UNIVERSITE DE NICE-SOPHIA ANTIPOLIS - UFR SCIENCES
Ecole Doctorale Sciences Fondamentales et Appliquées

THESE

pour obtenir le titre de

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

Spécialité : Sciences de l'Univers - Géophysique

présentée et soutenue par

Yaser GHOLAMI

Two-dimensional seismic imaging of anisotropic media by full waveform inversion

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

soutenue le 8 octobre 2012 devant le Jury

Michel DIETRICH	Directeur de Recherche CNRS	Examinateur
Gilles LAMBARÉ	Chercheur Sénior CGGVeritas	Rapporteur
Guust NOLET	Professeur Université Nice-Sophia Antipolis	Examinateur
Stéphane OPERTO	Chargé de Recherche CNRS	Directeur de Thèse
René-Edouard PLESSIX	Chercheur Sénior SHELL	Rapporteur
Alessandra RIBODETTI	Chargé de Recherche IRD	Directeur de Thèse
Jean VIRIEUX	Professeur Université Joseph Fourier	Examinateur

Acknowledgement

Résumé

L'exploration de la terre solide pour des applications à vocations socio-économiques (recherche d'hydrocarbures, évaluation des risques naturels et anthropiques, surveillance des zones de stockage de déchets) ou scientifiques (compréhension de la géodynamique de la planète Terre) constitue l'essence des méthodes d'imagerie sismique. Dans le domaine de la géophysique moderne, la méthode d'imagerie sismique par inversion des formes d'ondes complètes (FWI) a suscité une forte dynamique de recherche ces dernières années dans les domaines industriels et académiques, motivée par l'accroissement des moyens de calcul et le développement de nouvelles technologies d'acquisition (acquisition 'wide azimuth' multi-composantes, sources large bandes). L'objectif est de construire des modèles quantitatifs haute résolution de la structure du sous-sol, en particulier dans des zones complexes et/ou difficiles d'accès ('offshore' profond), pour lesquelles des approches conventionnelles d'imagerie peuvent se révéler inopérantes. Par essence, la FWI prend en compte toute l' information contenue dans les données sismiques (phase et amplitude de toutes les arrivées sismiques) pour imager les paramètres physiques qui gouvernent la propagation des ondes élastiques dans le sous-sol. Cela nécessite de modéliser l'ensemble des phénomènes de propagation d'ondes par résolution complète de l'équation d'ondes, cette tâche constituant le problème direct. Le problème inverse non linéaire est posé sous forme d'un problème d'optimisation local, visant à minimiser itérativement l'écart entre les données enregistrées et modélisées pour reconstruire les paramètres décrivant les propriétés physiques du milieu. Deux difficultés sont à surmonter: le cout numérique du problème direct pour des acquisitions sismiques impliquant plusieurs milliers de sources et la non linéarité du problème inverse illustrée par la présence de nombreux minimums locaux dans la fonction cout.

L'anisotropie décrit de manière générale les variations de la vitesse de propagation des ondes en fonction de la direction de propagation. L'anisotropie a de ce fait une influence évidente sur les temps de trajet des ondes mais aussi sur les amplitudes et les formes d'ondes. Une prise en compte incomplète des phénomènes liés à l'anisotropie peut de ce fait générer un positionnement incorrect et une dé-focalisation des réflecteurs dans les images sismiques. L'objectif de cette thèse est de prendre en compte l'anisotropie à symétrie transverse isotrope (VTI) dans la modélisation des ondes sismiques et dans la FWI. Cette forme d'anisotropie est la plus répandue dans les structures géologiques qui sont par nature stratifiées.

Le processus d'imagerie est entièrement formulé dans le domaine fréquentiel. La modélisation des ondes est effectuée avec la méthode des éléments finis dite de Galerkin Discontinu sur maillage triangulaire non structuré pour des milieux transverses isotropes présentant un axe de symétrie d'inclinaison arbitraire. Le problème inverse repose sur une approche locale d'optimisation où le gradient de la fonction cout est calculée avec la méthode de l'état adjoint, et les effets du Hessien sont pris en compte avec la méthode de quasi-Newton L-BFGS.

Les paramètres décrivant les milieux VTI peut être représentés par différentes combinaisons de paramètres: par exemple, les modules élastiques ou une ou plusieurs vitesses de propagation associées avec un ou plusieurs paramètres adimensionels de Thomsen (δ, ϵ, η). Dans ce contexte d'imagerie multi-paramètres non linéaire, le choix de la paramétrisation du sous-sol est de première importance pour sélectionner les paramètres ayant une influence suffisamment forte et découplée dans les données. Une analyse de sensibilité fondée sur l'analyse des diagrammes de rayonnement des paramètres anisotropes est proposée pour sélectionner la paramétrisation la plus adaptée. Deux paramétrisations sont tout d'abord proposées dans le cadre de l'approximation acoustique. La première combine une vitesse de propagation (par exemple, la vitesse verticale ou horizontale) avec les deux paramètres de Thomsen δ et ϵ. La vitesse de propagation a une influence dominante et uniforme sur toute la bande d'ouvertures angulaires. Pour cette paramétrisation, une inversion mono-paramètre permet de reconstruire la vitesse de propagation avec une bonne résolution, sous réserve que les modèles de référence des paramètres de Thomsen décrivent suffisamment précisément les grandes longueurs d'onde du milieu. Alternativement, deux vitesses de propagation (par exemple, les vitesses verticales et horizontales) peuvent être mis-à-jour conjointement, tandis que le paramètre de δ est maintenu fixe. Si des dispositifs d'acquisition dits grand-angle sont disponibles, les grandes longueurs d'onde de la vitesse horizontale sont alors reconstruites à partir des composantes grand-angles des données, tandis que les courtes longueurs d'onde de la vitesse verticale sont reconstruites à partie des angles d'ouvertures faibles. Cette approche est justifiée sous réserve que le modèle initial de la vitesse verticale décrit suffisamment précisément les grandes longueurs d'ondes du milieu, qui ne peuvent pas être mis à jour par FWI avec cette paramétrisation. Ces conclusions, qui révèlent l'importance de la géométrie des dispositifs d'acquisition et de l'information apriori contenue dans les modèles initiaux, s'étendent aux milieux élastiques, pour lesquels un paramètre supplémentaire, la vitesse de propagation des ondes S le long de l'axe de symétrie, doit être pris en compte.

Après avoir validé ces conclusions théoriques avec des tests d'inversion sur des modèles synthétiques simples, j'applique la FWI à des cas d'études synthétique et réel, représentatifs du champ pétrolier de Valhall à la fois dans le cadre des approximations visco-acoustiques et visco-élastiques. Dans le cas de l'inversion visco-élastique, je montre la faisabilité d'un protocole d'inversion hiérarchique, où j'image dans un premier temps les vitesses verticales de propagation des ondes P et S à partir des géophones en m'appuyant sur la paramétrisation de type 1, avant de mettre à jour conjointement les vitesses verticales et horizontales des ondes P à partir de la paramétrisation de type 2.

Dans ma thèse, je me suis attaché à comprendre la sensibilité des données sismiques aux paramètres anisotropes. Cette analyse, à partir de laquelle différentes paramétrisations anisotropes du sous-sol peuvent être proposées, doit s'accompagner d'une utilisation pertinente des méthodes d'optimisation numérique et des techniques de régularisation pour pouvoir reconstruire de manière robuste différentes classes de paramètres pouvant avoir une influence contrastée dans les données, ainsi que des unités et des ordres de grandeurs différents.

Abstract

Exploring the solid Earth for hydrocarbons, as social needs, is one of the main tasks of seismic imaging. As a domain of the modern geophysics, the seismic imaging by full waveform inversion (FWI) aims to improve and refine imaging of shallow and deep structures. Theoretically, the FWI method takes into account all the data gathered from subsurface in order to extract information about the physical parameter of the Earth. The kernel of the FWI is the full wave equation, which is considered in the heart of forward modeling engine. The FWI problem is represented as a least-squares local optimisation problem that retrieves the quantitative values of subsurface physical parameters.

The seismic images are affected by the manifested anisotropy in the seismic data as anomalies in travel time, amplitude and waveform. In order to circumvent mis-focusing and mis-positioning events in seismic imaging and to obtain accurate model parameters, as valuable lithology indicators, the anisotropy needs to be integrated in propagation/inversion work flows. In this context, the aim of this work is to develop two dimensional FWI for vertically transverse isotropic media (VTI). The physical parameters describing the Earth are elastic moduli or wavespeeds and Thomsen parameter(s). The forward modeling and the inversion are performed entirely in frequency domain. The frequency-domain anisotropic P-SV waves propagation modelling is discretized by the finite element discontinuous Galerkin method. The full waveform modeling (FWM) is performed for VTI and tilted transverse isotropic (TTI) media by various synthetic examples. The gradient of the misfit function is computed by adjoint-state method. The linearized inverse problem is solved with the quasi-Newton l-BFGS algorithm, which is able to compute an estimated Hessian matrix from a preconditionner and few gradients of previous iterations.

Three categories of parameterization type are proposed in order to parametrize the model space of the inverse problem. The sensitivity analysis on acoustic VTI FWI method is preformed by studying the partial derivative of pressure wavefield and the grid analysis of least-squares misfit functional. The conclusions inferred from the sensitivity analysis are verified by FWI experiment on a simple synthetic model. The anisotropic parameter classes that can be well retrieved by VTI FWI are recognized. Furthermore, the acoustic VTI FWI is applied on the realistic synthetic Valhall benchmark for a wide-aperture surface acquisition survey. The anisotropic acoustic and elastic FWI are performed on the three components of ocean bottom cable (OBC) data sets from Valhall oil/gas field, that is located in North sea.

Notations

2D: two-dimensional

3D: three-dimensional

4C: four components data (1 hydrophone plus 3 geophones)

API: American petroleum institute

A^t: transpose of matrix A

$\mathcal{C}$: misfit function

CIG: common image gather

c_{ij}: stiffness coefficients

DG : discontinuous Galerkin

δ: Thomsen parameter

ΔC: gradient of C function

ϵ: Thomsen parameter

ϵ_{ij}: strain tensors

η: anisotropy parameter related to δ and ϵ

FD: finite difference

FE: finite element

FWI: full waveform inversion

FWM: full waveform modeling

$\mathcal{G}$: gradient of misfit function

H: Hessian matrix

HTI: horizontal transverse isotropic

λ: regularization factor

$\parallel x \parallel^2$: L_2 norm of x vector in $\Re^n$, (n>1)

LoFS: life of field seismic

mD: 10^{-3} Darcy

ν: Poisson coefficient

OBC: ocean bottom cable

PML: perfectly matched layer

Q_P: quality factor for pressure wavespeed

Q_S: quality factor for shear wavespeed

$\mathbf{R}$: restriction operator

ρ : density

RTM: reverse time migration

σ_{ij}: stress tensors

*: conjugate operator

T_i: pressure wavefields

TTI: tilted transverse isotropic

V_h: horizontal pressure velocity

V_{NMO} : normal move-out pressure velocity

VSP: vertical seismic profile

V_{P_0} : vertical pressure velocity

V_P : isotropic pressure velocity

V_{SV} : vertical shear velocity

V_{SH} : horizontal shear velocity

v_x, v_z: horizontal and vertical component of velocity wavefields

VTI: vertically transverse isotropic

W_d: weighting matrix of data

W_m: weighting matrix of model parameters

Inverse Problem Theory has to be developed from the consideration of uncertainties (either experimental, or in physical laws), and the right (well-posed) equation to set is: given a certain amount of (a priory) information on some model parameters, and given an uncertain physical law relating some observable parameters to the model parameters, in which sense should I modify the a priory information, given the uncertain results of some experiments? In my opinion, this is the only approach allowing us to analyze "error and resolution" in the "solution" with a convenient degree of generality, even for nonlinear forward problems.
Albert Tarantola, 1986.

Contents

Introduction

In the framework of seismic exploration, the *imaged* values of Earth's physical parameters are not usually identical to the *true* values. This is for two reasons: 1) experimental uncertainties and 2) modeling errors. Unfortunately, experimental uncertainties will always exist; thus, we are forced to accept this and treat it based on our own estimate of the *data noise*. Fortunately, the complexity and accuracy of our modeling is something we do have control over. Early on in seismic exploration, the physical model describing the earth was assumed to be *homogeneous* and *isotropic*. This means that for a given point in the model, the physical parameters of the system were constant with respect to all other points in the model. In reality, the Earth is heterogeneous and anisotropic (Boore, 1972; Gung and Romanowicz, 2004; Montelli et al., 2004).

In seismic exploration, the anisotropy is defined to be the dependence of seismic velocity upon angle. It can be said that the anisotropy is the large-scale manifestation of ordered, small-scale heterogeneity (Thomsen, 1986). The assumption of an isotropic earth introduces modeling errors into seismic imaging results. Furthermore, we know this assumption is not valid with today computational resources and we should not restrict ourselves to this assumption. The early plate tectonics studies show that the Earth's crust is affected by continental movement and seafloor spreading. Furthermore, each plate is composed of geological formations, with each formation having its own diverse intrinsic seismic properties. Complicating things further, the Earth's crust contains complex and heterogeneous geometric structures, such as mountain belts. With this in mind, the Earth's crust is definitely heterogeneous (figure 1) with regards to seismic exploration. Therefore, we should use models in our seismic imaging techniques that can account for this variation.

Nowadays, most seismic imaging techniques take into account the complexity of the earth using more accurate modeling methods (e.g., Červený, 1985; Audebert et al., 1997; Bevc, 1997; Broto and Ehinger, 1998; Casadei et al., 2002; Casarotti et al., 2007; Brossier et al., 2009a; Etienne et al., 2010b). Furthermore, wave-equation-based seismic imaging techniques, such as the full waveform inversion (FWI) method, require the most accurate forward modeling methods available. The most accurate forward modeling in this context, means the most realistic simulation of the wave propagation through the Earth. Because heterogeneity and anisotropy are real properties of the subsurface, a forward modeling approach that acknowledges both offers us the best possibility to obtain the true Earth parameter values. Moreover, for a real data set, which is observed (acquired) from anisotropic subsurface, isotropic numerical modeling and inversion are not able to accurately explain the data. For example, the quantitative calculation of model space parameters is ambiguous and incorrect, especially for wide-aperture/wide-azimuth data set (Pratt et al., 2001; Plessix and Cao, 2011a; Prieux et al., 2011). To that end, we investigate the influence of incorporating anisotropy into our model. There are two main reasons for sub-

Figure 1: A Grand Canyon landscape, which is a steep-sided canyon located in the state of Arizona. The exposed geology shows the thick and thin sequences of ancient rocks. This colorful landscape is due to various lithologies and demonstrates geologic heterogeneity at different scales.

surface anisotropy: extrinsic and intrinsic. Extrinsic anisotropy originates from heterogeneities existing at smaller scales than the dominant seismic wavelength (Helbig, 1994). In many geological basins, heterogeneity caused by fine layering or inter-bedding causes *weak* extrinsic anisotropy when the heterogeneity is on a scale smaller than the dominant seismic wavelength (Backus, 1962; Berryman et al., 1999). Contrastingly, intrinsic anisotropy is due to lithological and mineralogical characteristics (Fedorov, 1968; Cara, 2002). Intrinsic anisotropy can range in degree from *weak* to *strong*. Anisotropic media are categorized based on the specific structure of the stiffness matrix (Tsvankin, 2001a). In figure 2 we define four different media based on their anisotropy and heterogeneity. The general heterogeneous anisotropic media, known as triclinic medium, (figure 2d) is described by 21 stiffness coefficients. We describe the more simple monoclinic and orthorhombic media with 13 and 9 independent stiffness, respectively, (Winterstein and Meadows, 1991; Bakulin et al., 2000a,b). The most simple anisotropic medium, in terms of independent coefficients, is the transversely isotropic (TI) medium having only five independent stiffness coefficients. The TI media (with a tilted, horizontal, or vertical symmetry axis) are the most common observed anisotropy. Therefore, the majority of seismic exploration studies are performed using TI media (Tsvankin, 2001a). For a two-dimensional elastic TI medium with vertical symmetry axis (VTI), the stress-strain constitutive equation contains four independent stiffness coefficients. The forward problem of a visco-elastic VTI medium contains four stiffness coefficients, plus density (ρ) and P-wave and S-wave attenuation. Therefore, the model space for a two-dimensional VTI model is composed of seven parameters. However, in seismic exploration, we prefer to estimate model parameters in terms of wavespeed. For this purpose, Thomsen (1986) introduced non-dimensional parameter classes denoted as δ, ϵ and γ for 3D elastic VTI media. The δ and ϵ parameters, in addition to compressional and shear wavespeeds, are convenient to describe the two-dimensional VTI model space in forward modeling and full waveform inversion.

In this PhD dissertation, I implement, validate, and apply a frequency-domain full waveform inversion method (Pratt, 1999) for imaging two-dimensional anisotropic visco-acoustic and visco-elastic media. The method is first validated using synthetic data and then applied to real data. I consider anisotropy corresponding to vertical transverse isotropic (VTI) media. The wave equation, which we use to model seismic data in a two-dimensional VTI subsurface model, is the first-order visco-elastic velocity-stress elastodynamic system with stiffness parameters as physical parameters (Duveneck et al., 2008). The full waveform inversion method relies on a linearized gradient method and is implemented in the frequency domain. The frequency-domain implementation allows us to design a multi-resolution imaging approach by proceeding from the low frequencies to the higher frequencies (Pratt and Worthington, 1990; Bunks et al., 1995). The waveform modelling method is based on a frequency-domain Discontinuous Galerkin finite-element method (Reed and Hill, 1973; Remaki and Fézoui, 1998; Cockburn et al., 2000; Cockburn, 2003; Rivière, 2008; de la Puente et al., 2008; Etienne et al., 2010a; Tago et al., 2010; Smith et al., 2010; Wenk et al., 2010). I describe both the seismic modeling and full waveform

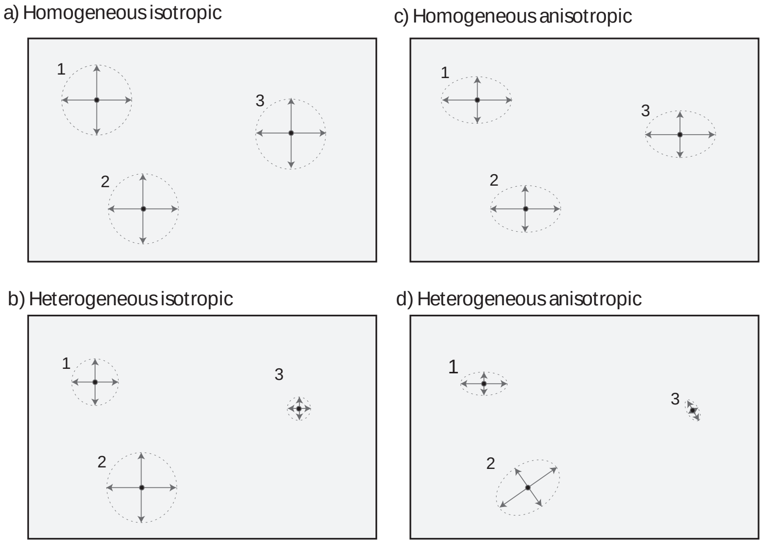


Figure 2: We define different media based on heterogeneity and anisotropy. (a) A homogeneous isotropic medium. The three points in the medium have the same properties and do not vary in space. (b) A heterogeneous isotropic medium. The medium properties are different at all three spatial points. However, at each point, the properties are constant for any direction. (c) A homogeneous anisotropic medium. The properties do not change with position, rather they change with angle. (d) A heterogeneous anisotropic medium. A complex structure where each spatial position has its own properties. A complete description of this medium requires a large number of parameters.

inversion below. Then I discuss how to combine these two approaches and follow with a brief overview of each chapter.

State of the art on FWI

Seismic imaging methods aim to provide precise positioning of reflectors (laterally and vertically), in order to accurately estimate the seismic attributes and explore and exploit hydrocarbon reservoirs. Starting in the beginning of 20th century and continuing up to present day, major seismic discoveries have been made by analyzing and investigating seismic wave traveltime information. Traveltime information is the sum of wave trajectory time from the source to a reflector and then on to the receiver. Around 1980s, seismologist moved beyond traveltimes and began to use amplitude information. This occurred because we gained enough knowledge about the theory of normal mode summation so that we could compute accurate seismograms (Woodhouse and Dziewonski, 1984; Woodhouse, 2007). The amplitude in the reflected wave seismogram directly corresponds to the reflectivity coefficient of the reflector point. Today, one of the predominant imaging techniques is known as full waveform inversion (FWI). This method incorporates both traveltime and amplitude information contained in the complete seismograms to reconstruct an image of the subsurface. The two fundamental elements of FWI are an efficient *full waveform modeling* method and an effective local *optimization* approach.

Seismic modeling

Nowadays, we seek for more advanced seismic imaging techniques in order to provide high resolution images of subsurface (the interest is the natural resources). In order to extract more information (highly resolved images) from the subsurface by seismic imaging techniques, more data needs to be acquired. Also, an accurate seismic modeling tool is necessary. The design of dense and multi-fold (wide-aperture/wide-azimuth) acquisition surveys helps to gather more information from the subsurface. The accurate seismic modeling tool for simulating the seismic waveform propagation through the Earth has been under investigation. The ray-base-method (based on the ray theory) has been often used. The wave-equation-based methods (use the wave equation) simulate the wave propagation. Both methods are used widely in seismic modeling, but the wave-equation-based methods are closer to the reality with respect to the ray-based methods (figure 3a). Various waveform modeling techniques (forward modeling) has been introduced (figure 3b), and are still under improvement. Modeling of the waveform by wave equation let us to observe the wavefield alterations such as reflections, transmissions, and refractions. The model of physical properties (such as velocity, density, etc.) provide the kinematic information required for focusing the waves inside the medium. The information of wave propagation through the subsurface is acquired by receiver and are demonstrated by seismograms.

There are various types of forward modeling techniques based on their discretization methods such as finite-difference methods (Virieux, 1986; Moczo, 1989; Moczo et al., 2001, 2007), finite-element methods (Marfurt, 1984; Zienkiewicz et al., 2005; Brossier et al., 2010a) and finite-volume methods (Dormy and Tarantola, 1995; Remaki, 2000; Ben Jemaa et al., 2007; Brossier et al., 2008, 2010b). Two kinds of boundary condition should be considered: (a) the free surface and (b) the absorbing boundary condition to mimic an infinite medium. The perfectly matched layer (PML) (Berenger, 1994; Chew and Liu, 1996; Hastings et al., 1996; Roden and Gedney,

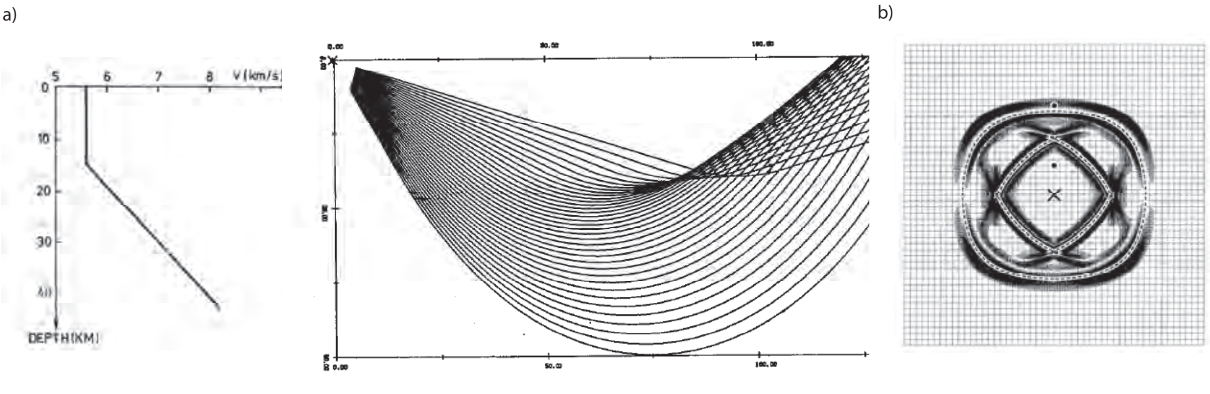


Figure 3: (a) An example of ray-based methods, after Červený et al. (1977, page 96), where the ray diagram of the refracted waves are shown for a medium with an interface between two half-spaces. (b) The modeling of wavefield propagation for elastic anisotropic medium, after Komatitsch et al. (2000, figure 1).

2000; Festa and Nielsen, 2003) has been introduced as an absorbing boundary condition to simulate the infinite wave propagation of seismic wavefield. The first application of full wave equation for seismic imaging was introduced in late 1970s and beginning of 1980s for depth migration of two-dimensional post-stack data by Baysal et al. (1983), McMechan (1983), and Whitmore (1983). The technique is the so-called reverse-time migration (RTM) method, which is a wave-equation-based migration method. On the other hand, the ray-based migration methods such as Kirchhoff migration and beam migrations used the asymptotic solution of the wave equation (French, 1975; Schneider, 1978; Beylkin, 1985). The Kirchhoff migration (as a ray-based method) has been one of the most familiar and dominant migration methods for decades, starting from late 1980s up to now. But in general, the methods based on wave equation are more accurate than ray-based methods, because rays represent an asymptotic solution to the wave equation (Etgen et al., 2009).

Seismic inversion

With the purpose of extracting data from seismograms, a data-fitting procedure based on least-squares optimization method is performed, where the waveform modeling is the core of the technique. The seismic data-fitting procedure was presented as a local optimization method by Lailly (1983) and Tarantola (1984). It is shown to be a least-squares minimization of the misfit between observed (recorded) and calculated (modeled) data, (figure 4). The seismic data fitting (or seismic inversion) appears to be nonlinear. Therefore, the methods were subjected to some iterative linearization. The reason for non-linearity of seismic inversion is the nonlinear relationship between the seismic data and the model parameters, in both time-space and in frequency-space domains in the forward modeling. The early developments of seismic inverse problem have started in 1980s in time-space domain (Tarantola, 1984, 1987; Gauthier et al., 1986; Mora, 1989). The seismic inversion method is called the full-waveform inversion (FWI), because the full information content in the seismogram are considered in the optimization. In the framework of the adjoint-state method, the gradient (first derivative) of the misfit functional is built by cross-correlating the incident modeled wavefield emitted from the source with

the back-propagated residual wavefields (Talagrand and Courtier, 1987; Tromp et al., 2005; Plessix, 2006). In time domain, the gradient computation is performed by summation over sources, receivers, and time. Lailly (1983) and Tarantola (1984) also showed the similarity between the image obtained by reverse-time migration (RTM) and the one obtained by the gradient of the first iteration of full waveform inversion (the difference is that, in seismic inversion, the residual wavefield is back-propagated instead of the observed wavefield in RTM). In full waveform inversion, the smooth background model is updated with the perturbation obtained by optimization algorithm (iterative nonlinear method). Then, the updated model is used as initial model for next iteration. This operation continues toward minimization of misfit function until the convergence to a local minimum. The seismic data in time domain are represented as temporal seismograms (figure 4a).
In general, the local optimization method has difficulties to converge toward the global minimum of misfit function. The main challenge is the accuracy of the starting model, also the lack of low frequencies of the acquired data, the presence of variable type of noises (which can have broad-band frequencies) and the approximations made for modeling the complex wave-propagation, are the other struggling issues. The linearization of the local optimization method was proposed and performed based on the Born approximation theory (single scattering theory) (Weglein and Gray, 1983; Weglein et al., 1986). The starting model of full waveform inversion is obtained by ray-tracing based methods (Červený, 2001) and is usually smooth. When the background is not smooth on wavelength scale (it contains rapid variation), the linearization is less accurate. In other words, the linearization accuracy depends on the degree of non-linearity of the least-squares misfit function. Another widely stressed aspect of FWI is the ill-posedness

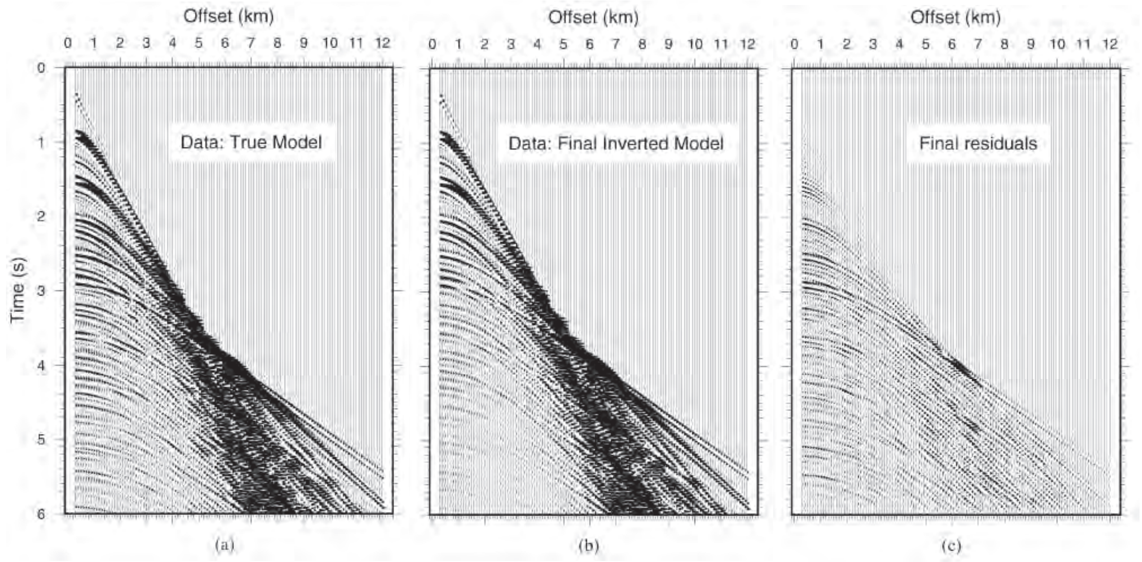


Figure 4: The time distance domain wavefield inversion experiment on a synthetic wide-aperture marine streamer data with 12 km offset, after Shipp and Singh (2002, figure 3). (a) An observed data in time domain shown by the temporal seismogram. (b) The final data computed by FWI, and (c) the residual between observed data and the FWI computed data. The forward problem in the wavefield inversion scheme is based on finite-difference solution of the 2-D elastic wave equation.

of the problem. This means an infinite number of solution models matches the data. In order to ease the ill-posed inversion problem, some conventional regularization (such as Tikhonov regularization) are applied to restrict the local optimum solution search near by a prior model (Menke, 1984; Tarantola, 1987; Scales et al., 1990).

Two fundamental elements of FWI are an efficient full waveform modeling engine (as the core) and an effective local optimization approach. Various local optimization algorithms are applied for solving the seismic inversion problem. The most frequent one is the conjugate-gradient method, which is used by Mora (1987) and Tarantola (1987). The Polak-Ribiére (Polak and Ribière, 1969) formula is one the known algorithm for conjugate-gradient method.

As mentioned, the temporal seismogram of seismic data have a broad band frequency contents, therefore the full information of data is considered in time-domain optimization problem. The intuitive idea of extracting only few frequencies of seismic data was the inspiring motivation to move from time-space to frequency-space domain. As a result, the possibility of separating the long and short seismic wavelength is practicable. In early 1990s, Pratt and Worthington (1990) and Pratt (1990) proposed the application of seismic inversion in frequency-space domain in order to analyze the long offset acquisition surveys. Later, the frequency-domain FWI became a widely applied seismic inversion method (see e.g. Pratt and Shipp (1999); Pratt and Symes (2002); Ravaut et al. (2004); Gelis et al. (2004); Operto and Virieux (2006); Malinowski and Operto (2008); Sirgue et al. (2007); Ben Hadj Ali (2009) and Plessix (2009)). The real temporal seismic data is transformed into frequency-space domain by Fourier transform (FT) and represented by complex numbers. Sirgue and Pratt (2004) showed that the frequency-domain inversion allows to limit the number of frequencies involved in the inversion without affecting the wavenumber sampling of the retrieved image. The hierarchical strategy of starting the seismic inversion from low discrete frequencies and moving upward to high discrete frequencies minimizes the non-linearity of the problem. The large-scale structures of the medium are progressively built at first by low frequency components of data then the smaller scale structures are built as the frequency increases (figure 5). Sirgue (2006) showed that the long offset data still remains challenging as the non-linearity increases by long distance wavefield propagation via various incident angles. The rate of convergence and quality of the frequency-domain inversion is highly dependent on the optimization algorithm. Pratt et al. (1998) showed that Newton or Gauss-Newton methods can reduce the number of convergence iterations with respect to gradient methods. They also gave a clear interpretation of the gradient and Hessian (second derivative of misfit function) using the compact matrix formalism of frequency-domain full waveform inversion. In frequency domain the gradient computation is performed by summation over sources, receivers, and frequencies. Shin et al. (2001) improved the gradient method by scaling the gradient with the diagonal terms of approximated Hessian, known as the pseudo-Hessian. By approximated Hessian, only the single scattered diffract ions are taken into account. The BFGS algorithm (named after its discoverers Broyden, Fletcher, Goldfarb, and Shanno), is a widespread used quasi-Newton optimization algorithm, which updates the approximated Hessian or its inverse at each iteration by taking into account the additional knowledge provided by gradients at previous iterations (S. G. Nash, 1991; Zhu et al., 1997). The approximated Hessian is constructed by a sequence of symmetric matrices, which in some way approximate the Hessian of the unconstrained function. The limited-memory version of the quasi-Newton BFGS method, known as the l-BFGS algorithm, is widely used for large-dimensional problems (Nocedal, 1980). At every step, the quasi-Newton matrix is updated continuously, and the oldest information contained in the matrix is discarded and replaced by new one (Nocedal, 1980). The application of l-BFGS algorithm in FWI requires

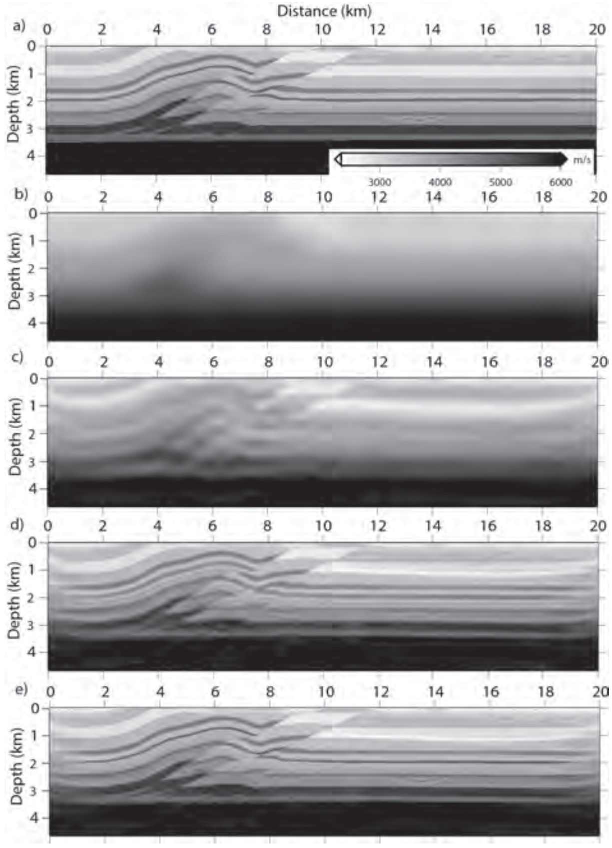


Figure 5: The successive inversion of single frequencies from low-to-high frequencies, after Sourbier et al. (2009, figure 6). (a) The true model of the Overthrust model, (b) the smooth starting model for FWI, (c) velocity model after inversion of the 3.5-Hz inversion, (d) after inversion of the 9.2-Hz inversion, and (d) final model after the inversion of the 20.6-Hz inversion. The resolution of retrieved model increases as the inversion moves up to higher frequencies.

an initial Hessian (as pre-conditioner), which is provided by the inverse of the diagonal of approximated Hessian (Brossier et al., 2009b). In this study we perform our inversion method using the l-BFGS optimization algorithm, which is considered to be the most efficient for our applications.

Anisotropy

During the course of the time, the presence and role of anisotropy was acknowledged in most of seismic processing and imaging methods (see e.g. Backus (1962); Thomsen (1986); Carcione et al. (1992); Farra (1989); Ben Menahem and Gibson (1990); Pratt and Chapman (1992); Carcione (1996); Chapman and Coates (1994); de Hoop et al. (1994); Helbig (1994); Alkhalifah and Tsvankin (1995); Tsvankin (1995); Alkhalifah (1996); Carcione (1996); Červený (1999); Alkhalifah (2000); Pratt et al. (2001); Červený and Pšenčík (2005); Zhou et al. (2006) and Barnes et al. (2008)). It became a well-known fact that the anisotropy is the reason of different interpretations obtained from short and long-offset data set. When Thomsen (1986) introduced the non-dimensional anisotropy parameters in relation with stiffness coefficients for vertically transverse isotropic media, the idea of describing the anisotropic wave equation by physically sensible anisotropic parameters became more feasible. Alkhalifah and Tsvankin (1995) showed that the traveltime curve of the pressure wave can be parametrized by anisotropic normal moveout velocity and the non-dimensional ellipticity parameter (η), which was an important step toward the moveout correction of the long-offset gathers. They showed that the acoustic approximation gives a correct estimation of the pressure wave kinematics. Later, the acoustic approximation of wave equation for VTI media was proposed by Alkhalifah (1998, 2000) by zeroing the shear-wave velocity on the symmetry axis. Zhou H. (2006) and Zhou et al. (2006) proposed the VTI and tilted transverse isotropic (TTI) acoustic wave equations respectively, similar to Alkhalifah (2000) equation, which also gives a good approximation to wave equation for elastic medium with ϵ - $\delta \geq 0$. Operto et al. (2009) followed the same road and proposed a frequency-domain visco-acoustic wavefield modeling for TTI media. Duveneck et al. (2008) presented another acoustic VTI wave equation derived from Hooke's law and the equation of motion, which is composed of first-order velocity-stress wavefields. They used this equation for anisotropic reverse-time migration applications. Beside the benefits of acoustic approximation, some instabilities were observed. The shear wave instability is a common phenomena when the acoustic approximation is performed on elastic wave equation. The shear waves characteristic in acoustic anisotropic media is investigated by Grechka et al. (2004), and they showed that the created shear waves in approximated acoustic wave propagation is the true shear wave (not an artifact), which is created for the angles of propagation deviated from vertical symmetry axis. The other origins of instability come from the absorbing boundary conditions in anisotropic acoustic medium (Bécache et al., 2003), and the sharp discontinuities in acoustic TTI modeling (Zhang and Zhang, 2008; Duveneck and Bakker, 2011).

The anisotropic forward modeling provided the possibility to take into account the anisotropy in full waveform inversion. The application of anisotropy in FWI leads to necessity of the choice of parameterization for the model space (Plessix and Cao, 2011b). The sensitivity analysis of relation between data and model parameters, defines the type of parameterization of the model space. The radiation pattern (scattering pattern) analysis of the virtual secondary sources originating from parameter's perturbation (Pao and Varatharajulu, 1976; Wu, 1989; Sato and Fehler, 1997; Carcione, 1998; Calvet et al., 2006) or the eigenvector decomposition of the Hessian matrix (Plessix and Cao, 2011b), or the principal component analysis of anisotropic sensitivity kernel (Sieminski et al., 2009) are the approaches used for sensitivity analysis. Yet, there are not enough studies on anisotropic FWI. Barnes et al. (2008) studied the feasibility of anisotropic full waveform inversion for a cross-well data. They showed the difficulties of reconstruction of the Thomsen parameters from cross-well data. Plessix and Cao (2011b) studied the parameterization of acoustic VTI medium by eigenvector/eigenvalue decomposition analysis of the Hessian matrix of the least-squares misfit. They showed that in the inversion process, the most reliable information is associated with the largest eigenvalues. They also showed the ambiguity between depth and the δ parameter, and recognized the recoverable anisotropic parameters from diving wave and reflection wave data set. Although, most of the few anisotropic FWI studies are limited to acoustic approximation, Lee et al. (2010) performed the elastic VTI FWI and reconstructed stiffness coefficients as anisotropic parameters. They showed the very weak and noisy reconstruction of the c_{13} parameter and proposed to keep it constant inside the inversion iterations.

Following with recent progresses in anisotropic FWI, this study aims to carry out and apply the 2D VTI FWI with clear objectives. This project includes four main objectives: (a) implementation of the anisotropic forward problem with a frequency-domain finite-element method, following by its validation, (b) implementation of anisotropic full waveform inversion with the frequency-domain finite-element method in an existing isotropic inversion framework, (c) a sensitivity study of the reconstruction of the anisotropic parameters (anisotropic wavespeeds and Thomsen parameters) via partial derivative wavefield analysis and grid analysis of misfit function of least-squares problem. The sensitivity analysis is the support for evaluating the results of synthetic examples of FWI in simple VTI media, (d) and finally, application of the method to realistic synthetic and real data set acquired in seismic exploration frameworks.

In **Chapter 1**, I discuss the anisotropic wave equation and discretisation method used for the forward modeling, as well as the validation of frequency-domain full-waveform modeling through anisotropic synthetic examples. Then, the theory of the full-waveform inversion is explained, following with an overview of parameterization types, which are implemented in the core of full-waveform inversion.

In **Chapter 2**, the Valhall field, located in North sea, is introduced. This field, which has anisotropic characteristics, is the real case study in my work. The geological description of the field and the description of seismic dataset are the ingredients of this chapter.

Chapter 3 is mainly a sensitivity study of the partial-derivative wavefield and the misfit function of least-squares method. The sensitivity analysis is performed to select the best parametrization for acoustic VTI FWI. Then, the VTI FWI is preformed for simple synthetic experimental setup. Through this chapter, the predominant and controlling aspects of anisotropic FWI are investigated and explained.

Chapter 4 contains the application of acoustic VTI FWI on synthetic and real Valhall dataset. At first the VTI FWI is applied on wide-aperture surface dataset for realistic synthetic Valhall model. Then, the two-dimensional data of three-dimensional hydrophone component dataset of Valhall field is inverted and the feasibility of acoustic VTI FWI is evaluated against a real case study.

Chapter 5 is devoted to elastic VTI FWI with a sensitivity analysis through synthetic examples, followed by the application of the method on the geophone dataset of the real Valhall field. The results of this chapter lead to a comparison between acoustic and elastic VTI FWI.

Chapter 1

Method

Contents

1.1 Anisotropic modeling

The seismic wave (generated by an explosive source or earthquake) propagates through the subsurface. The response of the Earth to the propagating wavefield is recorded (on surface or in the well), and is called 'the observed data'. A partial differential equation is used to simulate the seismic wave propagation and is known as the 'wave equation'. The wave equation can be derived under the framework of 'Newton' and 'Hooke' laws (as natural physical laws). Generally, the choice of the model space geometry (one, two or three dimensions), plus the type of propagating wavefield (pressure, velocity or particle displacement), together, introduce variety of wave equations. The method of simulating the wave propagation is known as the *forward problem*. The solution of the forward problem demonstrates the evaluation of the system by quantitative computation of wavefields from the known model parameters and source term(s). As mentioned in previous chapter, a wave equation, which acknowledges the heterogeneity and anisotropy of the subsurface, gives a more close simulation of wave propagation to the reality. Therefore, in this section the anisotropic seismic wave propagation is investigated. In the following, the anisotropic wave equation is represented. The wave equation is for two-dimensional vertically transverse isotropic (VTI) medium. The wave equation is derived using Newton and Hooke's laws in the time-space domain and is also represented in the frequency-space domain. Then, a short overview of discretization methods is discussed. The finite-element Discontinuous Galerkin (DG) discretization method (Brossier, 2009; Brossier et al., 2010b), is applied in forward problem (forward modeling). Some experiments of frequency-domain full waveform modeling (FWM), are performed on synthetic models (section 1.1.5). The validation of Discontinuous Galerkin frequency-domain VTI FWM against the finite-difference time-domain VTI FWM is performed on variety of synthetic models. The synthetic models are in the range of weakly anisotropic to highly anisotropic, and homogeneous to complex models.

