Formulation de la tomographie des temps de première arrivée à partir d'une méthode de gradient : un pas vers une tomographie interactive

Cédric Taillandier

ED n°398 : Géosciences et Ressources Naturelles

N° attribué par la bibliothèque
|__|__|__|__|__|__|__|__|__|__|__|

T H E S E

pour obtenir le grade de
DOCTEUR DE L'ECOLE NATIONALE SUPERIEURE DES MINES DE PARIS

Spécialité "Dynamique et Ressources des Bassins Sédimentaires"

présentée et soutenue publiquement par

Cédric TAILLANDIER

le 2 décembre 2008

Formulation de la tomographie des temps de première arrivée à partir d'une méthode de gradient : un pas vers une tomographie interactive

Jury

Jean VIRIEUX	Président
Henri CALANDRA	Examinateur
Gilles GRANDJEAN	Rapporteur
Gilles LAMBARE	Examinateur
Mark NOBLE	Examinateur
William SYMES	Rapporteur

Résumé

La tomographie des temps de première arrivée cherche à estimer un modèle de vitesse de propagation des ondes sismiques à partir des temps de première arrivée pointés sur les sismogrammes. Le modèle de vitesse obtenu peut alors permettre une interprétation structurale du milieu ou bien servir de modèle initial pour d'autres traitements de l'imagerie sismique. Les domaines d'application de cette méthode s'étendent, à des échelles différentes, de la géotechnique à la sismologie en passant par la géophysique pétrolière.
Le savoir-faire du géophysicien joue un rôle important dans la difficile résolution du problème tomographique non-linéaire et mal posé. De nombreuses recherches ont entrepris de faciliter et d'améliorer cette résolution par des approches mathématique ou physique. Dans le cadre de ce travail, nous souhaitons développer une approche pragmatique, c'est-à-dire que nous considérons que le problème tomographique doit être résolu par un algorithme interactif dont les paramètres de réglage sont clairement définis. L'aspect interactif de l'algorithme facilite l'acquisition du savoir-faire tomographique car il permet de réaliser, dans un temps raisonnable, de nombreuses simulations pour des paramétrisations différentes. Le but poursuivi dans cette thèse est de définir, pour le cas spécifique de la tomographie des temps de première arrivée, un algorithme qui réponde au mieux à ces critères.
Les algorithmes de tomographie des temps de première arrivée classiquement mis en œuvre aujourd'hui ne répondent pas à nos critères d'une approche pragmatique. En effet, leur implémentation ne permet pas d'exploiter l'architecture parallèle des supercalculateurs actuels pour réduire les temps de calcul. De plus, leur mise en œuvre nécessite une paramétrisation rendue complexe du fait de la résolution du système linéaire tomographique. Toutes ces limitations pratiques sont liées à la formulation même de l'algorithme à partir de la méthode de Gauss-Newton. Cette thèse repose sur l'idée de formuler la résolution du problème tomographique à partir de la méthode de plus grande descente pour s'affranchir de ces limitations.
L'étape clé de cette formulation réside dans le calcul du gradient de la fonction coût par rapport aux paramètres du modèle. Nous utilisons la méthode de l'état adjoint et une méthode définie à partir d'un tracé de rais *a posteriori* pour calculer ce gradient. Ces deux méthodes se distinguent par leur formulation, respectivement non-linéaire et linéarisée, et par leur mise en œuvre pratique.
Nous définissons ensuite clairement la paramétrisation du nouvel algorithme de tomographie et validons sur un supercalculateur ses propriétés pratiques : une parallélisation directe et efficace, une occupation mémoire indépendante du nombre de données observées et une mise en œuvre simple. Finalement, nous présentons des résultats de tomographie pour des acquisitions de type sismique réfraction, 2-D et 3-D, synthétiques et réelles, marines et terrestres, qui valident le bon comportement de l'algorithme, en termes de résultats obtenus et de stabilité. La réalisation d'un grand nombre de simulations a été rendue possible par la rapidité d'exécution de l'algorithme, de l'ordre de quelques minutes en 2-D.

Mots-clé : géophysique, sismique réfraction, tomographie, temps de première arrivée, équation eikonale, méthode de l'état adjoint, problème inverse, méthode de gradient, 3-D, parallélisation, algorithme interactif.

Abstract

First arrival traveltime tomography based on a gradient method:
a first step towards an interactive tomography algorithm

First arrival traveltime tomography aims at inferring a seismic wave propagation velocity model from first arrival traveltimes picked on seismograms. The velocity model inferred can be used directly to perform a structural interpretation of the subsurface or as an initial model for another seismic imaging method. This technique can be applied at different scales from geotechnical studies to seismology through oil exploration.
The geophysicist know-how plays an important role in the difficult resolution of the nonlinear and ill-posed tomographic problem. Numerous studies have tried to ease and improve this resolution considering a physical or mathematical approach. Within the scope of this work, we wish to develop a pragmatic approach, i.e. we consider that the tomographic problem should be solved using an interactive algorithm whose tuning parameters are clearly identified. The interactive aspect of the algorithm facilitates the acquisition of the tomographic know-how because it allows performing, within a reasonable time, many simulations for different kinds of parameterization. The goal pursued in this work is the definition of a first arrival traveltime tomography algorithm that fulfills these specifications at best.
Conventional first arrival traveltime tomography algorithms do not match our criteria of a pragmatic approach. Indeed, their implementation hardly takes benefit of the parallel architecture of current supercomputers in order to reduce the computation times. Moreover, their practical application implies a complex parameterization due to the resolution of the linear tomographic system. All these practical limitations issue from the original formulation of the algorithm based on a Gauss-Newton minimization method. The idea developed in this work is to formulate the resolution of the tomographic problem using a steepest descent method to overcome all these limitations.
The key step of this formulation is the computation of the gradient of the misfit function with respect to the model parameters. We use the adjoint state method and a method based on an *a posteriori* ray tracing to compute this gradient. These two methods differ from their formulation, respectively nonlinear and linearized, and their implementation.
Then, we clearly define the parameterization of the new algorithm and validate on a supercomputer its practical properties that are: a direct and efficient parallelization, a memory requirement independent of the amount of input data and a straightforward implementation.
Finally, in order to validate the tomographic behaviour of this new algorithm, in terms of obtained results and stability, we present tomography results for 2-D and 3-D, synthetics and real, marine and land, seismic refraction acquisitions. A great number of simulations have been carried out thanks to the fast execution time of the algorithm, typically few minutes for 2-D simulations.

Keywords: geophysics, seismic refraction, tomography, first arrival traveltimes, eikonal equation, adjoint state method, inverse problem, gradient method, 3-D, parallelization, interactive algorithm.

Remerciements

Ce travail a bénéficié des soutiens financier et informatique de Total. Je remercie Jérôme Guilbot, François Audebert et Henri Calandra pour leur encadrement et leur accueil lors de mes séjours à Pau.

Je remercie Stéphane Operto et Jean-Xavier Dessa de Géosciences Azur pour les données sismiques de la fosse de Nankai et leurs résultats de tomographie.

Je remercie Henri Calandra, Gilles Grandjean, Gilles Lambaré, Isabelle Lecomte, William Symes et Jean Virieux d'avoir examiné ce travail et siégé à mon jury de thèse.

Je remercie chaleureusement mon directeur de thèse, Mark Noble, de m'avoir fait confiance et confié ce sujet de recherche sur lequel j'ai pris un réel plaisir à travailler. Merci Mark, pour ta disponibilité, tes intuitions toujours justes et ton soutien tout au long de la thèse.

J'adresse des remerciements tout aussi chaleureux à Hervé Chauris, pour ses précieux conseils, à Pascal Podvin, pour son enthousiasme scientifique communicatif et à Véronique Lachasse pour son aide infaillible au quotidien.

Je garderai un souvenir inoubliable de ces années passées à Fontainebleau grâce à de très belles rencontres : Pierre-Yves, Olivier, Marie, Mathieu, Alexandre, Rose, Misha, Edwige, Claire, Elodie, Charles, Louise, Mondher, Béatrice, Ali, Serge, ... Merci pour votre amitié !

Mille mercis à mes parents, ma sœur et mon frère, ainsi qu'à toute ma famille, pour leur soutien et leurs encouragements qui m'accompagnent quels que soient mes choix et m'aident à aller toujours de l'avant.

Un immense merci à Magali qui est devenue, bien malgré elle, une experte en tomographie des temps de première arrivée.

"Je crois comprendre, dit-il d'une voix douce.
Tu es un petit garçon qui désire la lune afin
d'y boire comme à une coupe d'or. C'est pourquoi,
selon toute probabilité, tu seras un grand homme,
à condition que tu restes un enfant. Tous les grands
de ce monde ont été de petits garçons qui désiraient
la lune. A force de courir et de grimper, ils sont parfois
arrivés à attraper une luciole. Mais celui qui parvient à
avoir un cerveau d'homme voit nécessairement qu'il ne peut
pas obtenir la lune ; il s'abstiendrait de la désirer s'il le pouvait,
c'est pourquoi il n'attrape même pas une luciole."

-La Coupe d'Or-, John Steinbeck

Table des matières

Chapitre 1

Introduction générale

"*Seismic tomography : art or science ?*" telle est la question posée par Frederik J. Simons en introduction d'une présentation consacrée à la reconstruction tomographique [Simons, 2004]. La suite de cette présentation souligne la grande importance du savoir-faire du géophysicien dans la difficile résolution du problème tomographique non-linéaire et mal posé [Hadamard, 1902]. De nombreuses recherches ont entrepris de faciliter et d'améliorer la résolution de ce problème. Certaines d'entre elles ont adopté une approche mathématique, par exemple sous la forme d'une régularisation [Tikhonov and Arsenin, 1977], d'autres ont choisi une approche physique, par exemple en considérant les volumes de Fresnel [Červený and Soares, 1992].
Dans le cadre de ce travail, nous souhaitons développer une approche pragmatique, c'est-à-dire que nous considérons que le problème tomographique doit être résolu par un algorithme interactif dont les paramètres de réglage sont clairement définis. L'aspect interactif de l'algorithme facilite l'acquisition du savoir-faire tomographique car il permet de réaliser, dans un temps raisonnable, de nombreuses simulations pour des paramétrisations différentes. Le but poursuivi dans cette thèse est de définir, pour le cas spécifique de la tomographie des temps de première arrivée, un algorithme qui réponde au mieux à ces critères.

La tomographie des temps de première arrivée

La tomographie des temps de première arrivée cherche à estimer un modèle de vitesse de propagation des ondes sismiques à partir des temps de première arrivée pointés sur les sismogrammes. Le modèle de vitesse obtenu peut par exemple permettre une interprétation structurale du milieu [Improta et al., 2002], [Zelt et al., 2006] ou bien servir de modèle initial pour d'autres traitements de l'imagerie sismique [Operto et al., 2006], [Brenders and Pratt, 2007]. Les domaines d'application de cette méthode s'étendent, à des échelles différentes, de la géotechnique à la sismologie en passant par la géophysique pétrolière. Dans le cadre de cette dernière, la tomographie des temps de première arrivée est généralement associée à des acquisitions de sismique réflexion terrestre. Les temps de première arrivée contiennent alors essentiellement de l'information sur la subsurface proche. La détermination du modèle de vitesse de cette zone hétérogène permet d'en corriger les effets sur le signal sismique enregistré [Zhu et al., 1992]. Ces corrections, dites statiques, permettent alors d'améliorer l'imagerie des structures plus profondes.
La tomographie des temps de première arrivée est classiquement formulée comme la résolution d'un problème inverse linéarisé localement [Tarantola, 1987a] sous l'approximation haute fréquence de la théorie des rais. Le problème direct associé correspond au calcul des temps de

première arrivée pour une acquisition et un modèle de vitesse donnés. La solution du problème tomographique est obtenue par la minimisation de la fonction coût des moindres carrés définie entre les temps de première arrivée observés et ceux calculés pour un modèle de vitesse donné. Cette minimisation est généralement réalisée par la méthode de Gauss-Newton dont le schéma itératif repose sur la résolution d'un système linéaire liant les écarts des temps de première arrivée et la perturbation de vitesse recherchée [Spakman and Nolet, 1988].
La littérature associée à la tomographie sismique, avec ses nombreuses mises en œuvre et applications, est monumentale. Dans ce travail, nous nous appuyons sur quelques références qui nous semblent clés. Les travaux réalisés en sismologie par Aki et Lee [1976] sont à l'origine de la formulation du problème tomographique telle qu'elle est établie aujourd'hui. La résolution de ce problème par la méthode *Least Squares QR* (LSQR) [Paige and Saunders, 1982] fut introduite en tomographie par Nolet [1985]. Enfin, le calcul des temps de première arrivée par la résolution par différences finies de l'équation eikonale [Vidale, 1988], [Podvin and Lecomte, 1991] a permis d'améliorer la précision et la rapidité des calculs réalisés [Le Meur, 1994], [Baina, 1998], [Zelt and Barton, 1998].

L'idée

Les algorithmes de tomographie des temps de première arrivée classiquement mis en œuvre aujourd'hui ne répondent pas à nos critères d'une approche pragmatique. En effet, leur implémentation ne permet pas d'exploiter l'architecture parallèle des supercalculateurs actuels pour réduire les temps de calcul. De plus, leur mise en œuvre nécessite une paramétrisation rendue complexe du fait de la résolution du système linéaire tomographique. Toutes ces limitations pratiques sont causées par la formulation même du problème tomographique à partir de la méthode de Gauss-Newton.

Cette thèse repose sur l'idée de formuler la résolution du problème tomographique à partir d'une méthode de minimisation par gradient.

Le choix de cette formulation permet de s'affranchir des limitations pratiques énoncées précédemment [Nolet, 1987], [Tarantola, 1987b]. Nous évaluons dans ce travail si l'algorithme de tomographie ainsi défini vérifie les critères d'une approche pragmatique.

Le plan de la thèse

La présentation écrite de ce travail reflète presque fidèlement l'ordre chronologique dans lequel les différents points ont été abordés au cours de cette thèse.

- Nous nous intéressons tout d'abord à la formulation et à la mise en œuvre pratique d'un algorithme classique de tomographie des temps de première arrivée (Chapitre 2). Cette première partie nous permet d'en préciser les limitations pratiques notamment pour des acquisitions sismiques avec un nombre important de données observées et des modèles de vitesse de grande dimension.
- Dans un second temps, nous présentons la formulation du problème tomographique à partir d'une méthode de gradient (Chapitre 3). Nous étudions alors deux méthodes permettant de calculer le gradient de la fonction coût des moindres carrés par rapport aux paramètres du modèle : la méthode de l'état adjoint [Sei and Symes, 1995], [Leung and Qian, 2006] et

une méthode définie à partir d'un tracé de rais *a posteriori*. Les propriétés pratiques de l'algorithme de tomographie défini à partir de cette formulation sont évaluées sur un super-calculateur possédant une architecture parallèle.

– Finalement, pour étudier le comportement tomographique de ce nouvel algorithme nous l'appliquons sur des jeux de données complexes et des acquisitions de taille réaliste en 2-D et en 3-D (Chapitre 4). Cette dernière partie, nous permet aussi de préciser les paramètres déterminants dans la résolution du problème tomographique.

Nous concluons ce travail avec les limites et les possibilités d'extension du nouvel algorithme de tomographie des temps de première arrivée.

Chapitre 2

La formulation classique

La tomographie des temps de première arrivée a pour objectif de retrouver un modèle de vitesse de propagation des ondes sismiques à partir de temps de première arrivée mesurés à la surface. Mathématiquement, cela consiste à résoudre un problème inverse, à savoir retrouver les paramètres d'un modèle physique qui expliquent au mieux les données observées. La résolution de ce problème nécessite une étape de modélisation, appelée problème direct, qui permet, à partir d'un modèle donné et en utilisant les lois de la physique, de simuler des données observées (Fig. 2.1).

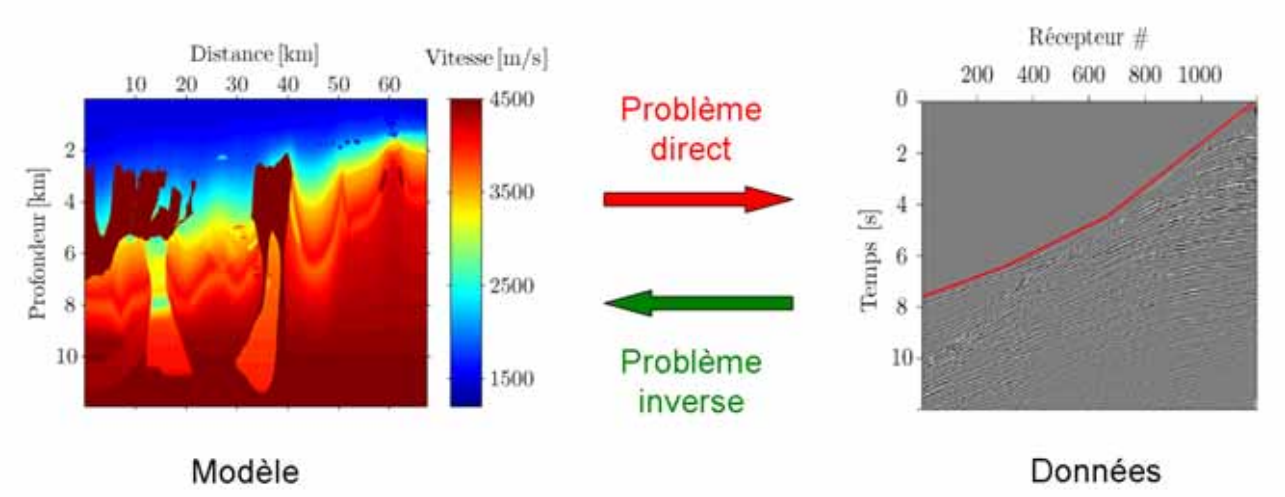

Fig. 2.1 – *Les problèmes direct et inverse en tomographie des temps de première arrivée relient un modèle de vitesse aux temps de première arrivée pointés (en ligne rouge) sur un sismogramme.*

Dans ce chapitre, nous développons tout d'abord la formulation mathématique classiquement utilisée pour la définition et la résolution des problèmes direct et inverse. Nous présentons ensuite la mise en œuvre pratique d'un algorithme de tomographie de temps de première arrivée défini à partir de cette formulation. Enfin, nous exposons les raisons qui nous ont conduits à considérer une nouvelle formulation pour la résolution du problème inverse.

2.1 Le problème direct

L'objectif de cette partie est de déterminer la relation physique $\mathbf{g}$ qui existe entre les données $\mathbf{d}$ et le modèle physique étudié $\mathbf{m}$

$$\mathbf{d} = \mathbf{g}(\mathbf{m}) . \tag{2.1}$$

Dans le cadre de ce travail, les données sont les temps de première arrivée et le modèle étudié est un modèle de vitesse de propagation des ondes sismiques. Nous exposons pour commencer, les hypothèses et les approximations faites pour permettre la définition de la loi de modélisation physique. Puis nous illustrons le développement de l'équation eikonale dont la résolution numérique permet le calcul des temps de première arrivée pour un modèle de vitesse donné.

2.1.1 Hypothèses et approximations

Nous nous intéressons dans ce travail à la propagation *asymptotique haute fréquence* d'ondes sismiques dans un milieu élastique *isotrope non homogène*. L'approximation asymptotique haute fréquence, aussi appelée approximation des rais, considère que la longueur d'onde des signaux est très faible devant les longueurs caractéristiques du milieu de propagation. Le champ d'onde peut alors être découplé entre un champ d'onde scalaire (ondes P) et un champ d'onde vectoriel (ondes S) qui se propagent indépendamment. Leur temps de trajet et leur amplitude sont alors définis à partir d'une solution particulière de l'équation de l'élastodynamique, à savoir respectivement l'équation eikonale et l'équation de transport [Červený, 2001], [Chapman, 2004].

2.1.2 L'équation eikonale

Nous choisissons d'illustrer le calcul de l'équation eikonale dans le cas acoustique [Baina, 1998]. La propagation d'une onde élastique P dans un milieu solide est alors assimilée à la propagation d'une onde de pression dans un fluide.

De l'équation des ondes à l'équation eikonale

Nous développons ici les calculs permettant d'obtenir l'équation eikonale à partir de l'équation d'onde scalaire en temps hors de la zone source [Červený, 2001]

$$K(\mathbf{x}) \, \nabla \cdot \left[\frac{1}{\rho(\mathbf{x})} \nabla P(\mathbf{x}, \tau) \right] = \frac{\partial^2}{\partial \tau^2} P(\mathbf{x}, \tau) \ , \tag{2.2}$$

où $P(\mathbf{x}, \tau)$ est la pression acoustique subie par la particule au point courant $\mathbf{x}$ à l'instant τ, $K(\mathbf{x})$ est le module d'incompressibilité et $\rho(\mathbf{x})$ est la masse volumique ou densité. Après transformation de Fourier sur la variable temporelle τ, on obtient dans le domaine fréquentiel l'équation d'Helmholtz

$$K(\mathbf{x}) \, \nabla \cdot \left[\frac{1}{\rho(\mathbf{x})} \nabla P(\mathbf{x}, \omega) \right] + \omega^2 P(\mathbf{x}, \omega) = 0 . \tag{2.3}$$

Si on adopte pour solution générale une solution analogue à celle obtenue dans un milieu homogène [Baina, 1998], on a

$$P(\mathbf{x}, \omega) = A(\mathbf{x}, \omega) \, e^{i\omega t(\mathbf{x})} , \tag{2.4}$$

où $A(\mathbf{x},\omega)$ et $t(\mathbf{x})$ sont respectivement l'amplitude et le temps de parcours de l'onde sismique. Cette solution une fois injectée dans l'équation (2.3), fournit après développement

$$|\nabla t(\mathbf{x})|^2 = \frac{1}{v^2(\mathbf{x})} + \omega^{-2}\frac{\nabla^2 A(\mathbf{x},\omega)}{A(\mathbf{x},\omega)} - \omega^{-2}\nabla\left(ln\rho(\mathbf{x})\right)\cdot\frac{\nabla A(\mathbf{x},\omega)}{A(\mathbf{x},\omega)} + \frac{i\omega^{-1}}{\rho(\mathbf{x},\omega)A^2(\mathbf{x},\omega)}\left[\nabla\left(\rho(\mathbf{x})A^2(\mathbf{x},\omega)\nabla t(\mathbf{x})\right)\right], \tag{2.5}$$

où $v(\mathbf{x}) = \sqrt{\frac{K}{\rho}(\mathbf{x})}$ est la vitesse de propagation des ondes de compression au point $\mathbf{x}$. En découplant la partie réelle de la partie imaginaire, on obtient l'équation de transport

$$\nabla\cdot\left(\rho(\mathbf{x})A^2(\mathbf{x},\omega)\nabla t(\mathbf{x})\right) = 0, \tag{2.6}$$

et l'équation dite hyper-eikonale [Menahem and Beydoun, 1985]

$$|\nabla t(\mathbf{x})|^2 = \frac{1}{v^2(\mathbf{x})}\left[1+\left(\frac{\omega_c}{\omega}\right)^2\right], \tag{2.7}$$

où l'on a utilisé la notation suivante :

$$\omega_c^2 = \frac{\nabla^2 A(\mathbf{x},\omega)}{A(\mathbf{x},\omega)} - \nabla\left(ln\rho(\mathbf{x})\right)\cdot\frac{\nabla A(\mathbf{x},\omega)}{A(\mathbf{x},\omega)}. \tag{2.8}$$

Le système d'équations couplées formé par (2.6) et (2.7) est équivalent à l'équation d'Helmholtz puisqu'il a été obtenu sans aucune approximation.

Une interprétation physique

L'équation de transport (2.6) exprime la conservation de l'énergie dans un tube de rais. L'équation hyper-eikonale (2.7) donne l'expression de la vitesse de phase en fonction de la fréquence de l'onde. Elle peut être vue comme une équation eikonale dans laquelle la lenteur effective du milieu serait corrigée par un terme d'autant plus important que la fréquence du signal est basse et que la variation locale de courbure des fronts d'onde est forte. Cette correction est une sorte de filtre lissant localement le modèle de lenteur et les fronts d'onde produits par la même occasion. Il en résulte que, dans un modèle donné, les singularités des fronts d'onde sont d'autant plus rares, et prennent leur origine d'autant plus loin de la position de la source que la fréquence du signal propagé est basse.

Application de l'approximation haute fréquence

L'approximation asymptotique de la théorie des rais est obtenue lorsque $\frac{\omega_c}{\omega} << 1$, c'est-à-dire lorsqu'on se place dans un régime haute fréquence $\omega \to \infty$. Dans ce cadre asymptotique l'équation hyper-eikonale se réduit à l'équation eikonale :

$$|\nabla t(\mathbf{x})|^2 = \frac{1}{v^2(\mathbf{x})}. \tag{2.9}$$

Les temps de trajet $t(\mathbf{x})$ deviennent alors indépendants de l'amplitude et le système d'équations découplées eikonale - transport peut alors être résolu d'une manière séquentielle. Notons que dans ce travail qui traite de la tomographie des temps de première arrivée, nous n'exploitons pas l'information contenue dans l'équation de transport.

2.1.3 Une relation non linéaire

Mises à part certaines lois explicites et particulières, comme celle à gradient constant de vitesse ou celle à gradient constant de lenteur au carré, il n'existe pas de solution analytique à l'équation eikonale [Červený, 2001]. Dans le cas général, il est indispensable de recourir à des méthodes de résolution numérique. Ces méthodes ont pour objectif de fournir une solution à l'équation eikonale (2.9) qui correspond aux temps de première arrivée. La relation $\mathbf{g}$ qui lie les temps de première arrivée au modèle de vitesse est une relation non-linéaire.

2.2 Le problème inverse

La résolution du problème inverse se propose de retrouver les paramètres physiques permettant d'expliquer au mieux les données observées. Tarantola et Valette [1982] posent les bases de la théorie stochastique de l'inversion. Le principe est de considérer le problème inverse en terme d'états de connaissance *a priori* et *a posteriori* sur les données et les modèles. Ce formalisme permet la prise en compte d'information *a priori* et des statistiques des incertitudes, expérimentales et théoriques, sur les modèles et les données. A partir d'hypothèses faites sur la nature de ces incertitudes il est possible de définir une fonction coût, dite des moindres carrés, qui mesure l'écart entre les données observées et celles théoriques obtenues pour un modèle donné. La minimisation de cette fonction coût permet alors de déterminer le modèle qui explique au mieux les données observées au sens des moindres carrés [Tarantola and Valette, 1982].

2.2.1 La fonction coût des moindres carrés

Nous présentons ici les hypothèses permettant de définir la fonction coût classiquement utilisée en tomographie des temps de première arrivée [Tarantola, 1987a]. Nous considérons comme *gaussiennes* les statistiques des incertitudes sur les modèles et les données. Cette hypothèse gaussienne est justifiée par le théorème central limite, mais elle ne peut pas toujours être vérifiée pour une expérience réelle, i.e. avec un nombre fini de réalisations des variables mises en jeu. Le modèle de vitesse $\mathbf{m}$ le plus vraisemblablement à l'origine des données observées $\mathbf{d}_{obs}$ se situe alors au minimum de la fonction coût $\mathcal{C}$, aussi appelée fonction des moindres carrés,

$$\mathcal{C}(\mathbf{m}) = \frac{1}{2}\Big[(\mathbf{g}(\mathbf{m}) - \mathbf{d}_{obs})^t \, {\mathbf{C}_{\mathcal{D}}}^{-1} (\mathbf{g}(\mathbf{m}) - \mathbf{d}_{obs}) + (\mathbf{m} - \mathbf{m}_{prior})^t \, {\mathbf{C}_{\mathcal{M}}}^{-1} (\mathbf{m} - \mathbf{m}_{prior}) \Big], \tag{2.10}$$

où $\mathbf{C}_{\mathcal{D}}$ et $\mathbf{C}_{\mathcal{M}}$ sont respectivement les matrices de covariance *a priori* sur l'espace des données $\mathcal{D}$ et l'espace des modèles $\mathcal{M}$, $\mathbf{m}_{prior}$ est un modèle connu *a priori* et ^{t} désigne l'opérateur transposée. Les méthodes d'inversion stochastiques cherchent donc à minimiser la fonction coût $\mathcal{C}$ par rapport au modèle $\mathbf{m}$.

2.2.2 La méthode de Gauss-Newton

L'approche choisie pour minimiser la fonction coût $\mathcal{C}$ est l'utilisation de méthodes itératives locales. Ces approches locales convergent par un cheminement dans l'espace des modèles. Pour la méthode de Gauss-Newton, ce cheminement est basé sur l'utilisation du gradient de la fonction

coût, qui est lié à la direction de descente, et du Hessien de la fonction coût qui lui en donne la courbure. Le schéma itératif, à l'itération $n+1$, de la méthode de Gauss-Newton s'écrit sous la forme [Tarantola, 1987a]

$$\mathbf{m}_{n+1} = \mathbf{m}_n - \mathbf{H}_n{}^{-1}\,\boldsymbol{\gamma}_n\,, \tag{2.11}$$

où $\boldsymbol{\gamma}_n$ représente le gradient,

$$\boldsymbol{\gamma}_n = \left(\frac{\partial \mathcal{C}}{\partial \mathbf{m}}\right)_{\mathbf{m}_n}, \tag{2.12}$$

et $\mathbf{H}_n$ est le Hessien de la fonction coût $\mathcal{C}$ au point $\mathbf{m}_n$

$$\mathbf{H}_n = \left(\frac{\partial^2 \mathcal{C}}{\partial \mathbf{m}^2}\right)_{\mathbf{m}_n}. \tag{2.13}$$

Schématiquement cela revient à faire un pas dans la direction donnée par le gradient à partir du modèle courant $\mathbf{m}_n$ jusqu'au minimum du paraboloïde défini par le Hessien tangent à la fonction coût en ce point (Fig. 2.2). La méthode de Gauss-Newton a pour propriété de converger en une itération lorsque la fonction coût à minimiser est quadratique. On choisit alors classiquement de linéariser localement le problème direct (2.1) autour du modèle courant $\mathbf{m}_n$, ce qui permet de rendre la fonction coût $\mathcal{C}$ à minimiser quadratique [Baina, 1998]. Ainsi à chaque itération, la méthode de Gauss-Newton fournit la solution exacte au problème localement linéarisé.

2.2.3 La solution des moindres carrés

La linéarisation du problème direct

La linéarisation du problème direct autour du modèle courant $\mathbf{m}_n$ s'écrit

$$\mathbf{g}\left(\mathbf{m}\right) \simeq \mathbf{g}\left(\mathbf{m}_n\right) + \mathbf{G}_n\,\left(\mathbf{m} - \mathbf{m}_n\right) = \mathbf{g}\left(\mathbf{m}_n\right) + \mathbf{G}_n\,\delta\mathbf{m}_n\,, \tag{2.14}$$

où l'opérateur linéaire $\mathbf{G}_n$, aussi appelé matrice des dérivées de Fréchet, représente la dérivée de $\mathbf{g}$ au point $\mathbf{m} = \mathbf{m}_n$

$$\mathbf{G}_n = \left(\frac{\partial \mathbf{g}}{\partial \mathbf{m}}\right)_{\mathbf{m}_n}, \tag{2.15}$$

et

$$\delta\mathbf{m}_n = \mathbf{m} - \mathbf{m}_n\,. \tag{2.16}$$

Le symbole $\simeq$ dans l'équation (2.14) signifie que les termes de second ordre peuvent être négligés par rapport aux erreurs d'observation et de modélisation, i.e. comparés aux valeurs d'écart-type et de corrélation présentes dans $\mathbf{C}_{\mathcal{D}}$. La fonction coût $\mathcal{C}$ donnée par (2.10) s'écrit alors

$$\begin{aligned}\mathcal{C}\left(\mathbf{m}\right) &= \mathcal{C}\left(\mathbf{m}_n + \delta\mathbf{m}_n\right)\\ &= \frac{1}{2}\Big[\left(\mathbf{G}_n\,\delta\mathbf{m}_n - \delta\mathbf{d}_n\right)^t \mathbf{C}_{\mathcal{D}}{}^{-1}\,\left(\mathbf{G}_n\,\delta\mathbf{m}_n - \delta\mathbf{d}_n\right)\\ &\qquad + \left(\delta\mathbf{m}_n - \boldsymbol{\Delta}\mathbf{m}_n\right)^t \mathbf{C}_{\mathcal{M}}{}^{-1}\,\left(\delta\mathbf{m}_n - \boldsymbol{\Delta}\mathbf{m}_n\right)\Big]\,,\end{aligned} \tag{2.17}$$

où l'on a utilisé les notations suivantes :

$$\delta\mathbf{d}_n = \mathbf{d}_{obs} - \mathbf{g}\left(\mathbf{m}_n\right)\,, \tag{2.18}$$

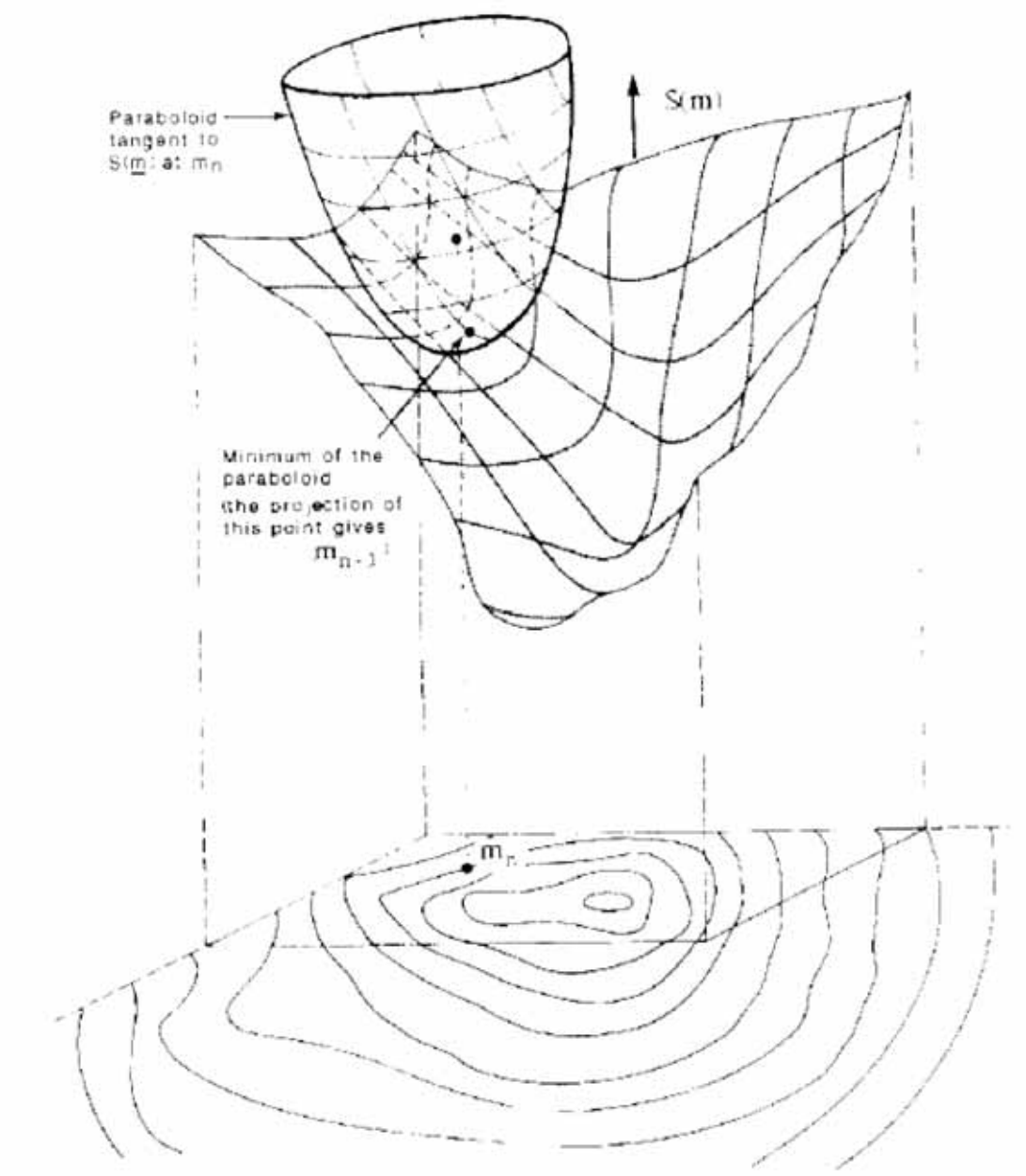

Fig. 2.2 – *Minimisation d'une fonction coût par une méthode de Gauss-Newton. Cette figure illustre, pour un problème à deux dimensions, les courbes de niveau et la surface représentant la fonction coût. Le paraboloïde tangent à la fonction coût au point courant est défini en utilisant l'information de courbure. Le point mis à jour correspond à la projection du minimum du paraboloïde, extrait de [Tarantola, 1987a].*

et

$$\Delta \mathbf{m}_n = \mathbf{m}_{prior} - \mathbf{m}_n \,, \tag{2.19}$$

La fonction coût $\mathcal{C}$ est alors quadratique en $\delta \mathbf{m}_n$.

