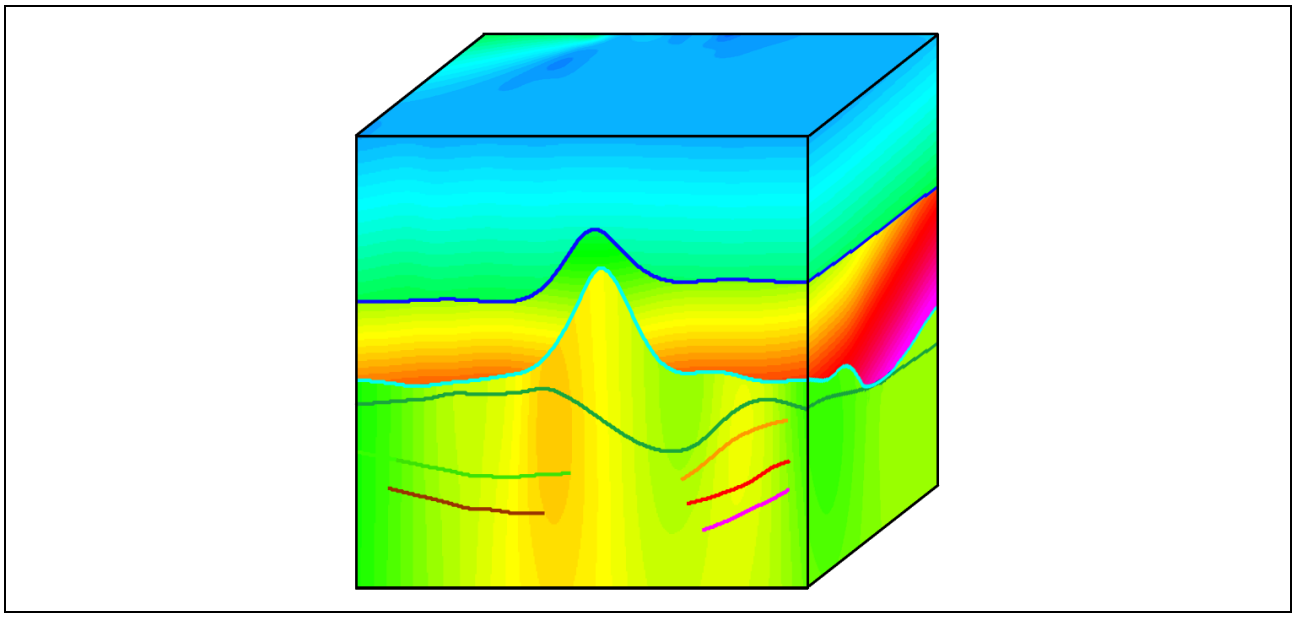

FIG. B.3 – Cube 3D représentant le modèle solution associé au modèle exact de la figure B.1 obtenu par une inversion **"layer stripping + inversion globale"** ($x_{min} = -2km$ et $x_{max} = 14km$, $y_{min} = -1km$ et $y_{max} = 11km$, $z_{min} = 0km$ et $z_{max} = 6km$).

Annexe C

FONCTIONS B-SPLINES

Une fonction B-spline cubique est générée par la sommation de polynômes du troisième degré. La forme générale d'une fonction B-spline cubique est :

$$f(x) = \sum_{m=1}^{N} \alpha_m \beta_m(x), \tag{C.1}$$

où les α_m sont les paramètres de B-spline, N est le nombre de paramètres de B-spline et les β_m sont les fonctions de base de la représentation B-spline (Figure C.1).

Les fonctions de base de la représentation B-spline cubique sont construites par morceaux et sont données par :

$$\beta_m(x) = \begin{cases} 0 & x \leq x_m \\ b_1(x_b) & x_m \leq x \leq x_{m+1} \\ b_2(x_b) & x_{m+1} \leq x \leq x_{m+2} \\ b_3(x_b) & x_{m+2} \leq x \leq x_{m+3} \\ b_4(x_b) & x_{m+3} \leq x \leq x_{m+4} \\ 0 & x_{m+4} \leq x, \end{cases} \tag{C.2}$$

où $x_b = \frac{x - x_m}{x_{m+1} - x_m}$ est l'abscisse relative ($0 \leq x_b \leq 1$) d'un point dans l'intervalle $[x_m, x_{m+1}]$ et

$$\begin{array}{rcl} b_1(x_b) & = & \frac{1}{6}x_b^3 \\ b_2(x_b) & = & -\frac{1}{2}x_b^3 + \frac{1}{2}x_b^2 + \frac{1}{2}x_b + \frac{1}{6} \\ b_3(x_b) & = & b_2(1 - x_b) \\ b_4(x_b) & = & b_1(1 - x_b). \end{array} \tag{C.3}$$

Par construction, les dérivées premières et secondes d'une fonction B-spline cubique sont continues. De plus, les fonctions de base des fonctions B-spline cubiques ont un support fini et égal à $[x_m, x_{m+4}]$. Une fonction B-spline peut donc être évaluée localement, en utilisant seulement quatre fonctions de base et quatre paramètres.

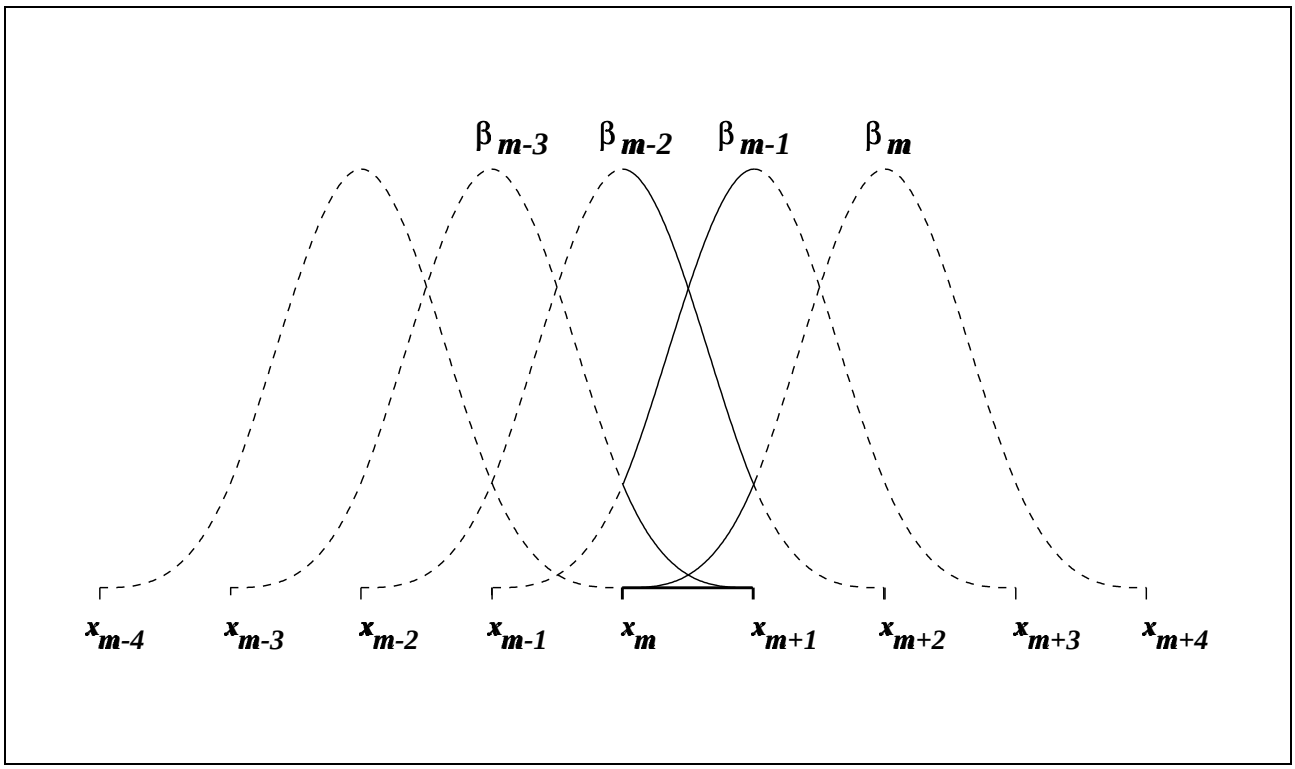

FIG. C.1 – Fonctions de base de la représentation B-spline cubique. Seules les parties en trait plein des quatre fonctions de base voisines β_{m-3}, β_{m-2}, β_{m-1}, et β_m interviennent dans l'évaluation de la fonction B-spline dans l'intervalle $[x_m, x_{m+1}]$.

Annexe D

CONSTRUCTION DE LA NORME INTERVENANT DANS LE TERME GÉOLOGIQUE : CAS ANISOTROPE

Lors de la construction du modèle a priori, la connaissance du vrai modèle en certains points peut-être plus précise qu'ailleurs. Par exemple, si on dispose de mesures dans certaines zones du modèle, ces informations sont précises et par conséquent, la variance en ces points devrait être plus faible qu'ailleurs. De la même façon, on peut envisager que les corrélations des incertitudes sur le modèle a priori soient plus faibles en certains points du milieu, ou que ces corrélations soient plus importantes dans certaines directions. Ces considérations vont nous amener à introduire des variations des variances et des longueurs de corrélations dans la fonction de covariance. Nous construisons donc une nouvelle fonction de covariance dérivée de la fonction de covariance exponentielle. Son expression est :

$$\mathcal{C}_m(x_1, x_2) = \sigma^2 \exp\left(- \parallel \Lambda(x_1) - \Lambda(x_2) \parallel\right), \tag{D.1}$$

où σ est la variance et $\Lambda(x_1)$ et $\Lambda(x_2)$ sont les longueurs de corrélations selon les deux directions principales d'anisotropie.

Nous cherchons donc à inverser l'opérateur de covariance (noté $\mathbf{C}$) dont le noyau est défini par 3.20.

Comme précédemment, on considère deux modèles m_1 et m_2 de $\mathcal{M}$ tels que m_1 soit l'image de m_2 par $\mathbf{C}$:

$$m_1(u) = \int_\Omega \sigma^2 \exp\left(- \parallel \Lambda(u) - \Lambda(v) \parallel\right) m_2(v) dv, \tag{D.2}$$

où $u \in \Omega$ est le point de coordonnées (x, y) avec $\Omega = \Omega_x \times \Omega_y$ en 2D et $\Omega = \Omega_x \times \Omega_y \times \Omega_z$ en 3D. On peut aussi écrire cette relation sous la forme :

$$\frac{m_1(u)}{\sigma^2} = \int_\Omega \exp\left(- \parallel \Lambda(u) - \Lambda(v) \parallel\right) m_2(v) dv. \tag{D.3}$$

Ensuite on pose :

$$\tilde{m}_1(u) = \frac{m_1(u)}{\sigma^2}, \tag{D.4}$$

et on obtient :

$$\tilde{m}_1(u) = \int_\Omega \exp\left(- \parallel \Lambda(u) - \Lambda(v) \parallel\right) m_2(v) dv. \tag{D.5}$$

On procède à ce stade au changement de variable suivant :

$$\begin{aligned} s &= \Lambda(u) \\ ds &= J(u)du \\ t &= \Lambda(v) \\ dt &= J(v)dv, \end{aligned} \tag{D.6}$$

où J est le jacobien de cette transformation. Notons alors :

$$\begin{aligned} U(s) &= \tilde{m}_1(u) \\ V(t) &= m_2(v). \end{aligned} \tag{D.7}$$

On a alors la relation suivante entre U et V :

$$U(s) = \int_\Omega \exp\left(- \parallel s - t \parallel\right) \frac{V(t)}{J(u(t))} dt. \tag{D.8}$$

Et finalement en posant :

$$\tilde{V}(t) = \frac{V(t)}{J(u(t))}, \tag{D.9}$$

on a :

$$U(s) = \int_\Omega \exp\left(- \parallel s - t \parallel\right) \tilde{V}(t) dt. \tag{D.10}$$

Ce qui signifie que U est l'image de $\tilde{V}$ par l'opérateur de covariance de longueur de corrélation et de variance égales à 1. Cet opérateur est inversible et en général son inverse est un opérateur différentiel (Dautray et Lions, 1984). On a donc :

$$\tilde{V}(s) = \sum_\alpha a_\alpha U(s)^{(\alpha)}, \tag{D.11}$$

où les α sont des entiers, les a_α sont des constantes réelles, et $U(s)^{(\alpha)}$ est la dérivée α-ième de U. Pour revenir aux variables initiales, on a successivement :

$$\begin{aligned} V(s) &= J(u(t)) \sum_\alpha a_\alpha U(s)^{(\alpha)} \\ m_2(u) &= J(u) \sum_\alpha a_\alpha \frac{d^\alpha \tilde{m}_1(u(s))}{ds^\alpha}, \end{aligned} \tag{D.12}$$

et finalement :

$$m_2(u) = \frac{J(u)}{\sigma^2} \sum_\alpha a_\alpha \frac{d^\alpha m_1(u(s))}{ds^\alpha}. \tag{D.13}$$

Cette dernière relation exprimant m_2 en fonction de m_1 est donc l'expression du noyau de l'opérateur inverse.
La norme associée à l'inverse de l'opérateur de covariance exponentielle est :

$$\parallel m_1 \parallel^2_{C_m^{-1}} = \int_\Omega m_1(u) m_2(u) du, \tag{D.14}$$

et avec les notations précédentes, on obtient :

$$\| m_1 \|^2_{C_m^{-1}} = \int_\Omega \frac{m_1(u)}{\sigma^2} \sum_\alpha a_\alpha \frac{d^\alpha}{ds^\alpha} (m_1(u))\, J du, \tag{D.15}$$

ou encore :

$$\boxed{\| m_1 \|^2_{C_m^{-1}} = \int_{\Lambda(\Omega)} \left(U(s) \sum_\alpha a_\alpha U(s)^\alpha \right) ds.} \tag{D.16}$$