1.1.1 Anisotropic elastic wave equation

The *'wave'* in nature and in physics, is defined as a disturbance that travels in space and time, usually accompanied by the transfer of energy from one spatial point to another spatial point in a certain Newtonian time period. If the energy of propagating wave does not cause any permanent displacement of the particles of the medium, then the medium is known as an elastic medium for the received range of strength of energy (stress). When the energy of the wave propagates in the medium both in compressional and in shear displacement then, the wave has elastic propagation. Some media such as fluids, do not bear any shear stress and displacement. This means, the shear displacement does not exist during wave propagation (acoustic media). These type of medium transfers energy only by compressional displacement of the particles. Waves are described by a wave equation, which sets out the disturbance behaviour in space proceeds over Newtonian time. Consider a general heterogeneous anisotropic media as the model space. The wave equation for this model should explain the transfer of energy of the disturbance that travels in a heterogeneous anisotropic space during the time. The wave equation is a partial differential equation (PDE). This PDE is derived from the second Newton's law applied to a volume $\Delta\mathbf{V}$ within a continuum. Expressing the contact forces that act across the surfaces of the $\Delta\mathbf{V}$ in terms of stress tensors σ_{ij}, and the displacement of the materials u_i, yields (Aki and Richards, 1980; Hanyga et al., 1984; Helbig, 1994; Tsvankin, 2001b; Červený, 2001):

$$\rho\frac{\partial^2 u_i}{\partial t^2} - \frac{\partial \sigma_{ij}}{\partial x_j} \quad = \quad f_i, \tag{1.1}$$

where ρ is the density of the medium, u_i stands for components of the particle displacement vector $u = (u_x, u_y, u_z)$, and f_i stands for components of the external force per unit volume f $= (f_x, f_y, f_z)$, t is the time and x_j represents the Cartesian coordinates. If the density of the medium and the applied body force *f(x)* are given, then the wave equation has two unknowns: the displacement field u and the stress tensor σ_{ij}. Therefore, we encounter one equation and two unknowns, which can not be solved without any other auxiliary information. The supplemented equation is the so-called 'constitutive relation' between stress and strain (displacement), which was introduced by Robert Hooke in 1678 in his famous '*of Spring*' article. A constitutive equation is a relation between two physical quantities that are specific to a material or substance, which approximates the response (strain) of that material to external forces (stress). Hooke described the constitutive relation in a very simple language: *The power of any spring is in the same proportion with the tension thereof: that is, if one power stretch or bend it one space, two will bend it two, and three will bend it three, and so forward. Now as the theory is very short, so the way of trying it is very easie.*

In the limit of sufficient stress on the surfaces of a volume $\Delta\mathbf{V}$, which does not cause any permanent deformation (i.e. the particles of the medium restore their place after reload of stress), there is a linear relation between stress and strain. Such a medium is known as elastic medium in this vicinity of applied stress (with respect to applied force the medium is in its elastic limit). Determining the constitutive equation of stress-strain relation provides the basis for the description of static and dynamic deformations of physical medium. The generalized Hooke's law gives the constitutive relation between stress and strain, for an anisotropic, heterogeneous and elastic solid medium. The relation is linear and can be expressed in matricial form. The most general expression of linearity between stress and strain that allows for a stress field in one direction to influence the displacement in all other directions is:

$$\sigma_{ij} = \sum_{kl} c_{ijkl}\epsilon_{kl}. \tag{1.2}$$

The above linear equation provides a general fourth-order tensor with $3^4 = 81$ stiffness elements. This is shown in matricial form as σ=$\mathbf{C}\epsilon$. Thanks to Voigt's convention and several symmetries (Fedorov, 1968; Aki and Richards, 1980; Helbig, 1994), in three-dimensional coordinate system, this relation is expressed in shortened matrix-notation.

$$\begin{bmatrix}\sigma_{xx}\\ \sigma_{yy}\\ \sigma_{zz}\\ \sigma_{yz}\\ \sigma_{zx}\\ \sigma_{xy}\end{bmatrix} = \begin{bmatrix} c_{11} & c_{12} & c_{13} & c_{14} & c_{15} & c_{16}\\ c_{12} & c_{22} & c_{23} & c_{24} & c_{25} & c_{26}\\ c_{13} & c_{23} & c_{33} & c_{34} & c_{35} & c_{36}\\ c_{14} & c_{24} & c_{34} & c_{44} & c_{45} & c_{46}\\ c_{15} & c_{25} & c_{35} & c_{45} & c_{55} & c_{56}\\ c_{16} & c_{26} & c_{36} & c_{46} & c_{56} & c_{66}\end{bmatrix} \begin{bmatrix}\epsilon_{xx}\\ \epsilon_{yy}\\ \epsilon_{zz}\\ 2\epsilon_{yz}\\ 2\epsilon_{zx}\\ 2\epsilon_{xy}\end{bmatrix} \tag{1.3}$$

The above relation describes a medium with lowest possible symmetry by a total of 21 stiffness elements. The tensor c_{ijkl} can be represented in the form of a 6×6 matrix. If the medium has single axis of symmetry and this axis would be vertical, then the medium is known as a vertically transverse isotropic (VTI) medium (Thomsen, 1986; Tsvankin, 2001b, for more details). The symmetry axis in vertically transverse isotropic (VTI) media is z-axis and can be expressed as following matrix with five non-zero independent stiffness coefficients:

$$\begin{bmatrix}\sigma_{xx}\\ \sigma_{yy}\\ \sigma_{zz}\\ \sigma_{yz}\\ \sigma_{zx}\\ \sigma_{xy}\end{bmatrix} = \begin{bmatrix} c_{11} & c_{12} & c_{13} & 0 & 0 & 0\\ c_{12} & c_{11} & c_{13} & 0 & 0 & 0\\ c_{13} & c_{13} & c_{33} & 0 & 0 & 0\\ 0 & 0 & 0 & c_{44} & 0 & 0\\ 0 & 0 & 0 & 0 & c_{44} & 0\\ 0 & 0 & 0 & 0 & 0 & c_{66}\end{bmatrix} \begin{bmatrix}\epsilon_{xx}\\ \epsilon_{yy}\\ \epsilon_{zz}\\ 2\epsilon_{yz}\\ 2\epsilon_{zx}\\ 2\epsilon_{xy}\end{bmatrix}, \tag{1.4}$$

where $c_{66} = \frac{1}{2}(c_{11} - c_{12})$. As we consider the two-dimensional wave propagation, the above stiffness matrix will be reduced to the below stiffness matrix in two-dimensional form. Denote that the horizontal and vertical coordinates are x and z-axis, respectively. Stress-strain relation matrix for 2D VTI media can be written as:

$$\begin{bmatrix}\sigma_{xx}\\ \sigma_{zz}\\ \sigma_{xz}\end{bmatrix} = \begin{bmatrix}c_{11} & c_{13} & 0\\ c_{13} & c_{33} & 0\\ 0 & 0 & c_{44}\end{bmatrix}\begin{bmatrix}\epsilon_{xx}\\ \epsilon_{zz}\\ 2\epsilon_{xz}\end{bmatrix}. \tag{1.5}$$

When the properties of a mediums does not change with respect to a variable direction of measurement, the medium is known as *isotropic*, like in fluids. We remind that in 2-D isotropic media the stress-strain relation is written as:

$$\begin{bmatrix}\sigma_{xx}\\ \sigma_{zz}\\ \sigma_{xz}\end{bmatrix} = \begin{bmatrix}\lambda+2\mu & \lambda & 0\\ \lambda & \lambda+2\mu & 0\\ 0 & 0 & \mu\end{bmatrix}\begin{bmatrix}\epsilon_{xx}\\ \epsilon_{zz}\\ 2\epsilon_{xz}\end{bmatrix}. \tag{1.6}$$

Stress-strain relation of two-dimensional isotropic and VTI media show that media properties are four coefficients for anisotropic media, while there are two coefficients to describe isotropic media. The strain tensor is related to the displacement and is defined as:

$$\epsilon_{kl} = \frac{1}{2}\left(\frac{\partial u_k}{\partial x_l} + \frac{\partial u_l}{\partial x_k}\right). \tag{1.7}$$

To solve the displacement vector in general wave equation, the Hooke's law (equation 1.2) and the definition of strain tensor (equation 1.7) substitute into equation 1.1. In the end, we obtain the general wave equation, which has only one unknown, the displacement vector u:

$$\rho\frac{\partial^2 u_i}{\partial t^2} - \frac{\partial}{\partial x_j}c_{ijkl}\frac{\partial u_k}{\partial x_l} = f_i. \tag{1.8}$$

Equation 1.8 is valid in three-dimensional, linearly elastic, heterogeneous and arbitrary anisotropic media. This means that we have a general wave equation, which can be applied in seismic wave propagation with any arbitrary assumption(s). For example, the wave equation for isotropic media can be obtained by replacing the stiffness coefficients with their isotropic form (λ and μ), or for two-dimensional media by reducing the entries of displacement and stress vectors to two dimensions, i.e. $u = (u_x,u_y,u_z)$ in 3D $\rightarrow u = (u_x,u_z)$ in 2D. In the following, we are going to derive the two-dimensional elastic VTI wave equation. We derive the first-order hyperbolic elasto-dynamics system for two-dimensional P-SV waves in VTI medium. Starting with Hooke's low, the relation between stress and strain vectors in two-dimensional VTI media is:

$$\begin{aligned}\sigma_{xx} &= c_{11}\epsilon_{xx} + c_{13}\epsilon_{zz},\\ \sigma_{zz} &= c_{13}\epsilon_{xx} + c_{33}\epsilon_{zz},\\ \sigma_{xz} &= c_{44}(2\epsilon_{xz}),\end{aligned} \tag{1.9}$$

and the relation between strain and displacement vectors in 2D VTI medium is:

$$\begin{aligned}\epsilon_{xx} &= \frac{\partial u_x}{\partial x},\\ \epsilon_{zz} &= \frac{\partial u_z}{\partial z},\end{aligned}$$

$$\epsilon_{xz} = \frac{1}{2}(\frac{\partial u_x}{\partial z} + \frac{\partial u_z}{\partial x}). \tag{1.10}$$

Inserting the equation 1.10 into the equation 1.9 yields a relation between stress and displacement:

$$\begin{aligned}
\sigma_{xx} &= c_{11}\frac{\partial u_x}{\partial x} + c_{13}\frac{\partial u_z}{\partial z}, \\
\sigma_{zz} &= c_{13}\frac{\partial u_x}{\partial x} + c_{33}\frac{\partial u_z}{\partial z}, \\
\sigma_{xz} &= c_{44}(\frac{\partial u_x}{\partial z} + \frac{\partial u_z}{\partial x}).
\end{aligned} \tag{1.11}$$

The stress and displacement induced from wave propagation are time dependant variables. Taking derivative of equation 1.11 with respect to time variable $\frac{\partial}{\partial t}$, yields:

$$\begin{aligned}
\frac{\partial \sigma_{xx}}{\partial t} &= c_{11}\frac{\partial}{\partial t}\frac{\partial u_x}{\partial x} + c_{13}\frac{\partial}{\partial t}\frac{\partial u_z}{\partial z}, \\
\frac{\partial \sigma_{zz}}{\partial t} &= c_{13}\frac{\partial}{\partial t}\frac{\partial u_x}{\partial x} + c_{33}\frac{\partial}{\partial t}\frac{\partial u_z}{\partial z}, \\
\frac{\partial \sigma_{xz}}{\partial t} &= c_{44}\frac{\partial}{\partial t}(\frac{\partial u_x}{\partial z} + \frac{\partial u_z}{\partial x}).
\end{aligned} \tag{1.12}$$

The time derivative of displacement vector is the particle velocity vector. Therefore, the spatially derived particle velocity vector is related to time variant stress vector by the stiffness coefficient matrix:

$$\begin{aligned}
\frac{\partial \sigma_{xx}}{\partial t} &= c_{11}\frac{\partial v_x}{\partial x} + c_{13}\frac{\partial v_z}{\partial z}, \\
\frac{\partial \sigma_{zz}}{\partial t} &= c_{13}\frac{\partial v_x}{\partial x} + c_{33}\frac{\partial v_z}{\partial z}, \\
\frac{\partial \sigma_{xz}}{\partial t} &= c_{44}(\frac{\partial v_x}{\partial z} + \frac{\partial v_z}{\partial x}),
\end{aligned} \tag{1.13}$$

keeping in mind that the stress-velocity equation 1.13 is derived from Hooke's law. From the general wave equation 1.1, the relation between displacement and stress vectors in 2D media is:

$$\begin{aligned}
\frac{\partial^2 u_x}{\partial t^2} &= \frac{1}{\rho(\mathbf{x})}\left\{\frac{\partial \sigma_{xx}}{\partial x} + \frac{\partial \sigma_{xz}}{\partial z}\right\} + \frac{1}{\rho(\mathbf{x})}f_x, \\
\frac{\partial^2 u_z}{\partial t^2} &= \frac{1}{\rho(\mathbf{x})}\left\{\frac{\partial \sigma_{xz}}{\partial x} + \frac{\partial \sigma_{zz}}{\partial z}\right\} + \frac{1}{\rho(\mathbf{x})}f_z.
\end{aligned} \tag{1.14}$$

Equation 1.14 can be written in terms of particle's velocity and stress:

$$\frac{\partial v_x}{\partial t} = \frac{1}{\rho(\mathbf{x})}\left\{\frac{\partial \sigma_{xx}}{\partial x} + \frac{\partial \sigma_{xz}}{\partial z}\right\} + \frac{1}{\rho(\mathbf{x})}f_x,$$

$$\frac{\partial v_z}{\partial t} = \frac{1}{\rho(\mathbf{x})}\Big\{\frac{\partial \sigma_{xz}}{\partial x} + \frac{\partial \sigma_{zz}}{\partial z}\Big\} + \frac{1}{\rho(\mathbf{x})} f_z. \tag{1.15}$$

Equations 1.13 and 1.15 together give the first-order hyperbolic elasto-dynamics system for 2D P-SV waves in transversely isotropic medium with vertical symmetry axis (VTI medium) in *time-space* domain.

$$\begin{aligned}
\frac{\partial v_x}{\partial t} &= \frac{1}{\rho(\mathbf{x})}\Big\{\frac{\partial \sigma_{xx}}{\partial x} + \frac{\partial \sigma_{xz}}{\partial z}\Big\} + \frac{1}{\rho(\mathbf{x})} f_x, \\
\frac{\partial v_z}{\partial t} &= \frac{1}{\rho(\mathbf{x})}\Big\{\frac{\partial \sigma_{xz}}{\partial x} + \frac{\partial \sigma_{zz}}{\partial z}\Big\} + \frac{1}{\rho(\mathbf{x})} f_z, \\
\frac{\partial \sigma_{xx}}{\partial t} &= c_{11}\frac{\partial v_x}{\partial x} + c_{13}\frac{\partial v_z}{\partial z}, \\
\frac{\partial \sigma_{zz}}{\partial t} &= c_{13}\frac{\partial v_x}{\partial x} + c_{33}\frac{\partial v_z}{\partial z}, \\
\frac{\partial \sigma_{xz}}{\partial t} &= c_{44}(\frac{\partial v_x}{\partial z} + \frac{\partial v_z}{\partial x}),
\end{aligned} \tag{1.16}$$

For tilted transverse isotropic (TTI) medium, the wave equation is subjected to a rotation of the symmetry axis. The derivation of the two dimensional TTI wave equation is shown in appendix A.

1.1.2 Anisotropic acoustic wave equation

In acoustic medium, only the compressional wave propagates and no shear wave can be transferred through the medium. The acoustic wave equation is a second order partial differential equation that describes the evolution of acoustic pressure P or particle velocity v as a function of time t and spatial position x. The general wave equation for acoustic media is:

$$\Delta^2 P(x,t) - \frac{1}{v(x)^2}\frac{\partial^2 P(x,t)}{\partial t^2} = S(x,t), \tag{1.17}$$

where Δ is the Laplacian operator, P is the acoustic pressure wavefield and v is the compressional velocity parameter of the medium. S is the source term, which is different from the source term of the wave equation in term of particle displacements. S in equation 1.17 is equivalent to the spatial derivative of f in equation 1.8 multiplied by the compressibility *K(x).* One should be aware of the meaning of source terms, when we write the wave equation in terms of particle displacement or pressure. In nature the acoustic medium is definitely isotropic. The acoustic wave equation 1.17 describes the evolution of pressure. The pressure is the compressional stress, which is equal in all directions for a spatial point (x_i,y_i,z_i) in model space (combination of principal stresses with equal quantity). As mentioned, in reality the acoustic medium is isotropic and anisotropic acoustic media does not exist. However, in seismic exploration it is favoured to explain the events by the lowest number of parameters and cost of numerical calculations. One of the most well-known assumptions in seismic wave propagation is to consider an elastic medium as an acoustic medium, in order to better explain the P-wave signature and interpretation of observed events. In consequence, the definition of anisotropic acoustic wave propagation has wide range of applications in seismic exploration, even though this is not the reality (Tsvankin, 1995; Alkhalifah, 2000; Grechka et al., 2004; Zhang et al., 2005; Zhou et al., 2006; Duveneck et al., 2008; Operto et al., 2009). The first application of anisotropic wave equation for VTI media with acoustic approximation was introduced by Alkhalifah (2000). He used the dispersion relation (Alkhalifah, 1998) for acoustic VTI media:

$$k_z^2 = \frac{V_{NMO}^2}{V_{P_0}^2}\left(\frac{\omega^2}{V_{NMO}^2} - \frac{\omega^2(k_x^2+k_y^2)}{\omega^2 - 2V_{NMO}^2\eta(k_x^2+k_y^2)}\right), \tag{1.18}$$

to derive the acoustic wave equation:

$$\frac{\partial^2 P}{\partial t^2} = (1+2\eta)V_{NMO}^2(\frac{\partial^2 P}{\partial x^2} + \frac{\partial^2 P}{\partial y^2}) + V_{P_0}^2\frac{\partial^2 P}{\partial z^2} - 2\eta V_{NMO}^2 V_{P_0}^2(\frac{\partial^4 F}{\partial x^2 \partial z^2} + \frac{\partial^4 F}{\partial y^2 \partial z^2}), \tag{1.19}$$

where

$$P = \frac{\partial^2 F}{\partial t^2}. \tag{1.20}$$

The Alkhalifah's equation is a fourth-order partial differential equation in time and space. He demonstrated the feasibility of the acoustic wave equation for simulating VTI wave propagation by numerical examples. Zhou et al. (2006) proposed a new anisotropic acoustic wave equation based on the dispersion relation of Alkhalifah (1998, 2000). They forced shear wavespeed to zero on the symmetry axis and introduced an auxiliary function, which allows the original fourth-order differential equation to become a coupled system of lower-order differential equations.

$$\begin{aligned} \frac{1}{V_{P_0}^2}\frac{\partial^2 p}{\partial t^2} - (1+2\delta)(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2})p - \frac{\partial^2 p}{\partial z^2} &= (1+2\delta)(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2})q, \\ \frac{1}{V_{P_0}^2}\frac{\partial^2 q}{\partial t^2} - 2(\epsilon-\delta)(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2})q &= 2(\epsilon-\delta)(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2})p. \end{aligned} \tag{1.21}$$

Equation 1.21 is easier to implement than the fourth-order Alkhalifah (2000) equation and provides a good kinematic approximation to the elastic wave equation for medium with $\epsilon-\delta \geq 0$. Operto et al. (2009) re-casted the second-order system of Zhou et al. (2006) in a first-order velocity-stress system by introducing the auxiliary particle-velocity wavefields. This allows them to introduce heterogeneous density and discretize the system with staggered-grid finite-difference stencils. The auxiliary velocity wavefields are removed after discretization to come back to a second-order system for numerical resolution in the frequency domain, following a parsimonious approach originally developed in the time domain (Luo and Schuster, 1990). Duveneck et al. (2008) proposed another acoustic wave equation, which is derived from Hooke's law and equations of motion. The acoustic approximation is implemented in their equation by forcing the shear-wave velocity to zero on the symmetry axis. The resulting equation is a first-order partial differential equation with three particle velocity components and two independent stress components in 3D:

$$\frac{\partial v_x}{\partial t} = \frac{1}{\rho}\frac{\partial \sigma_H}{\partial x},$$

$$\begin{aligned}
\frac{\partial v_y}{\partial t} &= \frac{1}{\rho}\frac{\partial \sigma_H}{\partial y}, \\
\frac{\partial v_z}{\partial t} &= \frac{1}{\rho}\frac{\partial \sigma_V}{\partial z}, \\
\frac{\partial \sigma_H}{\partial t} &= \rho V_{P_0}^2[(1+2\epsilon)\Big(\frac{\partial v_x}{\partial x}+\frac{\partial v_y}{\partial y}\Big)+\sqrt{1+2\delta}\frac{\partial v_z}{\partial z}], \\
\frac{\partial \sigma_V}{\partial t} &= \rho V_{P_0}^2[\sqrt{1+2\delta}\Big(\frac{\partial v_x}{\partial x}+\frac{\partial v_y}{\partial y}\Big)+\frac{\partial v_z}{\partial z}].
\end{aligned} \tag{1.22}$$

Equation 1.22 were used for modeling and reverse-time migration of P-wave in VTI media. I use an acoustic wave equation with a derivation similar to one performed by Duveneck et al. (2008). In our equation, the medium properties are the stiffness coefficients. By canceling the shear components of the equations 1.16 on the symmetry axis, it reduces to equation 1.23:

$$\begin{aligned}
\frac{\partial v_x}{\partial t} &= \frac{1}{\rho(\mathbf{x})}\Big\{\frac{\partial \sigma_{xx}}{\partial x}\Big\}+\frac{1}{\rho(\mathbf{x})}f_x, \\
\frac{\partial v_z}{\partial t} &= \frac{1}{\rho(\mathbf{x})}\Big\{\frac{\partial \sigma_{zz}}{\partial z}\Big\}+\frac{1}{\rho(\mathbf{x})}f_z, \\
\frac{\partial \sigma_{xx}}{\partial t} &= c_{11}\frac{\partial v_x}{\partial x}+c_{13}\frac{\partial v_z}{\partial z}, \\
\frac{\partial \sigma_{zz}}{\partial t} &= c_{13}\frac{\partial v_x}{\partial x}+c_{33}\frac{\partial v_z}{\partial z}.
\end{aligned} \tag{1.23}$$

Note that, in the acoustic approximation, we suppose that the source emits pressure pulses, the medium is fluid (which can be heterogeneous and/or anisotropic), and the receiver is a hydrophone.

1.1.3 Frequency-domain approach

In time-domain, the spatial wave propagation is observed through the proceeding time, while, in frequency-domain, the propagating wave in spatial space is shown by a band of the present frequencies over a range of frequencies. Conversion from time-domain to the frequency-domain is possible via the Fourier transform (FT) method (Bracewell, 1986). The inverse FT of FWM in frequency domain for all frequency bandwith of the source term is equivalent to the FWM in time domain. Therefore, the 2D VTI wave equation in time domain, can also be transformed into frequency domain by applying the Fourier transform. The first-order elasto-dynamics system for 2D P-SV waves in transversely isotropic medium with vertical symmetry axis (VTI medium) in the frequency domain, where both velocities (v_x, v_z) and stresses (σ_{xx}, σ_{zz}, σ_{xz}) are unknown quantities, is described by the following partial differential system:

$$\begin{aligned}
-\iota\omega v_x &= \frac{1}{\rho(\mathbf{x})}\Big\{\frac{\partial \sigma_{xx}}{\partial x}+\frac{\partial \sigma_{xz}}{\partial z}\Big\}+\frac{1}{\rho(\mathbf{x})}f_x, \\
-\iota\omega v_z &= \frac{1}{\rho(\mathbf{x})}\Big\{\frac{\partial \sigma_{xz}}{\partial x}+\frac{\partial \sigma_{zz}}{\partial z}\Big\}+\frac{1}{\rho(\mathbf{x})}f_z, \\
-\iota\omega \sigma_{xx} &= c_{11}\frac{\partial v_x}{\partial x}+c_{13}\frac{\partial v_z}{\partial z}-\iota\omega\sigma_{xx_0},
\end{aligned}$$

$$\begin{aligned}
-\iota\omega\sigma_{zz} &= c_{13}\frac{\partial v_x}{\partial x} + c_{33}\frac{\partial v_z}{\partial z} - \iota\omega\sigma_{zz_0}, \\
-\iota\omega\sigma_{xz} &= c_{44}\left\{\frac{\partial v_x}{\partial z} + \frac{\partial v_z}{\partial x}\right\} - \iota\omega\sigma_{xz_0}.
\end{aligned} \tag{1.24}$$

As for the time-domain wave equation, the c_{ij} (c_{11},c_{13}, c_{33} and c_{44}) are the stiffness coefficients for 2D VTI media, and the density is denoted by ρ. The time variable is replaced by the angular frequency (ω) variable. The source terms are either point forces (f_x, f_z) or applied stresses $(\sigma_{xx_0}, \sigma_{zz_0}, \sigma_{xz_0})$ as introduced in system (1.24). The Fourier transform follows the usual convention as $f(\omega) = \int_{-\infty}^{+\infty} f(t)e^{\iota\omega}dt$.

To develop a formulation for integration over a surface in 2D, we consider the following vector with three stress components; $\vec{T}^t = (T_1, T_2, T_3) = ((\sigma_{xx} + \sigma_{zz})/2, (\sigma_{xx} - \sigma_{zz})/2, \sigma_{xz})$. The T_1 stress component represents the pressure and the T_2 stress component represents the deviatoric stress. The new differential system is equivalent to system (1.24) and is written as:

$$\begin{aligned}
-\iota\omega\rho v_x &= \frac{\partial(T_1 + T_2)}{\partial x} + \frac{\partial(T_3)}{\partial z} + f_x, \\
-\iota\omega\rho v_z &= \frac{\partial T_3}{\partial x} + \frac{\partial(T_1 - T_2)}{\partial z} + f_z, \\
-\iota\omega T_1 &= \frac{c_{11} + c_{13}}{2}\frac{\partial v_x}{\partial x} + \frac{c_{13} + c_{33}}{2}\frac{\partial v_z}{\partial z} - \iota\omega T_1^0, \\
-\iota\omega T_2 &= \frac{c_{11} - c_{13}}{2}\frac{\partial v_x}{\partial x} + \frac{c_{13} - c_{33}}{2}\frac{\partial v_z}{\partial z} - \iota\omega T_2^0, \\
-\iota\omega T_3 &= c_{44}\frac{\partial v_z}{\partial x} + c_{44}\frac{\partial v_x}{\partial z} - \iota\omega T_3^0.
\end{aligned} \tag{1.25}$$

As for the time-domain, the acoustic approximation is implemented by setting the shear wave speed to 0 on the symmetry axis. Therefore, the equations 1.24 and 1.25 would be reduced to equations 1.26 and 1.27, respectively:

$$\begin{aligned}
-\iota\omega v_x &= \frac{1}{\rho(\mathbf{x})}\left\{\frac{\partial\sigma_{xx}}{\partial x}\right\} + \frac{1}{\rho(\mathbf{x})}f_x, \\
-\iota\omega v_z &= \frac{1}{\rho(\mathbf{x})}\left\{\frac{\partial\sigma_{zz}}{\partial z}\right\} + \frac{1}{\rho(\mathbf{x})}f_z, \\
-\iota\omega\sigma_{xx} &= c_{11}\frac{\partial v_x}{\partial x} + c_{13}\frac{\partial v_z}{\partial z} - \iota\omega\sigma_{xx_0}, \\
-\iota\omega\sigma_{zz} &= c_{13}\frac{\partial v_x}{\partial x} + c_{33}\frac{\partial v_z}{\partial z} - \iota\omega\sigma_{zz_0},
\end{aligned} \tag{1.26}$$

and

$$\begin{aligned}
-\iota\omega\rho v_x &= \frac{\partial(T_1 + T_2)}{\partial x} + f_x, \\
-\iota\omega\rho v_z &= \frac{\partial(T_1 - T_2)}{\partial z} + f_z, \\
-\iota\omega T_1 &= \frac{c_{11} + c_{13}}{2}\frac{\partial v_x}{\partial x} + \frac{c_{13} + c_{33}}{2}\frac{\partial v_z}{\partial z} - \iota\omega T_1^0, \\
-\iota\omega T_2 &= \frac{c_{11} - c_{13}}{2}\frac{\partial v_x}{\partial x} + \frac{c_{13} - c_{33}}{2}\frac{\partial v_z}{\partial z} - \iota\omega T_2^0.
\end{aligned}$$

(1.27)

The wave equation 1.24 is the system of equation in first-order representing the stress and velocity wavefields. This equation can be rewritten in second order form in term of velocity or stress wavefields for anisotropic medium. By inserting the first-order velocity wavefields of system 1.24 into first-order stress wavefields of the same system, we obtain:

$$\begin{aligned}
\omega^2\sigma_{xx} &= c_{11}\frac{\partial}{\partial x}\frac{1}{\rho}(\frac{\partial\sigma_{xx}}{\partial x}+\frac{\partial\sigma_{xz}}{\partial z})+c_{13}\frac{\partial}{\partial z}\frac{1}{\rho}(\frac{\partial\sigma_{xz}}{\partial x}+\frac{\partial\sigma_{zz}}{\partial z})+(c_{11}\frac{\partial}{\partial x}\frac{f_x}{\rho}+c_{13}\frac{\partial}{\partial z}\frac{f_z}{\rho})+\omega^2\sigma_{xx_0},\\
\omega^2\sigma_{zz} &= c_{13}\frac{\partial}{\partial x}\frac{1}{\rho}(\frac{\partial\sigma_{xx}}{\partial x}+\frac{\partial\sigma_{xz}}{\partial z})+c_{33}\frac{\partial}{\partial z}\frac{1}{\rho}(\frac{\partial\sigma_{xz}}{\partial x}+\frac{\partial\sigma_{zz}}{\partial z})+(c_{13}\frac{\partial}{\partial x}\frac{f_x}{\rho}+c_{33}\frac{\partial}{\partial z}\frac{f_z}{\rho})+\omega^2\sigma_{zz_0},\\
\omega^2\sigma_{xz} &= c_{44}\Big\{\frac{\partial}{\partial z}\frac{1}{\rho}(\frac{\partial\sigma_{xx}}{\partial x}+\frac{\partial\sigma_{xz}}{\partial z})+\frac{\partial}{\partial x}\frac{1}{\rho}(\frac{\partial\sigma_{xz}}{\partial x}+\frac{\partial\sigma_{zz}}{\partial z})+\frac{\partial}{\partial z}\frac{f_x}{\rho}+\frac{\partial}{\partial x}\frac{f_z}{\rho})\Big\}+\omega^2\sigma_{xz_0},
\end{aligned} \tag{1.28}$$

Furthermore, the second-order wave equation in term of velocity wavefields is performed by inserting the stress wavefields equations of system 1.24 into the velocity wavefields of the same systems. We obtain the system :

$$\begin{aligned}
\omega^2\rho v_x+\left(\frac{\partial}{\partial x}c_{11}\frac{\partial v_x}{\partial x}+\frac{\partial}{\partial z}c_{13}\frac{\partial v_z}{\partial z}\right)+\frac{\partial}{\partial z}c_{44}\left(\frac{\partial v_x}{\partial z}+\frac{\partial v_z}{\partial x}\right) &= \imath\omega(\frac{\partial\sigma_{xx0}}{\partial x}+\frac{\partial\sigma_{xz0}}{\partial z})+\imath\omega f_x,\\
\omega^2\rho v_z+\left(\frac{\partial}{\partial x}c_{13}\frac{\partial v_x}{\partial x}+\frac{\partial}{\partial z}c_{33}\frac{\partial v_z}{\partial z}\right)+\frac{\partial}{\partial x}c_{44}\left(\frac{\partial v_x}{\partial z}+\frac{\partial v_z}{\partial x}\right) &= \imath\omega(\frac{\partial\sigma_{zz0}}{\partial z}+\frac{\partial\sigma_{xz0}}{\partial x})+\imath\omega f_z.
\end{aligned} \tag{1.29}$$

Note that equations 1.28 and 1.29 are for heterogeneous medium. Therefore, the medium properties (stiffness coefficients and density) are included in the spatial derivations.

1.1.4 Discretization methods

In order to simulate the wave propagation by any numerical method, the continuous model and the wave equation are transferred into discrete counterparts. Various techniques for seismic wave modelling in realistic complex media have been developed. Such methods include wavenumber integration, e.g. the Reflectivity method (Muller, 1985), ray-tracing (Červený et al., 1977), finite elements (Chen, 1984), Fourier or pseudo-spectral methods (Kosloff and Baysal, 1982), hybrid methods (Emmerich, 1992), and finite differences (Alterman and Karal, 1968; Alford et al., 1974; Kelly et al., 1976). Two well-known and most applied methods of discretization are finite-difference (FD) and finite-element (FE) methods. These methods are applied on various applications of seismic imaging, such as migration methods, wavefield modeling, and any other applications that concern the simulation of seismic events via discretization of a system. Different applications of FD or FE have shown their particular advantages and drawbacks. The comparison between finite-difference and finite-element methods with accuracy analysis of both methods for modeling of scalar elastic wavefield is discussed by Marfurt (1984). He shows that in homogeneous media the explicit finite-element and finite-difference schemes are comparable when solving the scalar wave equation and when solving the elastic wave equations with Poisson's ratio ($\upsilon = \frac{1}{2}(V_p^2 - 2V_s^2)/(V_p^2 - V_s^2)$) less than 0.3. Discretization by finite-difference methods is straight forward and simple to apply. The drawback of FD methods is mostly recognized in complex structures and discontinuities such as irregular topographies. On the other hand, finite-element methods overcome this difficulty but the computation time (cost) of modeling by FE methods are higher than that of FD methods. The grid interval in FD methods is generally uniform, and at each node the derivation of the wavefield is approximated by wavefield values at neighbour nodes. On the other hand, the FE methods has the advantage to utilize elements of variable size (referred to as *h* adaptivity). Each element exchanges with neighbor elements by an interpolation formula, which yields a set of algebric equations. The reader is referred to Clough (1960); Zienkiewicz and Taylor (1967); Hughes (2003); Zienkiewicz et al. (2005); Brenner and Ridgway Scott (2008) for detailed discussion and basics of the finite-element methods. In general, the finite-element method is more complicated than the finite-difference method. A detailed and specialized application of FE method in seismology is presented by Lysmer and Drake (1972).

The Galerkin Method (DG) is a very popular form of FE methods for finding numerical solutions of partial differential equations (PDEs). The discontinuous Galerkin (DG) method allows the implementation of the p–adaptive strategy (the interpolation order of the shape function can vary from one element to the next). The discontinuous Galerkin (DG) method is formulated by a weak formulation (Zienkiewicz and Taylor, 1967; Brossier, 2009). The DG method works over a trial (h-p-adaptive) function that is piecewise discontinuous. Moczo et al. (2010a,b) compared the accuracy of finite-difference and finite-element methods, plus the evaluation of DG methods with respect to P-wave to S-wave speed ratio. Generally, the application of Galerkin method for a PDE along with its domain and boundary conditions needs to identify the solution space (a vectorial space). The PDE is rewritten as weak formulation, which utilises a function type (based on the order of the function) to approximate the solution. If the function is a piecewise completely discontinuous then, the method is known as the discontinuous Galerkin Method. In the following the h-p-adaptive discontinuous Galerkin discretization method applied on 2-D first-order VTI P-SV wave equation is explained. The method of discretization is very close to that developed by Brossier (2009) in isotropic media. However, the anisotropic wave equation cannot be written in a simple way in pseudo-conservative condition, unlike the isotropic counterpart. The pseudo-conservative form means that all of the model properties are in the coefficients of the mass matrix, a distinct advantage for discretization. On the other hand, the discretization is applied on each cell, which is considered as locally continuous.

1.1.4.1 Spatial discretisation of DG method

We consider the first-order hyperbolic elastodynamic system of equations 1.25 for two dimensional P-SV waves in transversely isotropic medium with vertical symmetry axis (VTI medium) in the frequency domain, where both velocities v_x and v_z, and stresses components T_1, T_2 and T_3, are the five unknown vectors. System 1.25 can be written in a vectorial form:

$$\begin{aligned} -\iota\omega\rho\vec{v} &= \overrightarrow{div(\vec{F}(\vec{T}))} - \frac{\partial s_x M_x}{\partial x}\vec{T} - \frac{\partial s_z M_z}{\partial z}\vec{T} + \rho\vec{f}, \\ -\iota\omega\Lambda\vec{T} &= \overrightarrow{div(\vec{G}(\vec{v}))} - \frac{\partial s_x N_x}{\partial x}\vec{v} - \frac{\partial s_z N_z}{\partial z}\vec{v} - \iota\omega\Lambda\vec{T}^0, \end{aligned} \tag{1.30}$$

with the following vectors:

- velocity vector $\vec{v} = (v_x, v_z)^t$,
- stress vector $\vec{T} = (T_1, T_2, T_3)^t$,

- external force vector $\vec{f} = (f_x, f_z)^t$,
- external stress vector $\vec{T^0} = (T_1^0, T_2^0, T_3^0)^t$,

and matrices:

$$\begin{aligned}
M_x &= \begin{bmatrix} 1 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \\
M_z &= \begin{bmatrix} 0 & 0 & 1 \\ 1 & -1 & 0 \end{bmatrix}, \\
N_x &= \begin{bmatrix} (c_{11}+c_{13})/2 & 0 \\ (c_{11}-c_{13})/2 & 0 \\ 0 & c_{44} \end{bmatrix}, \\
N_z &= \begin{bmatrix} 0 & (c_{13}+c_{33})/2 \\ 0 & -(c_{13}-c_{33})/2 \\ c_{44} & 0 \end{bmatrix}.
\end{aligned} \tag{1.31}$$

Vectors $\vec{F}$ and $\vec{G}$ are defined as

$$\begin{aligned}
\vec{F} &= (s_x M_x \vec{T}, s_z M_z \vec{T}), \\
\vec{G} &= (s_x N_x \vec{v}, s_z N_z \vec{v}).
\end{aligned}$$

The Discontinuous Galerkin h-p adaptive method for 2D elastodynamic isotropic media in the frequency domain has been developed by Brossier (2011a). For spatial discretization of 2D elastodynamic anisotropic media all the assumptions of DG isotropic method are valid. Considering the discretize Ω domain containing polygonal 'control cells', notations for boundary of domain $\partial\Omega$, boundary of cell $\partial\mathcal{T}_i$, area of the cell $A_i = \int_{\mathcal{T}_i} dV$, shared interface between cells $\mathcal{T}_{ik} = \mathcal{T}_i \bigcap \mathcal{T}_k$, and unit external normal vector, $\vec{n}_{ik}$, all defined for one control cell $\mathcal{T}_i$ are considered. Technique to develop spatial discretization formula for anisotropic case is quite similar to isotropic one, with minor differences. Such that the physical parameters are not in matrix Λ any more. On the other hand, the physical parameters of anisotropic formulation are placed in N_x and N_z matrices. As the physical parameters are constant inside the cell $\mathcal{T}_i$, minor modification is necessary to develop the anisotropic equations. Following the formulation of spatial discretization of isotropic medium, with little change in few assumptions, the formulation for DG elastodynamic anisotropic 2D P-SV method is performed. On each individual cell $\mathcal{T}_i$, velocity and stress fields (v_x, v_z, T_1, T_2 and T_3), are described based on a local vectorial base function $\vec{\varphi}_{ij}$ with $1 \leq r \leq d_i$ where d_i is the number of degrees of freedom in the cell $\mathcal{T}_i$. Number of degree of freedom depend on the order of local vectorial base function. Local vectorial base function, $\vec{\varphi}_{ij}$ of cell $\mathcal{T}_i$ is zero outside of cell. Velocity and stress fields inside cell $\mathcal{T}_i$ are represented as:

$$\begin{aligned}
\vec{v}_i &= \sum_{j=1}^{d_i} v_{ij} \vec{\varphi}_{ij}, \\
\vec{T}_i &= \sum_{j=1}^{d_i} T_{ij} \vec{\varphi}_{ij},
\end{aligned} \tag{1.32}$$

where v_{ir} and T_{ir} represent the j^{th} degree of freedom of $\vec{v}_i$ and $\vec{T}_i$ respectively. Reminding that all physical parameters inside the cell are taken as constant plus PML function for entire cell. The velocity and stress field in Ω domain is close to:

$$\begin{aligned}
\vec{v} \simeq \sum_i \vec{v}_i &= \sum_i \sum_{j=1}^{d_i} v_{ij} \vec{\varphi}_{ij}, \\
\vec{T} \simeq \sum_i \vec{T}_i &= \sum_i \sum_{j=1}^{d_i} T_{ij} \vec{\varphi}_{ij}.
\end{aligned} \tag{1.33}$$

A dot product of system of equations 1.30 with a test local vectorial function $\vec{\varphi}_{ij}$ (chosen in the same space than basis functions, which gives the Galerkin procedure) is applied and the system is integrated on a control cell $\mathcal{T}_i$:

$$\begin{aligned}
\int_{\mathcal{T}_i} -\iota\omega\rho\vec{\varphi}_{ij} \cdot \vec{v} &= \int_{\mathcal{T}_i} \vec{\varphi}_{ij} \cdot \overrightarrow{div(\vec{F}(\vec{T}))} + \int_{\mathcal{T}_i} \rho\vec{\varphi}_{ij} \cdot \vec{f}, \\
\int_{\mathcal{T}_i} -\iota\omega\vec{\varphi}_{ij} \cdot \vec{T} &= \int_{\mathcal{T}_i} \vec{\varphi}_{ij} \cdot \overrightarrow{div(\vec{G}(\vec{v}))} - \int_{\mathcal{T}_i} \iota\omega\vec{\varphi}_{ij} \cdot \vec{T}^0.
\end{aligned} \tag{1.34}$$

In N_x and N_z matrices, stiffness coefficients are constant inside each cell $\mathcal{T}_i$, as physical parameters. Second and third terms of right hand side of equations 1.30 are cancelled due to the constant function s_x and s_z on $\mathcal{T}_i$. Functions s_x and s_z are equal to 1 in center parts of the model and different from 1 in PML zone.