La solution des moindres carrés

On peut alors exprimer directement la solution des moindres carrés $\delta\mathbf{m}_n$ associée au problème linéarisé localement de minimisation de la fonction coût $\mathcal{C}$ [Tarantola, 1987a]

$$\delta\mathbf{m}_n = \left({\mathbf{G}_n}^t\,{\mathbf{C}_\mathcal{D}}^{-1}\,\mathbf{G}_n + {\mathbf{C}_\mathcal{M}}^{-1}\right)^{-1}\left({\mathbf{G}_n}^t\,{\mathbf{C}_\mathcal{D}}^{-1}\,\delta\mathbf{d}_n + {\mathbf{C}_\mathcal{M}}^{-1}\,\Delta\mathbf{m}_n\right)\,. \tag{2.20}$$

En choisissant $\mathbf{m} = \mathbf{m}_{n+1}$, on obtient alors

$$\mathbf{m}_{n+1} = \mathbf{m}_n + \delta\mathbf{m}_n\,. \tag{2.21}$$

Par identification avec (2.11), on peut en déduire l'expression du gradient

$$\boldsymbol{\gamma}_n = -\left({\mathbf{G}_n}^t\,{\mathbf{C}_\mathcal{D}}^{-1}\,\delta\mathbf{d}_n + {\mathbf{C}_\mathcal{M}}^{-1}\,\Delta\mathbf{m}_n\right)\,, \tag{2.22}$$

et du Hessien de la fonction coût $\mathcal{C}$ au point $\mathbf{m}_n$

$$\mathbf{H}_n = {\mathbf{G}_n}^t\,{\mathbf{C}_\mathcal{D}}^{-1}\,\mathbf{G}_n + {\mathbf{C}_\mathcal{M}}^{-1}\,. \tag{2.23}$$

Un schéma récapitulatif

Le schéma (Fig. 2.3) reprend les principaux calculs mathématiques réalisés pour une itération de la méthode de Gauss-Newton associée au problème localement linéarisé (2.21).

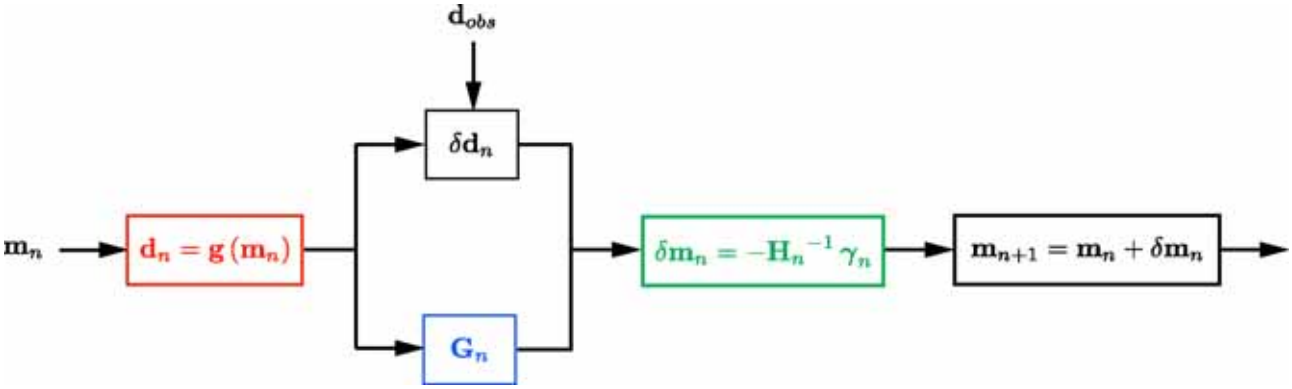

Fig. 2.3 – *Diagramme représentant les principaux calculs mathématiques réalisés pour une itération de la méthode de Gauss-Newton associée au problème localement linéarisé. Les notations mathématiques sont définies à la partie 2.2.*

Le processus itératif fournit alors un modèle $\mathbf{m}$ le plus vraisemblablement à l'origine des données observées $\mathbf{d}_{obs}$.

2.2.4 Application à la tomographie des temps de première arrivée

On vient de voir les principaux développements mathématiques de la théorie de l'inversion et de la méthode de minimisation de Gauss-Newton. On va maintenant appliquer ces différentes formules au problème de la tomographie des temps de première arrivée. Le modèle recherché $\mathbf{m}$ est donc un modèle de vitesse $\mathbf{v}$, ou indifféremment un modèle de lenteur $\mathbf{s}$ liés par la relation

$$\mathbf{v} = \frac{1}{\mathbf{s}}\,. \tag{2.24}$$

Le problème direct

Le schéma (2.4) est un exemple de discrétisation en grille cartésienne d'un modèle de lenteur $\mathbf{s} = (s^j)_j$. Nous avons aussi représenté sur ce schéma la trajectoire d'un rai entre une source et un récepteur. Un rai correspond, selon le principe de Fermat, à la trajectoire des temps extrema, ou d'un point de vue de la propagation des ondes, en milieu isotrope, à la trajectoire perpendiculaire aux fronts d'onde.

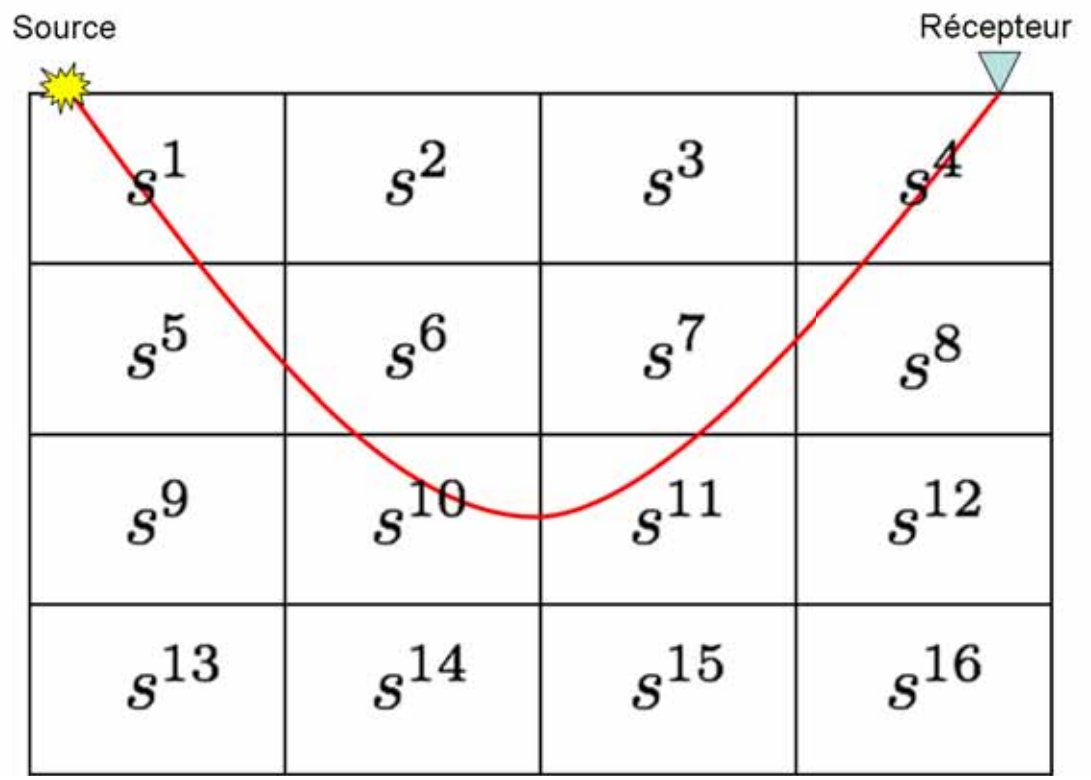

Fig. 2.4 – *Schéma illustrant la discrétisation en grille cartésienne d'un modèle de lenteur* $\mathbf{s} = (s^j)_j$ *et la trajectoire d'un rai entre une source et un récepteur.*

Les données observées $\mathbf{d}_{obs}$ correspondent aux temps de première arrivée pointés sur les sismogrammes $\mathbf{t}_{obs}$. Les données synthétiques $\mathbf{d}$ sont les temps de première arrivée $\mathbf{t} = (t^i)_i$ calculés pour un modèle $\mathbf{s}$ donné par la résolution de l'équation eikonale (2.9). On peut alors établir la relation entre le temps de première arrivée t^i, le modèle de lenteur $(s^j)_j$ et les longueurs des segments du rai $l^{i,j}$ dans chaque maille du milieu discrétisé

$$t^i = \sum_j l^{i,j}\, s^j \,. \tag{2.25}$$

En utilisant une notation matricielle, on obtient la formule suivante :

$$\mathbf{t} = \mathbf{L}\,\mathbf{s}\,, \tag{2.26}$$

avec $\mathbf{L} = (l^{i,j})_{i,j}$ la matrice des longueurs de rais. Cette équation correspond à la formulation du problème direct (2.1).

Le problème inverse

Nous faisons tout d'abord quelques choix permettant de simplifier l'expression de la fonction coût à minimiser (2.10). Nous ne prenons en compte aucune information *a priori*, i.e. $\mathbf{C}_{\mathcal{M}}{}^{-1} = \mathbf{0}$. De plus nous supposons que les erreurs sur les données sont indépendantes et possèdent la même variance égale à 1, i.e. $\mathbf{C}_{\mathcal{D}}{}^{-1} = \mathbf{I}$. Sous ces hypothèses, la fonction coût à minimiser s'écrit

$$\mathcal{C}(\mathbf{s}) = \frac{1}{2} (\mathbf{L}\,\mathbf{s} - \mathbf{t}_{obs})^t (\mathbf{L}\,\mathbf{s} - \mathbf{t}_{obs}) . \tag{2.27}$$

Considérons maintenant un modèle de lenteur de référence $\mathbf{s}_n$, la matrice des longueurs de rais $\mathbf{L}_n$ et les temps de première arrivée $\mathbf{t}_n$ qui lui sont associés. Soient $\mathbf{s} = \mathbf{s}_n + \delta\mathbf{s}_n$, avec $\delta\mathbf{s}_n$ une perturbation de lenteur, et $\mathbf{t}$ le temps correspondant. Le principe de Fermat permet d'établir qu'une petite perturbation $\delta\mathbf{s}_n$ de la loi de lenteur n'affecte les temps d'arrivée qu'au second ordre, c'est-à-dire que la trajectoire du rai n'est pas perturbée lorsque la variation de lenteur locale est faible, ce qui nous donne

$$\mathbf{t} \simeq \mathbf{L}_n\,\mathbf{s}_n + \mathbf{L}_n\,\delta\mathbf{s}_n = \mathbf{t}_n + \mathbf{L}_n\,\delta\mathbf{s}_n . \tag{2.28}$$

Cette étape correspond à la linéarisation du problème direct formulée à l'équation (2.14). On peut alors établir que $\mathbf{L}_n = \mathbf{G}_n$, c'est dire que la matrice des longueurs de rais correspond à la matrice des dérivées de Fréchet. En posant $\delta\mathbf{t}_n = \mathbf{t}^{obs} - \mathbf{t}_n$, on obtient

$$\delta\mathbf{t}_n = \mathbf{L}_n\,\delta\mathbf{s}_n . \tag{2.29}$$

La solution des moindres carrés associée au problème de minimisation de la fonction coût $\mathcal{C}$ est donnée par (2.20)

$$\delta\mathbf{s}_n \simeq \left(\mathbf{L}_n{}^t\,\mathbf{L}_n\right)^{-1} \mathbf{L}_n{}^t\,\delta\mathbf{t}_n . \tag{2.30}$$

Nous utilisons ici le symbole $\simeq$ car la matrice $\mathbf{L}_n$ est en général mal conditionnée et donc l'inverse de $\mathbf{L}_n{}^t\,\mathbf{L}_n$ est difficile à estimer. La solution donnée par (2.30) ne peut donc pas en général être calculée directement. On utilise alors des méthodes de résolution de système algébrique ou d'estimation de l'inverse généralisé correspondant associé au système linéaire (2.29) équivalent. Le gradient et le Hessien de la fonction coût $\mathcal{C}$ sont alors

$$\boldsymbol{\gamma}_n = -\mathbf{L}_n{}^t\,\delta\mathbf{t}_n , \tag{2.31}$$

et

$$\mathbf{H}_n = \mathbf{L}_n{}^t\,\mathbf{L}_n . \tag{2.32}$$

Le schéma itératif donné par la méthode de Gauss-Newton (2.21) permet alors de déterminer un modèle de lenteur $\mathbf{s}$ qui explique au mieux les temps de première arrivée observés à la surface $\mathbf{t}_{obs}$.

2.3 La mise en œuvre pratique

Nous venons de présenter la formulation mathématique conventionnelle des problèmes direct et inverse associés à la tomographie de temps de trajet de première arrivée. Nous nous intéressons maintenant à la mise en oeuvre pratique d'un algorithme de tomographie des temps de première arrivée défini à partir de cette formulation. Nous présentons dans cette partie les méthodes classiquement utilisées pour le calcul des temps de première arrivée $\mathbf{t}$, pour la construction de la matrice des dérivées de Fréchet $\mathbf{L}_n$ et pour la résolution du système linéaire qui permet la détermination de la perturbation de lenteur $\delta\mathbf{s}_n$.

2.3.1 La discrétisation du modèle de vitesse

Le choix de la discrétisation retenue pour décrire le modèle du vitesse dépend de l'application considérée et des algorithmes mis en œuvre. Dans le cadre de la tomographie des temps de première arrivée, le modèle de vitesse est classiquement discrétisé sur une grille cartésienne. Cette discrétisation permet la prise en compte de variations latérales de vitesse et elle ne nécessite aucune connaissance *a priori* du milieu, contrairement à la discrétisation par couches où un nombre donné de couches est considéré. De plus, elle est totalement adaptée aux algorithmes mis en œuvre pour le calcul des temps de première arrivée et pour la résolution du problème inverse.

2.3.2 Le calcul des temps de première arrivée

Le calcul des temps de première arrivée repose sur une résolution numérique de l'équation eikonale (2.9). L'approche la plus couramment utilisée est la résolution par différences finies introduite par Vidale [1988]. Les algorithmes ainsi définis calculent les temps de première arrivée sur des grilles cartésiennes, à partir d'un modèle de vitesse lui aussi défini sur une grille cartésienne avec le même espacement de grille.
L'implémentation d'un solveur eikonal peut être décomposée en deux problèmes :

- un problème local, qui a pour objectif de définir le schéma aux différences finies utilisé pour calculer le temps de première arrivée en un point donné de la grille en fonction des valeurs de temps des points voisins ;
- un problème global, qui permet de déterminer dans quel ordre les points de la grille doivent être considérés.

L'article de Vidale [1988] a rapidement été suivi par des dizaines d'articles introduisant des variantes de schémas aux différences finies développés pour améliorer la rapidité, la robustesse, la stabilité, ou la précision, par exemple [van Trier and Symes, 1991], [Podvin and Lecomte, 1991], [Qin et al., 1992], [Pica, 1997]. De nombreux articles continuent d'être publiés chaque année sur le sujet donnant ainsi naissance à des nouveaux noms de méthodes telles que la *Fast Marching Method* [Sethian, 1996], [Popovici and Sethian, 1998], ou la *Fast Sweeping Method* [Boué and Dupuis, 1999], [Zhao, 2005]. Des variantes de ces méthodes permettent aussi de prendre en compte par exemple l'anisotropie du milieu [Lecomte, 1993], [Qian et al., 2007] ou bien encore de réaliser une implémentation parallèle de l'algorithme [Zhao, 2006].

2.3.3 Le tracé de rais *a posteriori*

La résolution par différences finies de l'équation eikonale fournit une grille des temps de première arrivée en tout point du modèle. Il est alors ensuite possible de tracer les rais qui relient la source aux récepteurs de manière *a posteriori* [Podvin and Lecomte, 1991], [Baina, 1998]. Une méthode de calcul de la trajectoire des rais *a posteriori* repose sur le principe de stationnarité proposé par Vidale [1988]. On part du récepteur, pour une meilleure stabilité numérique, et on avance itérativement vers la source en suivant la direction opposée au gradient des temps. Ce gradient est estimé localement par un schéma aux différences finies centré sur un nœud du réseau ou bien sur le centre de la maille selon leur proximité de la position du point courant. Le tracé de rais *a posteriori* réalisé pour toutes les positions des couples {source - récepteur}

permet alors de calculer la matrice des dérivées de Fréchet $\mathbf{L}_n$. On calcule pour cela en chaque maille du modèle (Fig. 2.4) la longueur du segment du rai qui la traverse.

2.3.4 La résolution du système linéaire tomographique

La détermination de la perturbation de lenteur $\delta\mathbf{s}_n$ passe par la résolution au sens des moindres carrés du système linéaire tomographique (2.29). Plusieurs méthodes de résolution de système algébrique ou d'estimation de l'inverse généralisé correspondant existent. On peut en distinguer deux classes : les méthodes de résolution directe comme la factorisation de Cholesky, la décomposition LU ou la décomposition par valeurs singulières et les méthodes de résolution itérative comme les méthodes de projection et de reconstruction. Les caractéristiques de la matrice des dérivées de Fréchet $\mathbf{L}_n$ sont sa grande taille [*nombre de données* $\times$ *nombre de mailles*], le fait qu'elle soit creuse, c'est-à-dire qu'elle contient essentiellement des éléments nuls, et son mauvais conditionnement, du à la couverture inhomogène des rais. Ces caractéristiques font que les méthodes itératives sont les mieux adaptées à la résolution du système tomographique [van der Sluis and van der Vorst, 1987], [Baina, 1998].

Les méthodes de reconstruction

Les méthodes de reconstruction ont pour principe général de rétropropager à chaque itération les résidus des temps sur les mailles traversées par les rais calculés pour un modèle donné. Cette rétropropagation est pondérée par des coefficients liés à la couverture des rais dans chaque maille. La méthode de reconstruction la plus répandue est la méthode *Simultaneaous Iterative Reconstruction Technique* (SIRT). van der Sluis et van der Vorst [1987] ont prouvé théoriquement la convergence de cette méthode vers une solution au sens des moindres carrés pondérés. A partir de cette formulation commune, de nombreuses variantes ont été développées pour améliorer la rapidité de convergence [Zelt and Barton, 1998] ou bien prendre en compte les zones de Fresnel [Watanabe et al., 1999], [Grandjean and Sage, 2004]. L'avantage principal des méthodes de reconstruction est leur simplicité de mise en œuvre. Par contre, ces méthodes souffrent d'une renormalisation intrinsèque qui modifie le système linéaire à résoudre et donc la solution obtenue [van der Sluis and van der Vorst, 1987].

Les méthodes de projection

La méthode de projection la plus utilisée est la méthode *Least Squares QR* (LSQR) développée par Paige et Saunders [1982]. Dans son principe, cette méthode ressemble à la fois à la méthode itérative du gradient conjugué et à l'inversion par décomposition par valeurs singulières, de façon à tirer profit des avantages des deux méthodes simultanément [Baina, 1998]. La méthode LSQR, en calculant à chaque itération des estimateurs de la valeur du conditionnement, permet de prévenir les instabilités numériques. Nolet [1985] a montré, lors d'une étude comparative, que l'algorithme LSQR est supérieur aux algorithmes de reconstruction, tant du point de vue de la vitesse de convergence que de la stabilité numérique. De nombreux algorithmes de tomographie des temps de première arrivée utilisent l'algorithme LSQR pour déterminer la perturbation de lenteur $\delta\mathbf{s}_n$, par exemple [Spakman and Nolet, 1988], [Le Meur, 1994], [Baina, 1998], [Zelt and Barton, 1998].

2.3.5 Schéma de l'algorithme de tomographie

Le diagramme (Fig. 2.5) reprend les principales étapes algorithmiques aboutissant à la détermination de la perturbation de lenteur $\delta\mathbf{s}_n$ pour une itération de la méthode de Gauss-Newton. A partir d'un modèle de lenteur courant $\mathbf{s}_n$, la résolution du problème direct pour N points de tirs fournit les cartes des temps de première arrivée en tous points de la grille. Pour toutes les positions {source - récepteur}, un tracé de rais *a posteriori* permet la construction de la matrice des dérivées de Fréchet. Cette matrice et les résidus calculés à partir des temps observés à la surface sont ensuite utilisés pour la résolution itérative du système linéaire tomographique. Cette résolution permet de déterminer la perturbation de lenteur $\delta\mathbf{s}_n$ à appliquer au modèle courant.

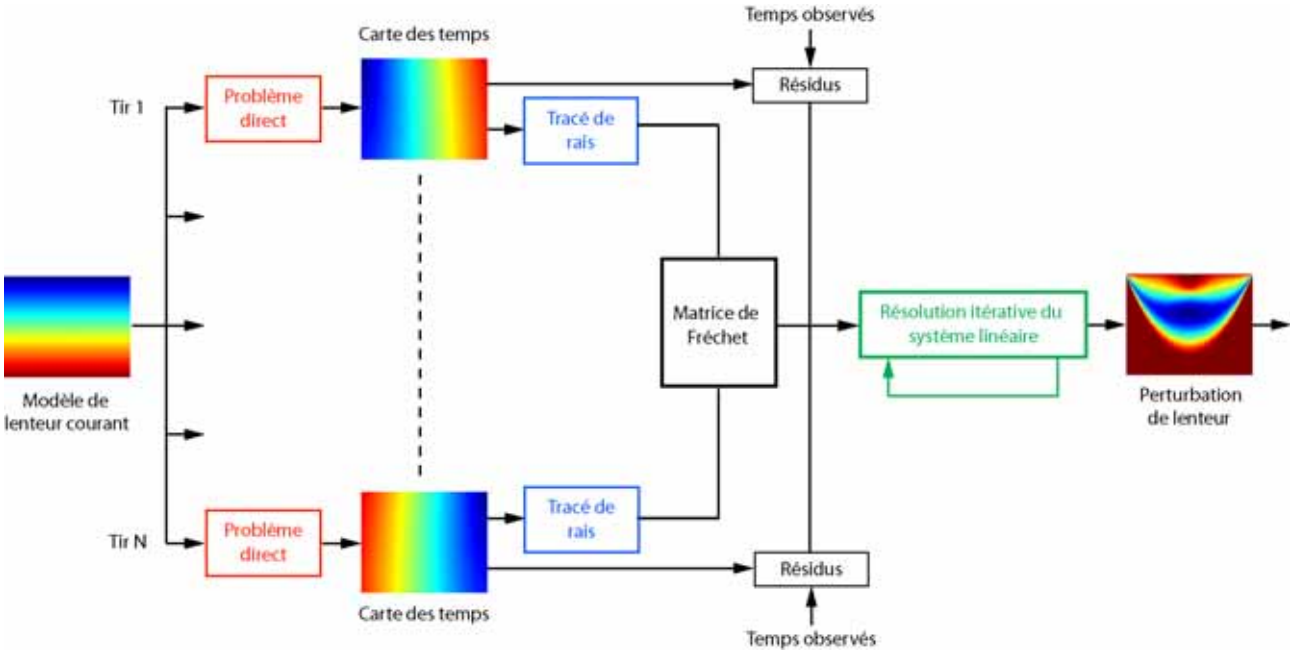

Fig. 2.5 – *Diagramme illustrant les principales étapes de l'algorithme de tomographie de temps de trajet de première arrivée défini à partir de la méthode de Gauss-Newton.*

Le schéma itératif donné par la méthode de Gauss-Newton (2.21) permet alors de déterminer un modèle de lenteur $\mathbf{s}$ qui explique au mieux les temps de première arrivée observés à la surface $\mathbf{t}_{obs}$.

2.4 Vers un changement de méthode mathématique

Nous venons de présenter la formulation et les principaux éléments d'un algorithme classiquement défini pour résoudre le problème de tomographie des temps de première arrivée. La formulation repose sur le choix de la méthode de Gauss-Newton pour la minimisation de la fonction coût des moindres carrés. Nous présentons dans cette partie le contexte, géophysique et informatique, de la tomographie des temps de première arrivée qui est en pleine évolution. Ce changement de contexte permet de révéler les limitations pratiques rencontrées par de nom-

breux algorithmes de tomographie actuels. Ces limitations pratiques sont essentiellement dues à l'estimation de l'inverse du Hessien nécessaire au schéma itératif de Gauss-Newton.

2.4.1 Un contexte qui évolue

Le contexte géophysique

La taille des dispositifs d'acquisition sismique et la dimension des volumes étudiés ne cessent de croitre. Vesnaver [2008] donne quelques chiffres caractéristiques des acquisitions sismiques 3-D actuelles et futures. Les systèmes d'enregistrement avec plusieurs milliers de récepteurs sont devenus des standards, et des acquisitions géophysiques pétrolières mettant en œuvre jusqu'à 36 000 canaux sont maintenant réalisées aussi bien sur mer que sur terre. Pour une acquisition sismique 3-D de taille raisonnable, par exemple sur une surface de 200 km^2 avec une densité de 250 points de tir par km^2, on obtient au total 50 000 points de tir. Si chaque point de tir est composé de 36 000 traces sismiques, 1.8 milliards de sismogrammes sont enregistrés. Si on considère uniquement les temps de première arrivée, l'algorithme de tomographie doit alors être capable de gérer 1.8 milliards de temps observés. De plus si on considère une profondeur étudiée de l'ordre de 5 km et une discrétisation du modèle de vitesse avec une grille cubique de 12.5 m, le nombre de mailles est d'environ 600 millions. Un algorithme de tomographie des temps de première arrivée doit donc être capable d'assimiler un nombre très important de données observées et de gérer des modèles de vitesse de grande dimension.

Le contexte informatique

Parallèlement à l'évolution du contexte géophysique, le contexte informatique connait lui aussi une évolution très rapide. Les moyens et la puissance de calcul disponibles ne cessent de croître. Cette augmentation de la puissance de calcul est essentiellement due à l'architecture parallèle des supercalculateurs actuels. Ils combinent en effet un nombre élevé de cœurs, ou unités de calcul, qui possèdent chacun une quantité de mémoire partagée ou distribuée et sont reliés entre eux par des connexions très rapides. La réalisation de calculs parallèles consistent en l'exécution d'un traitement partitionné en tâches élémentaires réparties entre plusieurs cœurs opérant simultanément. Il est ainsi possible de réduire significativement le temps de calcul nécessaire à l'exécution de ce traitement. Un algorithme de tomographie des temps de première arrivée, pour tirer pleinement profit de la puissance de calcul d'une architecture parallèle, doit donc pouvoir être facilement partitionné en tâches élémentaires.

2.4.2 Les limites pratiques de la méthode de Gauss-Newton

Le choix de la méthode de Gauss-Newton a des conséquences sur la mise en œuvre pratique d'un algorithme de tomographie des temps de première arrivée. Nous décrivons ici les limites pratiques dues au choix de cette méthode. Ces limitations sont essentiellement dues à la taille et au mauvais conditionnement de la matrice des dérivées de Fréchet $\mathbf{L}_n$. Cette matrice permet d'obtenir une estimation de l'inverse du Hessien ${\mathbf{L}_n}^t \mathbf{L}_n$ utilisé par le schéma itératif de la méthode de Gauss-Newton.

Des contraintes informatiques

Une caractéristique de la matrice des dérivées de Fréchet est sa grande taille [*nombre de données* × *nombre de mailles*], même si de par sa construction cette matrice est le plus souvent creuse. Les valeurs caractéristiques des acquisitions sismiques 3-D actuelles données à la partie (2.4), 1.8 milliards de temps pointés et 600 millions de mailles, permettent d'établir un ordre de grandeur de l'espace mémoire occupé par cette matrice. En considérant un codage des réels sur 4 octets, l'occupation mémoire des temps pointés est d'environ 7.2 Go. En supposant, qu'en moyenne, un rai est décrit sur 100 mailles, l'occupation mémoire des éléments non nuls de la matrice des dérivées de Fréchet est de l'ordre de 720 Go. Une telle quantité de données ne peut être que très difficilement maintenue en mémoire vive et nécessite donc un stockage sur disque. On peut noter que l'occupation mémoire dépend essentiellement du nombre de données observées. Pour résoudre le système linéaire tomographique, l'algorithme de tomographie doit alors effectuer un grand nombre d'accès disque qui pénalisent l'algorithme en temps de calcul. Pour limiter l'occupation mémoire, une solution consiste à décimer le nombre de données observées utilisées et/ou à réduire la discrétisation du modèle étudié. Mais cette solution peut provoquer des pertes d'information et/ou de résolution.
Par ailleurs, les méthodes itératives de résolution du système linéaire tomographique ne se prêtent guère à une parallélisation des calculs à réaliser. Ainsi la méthode LSQR est très difficilement parallélisable, pour un gain en temps très faible. Les méthodes de reconstruction possèdent *a priori* une formulation adaptée au calcul parallèle. Cependant, cette parallélisation des calculs se fait au niveau des mailles du modèle, ce qui nécessite de fait un temps de calcul très important quand la taille du modèle considéré est grande.

Une paramétrisation complexe

La résolution numérique du système linéaire tomographique est une tâche délicate à cause du mauvais conditionnement, synonyme de valeurs propres faibles, de la matrice des dérivées de Fréchet $\mathbf{L}_n$ qui rend l'estimation de l'inverse du Hessien ${\mathbf{L}_n}^t \mathbf{L}_n$ instable. La couverture inhomogène des rais est à l'origine du mauvais conditionnement de la matrice des dérivées de Fréchet. Pour prévenir les instabilités, il est nécessaire de procéder à des régularisations numériques. Ces régularisations peuvent prendre différentes formes [Tikhonov and Arsenin, 1977], [Yao and Roberts, 2002], l'approche la plus commune consiste à introduire un facteur d'amortissement, aussi appelé *damping*, dans la fonction coût à minimiser [Baina, 1998]. Par ailleurs, une paramétrisation adaptée des méthodes de résolution est indispensable pour permettre une estimation correcte de la perturbation de lenteur. Cette paramétrisation peut s'avérer complexe, par exemple les "boutons de réglage" de la fonction LSQR ou la définition des poids à appliquer au schéma itératif des méthodes de reconstruction.

2.4.3 Le choix d'une autre méthode

Les limitations pratiques rencontrées par les algorithmes de tomographie de première arrivée définis à partir de la méthode de Gauss-Newton sont essentiellement dues à l'estimation de l'inverse du Hessien de la fonction coût. Nous avons alors choisi d'utiliser une méthode de gradient pour la minimisation de la fonction coût des moindres carrés associée au problème de tomographie. L'utilisation d'une méthode de gradient permet généralement une mise en œuvre et une

résolution plus simples des problèmes inverses de grande taille [1]. Par exemple en géophysique, la résolution du problème d'inversion de forme d'onde [Lailly, 1984], [Tarantola, 1984], repose sur une méthode de gradient associée à la méthode de l'état adjoint [Lions, 1971] pour calculer le gradient de la fonction coût par rapport aux paramètres du modèle. Nous choisissons donc de formuler la résolution du problème de tomographie des temps de première arrivée à partir d'une méthode de gradient.

1. "The Newton method is not well adapted to large-sized inverse problems, which are more easily solved using gradient methods", d'après [Tarantola, 1987a].

Chapitre 3

Une nouvelle formulation

Nous exposons au chapitre précédent les raisons qui nous ont conduits à formuler la résolution du problème de tomographie des temps de première arrivée à partir d'une méthode de gradient pour la minimisation de la fonction coût des moindres carrés. Cette nouvelle formulation doit permettre la prise en compte d'un grand nombre de données observées et de modèles de vitesse de grande taille. Dans ce chapitre, nous présentons tout d'abord la méthode de gradient retenue et son application au problème de tomographie des temps de première arrivée. Ensuite, nous présentons deux approches permettant le calcul du gradient de la fonction coût par rapport aux paramètres du modèle : une formulation linéarisée à partir d'un tracé de rais *a posteriori* et une formulation non-linéaire par la méthode de l'état adjoint. La méthode de l'état adjoint permet le calcul du gradient de la fonction coût sans avoir besoin de calculer la matrice des dérivées de Fréchet. Nous illustrons alors la mise en œuvre pratique du nouvel algorithme de tomographie des temps de première arrivée défini à partir de la méthode de plus grande descente. Finalement, nous présentons les performances informatiques obtenues avec ce nouvel algorithme qui en font un outil adapté à la tomographie des temps de première arrivée.

3.1 Une méthode de gradient

Nous rappelons tout d'abord la fonction coût $\mathcal{C}$ des moindres carrés associée au problème inverse de la tomographie des temps première arrivée (2.10)

$$\mathcal{C}(\mathbf{m}) = \frac{1}{2}\Big[(\mathbf{g}(\mathbf{m}) - \mathbf{d}_{obs})^t \, {\mathbf{C}_{\mathcal{D}}}^{-1} (\mathbf{g}(\mathbf{m}) - \mathbf{d}_{obs}) + (\mathbf{m} - \mathbf{m}_{prior})^t \, {\mathbf{C}_{\mathcal{M}}}^{-1} (\mathbf{m} - \mathbf{m}_{prior}) \Big], \tag{3.1}$$

où $\mathbf{m}$ est un modèle donné, $\mathbf{g}$ est l'opérateur associé au problème direct, $\mathbf{C}_{\mathcal{D}}$ et $\mathbf{C}_{\mathcal{M}}$ sont respectivement les matrices de covariance *a priori* sur l'espace des données $\mathcal{D}$ et l'espace des modèles $\mathcal{M}$, $\mathbf{m}_{prior}$ est un modèle connu *a priori* et ^{t} désigne l'opérateur transposée. L'algorithme de tomographie des temps de première arrivée a pour objectif de minimiser $\mathcal{C}$ pour déterminer le modèle $\mathbf{m}$ qui explique au mieux les données observées $\mathbf{d}_{obs}$ au sens des moindres carrés.

3.1.1 Le schéma itératif

Il existe de nombreuses méthodes de minimisation par gradient [Tarantola, 1987a]. Dans le cadre de ce travail, nous choisissons la méthode de plus grande descente, qui a pour avantage

d'être simple à mettre en œuvre. Les développements mathématiques présentés dans cette partie sont néanmoins valables pour toutes les autres méthodes de gradient.

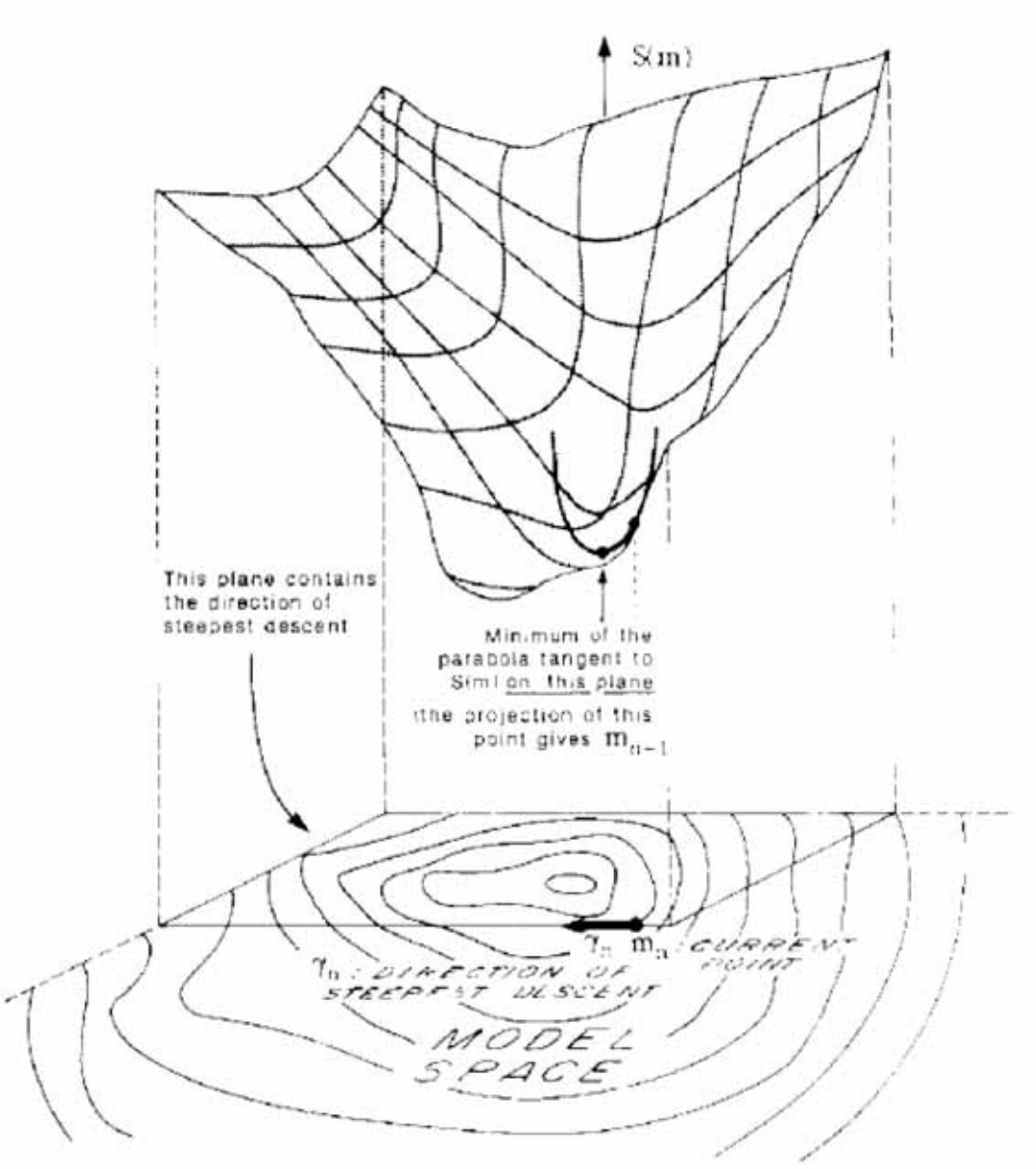

Fig. 3.1 – *Minimisation d'une fonction coût par la méthode de plus grande pente. Le minimum de la fonction coût est cherché dans la direction de plus forte descente. La méthode la plus simple pour définir le pas dans la direction de plus grande pente est de définir une parabole au point courant à l'intersection de la surface de la fonction coût et du plan contenant la direction de plus forte descente. Le point mis à jour correspond alors à la projection du minimum de la parabole, extrait de [Tarantola, 1987a]*

Le schéma itératif à l'itération $n + 1$, de la méthode de plus grande descente s'écrit sous la

forme [Tarantola, 1987a]

$$\mathbf{m}_{n+1} = \mathbf{m}_n - \mu_n \boldsymbol{\gamma}_n \,, \tag{3.2}$$

où μ_n est un réel positif défini de telle manière que $\mathcal{C}(\mathbf{m}_{n+1})$ soit inférieure à $\mathcal{C}(\mathbf{m}_n)$ et $\boldsymbol{\gamma}_n$ est le gradient de la fonction coût $\mathcal{C}$ au point $\mathbf{m}_n$. Schématiquement cela revient à faire un pas dans la direction donnée par le gradient à partir du modèle courant $\mathbf{m}_n$ (Fig.3.1). Ce schéma itératif est optimal pour des valeurs de μ_n infiniment petites ; dans ce cas, pour reprendre l'image donnée par Tarantola [1987a] : "*this would simulate the motion of a drop rain on the slope of a mountain*".
L'étape clé dans la mise en œuvre d'une méthode de gradient est le calcul du gradient de la fonction coût par rapport aux paramètres du modèle. Les performances d'un algorithme de tomographie des temps de première arrivée défini à partir de cette méthode dépendent directement de la précision et de la rapidité du calcul du gradient de la fonction coût.