Les expressions de cette norme selon la dimension de Ω (en 2D, $\Omega = \Omega_x \times \Omega_y$ et en 3D, $\Omega = \Omega_x \times \Omega_y \times \Omega_z$) sont respectivement :

$$\begin{aligned} \| m_1 \|^2_{\mathbf{C}_m^{-1}} |_{2D} &= \frac{1}{2\pi\sigma^2} \int_\Omega \frac{1}{\lambda_x \lambda_y} m_1(u)^2 du \\ &+ \frac{1}{\sigma^2 \pi} \int_\Omega \lambda_x \lambda_y \left(\frac{1}{\lambda_y^2} \left(\frac{\partial\, m_1}{\partial x} \right)^2 + \frac{1}{\lambda_x^2} \left(\frac{\partial m_1}{\partial y} \right)^2 \right) du \\ &+ \frac{1}{2\pi\sigma^2} \int_\Omega (\lambda_x \lambda_y)^3 \left(\frac{1}{\lambda_y^2} \frac{\partial^2 m_1}{\partial x^2} + \frac{1}{\lambda_x^2} \frac{\partial^2 m_1}{\partial y^2} \right)^2 du \end{aligned} \tag{D.17}$$

avec $\Omega = \Omega_x \times \Omega_y$ et λ_x et λ_y les longueurs de corrélations selon les deux directions principales d'anisotropie.
Et :

$$\begin{aligned} \| m_1 \|^2_{\mathbf{C}_m^{-1}} |_{3D} &= \frac{1}{8\pi\sigma^2} \int_\Omega \frac{1}{\lambda_x \lambda_y \lambda_z} m_1(u)^2 du \\ &+ \frac{1}{4\pi\sigma^2} \int_\Omega \lambda_x \lambda_y \lambda_z \left(\frac{1}{(\lambda_y \lambda_z)^2} \left(\frac{\partial m_1}{\partial x} \right)^2 + \frac{1}{(\lambda_x \lambda_z)^2} \left(\frac{\partial m_1}{\partial y} \right)^2 + \frac{1}{(\lambda_x \lambda_y)^2} \left(\frac{\partial m_1}{\partial z} \right)^2 \right) du \\ &+ \frac{1}{8\pi\sigma^2} \int_\Omega (\lambda_x \lambda_y \lambda_z)^3 \left(\frac{1}{(\lambda_y \lambda_z)^2} \frac{\partial^2 m_1}{\partial x^2} + \frac{1}{(\lambda_x \lambda_z)^2} \frac{\partial^2 m_1}{\partial y^2} + \frac{1}{(\lambda_x \lambda_y)^2} \frac{\partial^2 m_1}{\partial z} \right)^2 du \end{aligned} \tag{D.18}$$

avec $\Omega = \Omega_x \times \Omega_y \times \Omega_z$ et λ_x, λ_y et λ_z les longueurs de corrélations selon les trois directions principales d'anisotropie.

D.1 Calcul de l'expression de la norme en 2D

Pour traîter le cas 2D, nous utilisons l'approximation faite dans le cas isotrope (équation 3.21 de la section 3.3.2.2 du chapitre 3), c'est à dire que nous supposons :

$$\tilde{V}(s) = \frac{1}{2\pi} \left(U(s) - 2\Delta U(s) + \Delta\Delta U(s) \right). \tag{D.19}$$

et donc :

$$\| m_1 \|^2_{\mathbf{C}_M^{-1}} |_{2D} = \frac{1}{2\pi} \int_{\Lambda(\Omega)} \left(U(s)^2 + 2 \| \text{grad}\, U(s) \|^2 + (\Delta U(s))^2 \right) ds, \tag{D.20}$$

avec $\Omega = \Omega_x \times \Omega_y$.

On effectue maintenant le changement de variable $s = \Lambda(u) = (\Lambda_x(u), \Lambda_y(u))^T$ avec $u = (x, y)$. En effet, comme pour l'expression de la norme dans le cas 2D ou 3D dans le cas où on se place en coordonnées polaires (on définit un λ et un σ pour tout le domaine et pour toutes les directions), on peut choisir de définir des directions d'anisotropie pour les longueurs de corrélations. On pose alors :

$$\Lambda(u) = \begin{pmatrix} \frac{x\cos\theta + y\sin\theta}{\lambda_x} \\ \frac{-x\sin\theta + y\cos\theta}{\lambda_y} \end{pmatrix}, \tag{D.21}$$

avec λ_x et λ_y les longueurs de corrélations selon les deux directions principales d'anisotropie.

La matrice jacobienne de cette transformation s'écrit alors :

$$J = \begin{pmatrix} \frac{\partial\Lambda_x}{\partial x} & \frac{\partial\Lambda_x}{\partial y} \\ \frac{\partial\Lambda_y}{\partial x} & \frac{\partial\Lambda_y}{\partial y} \end{pmatrix} = \begin{pmatrix} \frac{\cos\theta}{\lambda_x} & \frac{\sin\theta}{\lambda_x} \\ \frac{-\sin\theta}{\lambda_y} & \frac{\cos\theta}{\lambda_y} \end{pmatrix}, \tag{D.22}$$

avec λ_x et λ_y les longueurs de corrélations selon les deux directions principales d'anisotropie.

On calcule le déterminant de J :

$$\begin{aligned} \det J &= \frac{\partial\Lambda_x}{\partial x}\frac{\partial\Lambda_y}{\partial y} - \frac{\partial\Lambda_y}{\partial x}\frac{\partial\Lambda_x}{\partial y} \\ &= \frac{1}{\lambda_x\lambda_y}. \end{aligned} \tag{D.23}$$

On a $\det J \neq 0$, donc J est inversible et J^{-1} est la matrice du changement de variable qui donne u en fonction de s. Les formules d'inversion d'une matrice 2×2 ($J^{-1} = \frac{1}{\det J}(\text{com } J)^T$) donnent alors :

$$J^{-1} = \frac{1}{\det J(u)} \begin{pmatrix} \frac{\partial\Lambda_y}{\partial y} & -\frac{\partial\Lambda_x}{\partial y} \\ -\frac{\partial\Lambda_y}{\partial x} & \frac{\partial\Lambda_x}{\partial x} \end{pmatrix} = \lambda_x\lambda_y \begin{pmatrix} \frac{\cos\theta}{\lambda_y} & \frac{-\sin\theta}{\lambda_x} \\ \frac{\sin\theta}{\lambda_y} & \frac{\cos\theta}{\lambda_x} \end{pmatrix}. \tag{D.24}$$

On a ensuite :

$$\text{grad } U(s) = J^{-1}\text{grad } \tilde{m}_1(u), \tag{D.25}$$

ou encore :

$$\text{grad } U(s) = \frac{1}{\det J(u)} \begin{pmatrix} \frac{\partial\Lambda_y}{\partial y}\frac{\partial\tilde{m}_1}{\partial x} - \frac{\partial\Lambda_x}{\partial y}\frac{\partial\tilde{m}_1}{\partial y} \\ -\frac{\partial\Lambda_y}{\partial x}\frac{\partial\tilde{m}_1}{\partial x} + \frac{\partial\Lambda_x}{\partial x}\frac{\partial\tilde{m}_1}{\partial y} \end{pmatrix} = \lambda_x\lambda_y \begin{pmatrix} \frac{\cos\theta}{\lambda_y}\frac{\partial\tilde{m}_1}{\partial x} + \frac{-\sin\theta}{\lambda_x}\frac{\partial\tilde{m}_1}{\partial y} \\ \frac{\sin\theta}{\lambda_y}\frac{\partial\tilde{m}_1}{\partial x} + \frac{\cos\theta}{\lambda_x}\frac{\partial\tilde{m}_1}{\partial y} \end{pmatrix}. \tag{D.26}$$

Donc :

$$\begin{aligned} \| \text{grad } U(s) \|^2 &= \frac{1}{\det J(u)^2}\left(\frac{\partial\Lambda_y}{\partial y}\frac{\partial\tilde{m}_1}{\partial x} - \frac{\partial\Lambda_x}{\partial y}\frac{\partial\tilde{m}_1}{\partial y}\right)^2 + \frac{1}{\det J(u)^2}\left(-\frac{\partial\Lambda_y}{\partial x}\frac{\partial\tilde{m}_1}{\partial x} + \frac{\partial\Lambda_x}{\partial x}\frac{\partial\tilde{m}_1}{\partial y}\right)^2 \\ &= (\lambda_x\lambda_y)^2\left(\frac{\cos\theta}{\lambda_y}\frac{\partial\tilde{m}_1}{\partial x} + \frac{-\sin\theta}{\lambda_x}\frac{\partial\tilde{m}_1}{\partial y}\right)^2 + (\lambda_x\lambda_y)^2\left(\frac{\sin\theta}{\lambda_y}\frac{\partial\tilde{m}_1}{\partial x} + \frac{\cos\theta}{\lambda_x}\frac{\partial\tilde{m}_1}{\partial y}\right)^2 \end{aligned} \tag{D.27}$$

On peut définir les variables suivantes :

$$a_{11} = \left(\frac{\partial\Lambda_y}{\partial y}\right)^2 + \left(\frac{\partial\Lambda_y}{\partial x}\right)^2 = \left(\frac{\cos\theta}{\lambda_y}\right)^2 + \left(\frac{\sin\theta}{\lambda_y}\right)^2 = \frac{1}{\lambda_y^2}, \tag{D.28}$$

$$a_{22} = \left(\frac{\partial\Lambda_x}{\partial y}\right)^2 + \left(\frac{\partial\Lambda_x}{\partial x}\right)^2 = \left(\frac{\sin\theta}{\lambda_x}\right)^2 + \left(\frac{\cos\theta}{\lambda_x}\right)^2 = \frac{1}{\lambda_x^2}, \tag{D.29}$$

$$a_{12} = -2\left(\frac{\partial\Lambda_y}{\partial y}\frac{\partial\Lambda_x}{\partial y} + \frac{\partial\Lambda_x}{\partial x}\frac{\partial\Lambda_y}{\partial x}\right) = -2\left(\frac{\cos\theta}{\lambda_y}\frac{\sin\theta}{\lambda_x} - \frac{\sin\theta}{\lambda_y}\frac{\cos\theta}{\lambda_x}\right) = 0. \tag{D.30}$$

On obtient :

$$\begin{aligned} \| \text{ grad } U(s) \|^2 &= \frac{1}{\det J(u)^2}\left(a_{11}\left(\frac{\partial\tilde{m}_1}{\partial x}\right)^2 + a_{12}\left(\frac{\partial\tilde{m}_1}{\partial x}\frac{\partial\tilde{m}_1}{\partial y}\right) + a_{22}\left(\frac{\partial\tilde{m}_1}{\partial y}\right)^2\right) \\ &= (\lambda_x\lambda_y)^2\left(\frac{1}{\lambda_y^2}\left(\frac{\partial\tilde{m}_1}{\partial x}\right)^2 + \frac{1}{\lambda_x^2}\left(\frac{\partial\tilde{m}_1}{\partial y}\right)^2\right). \end{aligned} \tag{D.31}$$

Et pour le Laplacien de U on a :

$$\begin{aligned} \Delta U(s) &= \text{div grad } U(s) \\ &= \sum_{i=x,y}\frac{\partial}{\partial\Lambda_i}\left(J^{-1}\text{grad } \tilde{m}_1(u)\right)_i \\ &= \sum_{i=x,y}\left(J^{-T}\text{grad } \left(J^{-1}\text{grad } \tilde{m}_1(u)\right)_i(u)\right)_i, \end{aligned} \tag{D.32}$$

avec :

$$\frac{\partial}{\partial\Lambda_i} = \sum_{j=x,y}\frac{\partial u_j}{\partial\Lambda_i}\frac{\partial}{\partial u_j} = \sum_{j=x,y}\left(J^{-T}\right)_{ij}\frac{\partial}{\partial u_j}, \tag{D.33}$$

et

$$u_{j_{j=x,y}} = (u_x, u_y) = (x, y). \tag{D.34}$$

Par conséquent, il vient :

$$\text{grad } \left(J^{-1}\text{grad } \tilde{m}_1(u)\right)_x(u) = \frac{1}{\det J(u)}\begin{pmatrix} \frac{\partial\Lambda_y}{\partial y}\frac{\partial^2\tilde{m}_1}{\partial x^2} - \frac{\partial\Lambda_x}{\partial y}\frac{\partial^2\tilde{m}_1}{\partial x\partial y} \\ -\frac{\partial\Lambda_y}{\partial x}\frac{\partial^2\tilde{m}_1}{\partial x^2} + \frac{\partial\Lambda_x}{\partial x}\frac{\partial^2\tilde{m}_1}{\partial x\partial y} \end{pmatrix} = \lambda_x\lambda_y\begin{pmatrix} \frac{\cos\theta}{\lambda_y}\frac{\partial^2\tilde{m}_1}{\partial x^2} + \frac{-\sin\theta}{\lambda_x}\frac{\partial^2\tilde{m}_1}{\partial x\partial y} \\ \frac{\sin\theta}{\lambda_y}\frac{\partial^2\tilde{m}_1}{\partial x^2} + \frac{\cos\theta}{\lambda_x}\frac{\partial^2\tilde{m}_1}{\partial x\partial y} \end{pmatrix}, \tag{D.35}$$

et

$$\text{grad } \left(J^{-1}\text{grad } \tilde{m}_1(u)\right)_y(u) = \frac{1}{\det J(u)}\begin{pmatrix} \frac{\partial\Lambda_y}{\partial y}\frac{\partial^2\tilde{m}_1}{\partial x\partial y} - \frac{\partial\Lambda_x}{\partial y}\frac{\partial^2\tilde{m}_1}{\partial y^2} \\ -\frac{\partial\Lambda_y}{\partial x}\frac{\partial^2\tilde{m}_1}{\partial x\partial y} + \frac{\partial\Lambda_x}{\partial x}\frac{\partial^2\tilde{m}_1}{\partial y^2} \end{pmatrix} = \lambda_x\lambda_y\begin{pmatrix} \frac{\cos\theta}{\lambda_y}\frac{\partial^2\tilde{m}_1}{\partial x\partial y} + \frac{-\sin\theta}{\lambda_x}\frac{\partial^2\tilde{m}_1}{\partial y^2} \\ \frac{\sin\theta}{\lambda_y}\frac{\partial^2\tilde{m}_1}{\partial x\partial y} + \frac{\cos\theta}{\lambda_x}\frac{\partial^2\tilde{m}_1}{\partial y^2} \end{pmatrix}, \tag{D.36}$$