An integration by part of the first term of right hand side of equations 1.34 gives:

$$\begin{aligned}
\int_{\mathcal{T}_i} -\iota\omega\rho\vec{\varphi}_{ij} \cdot \vec{v} &= -\int_{\mathcal{T}_i} \vec{\nabla}\vec{\varphi}_{ij} \colon \vec{F}(\vec{T}) + \int_{\partial\mathcal{T}_i} \vec{\varphi}_{ij} \cdot (\vec{F}(\vec{T})\vec{n}) + \int_{\mathcal{T}_i} \rho\vec{\varphi}_{ij} \cdot \vec{f}, \\
\int_{\mathcal{T}_i} -\iota\omega\vec{\varphi}_{ij} \cdot \vec{T} &= -\int_{\mathcal{T}_i} \vec{\nabla}\vec{\varphi}_{ij} \colon \vec{G}(\vec{v}) + \int_{\partial\mathcal{T}_i} \vec{\varphi}_{ij} \cdot (\vec{G}(\vec{v})\vec{n}) - \int_{\mathcal{T}_i} \iota\omega\vec{\varphi}_{ij} \cdot \vec{T}^0,
\end{aligned} \tag{1.35}$$

with

$$\vec{\nabla}\vec{w} = \begin{bmatrix} \partial_x w_1 & \partial_z w_1 \\ \vdots & \vdots \\ \partial_x w_d & \partial_z w_d \end{bmatrix}, \tag{1.36}$$

for all vector $\vec{w} = (w_1, \cdots, w_d)^t$, and

$$\mathcal{A} \colon \mathcal{B} = \sum_{\substack{1 \leq i \leq n \\ 1 \leq j \leq m}} a_{i,j} b_{i,j}, \tag{1.37}$$

for all matrices $\mathcal{A} = (a_{ij})_{\substack{1 \leq i \leq n \\ 1 \leq j \leq m}} \in \mathcal{M}_{n,m}$ and $\mathcal{B} = (b_{ij})_{\substack{1 \leq i \leq n \\ 1 \leq j \leq m}} \in \mathcal{M}_{n,m}$. Equation 1.35 can be approximated by

$$\begin{aligned}
\int_{\mathcal{T}_i} -\iota\omega\rho_i\vec{\varphi}_{ij} \cdot \vec{v}_i &= -\int_{\mathcal{T}_i} \vec{\nabla}\vec{\varphi}_{ij} \colon \vec{F}(\vec{T}_i) + \int_{\partial\mathcal{T}_i} \vec{\varphi}_{ij} \cdot (\vec{F}(\vec{T}_{/\partial\mathcal{T}_i})\vec{n}) + \int_{\mathcal{T}_i} \rho_i\vec{\varphi}_{ij} \cdot \vec{f}_i, \\
\int_{\mathcal{T}_i} -\iota\omega\vec{\varphi}_{ij} \cdot \vec{T}_i &= -\int_{\mathcal{T}_i} \vec{\nabla}\vec{\varphi}_{ij} \colon \vec{G}(\vec{v}_i) + \int_{\partial\mathcal{T}_i} \vec{\varphi}_{ij} \cdot (\vec{G}(\vec{v}_{/\partial\mathcal{T}_i})\vec{n}) - \int_{\mathcal{T}_i} \iota\omega\vec{\varphi}_{ij} \cdot \vec{T}_i^0,
\end{aligned} \tag{1.38}$$

where $\vec{F}(\vec{T}_{/\partial\mathcal{T}_i})$ and $\vec{G}(\vec{v}_{/\partial\mathcal{T}_i})$ are the approximation of $\vec{F}(\vec{T})$ and $\vec{G}(\vec{v})$ on the interface $\partial\mathcal{T}_i$. ρ_i and matrices N_x and N_z inside $\vec{G}(\vec{v_i})$ are the physical parameters constant on cell $\mathcal{T}_i$.
Centered flux is defined same as what defined by Brossier (2009), considering vectorial properties:

$$\begin{aligned}
\vec{F}(\vec{T}_{/\mathcal{T}_{ik}}) &= \vec{F}\Big(\frac{\vec{T}_i+\vec{T}_k}{2}\Big),\\
\vec{G}(\vec{v}_{/\mathcal{T}_{ik}}) &= \vec{G}\Big(\frac{\vec{v}_i+\vec{v}_k}{2}\Big),
\end{aligned} \tag{1.39}$$

where $\mathcal{T}_k$ is the neighbor cell, which shares the edge $\mathcal{T}_{ik}$ with $\mathcal{T}_i$.
As physical properties are constant on cell, so expression 1.39 can be rewritten as :

$$\begin{aligned}
\vec{F}(\vec{T}_{/\mathcal{T}_{ik}}) &= \frac{1}{2}\big(\vec{F}(\vec{T}_i)+\vec{F}(\vec{T}_k)\big),\\
\vec{G}(\vec{v}_{/\mathcal{T}_{ik}}) &= \frac{1}{2}\big(\vec{G}(\vec{v}_i)+\vec{G}(\vec{v}_k)\big).
\end{aligned} \tag{1.40}$$

This is one main assumption in all over DG formulas. Applying center flux assumption, system 1.38 can now be written as :

$$\begin{aligned}
\int_{\mathcal{T}_i} -\imath\omega\rho_i\vec{\varphi_{ij}}^t\vec{v_i} &= -\int_{\mathcal{T}_i}\vec{\nabla}\vec{\varphi}_{ij} : \vec{F}(\vec{T_i}) + \frac{1}{2}\sum_{k\in v(i)}\int_{\mathcal{T}_{ik}}\vec{\varphi_{ij}}^t(\vec{F}(\vec{T_i})+\vec{F}(\vec{T_k}))\vec{n}_{ik} + \int_{\mathcal{T}_i}\rho_i\vec{\varphi_{ij}}^t\vec{f_i},\\
\int_{\mathcal{T}_i} -\imath\omega\vec{\varphi_{ij}}^t\vec{T_i} &= -\int_{\mathcal{T}_i}\vec{\nabla}\vec{\varphi}_{ij} : \vec{G}(\vec{v_i}) + \frac{1}{2}\sum_{k\in v(i)}\int_{\mathcal{T}_{ik}}\vec{\varphi_{ij}}^t(\vec{G}(\vec{v_i})+\vec{G}(\vec{v_k}))\vec{n}_{ik})\\
&- \int_{\mathcal{T}_i}\imath\omega\vec{\varphi_{ij}}^t\vec{T}_i^0.
\end{aligned} \tag{1.41}$$

We replace $\vec{v_i}$ and $\vec{T_i}$ by their approximation given by expressions 1.33

$$\begin{aligned}
\sum_{k=1}^{d_i} -\imath\omega\rho_i v_{ik}\int_{\mathcal{T}_i}\vec{\varphi_{ij}}^t\vec{\varphi}_{ik} &= -\sum_{k=1}^{d_i} T_{ik}\int_{\mathcal{T}_i}\vec{\nabla}\vec{\varphi}_{ij} : \vec{F}(\vec{\varphi}_{ik}),\\
&+ \frac{1}{2}\sum_{k\in v(i)}\Big[\sum_{r=1}^{d_i} T_{ir}\int_{\mathcal{T}_{ik}}\vec{\varphi_{ij}}^t\vec{F}(\vec{\varphi}_{ir})\vec{n}_{ik} + \sum_{s=1}^{d_k} T_{ks}\int_{\mathcal{T}_{ik}}\vec{\varphi_{ij}}^t\vec{F}(\vec{\varphi}_{ks})\vec{n}_{ik}\Big]\\
&+ \sum_{k=1}^{d_i}\rho_i f_{ik}\int_{\mathcal{T}_i}\vec{\varphi_{ij}}^t\vec{\varphi}_{ik}\\
\sum_{k=1}^{d_i} -\imath\omega T_{ik}\int_{\mathcal{T}_i}\vec{\varphi_{ij}}^t\vec{\varphi}_{ik} &= -\sum_{k=1}^{d_i} v_{ik}\int_{\mathcal{T}_i}\vec{\nabla}\vec{\varphi}_{ij} : \vec{G}(\vec{\varphi}_{ik})\\
&+ \frac{1}{2}\sum_{k\in v(i)}\Big[\sum_{r=1}^{d_i} v_{ir}\int_{\mathcal{T}_{ik}}\vec{\varphi_{ij}}^t\vec{G}(\vec{\varphi}_{ir})\vec{n}_{ik} + \sum_{s=1}^{d_k} v_{ks}\int_{\mathcal{T}_{ik}}\vec{\varphi_{ij}}^t\vec{G}(\vec{\varphi}_{ks})\vec{n}_{ik}\Big]\\
&- \sum_{k=1}^{d_i}\imath\omega T_{ij}^0\int_{\mathcal{T}_i}\vec{\varphi_{ij}}^t\vec{\varphi}_{ik}.
\end{aligned} \tag{1.42}$$

System 1.42 can be expressed in a matrix form with these expressions

$$
\begin{aligned}
\vec{\nabla}\vec{\varphi}\colon \vec{F}(\vec{\varphi}) &= \sum_{\alpha\in\{x,z\}} (\partial_\alpha\vec{\varphi})^t s_\alpha M_\alpha \vec{\varphi},\\
\vec{\nabla}\vec{\varphi}\colon \vec{G}(\vec{\varphi}) &= \sum_{\alpha\in\{x,z\}} (\partial_\alpha\vec{\varphi})^t s_\alpha N_\alpha \vec{\varphi},
\end{aligned}
\tag{1.43}
$$

for each base function vector $\vec{\varphi}$. Moreover, if we define following matrices $\mathcal{P}$ and $\mathcal{Q}$

$$
\begin{aligned}
\mathcal{P}_{ik} &= \sum_{\alpha\in\{x,z\}} n_{ik_\alpha} s_{i_\alpha} M_\alpha,\\
\mathcal{Q}_{ik} &= \sum_{\alpha\in\{x,z\}} n_{ik_\alpha} s_{i_\alpha} N_\alpha,
\end{aligned}
\tag{1.44}
$$

as well as matrices $\mathcal{P}'$ and $\mathcal{Q}'$

$$
\begin{aligned}
\mathcal{P}'_{ik} &= \sum_{\alpha\in\{x,z\}} n_{ik_\alpha} s_{k_\alpha} M_\alpha,\\
\mathcal{Q}'_{ik} &= \sum_{\alpha\in\{x,z\}} n_{ik_\alpha} s_{k_\alpha} N_\alpha.
\end{aligned}
\tag{1.45}
$$

We have following equalities:

$$
\begin{aligned}
\vec{\varphi_{ij}}^t \vec{F}(\vec{\varphi}_{ir})\vec{n}_{ik} &= \vec{\varphi_{ij}}^t \mathcal{P}_{ik}\vec{\varphi}_{ir},\\
\vec{\varphi_{ij}}^t \vec{G}(\vec{\varphi}_{ir})\vec{n}_{ik} &= \vec{\varphi_{ij}}^t \mathcal{Q}_{ik}\vec{\varphi}_{ir},\\
\vec{\varphi_{ij}}^t \vec{F}(\vec{\varphi}_{ks})\vec{n}_{ik} &= \vec{\varphi_{ij}}^t \mathcal{P}'_{ik}\vec{\varphi}_{ks},\\
\vec{\varphi_{ij}}^t \vec{G}(\vec{\varphi}_{ks})\vec{n}_{ik} &= \vec{\varphi_{ij}}^t G'_{ik}\vec{\varphi}_{ks}.
\end{aligned}
\tag{1.46}
$$

If we note $\tilde{\mathcal{K}_i^1}$ and $\tilde{\mathcal{K}_i^2}$ local mass matrices of cell $\mathcal{T}_i$ for velocity and stress fields, whose terms are defined by :

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

and if we note $\mathbf{v}_i$ and $\mathbf{T}_i$ the column vectors $(v_{ij})_{1\le j\le d_i}$ and $(T_{ij})_{1\le j\le d_i}$ respectively, the system 1.42 can be written in the following form,

$$
\begin{aligned}
-\iota\omega\rho_i\left(\tilde{\mathcal{K}_i^1}\mathbf{v}_i\right)_j &= -\sum_{\alpha\in\{x,z\}}\sum_{k=1}^{d_i} T_{ik}\int_{\mathcal{T}_i}(\partial_\alpha\vec{\varphi})^t s_\alpha M_\alpha\vec{\varphi}\\
&+ \frac{1}{2}\sum_{k\in v(i)}\Big[\sum_{r=1}^{d_i} T_{ir}\int_{\mathcal{T}_{ik}} \vec{\varphi_{ij}}^t \mathcal{P}_{ik}\vec{\varphi}_{ir} + \sum_{s=1}^{d_k} T_{ks}\int_{\mathcal{T}_{ik}} \vec{\varphi_{ij}}^t \mathcal{P}'_{ik}\vec{\varphi}_{ks}\Big]\\
&+ \rho_i\left(\tilde{\mathcal{K}_i^1}\mathbf{f}_i\right)_j,
\end{aligned}
\tag{1.47}
$$

$$
\begin{aligned}
-\iota\omega\Big(\tilde{\mathcal{K}_i^2}\mathbf{T}_i\Big)_j &= -\sum_{\alpha\in\{x,z\}}\sum_{k=1}^{d_i} v_{ik}\int_{\mathcal{T}_i}(\partial_\alpha\vec{\varphi})^t s_\alpha N_\alpha\vec{\varphi} \\
&+ \frac{1}{2}\sum_{k\in v(i)}\Big[\sum_{r=1}^{d_i} v_{ir}\int_{\mathcal{T}_{ik}}\vec{\varphi_{ij}}^{\,t}\mathcal{Q}_{ik}\vec{\varphi}_{ir} + \sum_{s=1}^{d_k} v_{ks}\int_{\mathcal{T}_{ik}}\vec{\varphi_{ij}}^{\,t}\mathcal{Q}'_{ik}\vec{\varphi}_{ks}\Big] \\
&- \iota\omega\Big(\tilde{\mathcal{K}_i^2}\mathbf{T}_i^0\Big)_j, \qquad (1.48)
\end{aligned}
$$

using all above mentioned notations, where $\Big(\tilde{\mathcal{K}_i^1}\Big)_{jk} = \Big(\tilde{\mathcal{K}_i^2}\Big)_{jk}$.
Finally we define $\vec{\mathbf{v}}_i$ and $\vec{\mathbf{T}}_i$ vectors, which describe velocity and stress field in cell $\mathcal{T}_i$ defined by

$$
\begin{aligned}
\vec{\mathbf{v}}_i &= (\vec{v_{x_i}}, \vec{v_{z_i}})^t, \\
\vec{\mathbf{T}}_i &= (\vec{T_{1_i}}, \vec{T_{2_i}}, \vec{T_{3_i}})^t. \qquad (1.49)
\end{aligned}
$$

We assume that basis functions ($\vec{\varphi}$) are taken from the same type ($\vec{\varphi}_{ij}, 1 \le j \le d_i$) for representing each component of vector $\vec{\mathbf{v}}_i$ and $\vec{\mathbf{T}}_i$. This is also another main assumption for this method.
Systems 1.47 and 1.48 can be re-casted in a compact form as:

$$
\begin{aligned}
-\iota\omega\rho_i\Big(\mathcal{I}_2\otimes\mathcal{K}_i\vec{\mathbf{v}}_i\Big) &= -\sum_{\alpha\in\{x,z\}}\Big(s_\alpha M_\alpha\otimes\mathcal{E}_{i_\alpha}\vec{\mathbf{T}}_i\Big) \\
&+ \frac{1}{2}\sum_{k\in v(i)}\Big[\big(\mathcal{P}_{ik}\otimes\mathcal{F}_{ik}\vec{\mathbf{T}}_i\big) + \big(\mathcal{P}'_{ik}\otimes\mathcal{G}_{ik}\vec{\mathbf{T}}_k\big)\Big] \\
&+ \rho_i\Big(\mathcal{I}_2\otimes\mathcal{K}_i^1\vec{\mathbf{f}}_i\Big), \\
-\iota\omega\Big(\mathcal{I}_3\otimes\mathcal{K}_i\vec{\mathbf{T}}_i\Big) &= -\sum_{\alpha\in\{x,z\}}\Big(s_\alpha N_\alpha\otimes\mathcal{E}_{i_\alpha}\vec{\mathbf{v}}_i\Big) \\
&+ \frac{1}{2}\sum_{k\in v(i)}\Big[\big(\mathcal{Q}_{ik}\otimes\mathcal{F}_{ik}\vec{\mathbf{v}}_i\big) + \big(\mathcal{Q}'_{ik}\otimes\mathcal{G}_{ik}\vec{\mathbf{v}}_k\big)\Big] \\
&- \iota\omega\Big(\mathcal{I}_3\otimes\mathcal{K}_i\vec{\mathbf{T}}_i^0\Big). \qquad (1.50)
\end{aligned}
$$

Where the expression of the different matrix are :

- Mass matrix $\mathcal{K}_i$ is defined by a surface integration over cell $\mathcal{T}_i$:

$$
\Big(\mathcal{K}_i\Big)_{jk} = \int_{\mathcal{T}_i}\varphi_{ij}\varphi_{ik}d\mathcal{T} \qquad 1 \le j,k \le d_i.
$$

- Matrix $\mathcal{E}_{i_\alpha}$ is defined by a surface integration over cell $\mathcal{T}_i$:

$$
\Big(\mathcal{E}_{i_\alpha}\Big)_{jk} = \int_{\mathcal{T}_i}(\partial_\alpha\varphi_{ij})\varphi_{ik}d\mathcal{T} \qquad 1 \le j,k \le d_i.
$$

- Matrix $\mathcal{F}_{ik}$ and $\mathcal{G}_{ik}$ are defined by a line integration over the interface $\mathcal{T}_{ik}$ between cells $\mathcal{T}_i$ and $\mathcal{T}_k$:

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

and the tensorial product $\otimes$ is defined for every matrix $\mathcal{A} = (a_{ij})_{\substack{1\leq i\leq n \\ 1\leq j\leq m}} \in \mathcal{M}_{n,m}$ and $\mathcal{B} = (b_{ij})_{\substack{1\leq i\leq p \\ 1\leq j\leq q}} \in \mathcal{M}_{p,q}$

$$\mathcal{A} \otimes \mathcal{B} = \begin{bmatrix} a_{11}\mathcal{B} & \cdots & a_{1m}\mathcal{B} \\ \vdots & \vdots & \vdots \\ a_{n1}\mathcal{B} & \cdots & a_{nm}\mathcal{B} \end{bmatrix} \in \mathcal{M}_{np,mq}. \tag{1.51}$$

System 1.50 can be written under the extended following form

$$\begin{aligned}
-\imath\omega\rho_i\mathcal{K}_i\vec{v}_{x_i} &= -\Big[s_{i_x}\mathcal{E}_{i_x}(\vec{T}_{1_i}+\vec{T}_{2_i}) + s_{i_z}\mathcal{E}_{i_z}\vec{T}_{3_i}\Big] \\
&+ \frac{1}{2}\sum_{k\in V(i)} \mathcal{F}_{ik}\Big[(\vec{T}_{1_i}+\vec{T}_{2_i})s_{i_x}n_{ik_x} + \vec{T}_{3_i}s_{i_z}n_{ik_z}\Big] \\
&+ \frac{1}{2}\sum_{k\in V(i)} \mathcal{G}_{ik}\Big[(\vec{T}_{1_k}+\vec{T}_{2_k})s_{k_x}n_{ik_x} + \vec{T}_{3_k}s_{k_z}n_{ik_z}\Big] \\
&+ \rho_i\mathcal{K}_i\vec{F}_{x_i}, \\
-\imath\omega\rho_i\mathcal{K}_i\vec{v}_{z_i} &= -\Big[s_{i_x}\mathcal{E}_{i_x}\vec{T}_{3_i} + s_{i_z}\mathcal{E}_{i_z}(\vec{T}_{1_i}-\vec{T}_{2_i})\Big] \\
&+ \frac{1}{2}\sum_{k\in V(i)} \mathcal{F}_{ik}\Big[\vec{T}_{3_i}s_{i_x}n_{ik_x} + (\vec{T}_{1_i}-\vec{T}_{2_i})s_{i_z}n_{ik_z}\Big] \\
&+ \frac{1}{2}\sum_{k\in V(i)} \mathcal{G}_{ik}\Big[\vec{T}_{3_k}s_{k_x}n_{ik_x} + (\vec{T}_{1_k}-\vec{T}_{2_k})s_{k_z}n_{ik_z}\Big] \\
&+ \rho_i\mathcal{K}_i\vec{F}_{z_i}, \\
-\imath\omega\mathcal{K}_i\vec{T}_{1_i} &= -\Big[\frac{c_{11_i}+c_{13_i}}{2}s_{i_x}\mathcal{E}_{i_x}\vec{v}_{x_i} + \frac{c_{13_i}+c_{33_i}}{2}s_{i_z}\mathcal{E}_{i_z}\vec{v}_{z_i}\Big] \\
&+ \frac{1}{2}\sum_{k\in V(i)} \mathcal{F}_{ik}\Big[\frac{c_{11_i}+c_{13_i}}{2}\vec{v}_{x_i}s_{i_x}n_{ik_x} + \frac{c_{13_i}+c_{33_i}}{2}\vec{v}_{z_i}s_{i_z}n_{ik_z}\Big] \\
&+ \frac{1}{2}\sum_{k\in V(i)} \mathcal{G}_{ik}\Big[\frac{c_{11_i}+c_{13_i}}{2}\vec{v}_{x_k}s_{k_x}n_{ik_x} + \frac{c_{13_i}+c_{33_i}}{2}\vec{v}_{z_k}s_{k_z}n_{ik_z}\Big] \\
&- \mathcal{K}_i\vec{T}^0_{1_i}, \\
-\mathcal{K}_i\vec{T}_{2_i} &= -\Big[\frac{c_{11_i}-c_{13_i}}{2}s_{i_x}\mathcal{E}_{i_x}\vec{v}_{x_i} + \frac{c_{13_i}-c_{33_i}}{2}s_{i_z}\mathcal{E}_{i_z}\vec{v}_{z_i}\Big] \\
&+ \frac{1}{2}\sum_{k\in V(i)} \mathcal{F}_{ik}\Big[\frac{c_{11_i}-c_{13_i}}{2}\vec{v}_{x_i}s_{i_x}n_{ik_x} + \frac{c_{13_i}-c_{33_i}}{2}\vec{v}_{z_i}s_{i_z}n_{ik_z}\Big]
\end{aligned}$$

$$\begin{aligned}
&+ \quad \frac{1}{2}\sum_{k\in V(i)} \mathcal{G}_{ik}\Big[\frac{c_{11_i}-c_{13_i}}{2}\vec{v}_{x_k}s_{k_x}n_{ik_x} + \frac{c_{13_i}-c_{33_i}}{2}\vec{v}_{z_k}s_{k_z}n_{ik_z}\Big] \\
&- \quad \iota\omega\mathcal{K}_i\vec{T}^0_{2_i}, \\
-\iota\omega\mathcal{K}_i\vec{T}_{3_i} \quad &= \quad -c_{44_i}\Big[s_{i_x}\mathcal{E}_{i_x}\vec{v}_{z_i} + s_{i_z}\mathcal{E}_{i_z}\vec{v}_{x_i}\Big] \\
&+ \quad \frac{1}{2}\sum_{k\in V(i)} \mathcal{F}_{ik}\Big[c_{44_i}\vec{v}_{z_i}s_{i_x}n_{ik_x} + c_{44_i}\vec{v}_{x_i}s_{i_z}n_{ik_z}\Big] \\
&+ \quad \frac{1}{2}\sum_{k\in V(i)} \mathcal{G}_{ik}\Big[c_{44_i}\vec{v}_{z_k}s_{k_x}n_{ik_x} + c_{44_i}\vec{v}_{x_k}s_{k_z}n_{ik_z}\Big] \\
&- \quad \iota\omega\mathcal{K}_i\vec{T}^0_{3_i}, \qquad (1.52)
\end{aligned}$$

where $\mathcal{F}_{ik}$ is the flux matrix associated to cell i and $\mathcal{G}_{ik}$ is flux matrix associated to cell k. These matrices are equal if i and k have the same interpolation order but can be different in the other case. Physical properties in cell i is considered constant and not relate to interpolation order. Finally, we define following matrices :

$$\begin{aligned}
\tilde{\mathcal{E}}_{i_x} \quad &= \quad A_i\mathcal{K}_i^{-1}\mathcal{E}_{i_x}, \\
\tilde{\mathcal{E}}_{i_z} \quad &= \quad A_i\mathcal{K}_i^{-1}\mathcal{E}_{i_z}, \\
\tilde{\mathcal{F}}_{i1} \quad &= \quad A_i\mathcal{K}_i^{-1}\mathcal{F}_{i1}, \\
\tilde{\mathcal{F}}_{i2} \quad &= \quad A_i\mathcal{K}_i^{-1}\mathcal{F}_{i2}, \\
\tilde{\mathcal{F}}_{i3} \quad &= \quad A_i\mathcal{K}_i^{-1}\mathcal{F}_{i3}, \\
\tilde{\mathcal{G}}_{i1} \quad &= \quad A_i\mathcal{K}_i^{-1}\mathcal{G}_{i1}, \\
\tilde{\mathcal{G}}_{i2} \quad &= \quad A_i\mathcal{K}_i^{-1}\mathcal{G}_{i2}, \\
\tilde{\mathcal{G}}_{i3} \quad &= \quad A_i\mathcal{K}_i^{-1}\mathcal{G}_{i3}. \qquad (1.53)
\end{aligned}$$

The final system for DG method for anisotropic elastic 2D media would be:

$$\begin{aligned}
-\iota\omega\rho_i A_i\vec{v}_{x_i} \quad &= \quad -\Big[s_{i_x}\tilde{\mathcal{E}}_{i_x}(\vec{T}_{1_i}+\vec{T}_{2_i}) + s_{i_z}\tilde{\mathcal{E}}_{i_z}\vec{T}_{3_i}\Big] \\
&+ \quad \frac{1}{2}\sum_{k\in V(i)} \tilde{\mathcal{F}}_{ik}\Big[(\vec{T}_{1_i}+\vec{T}_{2_i})s_{i_x}n_{ik_x} + \vec{T}_{3_i}s_{i_z}n_{ik_z}\Big] \\
&+ \quad \frac{1}{2}\sum_{k\in V(i)} \tilde{\mathcal{G}}_{ik}\Big[(\vec{T}_{1_k}+\vec{T}_{2_k})s_{k_x}n_{ik_x} + \vec{T}_{3_k}s_{k_z}n_{ik_z}\Big] \\
&+ \quad \rho_i A_i\vec{F}_{x_i}, \\
\iota\omega\rho_i A_i\vec{v}_{z_i} \quad &= \quad -\Big[s_{i_x}\tilde{\mathcal{E}}_{i_x}\vec{T}_{3_i} + s_{i_z}\tilde{\mathcal{E}}_{i_z}(\vec{T}_{1_i}-\vec{T}_{2_i})\Big] \\
&+ \quad \frac{1}{2}\sum_{k\in V(i)} \tilde{\mathcal{F}}_{ik}\Big[\vec{T}_{3_i}s_{i_x}n_{ik_x} + (\vec{T}_{1_i}-\vec{T}_{2_i})s_{i_z}n_{ik_z}\Big] \\
&+ \quad \frac{1}{2}\sum_{k\in V(i)} \tilde{\mathcal{G}}_{ik}\Big[\vec{T}_{3_k}s_{k_x}n_{ik_x} + (\vec{T}_{1_k}-\vec{T}_{2_k})s_{k_z}n_{ik_z}\Big] \\
&+ \quad \rho_i A_i\vec{F}_{z_i}, \\
-\iota\omega A_i\vec{T}_{1_i} \quad &= \quad \Big\{ -\Big[(\frac{c_{11_i}+c_{13_i}}{2})s_{i_x}\tilde{\mathcal{E}}_{i_x}\vec{v}_{x_i} + (\frac{c_{13_i}+c_{33_i}}{2})s_{i_z}\tilde{\mathcal{E}}_{i_z}\vec{v}_{z_i}\Big]
\end{aligned}$$

$$
\begin{aligned}
&+ \frac{1}{2}\sum_{k\in V(i)} \tilde{\mathcal{F}}_{ik}\Big[\Big(\frac{c_{11_i}+c_{13_i}}{2}\Big)\vec{v}_{x_i}s_{i_x}n_{ik_x} + \Big(\frac{c_{13_i}+c_{33_i}}{2}\Big)\vec{v}_{z_i}s_{i_z}n_{ik_z}\Big] \\
&+ \frac{1}{2}\sum_{k\in V(i)} \tilde{\mathcal{G}}_{ik}\Big[\Big(\frac{c_{11_i}+c_{13_i}}{2}\Big)\vec{v}_{x_k}s_{k_x}n_{ik_x} + \Big(\frac{c_{13_i}+c_{33_i}}{2}\Big)\vec{v}_{z_k}s_{k_z}n_{ik_z}\Big]\Big\} \\
&- \iota\omega A_i\vec{T}^0_{1_i}, \\
-\iota\omega A_i\vec{T}_{2_i} = &\Big\{-\Big[\Big(\frac{c_{11_i}-c_{13_i}}{2}\Big)s_{i_x}\tilde{\mathcal{E}}_{i_x}\vec{v}_{x_i} + \Big(\frac{c_{13_i}-c_{33_i}}{2}\Big)s_{i_z}\tilde{\mathcal{E}}_{i_z}\vec{v}_{z_i}\Big] \\
&+ \frac{1}{2}\sum_{k\in V(i)} \tilde{\mathcal{F}}_{ik}\Big[\Big(\frac{c_{11_i}-c_{13_i}}{2}\Big)\vec{v}_{x_i}s_{i_x}n_{ik_x} + \Big(\frac{c_{13_i}-c_{33_i}}{2}\Big)\vec{v}_{z_i}s_{i_z}n_{ik_z}\Big] \\
&+ \frac{1}{2}\sum_{k\in V(i)} \tilde{\mathcal{G}}_{ik}\Big[\Big(\frac{c_{11_i}-c_{13_i}}{2}\Big)\vec{v}_{x_k}s_{k_x}n_{ik_x} + \Big(\frac{c_{13_i}-c_{33_i}}{2}\Big)\vec{v}_{z_k}s_{k_z}n_{ik_z}\Big]\Big\} \\
&- \iota\omega A_i\vec{T}^0_{2_i}, \\
-\iota\omega A_i\vec{T}_{3_i} = &\; c_{44_i}\Big\{-\Big[s_{i_x}\tilde{\mathcal{E}}_{i_x}\vec{v}_{z_i} + s_{i_z}\tilde{\mathcal{E}}_{i_z}\vec{v}_{x_i}\Big] \\
&+ \frac{1}{2}\sum_{k\in V(i)} \tilde{\mathcal{F}}_{ik}\Big[\vec{v}_{z_i}s_{i_x}n_{ik_x} + \vec{v}_{x_i}s_{i_z}n_{ik_z}\Big] \\
&+ \frac{1}{2}\sum_{k\in V(i)} \tilde{\mathcal{G}}_{ik}\Big[\vec{v}_{z_k}s_{k_x}n_{ik_x} + \vec{v}_{x_k}s_{k_z}n_{ik_z}\Big]\Big\} \\
&- \iota\omega A_i\vec{T}^0_{3_i}. \qquad (1.54)
\end{aligned}
$$

1.1.4.2 Wavefield solution: direct solver

The system of linear equation 1.54 can be expressed as matricial form:

$$A(m(x),\omega)u(x,\omega) = s(x,\omega), \qquad (1.55)$$

where A(m(x),ω) is the impedance matrix, m(x) denotes the model parameters, x denotes the spatial coordinate, ω is the angular frequency. The solution of this equation is u, the wavefield vector, and s is the source term. The system of linear equations in matricial form is solved by decomposition of matrix **A** into two upper-diagonal (**U**) and lower-diagonal (**L**) matrices (Press et al., 2007) in our applications. The method is known as direct solver method because the direct solution of wavefield is obtained (Amestoy et al., 2003; MUMPS-team, 2009). Therefore, the system of linear equation **Au=s** is transformed into **LUu=s**, which is solved in two steps. Firstly **Ly=s** and then **Ux=y**. The direct solver is able to solve the equation 1.55 for both symmetrical and unsymmetrical impedance matrices. The advantage of factorization of impedance matrix is that the system can be effectively solved for a large number of sources (Marfurt, 1984). The factorization is performed only one time with a multi-frontal method and the wavefield for multiple sources are computed effectively by forward and backward substitutions (Duff and Reid, 1983; Amestoy et al., 2000).

1.1.5 Anisotropic full waveform modeling results and numerical validation

The full waveform modeling (FWM) for 2D VTI medium is developed and applied for synthetic case studies. A simple illustration of elastic isotropic and anisotropic wave propagation are shown in figure 1.1. In isotropic medium, the propagating wavefront arrives at the same arrival time in any spatial point on a circular trajectory. In other words, the wave propagates in all direction with same speed (figure 1.1a). On the other hand, in anisotropic medium the propagating wavefield differs with respect to the angle of emission for a point source. Figure 1.1b illustrates pressure wavefield propagating in a homogeneous anisotropic medium with Thomsen parameters δ=0.1 and ϵ=0.2. Th/ wavefield trajectory is dominated by Thomsen parameters δ and ϵ on the near-vertical and horizontal axis, respectively. By deviating from principal axis, the wavefield trajectory is controlled by a combination relation of δ and ϵ. The footprint of anisotropy is more detectable as long as the wavefield propagates in longer distance. Figure 1.2 shows the seismograms computed in elastic isotropic and anisotropic models shown in figure 1.1. In (near) zero-offset arrival-time, the footprint of anisotropy is almost negligible and undetectable. A and A′ traces have the same arrival-times in isotropic and anisotropic seismograms. On the other hand, B and B′ traces demonstrate a Δt $\simeq$ 10ms time-difference of arrival-times, highlighting the fact that as wavefield propagates over longer distance, the anisotropy is more effective. This implies the necessity of implying anisotropy in case of long offset dataset, otherwise there would be error in estimation of model space parameters by seismic inversion method (Prieux et al., 2011, for more details). Note that, only 400m offset is considered in this example. In case of 10km offset the arrival time difference is much bigger and should be taken into account. The reflection coefficient amplitude versus offset (AVO), angle (AVA) and/or azimuth are the techniques, which describe the response of medium parameters when the wavefield propagates in far distance.

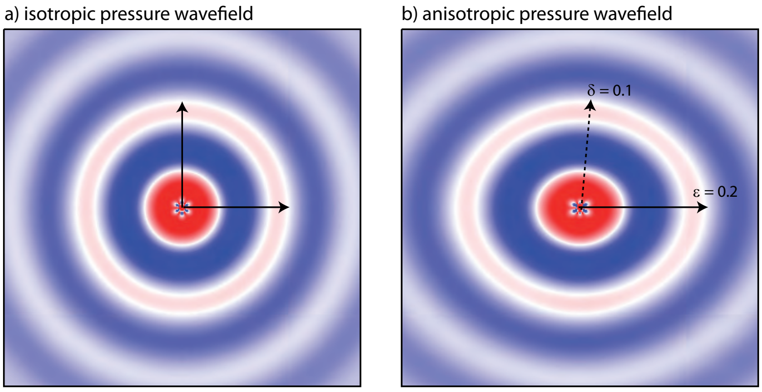


Figure 1.1: The pressure wavefield in elastic homogeneous (a) isotropic and (b) anisotropic (δ=0.1, ϵ=0.2) media. Notice the role of δ and ϵ in variation of wavefield with respect to incident angles. The source is located in the center of the homogeneous medium.

AVA responses of VTI elastic medium

The amplitude versus angle (AVA) response of elastic VTI parameter classes (V_{P_0}, V_{SV}, δ and ϵ) are computed for P-P reflection. The P-P reflection coefficient R_{pp} is computed by a linearized formula (Plessix and Bork, 2000):

$$R_{pp}(\theta) = B_0 + B_1 sin^2\theta + B_2 sin^2\theta tan^2\theta, \tag{1.56}$$

where θ is the incidence angle with respect to the vertical symmetry axis. B_0, B_1 and B_2 are function of elastic VTI parameter classes, which, for small contrast about the interface and

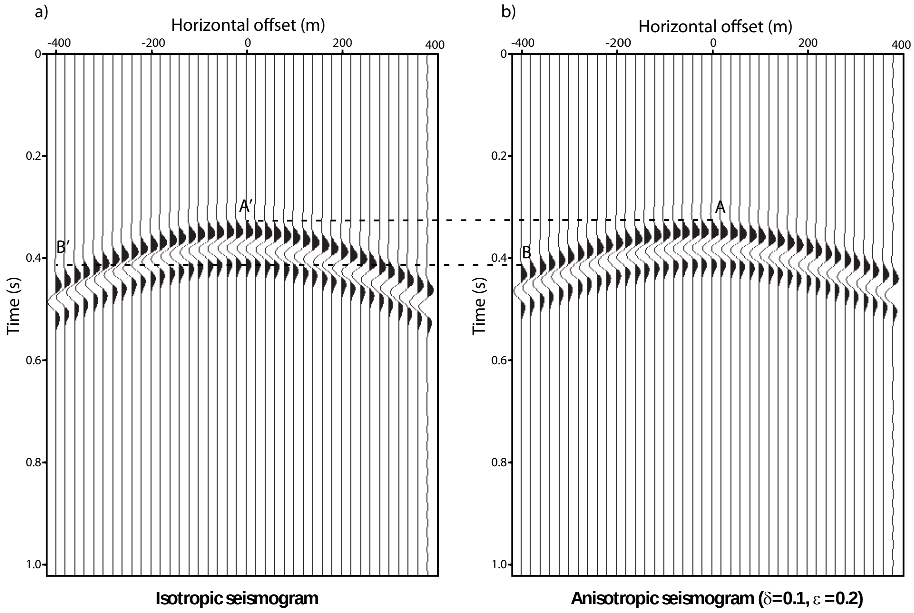


Figure 1.2: The pressure seismograms of elastic isotropic (a) and elastic anisotropic (b) media corresponding to figure 1.1. At zero-offset, there is no difference between first arrival time of A and A′ traces. The footprint of anisotropy is more sensible by longer arrival times where there is a Δt $\simeq$ 10ms difference between first arrival times of B and B′ traces.

	V_{p0}(m/s)	V_{SV}(m/s)	δ	ϵ	ρ(kg/m^3
layer 1	3000	2000	0.0	0.0	2000
layer 2	3300	2300	0.1	0.2	2000

Table 1.1: Values of two layer parameters of elastic VTI interface.

weak anisotropy, are:

$$\begin{aligned} B_0 &= \frac{1}{2}\frac{\Delta I_P}{\bar{I_P}}, \\ B_1 &= \frac{1}{2}(\frac{\Delta V_{P_0}}{\bar{V_{P_0}}} - 4\frac{\bar{V_{SV}^2}}{\bar{V_{P_0}^2}}\frac{\Delta G}{\bar{G}} + \Delta\delta), \\ B_2 &= \frac{1}{2}(\frac{\Delta V_{P_0}}{\bar{V_{P_0}}} + \Delta\epsilon), \end{aligned} \tag{1.57}$$

where $I_P = \rho V_{P_0}$ is the vertical P-wave impedance and $G = \rho V_{SV}^2$ is the vertical shear wave modulus. The bar ($^-$) stands for average operator, and Δ denotes the difference operator in the interface. We follow the experiment performed by Plessix and Bork (2000) and consider an interface between two layers (figure 1.3), which their properties are shown in table 1.1. The behaviour of R_{pp} is shown (figure 1.4), when only one parameter is varied on the second layer and other parameters are fixed. Note that the reflection coefficient is computed until critical angle ($\simeq 65^{\circ}$), and the density does not vary in both layers. The experiment is performed for each parameter class with three different values, when the other parameters are fixed to their true values. The entire AVA curve changes by change in P-wave value, demonstrating the high sensitivity of R_{pp} with respect to P-wave velocity (figure 1.4a). The R_{pp} at very short incident

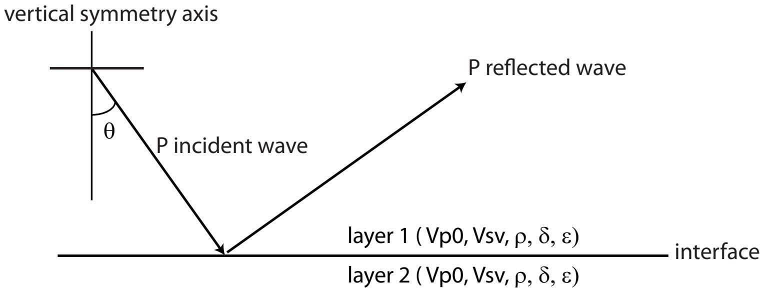


Figure 1.3: The geometry of two elastic VTI layers with horizontal interface. P-P reflected wave of P-P incident wave with θ incident angle is considered.

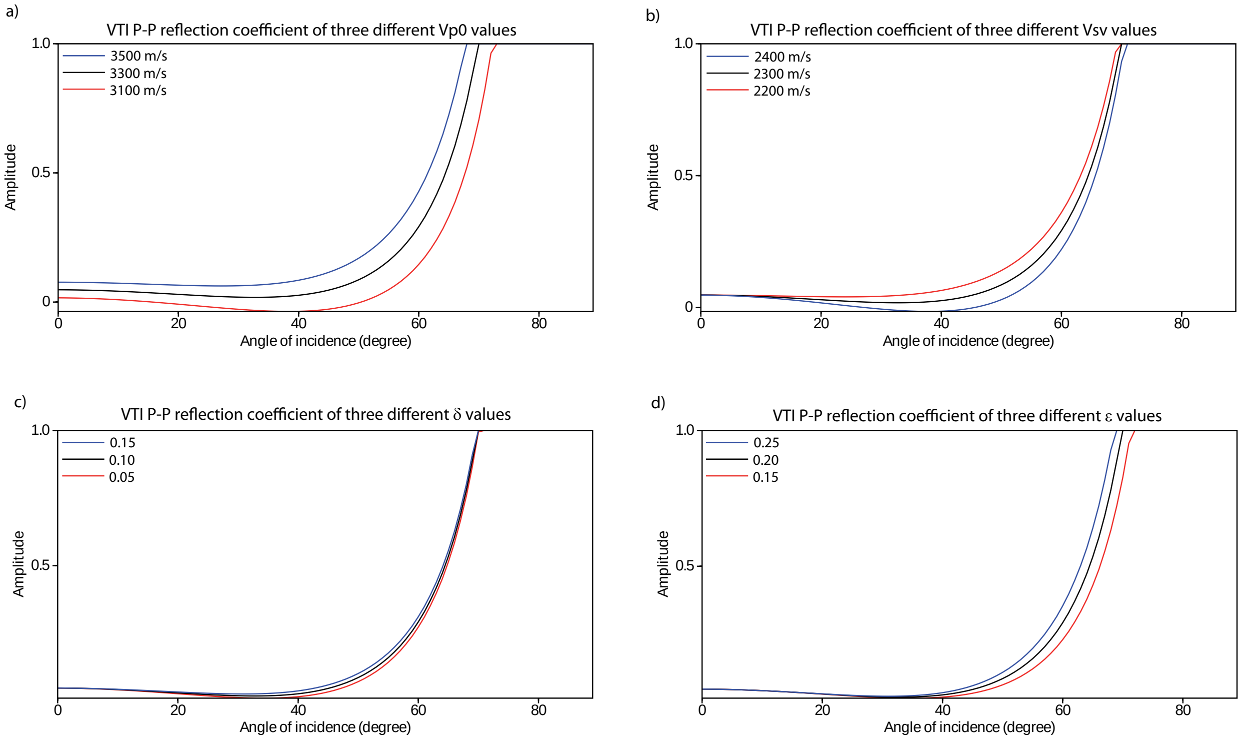


Figure 1.4: P-P reflection coefficient R_{pp} versus incident angle θ of an elastic VTI interface. At each graph three different values of a parameter class of the second layer is applied. When the other parameter are fixed for second layer. (a) V_{P_0} (b) V_{SV} (c) δ (d) ϵ.

angles is not influenced by variation of S-wave velocity. The variation becomes significant beyond 10° incidence angles, and demonstrates noticeable changes in amplitude curves (figure 1.4b). Of note, the critical angle does not vary significantly, when S-wave velocity is modified. Variation in Thomsen parameter δ has minor effect on AVA curves, which is observable only for intermediate angles (figure 1.4c). The critical angle does not change by variation in δ. On the other hand, Thomsen parameter ϵ shows more distinguishable change in AVA curves beyond the medium incidence angles and furthermore, the change in critical angle (figure 1.4d). Of note, the trade-off between δ and V_{SV} at intermediate incident angles and between ϵ and V_{P_0} at large aperture angles is remarkable. In the following, we will show the trade-off between VTI parameter classes by FWI specially between ϵ and V_{P_0} when acoustic VTI FWI is performed (chapter 3).

1.1.5.1 VTI seismic modeling with the discontinuous Galerkin method

Elastic modeling

The Discontinuous Galerkin frequency domain (DGFD) VTI 2D P-SV method is validated against 2D elastic anisotropic (O(Δt^2,Δx^4)) finite-difference time domain (FDTD) method for synthetic anisotropic models. Four types of synthetic models are used: a weakly anisotropic model (the sediment model), a highly anisotropic model (the Zinc model), a heterogeneous weakly anisotropic model (the anticline model), and a complex anisotropic model (the Overthrust model). This choice of synthetic models provides a variety of anisotropic environments, therefore, feasibility and difficulties of anisotropic FWM is investigated. In case of weak anisotropy the frequency-domain FWM succeeds to provide a good fit of seismograms with the 2D elastic finite-difference time-domain FWM. High values of anisotropy raise some difficulties in the absorbing boundaries implemented with PML (Bécache et al., 2004). The frequency-domain anisotropic 2D P-SV FWM has no difficulties in modeling the wave propagation in complex structures such as overthrust synthetic model, except some artifacts on the observed seismogram due to absorbing boundaries. In this study we applied Perfectly Matched Layer (PML) as absorbing boundary condition. The application, feasibility and drawbacks of different kinds of absorbing boundaries (PML, C-PML, ...) are discussed in details by Berenger (1994); Kuzuoglu and Mittra (1996); Hastings et al. (1996); Roden and Gedney (2000) and Festa and Nielsen (2003)for anisotropic, elastic or any arbitrary media. We first validate the method for homogeneous weakly anisotropic model and then homogeneous highly anisotropic media. Then, I move to heterogeneous model with two layers, the anticline, and finally the Overthrust model with many anisotropic layers.