3.1.2 Le gradient de la fonction coût

Dans le cas général, le gradient $\boldsymbol{\gamma}_n$ de la fonction coût $\mathcal{C}$ au point $\mathbf{m} = \mathbf{m}_n$ dérivé à partir de (3.1) s'écrit

$$\begin{aligned} \boldsymbol{\gamma}_n &= \left(\frac{\partial \mathcal{C}}{\partial \mathbf{m}} \right)_{\mathbf{m}_n} \\ &= \mathbf{G}_n{}^t \, \mathbf{C}_{\mathcal{D}}{}^{-1} \left(\mathbf{g}(\mathbf{m}_n) - \mathbf{d}_{obs} \right) + \mathbf{C}_{\mathcal{M}}{}^{-1} \left(\mathbf{m}_n - \mathbf{m}_{prior} \right) \,, \end{aligned} \tag{3.3}$$

où $\mathbf{G}_n$ est la matrice des dérivées de Fréchet. Elle représente la dérivée de $\mathbf{g}$ au point $\mathbf{m} = \mathbf{m}_n$

$$\mathbf{G}_n = \left(\frac{\partial \mathbf{g}}{\partial \mathbf{m}} \right)_{\mathbf{m}_n} . \tag{3.4}$$

Le diagramme (Fig. 3.2) reprend les principaux calculs mathématiques réalisés pour une itération de la méthode de plus grande descente.

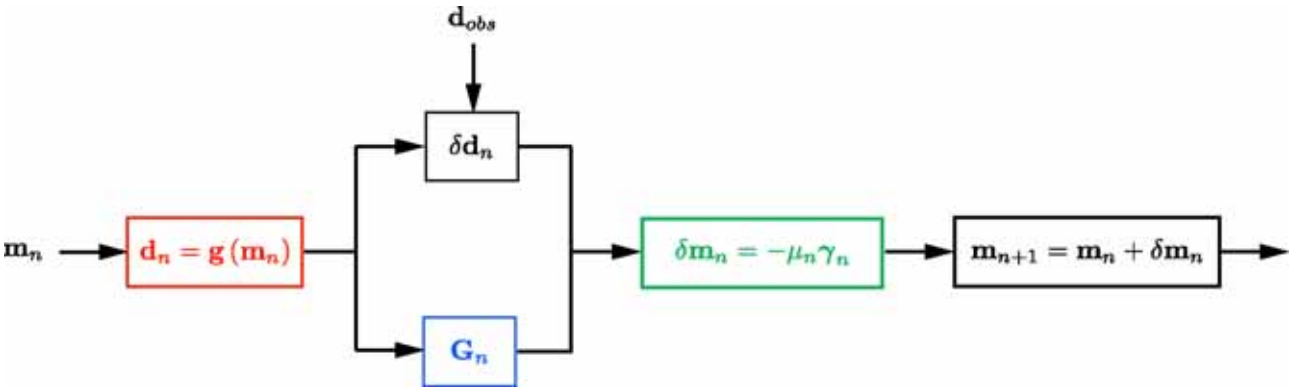

Fig. 3.2 – *Diagramme représentant les principaux calculs mathématiques réalisés pour une itération de la méthode de plus grande descente.*

Le schéma itératif donné par (3.2) fournit alors un modèle $\mathbf{m}$ le plus vraisemblablement à l'origine des données observées $\mathbf{d}_{obs}$.

3.1.3 Application à la tomographie des temps de première arrivée

Pour simplifier l'écriture des développements mathématiques présentés par la suite, nous ne prenons en compte aucune information *a priori*, i.e. ${\mathbf{C}_{\mathcal{M}}}^{-1} = \mathbf{0}$. De plus nous supposons que les erreurs sur les données sont indépendantes et possèdent la même variance égale à 1, i.e. ${\mathbf{C}_{\mathcal{D}}}^{-1} = \mathbf{I}$. L'introduction des matrices de covariance *a priori* dans les formules développées ne pose pas de problème particulier. Le modèle recherché $\mathbf{m}$ est un modèle de vitesse $\mathbf{v}$, ou indifféremment un modèle de lenteur $\mathbf{s}$. Les données observées $\mathbf{d}_{obs}$ correspondent aux temps de première arrivée pointés $\mathbf{t}_{obs}$ sur les sismogrammes. Sous ces hypothèses, la fonction coût à minimiser pour résoudre le problème de tomographie des temps de première arrivée s'écrit

$$\mathcal{C}(\mathbf{s}) = \frac{1}{2} \left(\mathbf{g}(\mathbf{s}) - \mathbf{t}_{obs}\right)^t \left(\mathbf{g}(\mathbf{s}) - \mathbf{t}_{obs}\right) . \tag{3.5}$$

Nous considérons maintenant deux approches pour le calcul du gradient de la fonction coût $\mathcal{C}$ par rapport aux paramètres du modèle.

Formulation linéarisée

Nous montrons au paragraphe 2.2.4, que la linéarisation du problème direct autour du point de référence $\mathbf{s}_n$, permet d'écrire le gradient de la fonction coût par rapport au modèle de lenteur sous la forme

$$\boldsymbol{\gamma}_n = -{\mathbf{L}_n}^t \, \delta\mathbf{t}_n \, , \tag{3.6}$$

où $\mathbf{L}_n = \mathbf{G}_n$ est la matrice des longueurs de rais, égale dans ce cas à la matrice des dérivées de Fréchet, et $\delta\mathbf{t}_n = \mathbf{t}_{obs} - \mathbf{t}_n$. La détermination du gradient de la fonction coût par rapport au modèle de lenteur nécessite donc le calcul des éléments de la matrice des dérivées de Fréchet. Les éléments de cette matrice correspondent aux longueurs des segments de rai dans chaque maille du milieu discrétisé (Fig. 2.4). Les trajectoires des rais sont généralement obtenues par un tracé de rais *a posteriori* réalisé à partir de la carte des temps de première arrivée fournie par la résolution numérique de l'équation eikonale.

Formulation non-linéaire

En considérant le problème direct non-linéaire, le gradient de la fonction coût $\boldsymbol{\gamma}_n$ dérivé à partir de (3.5) au point $\mathbf{s} = \mathbf{s}_n$ s'écrit

$$\boldsymbol{\gamma}_n = {\mathbf{G}_n}^t \left(\mathbf{g}(\mathbf{s}_n) - \mathbf{t}_{obs}\right) . \tag{3.7}$$

Nous présentons dans la prochaine partie la méthode de l'état adjoint qui permet de calculer le gradient de la fonction coût $\boldsymbol{\gamma}_n$ par rapport aux paramètres du modèle sans avoir besoin de déterminer la matrice des dérivées de Fréchet $\mathbf{G}_n$.

Gradients de la fonction coût

Les gradients de la fonction par rapport au modèle de vitesse v et par rapport au modèle de lenteur s sont liés par la relation

$$\frac{\partial \mathcal{C}}{\partial s}(s) = \frac{\partial \mathcal{C}}{\partial v}(s) \, \frac{\mathrm{d}v}{\mathrm{d}s} = -\frac{1}{s^2} \frac{\partial \mathcal{C}}{\partial v}(s) . \tag{3.8}$$

3.2 La méthode de l'état adjoint

La méthode de l'état adjoint est une méthode générale pour calculer le gradient d'une fonction coût qui dépend de variables d'état solutions d'un problème direct. On introduit alors la notion de variables de l'état adjoint solutions du problème adjoint associé. Ces variables peuvent être vues comme collectant une mesure globale de la perturbation du problème par rapport aux variables d'état [Plessix, 2006]. Numériquement cette approche est très attractive car le calcul du gradient de la fonction coût par rapport aux paramètres du modèle devient alors équivalent à la résolution du problème direct. La méthode de l'état adjoint a été introduite dans la théorie des problèmes inverses par Chavent [1974]. Cette approche venait de la théorie du contrôle [Lions, 1971]. Plusieurs auteurs ont déjà appliqué cette méthode en géophysique, par exemple [Lailly, 1984], [Tarantola, 1984], [Delprat-Jannaud et al., 1992], [Sei and Symes, 1994], [Chavent and Jacewitz, 1995], [Plessix et al., 1999], [Shen et al., 2003] et [Leung and Qian, 2006].

3.2.1 Une formulation par le Lagrangien

Pour développer les formules liées à la méthode de l'état adjoint, nous choisissons de formuler les équations sous une forme continue. Pour un point de tir, la fonction coût des moindres carrés $\mathcal{C}$ s'écrit

$$\mathcal{C}(v) = \frac{1}{2}\int_{\Omega_S} \left| t(v,\mathbf{x}) - t_{obs}(\mathbf{x}) \right|^2 \mathrm{d}\mathbf{x}\,, \tag{3.9}$$

définie entre les temps de premières arrivées t_{obs} mesurés à la surface Ω_S et les temps calculés $t(v)$ pour un modèle de vitesse v défini sur un ouvert Ω. Par souci de ne pas surcharger les formules, nous n'exprimons plus par la suite la dépendance spatiale des variables, ainsi $t(v,\mathbf{x})$ devient $t(v)$. Il existe une méthode générale [Plessix, 2006] qui permet d'établir systématiquement l'expression du gradient de la fonction coût par rapport aux paramètres du modèle et de l'équation de l'état adjoint dont une solution est la variable de l'état adjoint λ. Cette formulation repose sur la définition du Lagrangien $\mathcal{L}$

$$\mathcal{L}\big(\tilde{t},\tilde{\lambda},v\big) = \frac{1}{2}\int_{\Omega_S} \left|\tilde{t} - t_{obs}\right|^2 \mathrm{d}\mathbf{x} - \frac{1}{2}\int_{\Omega} \tilde{\lambda}\left(\left|\nabla\tilde{t}\right|^2 - \frac{1}{v^2}\right)\mathrm{d}\mathbf{x}\,, \tag{3.10}$$

où $\tilde{t}$ et $\tilde{\lambda}$ sont deux variables indépendantes de v. $\mathcal{L}$ peut être interprété comme le Lagrangien associé au problème de minimisation suivant : trouver le temps de trajet t qui minimise la fonction coût $\mathcal{C}$ tout en satisfaisant l'équation eikonale (2.9). Le gradient de la fonction coût $\mathcal{C}$ par rapport au modèle de vitesse v est alors donné par

$$\frac{\partial\mathcal{C}}{\partial v}(v) = \frac{\partial\mathcal{L}}{\partial v}\big(t,\lambda,v\big)\,, \tag{3.11}$$

l'équation de l'état adjoint sous la forme

$$\frac{\partial\mathcal{L}}{\partial\tilde{t}}\big(t,\lambda,v\big) = 0\,, \tag{3.12}$$

et l'équation eikonale est donnée par

$$\frac{\partial\mathcal{L}}{\partial\tilde{\lambda}}\big(t,\lambda,v\big) = 0\,. \tag{3.13}$$

3.2.2 Le gradient de la fonction coût

Nous détaillons ici le calcul du gradient de la fonction coût $\mathcal{C}$ par rapport au modèle de vitesse v donné par (3.11). Nous calculons tout d'abord la dérivée du Lagrangien $\mathcal{L}$ par rapport à v

$$\frac{\partial \mathcal{L}}{\partial v}(\tilde{t}, \tilde{\lambda}, v) = -\int_{\Omega} \frac{\tilde{\lambda}}{v^3} \, \mathrm{d}\mathbf{x} \,. \tag{3.14}$$

Nous rappelons que $\tilde{t}$ et $\tilde{\lambda}$ sont deux variables indépendantes de v. Pour un temps de trajet donné t et sa variable de l'état adjoint correspondante λ on a

$$\frac{\partial \mathcal{C}}{\partial v}(v) = -\int_{\Omega} \frac{\lambda}{v^3} \, \mathrm{d}\mathbf{x} \,. \tag{3.15}$$

L'équation (3.15) représente la forme continue du gradient de la fonction coût par rapport au modèle de vitesse. L'expression discrète du gradient de la fonction coût dépend de la paramétrisation du modèle choisie. Nous choisissons de discrétiser le modèle de vitesse par une grille cartésienne uniforme, on a alors

$$\frac{\partial \mathcal{C}}{\partial v}(v) = -\frac{\lambda}{v^3} \tag{3.16}$$

L'expression du gradient de la fonction coût par rapport au modèle de vitesse dépend simplement de la variable de l'état adjoint λ et du modèle de vitesse v. Dans le cas où N sources sont considérées, le gradient de la fonction coût par rapport au modèle de vitesse s'écrit

$$\frac{\partial \mathcal{C}}{\partial v}(v) = -\frac{1}{v^3} \sum_{i=1}^{N} \lambda^i \,, \tag{3.17}$$

où λ^i est la variable de l'état adjoint associée à la $i^{ème}$ source. Il faut maintenant déterminer λ^i pour chaque position de source.

3.2.3 L'équation de l'état adjoint

Nous présentons les étapes permettant de définir l'équation de l'état adjoint à partir de (3.12). Nous calculons tout d'abord la dérivée du Lagrangien $\mathcal{L}$ par rapport à $\tilde{t}$

$$\frac{\partial \mathcal{L}}{\partial \tilde{t}}(\tilde{t}, \tilde{\lambda}, v) = \int_{\Omega_S} \left(\tilde{t} - t_{obs}\right) \mathrm{d}\mathbf{x} - \int_{\Omega} \tilde{\lambda} \left(\nabla \tilde{t} \cdot \nabla\right) \mathrm{d}\mathbf{x} \,. \tag{3.18}$$

Ensuite nous intégrons par partie la seconde intégrale

$$\begin{aligned}
\frac{\partial \mathcal{L}}{\partial \tilde{t}}(\tilde{t}, \tilde{\lambda}, c) = & \int_{\Omega_S} \left(\tilde{t} - t_{obs}\right) \, \mathrm{d}\mathbf{x} \\
& - \int_{\Omega_S} \tilde{\lambda} \, \left(\mathbf{n} \cdot \nabla \tilde{t}\right) \, \mathrm{d}\mathbf{x} \\
& + \int_{\Omega} \nabla \cdot \left(\tilde{\lambda} \, \nabla \tilde{t}\right) \, \mathrm{d}\mathbf{x} \\
= & \int_{\Omega_S} \left[\left(\tilde{t} - t_{obs}\right) - \tilde{\lambda} \left(\mathbf{n} \cdot \nabla \tilde{t}\right)\right] \, \mathrm{d}\mathbf{x} \\
& + \int_{\Omega} \nabla \cdot \left(\tilde{\lambda} \, \nabla \tilde{t}\right) \, \mathrm{d}\mathbf{x} \,,
\end{aligned} \tag{3.19}$$

où $\mathbf{n}$ est le vecteur unitaire normal à la surface Ω_S. Pour satisfaire la condition (3.12), nous considérons que les deux intégrales définies sur des domaines différents sont nulles. Pour un temps de trajet donné t, on peut alors choisir λ solution de l'équation aux dérivées partielles suivante sur le domaine Ω

$$\nabla \cdot \left(\lambda \, \nabla t\right) = 0 \,, \tag{3.20}$$

et à la surface Ω_S

$$\lambda \left(\mathbf{n} \cdot \nabla t\right) = t - t_{obs} \,. \tag{3.21}$$

L'équation (3.20) est appelée l'équation de l'état adjoint et l'équation (3.21) représente sa condition limite à la surface Ω_S. La résolution de ce système d'équations permet de calculer la variable de l'état adjoint λ et ainsi de déterminer le gradient de la fonction coût par rapport au modèle de vitesse. Dans le cas où plusieurs sources sont considérées, l'équation de l'état adjoint doit être résolue pour chaque position de source pour permettre de calculer la variable de l'état adjoint λ^i associée à la $i^{ème}$ source.

3.2.4 Discussions

Nous abordons dans ce paragraphe quelques points de discussion concernant la méthode de l'état adjoint et sa mise en œuvre pratique.

- L'équation de l'état adjoint (3.20) et sa condition limite à la surface (3.21) peuvent être intuitivement interprétées comme la rétropropagation des résidus des temps calculés à la surface en suivant les gradients locaux des temps de première arrivée. Cette formulation est analogue à celle de l'inversion de forme d'onde non-linéaire où le champ d'onde des résidus des données est corrélé avec le champ d'onde direct pour calculer le gradient de la fonction coût associée [Tarantola, 1986], [Noble, 1992].
- La formulation du gradient de la fonction coût linéarisé (3.6) peut être obtenu par la méthode de l'état adjoint en considérant le problème direct

  $$\mathbf{t} = \mathbf{L}\,\mathbf{s} \,, \tag{3.22}$$

  et la fonction coût associée

  $$\mathcal{C}\left(\mathbf{s}\right) = \frac{1}{2}\left(\mathbf{L}\,\mathbf{s} - \mathbf{t}_{obs}\right)^t \left(\mathbf{L}\,\mathbf{s} - \mathbf{t}_{obs}\right) \,. \tag{3.23}$$

  Dans ce cas, la variable de l'état adjoint devient

  $$\boldsymbol{\lambda} = \mathbf{t} - \mathbf{t}_{obs} \,. \tag{3.24}$$

- D'un point de vue purement théorique, le schéma numérique utilisé pour déterminer la variable de l'état adjoint λ devrait être défini à partir du schéma numérique utilisé pour calculer les temps de première arrivée t. Cette dérivation peut se révéler être une tâche formidablement complexe [Bennett, 2002]. Par ailleurs, elle ne garantit pas que le schéma numérique adjoint, dérivé du schéma numérique direct, hérite de sa consistance ou de sa précision [Sei and Symes, 1995]. D'un point de vue pratique, nous choisissons donc de discrétiser directement l'équation de l'état adjoint (3.20).

3.3 La mise en oeuvre pratique

Nous nous intéressons dans cette partie à la mise en œuvre pratique d'un algorithme de tomographie des temps de première arrivée défini à partir de la méthode de plus grande descente pour la minimisation de la fonction coût. Nous présentons tout d'abord les algorithmes retenus pour la résolution numérique des équations eikonale (2.9) et de l'état adjoint (3.20). Nous illustrons ensuite le calcul du gradient de la fonction coût sous une forme non-linéaire, par la méthode de l'état adjoint, et sous une forme linéarisée, à partir d'un tracé de rais *a posteriori*. Les calculs des gradients obtenus par les deux formulations sont ensuite validés par le calcul du gradient par différences finies.

3.3.1 Le choix des algorithmes

Equations de Hamilton-Jacobi

Les équations eikonale (2.9) et de l'état adjoint (3.20) font partie de la même classe d'équations aux dérivées partielles appelées équations de Hamilton-Jacobi. Les temps de première arrivée et la variable de l'état adjoint correspondent alors à des solutions de viscosité. Crandall et Lions [1983] ont développé une notion de solution de viscosité des équations de Hamilton-Jacobi et ont montré que les schémas numériques stables convergent vers ce type de solution. Ces solutions peuvent présenter des discontinuités de gradient qui apparaissent naturellement lorsque les trajectoires extrémales se croisent. La solution de viscosité effectue alors automatiquement le choix optimal parmi les trajectoires rendant stationnaire le critère. L'étude théorique des solutions de viscosité des équations de Hamilton-Jacobi est aujourd'hui très complète et il existe un grand nombre de schémas numériques adaptés à leur calcul, par exemple la *Fast Marching Method* [Sethian, 1996], [Popovici and Sethian, 1998] ou bien la *Fast Sweeping Method* [Boué and Dupuis, 1999], [Zhao, 2005].

La résolution de l'équation eikonale

Nous choisissons pour la résolution numérique de l'équation eikonale (2.9), le solveur au premier ordre par différences finies proposé par Podvin et Lecomte [1991]. L'algorithme calcule les temps de première arrivée sur une grille cartésienne, à partir d'un modèle de vitesse lui aussi défini sur une grille cartésienne avec le même espacement de grille. Le schéma numérique de l'algorithme repose sur le principe de Huyghens qui considère chaque point de la surface d'une maille comme une source virtuelle qui va émettre un front d'onde secondaire. Chaque point de maille est alors éclairé par plusieurs fronts d'ondes dont seul le plus rapide est retenu [Le Meur, 1994]. Les avantages de cet algorithme sont la rapidité de calcul et la possibilité de gérer les contrastes de vitesse élevés [Podvin and Lecomte, 1991].

La résolution de l'équation de l'état adjoint

Nous choisissons pour la résolution numérique de l'équation de l'état adjoint (3.20) avec sa condition limite (3.21), la *Fast Sweeping Method* proposée par Zhao [2005]. La *Fast Sweeping Method* est une méthode itérative qui balaye alternativement le domaine étudié dans plusieurs directions pour résoudre l'équation de l'état adjoint discrétisée. L'idée principale de cet al-

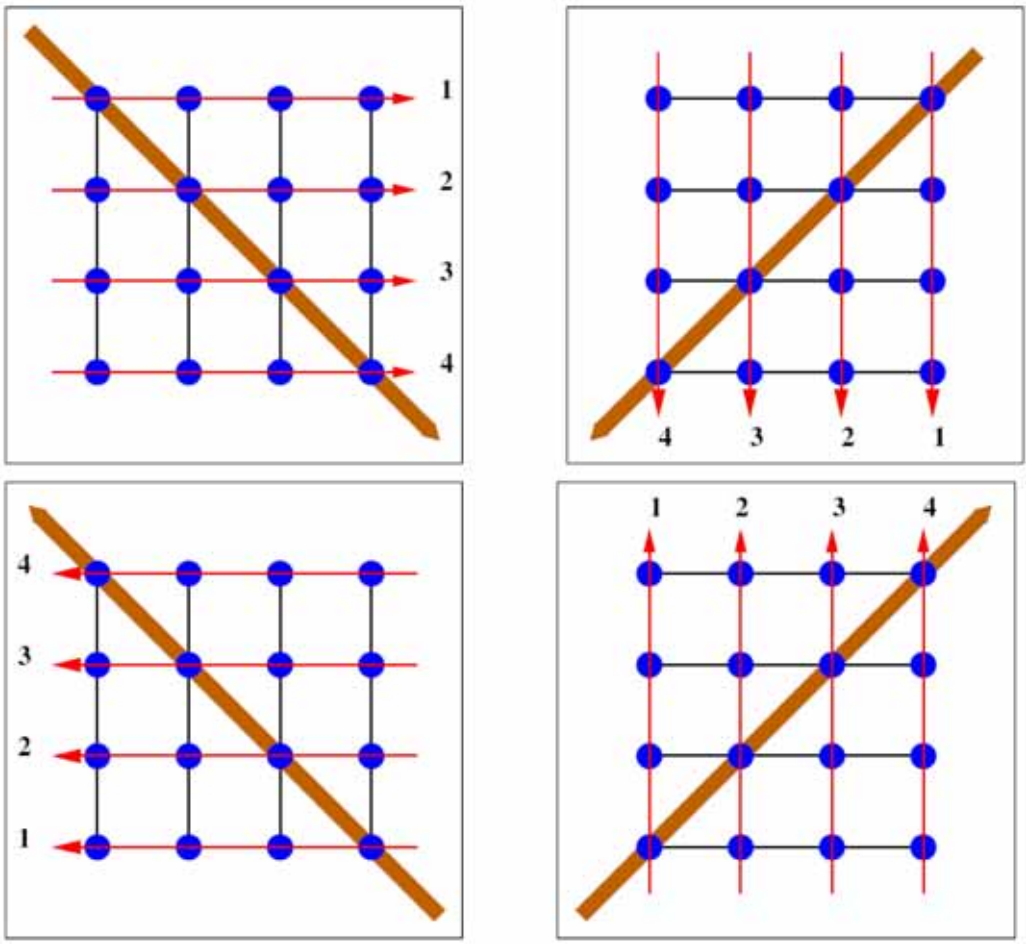

Fig. 3.3 – *Illustration des quatres balayages nécessaires en 2-D pour explorer toutes les directions possibles pour une itération de la Fast Sweeping Method, extrait de [Kuster, 2006].*

gorithme est que chaque balayage suit simultanément une famille de caractéristiques, liées à l'équation de l'état adjoint, dans une certaine direction. Intuitivement, on peut interpréter cette idée comme la rétropropagation des résidus en temps à travers la carte des temps de première arrivée en suivant les gradients locaux des temps. Pour explorer toutes les directions possibles, une itération de la *Fast Sweeping Method* comprend 4 balayages en 2-D (Fig. 3.3) et 8 balayages en 3-D. Pour un critère de convergence donné, le nombre d'itérations dépend de la complexité du modèle étudié mais est indépendant de la taille de la grille. Nous décrivons en Annexe A l'algorithme de la *Fast Sweeping Method* appliquée à la résolution de l'équation de l'état adjoint [Leung and Qian, 2006]. L'algorithme calcule la variable de l'état adjoint sur une grille cartésienne, à partir d'une carte des temps de première arrivée elle aussi définie sur une grille cartésienne avec le même espacement de grille. Les avantages de cette méthode sont une convergence rapide et la simplicité de mise en œuvre [Kuster, 2006].

3.3.2 Le calcul du gradient de la fonction coût

Nous illustrons le calcul du gradient de la fonction coût sous une forme non-linéaire, avec la méthode de l'état adjoint, et sous une forme linéarisée, à partir d'un tracé de rais *a posteriori*.

Ces calculs sont ensuite validés par le calcul du gradient de la fonction coût par différences finies.

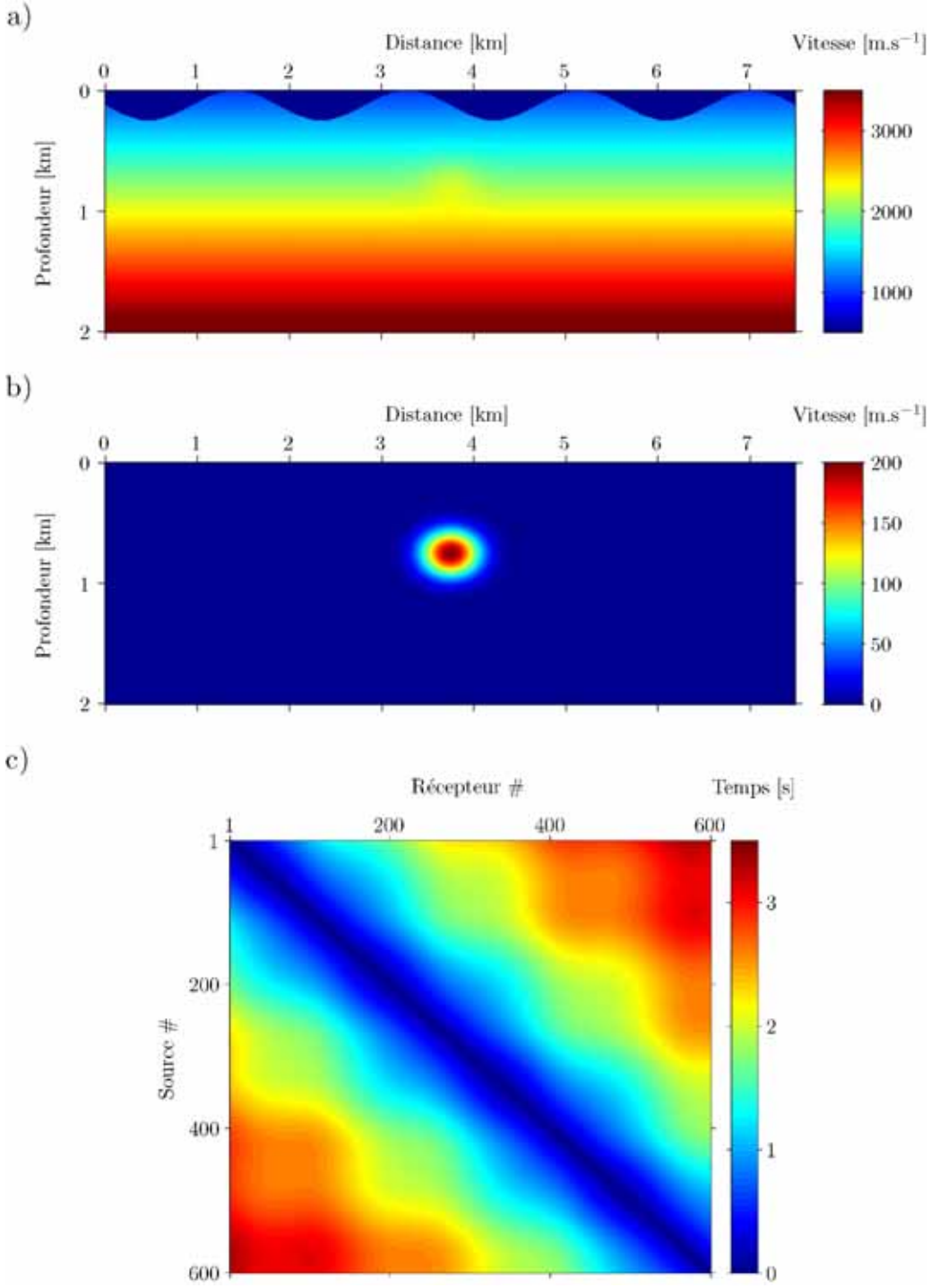

Fig. 3.4 – *a) Le modèle de vitesse de référence, b) la perturbation de vitesse recherchée, c) la carte des temps observés pour toutes les positions de sources et de récepteurs.*

Modèles de vitesse et temps de première arrivée

Nous choisissons pour illustrer le calcul du gradient, un modèle de vitesse de référence (Fig. 3.4 a) et une acquisition donnée. Le modèle de vitesse de référence est construit à partir d'une loi de vitesse linéaire, de 1000 m.s^{-1} à 3600 m.s^{-1}, en fonction de la profondeur, de 0 km à 2 km. On ajoute alors au modèle une anomalie de vitesse positive de 200 m.s^{-1} (Fig. 3.4 b). On définit sur la partie supérieure du modèle de vitesse une surface d'acquisition avec une topographie de forme sinusoïdale, la valeur de vitesse attribuée au dessus de cette surface est de 330 m.s^{-1}. Le modèle mesure 2 km en profondeur et 7.5 km en longueur. La discrétisation du modèle avec une maille carrée de 12.5 m donne une grille de taille [161 $\times$ 601]. Le dispositif d'acquisition est constitué de 600 sources spatialement réparties à la surface, chaque source possédant 600 récepteurs. Les temps de première arrivée observés (Fig. 3.4 c) sont calculés en utilisant le solveur eikonal de Podvin et Lecomte [1991]. On choisit pour modèle de vitesse courant (Fig. 3.5 a), un modèle possédant la même loi de vitesse linéaire en fonction de la profondeur que le modèle de référence. Ainsi la différence entre le modèle de référence et le modèle courant est la perturbation de vitesse (Fig. 3.4 b). Le calcul des temps de première arrivée pour le modèle courant (Fig. 3.5 b) permet de déterminer la carte des résidus entre les temps de première arrivée observés et calculés pour toutes les positions de sources et de récepteurs (Fig. 3.5 c). Ces résidus des temps et les cartes des temps de première arrivée calculées sont utilisés pour déterminer le gradient de la fonction par rapport au modèle de vitesse courant.

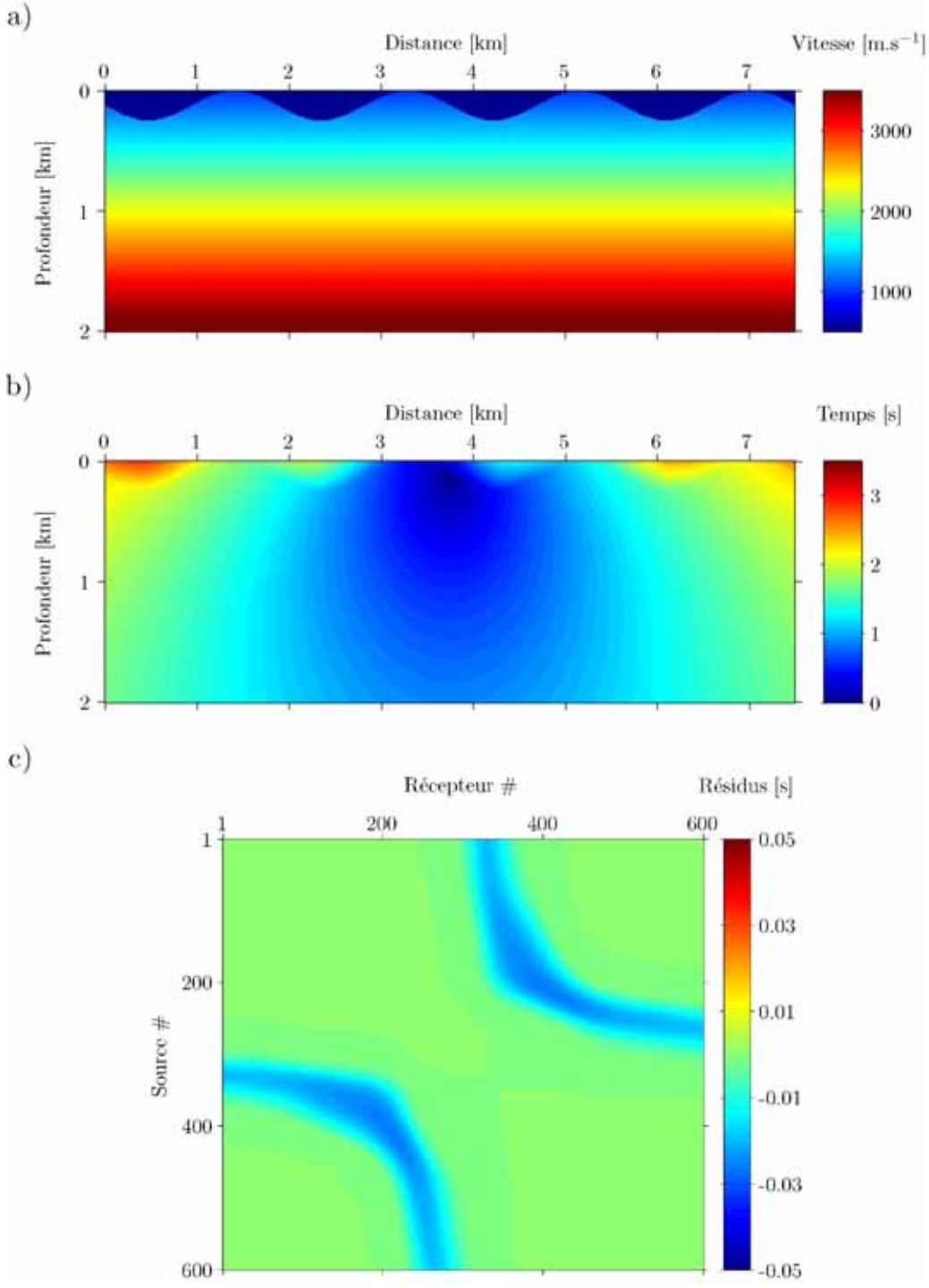

Fig. 3.5 – *a) Le modèle de vitesse courant, b) la carte des temps de première arrivée calculés par la résolution numérique de l'équation eikonale, pour une source située en (0.125 km, 3.75 km), c) la carte des résidus entre les temps de première arrivée observés et calculés pour toutes les positions de sources et de récepteurs.*

Formulation non-linéaire

Nous illustrons ici le calcul du gradient de la fonction coût par rapport au modèle de vitesse par la méthode de l'état adjoint pour une source de coordonnées (0.125 km, 3.75 km), correspondant respectivement à la profondeur et à la distance. Le calcul du gradient consiste essentiellement en la résolution de l'équation de l'état adjoint (3.20) avec sa condition limite à la surface (3.21). La méthode retenue pour résoudre cette équation est la *Fast Sweeping Method*.