Et alors (après simplifications) :

$$\begin{aligned} \Delta U(s) &= \sum_{i=x,y}\left(J^{-T}\text{grad } \left(J^{-1}\text{grad } \tilde{m}_1(u)\right)_i(u)\right)_i \\ &= \frac{1}{\det J(u)^2}\left(a_{11}\left(\frac{\partial^2\tilde{m}_1}{\partial x^2}\right) + a_{12}\frac{\partial^2\tilde{m}_1}{\partial x\partial y} + a_{22}\left(\frac{\partial^2\tilde{m}_1}{\partial y^2}\right)\right) \\ &= (\lambda_x\lambda_y)^2\left(\frac{1}{\lambda_y^2}\left(\frac{\partial^2\tilde{m}_1}{\partial x^2}\right) + \frac{1}{\lambda_x^2}\left(\frac{\partial^2\tilde{m}_1}{\partial y^2}\right)\right). \end{aligned} \tag{D.37}$$

En remplaçant ces expressions (D.31 et D.37) dans (D.20), on obtient la norme associée à l'inverse de l'opérateur de covariance exponentielle :

$$\begin{aligned} \| m_1 \|^2_{\mathbf{C}_M^{-1}} |_{2D} &= \frac{1}{2\pi}\int_\Omega \det J(u)\tilde{m}_1(u)^2 du \\ &+ \frac{1}{\pi}\int_\Omega \frac{1}{detJ(u)}\left(a_{11}\left(\frac{\partial \tilde{m}_1}{\partial x}\right)^2 + a_{12}\left(\frac{\partial \tilde{m}_1}{\partial x}\frac{\partial \tilde{m}_1}{\partial y}\right) + a_{22}\left(\frac{\partial \tilde{m}_1}{\partial y}\right)^2\right) du \\ &+ \frac{1}{2\pi}\int_\Omega \frac{1}{detJ(u)^3}\left(a_{11}\left(\frac{\partial^2 \tilde{m}_1}{\partial x^2}\right) + a_{12}\frac{\partial^2 \tilde{m}_1}{\partial x \partial y} + a_{22}\left(\frac{\partial^2 \tilde{m}_1}{\partial y^2}\right)\right)^2 du \end{aligned} \tag{D.38}$$

avec $\Omega = \Omega_x \times \Omega_y$.
C'est à dire :

$$\begin{aligned} \| m_1 \|^2_{\mathbf{C}_M^{-1}} |_{2D} &= \frac{1}{2\pi}\int_\Omega \frac{1}{\lambda_x\lambda_y}\tilde{m}_1(u)^2 du \\ &+ \frac{1}{\pi}\int_\Omega \lambda_x\lambda_y\left(\frac{1}{\lambda_y^2}\left(\frac{\partial \tilde{m}_1}{\partial x}\right)^2 + \frac{1}{\lambda_x^2}\left(\frac{\partial \tilde{m}_1}{\partial y}\right)^2\right) du \\ &+ \frac{1}{2\pi}\int_\Omega (\lambda_x\lambda_y)^3\left(\frac{1}{\lambda_y^2}\frac{\partial^2 \tilde{m}_1}{\partial x^2} + \frac{1}{\lambda_x^2}\frac{\partial^2 \tilde{m}_1}{\partial y^2}\right)^2 du \end{aligned} \tag{D.39}$$

avec $\Omega = \Omega_x \times \Omega_y$ et $\tilde{m}_1 = \frac{m_1}{\sigma^2}$.
On remarque que si on pose $\lambda_x = \lambda_y$ on obtient l'expression de la norme dans le cas isotrope.

D.2 Calcul de l'expression de la norme en 3D

Pour traîter le cas 3D, on utilise l'approximation faite dans le cas isotrope (équation 3.22 de la section 3.3.2.2 du chapitre 3), c'est à dire que nous supposons :

$$\tilde{V}(s) = \frac{1}{8\pi}\left(U(s) - 2\Delta U(s) + \Delta\Delta U(s)\right). \tag{D.40}$$

et donc :

$$\| m_1 \|^2_{\mathbf{C}_M^{-1}} |_{3D} = \frac{1}{8\pi}\int_{\Lambda(\Omega)} \left(U(s)^2 + 2 \| \text{grad } U(s) \|^2 + (\Delta U(s))^2\right) ds, \tag{D.41}$$

avec $\Omega = \Omega_x \times \Omega_y \times \Omega_z$.

On effectue maintenant le changement de variable $s = \Lambda(u) = (\Lambda_x(u), \Lambda_y(u), \Lambda_z(u))$ avec $u = (x, y, z)$. En effet, comme pour l'expression de la norme dans le cas 2D ou 3D dans le cas où on se place en coordonnées polaires (on définit un λ et un σ pour tout le domaine et pour toutes les directions), on peut choisir de définir des directions d'anisotropie pour les longueurs de corrélations. On pose alors :

$$\Lambda(u) = \begin{pmatrix} \frac{x\cos\theta + y\sin\theta}{\lambda_x} \\ \frac{-x\sin\alpha + y\cos\theta}{\lambda_y} \\ \frac{z}{\lambda_z} \end{pmatrix}, \tag{D.42}$$

avec λ_x, λ_y et λ_z les longueurs de corrélations selon les trois directions principales d'anisotropie.

La matrice jacobienne de cette transformation s'écrit alors :

$$J = \begin{pmatrix} \frac{\partial \Lambda_x}{\partial x} & \frac{\partial \Lambda_x}{\partial y} & \frac{\partial \Lambda_x}{\partial z} \\ \frac{\partial \Lambda_y}{\partial x} & \frac{\partial \Lambda_y}{\partial y} & \frac{\partial \Lambda_y}{\partial z} \\ \frac{\partial \Lambda_z}{\partial x} & \frac{\partial \Lambda_z}{\partial y} & \frac{\partial \Lambda_z}{\partial z} \end{pmatrix} = \begin{pmatrix} \frac{\cos\theta}{\lambda_x} & \frac{\sin\theta}{\lambda_x} & 0 \\ \frac{-\sin\theta}{\lambda_y} & \frac{\cos\theta}{\lambda_y} & 0 \\ 0 & 0 & \frac{1}{\lambda_z} \end{pmatrix}, \tag{D.43}$$

avec λ_x, λ_y et λ_z les longueurs de corrélations selon les trois directions principales d'anisotropie.

On calcule le déterminant de J :

$$\begin{aligned} \det J &= \frac{\partial \Lambda_x}{\partial x}\frac{\partial \Lambda_y}{\partial y}\frac{\partial \Lambda_z}{\partial z} - \frac{\partial \Lambda_x}{\partial x}\frac{\partial \Lambda_z}{\partial y}\frac{\partial \Lambda_y}{\partial z} - \frac{\partial \Lambda_x}{\partial y}\frac{\partial \Lambda_y}{\partial x}\frac{\partial \Lambda_z}{\partial z} \\ &+ \frac{\partial \Lambda_x}{\partial y}\frac{\partial \Lambda_z}{\partial x}\frac{\partial \Lambda_y}{\partial z} + \frac{\partial \Lambda_x}{\partial z}\frac{\partial \Lambda_y}{\partial x}\frac{\partial \Lambda_z}{\partial y} - \frac{\partial \Lambda_x}{\partial z}\frac{\partial \Lambda_z}{\partial x}\frac{\partial \Lambda_y}{\partial y} \\ &= \frac{1}{\lambda_x \lambda_y \lambda_z}. \end{aligned} \tag{D.44}$$

On a $\det J \neq 0$, donc J est inversible et J^{-1} est la matrice du changement de variable qui donne u en fonction de s. Les formules d'inversion d'une matrice 2×2 ($J^{-1} = \frac{1}{\det J}(\text{com } J)^T$) donnent alors :

$$\begin{aligned} J^{-1} &= \frac{1}{\det J(u)} \begin{pmatrix} \frac{\partial \Lambda_y}{\partial y}\frac{\partial \Lambda_z}{\partial z} - \frac{\partial \Lambda_y}{\partial z}\frac{\partial \Lambda_z}{\partial y} & -\left(\frac{\partial \Lambda_x}{\partial y}\frac{\partial \Lambda_z}{\partial z} - \frac{\partial \Lambda_z}{\partial y}\frac{\partial \Lambda_x}{\partial z}\right) & \frac{\partial \Lambda_x}{\partial y}\frac{\partial \Lambda_y}{\partial z} - \frac{\partial \Lambda_y}{\partial y}\frac{\partial \Lambda_x}{\partial z} \\ -\left(\frac{\partial \Lambda_y}{\partial x}\frac{\partial \Lambda_z}{\partial z} - \frac{\partial \Lambda_z}{\partial x}\frac{\partial \Lambda_y}{\partial z}\right) & \frac{\partial \Lambda_x}{\partial x}\frac{\partial \Lambda_z}{\partial z} - \frac{\partial \Lambda_z}{\partial x}\frac{\partial \Lambda_x}{\partial z} & -\left(\frac{\partial \Lambda_x}{\partial x}\frac{\partial \Lambda_y}{\partial z} - \frac{\partial \Lambda_y}{\partial x}\frac{\partial \Lambda_x}{\partial z}\right) \\ \frac{\partial \Lambda_y}{\partial x}\frac{\partial \Lambda_z}{\partial y} - \frac{\partial \Lambda_z}{\partial x}\frac{\partial \Lambda_y}{\partial y} & -\left(\frac{\partial \Lambda_x}{\partial x}\frac{\partial \Lambda_z}{\partial y} - \frac{\partial \Lambda_z}{\partial x}\frac{\partial \Lambda_x}{\partial y}\right) & \frac{\partial \Lambda_x}{\partial x}\frac{\partial \Lambda_y}{\partial y} - \frac{\partial \Lambda_y}{\partial x}\frac{\partial \Lambda_x}{\partial y} \end{pmatrix} \\ &= \lambda_x \lambda_y \lambda_z \begin{pmatrix} \frac{\cos\theta}{\lambda_y \lambda_z} & \frac{-\sin\theta}{\lambda_x \lambda_z} & 0 \\ \frac{\sin\theta}{\lambda_y \lambda_z} & \frac{\cos\theta}{\lambda_x \lambda_z} & 0 \\ 0 & 0 & \frac{1}{\lambda_x \lambda_y} \end{pmatrix}. \end{aligned} \tag{D.45}$$

On a ensuite :

$$\text{grad } U(s) = J^{-1} \text{grad } \tilde{m}_1(u), \tag{D.46}$$

ou encore :

$$\begin{aligned} \text{grad } U(s) &= \frac{1}{\det J(u)} \begin{pmatrix} \left(\frac{\partial \Lambda_y}{\partial y}\frac{\partial \Lambda_z}{\partial z} - \frac{\partial \Lambda_y}{\partial z}\frac{\partial \Lambda_z}{\partial y}\right)\frac{\partial \tilde{m}_1}{\partial x} - \left(\frac{\partial \Lambda_x}{\partial y}\frac{\partial \Lambda_z}{\partial z} - \frac{\partial \Lambda_z}{\partial y}\frac{\partial \Lambda_x}{\partial z}\right)\frac{\partial \tilde{m}_1}{\partial y} + \left(\frac{\partial \Lambda_x}{\partial y}\frac{\partial \Lambda_y}{\partial z} - \frac{\partial \Lambda_y}{\partial y}\frac{\partial \Lambda_x}{\partial z}\right)\frac{\partial \tilde{m}_1}{\partial z} \\ -\left(\frac{\partial \Lambda_y}{\partial x}\frac{\partial \Lambda_z}{\partial z} - \frac{\partial \Lambda_z}{\partial x}\frac{\partial \Lambda_y}{\partial z}\right)\frac{\partial \tilde{m}_1}{\partial x} + \left(\frac{\partial \Lambda_x}{\partial x}\frac{\partial \Lambda_z}{\partial z} - \frac{\partial \Lambda_z}{\partial x}\frac{\partial \Lambda_x}{\partial z}\right)\frac{\partial \tilde{m}_1}{\partial y} - \left(\frac{\partial \Lambda_x}{\partial x}\frac{\partial \Lambda_y}{\partial z} - \frac{\partial \Lambda_y}{\partial x}\frac{\partial \Lambda_x}{\partial z}\right)\frac{\partial \tilde{m}_1}{\partial z} \\ \left(\frac{\partial \Lambda_y}{\partial x}\frac{\partial \Lambda_z}{\partial y} - \frac{\partial \Lambda_z}{\partial x}\frac{\partial \Lambda_y}{\partial y}\right)\frac{\partial \tilde{m}_1}{\partial x} - \left(\frac{\partial \Lambda_x}{\partial x}\frac{\partial \Lambda_z}{\partial y} - \frac{\partial \Lambda_z}{\partial x}\frac{\partial \Lambda_x}{\partial y}\right)\frac{\partial \tilde{m}_1}{\partial y} + \left(\frac{\partial \Lambda_x}{\partial x}\frac{\partial \Lambda_y}{\partial y} - \frac{\partial \Lambda_y}{\partial x}\frac{\partial \Lambda_x}{\partial y}\right)\frac{\partial \tilde{m}_1}{\partial z} \end{pmatrix} \\ &= \lambda_x \lambda_y \lambda_z \begin{pmatrix} \frac{\cos\theta}{\lambda_y \lambda_z}\frac{\partial \tilde{m}_1}{\partial x} + \frac{-\sin\theta}{\lambda_x \lambda_z}\frac{\partial \tilde{m}_1}{\partial y} \\ \frac{\sin\theta}{\lambda_y \lambda_z}\frac{\partial \tilde{m}_1}{\partial x} + \frac{\cos\theta}{\lambda_x \lambda_z}\frac{\partial \tilde{m}_1}{\partial y} \\ \frac{1}{\lambda_x \lambda_y}\frac{\partial \tilde{m}_1}{\partial z} \end{pmatrix}. \end{aligned} \tag{D.47}$$