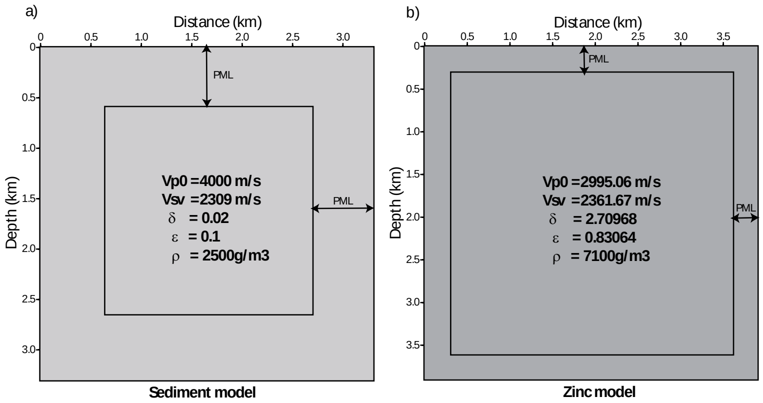


Figure 1.5: (a) The homogeneous weakly anisotropic sediment and (b) the highly anisotropic Zinc synthetic models.

The sediment model: weak anisotropy

The sediment model (figure 1.5a) is a homogeneous, weakly anisotropic synthetic model with size of 2700m in 2700m plus 600m PML layer surrounding four sides. The acquisition array is a surface array composed of a Ricker wavelet source (with dominant frequency of 17 Hz) located at the center of the medium and 101 receivers in a line 1 km above the source. The mesh grid interval is 5m corresponding to 10 cells per wavelength in P_0 mesh (Brossier, 2009). The grid interval is 15 m for finite-difference time-domain FWM. The difference in grid interval in FDTD method (with O(Δt^2, Δx^4) order of precision), with respect to DGFD method is due to the intrinsic properties of discretization of two methods related to dispersion analysis and numerical anisotropy. Grid interval size of model is chosen based on the lowest velocity value (the S_V-wave velocity), and maximum modeled frequency. For DGFD method, the proper number of grid-interval per wavelength is near to 10, while in case of FDTD method, only 4 grid points per wavelength is sufficient. The symmetry axis is vertical, representing a VTI media. The propagating wavefields of T_1, T_2 and T_3 for sediment model simulated at 10 Hz, are shown in figure 1.6. The T_2 wavefield, which is the difference between σ_{zz} and σ_{xx} stress wavefields, has higher amplitude in vertical and horizontal axis. The T_3 wavefield shows the shear wavefield and is zero in principal stresses directions. The comparison between FDTD and DGFD pressure seismograms shows a very good agreement (figure 1.8a). This suggests the validation of the method for a simple case study.

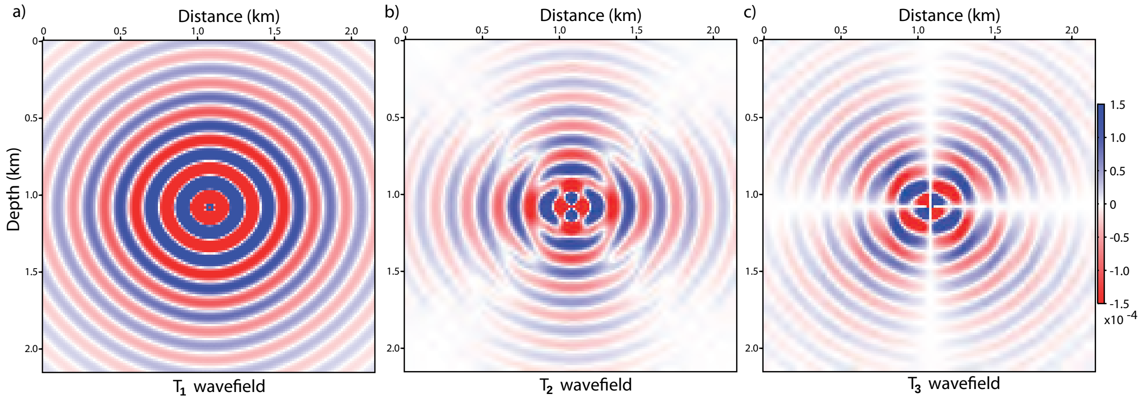


Figure 1.6: The monochromatic stress wavefields (a) T_1 (b) T_2 and (c) T_3 of the weakly anisotropic, elastic and homogeneous sediment model at 10 Hz frequency. The source is located at the center of the homogeneous medium. The model properties are V_{P_0}=4000 m/s, V_{SV}=2309 m/s, δ=0.02 , ϵ=0.1 and density=2500 g/cm^3.

The Zinc model: strong anisotropy

The Zinc crystal model (figure 1.5b) is an elastic homogeneous model with a strong anisotropy. Moreover, the value of δ is much higher than ϵ (δ-ϵ=-1.87904). Providing the condition of an abnormal geological anisotropy experiment, which is accomplished to validate the stability of the DGFD P-SV VTI modeling. Such kind of anisotropy with negative η value is treated and modeled by acoustic full waveform modeling (Operto et al., 2009). We aim to evaluate the stability of DGFD method and the degree of agreement of the method comparing with FDTD result. The model has the square dimension of 3300m×3300m. The grid interval is 5 m for both FDTD grid and DGFD mesh. The source is a Ricker wavelet with dominant frequency of 17 Hz. The symmetry axis is considered vertical, representing a VTI media. There is one shot located at the middle of the model and 101 receivers located on the top of the source at a distance of nearly 1 km.

The compressional pressure wavefields of T_1, T_2 and shear wavefield T_3 of Zinc model at 10 Hz, are shown in figure 1.7. The complexity of a highly anisotropic wavefield can be deduced from

T_2 wavefield, which demonstrates many ruptures in continuity of the wave propagation. The agreement between FDTD and DGFD pressure seismograms are shown in figure 1.8b.

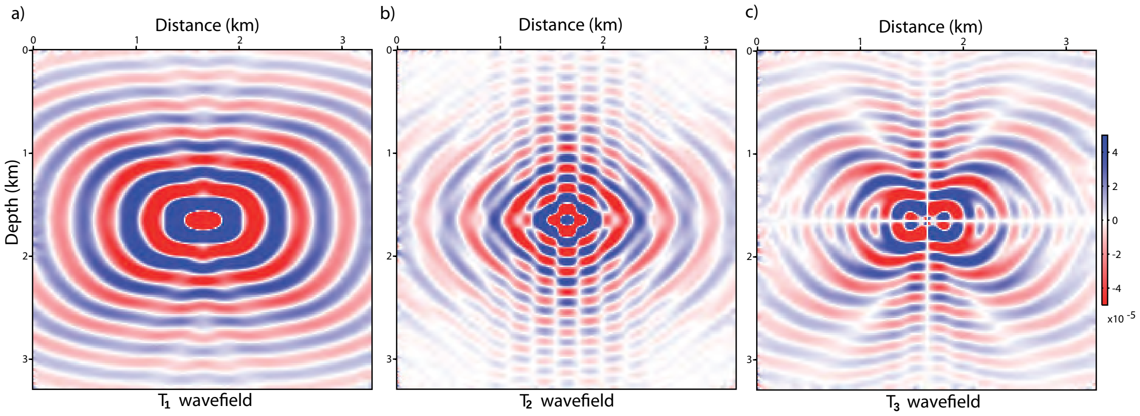


Figure 1.7: The monochromatic stress wavefields (a) T_1 (b) T_2 and (c) T_3 of highly anisotropic homogeneous Zinc model at 6Hz frequency. Note the artifacts on the edges due to instability of PML in the highly anisotropic medium. The source is located in the center of the homogeneous medium. The model properties are V_{P_0}=2955.06 m/s, V_{SV}=2361.67 m/s, δ=2.70968 , ϵ=0.83064 and density=7100 g/cm^3.

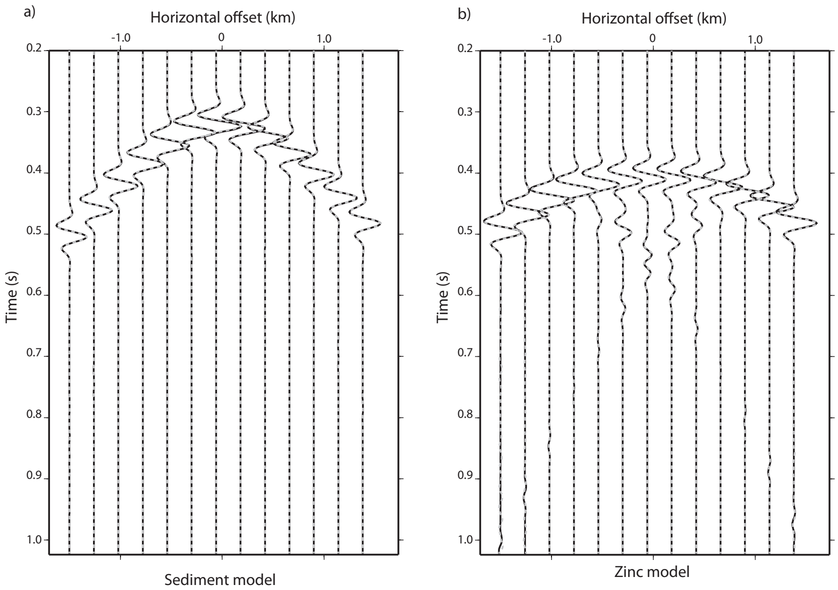
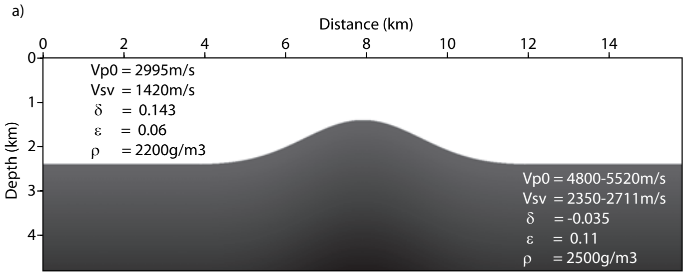


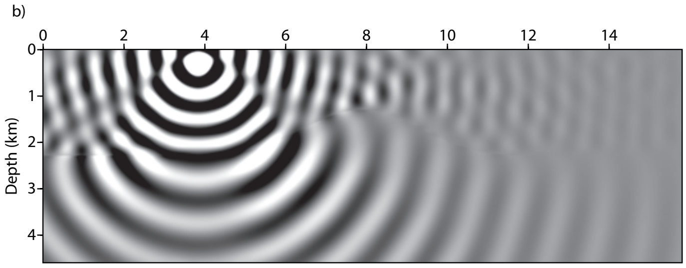
Figure 1.8: Comparison of pressure seismogram of (a) the sediment and (b) the Zinc models. Note the good agreement when the elastic FDTD (the dashed line) compared versus the elastic DGFD P-SV method (the grey line).

The anticline model: a heterogeneous anisotropic structure

The anticline model is an elastic, anisotropic and heterogeneous medium, composed of two anisotropic layers delineated by a bell-shaped interface (anticlinal form), figure(1.9a). The P-wave and S-wave velocity models are homogeneous on the top layer, whereas both increase linearly with respect to depth on the bottom layer. The other parameter classes (δ, ϵ, ρ) are homogeneous for the second layer. The source wavelet is an explosive Ricker wavelet with dominant frequency of 4 Hz. The model size is originally 16000 m in 5000 m plus the 200 m PML size, which is added to the four sides of the model. The grid size for both FDTD and DGFD is 10 m. Considering model size with PML, the source is located at the distance of 8.2 km, and the depth of 0.7 km. A line of 401 receivers are located every 40 m on the top of the source with the distance of 0.1 km. A 10-Hz monochromatic wavefield computed with DGFD method is shown in figure(1.9b). As anticline model has VTI anisotropy, dip of bell-shape interface between two layers is not along or perpendicular to symmetry axis.



Anticline model



T_1 wavefield

Figure 1.9: The anticline model (a) which composed of a homogeneous anisotropic layer plus a heterogeneous anisotropic layer with an anticlinal form. The model properties for each layer is shown in the figure. Second layer has an increasing gradient for V_{P_0} and V_{SV} wavespeeds. (b) the T_1 stress wavefield with a source located near 4 km distance. The wavefield pattern changes after entering in the second layer.

The validation of DGFD method is performed on anticline model by comparing the pressure, the x-velocity and the z-velocity components seismograms against the FDTD method. Figure (1.10) clearly shows how accurate is the coherency of pressure and velocity components seismograms of two methods.

The overthrust model: a complex anisotropic medium

The Overthrust synthetic model (figure 1.11) is a complex anisotropic model with dimension of 20 km in 4.38 km. The PML size around the model is 0.2 km. The model is composed of many layers with isotropic, elliptic and VTI properties. The Thomsen's parameters δ and ϵ range between -0.176602 and 0.06, and between 0 and 0.2, respectively. The P-wave velocity is between 1650 m/s and 6000 m/s, and the shear wave velocity is between 720 m/s and 3420 m/s. A Ricker wavelet with dominant frequency of 4 Hz is used as explosive source. The source

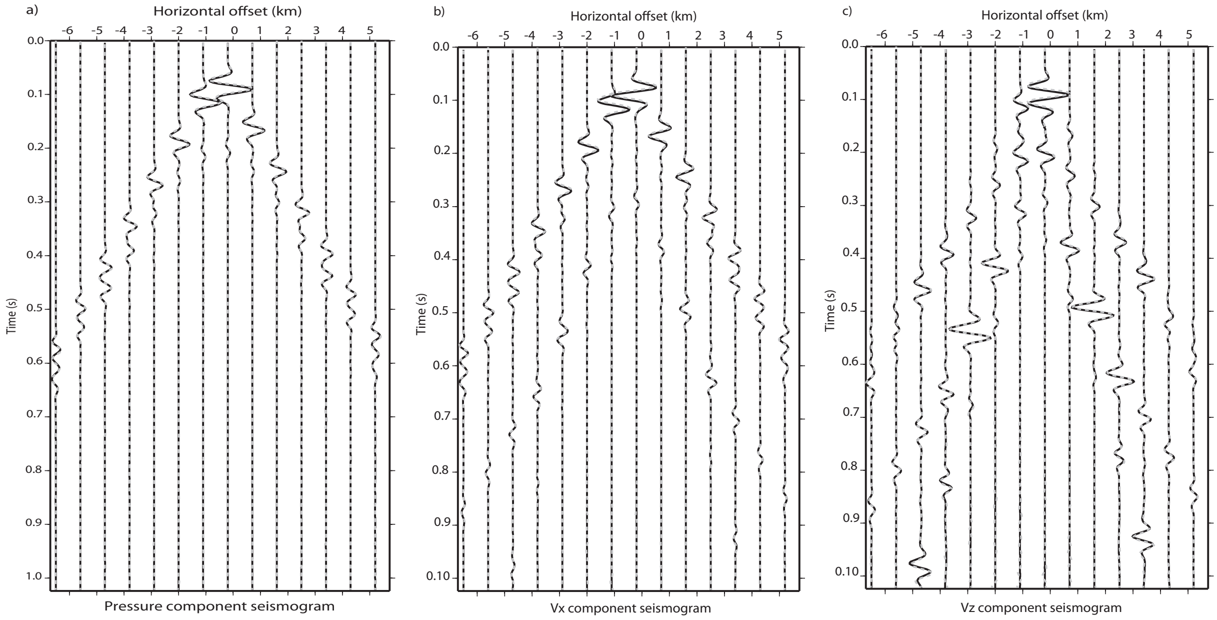


Figure 1.10: Comparison of (a) the pressure, (b) the v_x component and (c) the v_z component seismograms for the anticline model. The elastic FDTD (the dashed line) versus the elastic DG P-SV methods (the grey line).

is located at distance 10 km and depth of 0.4 km with a line of 491 receivers spaced at 40 m intervals at 300 m depth. The grid interval for both FDTD and DGFD methods are 10 m. The monochromatic v_x and v_z wavefields at 10 Hz (figure 1.12) demonstrate the complexity of wave propagation in such a geological structure. The comparison of pressure and velocity seismograms of Overthrust model are shown in figures (1.13). Except the artifacts in the z-component seismogram computed by DGFD, seismic interface arrivals show good agreement between FDTD and DGFD methods.

Acoustic modeling

Beside the advantages of acoustic approximation for anisotropic modeling and inversion, such as reducing the cost of computation, the stability of the forward modeling has been under question. Alkhalifah (2000) observed a diamond-shape wavefronts inside the modeled acoustic wavefield, and considered it as numerical artifacts. Later, Zhang et al. (2003) observed the same diamond-shape wavefronts and interpreted them, also, as numerical artifacts. Grechka et al. (2004) investigated on the stability of the acoustic anisotropic modeling and showed that the diamond-shape wavefronts, which were recognized as numerical artifacts are, in fact, the generated shear waves away from the vertical symmetry axis.

The stability of our acoustic full waveform modeling method is investigated for a homogeneous synthetic model (figure 1.14). We perform the full waveform modeling for an explosive point-source located at the center of three acoustic media, the isotopic, the elliptic and the VTI media. The T_1 wavefield in isotropic acoustic medium (figure 1.14a), shows the simulation of acoustic pressure wavefiled without any artifact. The T_1 wavefield in elliptic (δ=ϵ=0.10) acoustic medium (figure 1.14b), shows also, the simulation of acoustic pressure wavefiled without any artifact. But, the T_1 wavefield simulated in VTI (δ=0.10, ϵ=0.20) shown in figure

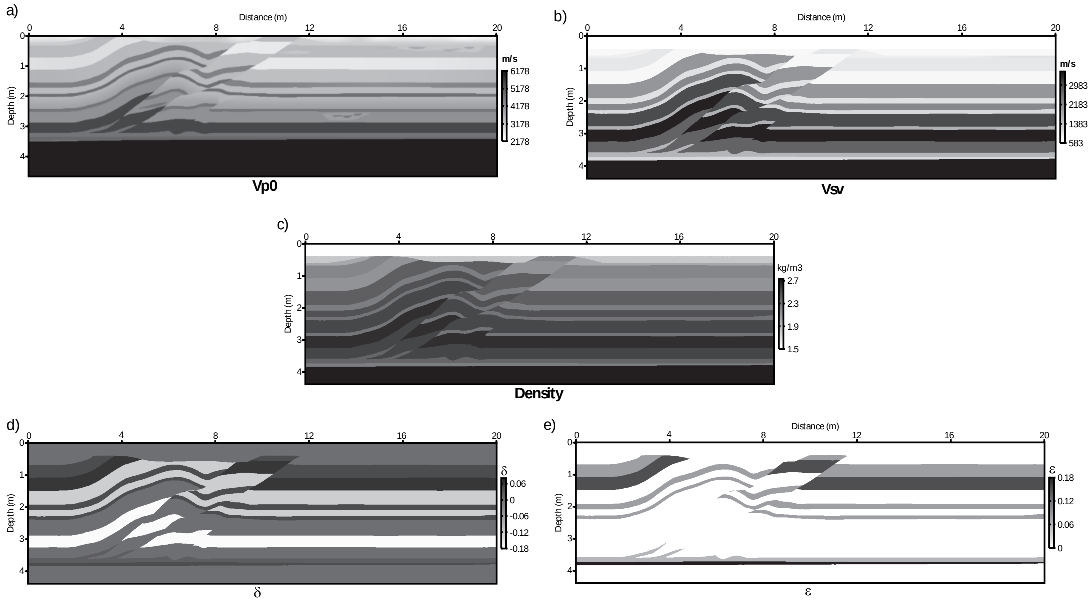
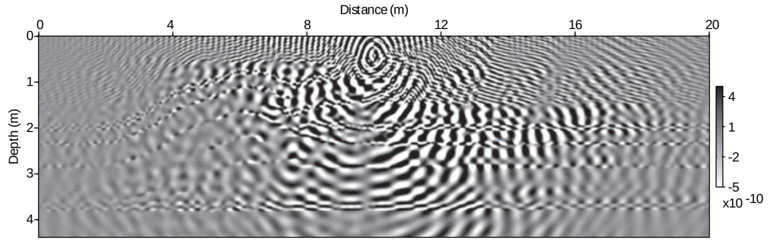
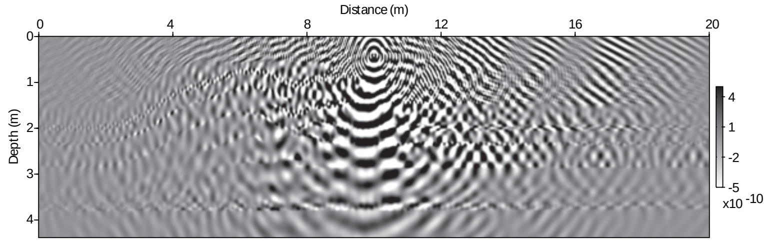


Figure 1.11: The Overthrust model, composed of an overthrust fault representing a heterogeneous anisotropic model. (a) V_{P_0}, (b) V_{SV}, (c) the density (d) Thomsen parameter δ, and (e) Thomsen parameter ϵ.



Vx monochromatic wavefield at10hz



Vz monochromatic wavefield at10hz

Figure 1.12: The velocity wavefields (a) v_x and (b) v_z at 10 Hz frequency of complex anisotropic Overthrust model.

1.14c, shows the diamond-shape wavefronts of shear wave generated away from the vertical and horizontal axis. These observations show that when the source is located in isotropic or elliptic medium the acoustic approximation does not generate any shear wavefront (refer to Grechka et al. (2004, equation 1)). The shear wavefronts are generated only when the source is located in VTI media. In order to demonstrate this claim, we performed an experiment when the source is embedded with a small circular area with elliptic property, inside the VTI

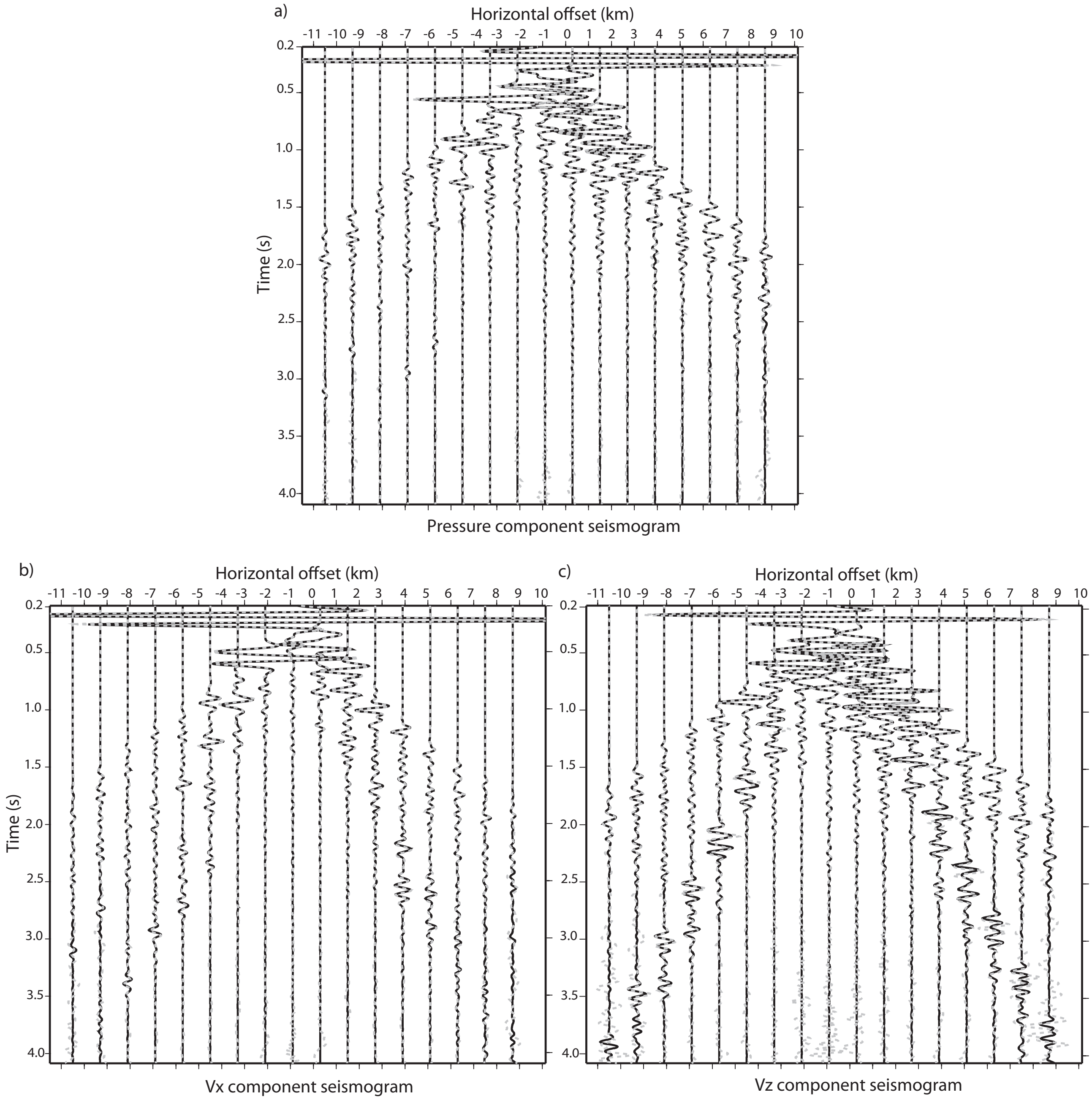


Figure 1.13: Comparison of (a) pressure, (b) v_x component and (c) v_z component seismograms of the Overthrust model, elastic FDTD versus elastic DG P-SV.

medium (figure 1.14d). No diamond-shape shear wavefront is generated. We perform another experiment for a model with VTI interfaces (figure 1.15a). The source is located in an isotropic acoustic layer (similar to water). The P-Sv converted waves are not generated as the wavefront passes through the VTI interface (figure 1.15b). This experiment demonstrates that the shear waves in anisotropic media with acoustic approximation, are produced, only if, the source is located in a VTI medium.

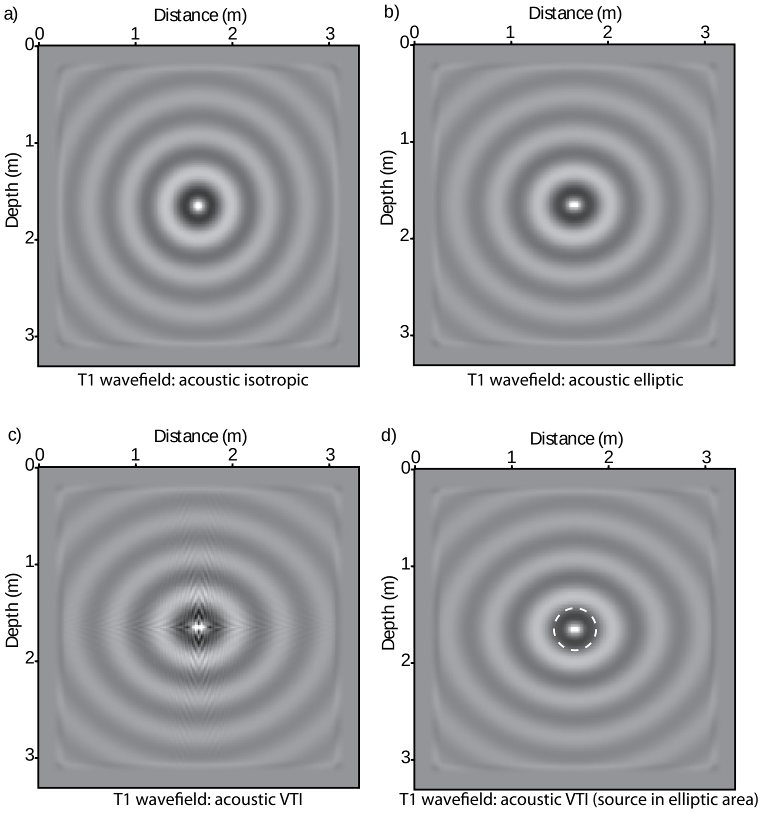


Figure 1.14: The acoustic full waveform modeling of T_1 wavefield for (a) isotropic, (b) elliptic and (c) VTI media. In (d), the source is located in small elliptic area in the center of the medium (the dashed white circle), and out of this area the medium is VTI. The model is shown with PML, which surrounds all four sides of the homogeneous model.

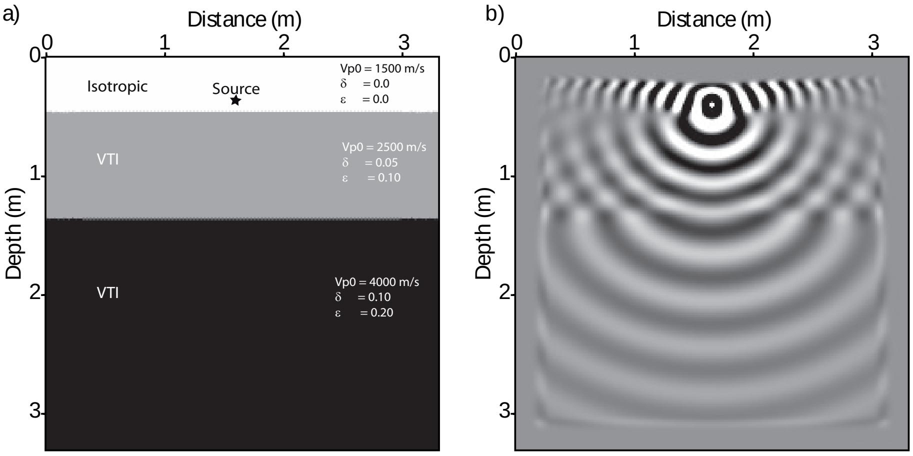


Figure 1.15: (a) The three layers model with one isotropic layer on the top and two VTI layers. (b) The T_1 acoustic pressure wavefield.

1.1.5.2 TTI Modeling

Recent works on tilted TI medium for acoustic reverse-time migration and wave propagation modeling demonstrate the difficulties of the TTI modeling. Zhang and Zhang (2008); Fletcher et al. (2009); Liu et al. (2009); Duveneck and Bakker (2011) and Bakker and Duveneck (2011) showed that the sharp discontinuities of angle of symmetry are the source of instability in acoustic TTI modeling. The instabilities are observed for wave equations, which are derived from dispersion relation and also, for wave equations, which are derived from the Hooke's law and the equation of motions. These sharp variations of angle of symmetry axis mostly happens on the salt flanks. Zhang and Zhang (2008) proposed to use smooth models for TTI modeling in order to remove the instabilities due to sharp discontinuities. Fletcher et al. (2009) and Liu et al. (2009) proposed to perform elastic TTI modeling with a very small shear wave speed, and hence drop the acoustic approximation for TTI modeling. Bakker and Duveneck (2011) analyzed the stability of wave propagation in acoustic TTI media by finite-difference modeling. They proposed an approach based on a modified finite-difference discretization scheme, which overcomes the instability problem.

Zhang et al. (2011) and Zhang Y. (2011) proposed a stable second-order TTI acoustic wave equation and applied in RTM applications. They introduced a self-adjoint differential operators in rotated coordinates to stabilize the TTI acoustic wave equations (in a sense that the energy of the acoustic wavefield is conserved during wave propagation). Their equation system does not add numerical complexity and is solved by central finite-difference scheme.

We apply the acoustic tilted transversely isotropic (TTI) full waveform modeling (FWM) on the purely 2D *2007 BP Anisotropic Velocity-Analysis Benchmark*. The P-wave velocity model is shown in figure 1.16, which is detailed and looks very much like to a real model but with completely invented values. This data set is created at the BP International Centre for Business and Research, Sunbury in 2007 by Hemang Shah as part of a project to study methods in anisotropic imaging and parameterization determination (`http://www.freeusp.org/2007_BP_Ani_Vel_Benchmark/listing.html`).

The structural geology is very complex in some areas, containing an overthrust, an anticlinal and two salt domes. The acoustic anisotropy parameters are shown in figure 1.17. Thomsen parameters δ (figure 1.17b) and ϵ (figure 1.17c) are very detailed for each layer and have zero value inside the salt domes, demonstrating the isotropic property of salt formation. The impure salt (salt mixed with clay, sands, etc.) shows the isotropic behaviour (Sun et al., 1991). The corresponding η model is shown in figure 1.17e. The main feature of this data set is the angle of symmetry axis model θ (figure 1.17d). The high values of deviation angle from vertical symmetry axis are highlighted on the flancs of salt domes and on the discontinuities produced by overthrust faults. These sharp heterogeneities of θ are the undesirable resource of numerical artifacts in TTI modeling.

We apply an experimental procedure for TTI FWM by increasing the complexity originating from the θ parameter class. The stress wavefields T_1 and T_2 are modeled with an explosive source located at near 5 km distance. The monochromatic wavefields are simulated at 4 Hz frequency. First experiment is performed by applying VTI FWM on the BP salt models (figure 1.18). The modeling does not produce any numerical artifacts. Note the zero value of T_2 inside the salt dome (due to isotropic properties of salt). Second experiment is to perform TTI FWM but assuming the θ model as homogeneous model with zero value in whole model parameter space (figure 1.19). We expect to have unique wavefields by VTI FWM and TTI FWM(θ=0). This is verified by computing the residual difference between T_1 wavefields of both modelings

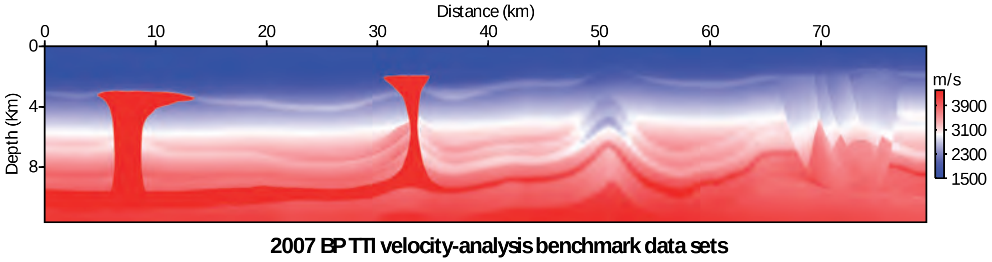


Figure 1.16: The 2D salt model of 2007 BP Anisotropic Velocity-Analysis Benchmark, which is created at the BP International Centre for Business and Research, Sunbury in 2007 by Hemang Shah.

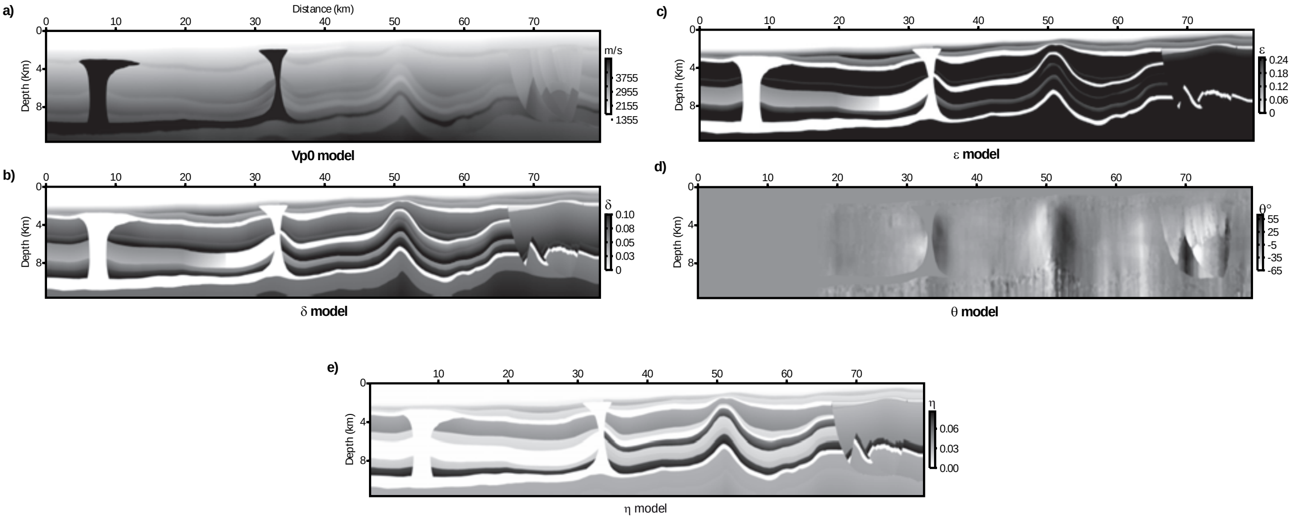


Figure 1.17: The salt models (a) V_{P_0} (b) δ (c) ϵ (d) θ and (e) η model compute from δ and ϵ values.

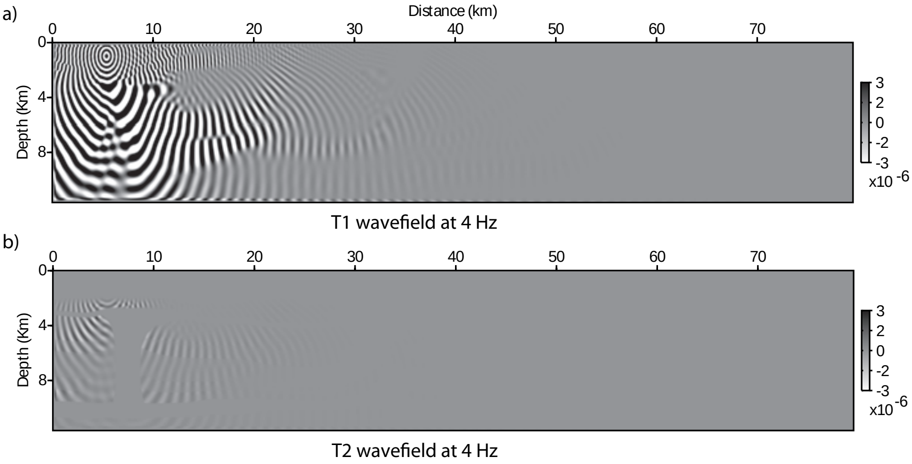
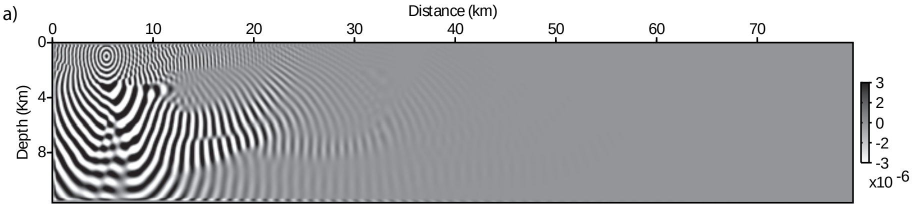


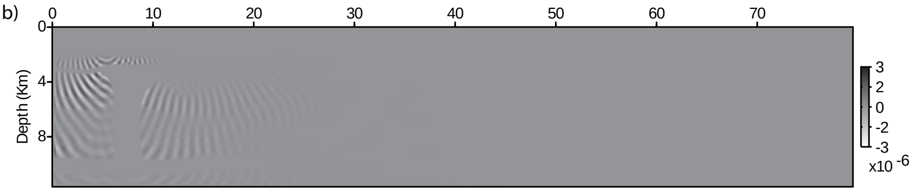
Figure 1.18: The monochromatic (a) T_1 and (b) T_2 pressure wavefiled simulated at 4 Hz frequency of VTI modeling experimental setup.

(figure 1.20). The residual difference is almost negligible, verifying the accuracy of TTI FWM.

As mentioned, the difficulty in TTI FWM raises when the θ parameter class is heterogeneous

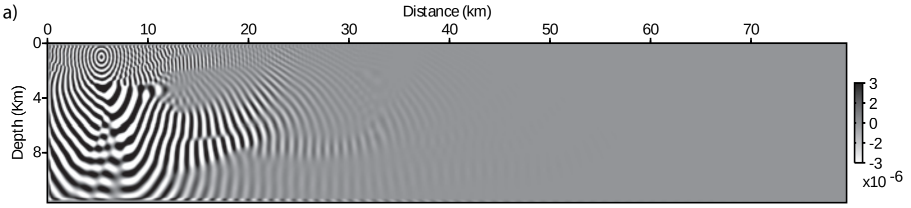


T1 wavefield at 4 Hz

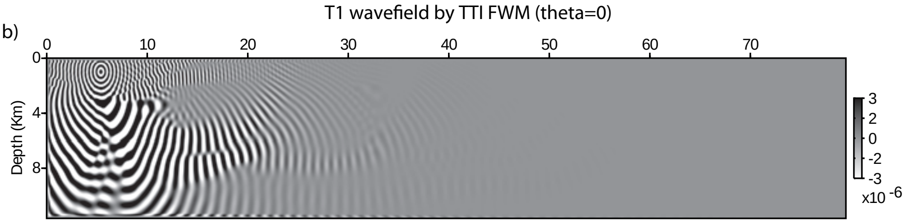


T2 wavefield at 4 Hz

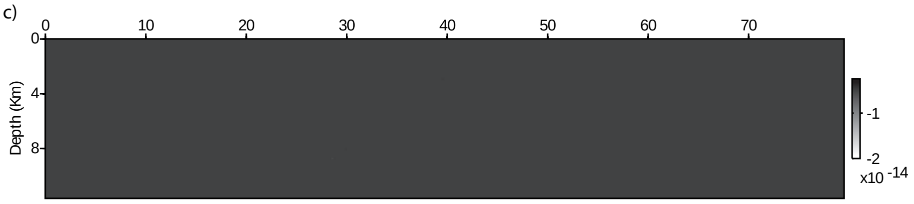
Figure 1.19: The monochromatic (a) T_1 and (b) T_2 pressure wavefiled simulated at 4 Hz frequency of TTI modeling experimental setup. The θ model is set to zero.



T1 wavefield by TTI FWM (theta=0)



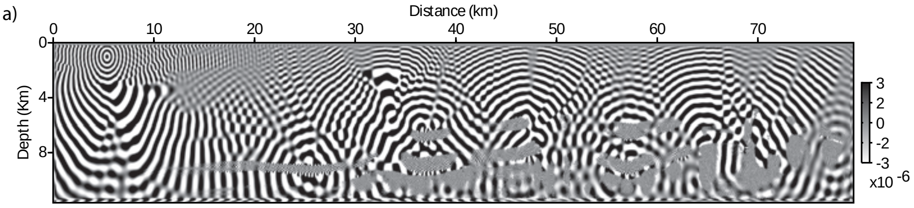
T1 wavefield by VTI FWM



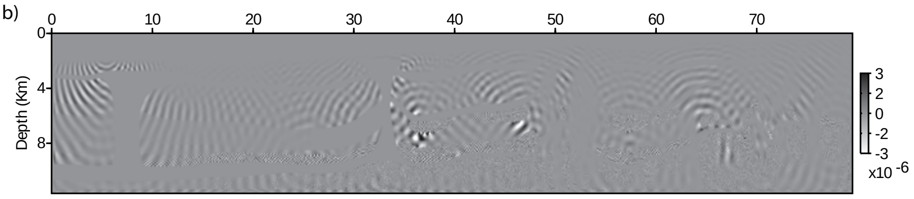
Residual of T1 wavefield between TTI (theta=0) and VTI modeling

Figure 1.20: Comparison between (a) VTI modeling and (b) TTI modeling, setting $\theta = 0$. (c) the difference between these two simulations has very negligible residuals. Demonstrating the accuracy of TTI FWM with $\theta = 0$ in this experimental setup to simulate VTI FWI.

with strong discontinuities, for example on the flancs of salt dome. This is shown in figure 1.21 with the original θ model is included in the modeling. Very high amplitude artifacts are present in the T_1 and T_2 wavefields. These artifacts happen on the discontinuities, which act as highly scatter points and look very much like a strong secondary source.



T1 wavefield at 4 Hz



T2 wavefield at 4 Hz

Figure 1.21: The monochromatic (a) T_1 and (b) T_2 pressure wavefiled simulated at 4 Hz frequency of TTI modeling experimental setup. The high amplitude artifacts act like secondary scattering sources.

These results show that there are instabilities in our TTI FWM. We use the first-order wave equation for TTI FWM, which its differential operator is not self-adjoint (refer to appendix A). The discretization of the second-order wave equation by Discontinuous Galerkin method is a difficult task. These instabilities are related to sharp discontinuities in the θ model. For the moment, the solution of this problem remains under investigation. The idea is to provide a more stable TTI wave equation for modeling with a self-adjoint differential operator.