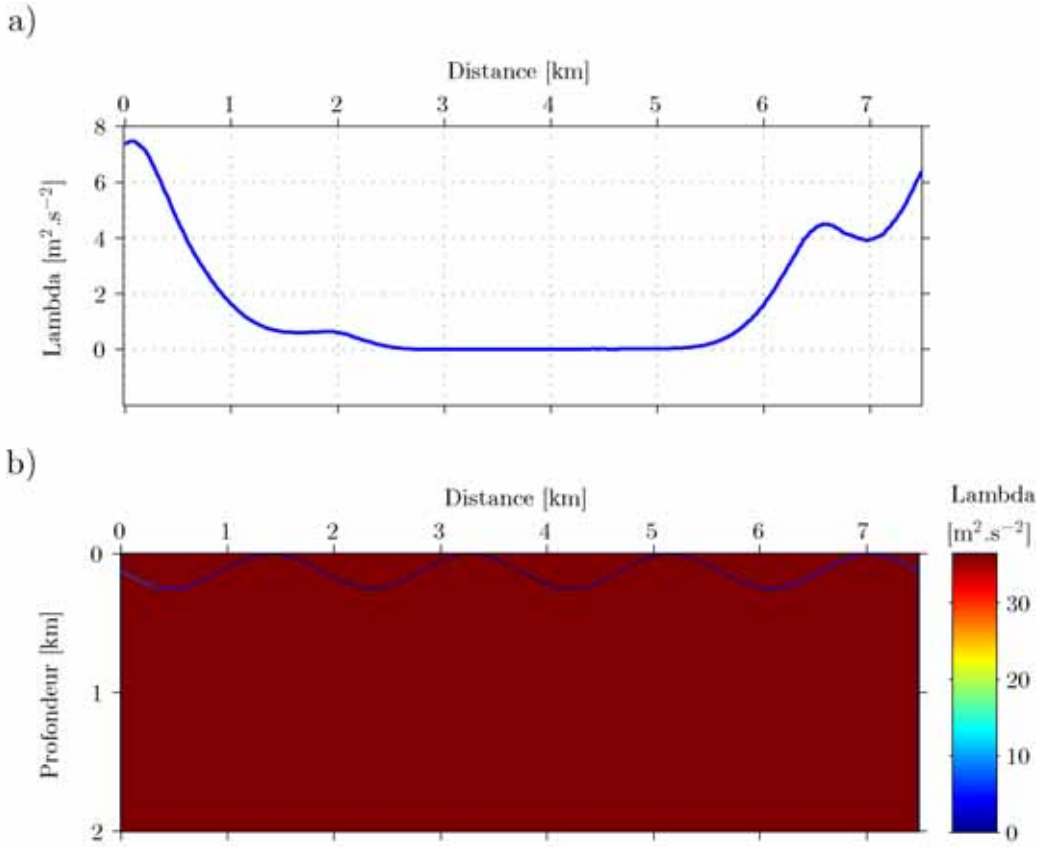

Fig. 3.6 – *a) Valeur de la variable de l'état adjoint calculée à la surface d'acquisition par la condition limite (3.21), pour une source située en (0.125 km, 3.75 km), b) initialisation d'une grille avec une valeur positive très élevée et initialisation de la variable de l'état adjoint à la surface d'acquisition.*

La condition limite de l'équation de l'état adjoint (3.21) permet de calculer la variable de l'état adjoint à la surface d'acquisition (Fig. 3.6 a) à partir des résidus des temps, des gradients des temps de première arrivée estimés localement et du vecteur unitaire normal à la surface d'acquisition. On définit une grille, de même taille que la carte des temps de première arrivée, avec une valeur très élevée. On initialise ensuite cette grille avec la variable de l'état adjoint calculée à la surface (Fig. 3.6 b). On balaye alors la grille dans différentes directions, et en chaque point, on calcule la variable de l'état adjoint candidate (A.12) que l'on compare avec la valeur courante en ce point. Dans l'exemple considéré (Fig. 3.7), le premier balayage s'effectue de gauche à droite et de haut en bas ; le second balayage de gauche à droite et de bas en

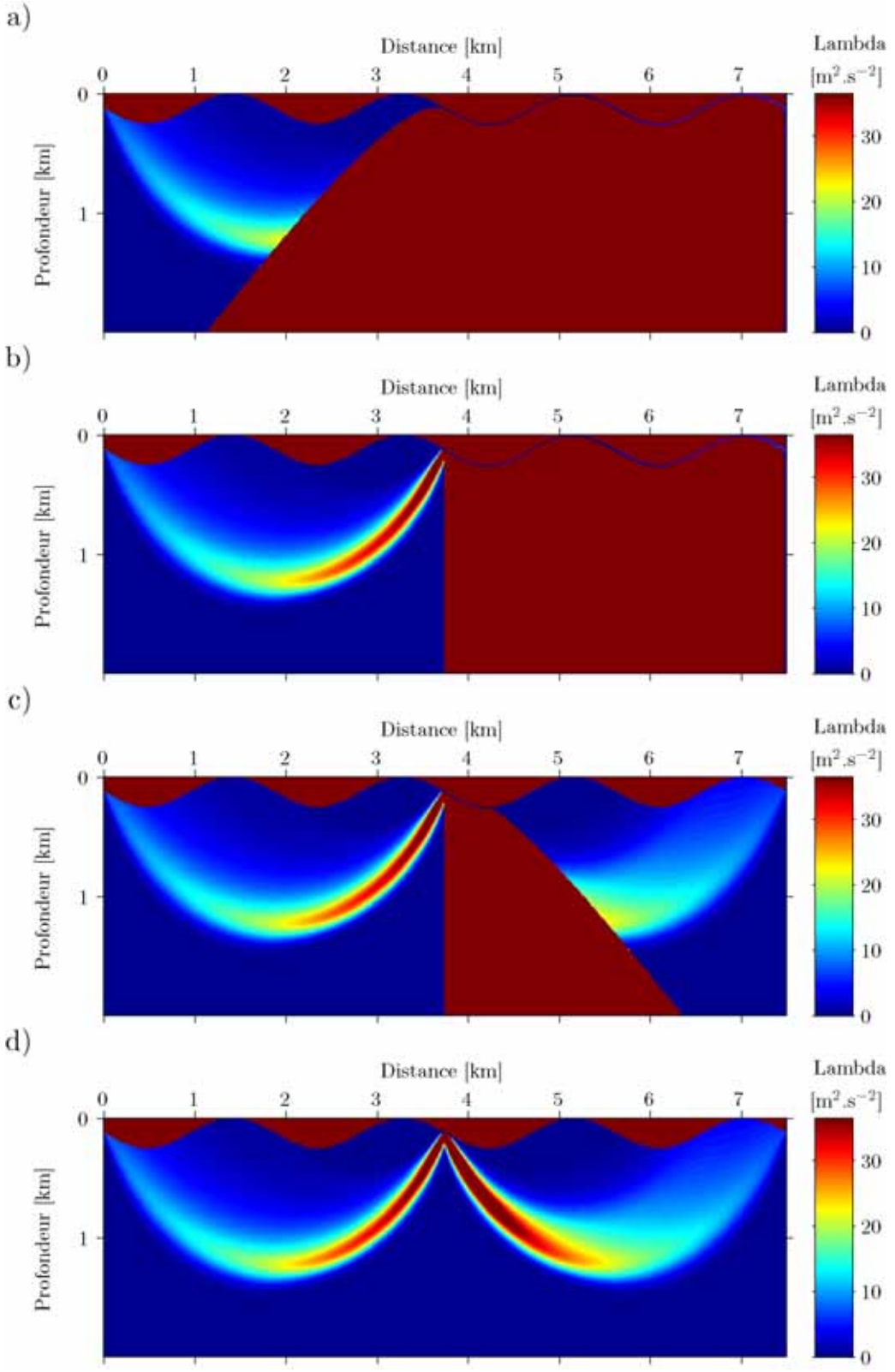

Fig. 3.7 – *Illustration des quatre balayages nécessaires pour la résolution de l'équation de l'état adjoint a) de gauche à droite et de haut en bas, b) de gauche à droite et de bas en haut, c) de droite à gauche et de haut en bas et d) de gauche à droite et de bas en haut.*

haut ; le troisième balayage de droite à gauche et de haut en bas et enfin le quatrième balayage de gauche à droite et de bas en haut. Une itération de la *Fast Sweeping Method*, constituée de quatre balayages en 2-D, est ici suffisante pour converger vers la solution de l'équation de l'état adjoint. L'ordre des balayages n'a pas d'importance, les itérations successives de la *Fast Sweeping Method* permettent de couvrir tous les cas possibles. Le calcul du gradient de la fonction coût par rapport au modèle de vitesse pour cette position de source (Fig. 3.8 a) s'effectue alors simplement en appliquant la formule donnée par la méthode de l'état adjoint (3.16). La sommation des gradients pour toutes les positions de sources fournit le gradient global de la fonction coût par rapport au modèle de vitesse (Fig. 3.8 b).

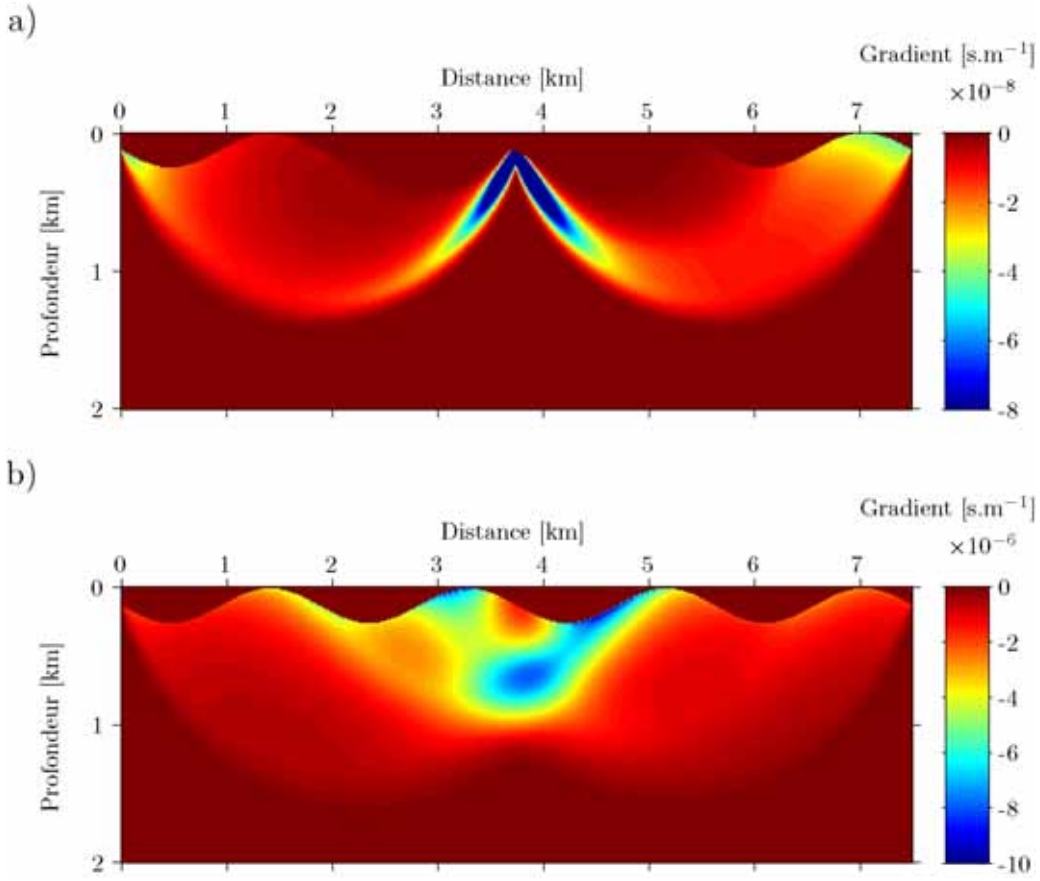

Fig. 3.8 – *a) Gradient de la fonction coût par rapport au modèle de vitesse calculé par la méthode de l'état adjoint pour une source de coordonnées (0.125 km, 3.75 km), b) gradient global de la fonction coût calculé par la méthode de l'état adjoint pour toutes les positions de sources.*

Formulation linéarisée

Le calcul du gradient de la fonction coût par rapport au modèle de vitesse peut aussi se faire à partir de la formule (3.6). L'application de cette formule nécessite de calculer les rais pour chaque position de source et de récepteur mais pas de construire directement la matrice des dérivées de Fréchet dans son intégralité. En effet, pour une position de source et de récepteur données on réalise tout d'abord un tracé de rais *a posteriori* (Fig. 3.9 a). On détermine ensuite les longueurs du rai dans chaque maille du modèle (Fig. 3.9 b), puis on multiplie par la valeur du résidu en temps associé. Le même calcul réalisé pour toutes les positions de récepteurs donne le gradient de la fonction coût pour une position de source (Fig. 3.10 a). La sommation des gradients pour toutes les positions de sources fournit le gradient global de la fonction coût par rapport au modèle de vitesse (Fig. 3.10 b).

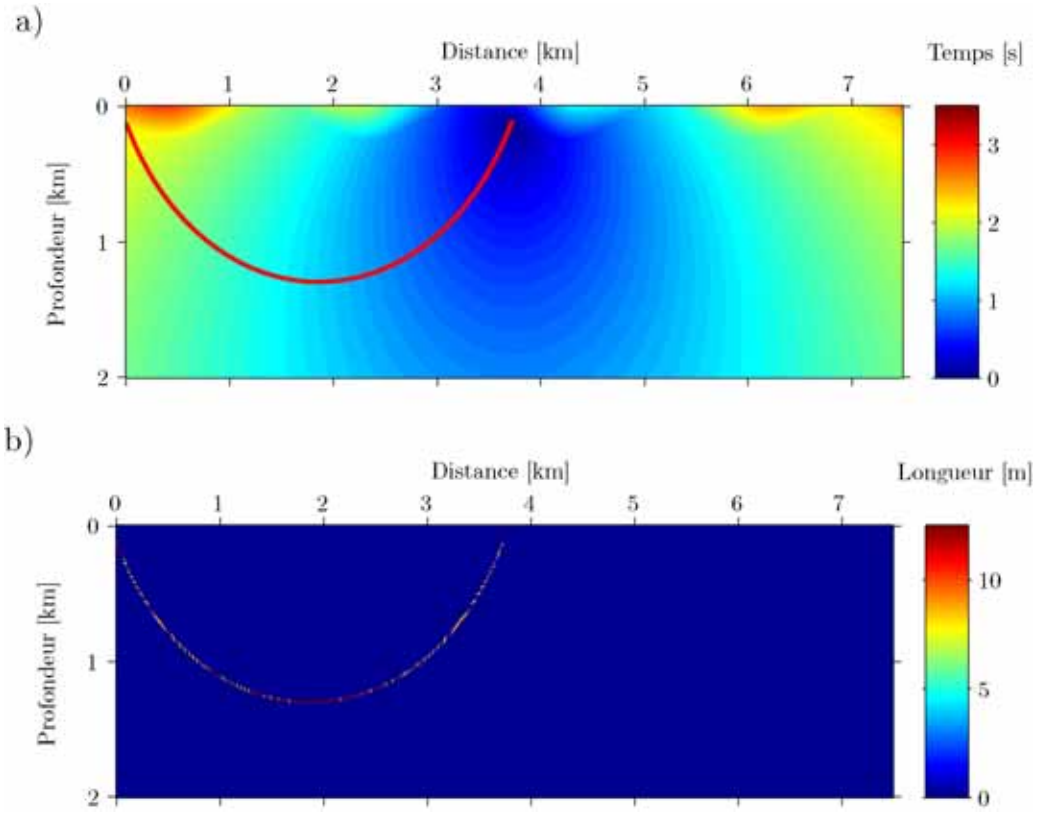

Fig. 3.9 – *a) Tracé de rais a posteriori réalisé à partir de la carte des temps de première arrivée pour une source de coordonnées (0.125 km, 3.75 km) et un récepteur de coordonnées (0.125 km, 0 km) b) carte des longueurs des segments du rai dans chaque maille de la grille.*

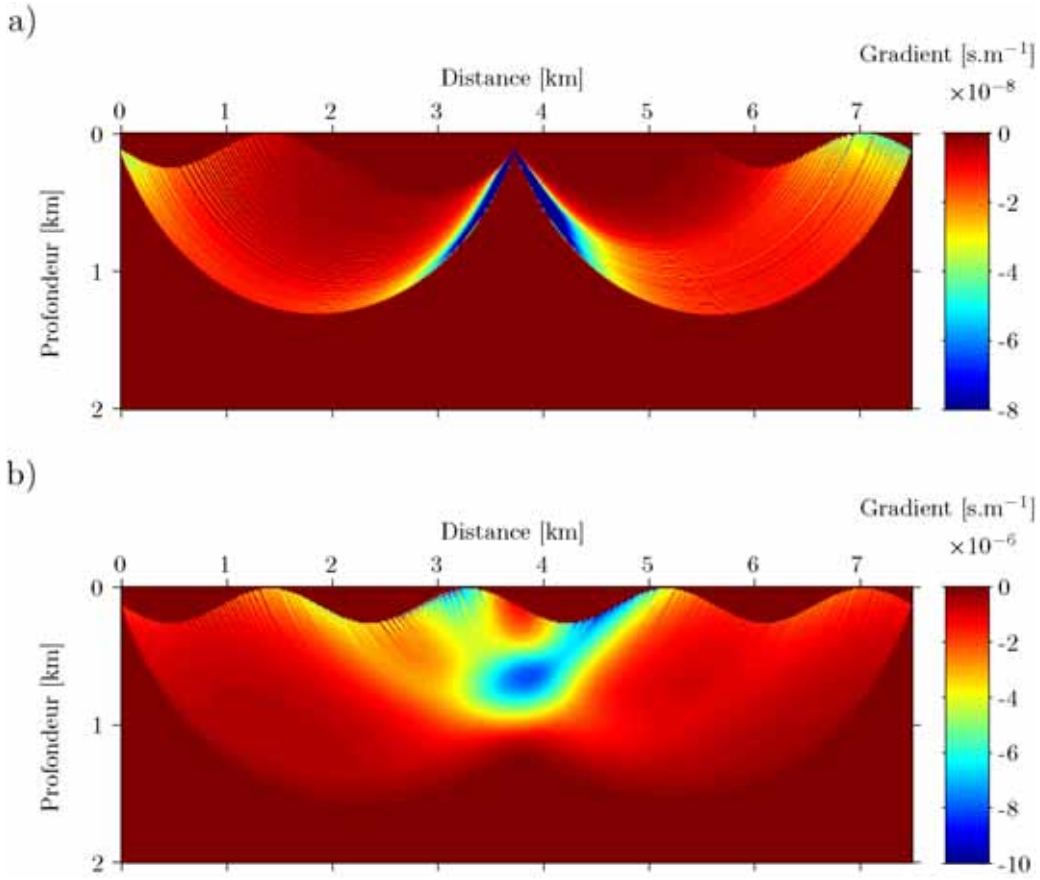

Fig. 3.10 – *a) Gradient de la fonction coût par rapport au modèle de vitesse calculé à partir d'un tracé de rais a posteriori pour une source de coordonnées (0.125 km, 3.75 km), b) gradient global de la fonction coût calculé à partir d'un tracé de rais a posteriori pour toutes les positions de sources.*

Validation des calculs

Pour valider les calculs du gradient de la fonction coût par rapport au modèle de vitesse sous une forme non-linéaire, par la méthode de l'état adjoint, et sous une forme linéarisée, par le calcul des rais *a posteriori*, nous calculons le gradient de la fonction coût par différences finies. Pour une position de source donnée, on ajoute une petite perturbation de vitesse à une maille du modèle de vitesse considéré. On calcule alors la fonction coût correspondante à ce modèle de vitesse perturbé. En considérant une perturbation de vitesse positive puis négative, de 10 m.s^{-1}, on calcule la valeur du gradient de la fonction coût au point considéré. En répétant les mêmes calculs pour toutes les mailles du modèle de vitesse, on détermine alors le gradient de la fonction coût par rapport au modèle de vitesse (Fig. 3.11 a). La sommation des gradients pour toutes les positions de sources fournit le gradient global de la fonction coût par rapport au modèle de vitesse (Fig. 3.11 b). Pour comparer les différents gradients de la fonction coût obtenus, on représente tout d'abord la valeur des gradients pour une source et une profondeur, 600 m, donnés (Fig. 3.12 a). On remarque que tous les gradients possèdent une structure et des valeurs similaires. Cependant, les gradients calculés à partir d'un tracé de rais *a posteriori* et par différences finies possèdent un aspect haute fréquence et des sauts de valeurs très importants ;

contrairement au gradient calculé par la méthode de l'état adjoint qui est lisse. En comparant les gradients calculés pour toutes les positions de sources et à la même profondeur (Fig. 3.12 b), tous les gradients calculés ont la même structure et les effets haute fréquence des gradients calculés par différences finies et par tracé de rais *a posteriori* ont quasiment disparu. De cette manière, on valide les calculs du gradient de la fonction coût par rapport au modèle de vitesse par la méthode de l'état adjoint et à partir d'un tracé de rais *a posteriori*. Le choix d'utiliser la formulation non-linéaire ou linéarisée du calcul du gradient de la fonction coût par rapport aux paramètres du modèle dépend de l'application considérée. Dans la suite de ce travail, nous utilisons indifféremment l'une ou l'autre des méthodes pour calculer le gradient de la fonction coût par rapport aux paramètres du modèle.

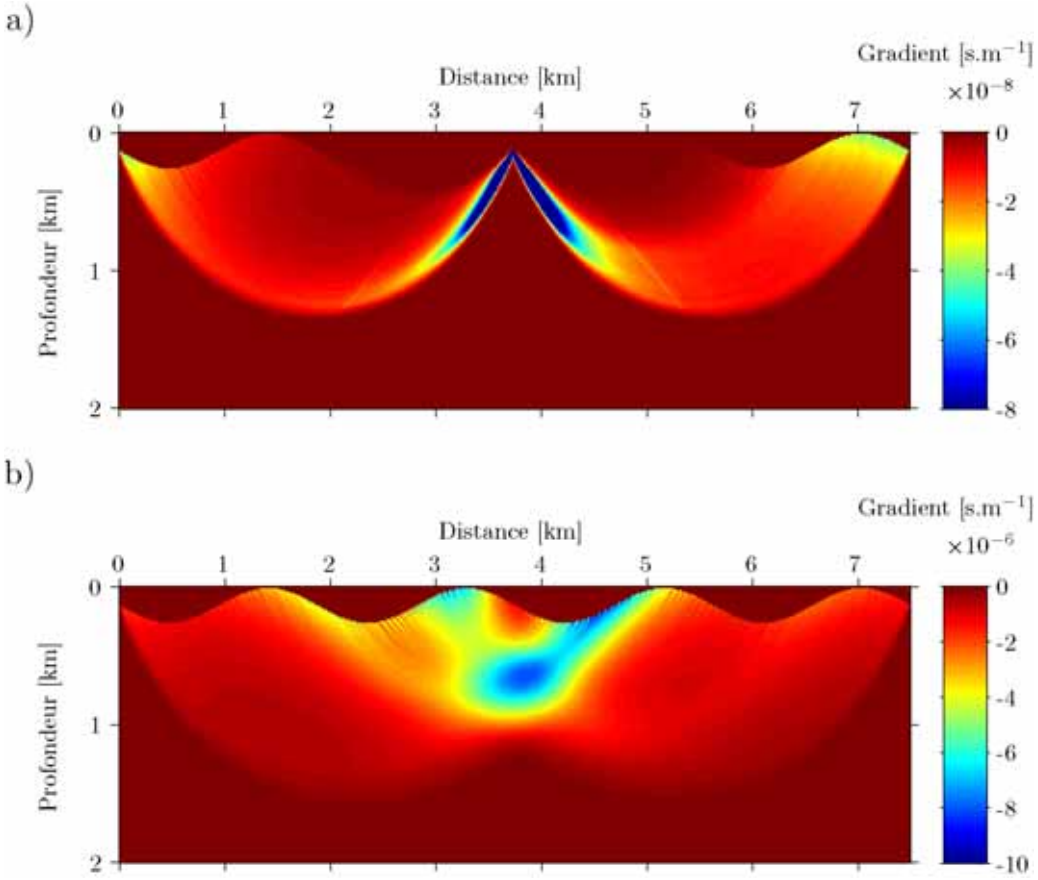

Fig. 3.11 – *a) Gradient de la fonction coût par rapport au modèle de vitesse calculé par différences finies pour une source de coordonnées (0.125 km, 3.75 km), b) gradient global de la fonction coût calculé par différences finies pour toutes les positions de sources.*

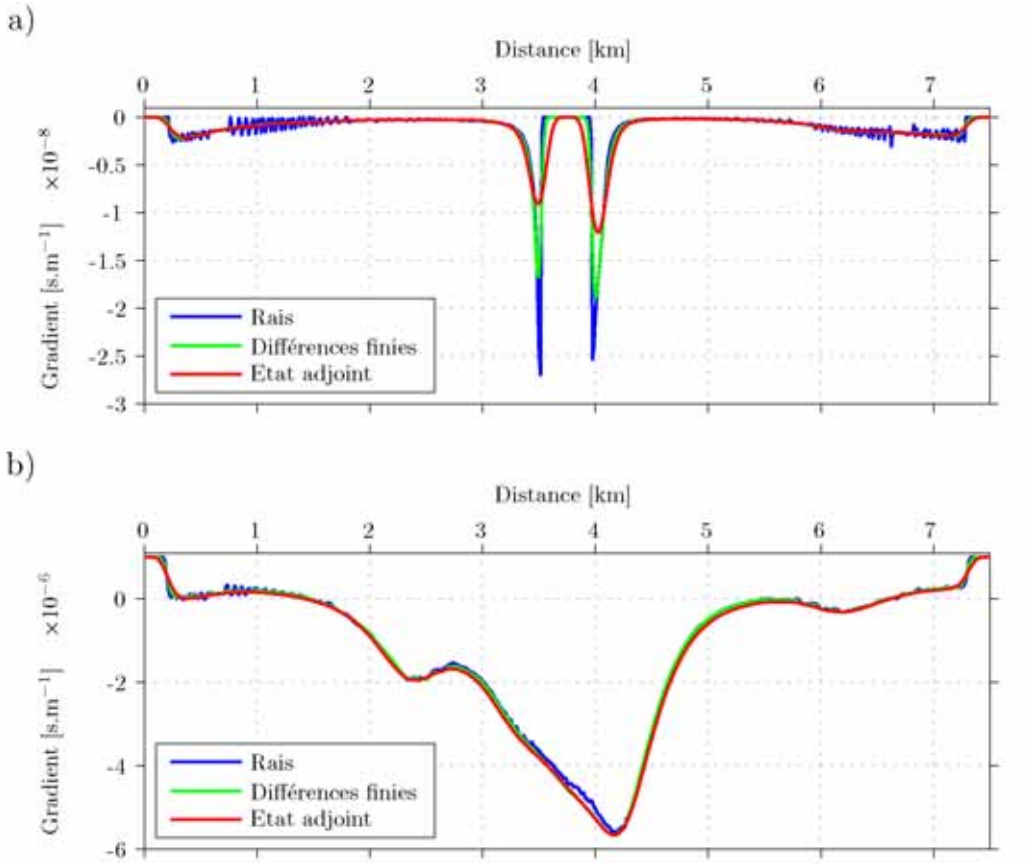

Fig. 3.12 – *Comparaison des différents gradients de la fonction coût par rapport au modèle de vitesse calculés par la méthode de l'état adjoint (en rouge), à partir du tracé de rais posteriori (en bleu) et par différences finies (en vert), à 600 m de profondeur, a) pour une source de coordonnées (0.125 km, 3.75 km), b) pour toutes les positions de sources.*

3.3.3 Le calcul du pas

Le schéma itératif de la méthode de plus grande descente (3.2) nécessite de déterminer un réel positif μ_n tel que $\mathbf{m}_{n+1} = \mathbf{m}_n - \mu_n \boldsymbol{\gamma}_n$. La valeur de μ_n doit avant toute chose assurer que la fonction coût diminue, $\mathcal{C}(\mathbf{m}_{n+1}) < \mathcal{C}(\mathbf{m}_n)$. Pour déterminer cette valeur de pas, nous choisissons une méthode d'interpolation [Tarantola, 1987a]. $\mathcal{C}(\mathbf{m}_{n+1})$ est calculée pour trois valeurs différentes de μ_n, une parabole est alors interpolée à partir des valeurs prises par la fonction coût. Dans le cadre de ce travail, pour assurer une rapide convergence et la stabilité de l'algorithme, nous choisissons généralement pour les trois valeurs distinctes de pas : 0 $m.s^{-1}$, 100 $m.s^{-1}$ et 200 $m.s^{-1}$. La valeur de μ_n, minimisant la parabole ainsi définie, est utilisée.

3.4 Un nouvel algorithme de tomographie

3.4.1 Description

Le nouvel algorithme de tomographie des temps de première arrivée défini à partir de la méthode de plus grande descente se compose, à chaque itération, de deux étapes principales : le calcul du gradient de la fonction coût par rapport au modèle de vitesse $\boldsymbol{\gamma}_n$ et le calcul du pas μ_n appliqué à ce gradient.

Le calcul du gradient de la fonction coût

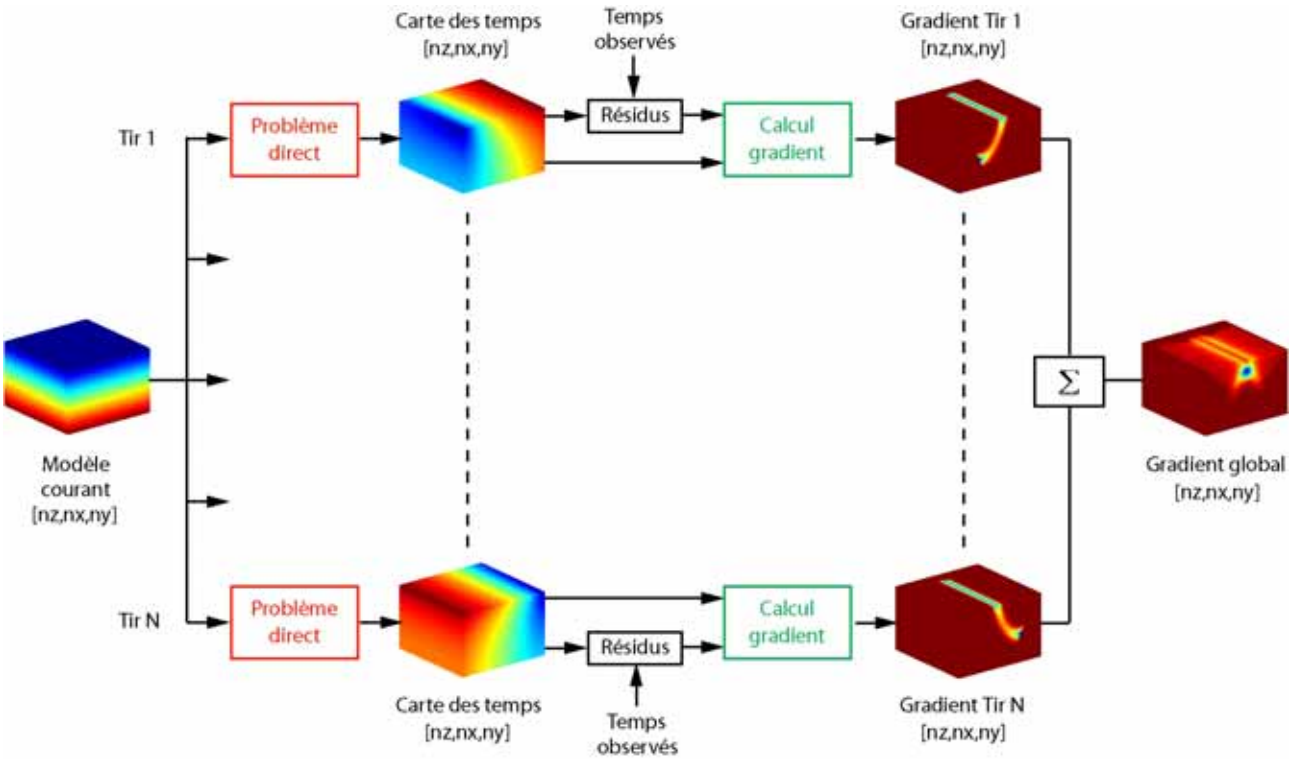

Fig. 3.13 – *Diagramme illustrant les principales étapes algorithmiques permettant le calcul du gradient de la fonction coût par rapport au modèle de vitesse $\boldsymbol{\gamma}_n$ pour une itération de la méthode de plus grande descente.*

Le diagramme (Fig. 3.13) reprend les principales étapes algorithmiques permettant la détermination du gradient de la fonction coût par rapport au modèle de vitesse $\boldsymbol{\gamma}_n$ pour une itération de la méthode de plus grande descente. A partir du modèle de vitesse courant, la résolution du problème direct pour N points de tirs fournit les cartes des temps de première arrivée en tous points de la grille. Ces cartes de temps de trajet et les résidus des temps calculés à la surface d'acquisition permettent alors de calculer le gradient de la fonction coût par rapport au modèle de vitesse. Le calcul du gradient peut se faire sous une forme non-linéaire, par la méthode de l'état adjoint, ou sous une forme linéarisée, à partir d'un tracé de rais *a posteriori*. La sommation des gradients calculés pour tous les points de tir permet de déterminer le gradient global de la fonction coût par rapport au modèle de vitesse courant.

Le calcul du pas

Le diagramme (Fig. 3.14) illustre les principales étapes algorithmiques permettant le calcul de la fonction coût pour une valeur de pas μ_n donnée. A partir du modèle de vitesse défini par le schéma itératif de la méthode de plus grande descente $\mathbf{m}_{n+1} = \mathbf{m}_n - \mu_n \boldsymbol{\gamma}_n$, la résolution du problème direct pour N points de tirs fournit les cartes des temps de première arrivée en tous points de la grille. La sommation des carrés des résidus, définis entre les temps observés à la surface et ceux calculés pour le modèle de vitesse considéré, permet de déterminer la valeur de la fonction coût $\mathcal{C}(\mathbf{m}_{n+1})$.

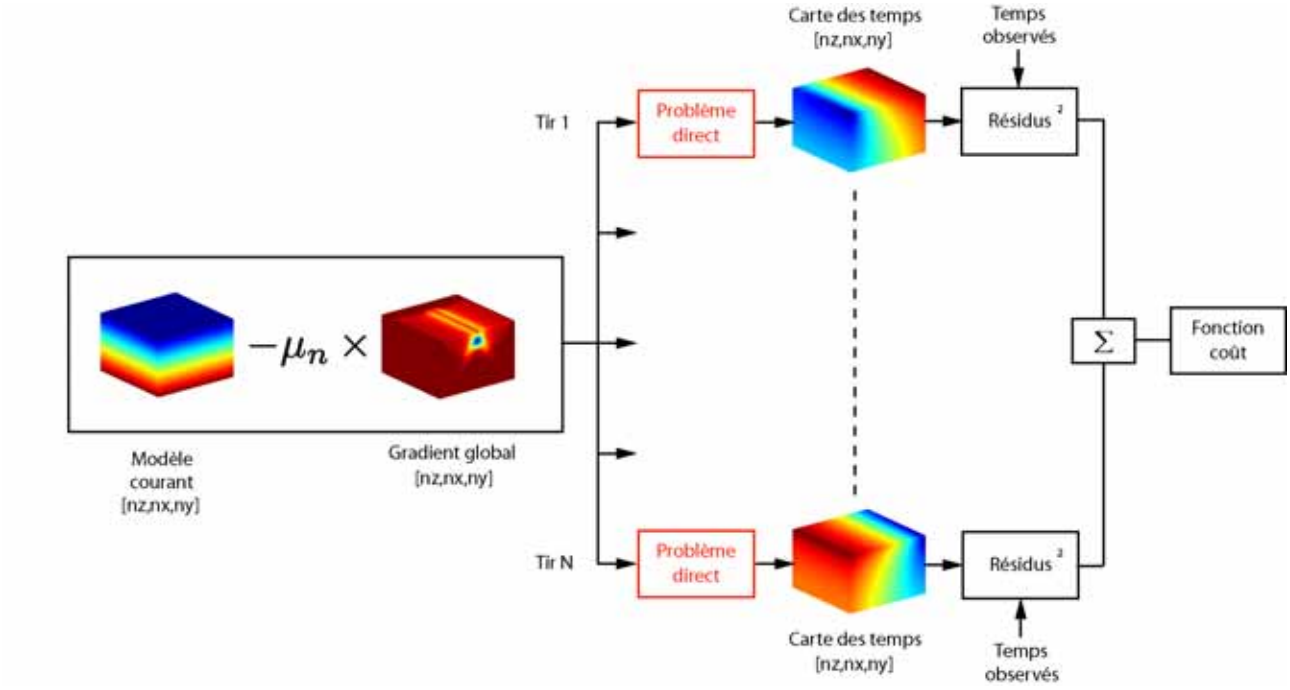

Fig. 3.14 – *Diagramme illustrant les principales étapes algorithmiques permettant le calcul de la fonction coût* $\mathcal{C}(\mathbf{m}_{n+1})$ *pour une valeur de pas* μ_n *donnée.*

Ce calcul répété pour trois valeurs différentes de μ_n permet ensuite d'interpoler une parabole à partir des trois valeurs de fonction coût calculées. La valeur de μ_n, minimisant la parabole ainsi définie, est utilisée.

Le schéma de l'algorithme de tomographie

Le schéma (Fig. 3.15) représente les principales étapes de l'algorithme de tomographie de temps de trajet de première arrivée défini à partir de la méthode de plus grande descente. Pour un modèle de vitesse $\mathbf{m}_n$, on calcule le gradient de la fonction coût $\boldsymbol{\gamma}_n$ puis la valeur du pas μ_n appliqué au gradient. Le schéma itératif de la méthode de plus grande descente donne la définition du modèle mis à jour $\mathbf{m}_{n+1} = \mathbf{m}_n - \mu_n \boldsymbol{\gamma}_n$. Un critère d'arrêt permet de stopper ou non le schéma itératif. Les paramètres de réglage de l'algorithme sont le modèle de vitesse initial, la valeur de μ_n et le critère d'arrêt.