Donc :

$$
\begin{aligned}
\| \operatorname{grad} U(s) \|^2 &= \frac{1}{\det J(u)^2}\left(\left(\frac{\partial\Lambda_y}{\partial y}\frac{\partial\Lambda_z}{\partial z}-\frac{\partial\Lambda_y}{\partial z}\frac{\partial\Lambda_z}{\partial y}\right)\frac{\partial\tilde{m}_1}{\partial x}-\left(\frac{\partial\Lambda_x}{\partial y}\frac{\partial\Lambda_z}{\partial z}-\frac{\partial\Lambda_z}{\partial y}\frac{\partial\Lambda_x}{\partial z}\right)\frac{\partial\tilde{m}_1}{\partial y}\right.\\
&+ \left.\left(\frac{\partial\Lambda_x}{\partial y}\frac{\partial\Lambda_y}{\partial z}-\frac{\partial\Lambda_y}{\partial y}\frac{\partial\Lambda_x}{\partial z}\right)\frac{\partial\tilde{m}_1}{\partial z}\right)^2\\
&+ \frac{1}{\det J(u)^2}\left(-\left(\frac{\partial\Lambda_y}{\partial x}\frac{\partial\Lambda_z}{\partial z}-\frac{\partial\Lambda_z}{\partial x}\frac{\partial\Lambda_y}{\partial z}\right)\frac{\partial\tilde{m}_1}{\partial x}+\left(\frac{\partial\Lambda_x}{\partial x}\frac{\partial\Lambda_z}{\partial z}-\frac{\partial\Lambda_z}{\partial x}\frac{\partial\Lambda_x}{\partial z}\right)\frac{\partial\tilde{m}_1}{\partial y}\right.\\
&- \left.\left(\frac{\partial\Lambda_x}{\partial x}\frac{\partial\Lambda_y}{\partial z}-\frac{\partial\Lambda_y}{\partial x}\frac{\partial\Lambda_x}{\partial z}\right)\frac{\partial\tilde{m}_1}{\partial z}\right)^2\\
&+ \frac{1}{\det J(u)^2}\left(\left(\frac{\partial\Lambda_y}{\partial x}\frac{\partial\Lambda_z}{\partial y}-\frac{\partial\Lambda_z}{\partial x}\frac{\partial\Lambda_y}{\partial y}\right)\frac{\partial\tilde{m}_1}{\partial x}-\left(\frac{\partial\Lambda_x}{\partial x}\frac{\partial\Lambda_z}{\partial y}-\frac{\partial\Lambda_z}{\partial x}\frac{\partial\Lambda_x}{\partial y}\right)\frac{\partial\tilde{m}_1}{\partial y}\right.\\
&+ \left.\left(\frac{\partial\Lambda_x}{\partial x}\frac{\partial\Lambda_y}{\partial y}-\frac{\partial\Lambda_y}{\partial x}\frac{\partial\Lambda_x}{\partial y}\right)\frac{\partial\tilde{m}_1}{\partial z}\right)^2\\
&= (\lambda_x\lambda_y\lambda_z)^2\left(\frac{\cos\theta}{\lambda_y\lambda_z}\frac{\partial\tilde{m}_1}{\partial x}+\frac{-\sin\theta}{\lambda_x\lambda_z}\frac{\partial\tilde{m}_1}{\partial y}\right)^2+(\lambda_x\lambda_y\lambda_z)^2\left(\frac{\sin\theta}{\lambda_y\lambda_z}\frac{\partial\tilde{m}_1}{\partial x}+\frac{\cos\theta}{\lambda_x\lambda_z}\frac{\partial\tilde{m}_1}{\partial y}\right)^2\\
&+ (\lambda_x\lambda_y\lambda_z)^2\left(\frac{1}{\lambda_x\lambda_y}\frac{\partial\tilde{m}_1}{\partial z}\right)^2. \qquad \text{(D.48)}
\end{aligned}
$$

On peut définir les variables suivantes :

$$
\begin{aligned}
a_{11} &= \left(\frac{\partial\Lambda_y}{\partial y}\frac{\partial\Lambda_z}{\partial z}-\frac{\partial\Lambda_y}{\partial z}\frac{\partial\Lambda_z}{\partial y}\right)^2+\left(\frac{\partial\Lambda_y}{\partial x}\frac{\partial\Lambda_z}{\partial z}-\frac{\partial\Lambda_z}{\partial x}\frac{\partial\Lambda_y}{\partial z}\right)^2+\left(\frac{\partial\Lambda_y}{\partial x}\frac{\partial\Lambda_z}{\partial y}-\frac{\partial\Lambda_z}{\partial x}\frac{\partial\Lambda_y}{\partial y}\right)^2\\
&= \left(\frac{\cos\theta}{\lambda_y\lambda_z}\right)^2+\left(\frac{\sin\theta}{\lambda_y\lambda_z}\right)^2=\left(\frac{1}{\lambda_y\lambda_z}\right)^2, \qquad \text{(D.49)}
\end{aligned}
$$

$$
\begin{aligned}
a_{22} &= \left(\frac{\partial\Lambda_x}{\partial y}\frac{\partial\Lambda_z}{\partial z}-\frac{\partial\Lambda_z}{\partial y}\frac{\partial\Lambda_x}{\partial z}\right)^2+\left(\frac{\partial\Lambda_x}{\partial x}\frac{\partial\Lambda_z}{\partial z}-\frac{\partial\Lambda_z}{\partial x}\frac{\partial\Lambda_x}{\partial z}\right)^2+\left(\frac{\partial\Lambda_x}{\partial x}\frac{\partial\Lambda_z}{\partial y}-\frac{\partial\Lambda_z}{\partial x}\frac{\partial\Lambda_x}{\partial y}\right)^2\\
&= \left(\frac{\sin\theta}{\lambda_x\lambda_z}\right)^2+\left(\frac{\cos\theta}{\lambda_x\lambda_z}\right)^2=\left(\frac{1}{\lambda_x\lambda_z}\right)^2, \qquad \text{(D.50)}
\end{aligned}
$$

$$
\begin{aligned}
a_{33} &= \left(\frac{\partial\Lambda_x}{\partial y}\frac{\partial\Lambda_y}{\partial z}-\frac{\partial\Lambda_y}{\partial y}\frac{\partial\Lambda_x}{\partial z}\right)^2+\left(\frac{\partial\Lambda_x}{\partial x}\frac{\partial\Lambda_y}{\partial z}-\frac{\partial\Lambda_y}{\partial x}\frac{\partial\Lambda_x}{\partial z}\right)^2+\left(\frac{\partial\Lambda_x}{\partial x}\frac{\partial\Lambda_y}{\partial y}-\frac{\partial\Lambda_y}{\partial x}\frac{\partial\Lambda_x}{\partial y}\right)^2\\
&= \left(\frac{1}{\lambda_x\lambda_y}\right)^2, \qquad \text{(D.51)}
\end{aligned}
$$

$$
\begin{aligned}
a_{12} &= -2\left(\left(\frac{\partial\Lambda_y}{\partial y}\frac{\partial\Lambda_z}{\partial z}-\frac{\partial\Lambda_y}{\partial z}\frac{\partial\Lambda_z}{\partial y}\right)\left(\frac{\partial\Lambda_x}{\partial y}\frac{\partial\Lambda_z}{\partial z}-\frac{\partial\Lambda_z}{\partial y}\frac{\partial\Lambda_x}{\partial z}\right)\right.\\
&+ \left(\frac{\partial\Lambda_y}{\partial x}\frac{\partial\Lambda_z}{\partial z}-\frac{\partial\Lambda_z}{\partial x}\frac{\partial\Lambda_y}{\partial z}\right)\left(\frac{\partial\Lambda_x}{\partial x}\frac{\partial\Lambda_z}{\partial z}-\frac{\partial\Lambda_z}{\partial x}\frac{\partial\Lambda_x}{\partial z}\right)\\
&+ \left.\left(\frac{\partial\Lambda_y}{\partial x}\frac{\partial\Lambda_z}{\partial y}-\frac{\partial\Lambda_z}{\partial x}\frac{\partial\Lambda_y}{\partial y}\right)\left(\frac{\partial\Lambda_x}{\partial x}\frac{\partial\Lambda_z}{\partial y}-\frac{\partial\Lambda_z}{\partial x}\frac{\partial\Lambda_x}{\partial y}\right)\right)\\
&= -2\left(\frac{\cos\theta}{\lambda_y\lambda_z}\frac{\sin\theta}{\lambda_x\lambda_z}-\frac{\sin\theta}{\lambda_y\lambda_z}\frac{\cos\theta}{\lambda_x\lambda_z}\right)=0, \qquad \text{(D.52)}
\end{aligned}
$$

$$\begin{aligned}
a_{13} &= 2\left(\left(\frac{\partial\Lambda_y}{\partial y}\frac{\partial\Lambda_z}{\partial z}-\frac{\partial\Lambda_y}{\partial z}\frac{\partial\Lambda_z}{\partial y}\right)\left(\frac{\partial\Lambda_x}{\partial y}\frac{\partial\Lambda_y}{\partial z}-\frac{\partial\Lambda_y}{\partial y}\frac{\partial\Lambda_x}{\partial z}\right)\right.\\
&+ \left(\frac{\partial\Lambda_y}{\partial x}\frac{\partial\Lambda_z}{\partial z}-\frac{\partial\Lambda_z}{\partial x}\frac{\partial\Lambda_y}{\partial z}\right)\left(\frac{\partial\Lambda_x}{\partial x}\frac{\partial\Lambda_y}{\partial z}-\frac{\partial\Lambda_y}{\partial x}\frac{\partial\Lambda_x}{\partial z}\right)\\
&+ \left.\left(\frac{\partial\Lambda_y}{\partial x}\frac{\partial\Lambda_z}{\partial y}-\frac{\partial\Lambda_z}{\partial x}\frac{\partial\Lambda_y}{\partial y}\right)\left(\frac{\partial\Lambda_x}{\partial x}\frac{\partial\Lambda_y}{\partial y}-\frac{\partial\Lambda_y}{\partial x}\frac{\partial\Lambda_x}{\partial y}\right)\right)\\
&= 0, \qquad \text{(D.53)}
\end{aligned}$$

$$\begin{aligned}
a_{23} &= -2\left(\left(\frac{\partial\Lambda_x}{\partial y}\frac{\partial\Lambda_z}{\partial z}-\frac{\partial\Lambda_z}{\partial y}\frac{\partial\Lambda_x}{\partial z}\right)\left(\frac{\partial\Lambda_x}{\partial y}\frac{\partial\Lambda_y}{\partial z}-\frac{\partial\Lambda_y}{\partial y}\frac{\partial\Lambda_x}{\partial z}\right)\right.\\
&+ \left(\frac{\partial\Lambda_x}{\partial x}\frac{\partial\Lambda_z}{\partial z}-\frac{\partial\Lambda_z}{\partial x}\frac{\partial\Lambda_x}{\partial z}\right)\left(\frac{\partial\Lambda_x}{\partial x}\frac{\partial\Lambda_y}{\partial z}-\frac{\partial\Lambda_y}{\partial x}\frac{\partial\Lambda_x}{\partial z}\right)\\
&+ \left.\left(\frac{\partial\Lambda_x}{\partial x}\frac{\partial\Lambda_z}{\partial y}-\frac{\partial\Lambda_z}{\partial x}\frac{\partial\Lambda_x}{\partial y}\right)\left(\frac{\partial\Lambda_x}{\partial x}\frac{\partial\Lambda_y}{\partial y}-\frac{\partial\Lambda_y}{\partial x}\frac{\partial\Lambda_x}{\partial y}\right)\right)\\
&= 0. \qquad \text{(D.54)}
\end{aligned}$$

On obtient :

$$\begin{aligned}
\| \text{grad } U(s) \|^2 &= \frac{1}{\det J(u)^2}\left(a_{11}\left(\frac{\partial\tilde{m}_1}{\partial x}\right)^2 + a_{22}\left(\frac{\partial\tilde{m}_1}{\partial y}\right)^2 + a_{33}\left(\frac{\partial\tilde{m}_1}{\partial z}\right)^2\right.\\
&+ \left. a_{12}\left(\frac{\partial\tilde{m}_1}{\partial x}\frac{\partial\tilde{m}_1}{\partial y}\right) + a_{13}\left(\frac{\partial\tilde{m}_1}{\partial x}\frac{\partial\tilde{m}_1}{\partial z}\right) + a_{23}\left(\frac{\partial\tilde{m}_1}{\partial y}\frac{\partial\tilde{m}_1}{\partial z}\right)\right)\\
&= (\lambda_x\lambda_y\lambda_z)^2\left(\frac{1}{(\lambda_y\lambda_z)^2}\left(\frac{\partial\tilde{m}_1}{\partial x}\right)^2 + \frac{1}{(\lambda_x\lambda_z)^2}\left(\frac{\partial\tilde{m}_1}{\partial y}\right)^2 + \frac{1}{(\lambda_x\lambda_y)^2}\left(\frac{\partial\tilde{m}_1}{\partial z}\right)^2\right).
\end{aligned} \qquad \text{(D.55)}$$