1.2 Full waveform inversion

Generally, the full waveform inversion (FWI) is represented as a least-squares local optimisation problem. The simplest view of FWI is based on the so-called length method (Menke, 1984). Another representation of inverse problem is performed on probabilistic maximum likelihood or generalized inverse formulation (Menke, 1984; Tarantola, 1984). Lailly (1983) and Tarantola (1984) represented the seismic inversion problem by recasting the migration imaging principle, introduced by Claerbout (1971, 1976), as a local optimisation problem. The reader is referred to Virieux and Operto (2009) for a fulfill and up-to-date overview on the full waveform inversion problem in exploration geophysics.

1.2.1 Local optimization

The generalized inversion problem is based on minimizing the difference between the observed data d_{obs}, and the data calculated in (an) estimated model(s) $d_{cal}(m_{est})$, using the (L_p, p=1,2,∞) norms. The estimated model m_{est} represents some physical parameters. The inverse problem recasts as a local optimisation problem, the optimum solution obtains when the method converges to the global/local minimum. In general, the inverse problem involves optimisation of a misfit functional, which measures the distance between observed and computed data. The misfit functional is defined based on the considered (L_p) norms of data space. In some applications, in addition to data space norm, the model space norm is considered as the second part of the misfit functional. Moreover, the inverse problem is considered as a least-squares criterion, in the manner that it measures the distance between solution and estimated domains in vectorial space. The theory of the generalized inverse problem (as a least-squares optimization), and different methods for model parameter estimation is established by Tarantola (1987). One of the main principle of the least-squares criterion of the misfit function underlies on the hypothesis that all uncertainties in the inverse problem are modelled with Gaussian distributions (Tarantola, 1987). In FWI, the residual or misfit vector $\mathbf{\Delta d}$, is defined as the difference between observed and computed data vectors at the receiver positions for each source (source-receiver pairs). The residual vector is the criteria showing the closeness of the computed data vector to the observed data vector:

$$\Delta d = d_{obs} - d_{cal}(m_{est}). \tag{1.58}$$

In full waveform imaging, the inverse problem represents a nonlinear relation between data and model spaces. Hence, the solution of the inversion is computed with linearized iterative methods. In ideal case, for iterative methods the convergence is achieved when the residual vector goes to zero. The seismic wavefield data in the time domain are represented by real-valued seismogram, and in the frequency domain by complex-valued data for each frequency component. The modelling of the data is performed by full wave equation, which provides a data vector with full information of the subsurface at receiver positions. The full information content of the seismograms in the time domain or complex-valued data in the frequency domain is considered in the optimisation. This leads to a full waveform inversion (FWI) method, which takes into account all of the information content, and optimizes the final model vector by iterative methods.

Let us consider n_s seismic sources and n_r receivers in the frequency domain, then the data vector (for one frequency), has the size of $[n_d = n_s \times n_r]$. The solution of forward problem

$A(m(x),w)u(x,w) = b(x,w)$, is the wavefield of each seismic source in model domain. Mathematically speaking, the data is obtained by applying a sampling operator $\mathcal{P}_{data}$ (corresponding to the receivers) on the solution vector u in model domain ($d_{cal} = \mathcal{P}_{data}u$). The objective of full waveform inversion is to seek the minimum of the misfit function in the vicinity of the initial model m_0 (FWI is a local optimisation problem). The L_2 norm misfit function of data in frequency domain FWI method is defined by the least-squares equation as (Tarantola, 1987):

$$\mathcal{C}(m) = \frac{1}{2} \parallel d_{obs} - d_{cal}(m) \parallel^2_D = \frac{1}{2}(d_{obs} - d_{cal}(m))^\dagger (d_{obs} - d_{cal}(m)) = \frac{1}{2}\Delta d^\dagger \Delta d, \qquad (1.59)$$

where † is the transpose (^{t}) and complex conjugate (*) operators together (the data in frequency domain is complex). The misfit vector components are the difference between each components of observed and computed data. Some weighting (more desirable to be related to the variance of data), may apply on the residual vector to equalize the role of each residual vector components in the misfit function. The covariance between observed and computed data, which shows the variance of the data components is worthy to be preferred. The covariance matrix is a square matrix $[n_d \times n_d]$, with variance components on the diagonal terms corresponding to data components $[n_d \times 1]$.

$$cov(d) \;=\; \begin{bmatrix} \sigma_1^2 & 0 & \cdots & 0 \\ 0 & \sigma_2^2 & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \cdots & 0 & \sigma_{n_d}^2 \end{bmatrix}. \qquad (1.60)$$

The convenient choice of weighing matrix is the inverse of covariance matrix, $W_d = \mathbf{cov(d)}^{-1}$.

$$W_d \;=\; \begin{bmatrix} \sigma_1^{-2} & 0 & \cdots & 0 \\ 0 & \sigma_2^{-2} & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \cdots & 0 & \sigma_{n_d}^{-2} \end{bmatrix}. \qquad (1.61)$$

One should note that, when the weighting matrix is not applied or is equal to identity matrix, it means same weight is applied to all data misfit components. For a seismic wide-aperture/wide-azimuth data, the weighting matrix can be applied based on the offset distance and/or aperture angle. The weighting coefficient on each components of misfit vector can penalise (reward) bad (good) data vector components. The weighing depends on the distribution of the misfit between the observed and computed data vectors. By taking into account the *a priori* information about the data using weighting matrix, the misfit function (equation 1.59) would be as weighted misfit function:

$$\mathcal{C}(m) = \frac{1}{2}\Delta d^\dagger W_d \Delta d. \qquad (1.62)$$

The misfit function changes as function of variation in model parameters. The observed data is always permanent and is observed/acquired from true model. The computed data (of the updated parameters), is the variable. In FWI, the number of estimated parameters is huge, moreover the non-linearity of the problem is a big difficulty of inversion. Therefore, FWI uses linearized and iterative optimisation method.

Lets consider the framework of Born approximation in the scattering theory (see e.g. Born and Wolf (1980); Hudson and Heritage (1981); Beydoun and Tarantola (1988); Beydoun and Mendes (1989); Coates and Chapman (1990); Born and Wolf (1993); Forgues and Lambaré (1997). Suppose the first model of computed data be m^0, then the model m^1 for the first iteration is updated by the model parameter perturbation Δm^0. Then, the updating procedure continues iteratively until the computed model is close enough to the true model (with a specified residual's threshold criteria). At each iteration k, we search for the local minimum of the misfit function $\mathcal{C}(m^k)$ departing from model m^{k-1} of previous iteration:

$$\mathcal{C}(m^k) = \mathcal{C}(m^{k-1} + \Delta m^k). \tag{1.63}$$

This is shown schematically in Figure 1.22. By assuming that the model perturbation vector is small enough with respect to model vector, we can develop the second-order Taylor expansion of misfit function:

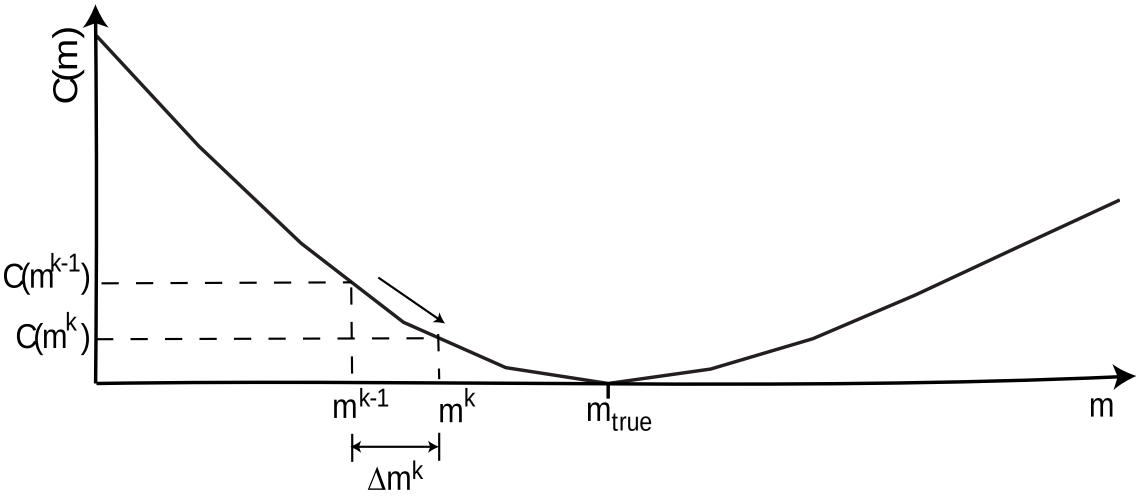


Figure 1.22: When the model parameter $\mathbf{m}^k$ is updated by $\mathbf{\Delta m}^k$ at iteration k, the quadratic misfit function is decreasing toward the minimum. The convergence occurs when the misfit function reaches to its minimum.

$$\begin{aligned}\mathcal{C}(m^{k-1} + \Delta m^k) &= \mathcal{C}(m^{k-1}) + \sum_{i=1}^{n} \frac{\partial \mathcal{C}(m^{k-1})}{\partial m_i} \Delta m_i^k \\ &+ \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} \frac{\partial^2 \mathcal{C}(m^{k-1})}{\partial m_i \partial m_j} \Delta m_i^k \Delta m_j^k + \mathcal{O}(m^3),\end{aligned} \tag{1.64}$$

where n is the size of model parameter vector. Taking the derivative of equation 1.64 with respect to the model parameter leads to:

$$\frac{\partial \mathcal{C}(m^k)}{\partial m_j} = \frac{\partial \mathcal{C}(m^{k-1})}{\partial m_j} + \sum_{i=1}^{n} \frac{\partial^2 \mathcal{C}(m^{k-1})}{\partial m_i \partial m_j} \Delta m_i^k. \tag{1.65}$$

The method of solving the equation 1.65 is known as the Newton method. The optimum solution is obtained, when the quadratic ($\mathcal{O}(m^3) = 0$) misfit function is tracked in the local minimum ($\frac{\partial \mathcal{C}(m^k)}{\partial m_j} = 0$):

$$\frac{\partial^2 \mathcal{C}(m^{k-1})}{\partial m^2}\Delta m^k = -\frac{\partial \mathcal{C}(m^{k-1})}{\partial m}. \tag{1.66}$$

The first order term $\frac{\partial \mathcal{C}(\mathbf{m}^{k-1})}{\partial \mathbf{m}}$ is the gradient vector of the misfit function and the second order term $\frac{\partial^2 \mathcal{C}(\mathbf{m}^{k-1})}{\partial \mathbf{m}^2}$ is the Hessian matrix. Equation 1.66 can be expressed in linear form as:

$$H^{k-1}\Delta m^k = -\mathcal{G}^{k-1}. \tag{1.67}$$

The gradient of misfit function $\mathcal{G}$, is the vector field whose components are the partial derivatives of misfit function $\mathcal{C}$:

$$\mathcal{G} = (\frac{\partial \mathcal{C}}{\partial m_1}, \frac{\partial \mathcal{C}}{\partial m_2}, \cdots, \frac{\partial \mathcal{C}}{\partial m_n}). \tag{1.68}$$

In other words, the negative of the gradient of misfit function ($\Delta\mathcal{C}$) is a vector (a differential operator) representing the direction of descending toward the local minimum. The Hessian matrix is the square matrix of second-order partial derivatives of misfit function:

$$H = \begin{bmatrix} \frac{\partial^2 \mathcal{C}}{\partial m_1^2} & \frac{\partial^2 \mathcal{C}}{\partial m_1 \partial m_2} & \cdots & \frac{\partial^2 \mathcal{C}}{\partial m_n^2} \\ \frac{\partial^2 \mathcal{C}}{\partial m_2 \partial m_1} & \frac{\partial^2 \mathcal{C}}{\partial m_2^2} & \cdots & \frac{\partial^2 \mathcal{C}}{\partial m_2 \partial m_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 \mathcal{C}}{\partial m_n \partial m_1} & \frac{\partial^2 \mathcal{C}}{\partial m_n \partial m_2} & \cdots & \frac{\partial^2 \mathcal{C}}{\partial m_n^2} \end{bmatrix}. \tag{1.69}$$

When the misfit function is continuous, then $\frac{\partial^2 \mathcal{C}}{\partial m_i \partial m_j} = \frac{\partial^2 \mathcal{C}}{\partial m_j \partial m_i}$ i.e. the Hessian matrix, is a symmetric square matrix. The Hessian matrix represents the curvature trend of the quadratic misfit function. The model perturbation vector is the solution of linear system 1.67, which is solved by iterative optimization methods.

Taking the first-order derivative of misfit function (equation 1.59) with respect to model parameter leads to an expression for gradient:

$$\begin{aligned} \mathcal{G}(m) &= -\frac{\partial \mathcal{C}(m)}{\partial m} \\ &= -[(\frac{\partial d_{cal}(m)}{\partial m})W_d(d_{obs} - d_{cal}(m))^* + (d_{obs} - d_{cal}(m))W_d(\frac{\partial d_{cal}(m)}{\partial m})^*] \\ &= -\Re[(\frac{\partial d_{cal}(m)}{\partial m})^\dagger W_d(d_{obs} - d_{cal}(m))] \\ &= -\Re[J^\dagger W_d \Delta d], \end{aligned} \tag{1.70}$$

where $\Re$ and $*$ represent the real part and conjugate of a complex number, respectively. **J** is the Jacobian or Fréchet derivative matrix (referred to Pratt et al. (1998) for more details). Jacobian matrix is the first-order derivative of data vector with respect to model vector:

$$J = \begin{bmatrix} \frac{\partial d_1}{\partial m_1} & \cdots & \frac{\partial d_1}{\partial m_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial d_n}{\partial m_1} & \cdots & \frac{\partial d_n}{\partial m_n} \end{bmatrix}. \tag{1.71}$$

By taking the second-order derivative of misfit function or first-order derivative of gradient with respect to model vector, the Hessian is expressed as:

$$H = -\frac{\partial^2 \mathcal{C}(m)}{\partial m^2} = \Re[J^{\dagger} W_d J] + \Re[\frac{\partial J^t}{\partial m^t} W_d (\Delta d^* \cdots \Delta d^*)]. \tag{1.72}$$

Replacing the gradient (equation 1.70) and the Hessian (equation 1.72) into the Newton equation 1.67, the linear system would be:

$$\Re[J^{\dagger} W_d J + \frac{\partial J^t}{\partial m^t} W_d (\Delta d^* \cdots \Delta d^*)]^{(k-1)} \Delta m^{(k)} = -\Re[J^{\dagger} W_d \Delta d]^{(k-1)}. \tag{1.73}$$

This equation is known as the Newton method.
One should note that, for linear problems (**d=G.m**), the second order derivative of the data with respect to model parameters is zero. Hence, the second term of equation 1.72 on the left-hand side would disappear. The first term of the Hessian is referred to as approximated Hessian.

$$H_a = \Re[J^{\dagger} W_d J] \tag{1.74}$$

The method, which solves model perturbation using the approximated Hessian H_a, is referred to as the Gauss-Newton method. Then equation 1.73 for a linear problem becomes:

$$\Re[J^{\dagger} W_d J]^{(k-1)} \Delta m^{(k)} = -\Re[J^{\dagger} W_d \Delta d]^{(k-1)}. \tag{1.75}$$

1.2.2 Regularization

The FWI is an ill-posed problem (infinite number of models fits the data) and often produces some artifacts and over fitting induced in the retrieved model, due to noise in data or high frequency components of the data. The over fitting generally occurs when a model space is extremely complex, such as an excessively multi-parameter model space (ex. a highly fractured overthrust structure) relative to number of observed data components (ex. purely distributed acquisition array with a weak coverage of offsets and apertures). Imposing certain prior distributions on model parameters brings more conform to the solution of ill-posed FWI problem by controlling the complexity of the solution space (restricting the solution space). On the other hand, applying some a priori information of model vector into misfit function induces an essential trade-off between the data and model spaces. We augment the misfit function of equation 1.62 with the Tikhonov regularization term (the L_2 norm) of model parameter (Tikhonov, 1963; Menke, 1984; Tarantola, 1987; Scales et al., 1990):

$$\begin{aligned} \mathcal{C}(m) &= \frac{1}{2}\left[\| d_{obs} - d_{cal} \|_D^2 + \lambda \| m - m_{prior} \|_M^2\right] \\ &= \frac{1}{2}\left[\Delta d^{\dagger} W_d d + \lambda (m - m_{prior})^{\dagger} W_m (m - m_{prior})\right], \end{aligned} \tag{1.76}$$

where $\lambda \parallel m - m_{prior} \parallel^2$ is the regularization term. λ is the weighting factor (regularization parameter) signifying the trade-off between data and model residuals. The regularization parameter gives relative weight to model optimisation term with respect to data optimisation term. W_d and W_m are the inverse of data and model covariance operators in the frame of Bayesian formulation (Tarantola, 1987). m_{prior} is a reference model, which is usually the initial model in FWI problem. The matrix W_m (equation 1.77) is the weighting operator applied on model space. The W_m (without any smoothing operation) is a diagonal square matrix where the entries on the diagonal are the inverse of covariance in model space:

$$W_m = \begin{bmatrix} \sigma_1^{-2} & 0 & \cdots & 0 \\ 0 & \sigma_2^{-2} & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \cdots & 0 & \sigma_{n_m}^{-2} \end{bmatrix}. \quad (1.77)$$

In seismic imaging, we generally seek the simplest or the smoothest model among all of the models that allow to match the data. In this framework, W_m is a roughness operator, which extract the contrasts of the vector $(m - m_{prior})$ in the misfit function (equation 1.76). Minimization of the second term of the misfit function, equation 1.76, aims to derive the inversion toward smooth models. This kind of regularization is generally referred to as Tikhonov regularization. Note that another kind of regularizations are emerging, that seek to preserve the edge of the subsurface model. These regularization are based on total variation of the model perturbation. They can be implemented with a L_1 norm (Anagaw and Sacchi, 2011) or as a L_2 multiplicative regularization (Abubakar et al., 2011). Other L_1-norm regularization approaches promote the sparsity of the solution in some suitable basis such as wavelet basis (Loris et al., 2007). While these L_1-norm regularizations were originally developed for electromagnetic applications, where the contrasts are sharp, the reliability for the earth model remains an open question.

The physical interpretation of equation 1.76 is that, the optimisation will perform for a model parameter m, which is not far from m_{prior}. If the regularization parameter is very big, the effect of data misfit will be small with respect to model misfit. It means we completely ignore the effect of data and noise in the data and try to minimize the optimization problem only by considering the default solution.

In the data space, one may use the norm L_1 or L_2, but in order to ensure the robustness, the L_2 norm is advised for the model space. Let us look at the regularized FWI problem with another viewpoint. The minimization of

$$\| Am - d \|^2 + \lambda \| m \|^2, \quad (1.78)$$

with respect to m leads to the following generalized normal equations:

$$(A^t A + \lambda I)m = A^t d, \quad (1.79)$$

where $(A^t A + \lambda I)$ is always an invertible matrix. This means, theoretically the addition of the regularization terms lead to an operator, which is invertible. Normally, a priori information is the information that is not already in operator matrix A. As mentioned before, the number of a priori information depends on the number of model parameter involved in inversion process. The minimization of the regularized misfit function for a linear problem change the model perturbation equation 1.75 to:

$$\Delta m = -\Re[J^\dagger W_d J + \lambda W_m]^{-1} \Re[J^\dagger W_d \Delta d + \lambda W_m (m - m_{prior})]. \quad (1.80)$$

The equation 1.80 can be rewritten as:

$$\Delta m = -\Re[W_m^{-1} J^\dagger W_d J + \lambda]^{-1} \Re[W_m^{-1} J^\dagger W_d \Delta d + \lambda (m - m_{prior})]. \quad (1.81)$$

Note that when W_m is roughness operator, W_m^{-1} is smoothing operator.
In case of multi-parameter inversion with regularization terms, the equation 1.76 forms as:

$$\begin{aligned}2\mathcal{C}(m) &= \parallel d_{obs} - d_{cal} \parallel_D^2 + \lambda_1 \parallel m_1 - m_{prior1} \parallel_M^2 + \lambda_2 \parallel m_2 - m_{prior2} \parallel_M^2 + \cdots \\ &= \Delta d^\dagger W_d d + \lambda_1 \Delta m_1^\dagger W_{m_1} \Delta m_1 + \lambda_2 \Delta m_2^\dagger W_{m_2} \Delta m_2 + \cdots \,, \qquad (1.82)\end{aligned}$$

where the model space regularization terms applied for each model parameter class involved

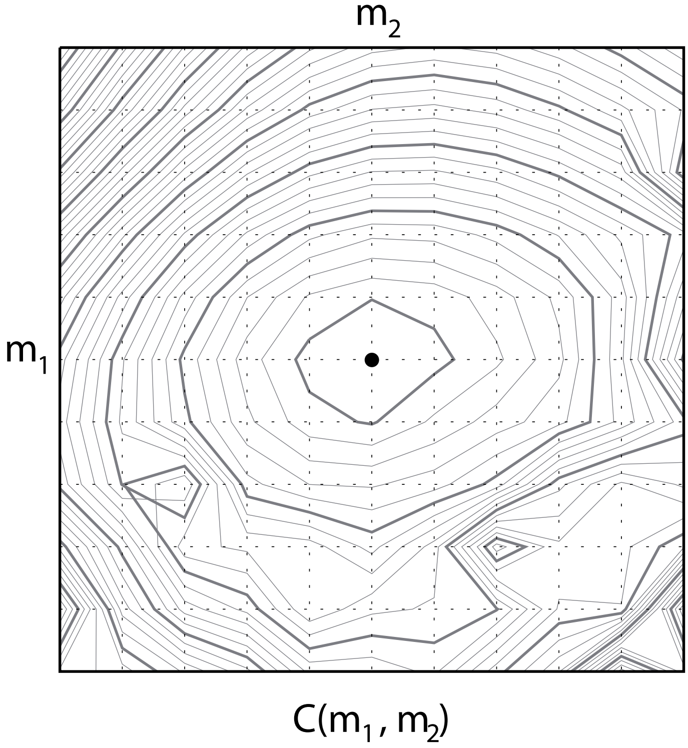


Figure 1.23: The misfit function as a function two model parameter, m_1 and m_2. The center of quadratic misfit function is the minimum (the true models).

in multi-parameter inversion. Figure 1.23 shows an example when the misfit function is a function of two parameters m_1 and m_2. Application of misfit function (equation 1.82) involves the regularization terms for both parameters with their corresponding regularization parameters and model weighting matrices. The difficulty in multi-parameter inversion with regularization terms comes from the choice of regularization parameters λ_i corresponding to each model parameter m_i (the choice of weighting matrices is straight forward). The choice of λ_i is crucial (Hansen, 1998) in the sense that, usually the data set have variable sensitivity with respect to each parameter class.
Figure 1.24 shows another example when the misfit function concerns only the data residual and data has different sensitivity with respect to two parameter classes. For the first parameter (figure 1.24a) the misfit function is well constrained and a small perturbation in model parameter has distinguishable influence on the variation of misfit function. On the other hand, the misfit function of the second parameter (figure 1.24b) is not well constrained, it means a considerable amount of perturbation in the model space is necessary to modify the data residuals. Application of multi-parameter inversion with regularization terms is dependent to the

choice of λ_i to control the influence of model norms in the misfit function and eventually on the final results.

It worth to repeat that the inverse problem is an ill-posed problem, it means many models match the data. The isotropic FWI of anisotropic data is a well suited example, it converges to a model parameter result but this solution does not have minimum length amongst infinite solutions that fit the data. Our results show that, in case of anisotropic data, the isotropic FWI reconstructs a model, which matches the data, but the anisotropic FWI is necessary to converge to a model, which gives the minimum length. Therefore, the anisotropic FWI solution is the proper solution that fits the anisotropic data and also have minimum length. Certainly, the anisotropic FWI solution does not exactly fit the data due to some uncertainties and estimations. The brief conclusion is that, in order to restrict the solution near the desired area of the research and, in order to avoid the effect of noises and artifacts, the choice of regularization terms is highly practical.

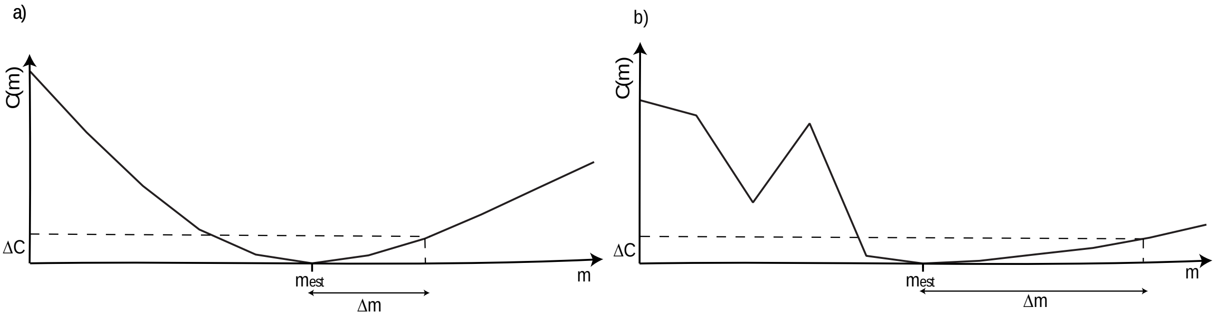


Figure 1.24: The sharper $\mathcal{C}(m)$ has well-defined minimum, which can be concluded that the minimum is well constrained, like (a) which stands for the misfit function curve of vertical velocity. Conversely, if $\mathcal{C}(m)$ has a broad, poorly defined minimum, then the solution is poorly constrained, like (b) which stands for the misfit function curve of Thomsen parameter δ . The estimated solution occurs at minimum of $\mathcal{C}(m)$, if the minimum is wide then a random function in $\mathcal{C}(m)$ lead to big error Δm in estimated solution m_{est}.

In our applications we used the Tikhonov regularization. There are other types of regularization of inverse problem such as total variation (TV) regularization (Rudin et al., 1992; Strong and Chan, 2003), or multiplicative regularization (van den Berg and Abubakar, 2001; van den Berg et al., 2003). The investigation on TV and multiplicative regularization methods is out of scope of this study.

1.2.3 The gradient and Hessian in FWI

The full Newton and Gauss-Newton methods in frequency-space seismic waveform inversion are presented by Pratt et al. (1998), with a clear interpretation of the gradient and Hessian. The gradient $\mathcal{G}$ of the misfit function $\mathcal{C}(m)$ is a vector with the size of model vector n_{mod}, which is the first derivative of misfit function with respect to model parameter m.

The computation of gradient by adjoint-state method

The adjoint-state method is a general method to compute the gradient of a functional that depends on a set of state variables, which are solutions of forward equations (Plessix, 2006).

The theory of adjoint-state method in inverse problem was introduced by Chavent (1974) in order to efficiently compute the gradient of misfit function. The concept of adjoint calculation was introduced by Talagrand and Courtier (1987) for computing the gradient of misfit function in meteorological inverse problem. In geophysics, also, the gradient of the misfit function is computed by adjoint-state method (Chavent and Jacewitz, 1995; Tromp et al., 2005; Plessix, 2006). To compute the gradient of misfit functional by adjoint-state method a new functional called the Lagrangian $\mathcal{L}$ is introduced. The Lagrangian function correspond to the misfit function (equation 1.76) subject to the constraint that the state equations are satisfied. The state equation is the forward problem equation ($Au = s$), which satisfies its solution as the state variables i.e. when $Au - s = 0$. The state variable are (pressure or velocity) wavefields and the d_{cal} at receiver positions. The Lagrangian function is:

$$\begin{aligned}\mathcal{L}\left(u, m, \beta\right) &= \frac{1}{2} < d_{cal} - d_{obs}, d_{cal} - d_{obs} > + \frac{1}{2}\lambda(m - m_{prior})W_m(m - m_{prior}) \\ &\quad - \Re < \beta, Au - s >,\end{aligned} \tag{1.83}$$

where $<,>$ denotes the product of two vectors. The Lagrange multiplier β is the adjoint-state variable (or adjoint field). First, let us define the adjoint operator. Consider a linear operator A that is applied to the vector m and gives the vector d. The adjoint operator $A^{\dagger}$ of A is defined as when:

$$\langle d, Am \rangle = \langle A^{\dagger} d, m \rangle. \tag{1.84}$$

The $\dagger$ is equal to transpose ^{t} and conjugate * together for a complex operator. The aim is to minimize the Lagrangian function (equation 1.83). The minimum must satisfy below conditions:

$$\begin{aligned}\frac{\partial \mathcal{L}}{\partial \beta} &= 0 \\ \frac{\partial \mathcal{L}}{\partial u} &= 0 \\ \frac{\partial \mathcal{L}}{\partial m} &= 0.\end{aligned} \tag{1.85}$$

The first condition gives:

$$\begin{aligned}\frac{\partial \mathcal{L}}{\partial \beta} &= 0 \\ Au - s &= 0.\end{aligned} \tag{1.86}$$

We choose the solution u in a way that this condition is satisfied (at the saddle point of the Lagrangian function). Therefore, at solution u the Lagrangian function $\mathcal{L}$ become equal to misfit function. Reminding that the $d_{cal} = \mathcal{P}_{data} u$, where $\mathcal{P}_{data}$ is the sampling operator at receiver positions. Therefore:

$$\begin{aligned}\frac{\partial \mathcal{L}}{\partial u} &= 0 \\ \mathcal{P}_{data}^{t} \Delta d &= A^{\dagger} \beta.\end{aligned} \tag{1.87}$$

Taking the conjugate of equation 1.87, we have:

$$A^{t} \beta^{*} = \mathcal{P}_{data}^{t} \Delta d^{*}. \tag{1.88}$$

An analogue with the forward problem, the equation 1.88 suggests that $\mathcal{P}^t_{data}\Delta d^*$ (sum of the residuals at receiver positions), is the source of adjoint field β^*. The residuals are back-propagated.
For the third condition, the derivative of $\mathcal{L}$ with respect to m, with inserting the equation 1.88 in the derivative, the gradient of misfit function is:

$$\begin{aligned}
\frac{\partial \mathcal{L}}{\partial m} &= \frac{\partial \mathcal{C}}{\partial m} \\
&= \Re[<\beta, \frac{\partial A}{\partial m} u> + \lambda < 1, W_m(m - m_{prior})>] \\
&= \Re[<(A^{-1})^{\dagger}\mathcal{P}^t_{data}\Delta d, \frac{\partial A}{\partial m} u> + \lambda < 1, W_m(m - m_{prior})>] \\
&= \Re[\{\frac{\partial A}{\partial m} u\}^{\dagger}(A^{-1})^{\dagger}\mathcal{P}^t_{data}\Delta d + \lambda W_m(m - m_{prior})] \\
&= \Re[u^{\dagger}(\frac{\partial A}{\partial m})^{\dagger}(A^{-1})^{\dagger}\mathcal{P}^t_{data}\Delta d + \lambda W_m(m - m_{prior})] \\
&= \Re[u^{t}(\frac{\partial A}{\partial m})^{t}(A^{-1})^{t}\mathcal{P}^t_{data}\Delta d^* + \lambda W_m(m - m_{prior})] \qquad (1.89)
\end{aligned}$$

The equation 1.89 is the gradient of the misfit function.

The physical interpretation of the gradient and Hessian

The physical interpretation of gradient comes from the forward problem $A(m(x),\omega)u(x,\omega)=b(x,\omega)$ expression. Taking the derivative of forward problem equation, for one source and one frequency, with respect to an element of model vector m_i, we obtain:

$$A\frac{\partial u}{\partial m_i} = -\frac{\partial A}{\partial m_i} u. \qquad (1.90)$$

Note that $\frac{\partial b}{\partial m_i} = 0$, because the source vector is not a function of model parameter. The equation 1.90 shows that, the partial derivative wavefield $\frac{\partial u}{\partial m_i}$ multiplied by impedance matrix operator A, is equivalent to a linear problem with the source term as $-\frac{\partial A}{\partial m_i}u$. This source term (known as virtual secondary source) is the production of multiplication of $\frac{\partial A}{\partial m_i}$ by the incident wavefield u. The matrix $\frac{\partial A}{\partial m_i}$ is the partial derivative of impedance matrix with respect to model parameter. The matrix $\frac{\partial A}{\partial m_i}$ is extremely sparse. Physically it is the signature of a diffraction comes from a scatter point (a perturbation point) localized on a spatial position in model m_i. This signature is called the *radiation pattern* of a perturbation point of a model parameter class. The analysis of radiation pattern as a function of incident angle, gives a direct insight into the knowledge of coupling and decoupling between parameter classes, and the influence of a parameter class on the data as a function of scattering angle. The radiation patterns of model parameters induce the opinion about the extent of its reconstruction during FWI (Wu and Aki, 1985b; Tarantola, 1986; Ribodetti and Virieux, 1996; Forgues and Lambaré, 1997).
An example for radiation patterns of partial-derivative wavefiled (PDW) is shown in figure 1.25. Three cases of PDW of scatter point are shown. Figure 1.25a shows a short-aperture PDW scattering, which is expected to have the signature of model parameter perturbation by short-scatter angles. The wide-aperture PDW scattering is shown in figure 1.25b. This means the long-offset acquisition array is necessary to capture the signature of this scattered wavefield.

The favorable PDW is shown in figure 1.25c, where the scattered wavefield has isotropic radiation patterns. Our examples show that the parameter class with isotropic radiation patterns is well retrieved by FWI (next chapter).
Although equation 1.90 leads to a clear physical interpretation of the gradient, the numerical implementation of this equation is impractical. The explicit computation of the sensitivity matrix would require computation of number of model parameters forward problem for each seismic source. This is a huge cost.
Let us go back to the gradient equation computed by adjoint-state method (equation 1.89). The columns of matrix A^{-1} correspond to Green function. When the matrix A^{-1} is symmetric, therefore, $A^{-1^t} = A^{-1}$. Then, the equation 1.89 for parameter class i, is equal to:

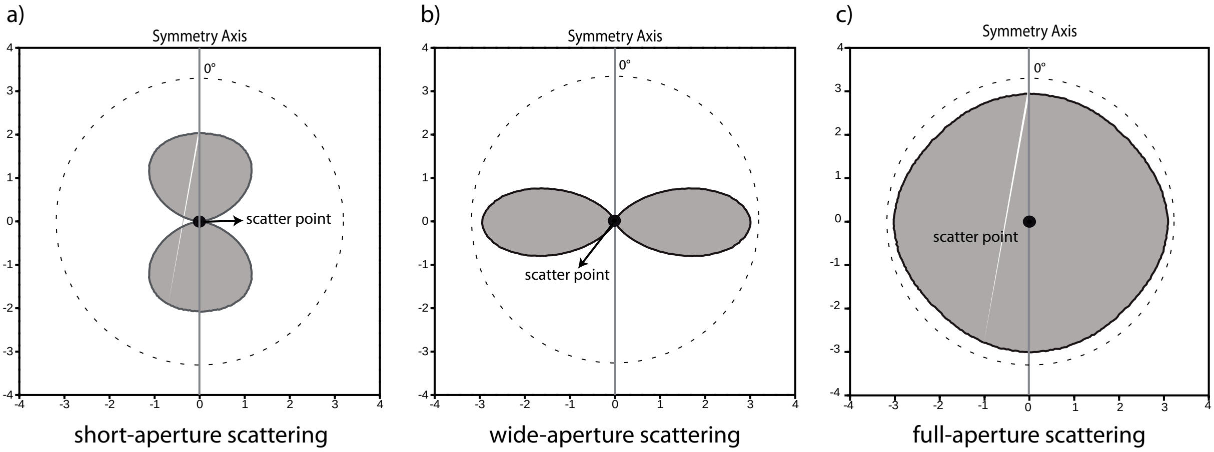
$$\begin{aligned} \frac{\partial \mathcal{C}}{\partial m_i} &= \Re\left\{ u^t \left(\frac{\partial A}{\partial m_i}\right)^t A^{-1}\mathcal{P}^t_{data}\Delta d^* + \lambda_i W_{m_i}(m_i - m_{i_{prior}})\right\} \\ &= \Re\left\{ u^t \left(\frac{\partial A}{\partial m_i}\right)^t \beta^* + \lambda_i W_{m_i}(m_i - m_{i_{prior}})\right\}, \end{aligned} \tag{1.91}$$

where β denotes to the back-propagated residual wavefield. The residual wavefield for every receivers corresponding to each source is back-propagated. In time domain, the residual wavefield is back-propagated by reversing the time. In the frequency domain, the back propagation is indicated by the conjugate operator. The principal of back propagating is similar to reverse time migration (RTM) method. The reader is referred to McMechan (1989) for a review on RTM method in acoustic seismic imaging. The equation 1.91 shows that only two forward modeling is necessary for gradient computation. One is the computation of the incident wavefield and another one is for back propagating the residuals. As is clear, the gradient is computed without explicitly building the PDW matrix. This method is referred to adjoint-wavefield approach, as the adjoint wavefield is built, not the Jacobian matrix (for more details about the adjoint approach refer to Chavent (2009), page 33). The gradient in frequency domain is computed for multiple frequency and multiple sources and receivers, by summing over the number of frequencies, sources:

$$\nabla \mathcal{C}_i = \sum_{k=1}^{N_\omega}\sum_{s=1}^{N_s} \Re\left\{ u^t \left(\frac{\partial A}{\partial m_i}\right)^t A^{-1^t}\mathcal{P}^t_{data}\Delta d^* + \lambda_i W_{m_i}(m_i - m_{i_{prior}})\right\}. \tag{1.92}$$

Note that the matrix A only depends on the frequency and does not depend on the source. Therefore, for each frequency (i.e. solving a monochromatic wavefield) any speedup method can be considered toward solving the multi-shot forward modeling. The gradient equation 1.92 is representing a general form and can be related to either acoustic or elastic wave equations. In acoustic assumption the wavefield is the pressure scalar wavefield. In the elastic medium the wave field is formed by the components of particle velocities and the pressure by first order elastic wave equation. In favorite case, the matrix A is self-adjoint, this means the matrix of the wave equation discretization is symmetric and A^{-1^t} can be replaced by A^{-1} in equation 1.92. In this case, the same modeling engine can be used to compute the incident and adjoint wavefields. In our applications, the gradient of the misfit function is computed with the adjoint-state method (Plessix, 2006), while the action of the Hessian on the gradient is computed by the quasi-Newton l-BFGS method (Nocedal, 1980; Nocedal and Wright, 1999), which provides an approximation of the product of the inverse of the Hessian from a few gradients of the previous iterations.

In FWI, the starting model m_0 (usually a smooth or homogeneous model) misses some



short-aperture scattering wide-aperture scattering full-aperture scattering

Figure 1.25: Three examples of radiation pattern for a scatter point. The black dot line is the scatter point located in center of a homogeneous medium. (a) partial derivative with respect to scatter point model parameter has short aperture angle radiations. This means that the receivers located at wide aperture do not capture any energy from scattering of this point. (b) a wide aperture scattering point. This means that wide aperture surface data or cross-well data will capture the energy of this scattered wavefield. (c) a full and isotropic scattered wavefield. The ideal case, which expected to have a considerable footprint on the data.

heterogeneities of the heterogeneous true model m_{true}. The scatter wavefields are produced due to these missed heterogeneities. Consider a series of closely-spaced large set of point-diffractors, then based on the Huygen's principal, the superposition of images of each point-diffractor produces the image of model perturbation. The same scenario is valid for perturbation wavefield, which is built by superposition of every scattered wavefield of each diffractor point (McMechan and Fuis, 1987).

When the relation between data and model is linear then the inversion converges in one iteration and does not depend on the initial guess (model). This is not the case for seismic inversion where the relation between data and model is non-linear, so the linearization of non-linear problem is the solution and iterative methods helps to accomplish this aim. As mentioned above, the role of Hessian is to apply the curvature trend of misfit function and facilitate the convergence of inverse problem. In addition, the Hessian acts as a tool to scale the gradient by removing the geometric amplitude of the partial-derivative wavefield and the residuals from the gradient. From equation 1.74 the approximated Hessian is formed by zero-lag correlation between partial-derivative wavefields. The zero-lag auto-correlation compose the diagonal terms of the approximate Hessian matrix, therefore represent the square of the amplitude of the partial-derivative wavefield. On the other hand, the off-diagonal terms of the H_a are computed by correlation between PDW of different parameter classes. An example of the approximate Hessian for a simple model shown in Virieux and Operto (2009, their figure 2).

1.2.4 Multi-scale FWI

The non-linearity property of FWI has inspired many hierarchical multi-scale approaches to mitigate this non-linearity. One of the multi-scale approaches is by performing successive inversion of increasing frequencies (it can be seen in the equation 1.92 of the gradient). The experimental setup of the inversion can be performed in single or multiple frequencies. There are three general levels of hierarchical approaches, which mitigate the non-linearity of FWI. The frequency dependant (Sirgue, 2003; Brenders and Pratt, 2007a; Brossier et al., 2009a), the time dependent (Shipp and Singh, 2002; Sears et al., 2008; Brossier et al., 2009a) and the offset dependent (Shipp and Singh, 2002; Wang and Rao, 2009) approaches: each can reduce the non-linearity of FWI. The frequency dependent hierarchical approach is more feasible by the frequency domain FWI (by building a loop over frequencies). On the other hand, time windowing approach is more applicable in time domain FWI. The time damping (time windowing) is possible in frequency-domain FWI by the use of complex-valued frequencies (equivalent to damping seismograms in time by an exponential decay (Shin et al., 2002; Brenders and Pratt, 2007b)). Proceeding from low to high frequencies helps to reduce the risk of cycle skipping. The time windowing allows to limit the inversion to specific arrivals (early arrivals, reflection) or to feed FWI with more complex late-arriving phases. The offset windowing coupled with time windowing allow to implement the layer-stripping approaches.

The multi-scale FWI algorithm (Brossier, 2011a) applied in this study for estimating the model parameter class is shown by algorithm 1.1. In our FWI algorithm, the forward modeling (computation of incident wavefield) is performed numerically with the first-order velocity/stress wave equation (algorithm 1.1,step 5). The first-order differential operators in matrix A are easier to discretize with the discontinuous Galerkin method for forward modeling. Of note, this first-order operator is not self-adjoint. In order to derive the expression of the gradient of the misfit function from the self-adjoint operator, the second-order velocity forward problem

Algorithm 1.1 The algorithm for frequency-domain full waveform inversion

```
1:  for frequency group = 1 to $n_{group}$ do
2:    for time damping = 1 to $n_{damping}$ do
3:      while (NOT convergence AND $n_{it} < n_{it_{max}}$) do
4:        for frequency = low to high do
5:          Compute incident wavefields $u$ from sources (by $1^{st}$-order modeling operator, $Au = s$)
6:          Compute residual vectors $\Delta d$ and cost function $\mathcal{C}^{(n_{it})}$
7:          Compute adjoint back-propagated wavefields from receivers
8:          Compute $\frac{\partial B}{\partial m}$ for parameter class $m$ (by $2^{nd}$-order modeling operator)
9:          Build gradient vector $\mathcal{G}^{(n_{it})}$
10:       end for
11:       Compute diagonal of pseudo-Hessian (Shin et al., 2001)
12:       Compute perturbation vector $\delta\mathbf{m_{n_{it}}}$ by $l$-BFGS Algorithm (Nocedal and Wright, 1999)
13:       Update model $\mathbf{m}^{(n_{it}+1)} = \mathbf{m}^{(n_{it})} + \delta\mathbf{m}^{(n_{it})}$
14:     end while
15:   end for
16: end for
```

operator B is used. Moreover, lower computational memory is necessary for the second-order operator with respect to first-order operator. The self-adjoint property of the operator matrices is more discussed in next chapter for two dimensional VTI acoustic and elastic wave equations.