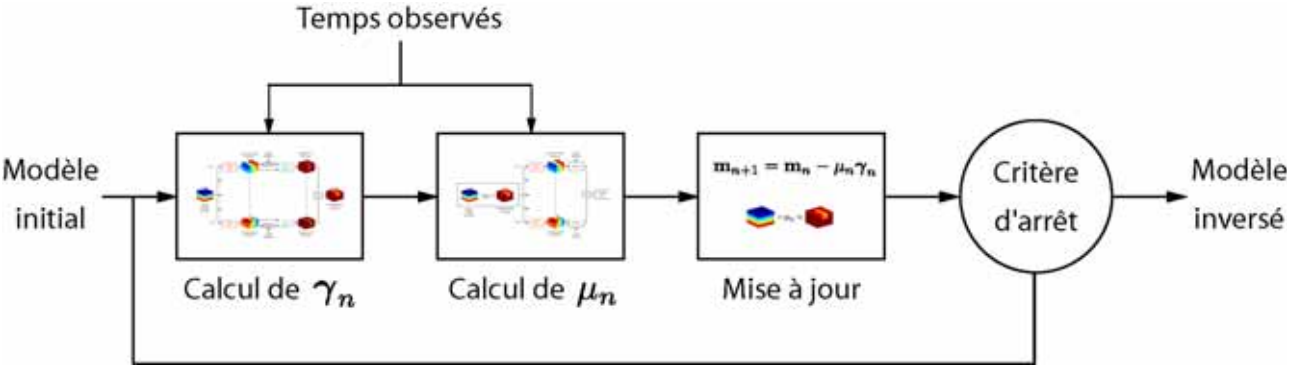

Fig. 3.15 – *Diagramme illustrant les principales étapes de l'algorithme de tomographie de temps de trajet de première arrivée défini à partir de la minimisation de la fonction coût par la méthode de plus grande descente.*

3.4.2 L'implémentation

L'algorithme de tomographie des temps de première arrivée défini à partir de la méthode de plus grande descente est implémenté en 2-D et en 3-D en langages C et Fortran 90. La parallélisation de l'algorithme est réalisée en utilisant l'environnement *Message Passing Interface* (MPI). Le nouvel algorithme de tomographie a été implémenté, testé et validé sur différents calculateurs, différents environnements, différentes architectures et en utilisant différents compilateurs.

3.4.3 Les performances informatiques

Les diagrammes (Fig. 3.13) et (Fig. 3.14) illustrent les deux propriétés clés d'un algorithme de tomographie des temps de première arrivée défini à partir d'une méthode de gradient. Pour chaque point de tir, l'occupation mémoire de l'algorithme dépend uniquement de l'occupation mémoire du modèle de vitesse considéré. En d'autres termes, l'occupation mémoire de l'algorithme est indépendante du nombre de données observées, du nombre de temps de première arrivée pointés. La seconde propriété est que les calculs peuvent être réalisés indépendamment par point de tir. On peut alors paralléliser l'algorithme de manière efficace et réduire ainsi le temps de calcul de façon significative. Nous présentons ici les performances informatiques du nouvel algorithme de tomographie des temps de première arrivée en 2-D et en 3-D. Les simulations sont réalisées sur un supercalculateur possédant 128 nœuds avec 8 Go de mémoire vive et 4 cœurs de fréquence 2.2 GHz. Les nœuds sont reliés entre eux par une connexion de 10 Go.s^{-1}.

Les données utilisées

Pour pouvoir étudier les performances informatiques du nouvel algorithme de tomographie nous utilisons les modèles de vitesse BP EAGE [Billette and Brandsberg-Dahl, 2005] en 2-D et Overthrust [Aminzadeh et al., 1997] en 3-D. Nous définissons des acquisitions spécifiques pour l'étude des performances informatiques. En 2-D, le nombre de mailles est d'environ 5.2 millions, le nombre de points de tir est de 8192 et le nombre de données observées est d'environ 19.7 millions. En 3-D, le nombre de mailles est d'environ 15.1 millions, le nombre de points de tir est de 512 et le nombre de données observées est d'environ 82.3 millions.

La parallélisation

Des variables spécifiques mesurent les performances d'un algorithme parallélisé : le gain en temps et l'efficacité. Le gain en temps mesure combien un calcul réalisé en parallèle est plus rapide que le même calcul réalisé en séquentiel. Il est donc défini comme le ratio entre le temps de calcul pour une opération donnée en utilisant n cœurs et le temps de calcul nécessaire pour la même opération en utilisant 1 cœur. La valeur idéale de gain en temps est de n en utilisant n cœurs . L'efficacité mesure la bonne utilisation des cœurs. Elle est définie comme le ratio entre le gain en temps pour n cœurs et le nombre de cœurs n considérés. La valeur idéale d'efficacité est de 1. La figure (Fig. 3.16) présente le gain en temps et la valeur de l'efficacité calculés pour le nouvel algorithme de tomographie en 2-D et en 3-D.

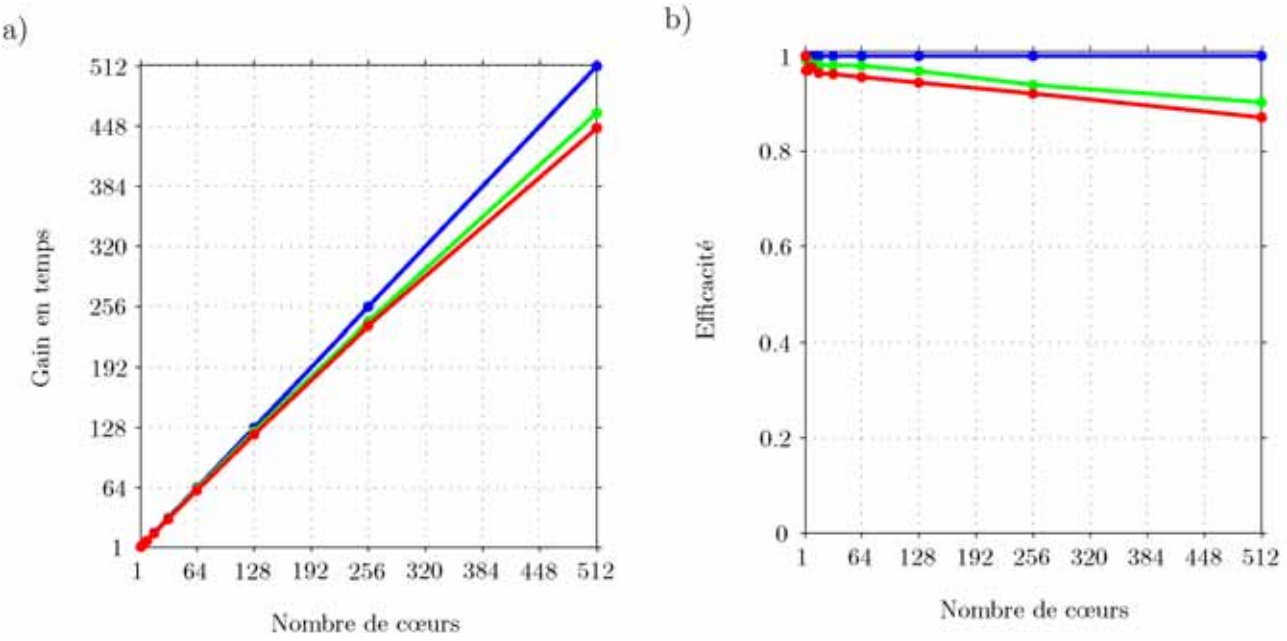

Fig. 3.16 – *Gain en temps a) et efficacité b) pour un nombre croissant de cœurs utilisés. La courbe bleue représente la performance idéale, la courbe rouge la performance du nouvel algorithme de tomographie en 2-D et la courbe verte la performance du nouvel algorithme de tomographie en 3-D.*

On peut noter que même en utilisant jusqu'à 512 cœurs les gains en temps restent très importants, environ 450, et les valeurs d'efficacité sont environ égales à 0.9. Les courbes de l'algorithme, en rouge (2-D) et en vert (3-D), tendent à s'éloigner des courbes idéales, en bleu, car les temps de calcul par cœur deviennent de moins en moins importants comparés aux temps de communication et aux temps de calcul des parties non parallélisables de l'algorithme.
Pour donner un ordre de grandeur du gain en temps obtenu grâce à la parallélisation de l'algorithme de tomographie des temps de première arrivée, une itération de méthode de plus grande descente pour les cas 2-D et 3-D définis ci-dessus dure environ 8 heures en utilisant 1 cœur et environ 1 minute en utilisant 512 cœurs. Ici, les temps de calcul pour les simulations en 2-D et 3-D sont à peu près équivalents car la simulation en 2-D comporte un grand nombre de points de tir contrairement à la simulation en 3-D où le nombre de sources est peu élevé.

L'occupation mémoire

L'occupation mémoire du nouvel algorithme de tomographie des temps de première arrivée est indépendante du nombre de données observées. Elle dépend uniquement de la taille du modèle considéré. Pour donner un ordre de grandeur de l'occupation mémoire de l'algorithme, qui dépend fortement de l'implémentation, nous considérons qu'elle est environ égale à huit fois celle nécessaire pour le modèle de vitesse. Pour les cas considérés ici, le modèle de vitesse occupe environ 21 Mo en 2-D et 60 Mo en 3-D. L'algorithme de tomographie en 2-D occupe donc 168 Mo de mémoire et en 3-D 480 Mo de mémoire. Pour les cas considérés, les calculs peuvent être distribués sur plusieurs cœurs sans que l'occupation mémoire soit un facteur limitant. Pour des modèles de vitesse de plus grande taille, il pourrait être nécessaire d'augmenter la mémoire attribuée à chaque cœur, ou à chaque nœud, ou bien d'utiliser des algorithmes par décomposition de domaines.

Chapitre 4

Applications

Nous présentons dans ce chapitre quelques exemples d'application de l'algorithme de tomographie des temps de première arrivée défini au chapitre précédent. La formulation de ce nouvel algorithme repose sur un changement de méthode pour la minimisation de la fonction coût associée au problème de tomographie. Ce changement de méthode ne modifie cependant pas la nature du problème à résoudre, à savoir un problème non-linéaire et mal posé [1]. En effet, la fonction coût à minimiser n'est généralement pas quadratique, l'algorithme de tomographie peut alors converger vers un minimum secondaire, et l'existence et l'unicité de la solution ne sont pas assurées. Toutes ces contraintes font de la paramétrisation du schéma d'inversion l'étape clé dans l'application d'un algorithme de tomographie des temps de première arrivée. Cette paramétrisation dépend à la fois de la nature des données, du type d'acquisition, du contexte de l'étude, de l'algorithme utilisé mais surtout de l'expérience et du savoir-faire du géophysicien. Pour pouvoir obtenir les résultats de tomographie présentés dans ce chapitre, nous avons dû définir et acquérir notre propre savoir-faire tomographique. L'acquisition de ce savoir-faire a nécessité de nombreuses simulations pour des jeux de données et de paramétrisation différents. La réalisation d'un grand nombre de simulations a été rendue possible par la rapidité d'exécution de l'algorithme de tomographie, typiquement de l'ordre de quelques minutes pour les simulations en 2-D présentées dans ce chapitre. Dans le cadre de ce travail, nous choisissons de présenter tout d'abord des résultats de tomographie en réfraction obtenus sur des jeux de données synthétiques en 2-D et 3-D qui nous permettent de valider le nouveau schéma d'inversion. Nous présentons enfin un exemple d'application sur un jeu de données réelles qui nous permet d'illustrer les propriétés du nouvel algorithme de tomographie des temps de première arrivée.

4.1 Un savoir-faire tomographique

Un algorithme de tomographie des temps de première arrivée est un outil qui permet de déterminer un modèle de vitesse le plus vraisemblablement à l'origine des temps de première arrivée mesurés à la surface. Le modèle obtenu dépend donc des temps de première arrivée observés mais surtout de la paramétrisation du schéma d'inversion définie par l'utilisateur, laquelle doit assurer la convergence du schéma itératif vers le minimum de la fonction coût. Les paramètres du schéma d'inversion sont le modèle de vitesse initial, le préconditionnement et la valeur du

1. "A change in mathematical method does not change the nature of physical difficulties to be solved, but may only provide more efficient tools", d'après [Podvin and Lecomte, 1991].

pas appliqués au gradient de la fonction coût à chaque itération et le critère d'arrêt. Le savoir-faire tomographique consiste à déterminer la meilleure paramétrisation du schéma d'inversion pour obtenir le meilleur modèle de vitesse possible.

4.1.1 Le modèle de vitesse initial

L'utilisation d'un schéma itératif local pour la minimisation de la fonction coût associée au problème de tomographie des temps de première arrivée implique de définir un modèle de vitesse initial. Le modèle de vitesse obtenu après inversion dépend directement du choix de ce modèle de vitesse initial. Le problème de la tomographie des temps de première arrivée est un problème non-linéaire, la fonction coût à minimiser peut alors posséder des minimums secondaires. C'est la raison pour laquelle, on cherche généralement à définir un modèle de vitesse initial le plus proche possible du modèle recherché. Dans le cadre de travail, nous construisons le modèle de vitesse initial avec pour seule contrainte que celui-ci soit plus rapide que le modèle de vitesse recherché, c'est-à-dire que les résidus initiaux des temps de première arrivée $(\mathbf{t}_{obs} - \mathbf{t})$ sont positifs. Nous justifions ce choix, en partie, par le fait qu'un modèle de vitesse initial plus lent serait à l'origine de perturbations de vitesse positives localisées proches de la surface qui ne permettraient pas ensuite d'expliquer le modèle en profondeur. Cette "recette" est valable pour des applications de tomographie en réfraction, mais elle ne s'applique pas nécessairement à tous les types d'acquisition. Par exemple, Baina [1998] montre qu'un modèle de vitesse initial plus lent que le modèle recherché est préférable pour des applications de tomographie entre puits.

4.1.2 Le préconditionnement

A chaque itération de la méthode de plus grande descente, nous choisissons d'appliquer un opérateur de préconditionnement [Crase et al., 1990], [Fomel and Claerbout, 2003] au gradient de la fonction coût. Ce préconditionnement a pour objectif d'ajuster les disparités causées par la géométrie d'acquisition et ainsi d'améliorer la rapidité de convergence du schéma d'inversion et la qualité de la solution obtenue. Dans le cadre de ce travail, le préconditionnement appliqué au gradient de la fonction coût consiste essentiellement en un filtrage gaussien. Nous choisissons de faire varier la longueur du filtrage appliqué au gradient de la fonction coût en deux étapes. Pour les premières itérations du schéma d'inversion, le gradient de la fonction coût est fortement lissé, de façon à répartir globalement la perturbation de vitesse appliquée au modèle. Ensuite, dans un second temps le filtrage est relâché, de façon à pouvoir déterminer des perturbations locales du modèle de vitesse.

4.1.3 La valeur maximale du pas

La valeur du pas appliqué à chaque itération au gradient de la fonction coût a pour objectif principal d'assurer la minimisation de la fonction coût le long de la direction donnée par le gradient. Théoriquement, le schéma itératif de la méthode de plus grande descente est optimal pour des valeurs de pas infiniment petites. En pratique, la valeur du pas est définie comme le minimum d'une parabole interpolée à partir de plusieurs valeurs de la fonction coût calculées pour différentes valeurs de ce pas [Tarantola, 1987a]. Dans le cadre de ce travail, nous choisissons de contraindre le schéma d'inversion en définissant une valeur maximale de pas à appliquer

au gradient. L'utilisation d'une valeur maximale permet de stabiliser l'inversion en contraignant l'algorithme à faire à chaque itération un petit pas dans la direction de descente. Cette contrainte peut aussi avoir pour conséquence d'augmenter le nombre d'itérations nécessaires à la convergence de l'algorithme.

4.1.4 Le critère d'arrêt

La définition du critère d'arrêt du schéma itératif d'inversion est directement liée à la définition de la qualité du modèle de vitesse obtenu. Le critère d'arrêt dépend alors essentiellement de l'expérience de l'utilisateur qui décide d'arrêter le processus itératif quand le modèle de vitesse obtenu lui semble satisfaisant. Dans la pratique, le critère d'arrêt peut être défini à partir de plusieurs valeurs : une valeur maximale d'itérations à réaliser, une valeur minimale de fonction coût à atteindre, une valeur minimale du pas appliqué au gradient de la fonction coût. Dans le cadre de ce travail, nous n'avons pas défini de critère d'arrêt préférentiel, nous utilisons un des critères définis ci-dessus en fonction de l'application considérée.

4.2 Données synthétiques 2-D

Dans cette partie, nous utilisons le modèle de vitesse synthétique 2-D BP EAGE [Billette and Brandsberg-Dahl, 2005]. A partir de ce jeu de données, nous avons réalisé un grand nombre de simulations qui nous ont permis d'acquérir un savoir-faire indispensable à l'application d'un algorithme de tomographie des temps de première arrivée.

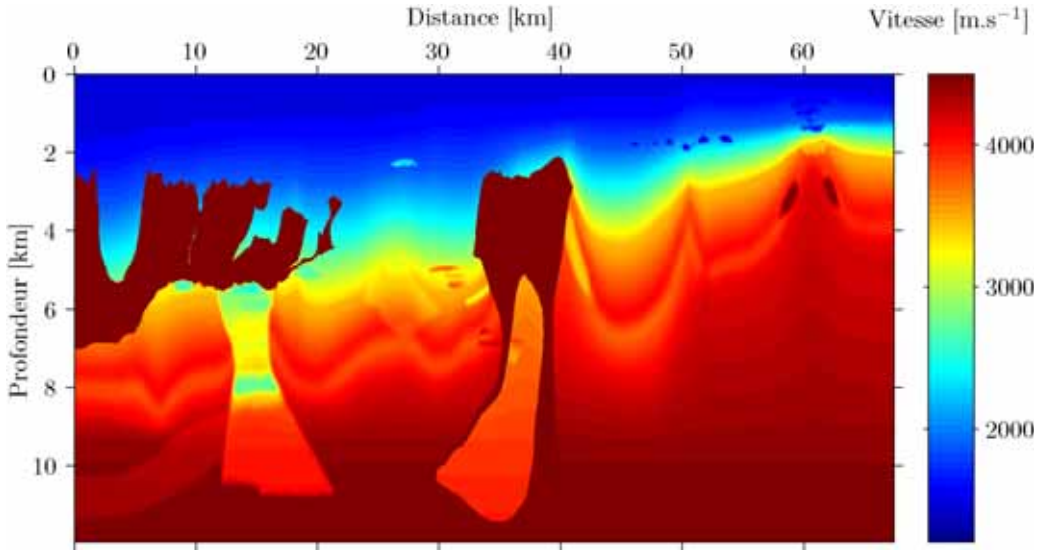

Fig. 4.1 – *Modèle de vitesse BP EAGE utilisé pour la génération de sismogrammes synthétiques, [Billette & Brandsberg-Dahl, 2005].*

Le modèle de vitesse BP EAGE défini par Billette et Brandsberg-Dahl [2005] mesure environ 67 km en longueur et 12 km en profondeur (Fig. 4.1). La discrétisation du modèle avec une maille carrée de 12.5 m donne une grille de taille [956 $\times$ 5395], soit environ 5.2 millions de mailles.

4.2.1 Modèle de vitesse observé et données synthétiques

Billette et Brandsberg-Dahl [2005] décrivent les principales étapes de construction du modèle de vitesse

- générer une couche d'eau avec une vitesse constante de 1485 $m.s^{-1}$,
- construire des couches de sédiments compactes et lisser les interfaces entre ces sédiments,
- construire et insérer des corps de sel dans le modèle, avec une vitesse constante de 4510 $m.s^{-1}$ pour le corps de sel situé à gauche et de 4790 $m.s^{-1}$ pour le corps de sel situé à droite (Fig. 4.1),
- définir manuellement un certain nombre d'anomalies et les insérer de manière lisse dans le modèle de vitesse,
- définir des valeurs de réflectivité à partir de données réelles et les appliquer à la densité du modèle.

La partie gauche du modèle est composée essentiellement d'un corps de sel complexe rugueux avec des inclusions de vitesse lentes se trouvant sous la structure saline. Cette partie du modèle est représentative de la géologie que l'on peut rencontrer en eaux profondes dans le Golfe du Mexique. La partie centrale du modèle est construite autour d'un dôme de sel profondément enraciné dans le milieu. Cette partie du modèle est aussi représentative du Golfe de Mexique et de l'Afrique de l'Ouest. La partie droite du modèle est presque exclusivement sans structure saline. La géologie de cette partie du modèle est commune à des zones telles que la Mer Caspienne et la Mer du Nord. Le champ de vitesse possède des variations significatives dans sa composante grande longueur d'onde et plusieurs petites anomalies de vitesse. La taille, la forme et la vitesse des anomalies sont variables.

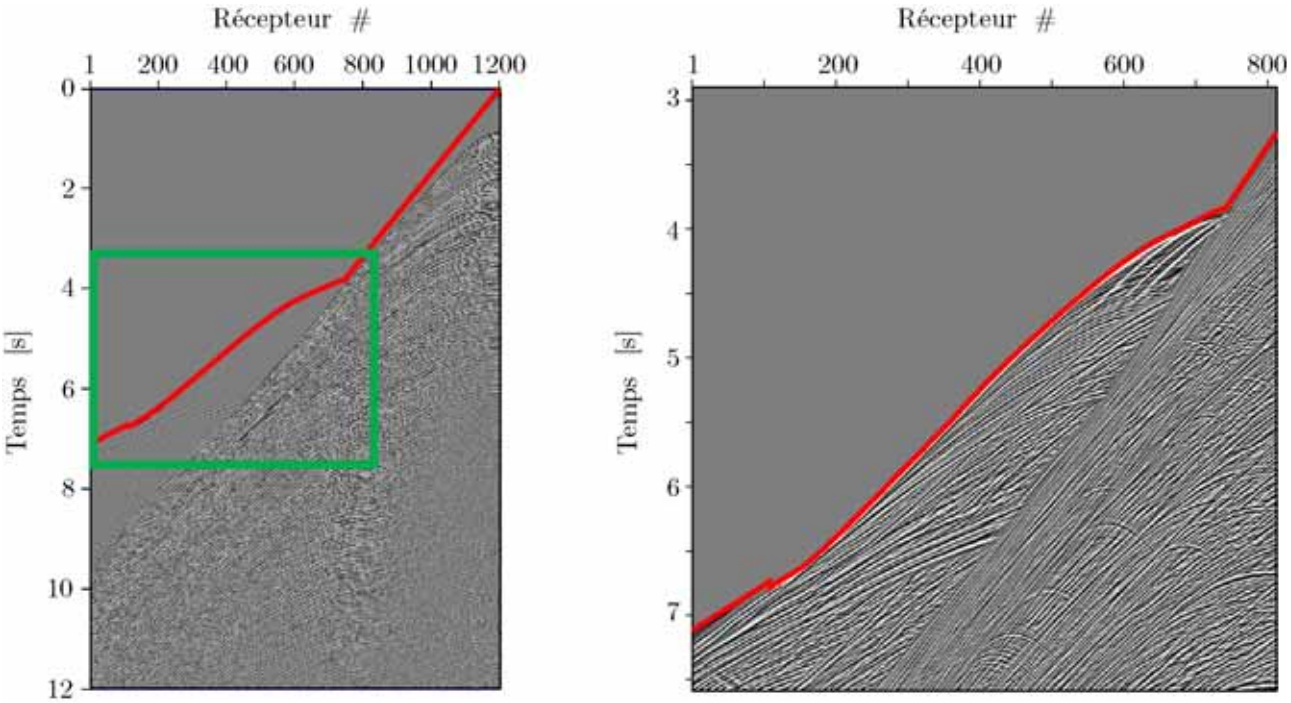

Fig. 4.2 – *a) Sismogramme synthétique généré par un algorithme de modélisation par différences finies acoustique temps à partir du modèle BP EAGE. La ligne rouge représente les temps de première arrivée pointés automatiquement à partir du premier mouvement, b) un zoom sur le sismogramme qui correspond au rectangle vert.*

Des sismogrammes synthétiques (Fig. 4.2) ont été générés à partir du modèle de vitesse BP EAGE par un algorithme de modélisation par différences finies acoustique temps. Ces sismogrammes simulent les enregistrements d'une acquisition marine, de type *streamer*, de 15 km de long avec un intervalle de 12.5 m entre récepteurs et de 50 m entre sources. Les récepteurs sont situés à gauche de la source, à une profondeur de 12.5 m. La première colonne à gauche du modèle de vitesse a été dupliquée pour permettre la génération des données synthétiques sur l'ensemble du modèle. 1348 points de tir ont ainsi été générés, chacun avec 1201 récepteurs.

4.2.2 Les temps de première arrivée

Nous utilisons un outil de pointé automatique, basé sur la détection du premier mouvement, pour déterminer la valeur des temps de première arrivée sur les sismogrammes (Fig. 4.2). Les sismogrammes ne sont pas bruités, ce qui facilite le pointé et donc la qualité de l'estimation des temps de première arrivée.

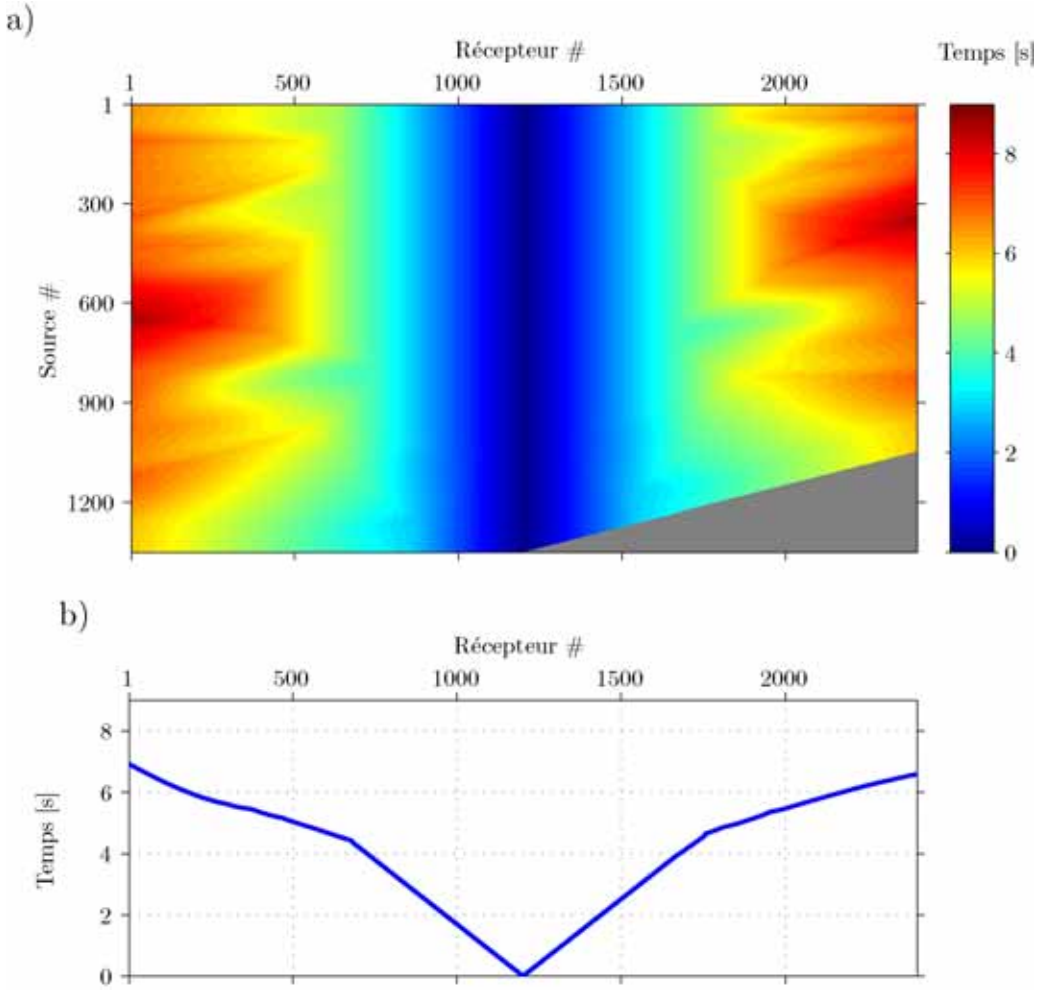

Fig. 4.3 – *a) Carte des temps de première arrivée observés pour 1348 sources et 2401 récepteurs. La partie grisée correspond aux acquisitions contenant seulement 1201 récepteurs placés à gauche de la source, b) les temps de première arrivée observés pour une source située en (12.5 m, 60 km) et tous les récepteurs.*

Pour étudier la capacité de l'algorithme de tomographie à gérer un grand nombre de données observées, nous choisissons d'augmenter artificiellement le nombre de récepteurs. Pour cela, nous appliquons le principe de réciprocité [Arntsen and Carcione, 2000] qui permet de déterminer les temps de première arrivée pour une acquisition symétrique de récepteurs positionnés à droite de la source. Le nombre total de données observées par point de tir est alors de 2401, et le nombre total de données observées est d'environ 3.2 millions. La figure (Fig 4.3 a) représente la carte des temps de première arrivée observés pour toutes les sources et tous les récepteurs. Les temps de première arrivée pour une source de coordonnées (12.5 m, 60 km) et 2401 récepteurs sont représentés (Fig 4.3 b). Les coordonnées des sources et des récepteurs correspondent respectivement à la profondeur et à la distance.

4.2.3 Paramétrisation du schéma d'inversion

Nous choisissons pour modèle de vitesse initial (Fig. 4.4), un modèle possédant une loi de vitesse linéaire, de 3000 m.s^{-1} à 6000 m.s^{-1}, en fonction de la profondeur, de 0 km à 12 km. Le calcul des temps de première arrivée à partir de ce modèle initial, en utilisant le solveur eikonal défini par Podvin et Lecomte [1991], permet de déterminer la carte des résidus initiaux entre les temps observés et calculés pour toutes les positions de sources et de récepteurs (Fig. 4.5). La valeur de la racine carrée de la moyenne des carrés des résidus des temps, aussi appelée valeur *Root Mean Square* (RMS), est environ de 1.99 s.

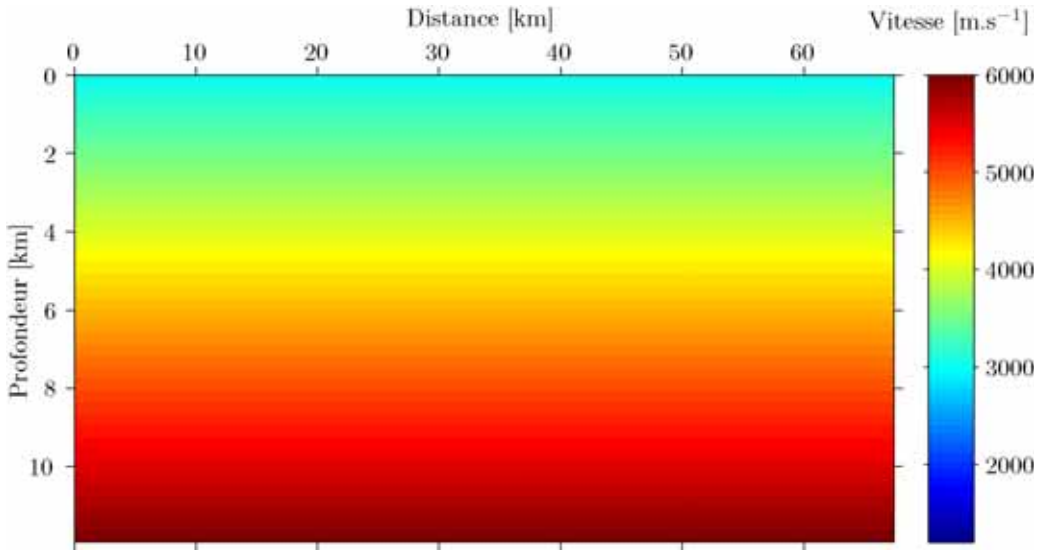

Fig. 4.4 – *Modèle de vitesse initial choisi pour le processus de minimisation.*

Le préconditionnement appliqué au gradient de la fonction coût consiste essentiellement en un filtrage gaussien. La longueur du filtrage appliqué au gradient de la fonction coût varie en deux étapes. Pour les premières itérations du schéma d'inversion, la longueur du filtre est de 6 km horizontalement et verticalement. Pour les dernières itérations, la longueur du filtre est de 6 km en longueur et de 1 km en profondeur. Pour assurer la stabilité et la convergence de l'algorithme, nous choisissons une valeur maximale du pas appliqué au gradient de la fonction

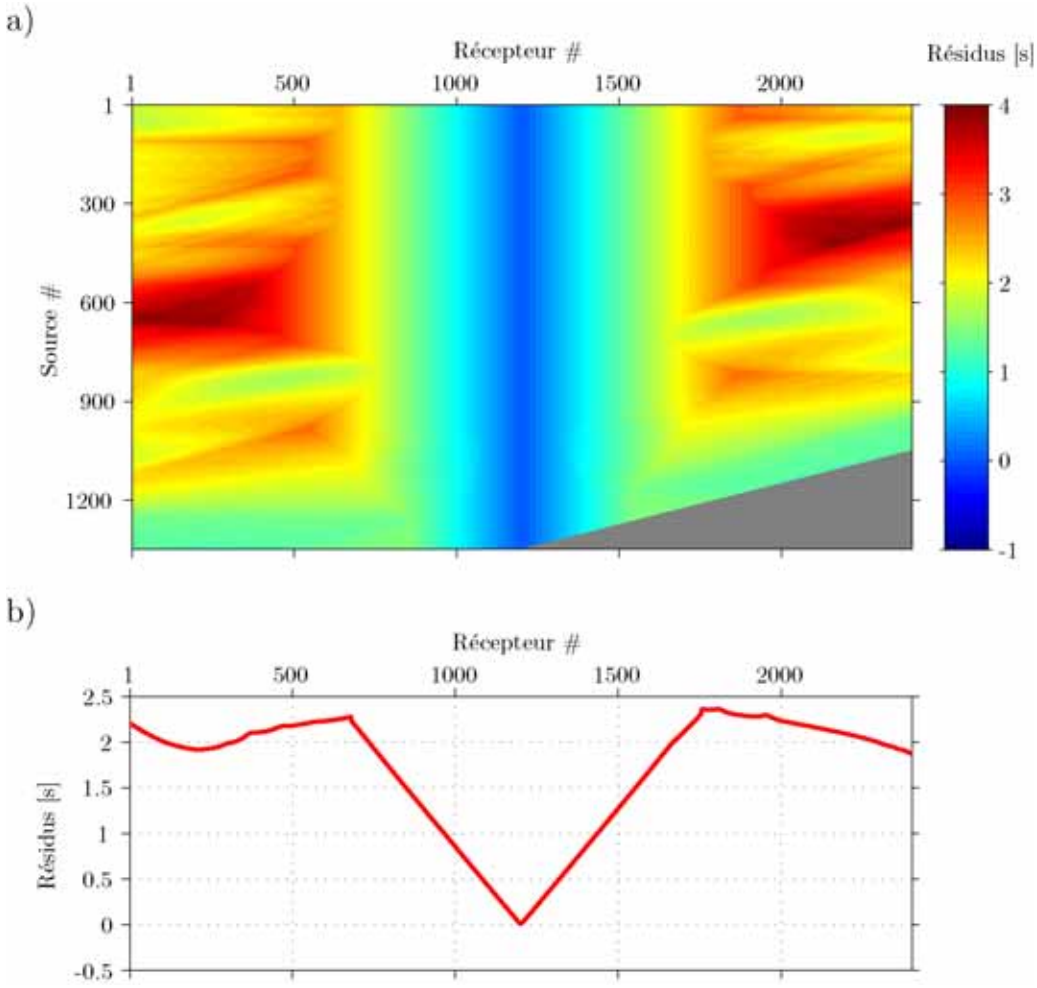

Fig. 4.5 – *a) Carte des résidus initiaux en temps pour toutes les positions de sources et de récepteurs. La partie grisée correspond aux acquisitions contenant seulement 1201 récepteurs placés à gauche de la source, b) les résidus initiaux en temps pour une source située en (12.5 m, 60 km) et tous les récepteurs.*

coût égale à 200 m.s^{-1}. Le schéma itératif est arrêté lorsque la valeur du pas calculée à chaque itération est nulle, i.e. lorsque l'algorithme de tomographie ne modifie plus le modèle courant.