Et pour le Laplacien de U on a :

$$\begin{aligned}
\Delta U(s) &= \text{div grad } U(s)\\
&= \sum_{i=x,y,z}\frac{\partial}{\partial\Lambda_i}\left(J^{-1}\text{grad } \tilde{m}_1(u)\right)_i\\
&= \sum_{i=x,y,z}\left(J^{-T}\text{grad } \left(J^{-1}\text{grad } \tilde{m}_1(u)\right)_i(u)\right)_i, \qquad \text{(D.56)}
\end{aligned}$$

avec :

$$\frac{\partial}{\partial\Lambda_i} = \sum_{j=x,y,z}\frac{\partial u_j}{\partial\Lambda_i}\frac{\partial}{\partial u_j} = \sum_{j=x,y,z}\left(J^{-T}\right)_{ij}\frac{\partial}{\partial u_j}, \qquad \text{(D.57)}$$

et

$$u_{j_{j=x,y,z}} = (u_x, u_y, u_z) = (x, y, z). \qquad \text{(D.58)}$$

Par conséquent, il vient :

$$
\begin{aligned}
&\text{grad}\ \left(J^{-1}\text{grad}\ \tilde{m}_1(u)\right)_x(u) \\
&= \frac{1}{\det J(u)}\begin{pmatrix}
\left(\frac{\partial\Lambda_y}{\partial y}\frac{\partial\Lambda_z}{\partial z}-\frac{\partial\Lambda_y}{\partial z}\frac{\partial\Lambda_z}{\partial y}\right)\frac{\partial^2\tilde{m}_1}{\partial x^2}-\left(\frac{\partial\Lambda_x}{\partial y}\frac{\partial\Lambda_z}{\partial z}-\frac{\partial\Lambda_z}{\partial y}\frac{\partial\Lambda_x}{\partial z}\right)\frac{\partial^2\tilde{m}_1}{\partial x\partial y}+\left(\frac{\partial\Lambda_x}{\partial y}\frac{\partial\Lambda_y}{\partial z}-\frac{\partial\Lambda_y}{\partial y}\frac{\partial\Lambda_x}{\partial z}\right)\frac{\partial^2\tilde{m}_1}{\partial x\partial z} \\
-\left(\frac{\partial\Lambda_y}{\partial x}\frac{\partial\Lambda_z}{\partial z}-\frac{\partial\Lambda_z}{\partial x}\frac{\partial\Lambda_y}{\partial z}\right)\frac{\partial^2\tilde{m}_1}{\partial x^2}+\left(\frac{\partial\Lambda_x}{\partial x}\frac{\partial\Lambda_z}{\partial z}-\frac{\partial\Lambda_z}{\partial x}\frac{\partial\Lambda_x}{\partial z}\right)\frac{\partial^2\tilde{m}_1}{\partial x\partial y}-\left(\frac{\partial\Lambda_x}{\partial x}\frac{\partial\Lambda_y}{\partial z}-\frac{\partial\Lambda_y}{\partial x}\frac{\partial\Lambda_x}{\partial z}\right)\frac{\partial^2\tilde{m}_1}{\partial x\partial z} \\
\left(\frac{\partial\Lambda_y}{\partial x}\frac{\partial\Lambda_z}{\partial y}-\frac{\partial\Lambda_z}{\partial x}\frac{\partial\Lambda_y}{\partial y}\right)\frac{\partial^2\tilde{m}_1}{\partial x^2}-\left(\frac{\partial\Lambda_x}{\partial x}\frac{\partial\Lambda_z}{\partial y}-\frac{\partial\Lambda_z}{\partial x}\frac{\partial\Lambda_x}{\partial y}\right)\frac{\partial^2\tilde{m}_1}{\partial x\partial y}+\left(\frac{\partial\Lambda_x}{\partial x}\frac{\partial\Lambda_y}{\partial y}-\frac{\partial\Lambda_y}{\partial x}\frac{\partial\Lambda_x}{\partial y}\right)\frac{\partial^2\tilde{m}_1}{\partial x\partial z}
\end{pmatrix} \\
&= \lambda_x\lambda_y\lambda_z\begin{pmatrix}
\frac{\cos\theta}{\lambda_y\lambda_z}\frac{\partial^2\tilde{m}_1}{\partial x^2}+\frac{-\sin\theta}{\lambda_x\lambda_z}\frac{\partial^2\tilde{m}_1}{\partial x\partial y} \\
\frac{\sin\theta}{\lambda_y\lambda_z}\frac{\partial^2\tilde{m}_1}{\partial x^2}+\frac{\cos\theta}{\lambda_x\lambda_z}\frac{\partial^2\tilde{m}_1}{\partial x\partial y} \\
\frac{1}{\lambda_x\lambda_y}\frac{\partial^2\tilde{m}_1}{\partial x\partial z}
\end{pmatrix},
\end{aligned}
\tag{D.59}
$$

et

$$
\begin{aligned}
&\text{grad}\ \left(\text{grad}\ \tilde{m}_1(u)\right)_y(u) \\
&= \frac{1}{\det J(u)}\begin{pmatrix}
\left(\frac{\partial\Lambda_y}{\partial y}\frac{\partial\Lambda_z}{\partial z}-\frac{\partial\Lambda_y}{\partial z}\frac{\partial\Lambda_z}{\partial y}\right)\frac{\partial^2\tilde{m}_1}{\partial x\partial y}-\left(\frac{\partial\Lambda_x}{\partial y}\frac{\partial\Lambda_z}{\partial z}-\frac{\partial\Lambda_z}{\partial y}\frac{\partial\Lambda_x}{\partial z}\right)\frac{\partial^2\tilde{m}_1}{\partial y^2}+\left(\frac{\partial\Lambda_x}{\partial y}\frac{\partial\Lambda_y}{\partial z}-\frac{\partial\Lambda_y}{\partial y}\frac{\partial\Lambda_x}{\partial z}\right)\frac{\partial^2\tilde{m}_1}{\partial y\partial z} \\
-\left(\frac{\partial\Lambda_y}{\partial x}\frac{\partial\Lambda_z}{\partial z}-\frac{\partial\Lambda_z}{\partial x}\frac{\partial\Lambda_y}{\partial z}\right)\frac{\partial^2\tilde{m}_1}{\partial x\partial y}+\left(\frac{\partial\Lambda_x}{\partial x}\frac{\partial\Lambda_z}{\partial z}-\frac{\partial\Lambda_z}{\partial x}\frac{\partial\Lambda_x}{\partial z}\right)\frac{\partial^2\tilde{m}_1}{\partial y^2}-\left(\frac{\partial\Lambda_x}{\partial x}\frac{\partial\Lambda_y}{\partial z}-\frac{\partial\Lambda_y}{\partial x}\frac{\partial\Lambda_x}{\partial z}\right)\frac{\partial^2\tilde{m}_1}{\partial y\partial z} \\
\left(\frac{\partial\Lambda_y}{\partial x}\frac{\partial\Lambda_z}{\partial y}-\frac{\partial\Lambda_z}{\partial x}\frac{\partial\Lambda_y}{\partial y}\right)\frac{\partial^2\tilde{m}_1}{\partial x\partial y}-\left(\frac{\partial\Lambda_x}{\partial x}\frac{\partial\Lambda_z}{\partial y}-\frac{\partial\Lambda_z}{\partial x}\frac{\partial\Lambda_x}{\partial y}\right)\frac{\partial^2\tilde{m}_1}{\partial y^2}+\left(\frac{\partial\Lambda_x}{\partial x}\frac{\partial\Lambda_y}{\partial y}-\frac{\partial\Lambda_y}{\partial x}\frac{\partial\Lambda_x}{\partial y}\right)\frac{\partial^2\tilde{m}_1}{\partial y\partial z}
\end{pmatrix} \\
&= \lambda_x\lambda_y\lambda_z\begin{pmatrix}
\frac{\cos\theta}{\lambda_y\lambda_z}\frac{\partial^2\tilde{m}_1}{\partial x\partial y}+\frac{-\sin\theta}{\lambda_x\lambda_z}\frac{\partial^2\tilde{m}_1}{\partial y^2} \\
\frac{\sin\theta}{\lambda_y\lambda_z}\frac{\partial^2\tilde{m}_1}{\partial x\partial y}+\frac{\cos\theta}{\lambda_x\lambda_z}\frac{\partial^2\tilde{m}_1}{\partial y^2} \\
\frac{1}{\lambda_x\lambda_y}\frac{\partial^2\tilde{m}_1}{\partial y\partial z}
\end{pmatrix},
\end{aligned}
\tag{D.60}
$$

et

$$
\begin{aligned}
&\text{grad}\ \left(\text{grad}\ \tilde{m}_1(u)\right)_z(u) \\
&= \frac{1}{\det J(u)}\begin{pmatrix}
\left(\frac{\partial\Lambda_y}{\partial y}\frac{\partial\Lambda_z}{\partial z}-\frac{\partial\Lambda_y}{\partial z}\frac{\partial\Lambda_z}{\partial y}\right)\frac{\partial^2\tilde{m}_1}{\partial x\partial z}-\left(\frac{\partial\Lambda_x}{\partial y}\frac{\partial\Lambda_z}{\partial z}-\frac{\partial\Lambda_z}{\partial y}\frac{\partial\Lambda_x}{\partial z}\right)\frac{\partial^2\tilde{m}_1}{\partial y\partial z}+\left(\frac{\partial\Lambda_x}{\partial y}\frac{\partial\Lambda_y}{\partial z}-\frac{\partial\Lambda_y}{\partial y}\frac{\partial\Lambda_x}{\partial z}\right)\frac{\partial^2\tilde{m}_1}{\partial z^2} \\
-\left(\frac{\partial\Lambda_y}{\partial x}\frac{\partial\Lambda_z}{\partial z}-\frac{\partial\Lambda_z}{\partial x}\frac{\partial\Lambda_y}{\partial z}\right)\frac{\partial^2\tilde{m}_1}{\partial x\partial z}+\left(\frac{\partial\Lambda_x}{\partial x}\frac{\partial\Lambda_z}{\partial z}-\frac{\partial\Lambda_z}{\partial x}\frac{\partial\Lambda_x}{\partial z}\right)\frac{\partial^2\tilde{m}_1}{\partial y\partial z}-\left(\frac{\partial\Lambda_x}{\partial x}\frac{\partial\Lambda_y}{\partial z}-\frac{\partial\Lambda_y}{\partial x}\frac{\partial\Lambda_x}{\partial z}\right)\frac{\partial^2\tilde{m}_1}{\partial z^2} \\
\left(\frac{\partial\Lambda_y}{\partial x}\frac{\partial\Lambda_z}{\partial y}-\frac{\partial\Lambda_z}{\partial x}\frac{\partial\Lambda_y}{\partial y}\right)\frac{\partial^2\tilde{m}_1}{\partial x\partial z}-\left(\frac{\partial\Lambda_x}{\partial x}\frac{\partial\Lambda_z}{\partial y}-\frac{\partial\Lambda_z}{\partial x}\frac{\partial\Lambda_x}{\partial y}\right)\frac{\partial^2\tilde{m}_1}{\partial y\partial z}+\left(\frac{\partial\Lambda_x}{\partial x}\frac{\partial\Lambda_y}{\partial y}-\frac{\partial\Lambda_y}{\partial x}\frac{\partial\Lambda_x}{\partial y}\right)\frac{\partial^2\tilde{m}_1}{\partial z^2}
\end{pmatrix} \\
&= \lambda_x\lambda_y\lambda_z\begin{pmatrix}
\frac{\cos\theta}{\lambda_y\lambda_z}\frac{\partial^2\tilde{m}_1}{\partial x\partial z}+\frac{-\sin\theta}{\lambda_x\lambda_z}\frac{\partial^2\tilde{m}_1}{\partial y\partial z} \\
\frac{\sin\theta}{\lambda_y\lambda_z}\frac{\partial^2\tilde{m}_1}{\partial x\partial z}+\frac{\cos\theta}{\lambda_x\lambda_z}\frac{\partial^2\tilde{m}_1}{\partial y\partial z} \\
\frac{1}{\lambda_x\lambda_y}\frac{\partial^2\tilde{m}_1}{\partial z^2}
\end{pmatrix}.
\end{aligned}
\tag{D.61}
$$

Et alors (après de lourds calculs) :

$$
\begin{aligned}
\Delta U(s) &= \sum_{i=x,y,z}\left(J^{-T}\text{grad}\ \left(J^{-1}\text{grad}\ \tilde{m}_1(u)\right)_i(u)\right)_i \\
&= \frac{1}{\det J(u)^2}\left(a_{11}\frac{\partial^2\tilde{m}_1}{\partial x}+a_{22}\frac{\partial^2\tilde{m}_1}{\partial y}+a_{33}\frac{\partial^2\tilde{m}_1}{\partial z}+a_{12}\frac{\partial^2\tilde{m}_1}{\partial x\partial y}+a_{13}\frac{\partial^2\tilde{m}_1}{\partial x\partial z}+a_{23}\frac{\partial^2\tilde{m}_1}{\partial y\partial z}\right) \\
&= (\lambda_x\lambda_y\lambda_z)^2\left(\frac{1}{(\lambda_y\lambda_z)^2}\frac{\partial^2\tilde{m}_1}{\partial x}+\frac{1}{(\lambda_x\lambda_z)^2}\frac{\partial^2\tilde{m}_1}{\partial y}+\frac{1}{(\lambda_x\lambda_y)^2}\frac{\partial^2\tilde{m}_1}{\partial z}\right).
\end{aligned}
\tag{D.62}
$$