1.2.5 Parameterization of 2D VTI FWI

The relative influence of the parameter classes on the data has major role in the model estimation by the (non-linear) FWI. In anisotropic FWI, some parameters has higher influence than the others on the data. Sometimes this relative influence is very different for two parameter classes. Usually, the FWI retrieve a better image of the parameter class with high influence on the data. The parameter class with low influence on the data is not likely to be reconstructed by multi-parameter FWI. The alternative solution to this problem is the change of parameterization. Parameterization is defined as a set of independent parameters that fully describe the subsurface properties. By change of parameterization, the relation between data and the parameter classes within new parametrization will change. Below, the different possible parametrization types for 2D anisotropic FWI are shown. The detailed sensitivity analysis of them is discussed in next chapter.

We present the alternative descriptions of 2D VTI model space by means of variable and specific parameter classes (wave speeds,Thomsen parameter, density, ..). We believe that these alternative parameterizations allow to improve the relative reconstruction of them by FWI. This means that, a parameter class within a specific parameterization might be retrieved differently when this parameter class is embedded in another parameterization. The ultimate goal is to improve the feasibility of multi-parameter reconstruction by FWI. The corresponding wave equation for VTI medium is described by stiffness coefficients and the density as physical parameters. But, in seismic exploration the tendency is to describe the medium by physical properties other than stiffness coefficients, such as compressional velocity (P-wave), shear velocity (S-wave), density (ρ), etc. Thomsen (1986) introduced the non-dimensional anisotropic parameters, δ, ϵ and γ, which describe the compressional and shear waves propagation for different angle of incidence for 3D elastic VTI medium. For a two dimensional VTI medium, the Thomsen parameters are related to stiffness coefficients (with weak anisotropic assumption) and are described as :

$$\begin{aligned} V_{P_0} &= \sqrt{\frac{c_{33}}{\rho}}, \\ V_{SV} &= \sqrt{\frac{c_{44}}{\rho}}, \\ \delta &= \frac{c_{13} - c_{33} + 2c_{44}}{c_{33}}, \\ \epsilon &= \frac{c_{11} - c_{33}}{2c_{33}}, \end{aligned} \tag{1.93}$$

where V_{P_0} is the vertical P-wave velocity and V_{SV} is the S-wave velocity along the vertical symmetry axis. Alkhalifah and Tsvankin (1995) introduced the normal move-out velocity (V_{NMO}), and non-ellipticity coefficient η, which provide the possibility to explain the seismic travel-time data for VTI medium with acoustic approximation. The horizontal velocity V_h, is defined as the pressure wavespeed in horizontal direction. These quantities are given by :

$$V_{NMO} = V_{P_0}\sqrt{1 + 2\delta},$$

$$\begin{aligned}
V_h &= V_{P_0}\sqrt{1+2\epsilon},\\
\eta &= \frac{\epsilon-\delta}{1+2\delta}.
\end{aligned} \tag{1.94}$$

In the following, we introduce three types of model space parameterization, which are proposed for 2D VTI FWI in this study. The density (ρ), the quality factor of compression and shear waves attenuations are the common parameters for all types of parameterization plus the shear-wave velocity in elastic medium. We arrange the parameter classes of each parameterization against the stiffness coefficients. In parameterization *type 0* the parameterization is based on 4 stiffness coefficients, $\{c_{11}, c_{13}, c_{44}, c_{44}\}$, which are the original parameter classes in the wave equation. In parameterization *type 1*, the parameterization is based on the shear wave velocity V_{SV}, two Thomsen parameters δ and ϵ or η and a pressure wavespeed, which can be either the vertical velocity V_{P_0}, the horizontal velocity V_h, or the normal moveout velocity V_{NMO}.
The parameterization *type 1-a* is composed of V_{P_0}, V_{SV}, δ and ϵ :

$$\begin{aligned}
c_{11} &= \rho V_{P_0}^2(2\epsilon+1),\\
c_{13} &= \rho V_{P_0}^2(\delta+1) - 2\rho V_{SV}^2,\\
c_{33} &= \rho V_{P_0}^2,\\
c_{44} &= \rho V_{SV}^2.
\end{aligned} \tag{1.95}$$

The parameterization *type 1-b* is composed of V_h, V_{SV}, δ, and ϵ :

$$\begin{aligned}
c_{11} &= \rho V_h^2,\\
c_{13} &= \rho V_h^2\frac{\delta+1}{1+2\epsilon} - 2\rho V_{SV}^2,\\
c_{33} &= \frac{\rho V_h^2}{1+2\epsilon},\\
c_{44} &= \rho V_{SV}^2.
\end{aligned} \tag{1.96}$$

The parameterization *type 1-c* is composed of V_{NMO}, V_{SV}, δ, and ϵ :

$$\begin{aligned}
c_{11} &= \rho V_{NMO}^2\frac{1+2\epsilon}{1+2\delta},\\
c_{13} &= \rho V_{NMO}^2\frac{\delta+1}{1+2\delta} - 2\rho V_{SV}^2,\\
c_{33} &= \frac{\rho V_{NMO}^2}{1+2\delta},\\
c_{44} &= \rho V_{SV}^2,
\end{aligned} \tag{1.97}$$

and the parameterization *type 1-d* is composed of V_{NMO}, V_{SV}, δ, and η :

$$\begin{aligned}
c_{11} &= \rho V_{NMO}^2(1+2\eta),\\
c_{13} &= (\rho V_{NMO}^2)\frac{(\delta+1)}{(2\delta+1)} - 2\rho V_{SV}^2,
\end{aligned}$$

$$\begin{aligned} c_{33} &= \frac{\rho V_{NMO}^2}{(2\delta+1)}, \\ c_{44} &= \rho V_{SV}^2. \end{aligned} \tag{1.98}$$

The third type of parametrisation is composed of shear velocity V_{SV}, Thomsen parameter δ and other two wavespeeds, which can be either the pair of V_{P_0} and V_h (*type 2-a*) or the pair of V_{NMO} and V_h (*type 2-b*). The parameterization *type 2-a* is composed of V_{P_0}, V_{SV}, δ, and V_h :

$$\begin{aligned} c_{11} &= \rho V_h^2, \\ c_{13} &= \rho V_{P_0}^2(\delta+1) - 2\rho V_{SV}^2, \\ c_{33} &= \rho V_{P_0}^2, \\ c_{44} &= \rho V_{SV}^2, \end{aligned} \tag{1.99}$$

and the parameterization *type 2-b* is composed of V_{NMO}, V_{SV}, δ, and V_h :

$$\begin{aligned} c_{11} &= \rho V_h^2, \\ c_{13} &= (\rho V_{NMO}^2)\frac{(\delta+1)}{(2\delta+1)} - 2\rho V_{SV}^2, \\ c_{33} &= \frac{\rho V_{NMO}^2}{(2\delta+1)}, \\ c_{44} &= \rho V_{SV}^2. \end{aligned} \tag{1.100}$$

By a quick insight into the 2-D P-SV VTI wave equation 1.25 of the full waveform modeling engine, one can conclude that the proper class of parameterization is necessary to apply for reconstruction of a parameter except the stiffness coefficients or density from VTI dataset.
One may note that, because of the dependency relation between V_{NMO}, V_h and η (equation 1.101) the conversion of these three parameters to stiffness coefficients, c_{ij} (or even to V_{P_0}, δ, ϵ) is not feasible.

$$V_h = V_{NMO}\sqrt{1+2\eta}. \tag{1.101}$$

Table 1.2 shows the above mentioned types of parameterization for elastic VTI media, which is described by seven parameter classes. Obviously for acoustic approximation (by canceling the shear components), the number of parameter classes reduces to five parameter classes for each parameterization type (table 1.3). Consider an acoustic medium, which is described by five parameter classes (V_{P_0}, δ, ϵ, ρ, q_p). Assume the density and the quality factor of attenuation as the parameter classes, which are fixed and are not updated during the inversion iterations (as this is the case in all over of this study). Then, it remains only three parameter classes, which fully describe the medium. Figure 1.26 shows schematically in vectorial space when we move from V_{P_0}, δ and ϵ parameterization to V_{P_0}, δ and V_h parameterization. This figure shows also metaphorically the change in the role of parameter classes within parameterization by the size of the corresponding model parameter vectors (for example V_h vector has higher length than

parameterization type 0	c_{33}	c_{13}	c_{11}	c_{44}	ρ	Q_p	Q_s
parameterization type 1a	V_{P_0}	δ	ϵ	V_{SV}	ρ	Q_p	Q_s
1b	V_h	δ	ϵ	V_{SV}	ρ	Q_p	Q_s
1c	V_{NMO}	δ	ϵ	V_{SV}	ρ	Q_p	Q_s
1d	V_{NMO}	δ	η	V_{SV}	ρ	Q_p	Q_s
parameterization type 2a	V_{P_0}	δ	V_h	V_{SV}	ρ	Q_p	Q_s
2b	V_{NMO}	δ	V_h	V_{SV}	ρ	Q_p	Q_s

Table 1.2: The three different possible types of parameterization for elastic VTI model space.

parameterization type 0	c_{33}	c_{13}	c_{11}	ρ	Q_p
parameterization type 1a	V_{P_0}	δ	ϵ	ρ	Q_p
1b	V_h	δ	ϵ	ρ	Q_p
1c	V_{NMO}	δ	ϵ	ρ	Q_p
1d	V_{NMO}	δ	η	ρ	Q_p
parameterization type 2a	V_{P_0}	δ	V_h	ρ	Q_p
2b	V_{NMO}	δ	V_h	ρ	Q_p

Table 1.3: The three different possible types of parameterization for acoustic VTI model space.

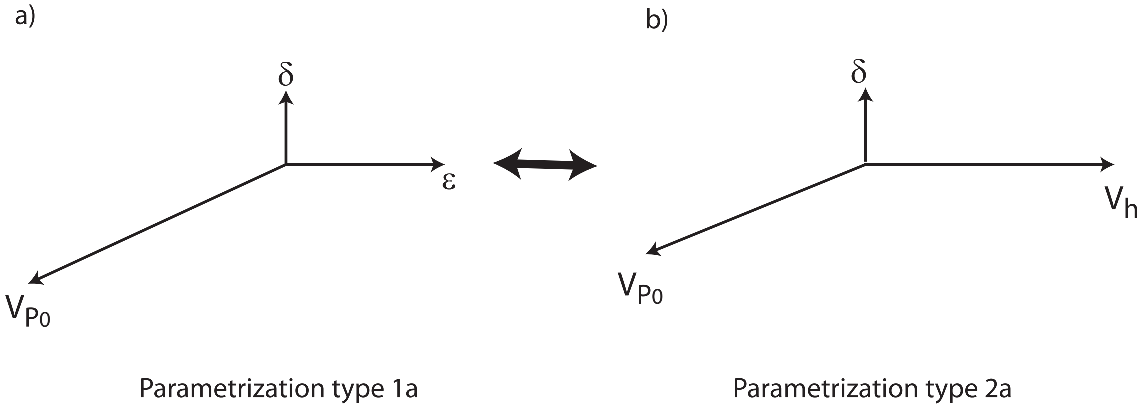


Figure 1.26: Schematic illustration of change of parameterization. We change the parameterization of an acoustic medium from (V_{P_0}, δ, ϵ) parameterization (a), to (b) the parameterization (V_{P_0}, δ, V_h). The reverse action is feasible as well.

ϵ parameter vector). We will explain this aspect of change of parameterization in more detail within next chapter where the sensibility analysis of different parameterizations is investigated. Note that this is just a schematic drawing in order to demonstrate the background idea of parameterization.

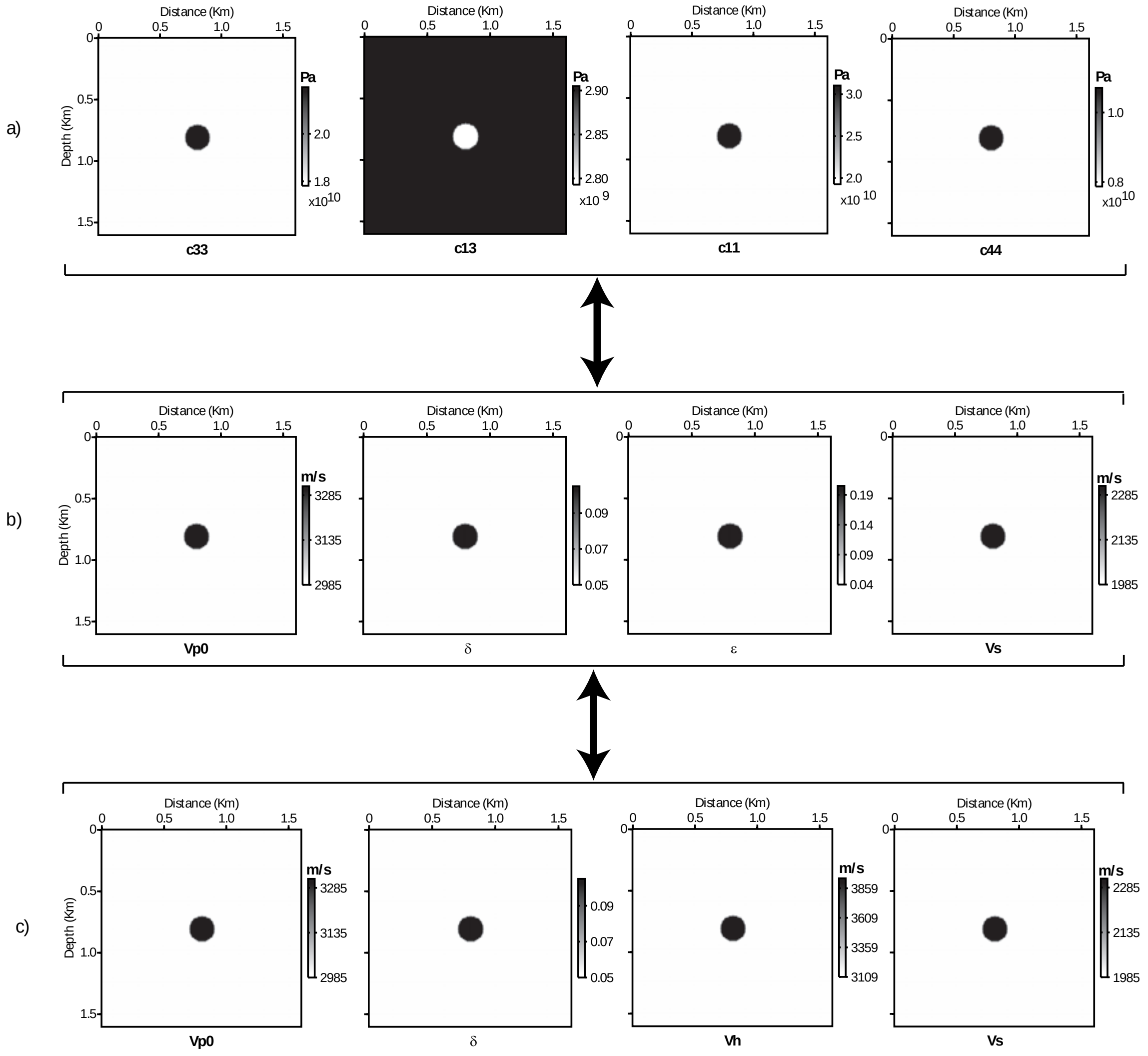


Figure 1.27: Three possible parameterization type of elastic 2D VTI medium shown for the synthetic inclusion model. (a) parameterization of model space by stiffness coefficients, which can be re-parametrized into V_{P_0}, δ, ϵ and V_{SV} parameter classes (b). Likewise, the model space can be parametrized into V_{P_0}, δ, V_h and V_{SV} (c). Density and quality factor are the common parameter classes.

Chapter 2

The Valhall field

Contents

2.1 Introduction

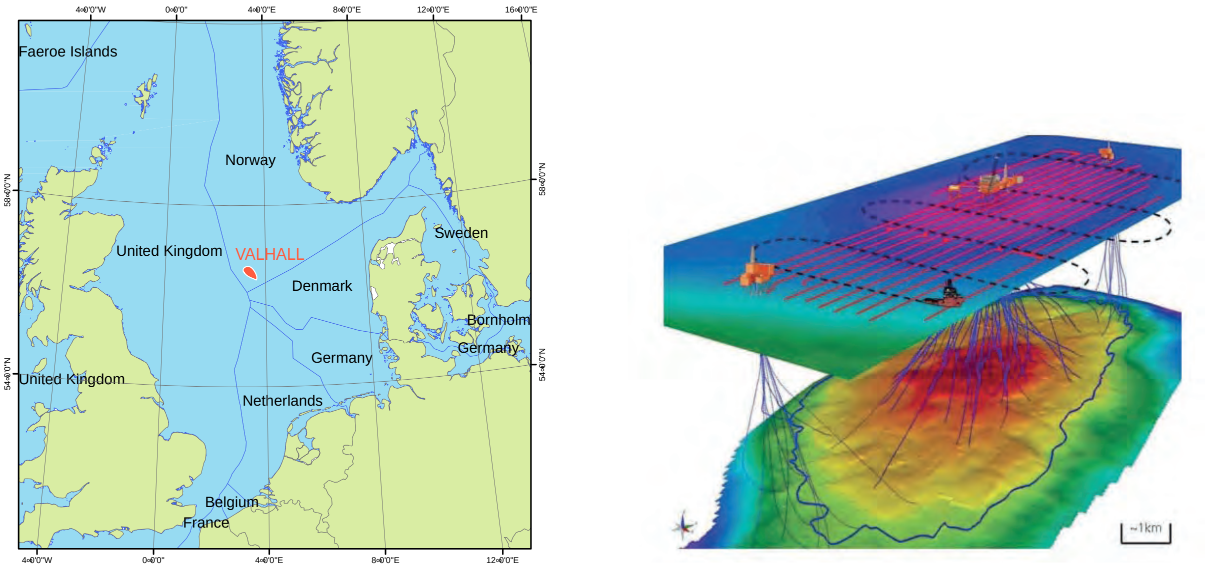
The seismic real dataset considered in this study was recorded in the Valhall oil and gas field, which is located in the North Sea (figure 2.1)in a shallow marine area (69 m is the depth of water). In terms of geological complexity, the field contains only an anticlinal form structure, accompanied with a chain of faults and fracturization. The geological layers are almost horizontal with some weak and strong intrinsic anisotropy, which is considered to be treated as VTI anisotropy in our 2D study. Due to faulting and fracturization of cap rock, the gas content of reservoir is escaped and trapped on upper layers in form of gas-filled layers. These gas-filled layers produce a zone of low-velocity P-wave interfaces, which raise major difficulties on imaging the underneath reservoir (as the target). The origin of anisotropy is mostly intrinsic (Thomsen et al., 1997), due to shaley formations. Intrinsic anisotropy is high in some areas (δ=0.06, ϵ=0.18) and is weak in shallow depth formations, and below the reservoir. A chain of faults on the anticline crest produce dense fracturization, which is a reason for extrinsic anisotropy beside the horizontal layering. Generally, detecting the extrinsic anisotropy due to fracturization depends on the degree of intensity of fracturization and the resolution of imaging the subsurface. Whenever, the size of fracturization is smaller than the resolution of associated dominant wavelengths, the fracturization can be interpreted as weak anisotropy in larger scales. We did not perform any investigation concerning the effect of fracturization on anisotropy of Valhall field in this study. Also the fine layering usually causes some (external) weak anisotropy. This kind of anisotropy is present in Valhall fields originating from interbedded shale (figure 2.2). Gas zone has strong effects on the amplitude decay of P-wave arrivals. Furthermore, ray tracing study shows that rays tend to do not pass through the low-velocity gas cloud (Prieux et al., 2011). Therefore, imaging the reservoir below the gas cloud using only P-wave component is subjected to some difficulties. Shear waves, on the other hand, are not affected by the gas zone. The intuitive idea is to place the receivers on sea bottom in order to capture the converted shear waves and P-wave arrivals. To overcome the challenge of seismic imaging due to gas layers a four components (4C) OBC is performed on the Valhall field. The 4C receivers are planted on the sea bottom as *life of field seismic* (LoFS) projects. Another main benefits of LoFS projects is the 4D study. The 4D seismic response appears to be very sensitive to production-induced depletion, compaction of the reservoir and subsidence of the overburden and seafloor (Barkved and Heavey, 2003).

In this chapter, we introduce the Valhall oil and gas field. The short overview of geology of the field is presented, including the geological structure and lithology of the reservoirs and the cap rock. Then, the previous and current works performed on the seismic data of the field are presented. The application of VTI FWI in more detail is discussed in chapters 4 and 5.

2.2 The geology of the Valhall field

We present a brief recall from the geology of the Valhall field. The Valhall field is a double plunging NNW-SSE trending anticline with an over-pressured, under-saturated Upper Cretaceous chalk reservoir and is located in the most southwestern corner of the Norwegian continental shelf. The field is discovered in 1975 and put in production in 1982. The geographical location of the field is show in figure 2.1a. The top view of the field with a schematic 3D acquisition design is shown in figure 2.1b. From the start of LoFS up to now, several seismic surveys have been acquired over this field (Hatchell et al., 2005).

The stratigraphic log of the flank of anticline is shown in figure 2.2. The lithology from sea



(a)

(b)

Figure 2.1: (a) Geographical location of the Valhall field situated in the North sea on the border of Norwegian territory (31N of UTM European Datum 1950 referential zone). (b) The top view of the Valhall field (van Gestel et al., 2008, figure 1). The red lines show the 4C OBC arrays. The dashed black lines are the cross-lines showing the trajectory of the boat in each survey. The anticlinal form of the reservoir at 3 km depth is shown by surface image. The wells are shown by blue lines, which most of them are drilled on the top of the anticline.

bottom until the erosion surface of unconfomity near 4700 m depth along with the geological time is shown in figure 2.2a. The first layer composed of interbedding of sandstone and shale, which causes a combination of intrinsic (shale) and layering induced (interbedding of shale and sandstone) anisotropy at shallow depth. The second formation is a mixed of siltstone, claystone and shale lithologies, which has the highest (intrinsic) anisotropy values. Claystone interbedded with limestone is the reason of anisotropy like interbedding of sandstone/shale withing fist layer. This formation is the cap rock of below reservoirs. The fracturing is a common phenomenon in limestone formations when confronted with tectonic stresses. The fracturization in limestone (due to faults on the crest of anticline) is one main characteristic of this field. Next formations are the reservoirs of Valhall field, the Upper Cretaceous Tor and the lower Hod formations. Tor is an argilious chalk and Hod has chalk/limestone lithology.

A cross section of the field from WSW to ENE direction is shown in figure 2.2b, which demon-

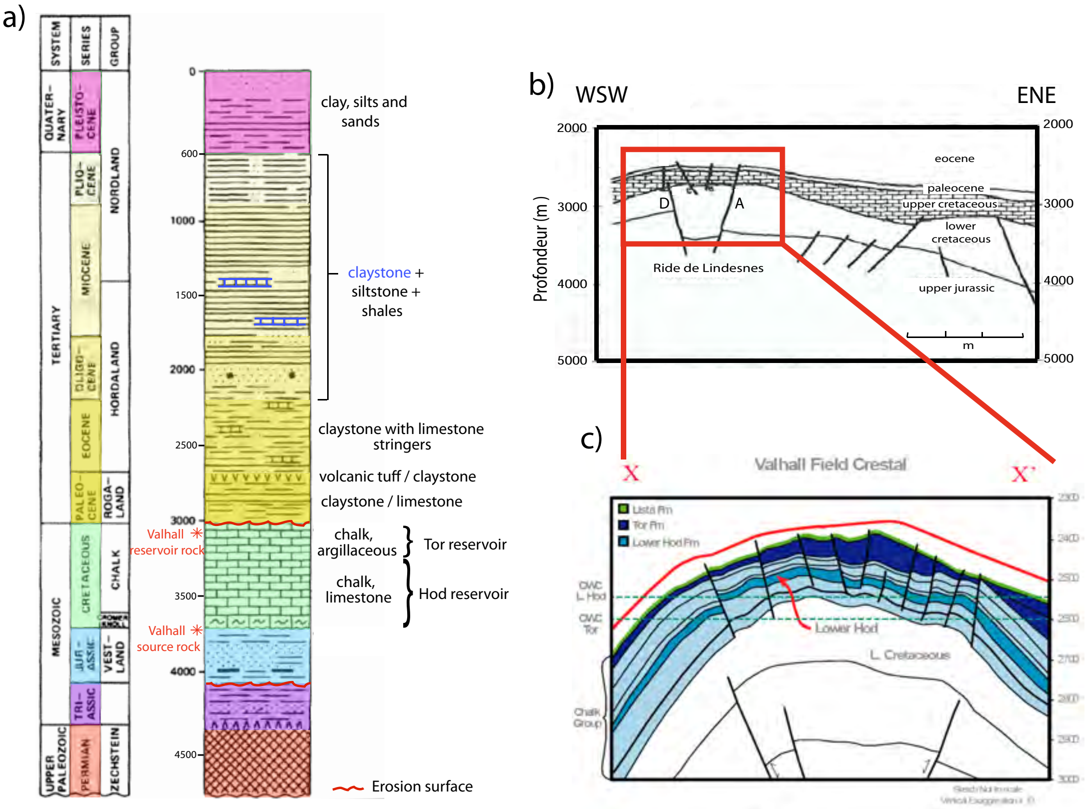
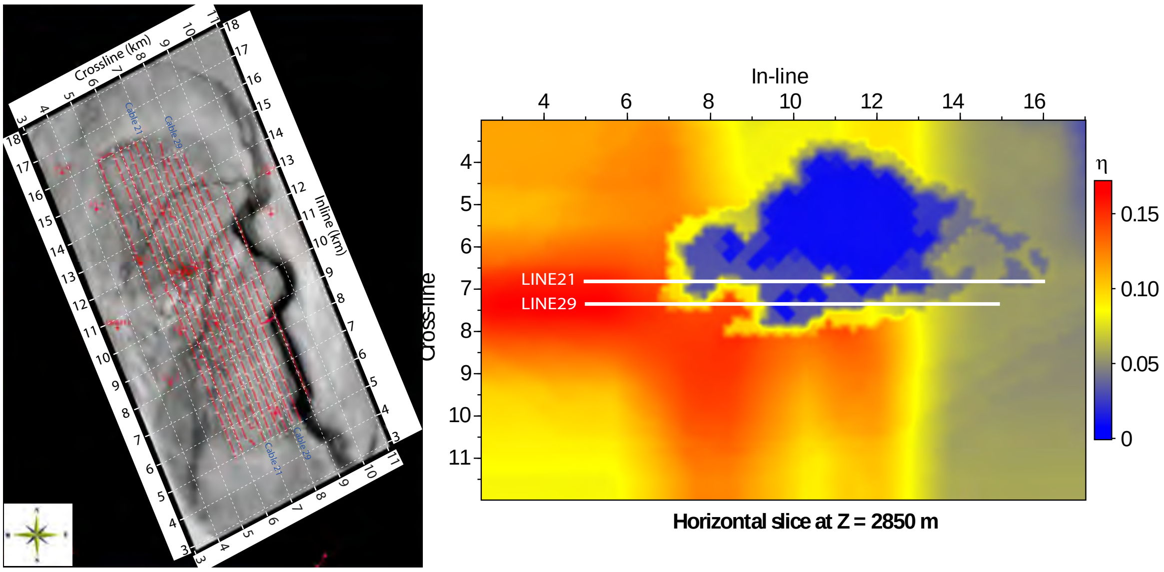


Figure 2.2: (a) The stratigraphic log of the Valhall field after Munns (1985, figure 5), and Olofsson et al. (2003). The log shows the geological formation mostly situated on the flank of the anticline. The cap rock is the Paleocene claystone/limestone formation. Different layers of shale, siltstone, and claystone represent the intrinsic anisotropy, which is treated as VTI anisotropy by 2D FWI. (b) A cross section of the anticline (Munns, 1985, figure 20) with graben structure representing the inversion of tectonic stress during geological periods. (c) The close-up from (b) representing the reservoirs composed of Tor and Hod formations. The reservoirs are partitioned by a series of faults on the crest.

strates the inversion of tectonic stresses during geological period. A zoom of the Tor and Hod

reservoirs shown by figure 2.2c, which illustrates the partitioning of the reservoir in the anticline crest by many faults that produce graben structure on the crest. The fractured zone on the crest is the reason why some gas content of reservoir escaped and made the gas cloud. In consequence, the porosity is exceeding 50 percent in places. Some fractured permeability is present, but matrix permeability is generally low, less than 10 mD. The reservoir is an over-pressured, under-saturated reservoir (Barkved and Heavey, 2003). Preservation of such high porosity is primarily due to over-pressured formation, which inhibits mechanical compaction of the chalk (Munns, 1985). The Tor formation is the primary reservoir and the secondary reservoir is from a unit within the Hod Formation. The thickness of Tor reservoir varies abruptly ranging from 0 to 80 m. Generally, the Tor Formation has higher porosity and permeability and provides the greatest reservoir volume within structural closure. The Hod reservoir is on average about 30 m thick.
Source rock is from the Upper Jurassic Mandal Formation (Kimmeridge Clay). Oil generation occurred during Early Miocene to the present time (Barkved and Heavey, 2003). The field contains paraffinic oil of 36 API gravity (Munns, 1985). As mentioned, the origin of anisotropy in Valhall field is a combination of intrinsic (shale,claystone and mudstone) and extrinsic (faults, fracturization, layering) anisotropy. A depth slice of η (anellipticity parameter) model at 2.850 km is shown in figure 2.3b. The η model shows the degree of anisotropy in this field.



(a) (b)

Figure 2.3: (a) The reference Cartesian map indicating the position of wells in red stars, and OBC receivers in red lines. This map is superimposed on the NW-SE direction and 150m depth horizontal slice of the 3D model obtained by 3D FWI (Sirgue et al., 2010) *(Copyright BP, used by permission)*. (b) The horizontal slice of η obtained by anisotropic reflection tomography. Note the position of lines 21 and 29, which are partly in highly anisotropic area (high values of η).

2.3 Presentation of the Valhall data

The Valhall reservoir is a classic example of a hydrocarbon field, which has difficulties to be imaged only by conventional P-wave techniques. The gas cloud, while present in uneconomic concentrations, has the effect of both lowering seismic velocities and of increasing seismic attenuation (Thomsen et al., 1997). Despite several 3D and 2D seismic surveys (including 1992 streamer survey, 1997 OBC survey, 2002 repeat streamer data, and the permanent OBC surveys that began regular acquisition in 2003 (Hatchell et al., 2005)), P-wave data quality remained poor. This was attributed to the gas cloud in overburden, which leads to a disruption in data continuity over a large area in the center of the field. The loss of P-wave continuity is attributed to attenuation in the gas charged zones resulting in very low P-wave amplitudes and complicated ray bending in the gas zones, effectively creating a complex velocity field, where the geologic structure is flat (O'Brien et al., 1999). The presence of anisotropy, reported by Thomsen et al. (1997), showed that it results in mistie between seismic depth sections and well data. The P-wave data set has been analyzed for amplitude versus offset and azimuth (AVOA) effects. These analyses revealed azimuthal anisotropy effects, which showed strong correlations to existing fault patterns (Barkved and Heavey, 2003). As shear waves are not affected (as like as P-wave) by gas cloud zone, the OBC project was performed to use (converted) shear wave information at receivers located at sea floor. Shear waves are converted waves, as shear waves do not propagate in the water. Converted shear waves are converted reflected energy from down-going compressional (P-wave) seismic energy, at the interfaces in the subsurface (Barkved and Heavey, 2003). The converted wave data has also been analysed for shear wave splitting. Early analysis pointed out the presence of shallow azimuthal anisotropy effects, probably related to subsidence, and detailed analysis demonstrated the orientation and magnitude of the shear wave splitting effects is closely related to the shape of the subsidence bowl at seafloor (Barkved and Heavey, 2003). Shear wave splitting usually is a good index for analysis of fracture orientation. The Valhall field is also an instrumental in developing and demonstrating the viability of imaging below gas using converted shear waves recorded by four component sensors placed on the seafloor (Barkved and Heavey, 2003).

A dense array of seismic OBC equipment (the so-called Life of Field Seismic (LoFS) array) is installed by BP in 2003 at Valhall. The LoFS is composed of 2500 4C sensors attached to 120 km of cables that cover a 45 sq. km area are connected to a recording platform. The OBC operation has been performed fast and easy. The advantage is the constant water depth of 70m in entire survey area and the hard sand ocean floor. Good data were recorded on vertical and two horizontal components and also on the hydrophone. Figure 2.4 shows raw data of gather of reflected events from the 4 components where both P-P and P-S events are clear. In particular a P-S converted event at the target reflector, shows up strongly on the in-line horizontal geophone (Kommedal et al., 1997). The LoFS array permits the detailed observation of field change by regular interval acquisitions during production, providing good information for 4D 4C applications. Other data than 4D data are recorded by the permanent receiver array, such as passive monitoring in periods of platform activities like drilling and injections, and also VSP data (Kommedal et al., 2004). When two or more seismic shots occur at the same time, the recorded data share a time interval, which causes the seismic interference (SI). It means the energy from one source is recorded on the other survey as noise. The characteristics of SI are controlled by the water depth (70 meters $\simeq$ 100 ms tow-way time), water properties, and the nature of the sea-floor (sand). The treatment of SI has been applied on Valhall data set. The grid of 50 m by 50 m for 3D acquisition survey is used. The main primary noise in the receiver

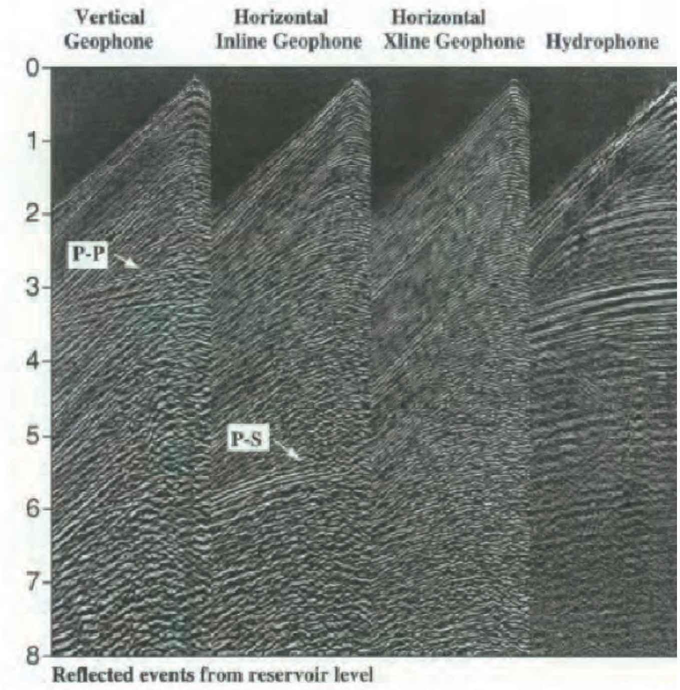


Figure 2.4: A gather of reflected events of 4C OBS survey, which was successful in obtaining the desired P-P and P-S events, from (Kommedal et al., 1997).

array comes from platform, which show up on the three closest cable, and varies a lot with time, and has frequency band more than 50 Hz (Kommedal et al., 2004).

Hall et al. (2002) studied the AVOA analysis on Valhall field seismic dataset with the aim to characterize fracture induced anisotropy throughout the field using surface AVOA data in an effort to gain insight into the 3D fracture distributions (figure 2.5). They focused on the local scale because anisotropy orientation can be sufficiently variable due to perturbation of stress field around faults or across fold structures. They concluded that at Valhall, AVOA analysis interpreted in terms of fracturing, shows spatial variability in both orientation and magnitude that correlates with mapped faults. Comparison of the inferred fracturing with an interpretation of faulting from 3D coherency analysis indicated a degree of correlation with larger scale faulting.

The shear wave splitting analysis is performed by Olofsson et al. (2003) to characterize the azimuthal anisotropy originating from fractures. Shear wave splitting is helpful in anisotropic medium, which faced to fracturization in order to characterize fractures orientations. Shear wave splitting separates the two vertical S_V, and horizontal S_H, shear waves, which are indexed with their polarization and velocities. The azimuthal anisotropy obtained by shear wave splitting method and the subsidence zone are compared in figure 2.5b demonstrating that the azimuthal anisotropy is in agreement with subsidence behaviour on the anticline crest. The OBC acquisition survey provides full-azimuth wide-offset data with high quality and lowest possible frequencies, which make it possible to apply FWI. Sirgue et al. (2010) performed frequency domain acoustic isotropic 3D FWI on Valhall seismic dataset, starting from 3.5 Hz and proceeding up to 7.0 Hz, and the initial velocity model is obtained by conventional tomography (figures 2.3a and 2.6a). To deal with the anisotropy of seismic dataset, the starting model of

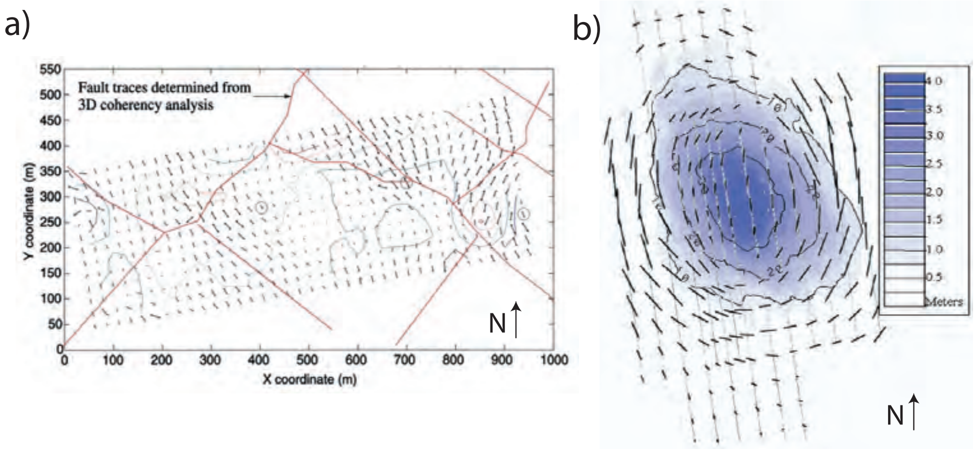


Figure 2.5: (a) The AVOA analysis of fracturization map corresponding to Valhall field by Hall and Kendall (2003). The dark arrows show the direction and amplitude of azimuthal anisotropy. The red line show the fault chain on the top of reservoir. Note that the fractures orientation follow the fault directions. (b) The shallow depth subsidence map is superimposed on the azimuthal anisotropy due to fracturization by (Olofsson et al., 2003, figure 1), which analysis the shear wave splitting in anisotropic medium.

isotropic 3D FWI is produced by converting the anisotopic velocity to an equivalent isotropic one that preserved moveout velocities (Sirgue et al., 2010). The final velocity model of 3D FWI improved migrated seismic images (figure 2.6b). Sirgue et al. (2010) finally showed that imaging the crest by only P-waves is feasible by wide-azimuth wide-offset data, when combined with a better understanding of the geometry of the crestal structure. The P-wave sonic log of well on line 21 is shown in figure 2.6c. Of note, in order to compare the P-wave sonic log with FWI final models, the proper up-scaling on well log is applied.

Figure 2.7 shows a 2D cross-section view of 3D P-wave velocity and anisotropy parameters

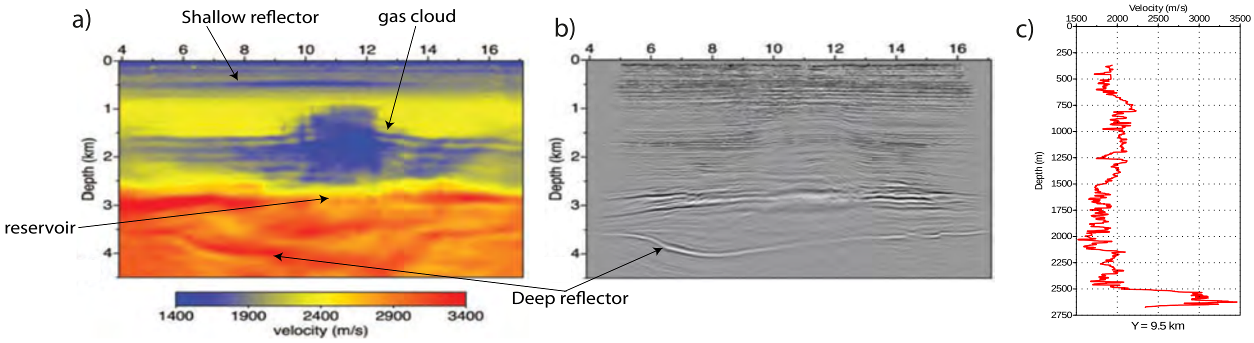


Figure 2.6: (a) Cross-section of the velocity model at position X = 5.575 m obtained by 3D FWI by Sirgue et al. (2010). (b) pre-stack migration image obtained by one-way wave equation calculated from model shown in (a). Note the resolution difference between macro-model obtained by FWI and model obtained by migration. (c) The vertical velocity sonic log placed on line 21 at position (X = 6.8 km ; Y = 9.5 km).

δ and ϵ obtained by reflection tomography, which corresponds to position of line 21 of 3D dataset. The low velocity gas cloud is obvious from 1.5 to 2.5 km depth and looks like a large blob in the velocity model (figure 2.7a). Additionally, the high velocity unconformity zone below the 2.6 km depth is distinguishable. The interface between each anisotropic layer is more clear in Thomsen parameters δ and ϵ models (2.7c-d). Below the reservoir formation is weakly anisotropic, while the most anisotropic layer has claystone, shale, and mudstone lithology between 1.1 and 1.7 km in depth.

The resolution is not high enough to visualise the chain of faults on the crest of the anticline, moreover the anticline flancs are not very deep in this cross section. However, the graben structures are detectable on the crest. The horizontal velocity model (figure 2.7b) is built via vertical velocity and ϵ models, which shows the low velocity gas cloud as a large blob like vertical velocity model.

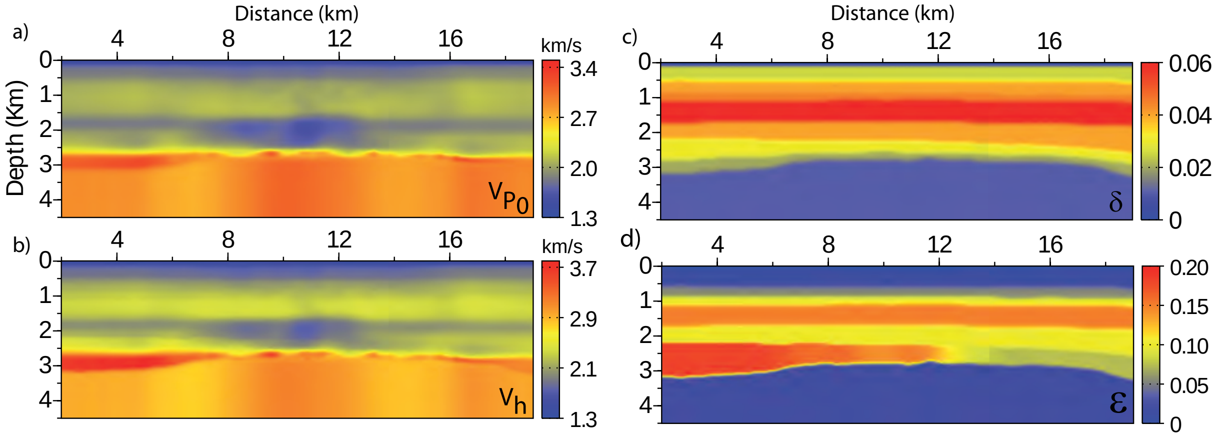


Figure 2.7: Two-dimensional section along position of cable 21 through anisotropic 3D models of the Valhall field. (a) Vertical velocity (V_{P_0}), (b) horizontal velocity (V_h), (c) Thomsen parameter δ, (d) Thomsen parameter ϵ. The V_{P_0}, δ and ϵ models were built by reflection tomography (courtesy of BP).