4.2.4 Résultats de tomographie

Avant de présenter les résultats obtenus par le nouvel algorithme de tomographie des temps de première arrivée sur ce jeu de données, nous définissons les notions de limite de couverture des rais et de modèle vrai lissé. Un tracé de rais *a posteriori* réalisé à partir du modèle de vitesse BP EAGE (Fig. 4.6 a) pour plusieurs positions de sources et de récepteurs permet de définir une zone grisée, pour laquelle les rais, et donc les temps de première arrivée, ne contiennent aucune information. On ne peut donc pas espérer retrouver cette partie du modèle avec un algorithme de tomographie des temps de première arrivée, c'est ce que nous appelons la limite de couverture des rais. Le modèle de vitesse BP EAGE possède des structures et des

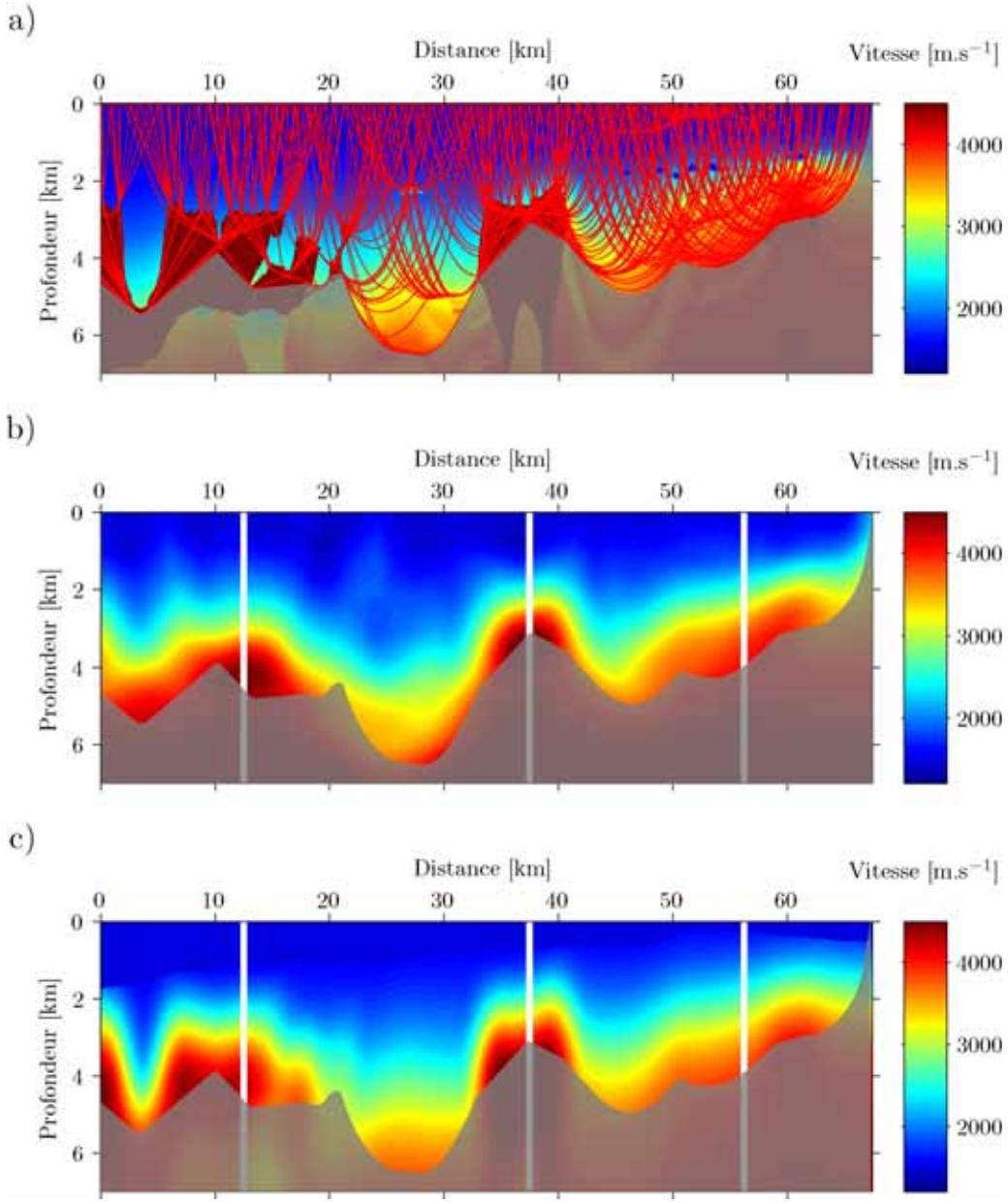

Fig. 4.6 – *a) Tracé de rais a posteriori réalisé à partir du modèle observé pour définir une zone grisée pour laquelle les rais ne contiennent pas d'information, b) modèle de vitesse obtenu par l'algorithme de tomographie des temps de première arrivée, c) modèle vrai lissé pour obtenir un contenu spectral équivalent au modèle obtenu après inversion. Les lignes verticales blanches marquent la position des profils 1-D représentés (Fig. 4.7).*

éléments de nature haute fréquence qu'un algorithme de tomographie des temps de première arrivée peut difficilement retrouver. Afin de pouvoir valider le résultat obtenu, nous choisissons de lisser le modèle de vitesse BP EAGE avec un filtre gaussien de longueurs horizontale et verticale d'environ 3 km, le résultat obtenu est appelé modèle vrai lissé (Fig. 4.6 c). Ce modèle vrai lissé possède approximativement le même contenu spectral que le modèle obtenu par le schéma d'inversion (Fig. 4.6 b). Le modèle inversé présenté ici est obtenu après 55 itérations

du schéma itératif de plus grande descente. Pour donner un ordre de grandeur, cela représente un temps de calcul d'environ 5 minutes en utilisant 512 cœurs sur le supercalculateur défini à la partie (3.4.3). La valeur finale de la racine de la moyenne des carrés des résidus en temps est environ de 0.06 s, la valeur initiale est de l'ordre de 1.99 s. On constate que les modèles inversé et vrai lissé ont une apparence très proche. On peut noter l'influence prépondérante de la couverture des rais dans le modèle observé (Fig. 4.6 a). En effet, sur la partie gauche du modèle, la couverture des rais est faible et ne permet pas une estimation correcte du modèle de vitesse par l'algorithme de tomographie des temps de première arrivée.

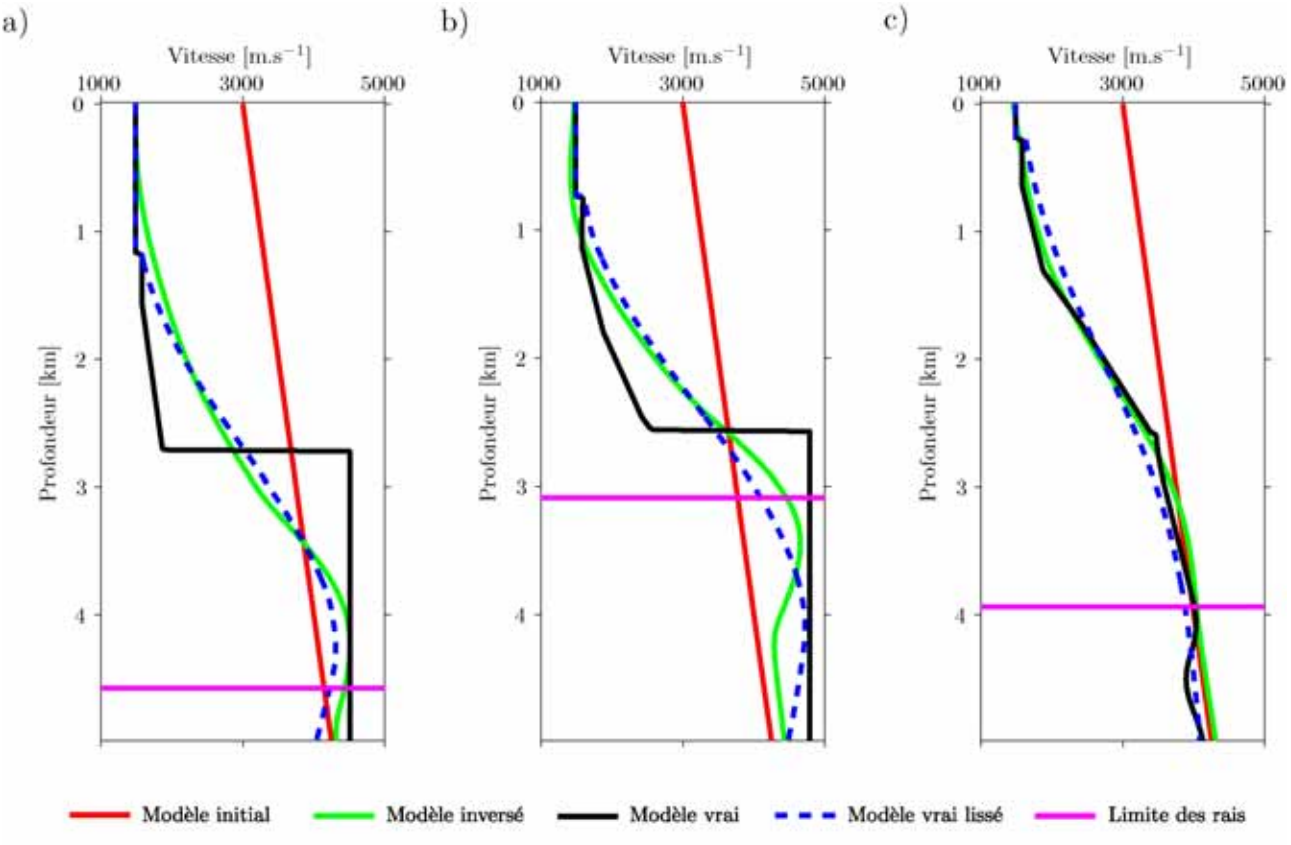

Fig. 4.7 – *Profils de vitesse 1-D à différentes positions dans les modèles de vitesse a) 12.5 km, b) 37.5 km, c) 56.25 km. La courbe rouge correspond au modèle initial, la courbe verte au modèle inversé, la courbe noire au modèle vrai, la courbe bleue au modèle vrai lissé et la courbe rose à la limite de la couverture des rais.*

Les profils de vitesse 1-D représentés à différentes positions, 12.5 km, 37.5 km et 56.25 km, dans les modèles (Fig. 4.7) permettent de valider le fait que les modèles de vitesse inversé et vrai lissé sont proches à condition que la couverture des rais soit suffisante. Le modèle inversé a été obtenu sans fixer préalablement le fond de l'eau. Les résidus finaux définis entre les temps observés et les temps calculés à partir du modèle inversé (Fig. 4.8) permettent de vérifier les temps de première arrivée observés sont bien expliqués par le modèle de vitesse inversé.

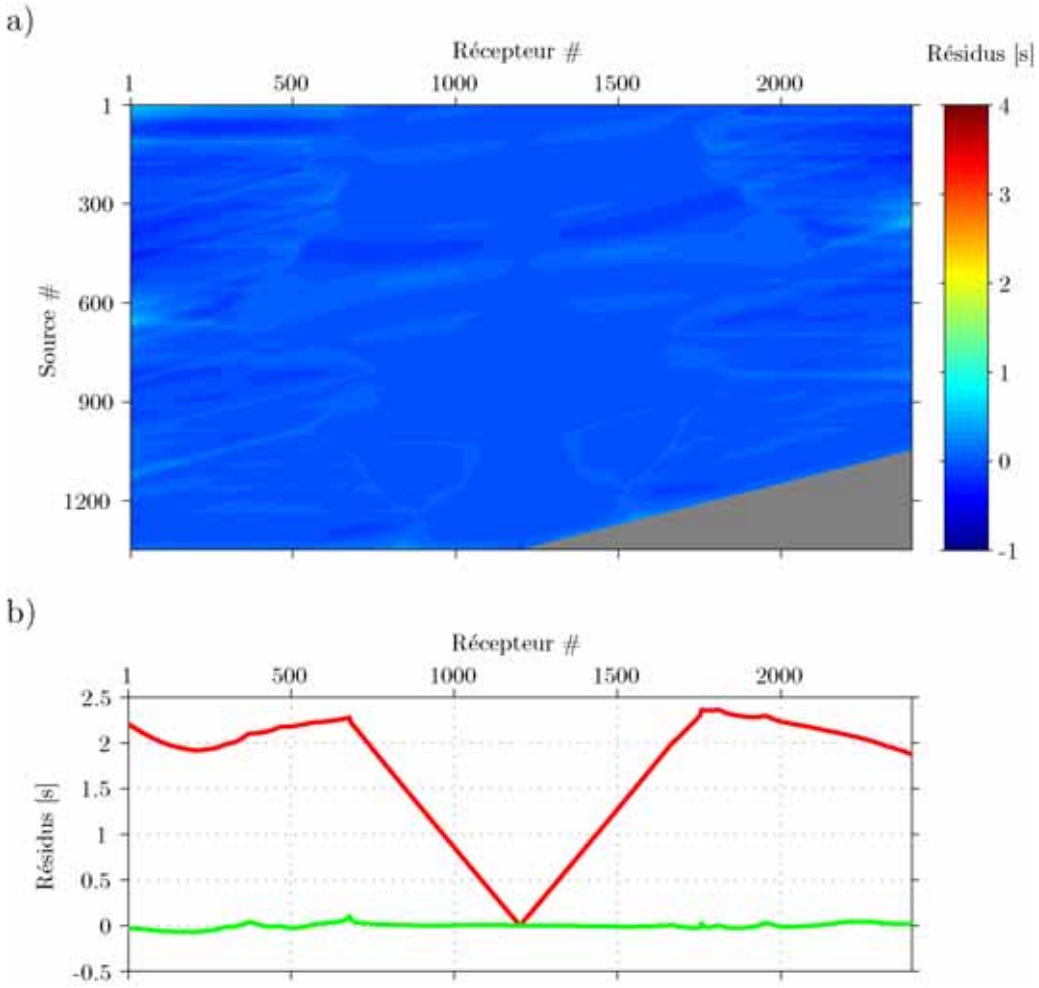

Fig. 4.8 – *a) Carte des résidus finaux en temps pour toutes les positions de sources et de récepteurs. La partie grisée correspond aux acquisitions contenant seulement 1201 récepteurs placés à gauche de la source, b) comparaison des résidus initiaux (en rouge) et des résidus finaux (en vert) en temps pour une source située en (12.5 m, 60 km) et tous les récepteurs.*

4.3 Données synthétiques 3-D

Dans cette partie, nous utilisons une version modifiée du modèle de vitesse synthétique 3-D Overthrust [Aminzadeh et al., 1997]. L'objectif de cette partie est de montrer que le nouvel algorithme de tomographie des temps de première arrivée peut gérer des modèles de grande dimension et des grandes quantités de données observées. Les résultats présentés ici ont été obtenus grâce au savoir-faire acquis pour les applications en 2-D.

4.3.1 Modèle de vitesse observé

Le modèle de vitesse Overthrust défini par Aminzadeh et al. [1997] mesure 20 km en longueur et en largeur et environ 4.5 km en profondeur (Fig. 4.9 a). La discrétisation du modèle avec une maille cubique de 50 m donne une grille de taille [94 $\times$ 401 $\times$ 401], soit environ 15.1 millions

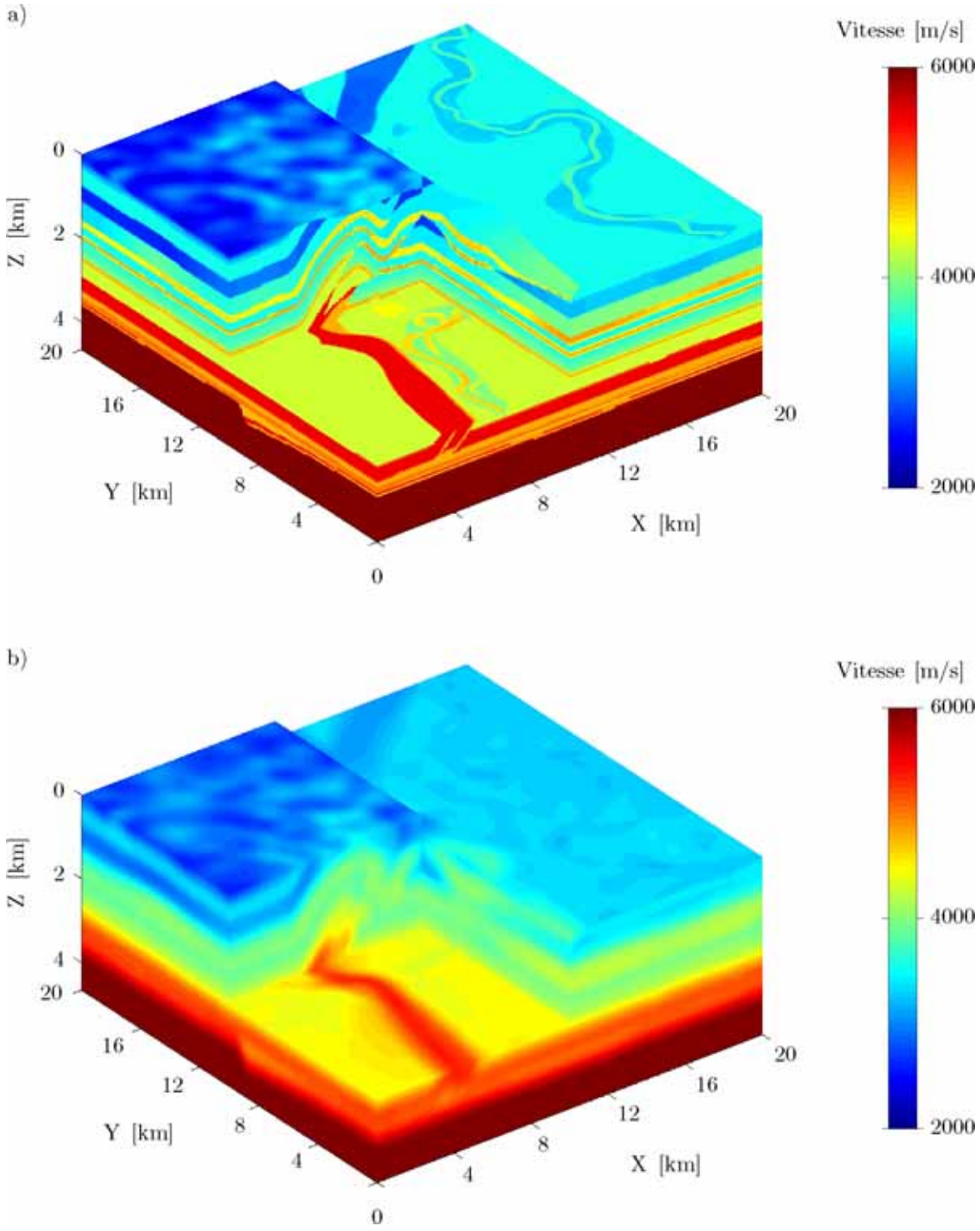

Fig. 4.9 – *a) Modèle de vitesse Overthrust original [Aminzadeh et al., 1995], b) modèle de vitesse Overthrust lissé défini pour l'application de tomographie des temps de première arrivée.*

de mailles. Aminzadeh et al. [1997] décrivent les principales caractéristiques du modèle de vitesse. Le modèle est défini à partir de 17 couches géologiques de complexités différentes et avec des vitesses allant de 2500 m.s^{-1} à 6000 m.s^{-1}. Le modèle présente deux chevauchements plus un chevauchement qui disparait latéralement. La partie haute du modèle représente une

zone érodée avec des variations de vitesse importantes. De par sa construction en couches géologiques, le modèle de vitesse possède une structure haute fréquence avec des sauts de vitesse très importants. Cette structure haute fréquence affecte la précision numérique du solveur eikonal et crée une couverture des rais peu favorable à l'application d'un code de tomographie des temps de première arrivée. C'est la raison pour laquelle, nous choisissons de lisser le modèle Overthrust original avec un filtre gaussien de 800 m dans les trois dimensions (Fig. 4.9 b).

4.3.2 Les temps de première arrivée

A partir du modèle lissé, nous simulons une acquisition sismique en réfraction avec 400 sources spatialement réparties à la surface du modèle, chaque source possèdent 401 × 401 récepteurs à la surface. L'*offset* maximal en longueur et en largeur est de 20 km. Les données observées (Fig. 4.10) sont calculées par le solveur eikonal défini par Podvin et Lecomte [1991]. Le nombre total de données observées est d'environ 64.3 millions. Les coordonnées des sources et des récepteurs correspondent respectivement aux dimensions Z, X et Y.

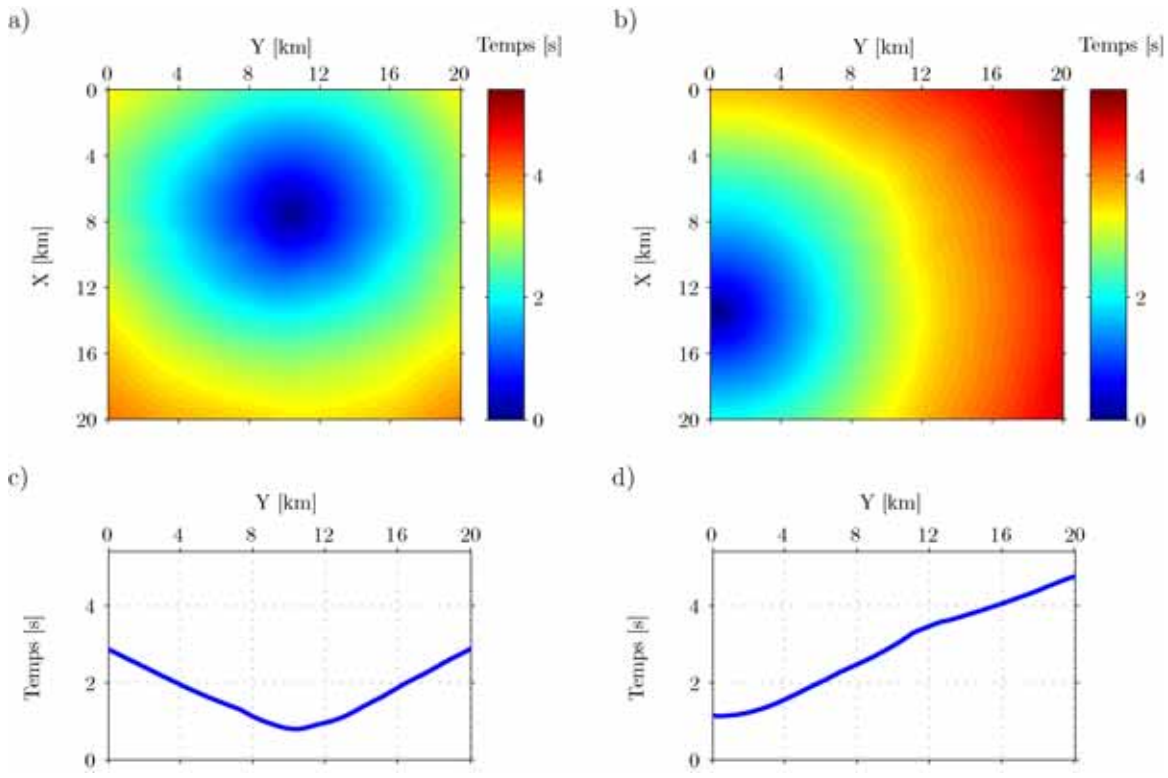

Fig. 4.10 – *Cartes des temps de première arrivée calculés pour une source située en (0 km, 7.4 km, 10.4 km) a) et en (0 km, 13.4 km, 0.4 km) b). Les temps de première arrivée observés pour un récepteur situé en (0 km, 10 km) et une source située en (0 km, 7.4 km, 10.4 km) c) et en (0 km, 7.4 km, 10.4 km) d).*

4.3.3 Paramétrisation du schéma d'inversion

Nous choisissons pour modèle de vitesse initial (Fig. 4.11), un modèle possédant une loi de vitesse linéaire, de 5000 m.s^{-1} à 6000 m.s^{-1}, en fonction de la profondeur, de 0 km à 4.5 km. Le calcul des temps de première arrivée à partir de ce modèle initial, en utilisant le solveur eikonal défini par Podvin et Lecomte [1991], permet de déterminer la carte des résidus initiaux entre les temps observés et calculés pour toutes les positions de sources et de récepteurs (Fig. 4.12). La valeur de la racine carrée de la moyenne des carrés des résidus des temps (RMS) est environ de 722.2 ms.

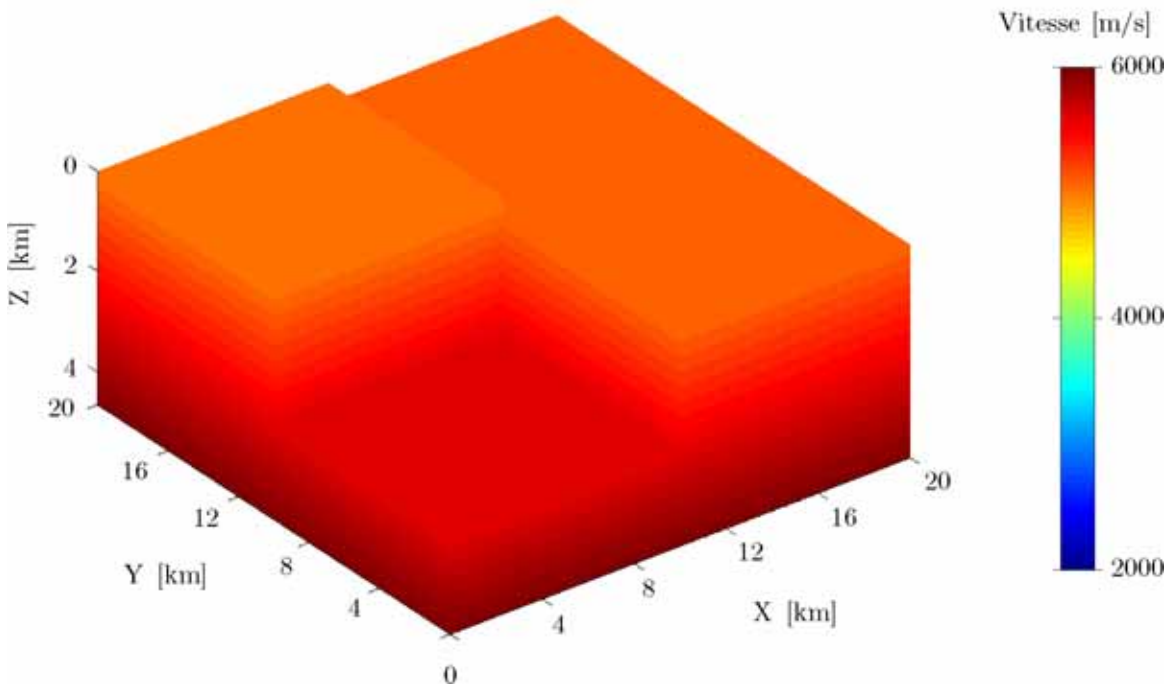

Fig. 4.11 – *Modèle de vitesse initial choisi pour le processus de minimisation.*

Le préconditionnement appliqué au gradient de la fonction coût consiste essentiellement en un filtrage gaussien. La longueur du filtrage appliqué au gradient de la fonction coût varie en deux étapes. Pour les premières itérations du schéma d'inversion, la longueur du filtre est de 6 km en longueur et en largeur et de 2 km en profondeur. Pour les dernières itérations, la longueur du filtre est de 2 km en longueur et en largeur et de 0.4 km en profondeur. Pour assurer la stabilité et la convergence de l'algorithme, nous choisissons une valeur maximale du pas appliqué au gradient de la fonction coût égale à 200 m.s^{-1}. Le schéma itératif est arrêté lorsque la valeur du pas calculée à chaque itération est nulle, i.e. lorsque l'algorithme de tomographie ne modifie plus le modèle courant.

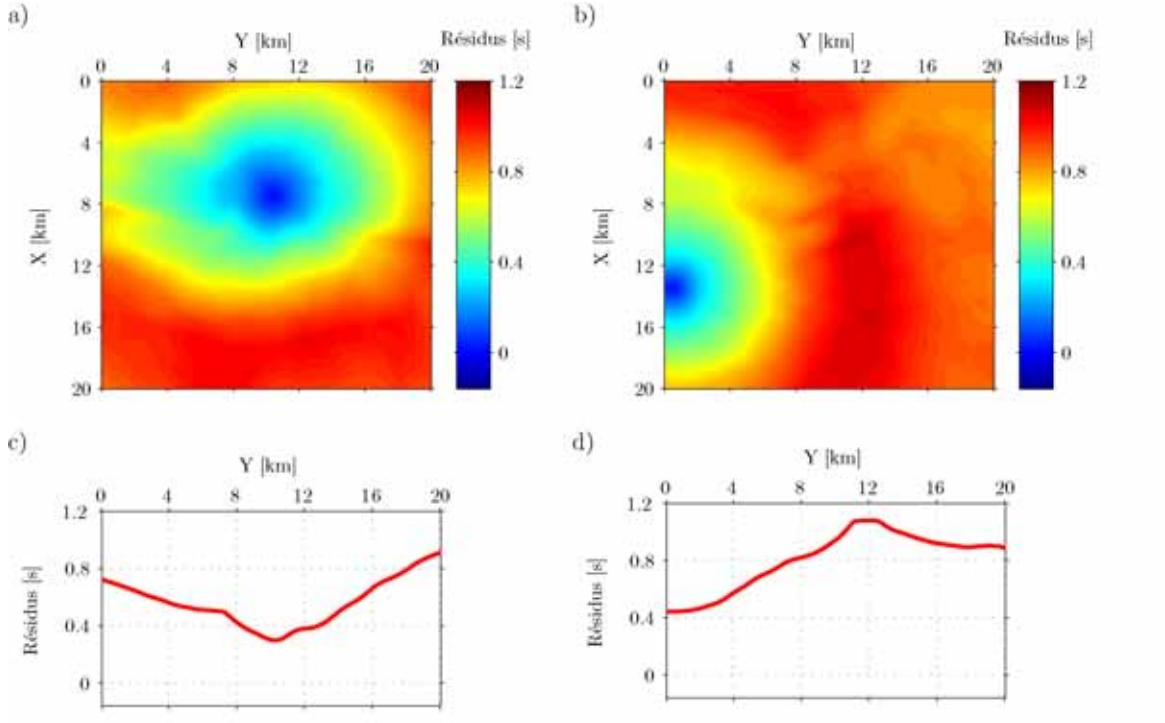

Fig. 4.12 – *Cartes des résidus initiaux en temps calculés pour une source située en (0 km, 7.4 km, 10.4 km) a) et en (0 km, 13.4 km, 0.4 km) b). Les résidus initiaux en temps pour un récepteur situé en (0 km, 10 km) et une source située en (0 km, 7.4 km, 10.4 km) c) et en (0 km, 6.7 km, 0.2 km) d).*

4.3.4 Résultats de tomographie

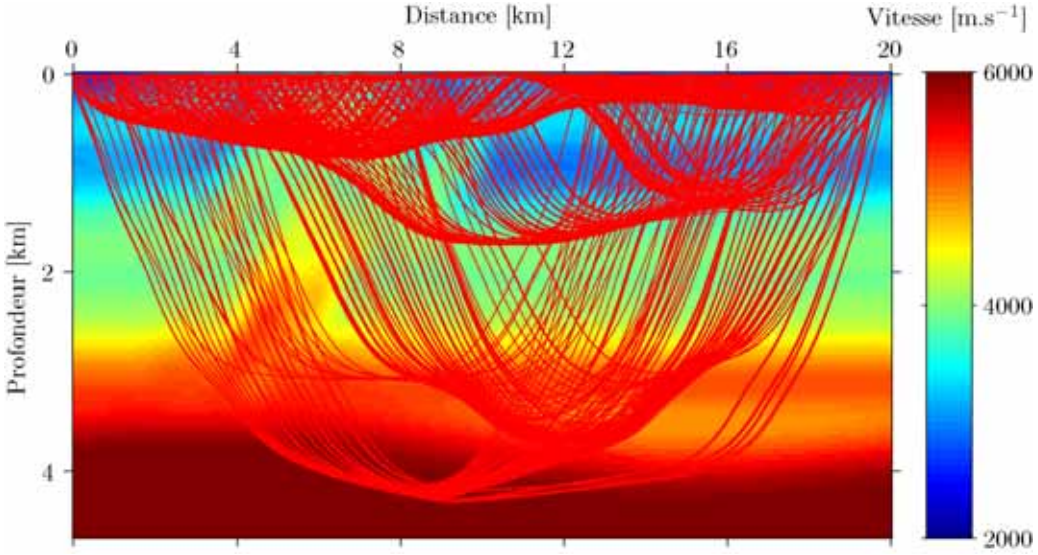

Fig. 4.13 – *Tracé de rais a posteriori réalisé à partir d'une coupe 2-D du modèle observé en $Y = 10$ km pour plusieurs positions de sources et de récepteurs positionnés à la surface.*

La figure (Fig. 4.13) représente un tracé de rais a posteriori réalisé à partir d'une coupe 2-D du modèle observé en Y = 10 km pour plusieurs positions de sources et de récepteurs positionnés à la surface. On peut noter la couverture inhomogène des rais sur une grande partie du modèle et la présence de nombreux rais subverticaux. Cette couverture inhomogène des rais est principalement due à la présence dans le modèle de structures à valeur de vitesse très élevée mais aussi à la présence de structures possédant des inversions de vitesse. Ces conditions rendent difficile l'estimation du modèle de vitesse à partir des temps de première arrivée. Nous choisissons de présenter les résultats de tomographie essentiellement sur la partie supérieure, de 0 km à 1 km environ, du modèle de vitesse où la couverture des rais est relativement homogène.

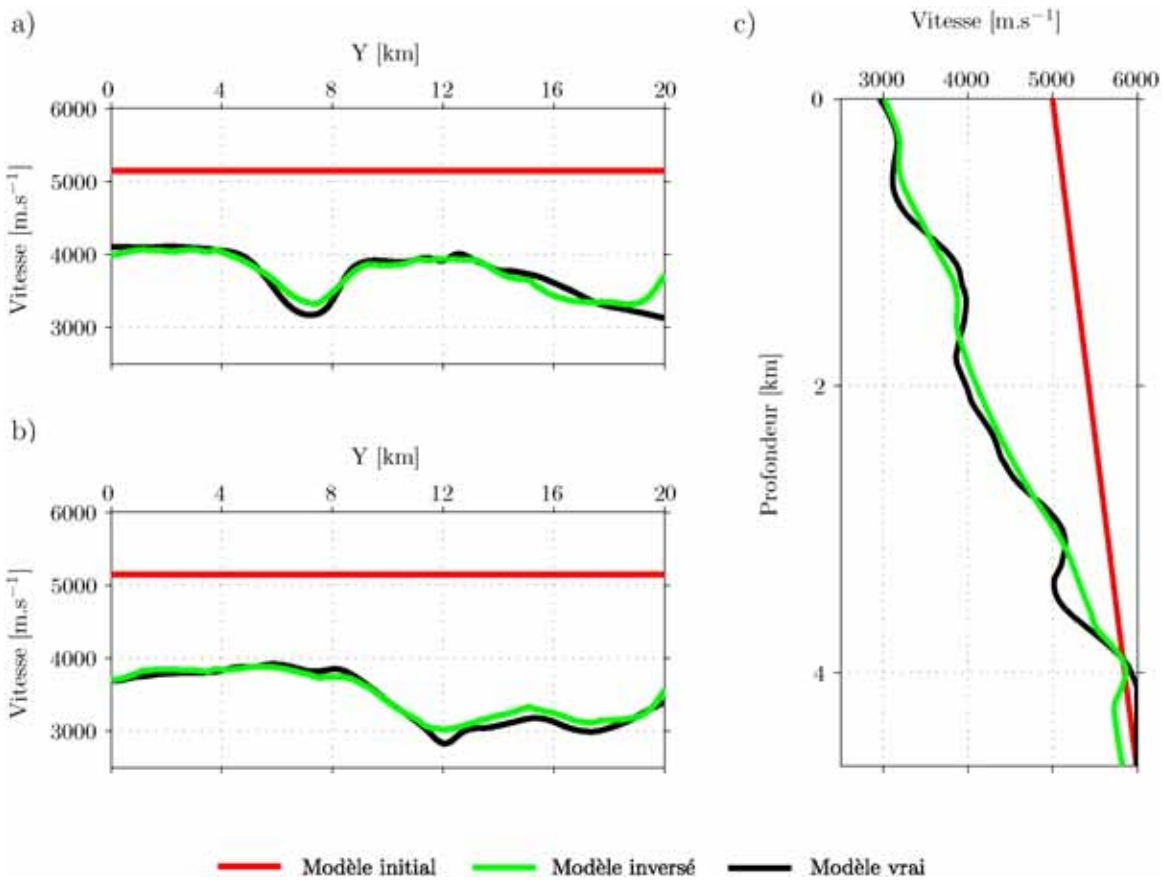

Fig. 4.14 – *Profils de vitesse horizontaux 1-D à une profondeur constante de Z = 700 m et à différentes positions dans les modèles de vitesse a) X = 10 km et b) X = 15 km, c) profil de vitesse vertical 1-D à X = 15 km et Y = 10 km. La courbe rouge correspond au modèle initial, la courbe verte au modèle inversé et la courbe noire au modèle recherché.*

Le modèle inversé (Fig. 4.15 a) est obtenu après 60 itérations du schéma itératif de plus grande descente. Pour donner un ordre de grandeur, cela représente un temps de calcul d'environ 1 heure en utilisant 400 cœurs sur le supercalculateur défini à la partie (3.4.3). La valeur finale de la racine de la moyenne des carrés des résidus en temps est environ de 0.4 ms, la valeur initiale est de l'ordre de 722.2 ms. Les modèles inversé et recherché (Fig. 4.15 b) ont une apparence très proche, cette tendance est confirmée par des profils de vitesse 1-D à différentes positions

à travers les modèles (Fig. 4.14). Les résidus finaux définis entre les temps de première arrivée observés et ceux calculés à partir du modèle inversé (Fig. 4.16) permettent de confirmer que l'algorithme de tomographie a bien expliqué les temps de première arrivée observés.