En remplaçant ces expressions (D.55 et D.62) dans (D.41), on obtient la norme associée à l'inverse de l'opérateur de covariance exponentielle :

$$\begin{aligned}
\| m_1 \|^2_{\mathbf{C}_M^{-1}} |_{3D} &= \frac{1}{8\pi}\int_\Omega \det J(u)\tilde{m}_1(u)^2 du \\
&+ \frac{1}{4\pi}\int_\Omega \frac{1}{detJ(u)}\left(a_{11}\left(\frac{\partial\tilde{m}_1}{\partial x}\right)^2 + a_{22}\left(\frac{\partial\tilde{m}_1}{\partial y}\right)^2 + a_{33}\left(\frac{\partial\tilde{m}_1}{\partial z}\right)^2\right) du \\
&+ \frac{1}{4\pi}\int_\Omega \frac{1}{detJ(u)}\left(a_{12}\left(\frac{\partial\tilde{m}_1}{\partial x}\frac{\partial\tilde{m}_1}{\partial y}\right) + a_{13}\left(\frac{\partial\tilde{m}_1}{\partial x}\frac{\partial\tilde{m}_1}{\partial z}\right) + a_{23}\left(\frac{\partial\tilde{m}_1}{\partial y}\frac{\partial\tilde{m}_1}{\partial z}\right)\right) du \\
&+ \frac{1}{8\pi}\int_\Omega \frac{1}{(detJ(u))^3}\left(a_{11}\frac{\partial^2\tilde{m}_1}{\partial x} + a_{22}\frac{\partial^2\tilde{m}_1}{\partial y} + a_{33}\frac{\partial^2\tilde{m}_1}{\partial z} + a_{12}\frac{\partial^2\tilde{m}_1}{\partial x\partial y} + a_{13}\frac{\partial^2\tilde{m}_1}{\partial x\partial z} + a_{23}\frac{\partial^2\tilde{m}_1}{\partial y\partial z}\right)^2 du
\end{aligned} \tag{D.63}$$

avec $\Omega = \Omega_x \times \Omega_y \times \Omega_z$.
C'est à dire :

$$\begin{aligned}
\| m_1 \|^2_{\mathbf{C}_M^{-1}} |_{3D} &= \frac{1}{8\pi}\int_\Omega \frac{1}{\lambda_x\lambda_y\lambda_z}\tilde{m}_1(u)^2 du \\
&+ \frac{1}{4\pi}\int_\Omega \lambda_x\lambda_y\lambda_z\left(\frac{1}{(\lambda_y\lambda_z)^2}\left(\frac{\partial\tilde{m}_1}{\partial x}\right)^2 + \frac{1}{(\lambda_x\lambda_z)^2}\left(\frac{\partial\tilde{m}_1}{\partial y}\right)^2 + \frac{1}{(\lambda_x\lambda_y)^2}\left(\frac{\partial\tilde{m}_1}{\partial z}\right)^2\right) du \\
&+ \frac{1}{8\pi}\int_\Omega (\lambda_x\lambda_y\lambda_z)^3\left(\frac{1}{(\lambda_y\lambda_z)^2}\frac{\partial^2\tilde{m}_1}{\partial x^2} + \frac{1}{(\lambda_x\lambda_z)^2}\frac{\partial^2\tilde{m}_1}{\partial y^2} + \frac{1}{(\lambda_x\lambda_y)^2}\frac{\partial^2\tilde{m}_1}{\partial z}\right)^2 du
\end{aligned} \tag{D.64}$$

avec $\Omega = \Omega_x \times \Omega_y \times \Omega_z$ et $\tilde{m}_1 = \frac{m_1}{\sigma^2}$.

On remarque que si on pose $\lambda_x = \lambda_y = \lambda_z$ et $\theta = 0$ on obtient l'expression de la norme dans le cas isotrope.

Annexe E

CONTINUITÉ DE $J(m_n)$

Le résultat démontré dans cette annexe ainsi que ceux des deux annexes suivantes ont déjà été démontrés par Lailly et Delprat-Jannaud, 1991.

On considère un triplet (source, réflecteur, récepteur) et le rayon associé $\mathcal{R}_j$ pour le modèle m_n. La composante de $J(m_n)$ correspondant à la vitesse étant

$$[J(m_n)\delta v]_j = -\int_{R_j} \frac{\delta v}{v_n^2} ds. \tag{E.1}$$

On a donc :

$$[J(m_n)\delta v]_j^2 = (\int_{R_j} \frac{\delta v}{v_n^2} ds)^2,$$

$$[J(m_n)\delta v]_j^2 \leq \frac{1}{v_{min}^2}(\int_{R_j} \delta v ds)^2,$$

$$[J(m_n)\delta v]_j^2 \leq \frac{C}{v_{min}^2}(\int_{R_j} \delta v^2 ds),$$

où $C = \int_{R_j} ds$ est une constante*. Le théorème de trace (Taylor, 1981) montre que :

$$\int_{R_j} \delta v^2 ds \leq K_1 \parallel \delta v \parallel_{H^{1+\epsilon}}^2, \tag{E.2}$$

où K_1 est une constante. Par conséquent

$$[J(m_n)\delta v]_j^2 \leq \frac{CK_1}{v_{min}^2} \parallel \delta v \parallel_{H^{1+\epsilon}}^2, \tag{E.3}$$

et donc

$$||J(m_n)\delta v||_{C_d^{-1}}^2 \leq K_2 \parallel \delta v \parallel_{H^{1+\epsilon}}^2, \tag{E.4}$$

*qui vient de l'inégalité de Cauchy-Schwartz :

$$\int_{R_j} \delta v ds \leq (\int_{R_j} ds)^{1/2} (\int_{R_j} \delta v^2 ds)^{1/2}$$

où K_2 est une constante. La composante de $J(m_n)$ correspondant à la profondeur étant

$$[J(m_n)\delta m]_j = \frac{\delta z(x_j)}{v(P_j)}(\cos\theta_k^{inc} + \cos\theta_k^{ref}), \tag{E.5}$$

où x_j est l'abscisse du point d'impact P_j du rayon sur le réflecteur non perturbé et $\cos\theta_k^{inc}$ et $\cos\theta_k^{ref}$ sont les angles entre l'axe vertical et respectivement, le rayon incident et réfléchi, il vient

$$[J(m_n)\delta z]_j^2 = \frac{\delta z(x_j)^2}{v(P_j)^2}(\cos\theta_k^{inc} + \cos\theta_k^{ref})^2,$$

$$[J(m_n)\delta z]_j^2 \leq 4\frac{\delta z(x_j)^2}{v((P_j)_{min})^2}.$$

Donc

$$||J(m_n)\delta z||^2_{C_d^{-1}} \leq 4\frac{\delta z(x_j)^2}{v((P_j)_{min})^2}. \tag{E.6}$$

Et en conséquence (théorème de trace)

$$||J(m_n)\delta z||^2_{C_d^{-1}} \leq K||\delta z||^2_{H^{1+\epsilon}}. \tag{E.7}$$

Pour des modèles 3D (avec des variations de vitesse latérale et verticale), la norme la plus faible qui rend $J(m_n)$ continu est donc $H^{1+\epsilon}$. Le jacobien $J(m_n)$ est donc continu pour la norme de H^2. De plus, il est facile de comprendre que $J(m_n)$ n'est pas borné pour la norme L^2.

Annexe F

UN PROBLÈME BIEN POSÉ

La continuité de $\tilde{C}$ pour la norme H^2 est déduite de la continuité de $J(m_n)$ mentionnée précédemment (Annexe B). Lailly et Delprat-Jannaud, 1991, ont montré que la solution du problème 3.32 existe et est unique, nous rappelons ici ces résultats.

Soit l'hypothèse de complétude des données (cette hypothèse sera toujours vérifiée chaque fois que l'on aura sufisamment pointé des temps de trajet) :

$$kerJ(m_n)\bigcap\mathcal{P}_1 = 0. \tag{F.1}$$

avec $\mathcal{P}_1$ le sous-espace des polynômes de degré 1 qui est l'espace de dimension finie des modèles décrits par des fonctions à gradient constant (vitesse à gradient constant et réflecteur plan). Le choix des polynômes du premier degré vient du fait que la régularisation par les dérivées secondes utilisée s'annule pour de tels polynômes.

On cherche à minimiser :

$$\begin{aligned}\tilde{C}(\delta m) &= \frac{1}{2}\langle J(m_n)\delta m, J(m_n)\delta m\rangle_{C_d^{-1}} - \langle J(m_n)\delta m, \delta T^{obs}\rangle_{C_d^{-1}} \\ &+ \frac{1}{2}\langle R_\epsilon\delta m, R_\epsilon\delta m\rangle_{L^2} + \langle R_\epsilon m_n, R_\epsilon\delta m\rangle_{L^2} + \frac{1}{2}\langle\delta T^{obs}, \delta T^{obs}\rangle_{C_d^{-1}} + \frac{1}{2}\langle R_\epsilon m_n, R_\epsilon m_n\rangle_{L^2}.\end{aligned} \tag{F.2}$$

On pose :

$$a(\delta m_1, \delta m_2) = \langle J(m_n)\delta m_1, J(m_n)\delta m_2\rangle_{C_d^{-1}} + \langle R_\epsilon\delta m_1, R_\epsilon\delta m_2\rangle_{L^2}, \tag{F.3}$$

$$\langle b, \delta m\rangle_{L^2} = 2\langle J(m_n)\delta m, \delta T^{obs}\rangle_{C_d^{-1}} - 2\langle R_\epsilon m_n, R_\epsilon\delta m\rangle_{L^2}. \tag{F.4}$$

Nous allons montrer que la forme bilinéaire symétrique $a(.,.)$ est continue et coercive sur $\mathcal{M}\times\mathcal{M}$ ($\mathcal{M} = H^2$).

La continuité de $a(.,.)$ provient de la continuité de $J(m_n)$ (voir Annexe E.

La coercivité se démontre de la manière suivante : Sous l'hypothèse F.1, le résultat de coercivité est une conséquence directe du théorème de Necăs (Necăs, 1967, théorème 7.3, chapitre 2) qui est le suivant :

<u>Théorème :</u> Soit Ω un ouvert, supposons que l'application identique de $W_p^{(1)}(\Omega)$ dans $L_p(\Omega)$ soit complètement continue, $P\subset P_{k-1}$ et f_i, $i = 1, 2, ..., l$ des fonctionnelles sur $W_p^{(k)}(\Omega)$,

k un entier ≥ 1 et $p \geq 1$, telles que si $v \in P_{k-1}$

$$\sum_{i=1}^{l} \mid f_i v \mid^p = 0 \Leftrightarrow p \in P. \tag{F.5}$$

Alors

$$c_1 \mid u \mid_{W_p^{(k)}/P} \leq \left(\sum_{|\alpha|=k} \mid D^\alpha u \mid_{L^p} + \sum_{i=1}^{l} \mid f_i u \mid^p \right)^{1/p} \leq c_2 \mid u \mid_{W_p^{(k)}/P}. \tag{F.6}$$

Si $p = 2$, $W_2^{(k)}/P$ est un espace de Hilbert dont le produit scalaire est

$$(\tilde{v}, \tilde{u}) = \sum_{|\alpha|=k} \int_\Omega D^\alpha v D^\alpha \bar{u} dx + \sum_{i=1}^{l} f_i v \bar{f_i} u. \tag{F.7}$$

On applique ce théorème et on obtient alors la coercivité :

$$a(\delta m, \delta m) = \langle J(m_n)\delta m, J(m_n)\delta m \rangle_{C_d^{-1}} + \langle R_\epsilon \delta m, R_\epsilon \delta m \rangle_{L^2} \geq K ||\delta m||^2_{H^2}. \tag{F.8}$$

Et donc on a bien :

$$\exists K > 0, \forall \delta m \in H^2, a(\delta m, \delta m) \geq K ||\delta m||^2_{H^2}. \tag{F.9}$$

D'autre part, pour la forme linéaire b on a le résultat suivant :

$$b \in (H^2)', \tag{F.10}$$

où $(H^2)'$ est le dual de H^2. Pour cela, il suffit de montrer la continuité de b.

$$|\langle b, \delta m \rangle_{L^2}| \leq K ||J(m_n)\delta m||_{L^2}. \tag{F.11}$$

La continuité de $J(m_n)$ entraîne la continuité de b. b est donc un élément de $(H^2)'$.

Pour finir, à partir de ces deux résultats (coercivité et $b \in (H^2)'$) on applique le théorème de Lax-Milgram :

<u>Théorème (Lax-Milgram) :</u> Soit a(u, v) une forme bilinéaire symétrique, continue et coercive sur un Hilbert $\mathcal{M}$. Soit b appartenant à $\mathcal{M}'$, une forme linéaire continue sur $\mathcal{M}$. Alors il existe un unique $\widehat{m}$ appartenant à $\mathcal{M}$ tel que :

$$\frac{1}{2} a(\widehat{m}, \widehat{m}) - \langle b, \widehat{m} \rangle \leq \frac{1}{2} a(m, m) - \langle b, m \rangle, \forall m \in \mathcal{M}. \tag{F.12}$$

Ce qui nous permet de montrer que le problème de minimisation de $\tilde{C}(\delta m)$ 3.32 admet une unique solution $\tilde{\delta m} \in H^2$. En effet, d'après le thórème précédent, on obtient :

$$\exists! \tilde{\delta m} \in H^2 /$$

$$\frac{1}{2} a(\tilde{\delta m}, \tilde{\delta m}) - \langle b, \tilde{m} \rangle_{L^2} \leq \frac{1}{2} a(\delta m, \delta m) - \langle b, \delta m \rangle_{L^2}, \forall \delta m \in H^2,$$

$$\frac{1}{2}a(\tilde{\delta m}, \tilde{\delta m}) - \left\langle b, \tilde{\delta m} \right\rangle_{L^2} + K \leq \frac{1}{2}a(\delta m, \delta m) - \langle b, \delta m \rangle_{L^2} + K.$$

On choisit $K = \frac{1}{2}\langle \delta T^{obs}, \delta T^{obs} \rangle_{C_d^{-1}} + \frac{1}{2}\langle R_\epsilon m_0, R_\epsilon m_0 \rangle_{L^2}$. On a alors :

$$\frac{1}{2}a(\delta m, \delta m) - \langle b, \delta m \rangle_{L^2} + K = \tilde{C}(\delta m).$$

$$\exists! \tilde{\delta m} \in H^2 / \tilde{C}(\tilde{\delta m}) \leq \tilde{C}(\delta m), \forall \delta m \in H^2.$$

On a donc montré que le problème linéarisé admettait une et une seule solution dans H^2 sous une certaine hypothèse F.1 concernant le modèle de référence.