The acquisition array in 2D (figure 2.8) for computed data (d_{cal}) is composed of 320 sources and 217 receivers spacing each 50 m. The sources are located 6 m below the water surface and receivers on the sea floor representing an OBC array. The maximum offset is up to 13 km. This acquisition array provides the wide-aperture data where the footprint of anisotropy is distinguishable by long distance wave propagation (Prieux et al., 2011). The seismogram of the hydrophone component of the data for line 21 is shown in figure 2.9. The first-arrivals (D_1, D_2), the reflection from shallow reflector (R_s), the reflection from top of gas layers (R_g), the reflection from the top of the reservoir (R_r) and the shingling dispersive guided waves (SW) are indicated, (after (Prieux et al., 2011, their figure 3a)). The R_{gss} and R_{gls} stand for short spread and long spread reflection from top of gas layers. The R_{rss} and R_{rls} stand for short spread and long spread reflection from top of the reservoir. The shot-gather of horizontal and vertical components of the geophone data for this line are shown in figure 2.10. The P-S converted and reflected arrivals are contaminated by the coherent noise of the platform located near to this line. On the ocean bottom, at fluid-solid interface, the Scholte wave (Cagniard, 1962), which travels mainly in water, and the leacky Rayleigh wave (Roever et al., 1959), which travel mainly in the sediment, can be observed (shown by red box in figure 2.10b). The analysis of the fluid/solid interface waves is beyond the scope of this study. The reader is referred to (Carcione and Helle, 2004) for detailed investigation of the physics of wave propagation at the ocean bottom.

In following next chapters, we perform 2D acoustic and elastic anisotropic FWI on hydrophone and geophone components of the dataset, respectively, corresponding to lines 21 and 29. We demonstrate the maximum possible resolution of the recovered images of anisotropic parameter classes. The investigation on the possibility of inverting different anisotropic parameter classes is featured in companion with sensitivity analysis and parameterization concept.

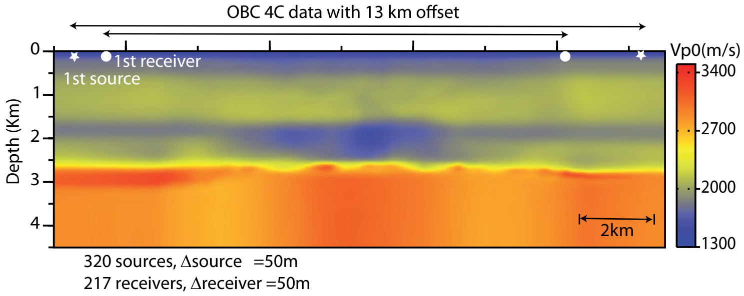


Figure 2.8: The 2D cross-section of the 3D velocity model dataset representing the line 21. The acquisition line is shown in black line for 320 sources and 217 receivers. The maximum offset is 13km with shot and receiver spacing of 50m.

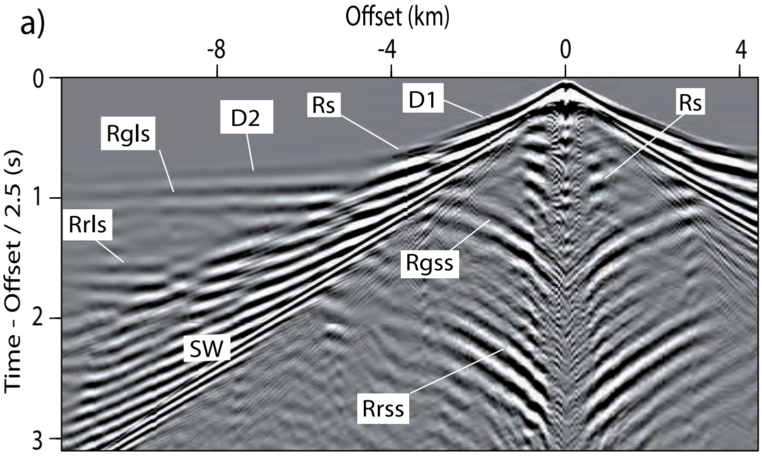


Figure 2.9: The shot-gather of observed data located at line 21 for hydrophone component of the data. The first-arrivals (D_1, D_2) and the main reflection arrivals (R_s, R_g, R_r), are indicated (after (Prieux et al., 2011, figure 3a)).

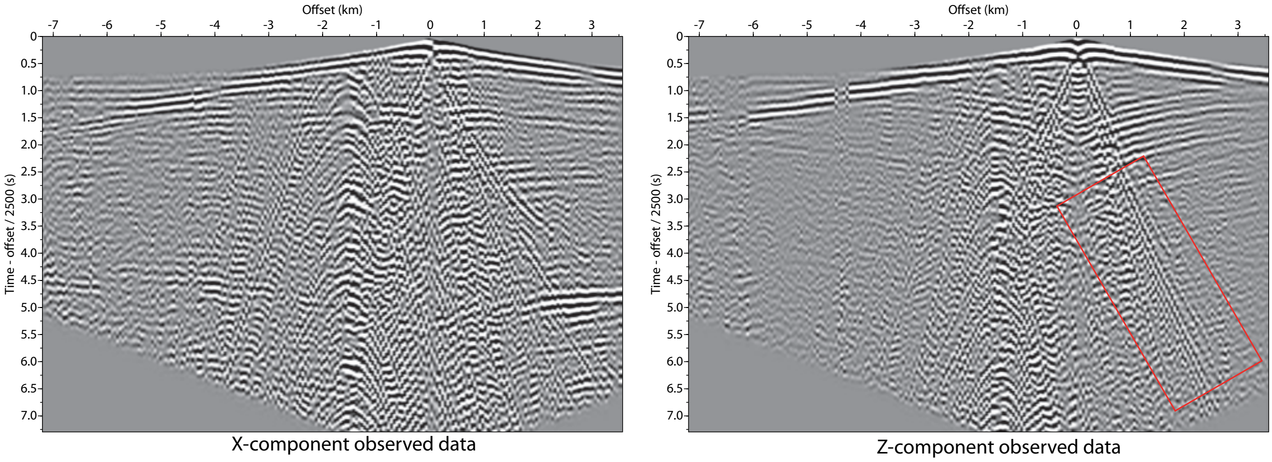


Figure 2.10: The shot-gather of observed data located at line 21 for (a) x-component and (b) z-component of the geophone data. The coherent noises originated from a platform are distinguishable by the hyperbolas on the left side of the source location. The fluid/solid interface (Scholte and Leaky Rayleigh waves), and the shear waves arrivals are highlighted inside the inclined red box (see Carcione and Helle (2004)).

Chapter 3

Sensitivity analysis of acoustic anisotropic FWI

Contents

3.1 Introduction

In this chapter we present a detailed discussion on sensitivity analysis of acoustic anisotropic FWI. The anisotropic medium is the 2D VTI medium. The context contains the article, which has been submitted for publication in *Geophysics* journal. The article is composed of three principal evaluation of sensitivity analysis. An introduction on state of the arts on FWI and specially on some aspects of anisotropic FWI is presented, such as parameterization of model space and partial differential wavefield of scattering heterogeneity. A description of anisotropic FWI method, which is developed in this study is presented plus the theoretical aspects of multi-parameter FWI. The applied wave equation for forward modeling is the first order velocity/stress wave equation, which is discretized with the Discontinuous Galerkin finite element method. The wave equation of partial derivative wavefield as the core of adjoint method is the second order velocity equation. This equation provides a self-adjoint impedance matrix, which facilitates the application of the adjoint-state method. Three investigation of sensitivity and trade-off analysis on possible parameterization types of 2D VTI FWI are performed. The 2D acoustic VTI medium is parametrized by three kinematic parameters without considering the density and attenuation.

Among the possible parameterization, we proposed to investigate three kinds of parameterization: (1) referred to as parameterization of type 1, involves one wavespeed and two Thomsen parameters, (2) referred to as parameterization of type 2, involves two wavespeeds and one Thomsen parameter and (3) referred to as parameterization of type 3, involves the elastic coefficients c_{11}, c_{33} and c_{13}.

The three sensitivity and trade-off analysis consist of (a) the analysis of radiation pattern of virtual sources, (b) the grid analysis of the misfit function, and (c) synthetic examples of full waveform inversion using a simple inclusion model. All three sensitivity analysis are performed for each parameterization type. Therefore, within each parameterization the parameter classes, that can be reconstructed by mono-parameter and/or multi-parameter inversion, are recognized. Moreover, a comparison between each parameterization type is obtained, which leads to know the benefits and drawbacks of each parameterization type.

We computed the monochromatic partial derivative of the wavefield with respect to model parameters in a finite difference sense. The amplitudes of the partial derivative wavefield (PDW) around the parameter of concern give an insight on the sensitivity of the data to the parameter as a function of the scattering angle θ. Later on, the variations of the modulus of the PDW with respect to a parameter class as a function of scattering angle is referred to as the radiation pattern of the parameter class.

The complement of the radiation patterns is the grid analysis of the misfit function. The behaviour of least-squares misfit functional is observed when three parameter's perturbation vary in the center of an elliptic homogeneous background model. This observation leads to understanding the sensitivity of data with respect to each parameter classes within the concerned parameterization type.

In the end, the conclusions of the radiation pattern and the grid analysis are validated against synthetic examples of FWI. The FWI seeks to image an inclusion in a homogeneous elliptic background model from a perfect acquisition geometry surrounding the target. One can decide about using which parametrisation for acoustic anisotropic FWI by these sensitivity analysis investigations and related interpretation of each part. As much as the analysis of the observed data is important, on the other hand, the analysis of response of the physical parameter to the data is also a crucial task. Strictly speaking, when more data are gathered from subsurface, more information can be extracted. The aim of this study is to earn more knowledge of the relations between the seismic data and the physical parameters, in order to be able to extract more accurate information about the model parameters of the subsurface.

3.2 Which parameterization for acoustic vertical transverse isotropic full waveform inversion? - Part 1: sensitivity and trade-off analysis

Which parameterization for acoustic vertical transverse isotropic full waveform inversion? - Part 1: sensitivity and trade-off analysis

Yaser Gholami, Romain Brossier, Stéphane Operto, Alessandra Ribodetti and Jean Virieux

Geophysics, 2012, submitted

3.2.1 Summary

In most geological environments, accounting for anisotropy is necessary to perform acoustic full waveform inversion (FWI) of wide-azimuth and wide-aperture seismic data because of the potential difference between horizontal and vertical wavespeeds. In the framework of multiparameter FWI, the subsurface parameterization controls the influence of the different parameter classes on the modeled seismic data as a function of the scattering angle, and hence the resolution with which the parameters can be reconstructed and the potential trade-off between different parameters. We propose a numerical procedure based on computation of the scattering patterns of the different parameters, on grid analysis of the misfit function, and on synthetic examples of FWI for simple models, to assess the sensitivity of the seismic data to different parameterizations of vertical transverse isotropic acoustic media. Among the different categories we have tested, a mono-parametric FWI is proposed for imaging one wavespeed with a broad wavenumber content, keeping the Thomsen parameters fixed, which have a small influence on the data relative to the wavespeed. This raises the question of the initial information required in the background models of the Thomsen parameters to perform reliable monoparameter FWI. Alternatively, inverting simultaneously the horizontal and vertical wavespeeds introduces limited trade-off effects, as these wavespeeds have significant influence on the data for distinct ranges of scattering angles, while the influence of the Thomsen parameter δ remains weak. As the influence of the wavespeeds on the data is over a narrower range of scattering angles, the resolution is degraded, unless the initial wavespeed models already have the wavenumber content, which cannot be recovered by the multiparameter FWI. We conclude that the choice of the subsurface parameterization can be driven by the acquisition geometry, which controls the aperture-angle coverage and hence the resolving power of FWI, and by the resolution of the available initial FWI models. Indeed, wide-aperture acquisition geometries increase the resolving power of the FWI, and hence provide more flexibility to choose the most suitable parameterization among all available ones.

Keywords

Anisotropy, wave propagation, frequency-domain, full waveform inversion, vertical transverse isotropic

3.2.2 Introduction

The potential of full waveform inversion (FWI) for high-resolution imaging of complex media from low-frequency, wide-aperture/ wide-azimuth data has become apparent over the last decade (e.g., Ravaut et al., 2004; Sirgue et al., 2010; Plessix et al., 2012). FWI seeks to exploit the full information content of the seismic wavefield recorded over a broad range of incidence angles, to build subsurface models with a broad wavenumber content (Virieux and Operto, 2009). Diving waves, pre-critical and super-critical reflections, and diffraction potentially carry information of the subsurface at different resolution powers. Suitable multiscale approaches are generally designed to reduce the non-linearity of the inversion that result from the complex nature of the full seismic wavefield. These multiscale strategies generally rely on a hierarchical inversion of increasing frequencies (Pratt and Worthington, 1990; Bunks et al., 1995), which can be combined with a hierarchical inversion of increasing offsets and traveltimes, or of decreasing aperture angles (e.g., Shipp and Singh, 2002; Brossier et al., 2009b; Shin and Cha, 2009). FWI of wide-aperture and wide-azimuth data raises the issue, however, of the footprint of anisotropy in the imaging, because of the potential variation in the wavespeed with respect to the direction of propagation. Recent case studies of acoustic FWI have shown that interpretable geological structures can be imaged in the isotropic approximation at high resolution even if the subsurface shows significant anisotropy, as in the Valhall field (Sirgue et al., 2010). However, in a Valhall case study, Prieux et al. (2011) highlighted the bias in the velocity estimation that is created when isotropic FWI is performed in anisotropic environments: the imaging can be locally driven towards the reconstruction of the horizontal velocities or the vertical velocities, depending on the local reflection-angle illumination, and this bias in the velocity estimation can be accompanied by significant mispositioning of the reflectors. Other possible artifacts are related to extrinsic anisotropy, which is artificially generated to account for the difference between the horizontal and the vertical wavespeeds (Pratt et al., 2001). Therefore, there is a clear need to better account for anisotropy in acoustic FWI, in particular for the processing of wide-azimuth data (Plessix and Cao, 2011b). In this framework, a suitable parameterization of the subsurface needs to be identified, and the number of parameter classes within the chosen parameterization, which can be reliably reconstructed by FWI, should be defined. Parameterization should be understood as the definition of a set of independent parameters, that fully describe the subsurface properties. For these objectives, we need to assess the influence of each parameter class on the data as a function of the scattering angle, which will in turn provide insights into the trade-off between the parameters and the resolution with which they can be reconstructed. This is the central aim of this study. Sensitivity and trade-off analysis of multiparameter waveform inversion can be performed in the framework of inverse scattering theory, through the analysis of the wavefield scattered by subsurface heterogeneities. This analysis can be performed analytically in the framework of the asymptotic high-frequency approximation, or numerically (the present study). The governing idea relies on the analysis of the scattering (or radiation) patterns of the secondary scattering sources located at the heterogeneity position. The scattering pattern provides clear insights into the influence of the parameter on the data as a function of the scattering angle. Such analysis were presented for elastic media in Wu and Aki (1985a), Wu and Aki (1985b), Tarantola (1986) and Forgues and Lambaré (1997) and for attenuating media in Sato (1984) and Ribodetti and Virieux (1996). Scattering patterns in elastic anisotropic transverse isotropic media were developed in Eaton and Stewart (1994),Burridge et al. (1998), Bostock (2003), Foss et al. (2005) and Calvet et al. (2006), who showed that the scattering pattern depends both on the angle between the incident-wave propagation direction and the symmetry axis, and on the angle between the scattered-wave polarization and the symmetry axis. More qualitative analyses of the sensitivity of the waveform inversion to the parameterization have also been proposed from the numerical computation of the finite-frequency sensitivity kernels of the waveform inversion (i.e., the partial derivative of the wavefield with respect to the model parameters) (Zhou and Greenhalgh, 2009; Sieminski et al., 2009). Other approaches rely on singular value decomposition (SVD) of the Hessian operator. Kiyashchenko et al. (2004) carried out a sensitivity analysis of amplitude-preserving migration to the anisotropic parameters for simple layered examples, to determine which combinations of parameters control the amplitude-versus-offset of the reflections. They have concluded that amplitude-preserving migration is mainly governed by the normal moveout (NMO) velocity and the anellipticity parameter η as already proposed by Alkhalifah and Tsvankin (1995). Parameterization analysis of the acoustic vertical transverse isotropic (VTI) FWI was performed by Plessix and Cao (2011b). The eigenvector decomposition of the Hessian of the FWI misfit function shows that the traveltimes of the diving waves are predominantly sensitive to the horizontal velocity, while the reflection waves at short offsets are predominantly sensitive to the NMO velocity. Moreover, they show from a theoretical viewpoint that the long-wavelength variations of the δ parameter cannot be retrieved from surface seismic data because of the intrinsic ambiguity between δ and depth. The present study is the first of a two-part series that deals with a sensitivity and trade-off analysis of VTI acoustic FWI. Although the aim of this study is similar to that presented by Plessix and Cao (2011b), we shall use different tools to gain more insights into the physical sense underlying the sensitivity of the seismic data to the model parameterization: our sensitivity and trade-off analysis relies on the numerical computation of the scattering patterns of different VTI parameterizations, a grid analysis of the misfit function, and numerical examples of FWI on simple models. The conclusions inferred from this study will be validated against a realistic synthetic example and a real data case study from the Valhall field in a companion report (Gholami et al., 2012a), referred to as Paper II in the following. In the first part of this study, we review the seismic modeling method and the multiparameter FWI algorithm, which we use to perform the FWI applications shown in this report and in the companion report (Gholami et al., 2012a). We also review the relationships between the partial derivative wavefield in FWI and the scattering pattern of their virtual sources. In the second part of this study, we numerically compute the scattering patterns for different VTI acoustic parameterizations. This allows us to discuss the influence of the different parameter classes on the data as a functions of the scattering angle from a theoretical viewpoint. Then, this analysis is validated against a grid analysis of the misfit function. The last part of this study presents the results of multiparameter FWI applied to simple synthetic models that correspond to an inclusion in a homogeneous background. We discuss the consistency between the scattering patterns of the different parameterizations and the results of these numerical experiments, in terms of the spatial resolution, the trade-off between parameters, and the amplitude estimation of the model parameters. The conclusions of this study are validated against a realistic synthetic example that is representative of Valhall, and a real-data case-study from this oil field in the Paper II.

3.2.3 Multiparameter full waveform inversion

We review some key theoretical and algorithmic aspects of multiparameter FWI, which are implemented to perform the numerical experiments of this study and the Valhall case studies of the Paper II.

3.2.3.1 The forward problem

We model the wave propagation in two-dimensional visco-acoustic VTI media with a frequency-domain finite-element discontinuous Galerkin method (Brossier et al., 2008, 2010b; Brossier, 2011a). Although the acoustic VTI wave equation has no physical reality, Operto et al. (2009) showed that the acoustic approximation allows the computing of sufficiently accurate pressure wavefields for FWI in moderate anisotropic media. The subsurface is discretized on unstructured triangular mesh, with piecewise constant medium properties per cell. Absorbing boundary conditions are implemented with perfectly matched layers (Berenger, 1994). We solve the system of linear equations, that results from the discretization of the wave equation in the frequency domain, with the sparse MUMPS direct solver (Amestoy et al., 2001; MUMPS-team, 2009). More details on the implementation of the acoustic VTI wave equation is provided in the Appendix A.

The inverse problem

We seek to minimize a misfit function $\mathcal{C}$ given by:

$$\mathcal{C}(\mathbf{m}) = \frac{1}{2}\Delta\mathbf{d}^{\dagger}\mathbf{W}_d\Delta\mathbf{d} + \frac{1}{2}\sum_{i=1}^{N_p}\lambda_i\left(\mathbf{m}_i - \mathbf{m}_{prior_i}\right)^{\dagger}\mathbf{W}_{m_i}\left(\mathbf{m}_i - \mathbf{m}_{prior_i}\right), \tag{3.1}$$

where $\Delta\mathbf{d} = \mathbf{d}_{cal}(\mathbf{m}) - \mathbf{d}_{obs}$ denotes the data residual vector, the difference between the modeled data $\mathbf{d}_{cal}(\mathbf{m})$ and the recorded data $\mathbf{d}_{obs}$. The conjugate transpose is denoted by the sign $\dagger$. In the present study, we consider only pressure wavefields recorded by the hydrophone component. The multiparameter subsurface model is denoted by $\mathbf{m} = \left(\mathbf{m}_1, ..., \mathbf{m}_{N_p}\right)$, where N_p denotes the number of parameter classes to be up-dated during the FWI. The acoustic VTI medium is parameterized by three classes of parameter, as the vertical velocity and the Thomsen parameters δ and ϵ (Thomsen, 1986), and therefore $N_p \leq 3$. The density is set to be constant, and no attenuation is considered. As FWI is an ill-posed problem, regularization terms are introduced into the model space through the weighting matrix $\mathbf{W}_{m_i}$. These weighting matrices penalize the roughness of the differences between the model $\mathbf{m}$ and the prior model $\mathbf{m}_{prior}$. Data preconditioning can be applied through the matrix $\mathbf{W}_d$, which weights each component of the data misfit vector. The scalar hyper-parameters λ_i, the values of which can be adapted to each parameter class, control the respective weight of the data-space and model-spaced misfit functions in the Equation 3.1.

Minimization of locally quadratic misfit functions gives the Newton descent direction at iteration k:

$$\mathbf{p}_k = -\left[\frac{\partial^2 C(\mathbf{m}_k)}{\partial \mathbf{m}^2}\right]^{-1}\frac{\partial C(\mathbf{m}_k)}{\partial \mathbf{m}}, \tag{3.2}$$

along which the model up-date is searched:

$$\mathbf{m}_{k+1} = \mathbf{m}_k + \gamma_k \mathbf{p}_k. \tag{3.3}$$

This local quadratic approximation of the misfit function requires the estimation of the step length γ_k, which is performed by line search through the parabolic fitting of the misfit function. The first and second derivatives of the misfit function on the right-hand side of the Equation 3.2 are the gradient and the Hessian of the misfit function.

The expression of the descent direction as a function of the sensitivity, or the Fréchet derivative matrix $\mathbf{J}$, is given by:

$$\mathbf{p}_k = \Re\left(\widehat{\mathbf{W}}_m^{-1}\mathbf{J}_k^{\dagger}\mathbf{W}_d\mathbf{J}_k + \widehat{\mathbf{W}}_m^{-1}\left(\frac{\partial \mathbf{J}_k^T}{\partial \mathbf{m}^T}\right)\left(\Delta\mathbf{d}_k^{*}...\Delta\mathbf{d}_k^{*}\right) + \Lambda\right)^{-1}$$
$$\Re\left(\widehat{\mathbf{W}}_m^{-1}\mathbf{J}_k^{T}\mathbf{W}_d\Delta\mathbf{d}_k^{*} + \Lambda\left(\mathbf{m}_k - \mathbf{m}_{prior}\right)\right), \tag{3.4}$$

where Λ is a block diagonal damping matrix:

$$\Lambda = \begin{pmatrix} \lambda_1\mathbf{I}_M & ... & \mathbf{0} \\ ... & ... & ... \\ \mathbf{0} & ... & \lambda_{N_p}\mathbf{I}_M \end{pmatrix}, \tag{3.5}$$

and $\mathbf{I}_M$ is the identity matrix of dimension M, where M denotes the number of nodes in the computational mesh. In Equation 3.4, $\Re$,^{T} and * denote the real part of a complex number, the transpose of a matrix, and the complex conjugate, respectively. The matrix $\widehat{\mathbf{W}}_m$ is a $N_p \times N_p$ block diagonal matrix, where each block is formed by the $\mathbf{W}_{m_i}$ matrices. In the expression of the descent direction, Equation 3.4, the term to be inverted is the full Hessian, which contains three terms. The first one partially corrects the gradient for linear effects, such as the limited bandwidth of the source, the limited spread of the acquisition geometry, and the geometrical spreading of the data. As such, the aim of the Hessian is to retrieve the true values of the model parameters from the misfit-function gradients, which have not the units of these parameters. The second term evaluates the double-scattering effects related to nonlinear propagation effects (Pratt et al., 1998). Finally, the third regularization term damps the deconvolution actions of the first two terms of the Hessian to improve the conditioning of the Hessian matrix. The right-hand-most term in Equation 3.4 is the gradient of the misfit function, which is composed of two terms: the first one represents the contributions of the data and is formed by the zero-lag correlations of the partial derivative wavefields at the receiver positions with the data residuals, while the second term is the model-space regularization term.

It is worth remembering that the partial derivative wavefields satisfy the equation:

$$\mathbf{B}\left(\omega, \mathbf{m}(\mathbf{x})\right)\frac{\partial \mathbf{v}}{\partial m_i} = -\frac{\partial \mathbf{B}\left(\omega, \mathbf{m}(\mathbf{x})\right)}{\partial m_i}\mathbf{v}. \tag{3.6}$$

The right-hand side of Equation 3.6 is the secondary virtual source of the partial derivative wavefield, the spatial support and the temporal support of which are centered on the position of the diffractor m_i and on the arrival time of the incident wavefield at the diffractor m_i, respectively (Pratt et al., 1998). The scattering or radiation pattern of this virtual source is given by $\partial\mathbf{B}\left(\omega, \mathbf{m}(\mathbf{x})\right)/\partial m_i$, and it gives some insight into the sensitivity of the data to the parameter m_i as a function of the scattering (or aperture) angle. The radiation pattern can be computed analytically in the framework of the asymptotic ray+Born approximation (e.g. Wu and Aki, 1985a,b; Forgues and Lambaré, 1997; Ribodetti and Virieux, 1996).

We use the quasi-Newton L-BFGS optimization algorithm for solving the equation 3.4 (Nocedal, 1980; Nocedal and Wright, 1999). The L-BFGS algorithm recursively computes an approximation of the product of the inverse of the Hessian with the gradient, from a few gradients and a few solution vectors from previous iterations. As an initial guess of the inverse of the Hessian, we use a diagonal approximation of the approximate Hessian (the linear term) damped by the Λ matrix,

$$\mathbf{H}_0 = \left(\widehat{\mathbf{W}}_m^{-1} diag \left\{ \mathbf{J}^{(k)^{\dagger}} \mathbf{W}_d \mathbf{J}^{(k)} \right\} + \Lambda \right)^{-1} . \tag{3.7}$$

We minimize the misfit function, equation 3.1, with respect to normalized parameter classes: each class of parameter is divided by the mean value of the parameter in the initial model, such that all of the normalized parameter classes have the same order of magnitude. This parameter normalization behaves as a preconditioning of the L-BFGS optimization, which recursively builds the product between the inverse of the Hessian and the gradient, through a sequence of inner products (Nocedal and Wright, 2006, page 178). This normalization is not neutral in the framework of multiparameter FWI, as it modifies the condition number of the Hessian matrix, as we shall discuss later.

The gradient of the misfit function in equation 3.2 is computed with the adjoint-state method (Plessix, 2006), and is given by (Appendix B):

$$\nabla \mathcal{C}_{\mathbf{m}} = \sum_{\omega} \sum_{s} \sum_{r} \Re \left\{ \left(\mathbf{P} \mathbf{R}_v \left(\mathbf{A}^{-1} \mathbf{s}_A \right) \right)^T \left(\frac{\partial \mathbf{B}}{\partial \mathbf{m}} \right)^T \mathbf{P} \mathbf{R}_v \left(\mathbf{A}^{-1} \frac{b}{\imath \omega} \mathbf{R}_p^T \mathbf{R}^T \left(\mathbf{R} \mathbf{R}_p \mathbf{A}^{-1} \mathbf{s}_A - \mathbf{d}_{obs} \right)^{*} \right) \right\} . \tag{3.8}$$

In equation 3.8, **A** and **B** denote first-order and second-order forward modeling operators, respectively (Appendix A). The self-adjoint operator **B** is used to derive the gradient of the misfit function from the particle-velocity wavefields, while the velocity-stress operator **A** is used to perform seismic wave modeling. The restriction operators $\mathbf{R}_v$ and $\mathbf{R}_p$ extracts the velocity components and the pressure component from the velocity-stress vector, respectively. The restriction operator **R** extracts the values of the pressure wavefield at the receiver positions.

3.2.3.2 Resolution power of full waveform inversion

We explicitly introduce the summation over the frequency, source and receiver in the expression of the gradient of the misfit function, Equation 3.8, to remember the factors that impact the resolution power of FWI (the summation over the receiver should be absent in the framework of the adjoint-state method, because all of the receivers associated with one shot are processed in one go in the source term of the adjoint-state equation). Among others, Wu and Toksöz (1987) and Sirgue and Pratt (2004) showed that in the framework of diffraction tomography, the gradient of the misfit function can be recast as a truncated Fourier series, where the arguments of the basis function are the wavenumber vectors spanned in the subsurface medium. The truncation of the Fourier series, which is controlled by the frequency f and the scattering angle θ bandwidths, limits the resolution with which the subsurface is imaged. The range of scattering angles is itself controlled by the acquisition geometry that is represented by the summations over sources and receiver in equation 3.8. The wavenumber vector **k** is related to the frequency f, the scattering angle θ, and the local wavespeed c by

$$\mathbf{k} = \frac{\omega}{2\,c} cos \left(\theta / 2 \right) \mathbf{n}, \tag{3.9}$$

where **n** is a unit vector in the direction of the vector formed by the sum of the slowness vectors associated with the rays that connect the source and the receiver to the diffractor, and the scattering angle θ is the angle formed by the two slowness vectors (Thierry et al., 1999, their figure 1). Therefore, the limited bandwidth of the source and the limited spread of the acquisition geometry apply bandpass filtering to the wavenumber spectrum of the gradient of the misfit function as shown by Equation 3.9. The radiation pattern of the model parameters, represented by $(\partial \mathbf{B}/\partial \mathbf{m})$ in Equation 3.8, is an additional factor that impacts on the resolution of the FWI, in addition to the frequency bandwidth and the limited spread of the acquisition. As mentioned above, this radiation pattern represents the directivity of the virtual source of the partial derivative wavefield, Equation 3.6. As the gradient of the misfit function is formed by the zero-lag correlation of these partial derivative wavefields at the receiver position with the data residuals, the amplitude variations of the partial derivative wavefields with the scattering angles θ that result from the directivity of the virtual sources also act as a bandpass filtering of the wavenumber spectrum of the gradient, by virtue of Equation 3.9. The radiation patterns of model parameters in multiparameter FWI are generally not isotropic, and therefore they impact significantly on the resolution of the multiparameter imaging, as well as on the trade-off between parameters. This highlights the role of the subsurface parameterization in multiparameter FWI, as we shall show in the following section of this study.

3.2.4 Sensitivity and trade-off analysis of acoustic VTI FWI

An acoustic VTI medium can be parametrized by three kinematic parameters, without considering the density and attenuation. Among the possible parameterization, we propose to investigate here three kinds: the first one, which is referred to as type 1 parameterization, involves one wavespeed and two dimensionless Thomsen parameters; the second one, referred to as type 2 parameterization, involves two wavespeeds and one Thomsen parameter and the last one, referred to as type 3 parameterization, involves the elastic coefficients c_{11}, c_{33} and c_{13}. As velocity, we consider the vertical velocity V_{P0}, the horizontal velocity V_h, and the NMO velocity V_{NMO}, while we shall consider the Thomsen parameters δ and ϵ, as well as a combination of both, which corresponds to the η parameter.

The relationships between these parameters in the acoustic approximation is given by:

$$c_{33} = \rho V_{P0}^2, \quad c_{11} = \rho V_h^2, \quad c_{13} = \rho V_{P_0}^2 \left(\delta + 1\right).$$

$$V_{NMO} = V_{P0}\sqrt{1+2\delta}, \quad V_h = V_{P0}\sqrt{1+2\epsilon} = V_{NMO}\sqrt{1+2\eta}, \quad \eta = \frac{(\epsilon - \delta)}{1+2\delta}. \tag{3.10}$$

3.2.4.1 Radiation patterns of virtual sources

In this section, we want to gain some insights on the sensitivity of the data to different parameter classes, when they are combined each other in the subsurface parameterization. To achieve this goal, we compute the partial derivative of the wavefield with respect to the model parameters, $\partial \mathbf{p}/\partial m$, in a finite-difference sense for the above-mentioned three subsurface parameterizations. From these partial derivatives, we compute the wavefield perturbations $\Delta \mathbf{p}$ resulting from a point model perturbation Δm_i associated with each parameter class i:

$$\Delta \mathbf{p} \approx \left(\frac{\partial \mathbf{p}}{\partial m_i}\right) \Delta m_i \, \bar{\delta}(\mathbf{x} - \mathbf{x}_j). \tag{3.11}$$

where $\bar{\delta}$ denotes the Dirac delta function. The model perturbation Δm_i is defined as a percentage of the value of the parameter in the background model: $\Delta m_i = perc \times m_{0_i}$. The model perturbation is applied on the node j of the finite-difference grid for the parameter class i, and the procedure is repeated independently for each parameter class. The same value of *perc* is used whatever the parameter class. As we are interested only in the relative amplitudes of the wavefield perturbations, *perc* is set to one. In this case, the wavefield perturbations, equation 3.11, are equivalent to partial derivative wavefields with respect to normalized parameters. The normalization consists in scaling each parameter class by its value in the background model, such that all of the parameter classes have the same range of values. Of note, this scaling is also applied to the Thomsen parameters, although they are dimensionless, and it is reminded that this normalization strategy is also used as a preconditioning of the L-BFGS optimization (see section *Multiparameter full waveform inversion*). With this setting, the relative amplitudes of the wavefield perturbations associated with each parameter class should represent the real (i.e., physical) influence of the parameters on the data. The amplitude variation of the wavefield perturbations around the diffractor point Δm_i give some insight into the sensitivity of the data to the parameter as a function of the scattering angle θ. By a slight abuse of language, the variations in the modulus of the perturbation wavefield with respect to one parameter class as a function of scattering angle will be referred to as the radiation pattern of the parameter class.

To compute the partial derivative wavefields in a finite-difference sense, we consider a homogeneous VTI background model defined by $V_{P_0} = 4$ km/s, $\delta = 0.05$, $\epsilon = 0.10$ and small parameter perturbations: $\delta V_{P_0} = 0.2$ km/s, $\delta\delta = 0.05$, $\delta\epsilon = 0.1$. The modeled frequency is 20 Hz. The modulus of the wavefield perturbations for three incidence angles with respect to the vertical symmetry axis are shown in Figures 3.1 and 3.2 for the (V_{P_0}, δ, ϵ) and (V_{P_0}, δ, V_h) parameterizations, respectively. We first show how the radiation pattern can vary with the incidence angle (Calvet et al., 2006) (Figures 3.1 and 3.2, top to bottom panels). For example, the sensitivity of the data to δ and ϵ is quite small in the (V_{P0},δ,ϵ) parameterization when the source is located on the symmetry axis as shown by the numerical noise in Figure 3.1b,c and increases progressively as the incidence angle increases. The same comment applies to δ and V_h for the (V_{P0},δ,V_h) parameterization (Figure 3.2b,c). Secondly, we show how the radiation pattern of one parameter class can vary depending on the other parameter classes involved in the subsurface parameterization. For example, the radiation pattern of the vertical wavespeed does not change with the incidence angle in the (V_{P0},δ,ϵ) parameterization, and spans over the full range of scattering angles (Figures 3.1)a,d,g). A imaging of this parameter with a broad wavenumber content is expected when the (V_{P0},δ,ϵ) parameterization is used. Note, however, that the wavefront of the partial derivative wavefields are not circles like in isotropic media, but have a diamond shape, because of the imprint of the VTI background medium. In contrast, the radiation pattern of the V_{P0} parameter shows some notches as the incidence angle increases, when the (V_{P_0}, δ, V_h) parameterization is used (Figure 3.2g). Significant limited bandwidth effects are expected during the imaging of V_{P_0} for an incidence angle of 90° and scattering angles of 0° and 180°. These range of incidence and scattering angles correspond to normal incidence reflection and pure transmission in the horizontal direction. In this case, the data are sensitive to the horizontal wavespeed only (Figure 3.2i). The opposite scenario is shown for a source located on the vertical symmetry axis, for which the data are sensitive to the vertical velocity only (Figure 3.2a,b,c). We plot the radiation patterns for several parameterizations in a more compact form in Figures 3.3 and 3.4, to facilitate the analysis of the sensitivity of the data to the parameterization. We consider the case of the specular reflection angle: namely, the incidence angle is half of the scattering angle, $\varphi = \theta/2$, when the incidence angle is defined with respect to the vertical and horizontal axes, respectively. The first case (Figure 3.3) is

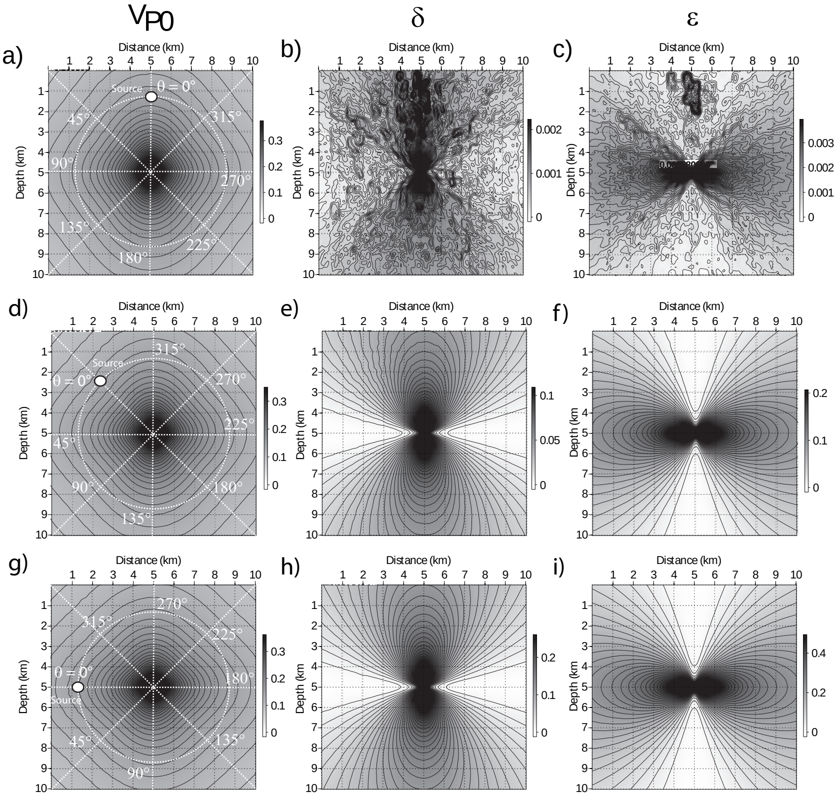


Figure 3.1: Radiation patterns for parameterization (V_{P_0}, δ, ϵ) for incidence angles 0° (a-c), 45° (d-f) and 90° (g-i). Here, radiation pattern refers to the modulus of the partial derivative of a 20-Hz monochromatic wavefield with respect to one model parameter perturbation (see text for more details). These radiation patterns are shown for the V_{P0} (a,d,g), δ (b,e,h) and ϵ (c,f,i) model perturbations, respectively. The incidence angle φ is defined with respect to the vertical symmetry axis, and this defines the source position in the finite-difference modeling, as denoted by the white circle. The diffractor point associated with the model perturbation of each parameter class is located in the center of the finite-difference grid. The source-receiver scattering angle θ is labeled around the dashed white circle for each incidence angle.

representative of reflections from a horizontal reflector recorded by a surface acquisition, while the second case is representative of reflections from a vertical reflector recorded by a vertical source-receiver array, such as vertical seismic profiling (VSP). We show that the radiation patterns in Figure 3.4 are the mirror image of the radiation patterns in Figure 3.3, with respect to the horizontal axis $\theta = 90\,;\,270°$. This is because the direction of propagation is rotated by 90° from Figure 3.3 to Figure 3.4. Therefore, a parameter that shows some influence on the data for a scattering angle θ in Figure 3.3 will show the same influence on the data for a scattering angle of $180° - \theta$ in Figure 3.4. These radiation patterns of Figures 3.3 and 3.4 will be useful in the interpretation of the results of the synthetic experiments shown at the end of this study.

Type 1 parameterizations: one wavespeed + two Thomsen parameters

The specular radiation patterns for the (V_{P0},δ,ϵ), (V_h,δ,ϵ) and (V_{NMO},δ,η) parameterizations

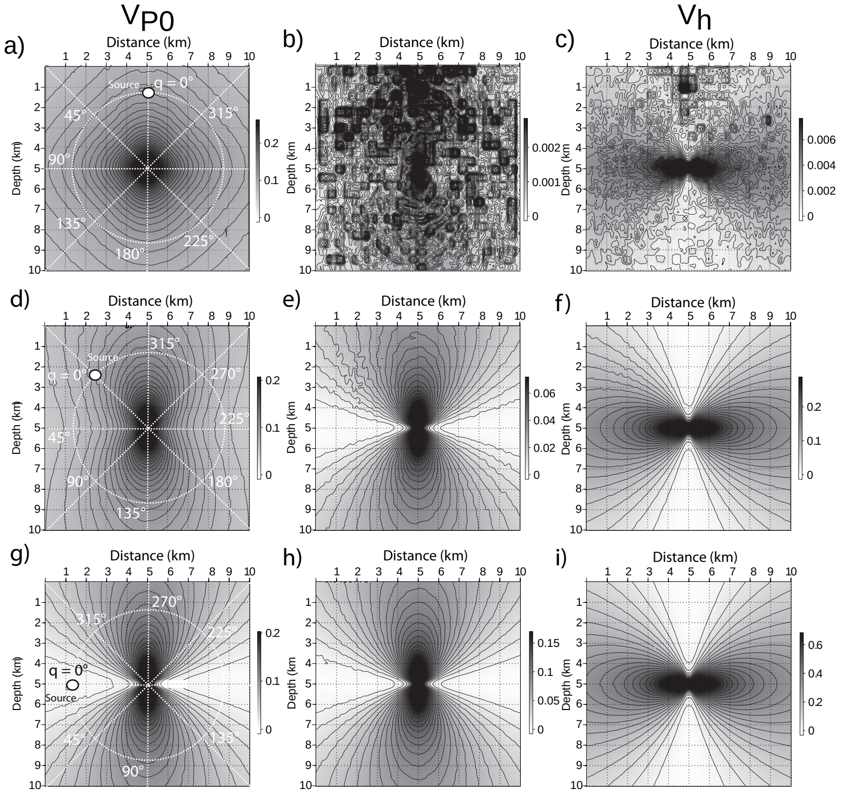


Figure 3.2: As for Figure 3.1, for the (V_{P_0}, δ, V_h) parameterization. The radiation patterns of V_{P_0} (a, d, g), δ. (b, e, h) and V_h (c, f, i). Note how the radiation pattern of V_{P_0} differs from that in Figure 3.1.

are shown in Figures 3.3b-d and 3.4b-d. The first conclusion is that the radiation patterns of the wavespeeds are similar for all of the three parameterizations, and they show significant influence of the wavespeed on the data for all of the scattering angles. Of note, the radiation patterns of the wavespeed are not circular, as they would be in isotropic media, but they show an elongated shape, which results from the footprint of the VTI anisotropy (Figures 3.3b-d and 3.4b-d, gray lines). A reconstruction of the wavespeed with a broad wavenumber content is therefore expected, as the radiation pattern of the wavespeed has high amplitudes whatever the scattering angle is. As the wavespeed parameter has an influence on the data over the full range of scattering angles, the trade-off between the wavespeed and the two other parameters of the parameterization are unavoidable. A second major conclusion is that the amplitudes of the radiation patterns of the Thomsen parameters are much smaller than those of the wavespeed.