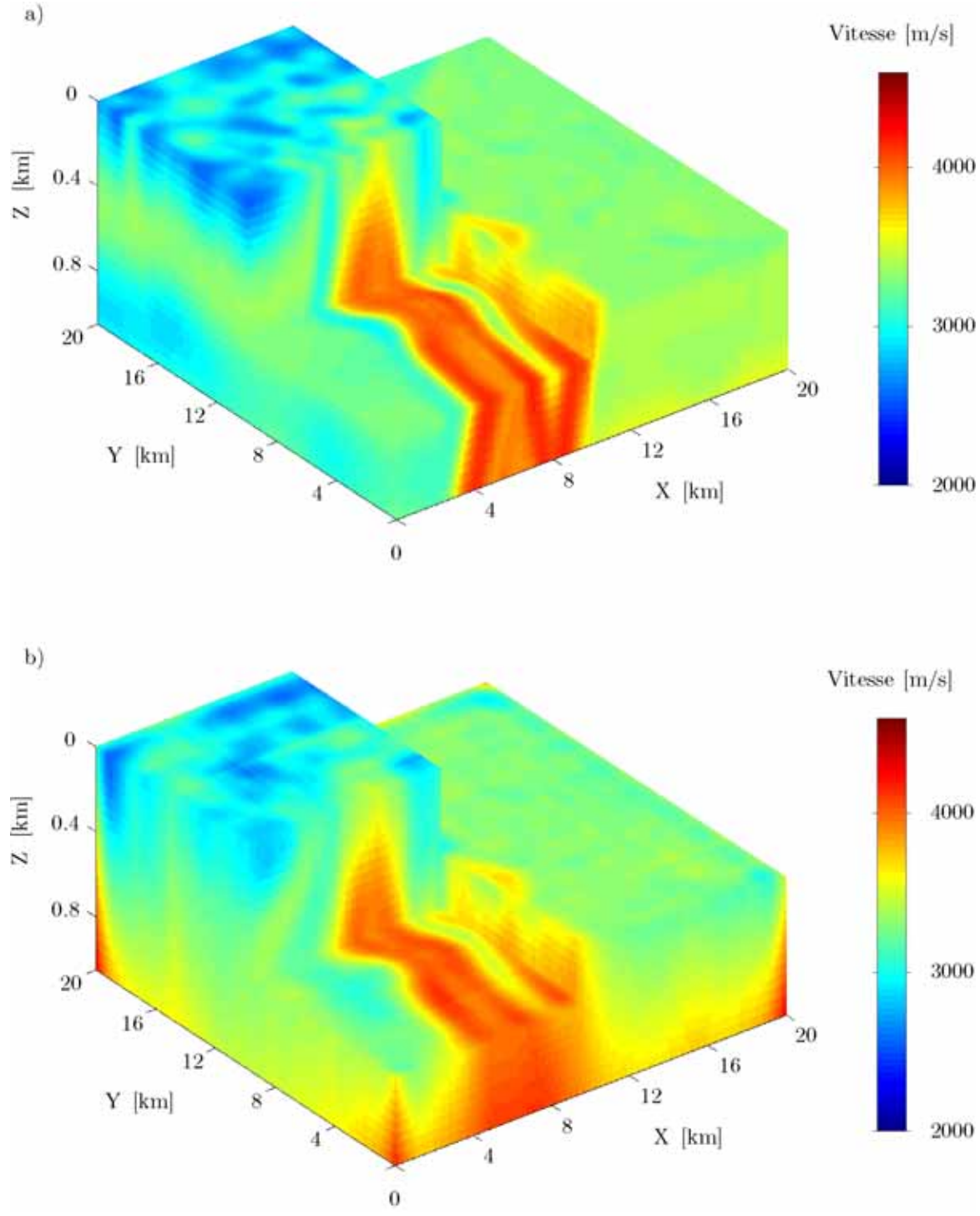

Fig. 4.15 – *a) Modèle de vitesse obtenu par l'algorithme de tomographie des temps de première arrivée, b) modèle de vitesse que l'on cherche à déterminer. Les modèles de vitesse sont représentés pour une profondeur maximale de 1 km.*

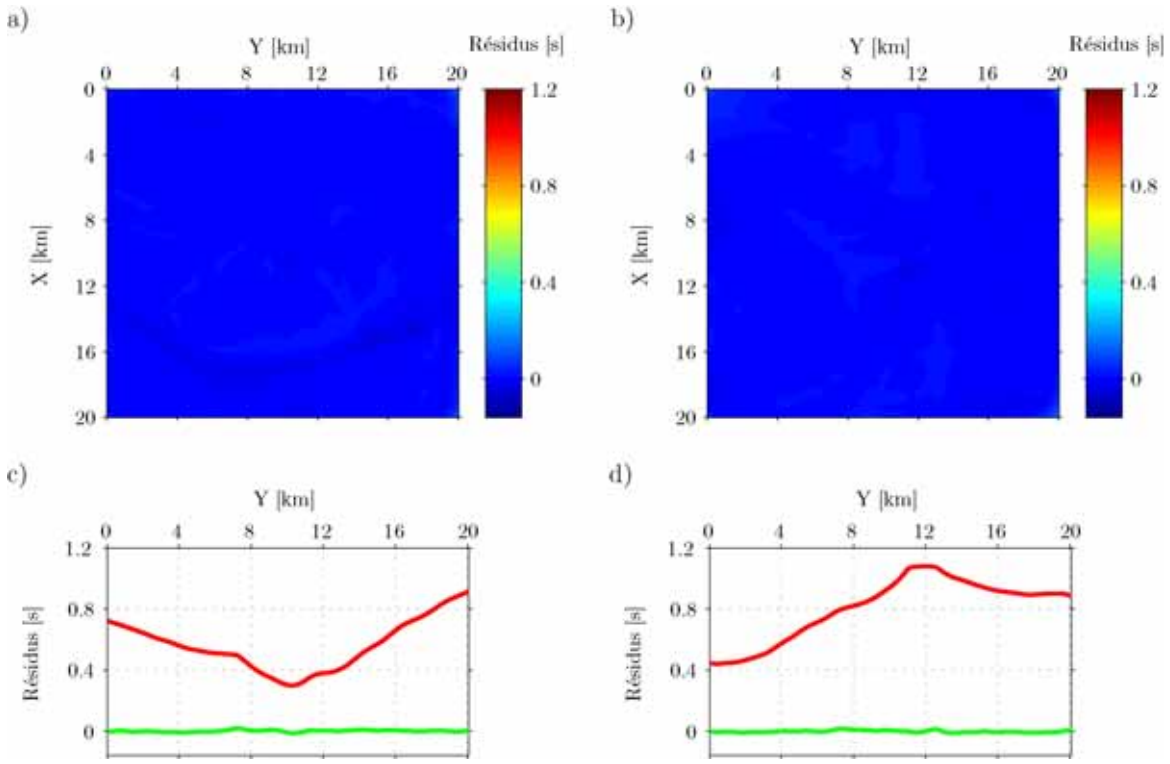

Fig. 4.16 – *Cartes des résidus finaux en temps calculés pour une source située en (0 km, 7.4 km, 10.4 km) a) et en (0 km, 13.4 km, 0.4 km) b). Les résidus finaux en temps pour un récepteur situé en (0 km, 10 km) et une source située en (0 km, 7.4 km, 10.4 km) c) et en (0 km, 13.4 km, 0.4 km) d).*

4.4 Données réelles 2-D

Nous présentons ici un exemple d'application du nouvel algorithme de tomographie des temps de première arrivée à un jeu de données réelles. Les objectifs de cette partie sont multiples. Tout d'abord nous voulons montrer que le savoir-faire acquis à partir de données synthétiques est transposable à des données réelles. Ensuite nous voulons étudier la stabilité du nouveau schéma d'inversion sur des données réelles bruitées, c'est à dire par exemple avec des erreurs de pointés, et du bruit ambiant. Enfin, nous voulons illustrer le fait que l'algorithme et la méthodologie mis en œuvre permettent d'utiliser un modèle de vitesse initial défini sans aucune connaissance *a priori*. Les temps de première arrivée pointés et les modèles de vitesse que nous utilisons pour valider les résultats que nous avons obtenus nous ont été fournis par Stéphane Operto et Jean-Xavier Dessa de l'Unité Mixte de Recherche Géosciences Azur.

4.4.1 L'acquisition des données

Les données utilisées dans cette partie ont été acquises pendant la campagne d'acquisition sismique KY0106 réalisée par l'*Institute for Frontier Research on Earth Evolution* (IFREE) à bord du R/V Kaiyo et avec pour maître d'œuvre le *Japan Marine Science and Technology* (JAM-

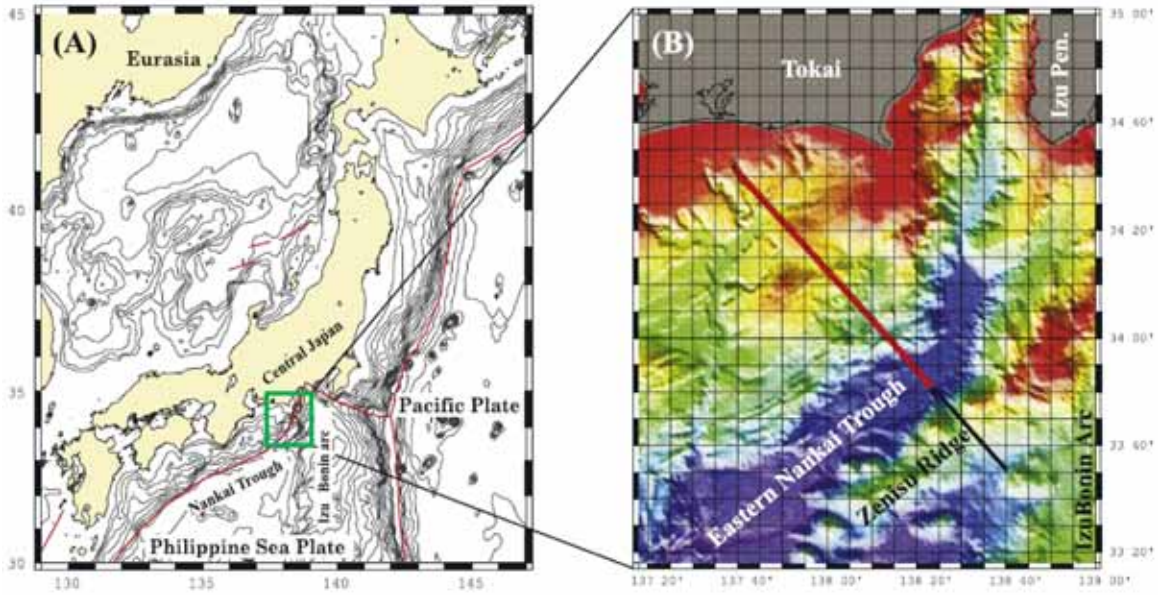

Fig. 4.17 – *a) Contexte géodynamique général autour de la fosse de Nankai, b) zoom sur le rectangle vert avec les principales structures de la zone étudiée. La ligne noire correspond à la position des points de tir et la ligne rouge coïncidente à la position des sismomètres sous-marins, extrait de [Operto et al., 2006].*

STEC). Cette campagne d'acquisition s'inscrivait dans le cadre du projet franco-japonais *Seize France Japan* (SFJ) dont un des objectifs était d'obtenir une image structurale du segment de Tokai situé à l'Est de la fosse de Nankai (Fig. 4.17). La fosse de Nankai, située au Sud-Ouest des côtes du Japon, correspond à la subduction de la plaque Philippine sous la plaque Eurasie. Le segment de Tokai est l'unique segment de la fosse de Nankai a ne pas avoir rompu lors des derniers tremblements de terre. Un séisme de forte amplitude est donc attendu dans les prochaines années dans cette zone qui est devenue l'objet de très nombreuses études. Une description détaillée de la zone étudiée est réalisée dans [Dessa et al., 2004] et [Operto et al., 2006]. Un dispositif dense de 100 sismomètres sous-marins, *Ocean Bottom Seismometer* (OBS), espacés de 1 km, a été déployé au fond de l'eau suivant une direction perpendiculaire aux structures de subduction (Fig. 4.17). Une acquisition de 140 km de long, avec un espacement de 100 m entre les points de tir a été effectuée.

4.4.2 Les temps de première arrivée observés

Les temps de première arrivée ont été pointés à partir des enregistrements des sismomètres (Fig. 4.18) avec un traitement semi-automatique réalisé et décrit par Dessa et al. [2004]. Parmi tous les enregistrements effectués, 91 sismomètres ont fourni des données exploitables. Les temps de première arrivée qui n'ont pu être pointés avec une précision suffisante ne sont pas considérés.

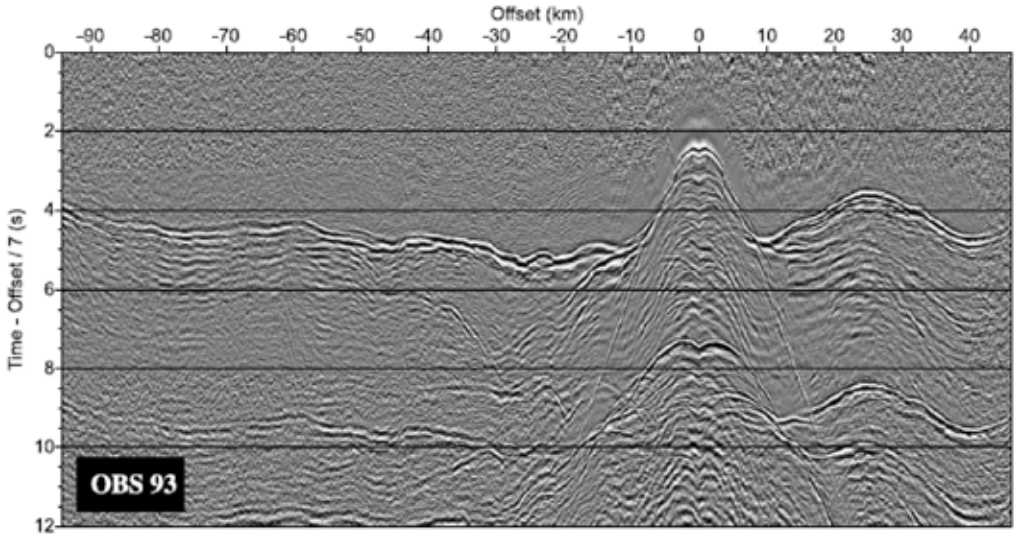

Fig. 4.18 – *Un exemple d'enregistrement réalisé par un sismomètre sous-marin, extrait de [Dessa et al., 2004].*

Le temps d'exécution du nouvel algorithme de tomographie des temps de première arrivée dépend du nombre de points de tir considérés. En vertu du principe de réciprocité , on interchange alors *virtuellement* les sources et les récepteurs [Arntsen and Carcione, 2000]. Ainsi, le nombre de sources est de 91 et le nombre maximal de récepteurs est de 1398. Les sources sont considérées comme étant situées sur le fond de l'eau et les récepteurs à 10 m de profondeur. Le nombre total de données observées considérés pour l'application de l'algorithme de tomographie des temps de première arrivée est environ de 100 000 (Fig. 4.19). Les coordonnées des sources et des récepteurs correspondent respectivement à la profondeur et à la distance.

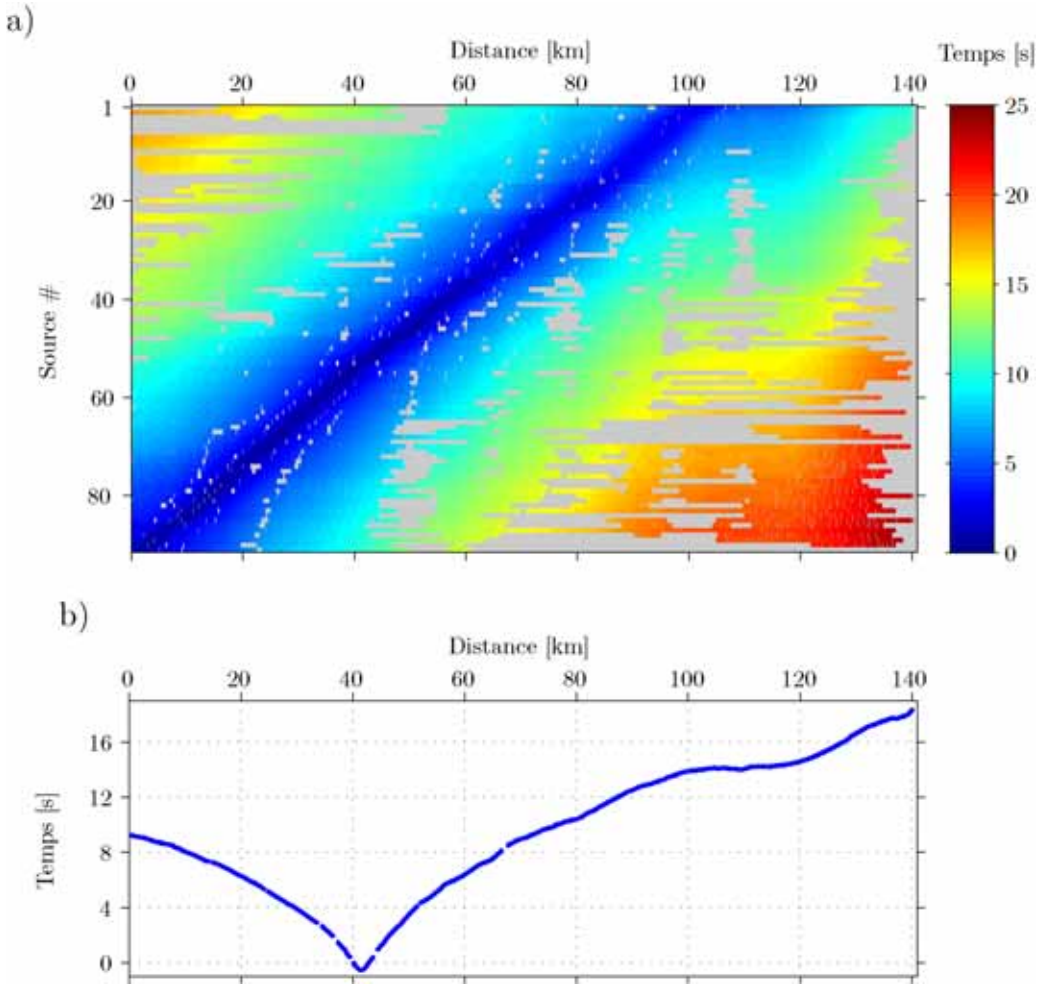

Fig. 4.19 – *a) Carte des temps de première arrivée observés obtenus pour toutes les sources et les récepteurs disponibles. Les parties grisées correspondent aux temps de première arrivée qui n'ont pu être pointés avec une précision suffisante. b) Les temps de première arrivée observés pour une source située en (0.75 km, 41.5 km) et tous les récepteurs disponibles.*

4.4.3 Paramétrisation du schéma d'inversion

Nous choisissons pour modèle de vitesse initial (Fig. 4.20), un modèle possédant une loi de vitesse linéaire, de 8000 m.s^{-1} à 9000 m.s^{-1}, en fonction de la profondeur, de 0 km à 40 km. La taille du modèle considéré est de 141 km en longueur et de 40 km en profondeur. La discrétisation du modèle avec une maille carrée de 25 m donne une grille de taille [1601 $\times$ 5641], soit environ 9 millions de mailles. Le calcul des temps de première arrivée à partir de ce modèle initial, en utilisant le solveur eikonal défini par Podvin et Lecomte [1991], permet de déterminer la carte des résidus initiaux entre les temps observés et calculés pour toutes les positions de sources et de récepteurs disponibles (Fig. 4.21). La valeur de la racine carrée de la moyenne des carrés des résidus des temps (RMS) est environ de 5.35 s.

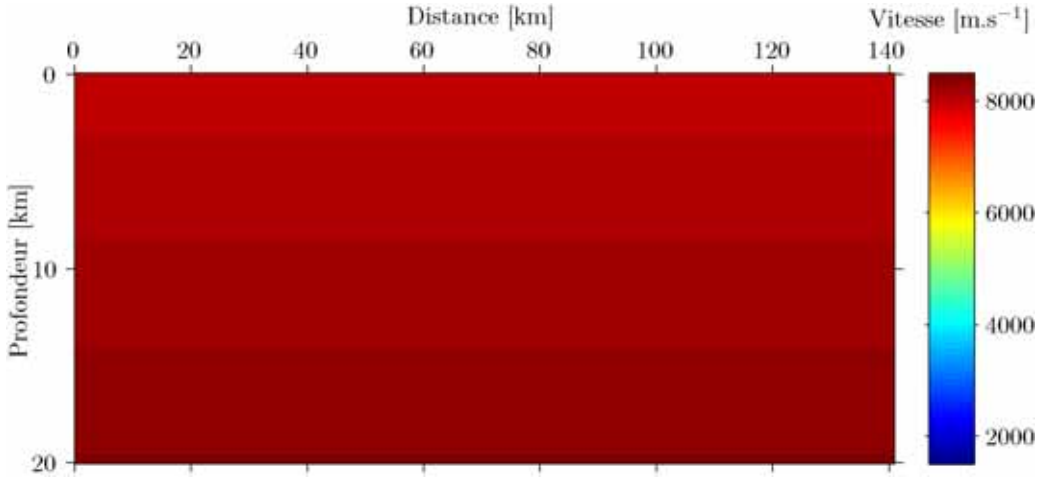

Fig. 4.20 – *Modèle de vitesse initial choisi pour le processus de minimisation, représenté pour une profondeur allant de 0 km à 20 km.*

Le préconditionnement appliqué au gradient de la fonction coût consiste essentiellement en un filtrage gaussien. La longueur du filtrage appliqué au gradient de la fonction coût varie en deux étapes. Pour les premières itérations du schéma d'inversion, la longueur du filtre est de 48 km en longueur et de 22 km en profondeur. Pour les dernières itérations, la longueur du filtre est de 18 km en longueur et de 6 km en profondeur. Pour assurer la stabilité et la convergence de l'algorithme, nous choisissons une valeur maximale du pas appliqué au gradient de la fonction coût égale à 200 m.s^{-1}. Le schéma itératif est arrêté lorsque la valeur du pas calculée à chaque itération est nulle, i.e. lorsque l'algorithme de tomographie ne modifie plus le modèle courant.

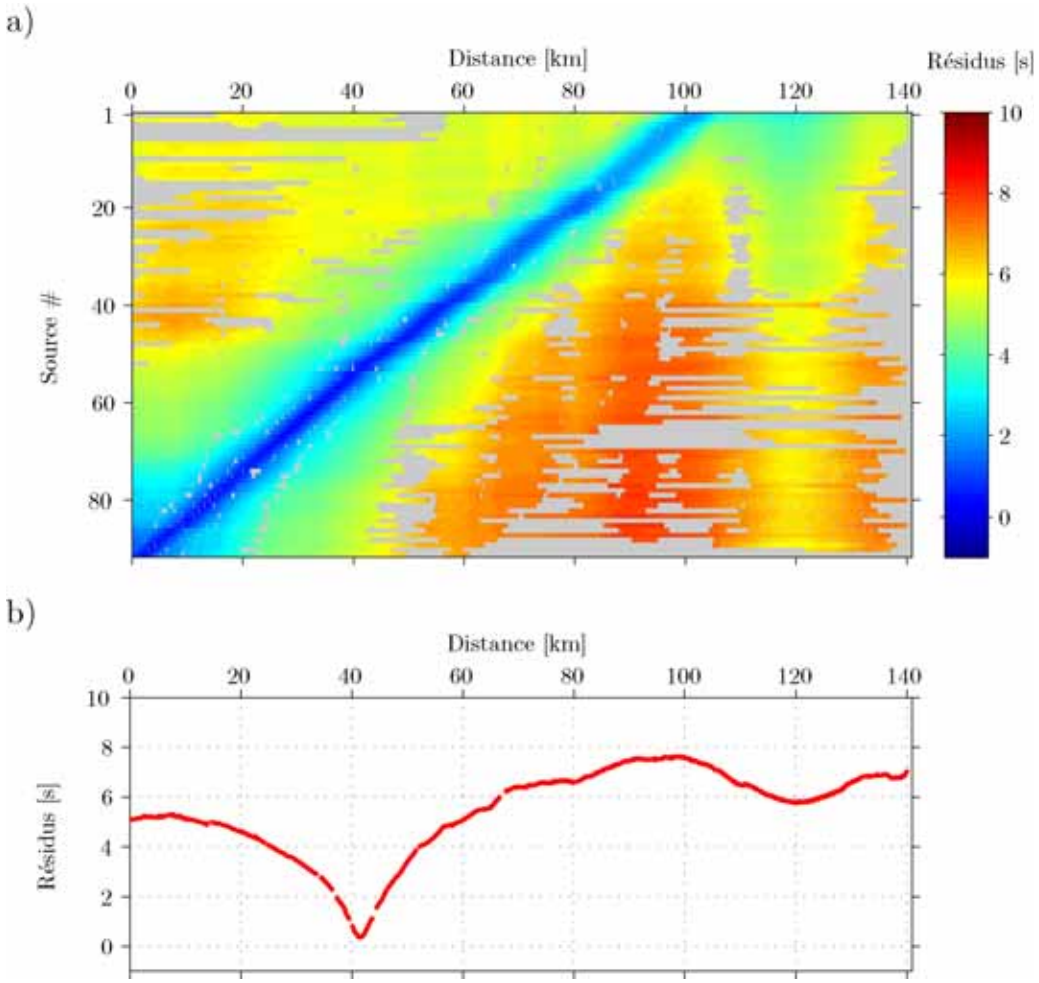

Fig. 4.21 – *a) Carte des résidus initiaux en temps pour toutes les positions de sources et de récepteurs disponibles. Les parties grisées correspondent aux temps de première arrivée qui n'ont pu être pointés avec une précision suffisante. b) Les résidus initiaux en temps pour une source située en (0.75 km, 41.5 km) et tous les récepteurs disponibles.*

4.4.4 Résultats de tomographie

Le modèle de vitesse inversé par le nouvel algorithme de tomographie des temps de première arrivée est obtenu après environ 100 itérations du schéma itératif de plus grande descente. Pour donner un ordre de grandeur, cela représente un temps de calcul d'environ 10 minutes en utilisant 91 cœurs sur le supercalculateur défini à la partie (3.4.3). Un tracé de rais *a posteriori* réalisé à partir de ce modèle inversé (Fig. 4.22) permet d'illustrer la couverture inhomogène des rais, particulièrement entre 0 km et 10 km et entre 100 km et 140 km avec la présence de nombreux rais subverticaux. Cette couverture inhomogène des rais est essentiellement due aux caractéristiques du dispositif d'acquisition. La détermination du modèle de vitesse en utilisant la tomographie des temps de première arrivée est une tâche difficile dans ces zones mal contraintes.

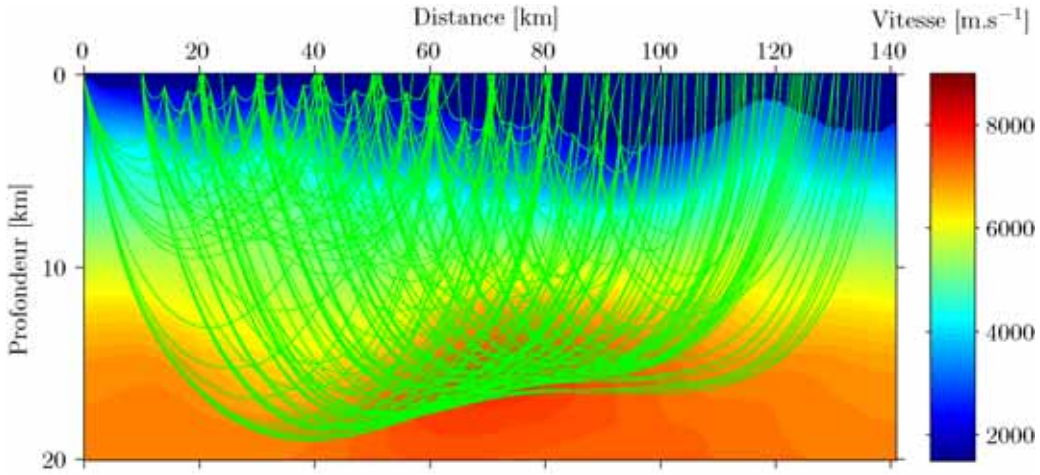

Fig. 4.22 – *Tracé de rais a posteriori réalisé à partir du modèle inversé qui permet d'illustrer la couverture inhomogène des rais, entre entre 0 km et 10 km et entre 100 km et 140 km, avec la présence de nombreux rais subverticaux.*

Nous choisissons alors de présenter le modèle de vitesse obtenu après inversion de 0 km à 105 km (Fig. 4.23 a). La partie grisée correspond à la partie du modèle de vitesse inversé qui n'est pas traversée par les rais. Nous choisissons de valider le modèle de vitesse obtenu par le nouveau schéma d'inversion en le comparant au modèle de vitesse obtenu à partir du même jeu de données par Dessa et al. [2004] (Fig. 4.23 b). Leur algorithme de tomographie des temps de première arrivée est défini à partir d'une méthode de projection. Le modèle de vitesse inversé qu'ils ont obtenu a permis de réaliser une interprétation structurale du segment Est de la fosse de Nankai. Il a aussi été utilisé comme modèle de vitesse initial pour réaliser une inversion de forme d'onde [Operto et al., 2006]. Ce modèle de vitesse inversé nous sert donc de référence pour valider notre résultat. On peut noter que les deux modèles de vitesse possèdent une structure et des valeurs de vitesse similaires.
Ce résultat est confirmé par des profils de vitesse 1-D à différentes positions à travers les modèles (Fig. 4.24). Nous représentons sur ces profils de vitesse le modèle de vitesse initial que nous avons utilisé pour l'inversion et le modèle de vitesse initial utilisé par Dessa et al. [2004]. Ce modèle initial a été défini à partir d'information géologique *a priori*. On peut noter que malgré un modèle de vitesse initial très "éloigné", le nouveau schéma d'inversion parvient à converger vers le modèle de vitesse de référence. Par ailleurs, le modèle inversé a été obtenu sans fixer

préalablement le fond de l'eau.
Les résidus définis entre les temps de première arrivée observés et ceux calculés à partir du modèle inversé (Fig. 4.16) permettent de confirmer que l'algorithme de tomographie a bien expliqué les temps de première arrivée observés. La valeur finale de la racine de la moyenne des carrés des résidus en temps est d'environ 0.16 s, la valeur initiale est de l'ordre de 5.35 s.

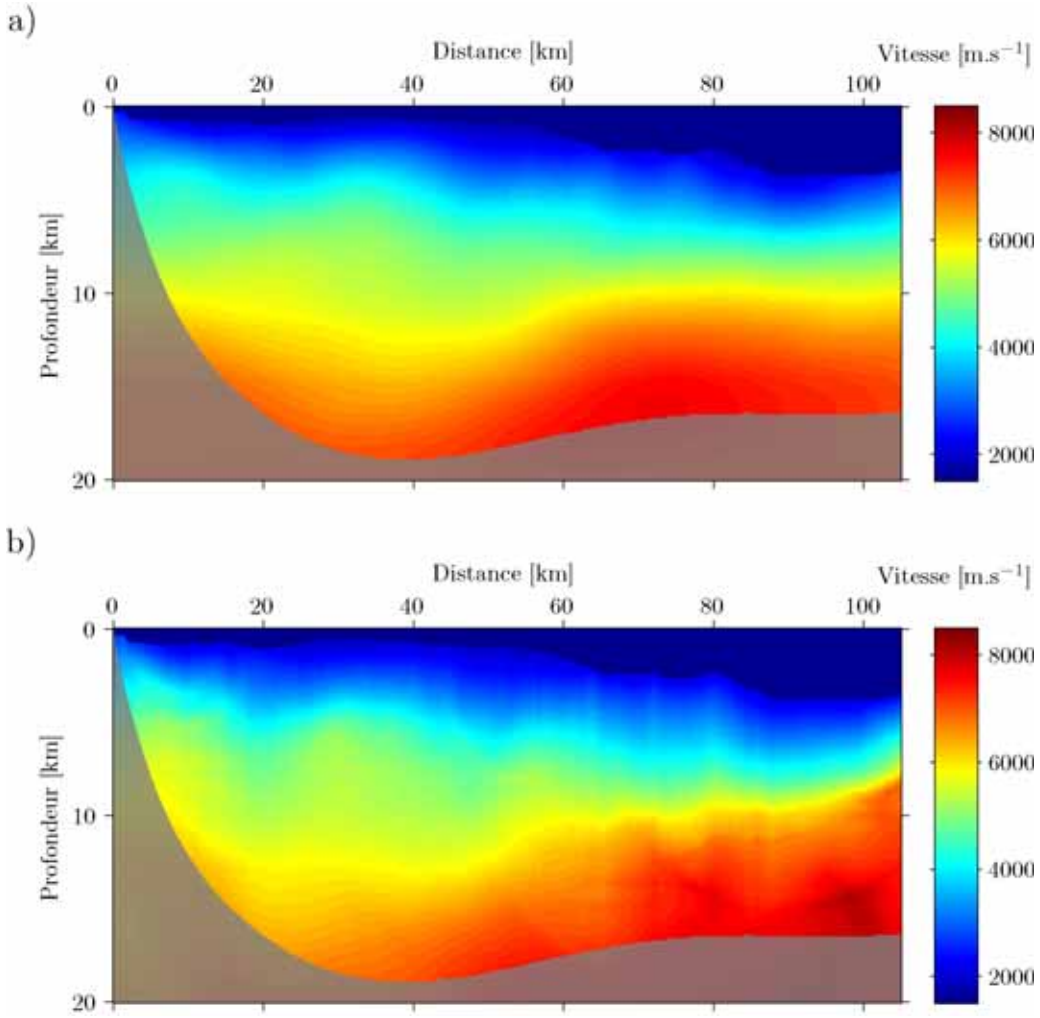

Fig. 4.23 – *a) Modèle de vitesse inversé obtenu en utilisant le nouvel algorithme de tomographie des temps de première arrivée, b) modèle de vitesse inversé obtenu par Dessa et al. [2004].*

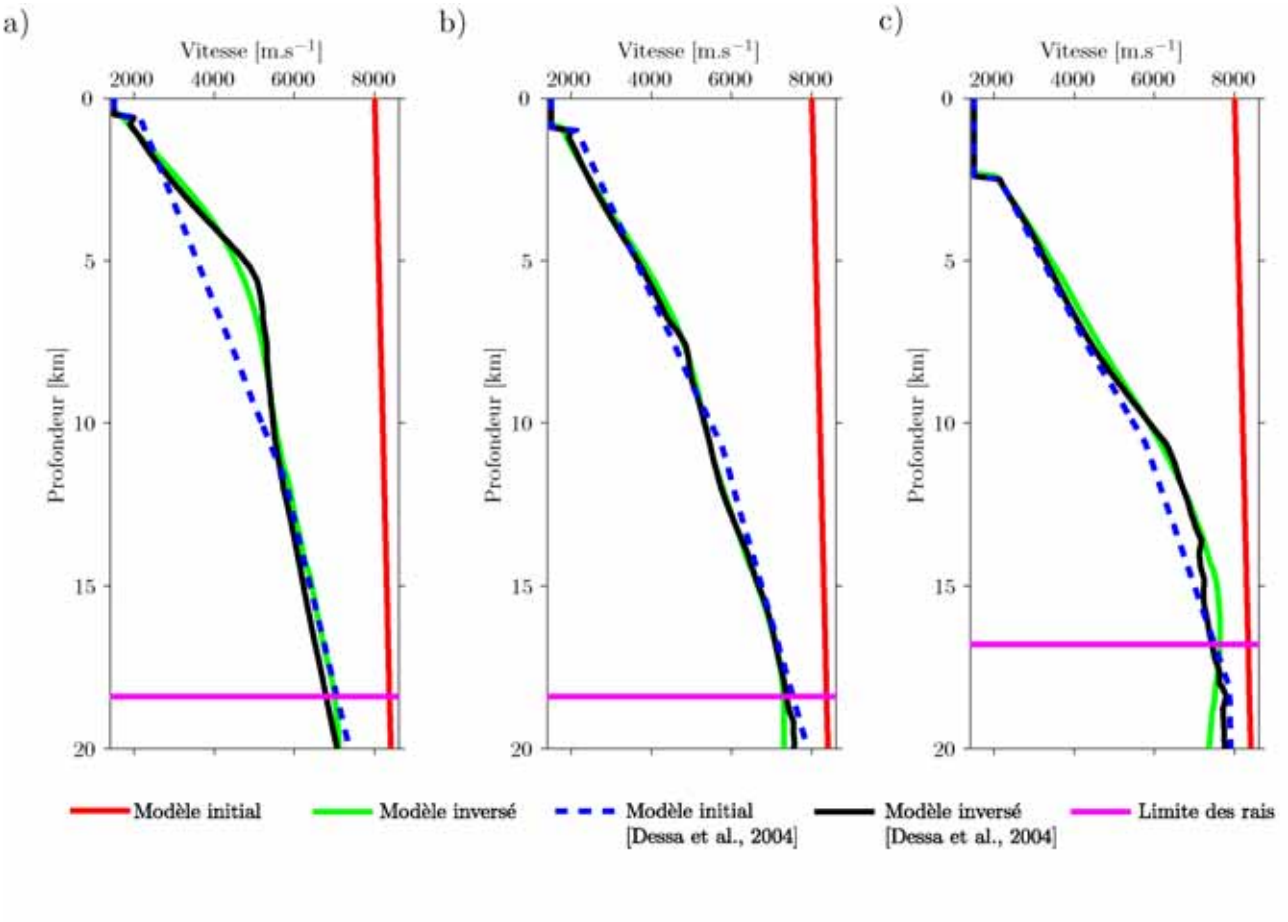

Fig. 4.24 – *Profils de vitesse 1-D à différentes positions dans les modèles de vitesse a) 30 km, b) 50 km, c) 70 km. La courbe rouge correspond au modèle initial, la courbe verte au modèle inversé, la courbe bleue au modèle initial utilisé par Dessa et al. [2004], la courbe noire à leur modèle inversé, et la courbe rose à la limite de la couverture des rais.*

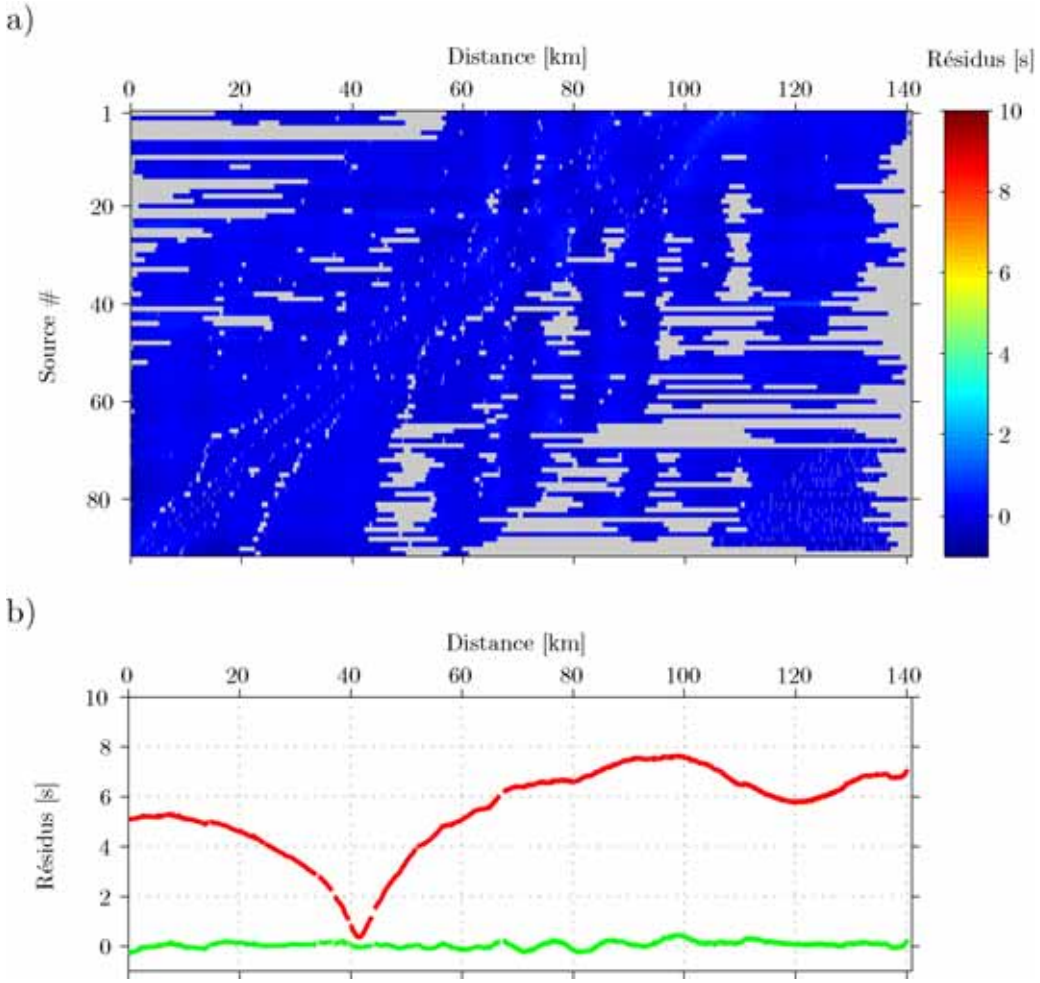

Fig. 4.25 – *a) Carte des résidus finaux en temps pour toutes les positions de sources et de récepteurs disponibles. Les parties grisées correspondent aux temps de première arrivée qui n'ont pu être pointés avec une précision suffisante. b) Comparaison des résidus initiaux (en rouge) et des résidus finaux (en vert) en temps pour une source située en (0.75 km, 41.5 km) et tous les récepteurs disponibles.*

Chapitre 5

Conclusions et perspectives

Conclusions

Dans le cadre d'une approche pragmatique pour la résolution du problème tomographique, nous avons étudié les propriétés pratiques et le comportement d'un algorithme de tomographie des temps de première arrivée défini à partir de la méthode de plus grande descente. Cette méthode est préconisée pour la résolution des problèmes inverses de grande taille [Tarantola, 1987a] car elle permet de s'affranchir de la résolution du système linéaire associé à la méthode de Gauss-Newton classiquement utilisée en tomographie.