Annexe G

L'APPROCHE BAYESIENNE

L'approche stochastique de la théorie du problème inverse est une approche très attractive mais il faut savoir que utiliser l'approche stochastique pour résoudre le problème inverse sismique n'est généralement pas concevable vu la taille des espaces du modèle sismique et des données. Dans ce contexte, la simplification de l'approche au cas de probabilités gaussiennes offre une sortie «providentielle». En effet, l'approche bayesienne est une approche cohérente pour la résolution des problèmes inverses car elle permet de prendre en compte et de traiter de la même manière l'information a priori sur la solution et sur les données.

Comme W. Gouveia et J. Scales, 1997, l'ont mentionné dans leur article, il y a de fortes similarités entre l'approche bayesienne et la formulation régularisée (3.13). Dans l'approche bayesienne, les distributions de probabilité décrivent l'information (incertaine) sur les données et l'information a priori sur le modèle. La régularisation peut être interprétée comme une information a priori. On retrouve en effet la principale idée de la régularisation qui consiste à combiner l'information contenue dans les données et celle contenue dans l'a priori, la fusion de ces deux sources d'information s'effectue par la règle de Bayes. On peut alors dire que cette appproche généralise la régularisation déterministe. Notons que A. Tarantola, 1987, a démontré que la solution du problème inverse stochastique était en fait la solution provenant du problème classique de minimisation de la fonction coût pour la norme L^2.

Soit D un vecteur aléatoire dont chaque réalisation d est un vecteur qui contient les données discrétisées. On définit aussi un vecteur aléatoire M dont chaque réalisation est un vecteur qui contient les paramètres du milieu discrétisé. Les techniques d'inversion bayesienne ont toutes pour fondement le théorème de Bayes (Saporta, 1990), qui permet d'écrire la densité de probabilité conditionnelle de M sachant D en fonction des densités de probabilités marginales de M et de D, ainsi que de leur densité jointe sous la forme :

$$f_{M|D}(m|d) = \frac{f_{D|M}(d|m) f_M(m)}{f_D(d)}. \tag{G.1}$$

La fonction $f_{M|D}(m|d)$ est appelée densité de probabilité a posteriori. C'est la densité de probabilité de m connaissant les données d mesurées. La fonction $f_{D|M}(d|m)$ est la densité de probabilité conditionnelle de d sachant m. Cette fonction contient toutes les informations dont on dispose concernant les relations théoriques entre les paramètres et les données ; elle contient aussi les informations dont on dispose sur le bruit contenu dans les données. La fonction $f_M(m)$ est la densité de probabilité marginale du vecteur M. Cette fonction traduit l'ensemble des informations dont on dispose sur les modèles en dehors de celles provenant des données. Par conséquent, cette fonction traduit l'ensemble

des connaissances a priori sur les paramètres du milieu. Le terme $f_D(d)$ ne dépend pas des modèles et il sera considéré comme une constante pour le problème inverse.

Il est important de remarquer que la fonction de densité de probabilité a posteriori contient toutes les informations disponibles sur les paramètres du milieu sachant les données observées. Par suite, cette fonction est en fait la solution du problème inverse stochastique. Cependant, il n'est généralement pas possible de représenter cette fonction. Aussi on se contente d'en dériver une estimation du vrai modèle, par exemple en calculant la moyenne a posteriori de M. En général, les densités de probabilité qui sont utilisées pour le problème inverse stochastique sont des densités de probabilités gaussiennes. Ceci est dû au fait que dans ce cas, la densité de probabilité a posteriori est, elle aussi, gaussienne, et qu'on peut la caractériser complètement ; de plus, dans le cas gaussien, le calcul de la moyenne a posteriori conduit à la minimisation d'un critère des moindres carrés.

Revenons à notre problème et explicitons la formule (G.1). Soient M et D les deux vecteurs aléatoires de dimensions respectives N_M et N_D.

Sous l'hypothèse gaussienne, la densité de probabilité marginale de M s'écrit :

$$f_M(m) = \frac{1}{\sqrt{(2\pi)^{N_M} \det C_m}} \exp\left(-\frac{1}{2}(m - m^{prior})^T C_m^{-1}(m - m^{prior})\right), \tag{G.2}$$

où m^{prior} est le modèle a priori et C_m est la matrice de covariance.

Le modèle a priori peut être obtenu en utilisant les informations dont on dispose sur la zone étudiée comme par exemple la connaissance géologique de la zone. La matrice de covariance permet de décrire les erreurs sur ce modèle a priori.

La densité de probabilité conditionnelle de D sachant M, en supposant qu'elle soit gaussienne est :

$$f_{D|M}(T^{obs}|m) = \frac{1}{\sqrt{(2\pi)^{N_D} \det C_d}} \exp\left(-\frac{1}{2}(T(m) - T^{obs})^T C_d^{-1}(T(m) - T^{obs})\right), \tag{G.3}$$

où C_d est la matrice de covariance qui décrit les erreurs de modélisation et les erreurs de mesures des données T^{obs}.

Une fois ces lois définies, il ne nous reste plus qu'à calculer la densité de probabilité a posteriori, c'est à dire la densité de probabilité conditionnelle de M sachant qu'une réalisation de D est T^{obs} (les données observées). On peut montrer (Tarantola, 1987) que cette densité de probabilité est encore gaussienne et que son expression est :

$$f_{M|D}(m|T^{obs}) \alpha \exp\left(-\frac{1}{2}\left((T(m) - T^{obs})^T C_d^{-1}(T(m) - T^{obs}) + (m - m^{prior})^T C_m^{-1}(m - m^{prior})\right)\right), \tag{G.4}$$

le signe de proportionnalité (α) étant dû à la fonction $f_D(T^{obs})$. Nous linéarisons ensuite l'opérateur direct T autour du modèle courant et à la convergence (Tarantola, 1987) on obtient alors l'expression de la densité de probabilité a posteriori :

$$f_{M|D}(m|T^{obs}) \alpha \exp\left(-\frac{1}{2}(m - \bar{m})^T C_m'^{-1}(m - \bar{m})\right), \tag{G.5}$$

avec

$$\bar{m} = \left(J^T C_d^{-1} J + C_m^{-1}\right)^{-1} \left(J^T C_d^{-1} T^{obs} + C_m^{-1} m^{prior}\right), \tag{G.6}$$

et

$$C_m' = \left(J^T C_d^{-1} J + C_m^{-1}\right)^{-1}. \tag{G.7}$$

où $J = \frac{\partial T^{cal}}{\partial m}$.

Il s'agit d'une densité de probabilité gaussienne de moyenne $\bar{m}$ et de matrice de covariance C'_m et la fonction $S(m)$ définie sur l'espace des modèles par :

$$S(m) = \frac{1}{2}(T(m) - T^{obs})^T C_d^{-1}(T(m) - T^{obs}) + (m - m^{prior})^T C_m^{-1}(m - m^{prior}), \tag{G.8}$$

est la fonction coût de l'inversion.

Le maximum (ou mode) de la densité de probabilité a posteriori est alors confondu avec la moyenne $\bar{m}$ (estimateur central de la gaussienne (G.5), appelé aussi modèle le plus probable) donnant ainsi une estimation correcte de la solution. Celui-ci peut donc être calculé soit par son expression analytique (G.6), soit en minimisant la forme quadratique $S(m)$ (critère des moindres carrés).

L'espace des modèles admissibles peut être caractérisé par les lignes de contour de la fonction de densité de probabilité Gaussienne a posteriori définies par la matrice de covariance a posteriori : ce sont des ellipsoïdes de centre $\bar{m}$

$$(m - \bar{m})^T C_m'^{-1}(m - \bar{m}) = \text{constante}. \tag{G.9}$$

Les termes diagonaux de C'_m sont les incertitudes sur chaque paramètre décrivant le modèle et les termes hors-diagonaux sont les corrélations entre ces incertitudes. Par exemple, la probabilité que les paramètres p_i du vrai modèle vérifient $-2C'_{mi,i} \leq p_i - p_{\bar{m}} \leq 2C'_{mi,i}$, indépendamment de la valeur des autres paramètres du modèle, est d'environ 95%. Pour prendre en compte les corrélations entre les paramètres, nous devrons étudier l'ellipse de confiance 95%. Les axes des ellipsoïdes (G.9) sont définis par les vecteurs propres de C'_m, la racine carrée des valeurs propres donne les incertitudes sur le vecteur propre associé.

Cette approche est valide quand le problème inverse est faiblement non linéaire, la matrice de covariance a posteriori étant seulement valide dans un voisinage du modèle solution (cadre linéarisé).

Annexe H

PUBLICATIONS/COMMUNICATIONS À CONGRÈS

K. Broto, F. Delbos, C. Duffet and T. Perdrizet, *Prestack Traveltime Tomography : towards the geology/geophysics integration for complex structure imaging*, AAPG 2004, Prague, 2004.

C. Duffet and D. Sinoquet, *Quantifying uncertainties on the solution model of seismic tomography*, 4th European Congress on Computational Methods in Applied Sciences and Engineering (ECCOMAS 2004), http ://www.mit.jyu.fi/eccomas2004/, Jyväskylä, Finlande, 2004.

C. Duffet et D. Sinoquet, *Estimation des incertitudes sur le modèle solution de la tomographie sismique*, 35ème Congrès National d'Analyse Numérique (CANUM 2003), La Grande-Motte, Languedoc-Roussillon, 2003.

C. Duffet and D. Sinoquet, *Quantifying geological uncertainties in velocity model building*, 72nd Annual International Meeting and Exposition, Soc. Expl. Geophys.,Expanded Abstracts, 926-929, Salt Lake City, Utah, 2002.

D. Sinoquet, joint work with C. Duffet, *Uncertainty analysis of the solution model of 3D seismic reflection tomography* Workshop IMA, Inverse Problems and Quantification of Uncertainty, Minneapolis, USA, 2002.

C. Duffet and D. Sinoquet, *Uncertainty on the solution model of reflection tomography*, KIM 2001 Annual Report, Institut Français du Pétrole, Rueil-Malmaison, France, 2001.

Bibliographie

Apostoiu-Marin, I., and Delprat-Jannaud, F., 1995, Application of the SMART method on the PICROCOL data set : PSI Annual Report, Institut Français du Pétrole, Rueil Malmaison, France.

Backus, G. E., and Gilbert, F., 1970, Uniqueness in the inversion of inaccurate gross data : Phil. Trans. Roy. Soc. London, **266**, 123–192.

Bamberger, A., Chavent, G., Hemon, C., and Lailly, P., 1982, Inversion of normal incidence seismograms : Geophysics, **47**, no. 5, 757–770.

Berkhout, A. J., 1980, Seismic migration-imaging of acoustic energy by wave field extrapolation : Elsevier Science Publ. Co., Inc.

Bickel, S. H., 1990, Velocity-depth ambiguity of reflection traveltimes : Geophysics, **55**, no. 3, 266–276.

Billette, F., and Lambaré, G., 1998, velocity macro-model estimation from seismic reflection data by stereotomography : Geophys. J. Int., **135**, no. 2, 671–680.

Bishop, T. N., Bube, K. P., Cutler, R. T., Langan, R. T., Love, P. L., Resnick, J. R., Shuey, R. T., Spindler, D. A., and Wyld, H. W., 1985, Tomographic determination of velocity and depth in laterally varying media : Geophysics, **50**, no. 6, 903–923.

Bois, P., La-Porte, M., Lavergne, M., and Thomas, G., 1971, Essai de détermination automatique des vitesses sismiques par mesures entre puits : Geophys. Prospect., **19**, 42–83.

Bording, R. P., Gersztenkorn, A., Lines, L., Scales, J. A., and Treitel, S., 1987, Application of seismic travel-time tomography : Geophys. J. Roy. Astr. Soc., **90**, 285–303.

Bourgeois, A., Lailly, P., Collino, F., Dumont, F., and Macé, D., 1989, The linearized seismic inverse problem : an attractive method for a sharp description of strategic traps. : 59th Ann. Internat. Mtg., Soc. Expl. Geophys., Expanded Abstract, pages 973–976.

Brac, J., Dequirez, P., Hervé, H., Jacques, C., Lailly, P., Richard, V., and Tran-Van-Nhieu, D., 1992, Inversion with a priori information : an approach to integrated stratigraphic interpretation *in* Sheriff, R., Ed., Reservoir Geophysics, Investigations in Geophysics,SEG publication :, 251–258.

Brezis, H., 1987, Analyse fonctionelle : théorie et applications : Masson.

Broto, K., and Ehinger, A., 1998, How to access 3D prestack kinematic information in case of complex geologies ? : 68th Ann. Internat. Mtg., Soc. Expl. Geophys., Expanded Abstracts, pages 1401–1404.

Broto, K., and Ehinger, A., 2001, Joint tomographic inversion of PP and PS reflections : a 2d real case study : KIM Annual Report, Institut Français du Pétrole, Pau, France.

Broto, K., 1996, Application of the SMART method to a 2D North Sea data set : KIM Annual Report, Institut Français du Pétrole, Pau, France.

Bube, K. P., and Langan, R. T., 1994, A continuation approach to regularization for traveltime tomography : 64th Ann. Internat. Mtg., Soc. Expl. Geophys., Expanded Abstracts, pages 980–983.

Bube, K. P., 1995, Uniqueness of reflector depths and characterization of the slowness null space in linearized seismic reflection tomography : SIAM J. Appl. Math., **55**, no. 1, 255–266.