L'étape clé dans la mise en œuvre d'une méthode de plus grande descente réside dans le calcul du gradient de la fonction coût des moindres carrés par rapport aux paramètres du modèle. Nous avons utilisé deux méthodes pour calculer ce gradient : la méthode de l'état adjoint [Sei and Symes, 1995], [Leung and Qian, 2006] et une méthode définie à partir d'un tracé de rais *a posteriori*. Ces deux méthodes se distinguent par leur formulation, respectivement non-linéaire et linéarisée, et par leur mise en œuvre pratique. A partir de la carte des temps de première arrivée obtenue par la résolution numérique de l'équation eikonale [Podvin and Lecomte, 1991], le calcul du gradient par la méthode de l'état adjoint repose sur une autre équation aux dérivées partielles résolue par différences finies en utilisant la *Fast Sweeping Method* [Zhao, 2005]. Le calcul du gradient linéarisé lui, repose sur un tracé de rais *a posteriori* [Baina, 1998] entre la source et les récepteurs réalisé à partir de la carte des temps de première arrivée. Ces deux méthodes fournissent des résultats équivalents que nous avons validés en les comparant au gradient de la fonction coût calculé par différences finies.

Des tests de performances réalisés sur un supercalculateur nous ont permis d'établir les propriétés pratiques du nouvel algorithme de tomographie des temps de première arrivée défini à partir de la méthode de plus grande descente :

– le nouvel algorithme est dit "embarrassingly parallel", c'est-à-dire que grâce à sa formulation l'algorithme est directement parallélisable. En effet, tous les calculs sont réalisés par point de tir, ils peuvent être alors distribués sur plusieurs cœurs. Il est ainsi possible de réduire significativement le temps de calcul. Par exemple, le gain en temps obtenu en utilisant 512 cœurs est d'environ 450 et la valeur d'efficacité est de 0.9. Cela se traduit par exemple par un temps d'exécution, pour une itération de plus grande descente, de 8 heures en utilisant 1 cœur et de 1 minute environ en utilisant 512 cœurs,

– l'occupation mémoire de l'algorithme est indépendante du nombre de données observées, elle dépend uniquement de la taille du modèle de vitesse considéré. Cette propriété permet de pouvoir prendre en compte toutes les données observées et de définir des modèles de vitesse avec une grille fine,
– ces tests nous ont aussi permis de mettre en évidence les bonnes propriétés pratiques de l'algorithme de tomographie défini à partir du calcul du gradient de la fonction coût par tracé de rais *a posteriori*. En effet, en comparaison avec l'algorithme défini à partir de la méthode de l'état adjoint, il nécessite une occupation mémoire moindre pour un temps de calcul équivalent et il se révèle être plus simple à mettre en œuvre, notamment pour la prise en compte de la topographie.

La paramétrisation du nouvel algorithme de tomographie est clairement définie et dépend directement du schéma itératif de minimisation de la fonction coût des moindres carrés : le modèle de vitesse initial, le préconditionnement et la valeur maximale du pas appliqués au gradient de la fonction coût à chaque itération et le critère d'arrêt. Le savoir-faire tomographique consiste à savoir déterminer la paramétrisation du schéma d'inversion pour obtenir le meilleur résultat possible. Pour acquérir notre propre savoir-faire tomographique, nous avons réalisé de nombreuses simulations pour des acquisitions de type sismique réfraction, 2-D et 3-D, synthétiques et réelles, marines et terrestres. La réalisation d'un grand nombre de simulations a été rendue possible par la rapidité d'exécution de l'algorithme, de l'ordre de quelques minutes en 2-D. Nous avons pu ainsi définir une paramétrisation type pour les applications que nous avons réalisées :
– le modèle de vitesse initial est construit de façon à être "plus rapide" que le modèle recherché, c'est à dire que les résidus initiaux des temps de première arrivée ($\mathbf{t}_{obs} - \mathbf{t}$) sont positifs,
– le préconditionnement appliqué au gradient de la fonction coût consiste essentiellement en un filtrage gaussien. La longueur du filtrage appliqué au gradient varie en deux étapes : pour les premières itérations du schéma d'inversion, le gradient est fortement lissé, de façon à répartir globalement la perturbation de vitesse appliquée au modèle. Pour les dernières itérations, le filtrage est relâché, de façon à pouvoir déterminer des perturbations locales du modèle de vitesse
– pour assurer la stabilité et la convergence de l'algorithme, nous choisissons une valeur maximale du pas appliqué au gradient de la fonction coût égale à 200 $\mathrm{m.s^{-1}}$,
– le schéma itératif est arrêté lorsque la valeur du pas calculée à chaque itération est nulle, c'est-à-dire lorsque l'algorithme de tomographie ne modifie plus le modèle courant.

Après avoir évalué ses performances informatiques, nous avons évalué le comportement du nouvel algorithme et les résultats obtenus par tomographie. Dans un premier temps, nous avons appliqué le nouveau schéma d'inversion sur des jeux de données en sismique réfraction définis à partir de modèles de vitesse synthétiques en 2-D, BP EAGE [Billette and Brandsberg-Dahl, 2005], et 3-D, Overthrust [Aminzadeh et al., 1997]. Le nombre de temps de première arrivée observés et de mailles des modèles de vitesse considérés sont de plusieurs millions. Les modèles de vitesse recherchés possèdent des structures complexes, des couvertures inhomogènes de rais et des inversions de vitesse. Les résultats de tomographie obtenus en appliquant le nouvel algorithme et la paramétrisation que nous avons définie sont très satisfaisants. Les modèles de vitesse obtenus après inversion sont proches des modèles de vitesse recherchés si la couverture des rais est suffisante et les temps de première arrivée observés sont bien expliqués par le modèle de vitesse inversé. Nous avons ensuite appliqué

le nouvel algorithme de tomographie sur un jeu de données réelles et une acquisition de type *Ocean Bottom Seismometer* réalisée pour étudier la fosse de Nankai [Dessa et al., 2004], [Operto et al., 2006]. Le modèle de vitesse obtenu par notre schéma d'inversion est très proche de celui obtenu à partir du même jeu de données par Dessa et al. [2004] en utilisant un algorithme classique de tomographie des temps de première arrivée. Cette application nous a permis de mettre en évidence le fait que l'algorithme et la méthodologie mis en œuvre permettent d'utiliser un modèle de vitesse initial défini sans aucune connaissance *a priori* géologique. D'une manière générale, les applications présentées dans cette partie valident le bon comportement tomographique de l'algorithme, en termes de résultats obtenus et de stabilité.

Le travail effectué au cours de cette thèse nous a permis de définir, mettre en œuvre et valider un nouvel algorithme de tomographie des temps de première arrivée formulé à partir de la méthode de plus grande descente. Les propriétés pratiques de cet algorithme, en termes de parallélisation, d'occupation mémoire et de simplicité de mise en œuvre, en font un outil adapté pour une résolution pragmatique du problème tomographique. Pour le traitement d'acquisitions de grande dimension, il reste cependant indispensable de mettre en œuvre de puissants moyens de calcul pour permettre une résolution interactive.

Perspectives

Dans le cadre de cette thèse, nous avons appliqué le nouvel algorithme de tomographie des temps de première arrivée à des acquisitions de type sismique réfraction. Toutefois, de par sa formulation et sa mise en œuvre générales, il peut être également utilisé pour d'autres formes d'acquisition telles que entre-puits ou bien puits-surface. Pour les mêmes raisons il s'applique aussi bien aux ondes de compression (ondes P), comme illustré dans cette thèse, qu'aux ondes de cisaillement (ondes S), par exemple pour des études sur la microsismicité [Vanorio et al., 2005].

Nous présentons trois voies de développement possibles du travail réalisé, elles concernent la formulation de l'algorithme mais elles ne nécessitent pas d'en changer la structure générale :

- l'introduction d'information *a priori* dans la définition de la fonction coût des moindres carrés à minimiser sous la forme de matrices de covariance *a priori* sur les espaces des données et des modèles. La prise en compte des incertitudes sur les temps de première arrivée et sur les modèles de vitesse permet de mieux contraindre et stabiliser le résultat de tomographie obtenu [Tarantola, 1987a],
- la définition d'une fonction coût fondée sur des statistiques robustes qui permettent de diminuer l'influence des valeurs aberrantes dans les données observées sur les résultats obtenus [Crase et al., 1990],
- l'utilisation pour la minimisation de la fonction coût d'une méthode de gradient conjugué ou d'une méhode de plus grande descente préconditionnée. Cela pourrait permettre de réduire le nombre d'itérations nécessaires au schéma d'inversion pour converger et ainsi diminuer le temps de calcul.

Pour finir, nous attirons l'attention sur le fait que le savoir-faire du géophysicien s'étend au delà de la résolution du problème tomographique. Il concerne aussi le pointé des temps de première

arrivée sur les sismogrammes et la validation des résultats obtenus après inversion. En effet, l'introduction d'erreurs de mesure dans les valeurs des temps de première arrivée observés affecte directement la qualité du modèle de vitesse inversé. Des traitements appliqués aux données et des corrections faites sur les temps pointés permettent de limiter ces erreurs de mesure. Dans le cadre de ce travail, nous avons considéré les statistiques des incertitudes sur les mesures des temps de première arrivée gaussiennes et indépendantes. Par ailleurs, la détermination de la qualité d'un modèle de vitesse et de sa profondeur de résolution est une tâche délicate. Un modèle de vitesse obtenu après inversion peut être validé grâce aux connaissances *a priori* du milieu étudié, expérience du géologue ou forage d'un puits. Si ce type d'information n'est pas disponible, de nombreuses méthodes ont été développées pour aider à la validation des résultats. Parmi les plus utilisées, on peut citer le test à damier, *checkboard test*, [Zelt and Barton, 1998] et le calcul des matrices de résolution [Berryman, 1994]. En pratique, les résultats obtenus par ces méthodes restent soumis à l'interprétation du géophysicien.

Annexe A

La *Fast Sweeping method*

Nous nous intéressons ici la formulation mathématique et à la mise en œuvre pratique de la *Fast Sweeping Method* [Zhao, 2005] appliquée à la résolution de l'équation de l'état adjoint [Leung and Qian, 2006] définie sur le domaine Ω par

$$\nabla \cdot \big(\lambda \, \nabla t\big) = 0 \,, \tag{A.1}$$

et avec pour condition initiale à la surface Ω_S

$$\lambda \, \big(\mathbf{n} \cdot \nabla t\big) = t - t_{obs} \,, \tag{A.2}$$

où λ représente la variable de l'état adjoint, t le temps de première arrivée correspondant, t_{obs} le temps de première arrivée observés à la surface Ω_S et $\mathbf{n}$ est le vecteur unitaire normal à la surface d'acquisition.

La formulation mathématique

Afin de simplifier la présentation des calculs, nous développons les expressions mathématiques en 2-D, l'extension en 3-D peut être réalisée directement de manière équivalente. L'équation de l'état adjoint (A.1) est alors reformulée sous la forme

$$\frac{\partial}{\partial \, x} \, (a \, \lambda) + \frac{\partial}{\partial \, z} (b \, \lambda) = 0 \,, \tag{A.3}$$

où $a = \dfrac{\partial \, t(x,z)}{\partial \, x}$ et $b = \dfrac{\partial \, t(x,z)}{\partial \, z}$ représentent les coordonnées du gradient de t. On discrétise le domaine Ω sur une grille rectangulaire avec des mailles de tailles Δx et Δz. On note $\lambda_{j,i}$ la variable de l'état adjoint au point (z_j, x_i) du domaine Ω où j et i représentent respectivement les indices pour les directions z et x. L'équation (A.3) est discrétisée

$$\frac{1}{\Delta x} \left(a_{j,i+\frac{1}{2}} \, \lambda_{j,i+\frac{1}{2}} - a_{j,i-\frac{1}{2}} \, \lambda_{j,i-\frac{1}{2}} \right) + \frac{1}{\Delta z} \left(b_{j+\frac{1}{2},i} \, \lambda_{j+\frac{1}{2},i} - b_{j-\frac{1}{2},i} \, \lambda_{j-\frac{1}{2},i} \right) = 0 \,. \tag{A.4}$$

Les valeurs des fonctions a et b ne sont pas spécifiées au centre des cellules (z_j, x_i) mais aux différentes interfaces $(z_{j\pm\frac{1}{2}}, x_i)$ et $(z_j, x_{i\pm\frac{1}{2}})$, on a

$$a_{j,i-\frac{1}{2}} = \frac{t_{j,i} - t_{j,i-1}}{\Delta x} \quad , \quad a_{j,i+\frac{1}{2}} = \frac{t_{j,i+1} - t_{j,i}}{\Delta x} \,, \tag{A.5}$$

$$b_{j-\frac{1}{2},i} = \frac{t_{j,i} - t_{j-1,i}}{\Delta z} \quad , \quad b_{j+\frac{1}{2},i} = \frac{t_{j+1,i} - t_{j,i}}{\Delta z} \,. \tag{A.6}$$

Les valeurs de la variable de l'état adjoint λ aux interfaces, $\lambda_{j,i\pm\frac{1}{2}}$ et $\lambda_{j\pm\frac{1}{2},i}$ sont déterminées suivant la direction de propagation des ondes. Par exemple, dans le cas où $a_{j,i+\frac{1}{2}} > 0$, c'est-à-dire lorsque le temps augmente de la gauche vers la droite, on choisit $\lambda_{j,i+\frac{1}{2}} = \lambda_{j,i}$; dans le cas contraire, on considère $\lambda_{j,i+\frac{1}{2}} = \lambda_{j,i+1}$. Les termes $\lambda_{j,i-\frac{1}{2}}$ et $\lambda_{j\pm\frac{1}{2},i}$ sont définis de la même manière.
On introduit les notations suivantes

$$a^{\pm}_{j,i+\frac{1}{2}} = \frac{a_{j,i+\frac{1}{2}} \pm \left| a_{j,i+\frac{1}{2}} \right|}{2}\,, \tag{A.7}$$

$$a^{\pm}_{j,i-\frac{1}{2}} = \frac{a_{j,i-\frac{1}{2}} \pm \left| a_{j,i-\frac{1}{2}} \right|}{2}\,, \tag{A.8}$$

$$b^{\pm}_{j+\frac{1}{2},i} = \frac{b_{j+\frac{1}{2},i} \pm \left| b_{j+\frac{1}{2},i} \right|}{2}\,, \tag{A.9}$$

$$b^{\pm}_{j-\frac{1}{2},i} = \frac{b_{j-\frac{1}{2},i} \pm \left| b_{j-\frac{1}{2},i} \right|}{2}\,. \tag{A.10}$$

L'équation (A.4) s'écrit alors

$$\begin{aligned}\frac{1}{\Delta x}\left[\left(a^{-}_{j,i+\frac{1}{2}}\,\lambda_{j,i} + a^{+}_{j,i+\frac{1}{2}}\,\lambda_{j,i+1} \right) - \left(a^{-}_{j,i-\frac{1}{2}}\,\lambda_{j,i-1} + a^{+}_{j,i-\frac{1}{2}}\,\lambda_{j,i} \right) \right] \\ + \frac{1}{\Delta z}\left[\left(b^{-}_{j+\frac{1}{2},i}\,\lambda_{j,i} + b^{+}_{j+\frac{1}{2},i}\,\lambda_{j+1,i} \right) - \left(b^{-}_{j-\frac{1}{2},i}\,\lambda_{j-1,i} + b^{+}_{j-\frac{1}{2},i}\,\lambda_{j,i} \right) \right] = 0\,,\end{aligned} \tag{A.11}$$

et elle peut être mise sous la forme

$$\begin{aligned}\left(\frac{a^{-}_{j,i+\frac{1}{2}} - a^{+}_{j,i-\frac{1}{2}}}{\Delta x} + \frac{b^{-}_{j+\frac{1}{2},i} - b^{+}_{j-\frac{1}{2},i}}{\Delta z} \right) \lambda_{j,i} = \\ \frac{a^{-}_{j,i-\frac{1}{2}}\lambda_{j,i-1} - a^{+}_{j,i+\frac{1}{2}}\lambda_{j,i+1}}{\Delta x} + \frac{b^{-}_{j-\frac{1}{2},i}\lambda_{j-1,i} - b^{+}_{j+\frac{1}{2},i}\lambda_{j+1,i}}{\Delta z}\,.\end{aligned} \tag{A.12}$$

L'équation (A.12) permet d'exprimer $\lambda_{j,i}$ en fonction de ses plus proches voisins $\lambda_{j,i\pm1}$ et $\lambda_{j\pm1,i}$, elle est utilisée pour construire le schéma itératif de la *Fast Sweeping Method*.

L'algorithme

1. A partir de la carte des temps de première arrivée obtenue par la résolution numérique de l'équation eikonale [Podvin and Lecomte, 1991], on calcule $\left(\mathbf{n} \cdot \nabla t\right)$ qui permet de déterminer les valeurs initiales de $\lambda^{(0)}$ à la surface Ω_S. Ces valeurs sont fixées pour tous les calculs suivants. Les bords du domaine Ω sont initialisés avec une valeur nulle et on attribue une valeur positive élevée à tous les autres points de la grille, dont les valeurs seront mises à jour par le schéma itératif.
2. A l'itération n, pour chaque point de la grille on calcule la solution candidate $\lambda^{cand}_{j,i}$ à partir de la formule (A.12). On définit alors la nouvelle valeur $\lambda^{(n)}_{j,i}$ comme étant la plus petite valeur entre la valeur courante $\lambda^{(n-1)}_{j,i}$ et la valeur candidate $\lambda^{cand}_{j,i}$, i.e. $\lambda^{(n)}_{j,i} = min\left(\lambda^{(n-1)}_{j,i}, \lambda^{cand}_{j,i}\right)$. On balaye alors itérativement le domaine Ω dans 4 ordres successifs :

- $i = 1 : I$ et $j = 1 : J$,
- $i = 1 : I$ et $j = J : 1$,
- $i = I : 1$ et $j = 1 : J$,
- $i = I : 1$ et $j = J : 1$,

où i et j sont respectivement les indices sur les axes x et z.

3. Pour un critère de convergence donné $\epsilon > 0$, on vérifie si $\|\lambda^{(n)} - \lambda^{(n-1)}\| \leq \epsilon$. On réitère le point 2 jusqu'à obtenir la convergence désirée.

Bibliographie

[Aki and Lee, 1976] Aki, K. and Lee, W. H. K. (1976). Determination of 3-dimensional velocity anomalies under a seismic array using 1st-P arrival times from local earthquakes. 1. A homogeneous initial model. *J. Geophys. Res.*, 81(23) :4381–4399.

[Aminzadeh et al., 1997] Aminzadeh, F., Brac, J., and Kunz, T. (1997). *3-D Salt and Overthrust models*. Society of Exploration Geophysicists.

[Arntsen and Carcione, 2000] Arntsen, B. and Carcione, J. (2000). A new insight into the reciprocity principle. *Geophysics*, 65(5) :1604–1612.

[Baina, 1998] Baina, R. (1998). *Tomographie sismique entre puits : mise en œuvre et rôle de l'analyse a posteriori ; vers une prise en compte de la bande passante*. PhD thesis, Université de Rennes I.

[Bennett, 2002] Bennett, A. (2002). *Inverse Modeling of the Ocean and Atmosphere*. Cambridge University Press, Cambridge, UK.

[Berryman, 1994] Berryman, J. G. (1994). Tomographic resolution without singular value decomposition. In *Mathematical Methods in Geophysical Imaging II*, pages 1–13.

[Billette and Brandsberg-Dahl, 2005] Billette, F. and Brandsberg-Dahl, S. (2005). The 2004 BP velocity benchmark. In *Expanded Abstracts, 67th Annual EAGE Meeting*, page B035. Eur. Ass. Geosc. Eng.

[Boué and Dupuis, 1999] Boué, M. and Dupuis, P. (1999). Markov chain approximations for deterministic control problems with affine dynamics and quadratic cost in the control. *SIAM J. Numer. Anal*, 36 :667–695.

[Brenders and Pratt, 2007] Brenders, A. J. and Pratt, R. G. (2007). Efficient waveform tomography for lithospheric imaging : implications for realistic, two-dimensional acquisition geometries and low-frequency data. *Geophys. J. Int.*, 168(1) :152–170.

[Chapman, 2004] Chapman, C. (2004). *Fundamentals of Seismic Wave Propagation*. Cambridge University Press, Cambridge, UK.

[Chauris et al., 2008] Chauris, H., Noble, M., and Taillandier, C. (2008). What initial velocity model do we need for full waveform inversion ? In *Workshop 11 : Full Waveform Inversion - Current Status and Perspectives , 70th Annual EAGE Meeting*. Eur. Ass. Geosc. Eng.

[Chavent, 1974] Chavent, G. (1974). Identification of Functional Parameters in Partial Differential Equations. In Goodson, R. E. and Polis, M., editors, *Identification of Parameters in Distributed Systems*, pages 31–48. ASME, New York.

[Chavent and Jacewitz, 1995] Chavent, G. and Jacewitz, C. (1995). Determination of background velocities by multiple migration fitting. *Geophysics*, 60(02) :476–490.

[Crandall and Lions, 1983] Crandall, M. G. and Lions, P.-L. (1983). Viscosity solutions of Hamilton-Jacobi equations. *Transactions of the American Mathematical Society*, 277(1) :1–42.

[Crase et al., 1990] Crase, E., Pica, A., Noble, M., McDonald, J., and Tarantola, A. (1990). Robust elastic nonlinear waveform inversion : Application to real data. *Geophysics*, 55(5) :527–538.

[Delprat-Jannaud et al., 1992] Delprat-Jannaud, F., Lailly, P., Becache, E., and Chovet, E. (1992). Reflection tomography : How to cope with multiple arrivals ? In *Expanded Abstracts, 62nd Annual SEG Meeting and Exposition*, pages 741–744. Soc. of Expl. Geophys.

[Dessa et al., 2004] Dessa, J.-X., S. Operto, S. K., Nakanishi, A., Pascal, G., Uhira, K., and Kaneda, Y. (2004). Deep seismic imaging of the eastern nankai trough, japan, from multifold ocean bottom seismometer data by combined travel time tomography and prestack depth migration. *J. Geophys. Res.*, 109 :B02111.

[Fomel and Claerbout, 2003] Fomel, S. and Claerbout, J. (2003). Multidimensional recursive filter preconditioning in geophysical estimation problems. *Geophysics*, 68(2) :577–588.

[Grandjean and Sage, 2004] Grandjean, G. and Sage, S. (2004). Jats : a fully portable seismic tomography software based on fresnel wavepaths and a probabilistic reconstruction approach. *Comput. Geosci.*, 30 :925–935.

[Hadamard, 1902] Hadamard, J. (1902). Sur les problèmes aux dérivées partielles et leur signification physique. In *Princeton University Bulletin*, pages 49–52.

[Improta et al., 2002] Improta, L., Zollo, A., Herrero, A., Frattini, R., Virieux, J., and Dell'Aversana, P. (2002). Seismic imaging of complex structures by non-linear traveltime inversion of dense wide-angle data : application to a thrust belt. *Geophys. J. Int.*, 151(1) :264 – 278.

[Kuster, 2006] Kuster, C. M. (2006). *Fast Numerical Methods for Evolving Interfaces*. PhD thesis, North Carolina State University.

[Lailly, 1984] Lailly, P. (1984). The seismic inverse problem as a sequence of before stack migrations. In Bednar, R. and Weglein, editors, *Conference on Inverse Scattering, SIAM, Philadelphia*. SIAM.

[Le Meur, 1994] Le Meur, H. (1994). *Tomographie tridimensionnelle à partir des temps des premières arrivées des ondes P et S*. PhD thesis, Université Paris VII.

[Lecomte, 1993] Lecomte, I. (1993). Finite difference calculation of first traveltimes in anisotropic media. *Geophys. J. Int.*, 113 :318–342.

[Leung and Qian, 2006] Leung, S. and Qian, J. (2006). An adjoint state method for three-dimensional transmission traveltime tomography using first arrivals. *Geophys. J. Int.*, 4 :249–266.

[Lions, 1971] Lions, J. L. (1971). *Optimal control systems governed by partial differential equations*. Springer Verlag, Berlin.

[Menahem and Beydoun, 1985] Menahem, A. B. and Beydoun, W. (1985). Range of validity of seismic ray and beam methods in general inhomogeneous media. 1: General theory. *Geophys. J. Int.*, 82 :207–234.

[Noble, 1992] Noble, M. S. (1992). *Inversion non linéaire de données de prospection pétrolière*. PhD thesis, Université Paris 7.

[Nolet, 1985] Nolet, G. (1985). Solving or resolving inadequate and noisy tomographic systems. *J. Comput. Phys.*, 61 :463–482.

[Nolet, 1987] Nolet, G. (1987). Waveform tomography. In Nolet, G., editor, *Seismic Tomography*, pages 301–322. Reidel, Dordrecht, The Netherlands.

[Operto et al., 2006] Operto, S., Virieux, J., Dessa, J.-X., and Pascal, G. (2006). Crustal seismic imaging from multifold ocean bottom seismometer data by frequency domain full waveform tomography : Application to the eastern nankai trough. *J. Geophys. Res.*, 111 :B09306.

[Paige and Saunders, 1982] Paige, C. and Saunders, M. A. (1982). LSQR : Sparse linear equations and least squares problems, Part I and Part II. *ACM Trans. math. Soft.*, 8 :43–71.

[Pica, 1997] Pica, A. (1997). Fast and accurate finite-difference solutions of the 3-D eikonal equation parameterized in celerity. In *Expanded Abstracts, 67th Annual SEG Meeting and Exposition*, pages 1774–1777. Soc. Expl. Geophys.

[Plessix, 2006] Plessix, R.-E. (2006). A review of the adjoint-state method for computing the gradient of a functional with geophysical applications. *Geophys. J. Int.*, 167 :495–503.

[Plessix et al., 1999] Plessix, R.-E., De Roeck, Y.-H., and Chavent, G. (1999). Waveform inversion of reflection seismic data for kinematic parameters by local optimization. *SIAM J. Sci. Comput.*, 20(3) :1033–1052.

[Podvin and Lecomte, 1991] Podvin, P. and Lecomte, I. (1991). Finite difference computation of traveltimes in very contrasted velocity model : a massively parallel approach and its associated tools. *Geophys. J. Int.*, 105 :271–284.

[Popovici and Sethian, 1998] Popovici, M. and Sethian, J. (1998). Three dimensional traveltimes using the fast marching method. In *Expanded Abstracts, 60th Annual EAGE Meeting*, pages 1–22. Eur. Ass. Geosc. Eng.

[Qian et al., 2007] Qian, J., Zhang, Y., and Zhao, H. (2007). A fast sweeping method for static convex hamilton-jacobi equations. *J. Sci. Comput.*, 31(1-2) :237–271.

[Qin et al., 1992] Qin, F., Luo, Y., Olsen, K. B., Cai, W., and Schuster, G. T. (1992). Finite-difference solution of the eikonal equation along expanding wavefronts. *Geophysics*, 57(03) :478–487.

[Schuster and Quintus-Bosz, 1993] Schuster, G. T. and Quintus-Bosz, A. (1993). Wavepath eikonal traveltime inversion : Theory. *Geophysics*, 58(9) :1314–1323.

[Sei and Symes, 1994] Sei, A. and Symes, W. W. (1994). Gradient calculation of the traveltime cost function without ray tracing. In *Expanded Abstracts, 65th Annual SEG Meeting and Exposition*, pages 1351–1354. Soc. Expl. Geophys.

[Sei and Symes, 1995] Sei, A. and Symes, W. W. (1995). A note on consistency and adjointness for numerical schemes. Technical Report TR95-04, Department of Computational and Applied Mathematics, Rice University.

[Sethian, 1996] Sethian, J. A. (1996). *Level Set Methods*. Cambridge University Press.

[Shen et al., 2003] Shen, P., Symes, W. W., and Stolk, C. (2003). Differential semblance velocity analysis by wave-equation migration. In *Expanded Abstracts, 74th Annual SEG Meeting and Exposition*, pages 2132–2135. Soc. Expl. Geophys.

[Simons, 2004] Simons, F. J. (2004). Seismic tomography : art or science ? 1st Meeting of Young Researchers in the Earth Sciences, La Jolla CA.

[Spakman and Nolet, 1988] Spakman, W. and Nolet, G. (1988). *Imaging algorithms, accuracy and resolution in delay time tomography*, pages 155–188. Vlaar, N.J., Nolet, G., Wortel, M.J.R. and S.A.P.L. Cloetingh, Reidel, Dordrecht.

[Taillandier and Noble, 2008] Taillandier, C. and Noble, M. (2008). 2-D and 3-D seismic refraction travel-time tomography based on the adjoint state method. In *Geophysical Research Abstracts, EGU General Assembly*, volume 10, pages A–07485.

[Taillandier et al., 2008a] Taillandier, C., Noble, M., and Calandra, H. (2008a). A massively parallel 3-D refraction traveltime tomography algorithm. In *Workshop 8*: *Computing Trends in O&G and Earth Sciences, 70th Annual EAGE Meeting*. Eur. Ass. Geosc. Eng.

[Taillandier et al., 2008b] Taillandier, C., Noble, M., Calandra, H., Chauris, H., and Podvin, P. (2008b). 3-D refraction traveltime tomography algorithm based on adjoint state techniques. In *Expanded Abstracts, 70th Annual EAGE Meeting*, page P163. Eur. Ass. Geosc. Eng.

[Taillandier et al., 2007a] Taillandier, C., Noble, M., Chauris, H., Podvin, P., and Calandra, H. (2007a). Refraction travel-time tomography based on adjoint state techniques. In *Expanded Abstracts, 7th SIAM Conference on Mathematical and Computational Issues in the Geosciences*, page CP6. SIAM.

[Taillandier et al., 2007b] Taillandier, C., Noble, M., Chauris, H., Podvin, P., Calandra, H., and Guilbot, J. (2007b). Refraction traveltime tomography based on adjoint state techniques. In *Expanded Abstracts, 69th Annual EAGE Meeting*, page P054. Eur. Ass. Geosc. Eng.

[Tarantola, 1984] Tarantola, A. (1984). Inversion of seismic reflection data in the acoustic approximation. *Geophysics*, 49(8) :1259–1266.

[Tarantola, 1986] Tarantola, A. (1986). A strategy for non linear inversion of seismic reflection data. *Geophysics*, 51(10) :1893–1903.

[Tarantola, 1987a] Tarantola, A. (1987a). *Inverse problem theory : methods for data fitting and model parameter estimation*. Elsevier, Netherlands.

[Tarantola, 1987b] Tarantola, A. (1987b). Inversion of travel times and seismic wavefront. In Nolet, G., editor, *Seismic Tomography*, pages 135–158. Reidel, Dordrecht, The Netherlands.

[Tarantola and Valette, 1982] Tarantola, A. and Valette, B. (1982). Generalized nonlinear inverse problems solved using the least square criterion. *Reviews of Geophys. and Space Phys.*, 20 :219–232.

[Tikhonov and Arsenin, 1977] Tikhonov, A. and Arsenin, V. (1977). *Solution of ill-posed problems*. Winston, Washington, DC.

[Červený, 2001] Červený, V. (2001). *Seismic Ray Theory*. Cambridge University Press, Cambridge, UK.

[Červený and Soares, 1992] Červený, V. and Soares, J. (1992). Fresnel volume ray tracing. *Geophysics*, 57(07) :902–915.

[van der Sluis and van der Vorst, 1987] van der Sluis, A. and van der Vorst, H. (1987). Numerical solution of large sparse linear equations and least squares problems. In Nolet, G., editor, *Seismic Tomography*, pages 49–83. Reidel, Dordrecht, The Netherlands.

[van Trier and Symes, 1991] van Trier, J. and Symes, W. (1991). Upwind finite-difference calculation of traveltimes. *Geophysics*, 56(6) :812–821.

[Vanorio et al., 2005] Vanorio, T., Virieux, J., Capuano, P., and Russo, G. (2005). Three-dimensional seismic tomography from p wave and s wave microearthquake travel times and rock physics characterization of the campi flegrei caldera. *J. Geophys. Res.*, 110(B03201).

[Vesnaver, 2008] Vesnaver, A. (2008). Yardsticks for industrial tomography. *Geophysical Prospecting*, 56(4) :457–465.

[Vidale, 1988] Vidale, J. (1988). Finite-difference calculation of travel time. *Bull. seism. Soc .Am.*, 78 :2062–2076.

[Watanabe et al., 1999] Watanabe, T., Matsuoka, T., and Y.Ashida (1999). Seismic traveltime tomography using fresnel volume approach. In *Expanded Abstracts, 69th Annual SEG Meeting and Exposition*, pages 1402–1405. Soc. Expl. Geophys.

[Yao and Roberts, 2002] Yao, Z. S. and Roberts, R. G. (2002). A practical regularization for seismic tomography. *Geophys. J. Int.*, 138(2) :293–299.

[Zelt et al., 2006] Zelt, C. A., Azaria, A., and Levander, A. (2006). 3d seismic refraction traveltime tomography at a groundwater contamination site. *Geophysics*, 71(5) :H67–H78.

[Zelt and Barton, 1998] Zelt, C. A. and Barton, P. J. (1998). Three-dimensional seismic refraction tomography : A comparison of two methods applied to data from the Faeroe Basin. *J. Geophys. Res.*, 103(B4) :7187–7210.

[Zhao, 2005] Zhao, H. (2005). A fast sweeping method for eikonal equations. *Math. Comp.*, 74 :603–627.

[Zhao, 2006] Zhao, H. (2006). Parallel implementations of the fast sweeping method. *UCLA CAM06-13*.

[Zhu et al., 1992] Zhu, X., Sixta, D. P., and Angstman, B. G. (1992). Tomostatics : turning-ray tomography + static corrections. *The Leading Edge*, 11 :15–23.

Table des figures