Cavalca, M., and Lailly, P., 2004, A traveltime tomography dedicated to the inversion of prismatic reflections : 66th Conference and Exhibition, EAGE, Expanded Abstracts.

Červený, V., Molotkov, I. A., and Pšenčík, I., 1977, Ray method in seismology : Univerzita Karlova, Praha.

Červený, V., 2001, Seismic ray theory : Cambridge Univ. Press, New York.

Ciarlet, P. G., 1988, Introduction à l'analyse numérique matricielle et à l'optimisation : Masson.

Claerbout, J. F., 1976, Fundamentals of geophysical data processing : Mc Graw–Hill Book Co.

Clarke, R. A., and Jannaud, L. R., 1996, Raytracing in the overthrust model : 66th Ann. Internat. Mtg., Soc. Expl. Geophys., Expanded Abstracts, pages 687–690.

Clarke, R. A., 1996a, Raytracing and tomography friendly geological models : KIM Annual Report, Institut Français du Pétrole, Pau, France.

Clarke, R. A. 1996b, Two point raytracing in 3D blocky models : KIM Annual Report, Institut Français du Pétrole, Pau, France.

Clarke, R. A., 1997, Modélisation et inversion de données cinématiques complexes en 3D (in English) : Ph.D. thesis, Université de Pau et des Pays de l'Adour, France.

Comby, I., 2004, Etude de méthodes d'optimisation pour la tomographie de réflexion (rapport de stage) : IFP Report, Institut Français du Pétrole, Rueil Malmaison, France, , no. 58072.

Constable, S., Parker, R. L., and Constable, C. G., 1987, Occam's inversion : A practical algorithm for generating smooth models from electromagnetic sounding data : Geophysics, **52**, 289–300.

Crase, E., Pica, A., Noble, M., McDonald, J., and Tarantola, A., 1990, Robust elastic waveform inversion : Application to real data : Geophysics, **55**, 527–538.

Dautray, R., and Lions, J., 1984, Analyse mathématique et calcul numérique pour les sciences et techniques :, volume I Masson.

de Boor, C., 1978, A practical guide to splines : Springer.

Deal, M. M., and Nolet, G., 1996, Comment on 'estimation of resolutionand covariance for large matrix inversions' by j. zhang and g. a. mcmechan : Geophys. J. Int., **127**, 245–250.

Delbos, F., Gilbert, J. C., Glowinski, R., and Sinoquet, D., 2004, Constrainted optimization in seismic reflection tomography : an sqp augmented lagrangian approach : soumis à Geophys. J. Int.

Delprat-Jannaud, F., and Lailly, P., 1991, Traveltime inversion of a 3D heterogeneous model : PSI Annual Report, Institut Français du Pétrole, Rueil Malmaison, France.

Delprat-Jannaud, F., and Lailly, P., 1992a, Reflections tomography : how to handle multiple arrivals ? : Expanded Abstacts, **62**, 741–744.

Delprat-Jannaud, F., and Lailly, P. 1992b, What information on the earth model do reflection travel times hold ? : Journal of Geophysical Research, **97**, 19827–19844.

Delprat-Jannaud, F., and Lailly, P., 1993, Ill-posed and well-posed formulations of the reflection travel time tomography problem : Journal of Geophysical Research, **98**, 6589–6605.

Delprat-Jannaud, F., 1991, Tomographie de reflection : quelle est l'information contenue dans les temps d'arrivée ? (in English) : Ph.D. thesis, Université de Paris-Sud Orsay.

der Sluis, A. V., and der Vost, H. A. V., 1987, Seismic tomography : , D. Reidel Publishing Company.

Deregowski, S. M., 1986, What is DMO : First Break, **4**, no. 7, 7–24.

Duffet, C., and Sinoquet, D., 2001, Uncertainty on the solution model of reflexion tomography : KIM Annual Report, Institut Français du Pétrole, Rueil Malmaison, France.

Duffet, C., and Sinoquet, D., 2002, Quantifying geological uncertainties in velocity model building : 72nd Ann. Internat. Mtg., Soc. Expl. Geophys., Expanded Abstracts, 926–929.

Duffet, C., and Sinoquet, D., 2003, Estimation des incertitudes sur le modèle solution de la tomographie sismique : 35ème Congrès National d'Analyse Numérique (CANUM 2003).

Duffet, C., and Sinoquet, D., 2004, Quantifying uncertainties on the solution model of seismic tomography : 4th European Congress on Computational Methods in Applied Sciences and Engineering (ECCOMAS 2004), http ://www.mit.jyu.fi/eccomas2004/.

Duquet, B., 1996, Amélioration de l'imagerie sismique de structures géologiques complexes : Ph.D. thesis, Université Paris XIII.

Ehinger, A., Broto, K., Jardin, A., and the KIMASI team, 2001, 3D tomographic velocity model determination for two North Sea case studies : 63rd Conference and Technical Exhibition, EAGE, Expanded Abstracts.

Farra, V., and Madariaga, R., 1988, Non-linear reflection tomography : Geophys. J. Roy. Ast. Soc., **95**, 135–147.

French, W. S., 1975, Computer migration of oblique seismic reflection profiles : Geophysics, **40**, 961–980.

Gardner, G. H., French, W. S., and Matzuk, T., 1974, Elements of migration and velocity analysis : Geophysics, **39**, 811–825.

Gouveia, W., and Scales, J., 1997, Resolution of seismic waveform inversion : Bayes versus occam : Inverse problems, pages 323–349.

Grenié, D., and Le Stunff, Y., 2000, Computation of a posteriori uncertainties in 3D traveltime reflection tomography : 70th Ann. Internat. Mtg., Soc. Expl. Geophys., Expanded Abstracts, **2**, 2261–2264.

Grenié, D., 2001, Tomographie de temps de trajet adaptée au problème de l'imagerie 3D en exploration pétrolière : théorie, mise en oeuvre logicielle et applications : Ph.D. thesis, Université Paris VII.

Guillaume, P., Audebert, F., Berthet, P., David, B., Herrenschmidt, A., and Zhang, X., 2001, 3d finite-offset tomographic inversion of crp-scan data, with or without anisotropy : 71st Ann. Internat. Mtg., Soc. Expl. Geophys., Expanded Abstracts, pages 718–721.

Guiziou, J. L., 1993, Estimation de modèles de vitesse en 3D par tomographie de sismique réflexion : développements d'outils sur la base du modeleur géométrique Gocad : Ph.D. thesis, Université de Paris VII.

Hanke, M., 1995, Conjugate gradient type methods for ill–posed problems.

Hansen, P. C., 2001, The l-curve and its use in the numerical treatment of inverse problems : invited chapter in P. Johnston (Ed.), Computational Inverse Problems in Electrocardiology, WIT Press, Southampton, 119–142.

Jannaud, L. R., 1992, Two-point raytracing in complex media : PSI Annual Report, Institut Français du Pétrole, Rueil-Malmaison, France.

Jurado, F., Sinoquet, D., and Ehinger, A., 1996a, 3D reflection tomography designed for complex structures : 66th Ann. Internat. Mtg., Soc. Expl. Geophys., Expanded Abstracts, 711–714.

Jurado, F., Sinoquet, D., and Lailly, P. 1996b, Jerry : a 3D reflection tomography designed for complex structures : KIM Annual Report, Institut Français du Pétrole, Pau, France.

Jurado, F., Sinoquet, D., Ehinger, A., Lailly, P., and Monestié, C., 1997, Improvements in the reflection tomography software jerry : KIM Annual Report, Institut Français du Pétrole, Pau, France.

Jurado, F., Lailly, P., and Ehinger, A., 1998, Fast 3D two-point raytracing for traveltime tomography : Proceedings of SPIE, Mathematical Methods in Geophysical Imaging V, **3453**, 70–81.

Kelley, C. T., 1999, Iterative methods for optimization : Siam, Philadelphia.

Kolb, P., Collino, F., and Lailly, P., 1986, Pre–stack inversion of a 1d medium : Proc. IEEE, **74**, 498–508.

Lailly, P., and Sinoquet, D., 1996, Smooth velocity models in reflection tomography for imaging complex geological structures : Geophys. J. Int., **124**, 349–362.

Lambaré, G., Virieux, J., Madariaga, R., and Jin, S., 1992, Iterative asymptotic inversion in the acoustic approximation : Geophysics, **57**, no. 9, 1138–1154.

Langan, R. T., Lerche, I., and Cutler, R. T., 1985, Tracing of rays through heterogeneous media : an accurate and efficient procedure : Geophysics, **50**, 1456–1465.

Lavergne, M., 1986, Méthodes sismiques : Technip.

Lawson, C. L., and Hanson, R., 1974, Solving least-squares problems : Prentice-hall.

Lines, L. R., and Treitel, S., 1984, Tutorial : A review of least–squares inversion and its application to geophysical problems. : Geophysical prospecting, **32**, 159–186.

Lines, L. R., and Treitel, S., 1985, Inversion with a grain of salt : Geophysics, **50**, 99–109.

Mensch, T., Farra, V., and Satish, C., 1998, 3D seismic traveltime tomography in orthorhombic media : 68th Ann. Internat. Mtg., Soc. Expl. Geophys., Expanded Abstracts, pages 1863–1866.

Mitsuhata, Y., 2004, Adjustement of regularization in ill-posed linear inverse problems by the empirical bayes approach : Geophysical Prospecting, **52**, 213–239.

Necăs, J., 1967, Les méthodes directes en théorie des équations elliptiques : Masson, Paris.

Noble, M., and Tarantola, A., 1991, Nonlinear waveform inversion of seismic reflection data : the marmousi model : Eur. Assn. Expl. Geophys., Zeist, 113–125.

Nocedal, J., and Wright, S. J., 1999, Numerical optimization : Springer.

Nolet, G., Montelli, R., and Virieux, J., 1999, Explicit, approximate expressions for the resolution and *a posteriori* covariance of massive tomographic systems : Geophys. J. Int., **138**, 36–44.

Parker, R., 1994, Geophysical inverse theory : Springer, Berlin.

Pellé, L., 2003, Inversion linéarisée simultanée des réflexions primaires et des réflexions multiples. : Ph.D. thesis, Université de Montpellier II.

R. G. ClappB. Biondi, S. F., and Claerbout, J. F., 1998, Regularizing velocity estimation using geologic dip information : Soc. Expl. Geophys., Expanded Abstracts, pages 1851–1854.

Renard, F., and Lailly, P., 1999, Robust and accurate determination of complex velocity/depth models by reflection travel time tomography : KIM Annual Report, Institut Français du Pétrole, Rueil-Malmaison, France.

Renard, F., and Lailly, P., 2001, How to handle correlated noise in seismic inversion : 71th Ann. Internat. Mtg., Soc. Expl. Geophys., Expanded Abstracts, 690–693.

Renard, F., 2002, Prise en compte des corrélations dans l'inversion de données sismiques. : Ph.D. thesis, Université de Monpellier.

Saporta, G., 1990, Probabilités, Analyse des données et Statistique : Éditions Technip.

Schneider, W. A., 1978, Integral formulation for migration in two and three dimensions : Geophysics, **43**, 49–76.

Sinoquet, D., 1995, Utilisation de modèles lisses pour l'inversion tomographique de données sismiques (in English) : Ph.D. thesis, Université Paris XIII, France.

Stolt, R., 1978, Migration by fournier tranform. : Geophysics, **43**, 23–48.

Stopin, A., and Ehinger, A., 2001, Joint PP PS tomographic inversion of the Mahogany 2D-4C OBC seismic data : 71th Ann. Internat. Mtg., Soc. Expl. Geophys., Expanded Abstracts.

Stopin, A., 2000, Comparison of $v(\theta)$ equations in TI medium : 9th IWSA abstracts, page 57.

Stork, C., and Clayton, R. W., 1991, Linear aspects of tomographic velocity analysis : Geophysics, **56**, 483–495.

Tarantola, A., and Valette, B., 1982, Inverse Problems = Quest for information : Journal of Geophysics, **50**, 159–170.

Tarantola, A., 1987, Inverse problem theory : Elsevier.

Taylor, M. E., 1981, Pseudodifferential operators : Princeton University Press.

Thisse, P., 1994, Tracé de rayons et migration 3D pour des interfaces triangulées : conception orientée objets et application à la sismique réflexion profonde : Ph.D. thesis, Université Louis Pasteur de Strasbourg.

Thomsen, L., 1986, Weak elastic anisotropy : Geophysics, **51**, 1954–1966.

Tikhonov, A., and Arsenin, V., 1977, Solution of ill–posed problems : Winston.

Tonellot, T., 2000, Introduction d'informations a priori dans l'inversion linéarisée élastique de données sismiques de surface avant sommation : Ph.D. thesis, Université René Descartes.

Virieux, J., and Farra, V., 1991, Ray tracing in 3–D complex isotropic media : an analysis of the problem : Geophysics, **56**, no. 12, 2057–2069.

Virieux, J., Farra, V., and Madariaga, R., 1988, Ray tracing in laterally heterogeneous media for earthquake location : J. Geophys. Res., **93**, 6585–6599.

Vogel, C. R., 1996, Non-convergence of the l-curve regularization method for nonlinear ill-posed problems : Inverse Problems, **12**, 535–547.

Woodward, M., Farmer, P., Nichols, D., and Charles, S., 1998, Automated 3d tomographic velocity analysis of residual moveout in prestack depth migrated common image point gathers : 68th Ann. Internat. Mtg. Soc. Expl. Geophys., Expanded Abstracts, pages 1218–1221.

Zabalza-Mezghani, I., 2000, Analyse statistique et planification d'expérience en ingénierie de réservoir (in french) : Ph.D. thesis, Université de Pau et des Pays de l'Adour.

Zhang, J., and McMechan, G. A., 1995, Estimation of resolution and covariance for large matrix inversions : Geophys. J. Int., **121**, 409–426.