When combined with either the vertical or horizontal wavespeeds, the Thomsen parameter δ has a small influence on the data at intermediate aperture angles, and has no influence on the vertical and horizontal wave paths (Figures 3.3b,c and 3.4b,c, solid black lines). The radiation pattern of δ is the same, when the incidence angle is defined with respect to the vertical and horizontal axis (in other words, the radiation pattern of δ is symmetric with respect to the horizontal axis $\theta = 90\,; 270°$). This implies that the image of a horizontal δ reflector built from a surface acquisition will be the same as that of a vertical reflector built from a vertical acquisition. The Thomsen parameter ϵ has an influence on the data for large scattering angles,

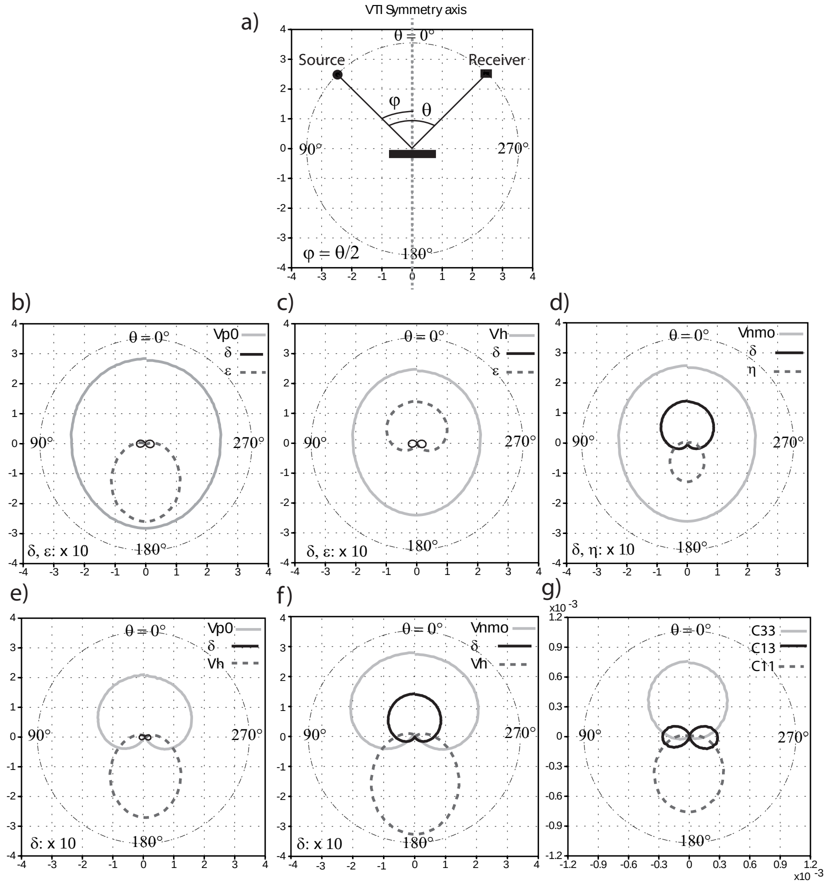


Figure 3.3: Radiation patterns for a specular reflection on a horizontal reflector. (a) Procedure to extract the value of the partial derivative wavefields (Figures 3.1 and 3.2) at a receiver position (black square) for the specular reflection on a horizontal reflector (thick black segment), given the source position (black circle). The incidence angle φ and the reflection angle θ are defined with respect to the vertical axis. This extraction is repeated for a source moving on the dashed circle around the diffractor point, located in the middle of the panel. The assemblage of the values of the partial derivative wavefields at the specular positions for each shot position is shown in (b-g) for different parameterizations. (b-g) Radiation patterns of the (V_{P_0}, δ, ϵ) (b), (V_h, δ, ϵ) (c), (V_{NMO}, δ, η) (d), (V_{P0}, δ, V_h) (e), (V_{NMO}, δ, V_h) (f), (c_{33}, c_{13}, c_{11}) parameterizations with respect to the reflection angle θ as defined in (a). Note that the radiation patterns of the Thomsen parameters are magnified by a factor of 10 in b, c, d, e and f.

with a maximum of sensitivity for the pure transmission regime (θ=180°, φ=90°), when a surface acquisition is considered and the vertical wavespeed is used in the parameterization (Figure 3.3b, dashed line). For the vertical acquisition, ϵ has a maximum influence on the data for small scattering angles (θ=0°, φ=0°) (Figure 3.4b, dashed line). In both case, the maximum influence on the data is shown for the horizontal wave paths, which is consistent with the close relationship between ϵ and the horizontal wave speed. When ϵ is combined with the horizontal wavespeed rather than with the vertical wavespeed, the opposite scenario is shown: the Thomsen parameter ϵ has influence mainly on wave paths propagating near

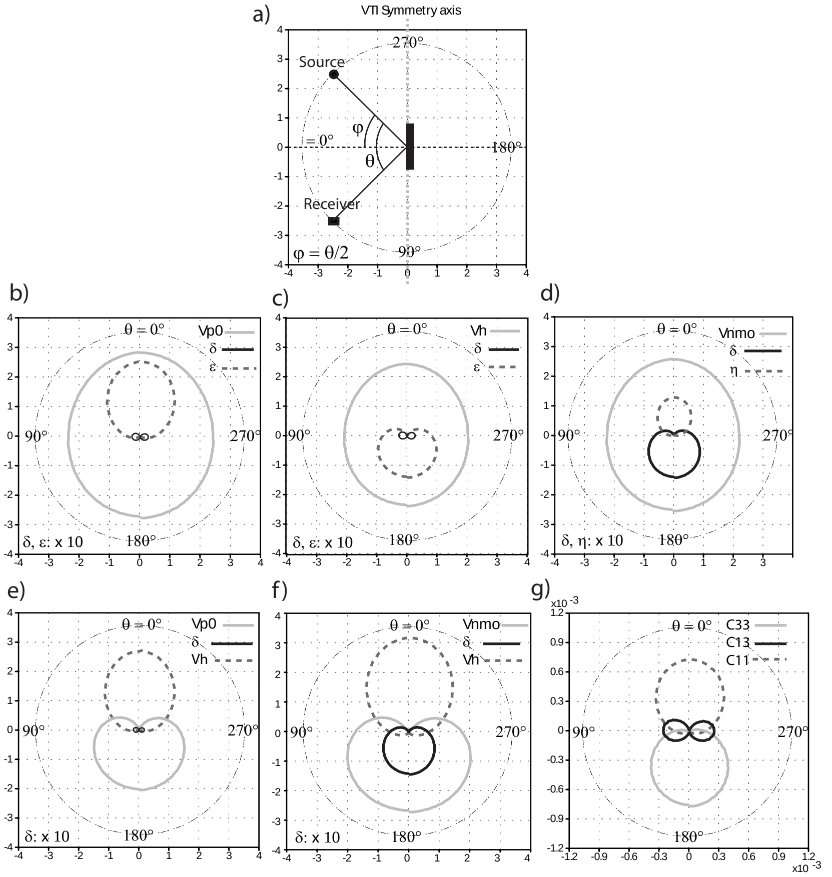


Figure 3.4: As for Figure 3.3, except that specular reflections from a vertical interface (thick black segment) are considered. The incidence and aperture angles are defined with respect to the horizontal axis. Of note, the radiation patterns shown here are the mirror images of those in Figure 3.3.

vertically (Figures 3.3c and 3.4c, dash line).

The different radiation patterns of ϵ in the (V_{P0},δ,ϵ) and (V_h,δ,ϵ) parameterizations, as well as the weaker influence of ϵ on the data relative to the wavespeeds, raise the following comment. If one parameter has a small influence on the data and if some prior information provides benefits for this parameter coming from well or traveltime tomographic methods, only the parameter with the major influence on the data during FWI might be viewed for up-date, keeping the background models of the parameters with the minor influence on the data fixed. Generally, this prior information corresponds to the large wavelengths of the parameter built by traveltime tomography. In this case, it would be preferable to use a parameterization for which the parameter kept fixed during FWI has an influence on the data for large scattering angles because these large scattering angles govern the reconstruction of the large wavelengths of the medium, Equation 3.9. In this case, the influence of the fixed parameter on the data should be predicted sufficiently accurately by the large-wavelength background model. According to this reasoning, the (V_{P0},δ,ϵ) should be chosen for surface acquisition because ϵ scatters energy at wide scattering angles. Indeed, this strategy that consists of keeping fixed some parameter classes during FWI raises the question of the required accuracy of the background models to obtain a reliable model of the wavespeed.

The (V_{NMO},δ,η) parameterization for surface acquisition shows an uncoupled influence of the parameters δ and η at short and wide scattering angles, respectively (Figure 3.3d). Both influences have the same order of magnitude although they remain small relative to that of the NMO velocity. A significant trade-off between the NMO velocity and the two anisotropic parameter δ η is expected with this parameterization.

Type 2 parameterizations: two wavespeeds + one Thomsen parameter

The radiation pattern for the (V_{P0},V_h,δ) and (V_{NMO},V_h, δ) parameterizations are shown in Figure 3.3e,f and 3.4e,f. The radiation pattern of the horizontal wavespeed has a significant influence on the horizontal wave paths, and has the same order of magnitude as the radiation pattern of V_{P0} for vertical wavepaths. Compared to the (V_{P0},δ,ϵ) parameterization, the vertical wave speed has no influence on horizontal wavepaths. Therefore, a limited trade-off between the vertical and horizontal wavespeeds can be expected with the (V_{P0},V_h,δ) parameterization, as these two parameters have radiation patterns that do not overlap significantly. The counterpart is that the range of scattering angles spanned by the radiation pattern of the vertical velocity is narrower than with the previous parameterization. Therefore, a more limited bandwidth reconstruction of the vertical velocity is expected with the (V_{P0},V_h,δ) parameterization. For surface acquisitions, long wavelengths of V_h should be built from the wide scattering angles (i.e., diving waves and super-critical reflections), while the short wavelengths of V_{P0} should be built from the short scattering angles (i.e., short-spread reflections). The radiation pattern of δ shows the same trend as for the previous parameterization, with an even smaller influence on the data (Figure 3.3e). It is unlikely that the δ parameter can be reconstructed by FWI of noisy data, because the influence of this parameter on the data will be dominated by the noise.

It is worth noting that the δ parameter has a stronger influence on the data when it is combined with the NMO velocity rather than with the vertical velocity, with both the (V_{NMO},δ,η) or (V_{NMO},δ,V_h) parameterizations (Figure 3.3d,f). This highlights the trade-off between V_{NMO} and δ, Equation 3.10. A significant trade-off between the NMO velocity and δ is expected with both parameterizations. If δ is kept fixed during FWI, a parameterization that minimizes the influence of δ should be favored, which directs us towards a parameterization involving the vertical velocity rather than the NMO velocity.

Type 3 parameterization: elastic moduli

The radiation patterns for the (c_{33},c_{13},c_{11}) parameterization show a similar trend to the radiation patterns of the previous parameterization (Figures 3.3g and 3.4g), which is consistent with the close relationship between the elastic moduli and the wave speeds, Equation 3.10. The two parameters c_{11} and c_{33}, which are closely related to the horizontal and vertical wave speeds, have partial derivative wavefields of dominant amplitudes relative to the c_{13} parameter. The c_{33} parameter, which is related to the vertical velocity, has a maximum influence on the vertical wave paths, while the c_{11} parameter, which is related to the horizontal velocity, has a maximum influence on the horizontal wave paths. The c_{13} parameter has a maximum imprint in the data at intermediate scattering angles as δ, which is consistent with the relationship between these elastic moduli and the Thomsen parameter δ. As for the vertical and horizontal wavespeeds, the partial derivative wavefields of c_{11} and c_{33} have amplitudes of the same order of magnitude for distinct ranges of incident angles. Therefore, a significant influence of these parameters on the data is expected, as well as a limited trade-off between them. The amplitude of the partial derivative wavefields of the c_{13} parameter is around three-fold less than that of the c_{11} and c_{33} parameters, and shows a greater influence on the data than δ in the two previous parameterizations.

3.2.4.2 Grid analysis of the misfit function

We complement the previous analysis of the radiation patterns with a grid analysis of the misfit function. We consider an anisotropic model corresponding to an inclusion in a homogeneous elliptic background model (Figure 3.5a). An elliptic background model is used to prevent excitation of unwanted shear waves at the source position, which would impact on the values of the misfit function (Grechka et al., 2004). The model space is parametrized by three model parameters, which describe the anisotropic properties of the homogeneous inclusion, the geometry of which is assumed to be known. We analyze the variations of the misfit function for the three kinds of parameterizations described in the previous section, when the values of the three parameters in the inclusion deviate from the true ones. The vertical velocity and the Thomsen's parameters δ and ϵ are (3 km/s, 0.05, 0.05) in the background and (3.3 km/s, 0.1, 0.2) in the inclusion. Nine frequencies between 4.8 Hz and 19.5 Hz are used for the computation of the misfit function. The radius of the inclusion is 300 m. An ideal acquisition is used, where sources and receivers surround the inclusion, hence providing a complete seismic illumination of the target in terms of incidence (φ) and scattering (θ) angles (Figure 3.5a). The maximum deviations from true parameters in the inclusion are ±0.5 km/s, ±0.1 and ±0.2 for V_{P_0}, δ and ϵ, respectively. These values are representative of the model perturbations that are expected to be found in realistic geological targets by FWI. Of note, the same model space as defined by the minimum and maximum values of a set of model parameters (for example, c_{11}, c_{13}, c_{33}) is explored in the following, whatever the parameterization is, hence providing a fair comparison of the different parameterizations.

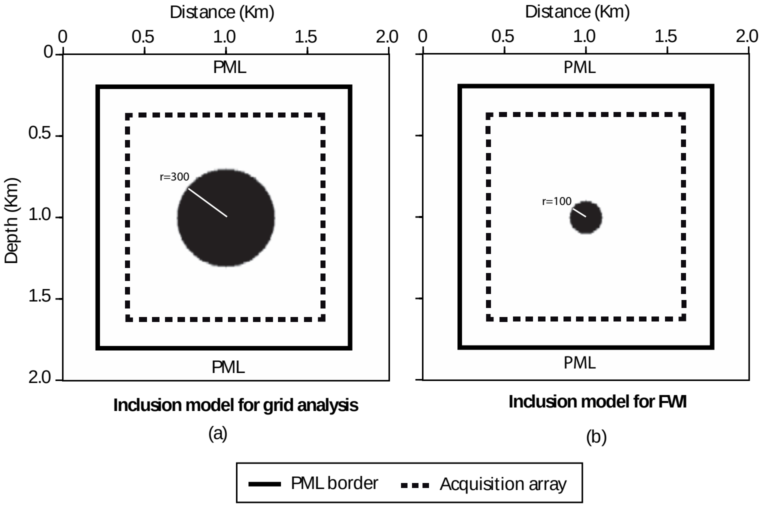


Figure 3.5: Inclusion model parametrized by the anisotropic properties of the inclusion used as the synthetic model for the grid analysis of the misfit function (a), and for VTI FWI (b). Both sources and receivers surround the inclusion, providing complete seismic illumination.

The contours of the misfit function for the (V_{P_0}, δ, ϵ) parameterization are shown in Figure 3.6a,b with their corresponding one-dimensional profiles shown in Figure 3.7a,b. The misfit function in the $(V_{P_0}, \epsilon)_{\delta=0.1}$ plane shows that V_{P_0} has a greater influence on the data than ϵ, as the contours of the misfit function become tighter and tilt towards the ϵ-axis (Figure 3.6a). The misfit function in the $(V_{P_0}, \delta)_{\epsilon=0.2}$ plane shows that the influence of δ is negligible, as the contours of the misfit function are parallel to the δ-axis (Figure 3.6b). The presence of local minima are shown in the profiles of the misfit function plotted as a function of δ for the true values of V_{P0} and ϵ (Figure 3.7b, lower panel). The dominant influence of V_{P_0} relative to ϵ and more significantly to δ is consistent with the radiation pattern analysis presented in the previous section, Figure 3.3b. The misfit function for the (V_{P_0}, V_h, δ) parameterization is shown in Figure 3.6c,d with their corresponding one-dimensional profiles shown in Figure 3.7c,d. The misfit function in the $(V_{P_0}, V_h)_{\delta=0.1}$ plane shows that the influences of V_{P_0} and V_h are of the same order of magnitude, as the contours of the misfit function are almost circles in the (V_{P_0}, V_h) plane. This is also consistent with the radiation patterns of V_{P_0} and V_h shown in Figure 3.3e, and confirms that the two parameter classes can be jointly involved in FWI, as their radiation pattern do not overlap significantly. The misfit function in the $(V_h, \delta)_{V_{p0}=3300m/s}$ plane shows that the influence of δ is negligible, as the contours of the misfit function are parallel to the δ-axis. The contours of the misfit function for the parameterization (c_{11}, c_{13}, c_{33}) are shown in Figure 3.6e,f with their one-dimensional profile shown in Figure 3.7e,f. The cross-section of the misfit function in the $(c_{11}, c_{33})_{c_{13}=23.958e+9pa}$ plane defined by the true value of c_{13} (Figure 3.6e) confirms that c_{11} and c_{33} have equal influence on the data, as the contours are almost circular. The profiles of the misfit function with respect to c_{11} and c_{33} (Figure 3.7e) confirm the equal sensitivity of the misfit function with respect to these two parameters. The elastic coefficient c_{13} has a similar influence on the data as c_{33} in the vicinity of the minimum of the misfit function (Figure 3.6f). However, a lack of sensitivity of the misfit function can be noted when both c_{13} and c_{33} increase (Figures 3.6f and 3.7f, top panel). The same trend is shown for the (V_{P_0}, δ, ϵ) and (V_{P_0}, V_h, δ) parameterizations, where the sensitivity of the misfit function decreases as δ increases (Figure 3.7d,f).

3.2.4.3 Synthetic examples of full waveform inversion: inclusion model

In this section, we validate the conclusions of the radiation pattern and the grid analysis against synthetic examples of FWI, where we seek to image an inclusion in a homogeneous background model (Figure 3.5b) from a perfect acquisition geometry surrounding the target, following a similar configuration as for the grid analysis. Although a perfect acquisition geometry may not be representative of real acquisitions, it allows us to study the sensitivity of the data to the model parameters over the full range of scattering angles, and validate more easily the conclusions of the radiation-pattern analysis. To perform FWI, we consider the same parameter scaling than for the radiation pattern analysis. Therefore, the misfit function is minimized with respect to parameters that are scaled by their mean value in the background model (see section *Analysis of radiation patterns of virtual sources*). For regularization, we use $\mathbf{m}_{prior} = \mathbf{m}_k$ in Equation 3.4. In this case, the regularization term in the gradient of the misfit function vanishes, and the regularization reduces to the smoothing of the gradient and the Hessian with the operator $\hat{\mathbf{W}}_m^{-1}$. The same value of the damping coefficient λ_i is used for each normalized parameter class and is set to 0.1% of the highest diagonal coefficient of the Hessian. A factor of 0.1% is of the order of magnitude of the damping term used in previous mono-parameter FWI application (Ravaut et al., 2004). Of note, if the diagonal coefficients of the Hessian

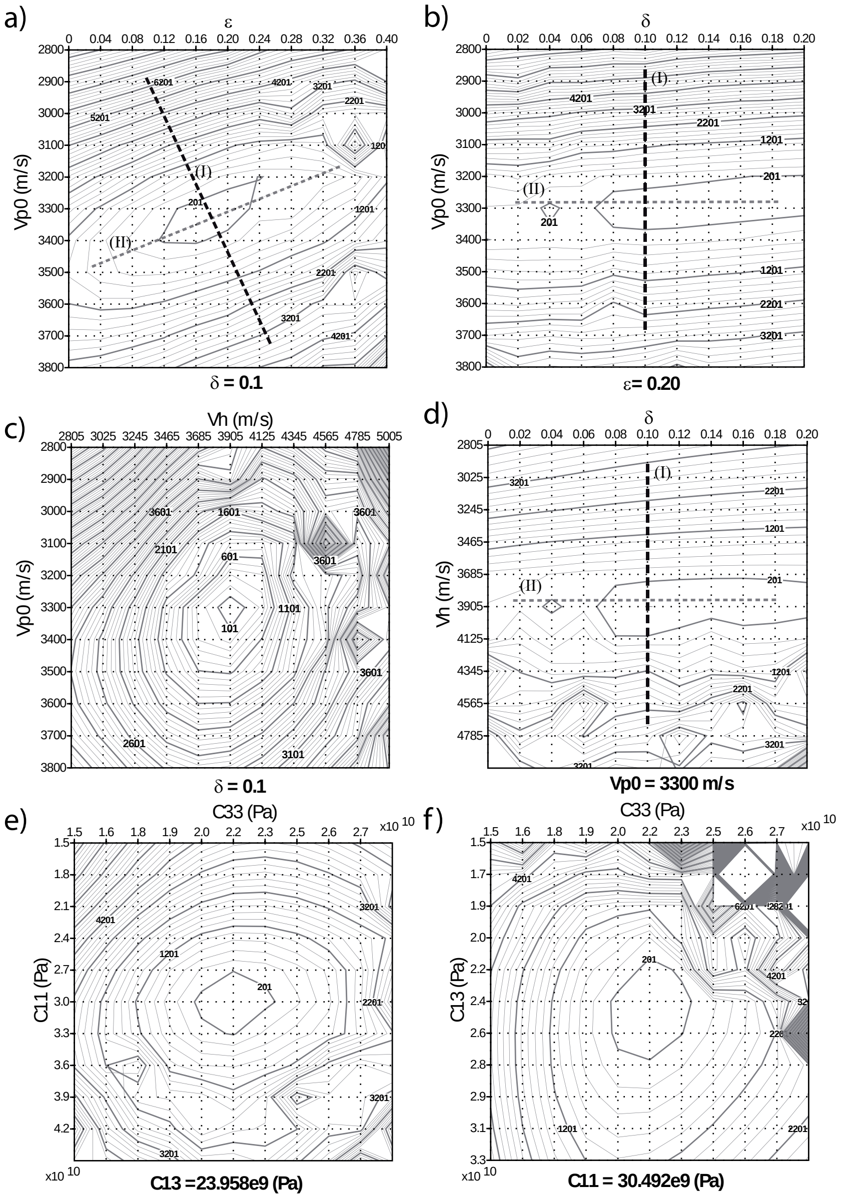


Figure 3.6: Grid analysis of the misfit function. (a-b) parameterization (V_{P0},δ,ϵ). In the cross-section of the misfit function in (a), the (V_{P0}-ϵ) plane is defined by the true value of δ ($\delta = 0.1$), and the (V_{P0}-δ) plane is defined by the true value of ϵ ($\epsilon = 0.2$). The dashed lines show the directions of the maximum (black) and minimum (gray) sensitivities, respectively. (c-d) As for (a, b) for the (V_{P0},V_h) (c) and (δ,V_h) (d) planes for the (V_{P0},δ,V_h) parameterization. (e, f) As for (a, b) for the (c_{11},c_{33}) (e) and (c_{13},c_{33}) (f) planes for the (c_{11},c_{33},c_{13}) parameterization.

associated with one parameter class are much smaller than the selected damping (here, 0.1% of the maximum coefficient), it is unlikely that the inversion will succeed in correctly scaling the model perturbation associated with this parameter class, because the high damping term will cancel out the scaling action of the Hessian.

The initial FWI model is the homogeneous background model. Nine frequencies between

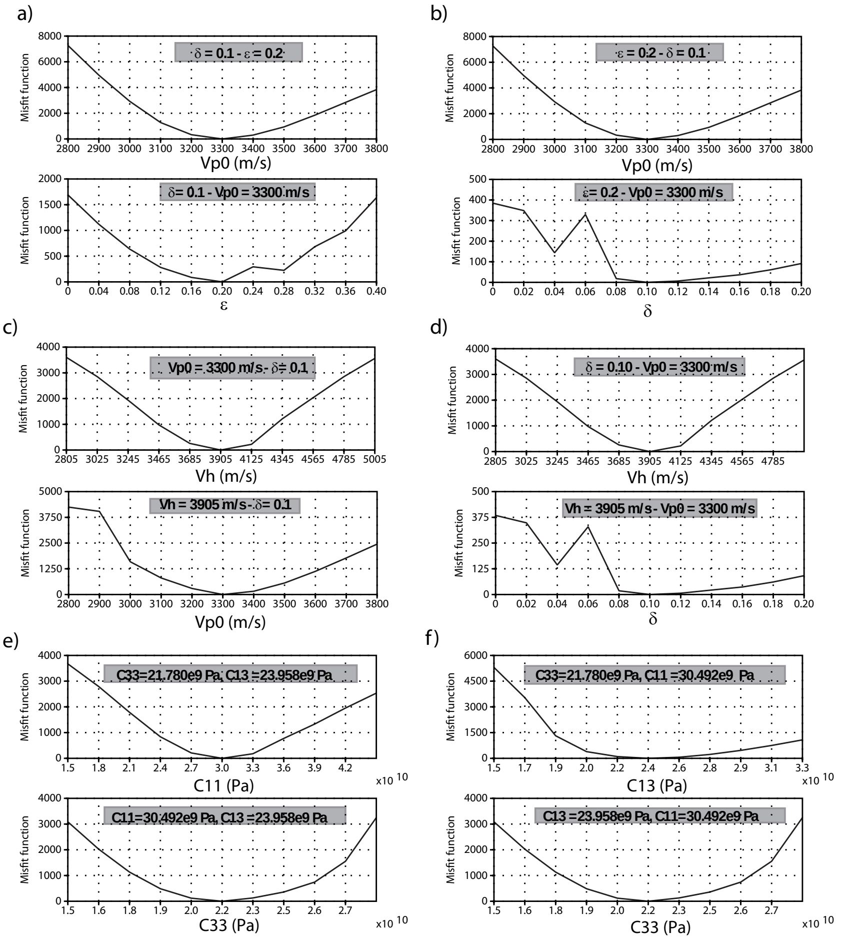


Figure 3.7: Profiles across the planes of the misfit function shown in Figure 3.6. The profiles are shown for the true values of the two parameters kept fixed. (a) V_{P0} and ϵ profiles across the planes of Figure 3.6a. (b) V_{P0} and δ profiles across the planes of Figure 3.6b. (c) V_{P0} and V_h profiles across the planes of Figure 3.6c. (d) δ and V_h profiles across the planes of Figure 3.6d. (e) c_{11} and c_{33} profiles across the planes of Figure 3.6e. (f) c_{13} and c_{33} profiles across the planes of Figure 3.6f.

4.8 Hz and 19.5 Hz are inverted sequentially. The maximum value of the diagonal blocks of the Hessian associated with each normalized parameter class are provided for three parameterizations ((V_{P0}, δ, ϵ), (V_{P0}, δ, V_h) and (c_{11}, c_{13}, c_{33}) in Table 3.1). To highlight the role of the Hessian in FWI and the scaling of the parameter classes, we show the Hessian for the

(V_{P0},δ,ϵ)	5e+25 (V_{P0})	1e+22 (δ)	2.5e+22 (ϵ)
(V_{P0},δ,V_h)	8e+24 (V_{P0})	2.5e+21 (δ)	3e+24 (V_h)
(c_{33},c_{13},c_{11})	1.8e+24 (c_{33})	4.5e+24 (c_{13})	3e+24 (c_{11})

Table 3.1: Maximum value of the diagonal coefficients of the Hessian for the three parametrizations (V_{P0},δ,ϵ), (V_{P0},δ,V_h) and (c_{33},c_{13},c_{11}).

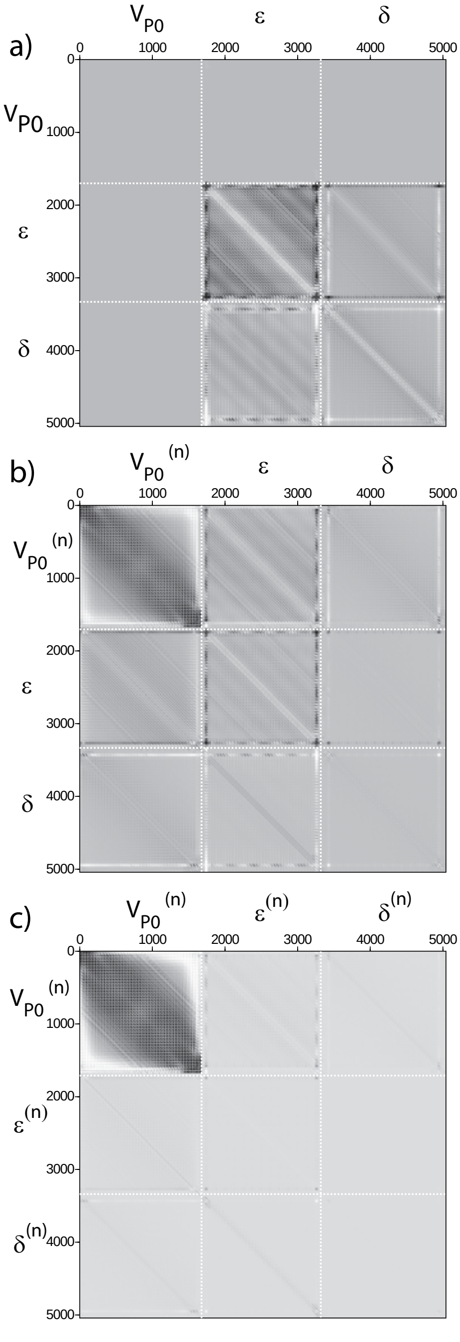


Figure 3.8: Hessian for the (V_{P0},δ,ϵ) parameterization. The upper-left diagonal, middle, and lower-right diagonal blocks are associated with V_{P0}, ϵ and δ, respectively. (a) No normalization is applied to the three parameter classes during minimization of the misfit function. (b) Only V_{P0} is normalized during minimization of the misfit function. (c) The three parameter classes are normalized during the minimization of the misfit function. In the present study, normalization of one parameter class amounts to multiply the Hessian coefficients by the square of the value of the parameter in the background model.

parameterization (V_{P0}, δ, ϵ) and for different normalization of the parameter classes (Figure 3.8). As three parameter classes are considered, the Hessian is a 3 x 3 block matrix, where the off-diagonal blocks represent the trade-off between the different parameter classes. When no normalization is applied, the partial derivative wavefields with respect to vertical velocity have very small amplitudes (it is reminded that the coefficients of the Hessian mostly represents the zero-lag correlation of the partial derivative wavefields, equation 3.4), although this parameter has the strongest influence in the data (Figure 3.8a). In this case, it is unlikely that the deconvolution performed by the Hessian will provide stable results. When only the vertical velocity is scaled (a possible choice as the Thomsen parameters are dimensionless and are proportional to wavespeed ratio), the Hessian shows more uniform coefficients from one block to the next (Figure 3.8b). In this case, the Hessian might reconstitute the true amplitude of the model parameters, if the influence of the secondary parameters on the data is sufficiently high. When the three parameter classes are normalized (Figure 3.8c), the highest coefficients are associated with the parameter that has the strongest influence in the data (the vertical velocity), while the coefficients associated with Thomsen parameters are small. In this case, the risk is to reconstruct overestimated perturbations of the dominant parameter and underestimated perturbations of the secondary parameters, because trade-off effects are not properly corrected for. In the following, the FWI results are obtained when all of the three parameter classes are normalized, unless otherwise mentioned (Figure 3.8c). Although it might not give the most reliable FWI results for some of the parameterizations, this setting allows us to reflect the real sensitivity of the data to the parameters, and cross-validate the conclusions of the radiation pattern analysis with synthetic examples. This approach is also quite general in the sense that it forces of all the parameter classes to have the same range of values, whatever the parameters are dimensionless or not.

Type 1 parameterizations: one wavespeed + two Thomsen parameters

We first perform the joint reconstruction of V_{P0}, δ, and ϵ, when the true medium contains an inclusion perturbation for each model parameter class (Figure 3.9). The extraction of vertical and horizontal profiles across the center of the inclusion allows us to determine how the resolution of the reconstruction is impacted upon by the radiation pattern of the model parameters as a function of the incidence angle. Of note, given the symmetry of the target and the complete coverage provided the acquisition geometry, the horizontal and vertical profiles of a velocity inclusion reconstructed by FWI should be strictly identical in isotropic media, unlike in VTI media. We first show that the amplitudes of the wavespeed perturbations are overestimated (Figure 3.9a-c), while the perturbations of the Thomsen's parameter perturbations are strongly underestimated (Figure 3.9d-i). This results from the combination of several factors: the dominant influence of V_{P0} on the data, the chosen scaling of the parameter classes (Figure 3.8c), and the chosen regularization damping in the Hessian. From Table 3.1, we have: $H^{(\delta)}_{max} = 0.02\% H^{(V_{P0})}_{max}$ and $H^{(\epsilon)}_{max} = 0.04\% H^{(V_{P0})}_{max}$, where $H^{(V_{P0})}_{max}$, $H^{(\delta)}_{max}$ and $H^{(V_{P0})}_{max}$ denote the maximum diagonal coefficients of the Hessian for the V_{P0}, δ and ϵ parameters, respectively. The chosen damping term is higher by one order of magnitude than the maximum coefficient of the Hessian diagonal blocks associated with the Thomsen parameters (Figure 3.8). In this case, the action of the Hessian on the gradient, which consists of correctly scaling the units of the model perturbations, is strongly reduced leading to underestimated perturbations of δ and ϵ. The reconstruction of the wavespeed differs significantly in the vertical and horizontal profiles (Figure 3.9b-c), while they should be nearly identical according to the almost isotropic

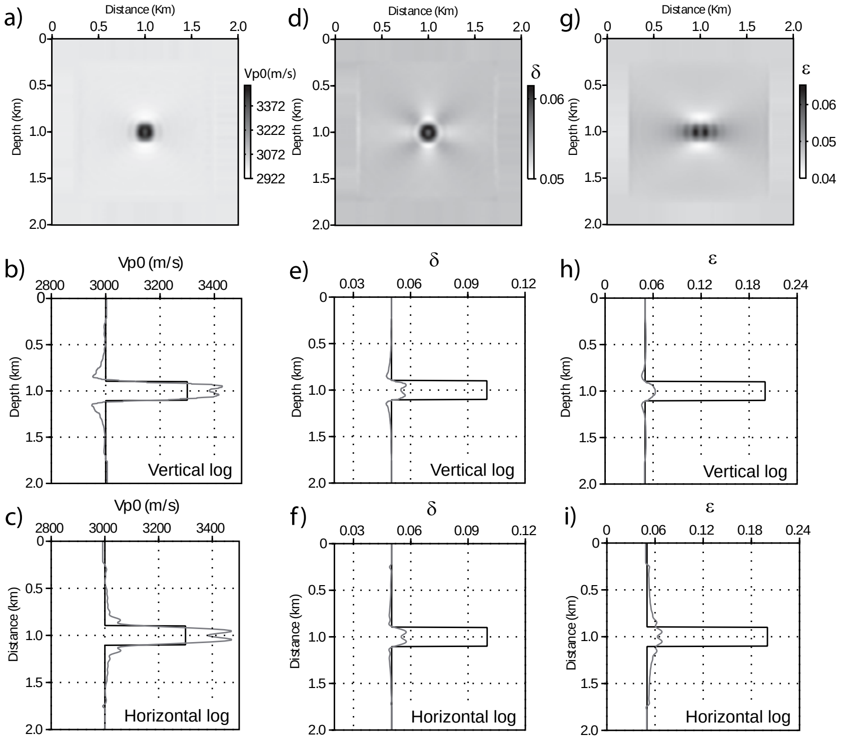


Figure 3.9: Inclusion-model FWI test. Joint update of V_{P0} (a-c), δ (d-f) and ϵ (g-i), when (V_{P0},δ,ϵ) parameterization is used. (a,d,g) Final FWI models. (b,e,h) Vertical profiles across the true inclusion (black) and the reconstructed inclusion (gray). (c,f,i) As for (b,e,h) for the horizontal profiles.

radiation pattern of V_{P0} in the (V_{P0},δ,ϵ) parameterization (Figures 3.3b and 3.4b). The vertical profile shows a deficit of small wavenumbers in the reconstructed V_{P0} model (a high-pass filtering of the inclusion profile would give a similar shape to the reconstructed model), while the horizontal profile shows a more complex shape, which suggests a deficit of intermediate wavenumbers in the spectrum of the reconstructed model. The incorrect shape and amplitudes of the velocity perturbations are interpreted as the footprint of the trade-off between the wavespeed and the Thomsen parameter ϵ. To verify this interpretation, we first perform three mono-parameter FWIs for V_{P0}, δ and ϵ, where the true models contain an inclusion associated with the parameter to be reconstructed (Figure 3.10). The background models of the parameters kept fixed during the inversion match the true models and are homogeneous. The damping term of the Hessian is scaled to the maximum coefficient of the mono-parameter Hessian. The reconstruction of the V_{P0} parameter is now the same along the vertical and horizontal directions (Figure 3.10a-c), which is consistent with the isotropic radiation pattern of the vertical wavespeed in the (V_{P0}, δ, ϵ) parameterization (Figures 3.3b and 3.4b). Moreover, the amplitude of the model perturbations match those of the true inclusion. The mono-parameter inversion for δ succeeds in reconstructing the true peak-to-peak amplitude of the model perturbations, because the damping in the Hessian is scaled to the maximum value of the mono-parameter Hessian associated with δ (Figure 3.10(d-f). The horizontal and vertical profiles show a deficit of small and high wavenumbers, which is consistent with the influence of δ at intermediate angles (Figures 3.3b and 3.4b). Of note, the footprint of the radiation pattern of δ is clearly

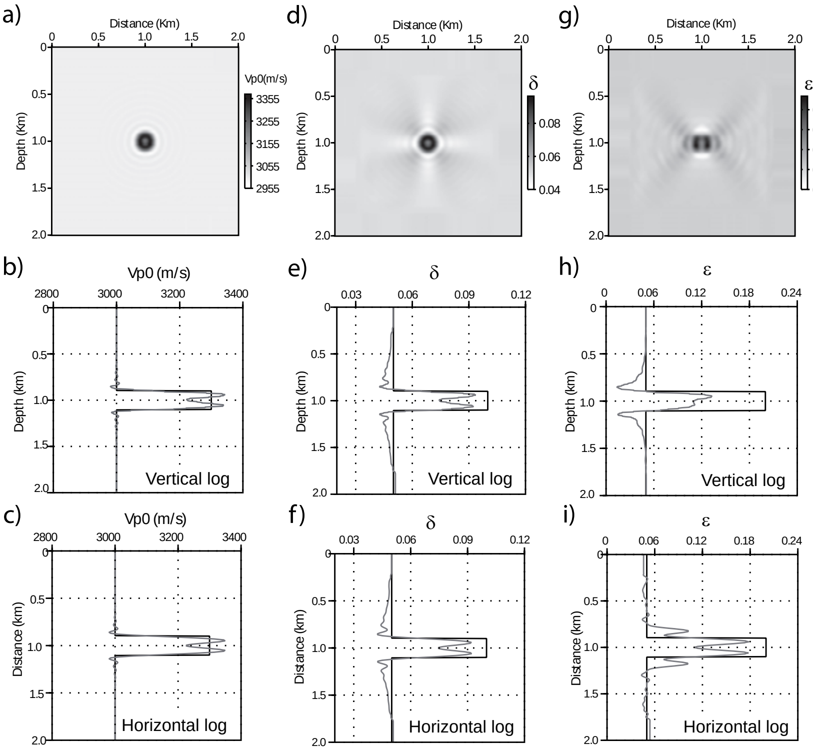


Figure 3.10: Inclusion-model FWI test. Results of the mono-parameter FWI for V_{P0} (a-c), δ (d-f) and ϵ (g-i) when the (V_{P0},δ,ϵ) parameterization is used. (a,d,g) Final FWI models. (b,e,h) Vertical profiles across the true inclusion (black) and the reconstructed inclusion (gray). (c,f,i) As for (b,e,h) for the horizontal profiles. Comparison between these results and those shown in Figure 3.9 allows an assessment of the effects of the trade-off between the parameters.

visible in the FWI model of δ. The horizontal and vertical profiles of δ are strictly identical, which is consistent with the symmetry of the radiation patterns of δ, with respect to the axis $\theta = 90\,;\,270°$ (Figures 3.3 and 3.4). The mono-parameter reconstruction of ϵ shows a more complex anisotropic reconstruction, with however the correct peak-to-peak amplitudes of the model perturbations (Figure 3.10g-i). The vertical profile shows a deficit of small and high wavenumbers, which is consistent with the lack of influence of ϵ on vertical normal-incidence reflection and vertical transmitted wavepaths, respectively. The complex shape of the horizontal profile might result from a deficit of intermediate wavenumbers. Indeed, the large wavelengths of the horizontal profile of the ϵ perturbation are reconstructed from transmitted horizontal wavepaths, while the short wavelengths are reconstructed from short-spread reflections that propagate sub-horizontally. A lack of sensitivity of the wide-spread reflections to ϵ might explain the lack of intermediate wavenumbers in the reconstruction of the horizontal profile of ϵ (Figure 3.1f-i). To test some possible trade-off artifacts, we perform a joint update of the three parameters, when the true medium is a homogeneous medium with an ϵ inclusion (Figure 3.11). Most of the model perturbations have been taken by the vertical velocity model, hence highlighting the dominant influence of this parameter for the chosen parameterization and scaling, and the trade-off between the vertical velocity and ϵ. Moreover, the wavenumber content of the V_{P0} perturbations (Figure 3.11b,c) looks like that of the artificial features, which are superimposed on the horizontal and vertical profiles of V_{P0} reconstructed during the joint

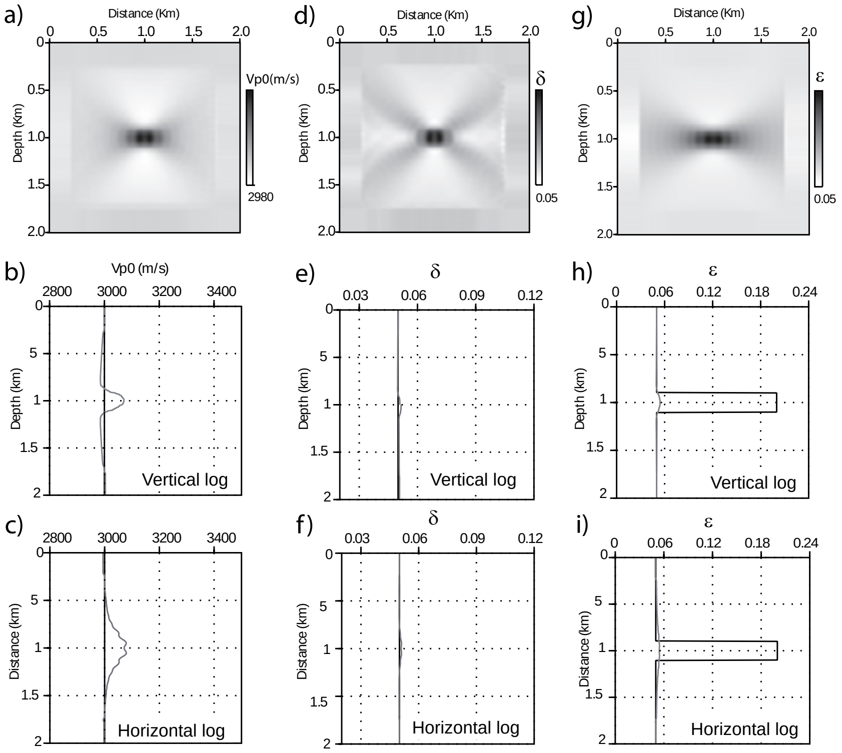


Figure 3.11: Inclusion-model FWI test. Joint up-date of V_{P0} (a-c), δ (d-f) and ϵ (g-i) when the (V_{P0},δ,ϵ) parameterization is used. The true model is homogeneous in V_{P0} and δ and contains an inclusion in ϵ. (a,d,g) Final FWI models. (b,e,h) Vertical profiles across the true inclusion (black) and the reconstructed inclusion (gray). (c,f,i) As for (b,e,h) for the horizontal profiles.

update of the three parameters in the case of three inclusions (compare Figures 3.9b,c and 3.10b,c to assess the artificial features superimposed on the velocity reconstruction in Figure 3.9b,c). This is an additional evidence of the trade-off between the vertical velocity and the parameter ϵ. To highlight the importance of the parameter scaling, we show in Figure 3.12 the FWI results for the joint reconstruction of V_{P0}, δ, and ϵ, when the true medium contains an inclusion perturbation for each model parameter class, and when only the vertical wavespeed is normalized during FWI (Figure 3.8b). The same regularization strategy is used that consists in using the damping terms λ_i of the Hessian of 10% of the maximum Hessian coefficient, whatever i. Compared to the results shown in Figure 3.9 where all of the parameter classes were normalized, the δ model shows overestimated perturbations. In case of noisy data, this will lead to quite noisy reconstruction, because the influence of δ on the data is small. The V_{P0} perturbations are now underestimated, while the ϵ perturbations show reliable peak-to-peak perturbations. For this scaling of the model parameters, the most dominant trade-off effects seem to occur between V_{P0} and δ. Comparison between the results shown in Figures 3.9 and 3.12 shows how the FWI results are sensitive to the way the different parameter classes are scaled during optimization.

Type 2 parameterizations: two wavespeeds + one Thomsen parameter

The joint reconstruction of V_{P0}, δ and V_h performed with the (V_{P0},δ,V_h) parameterization

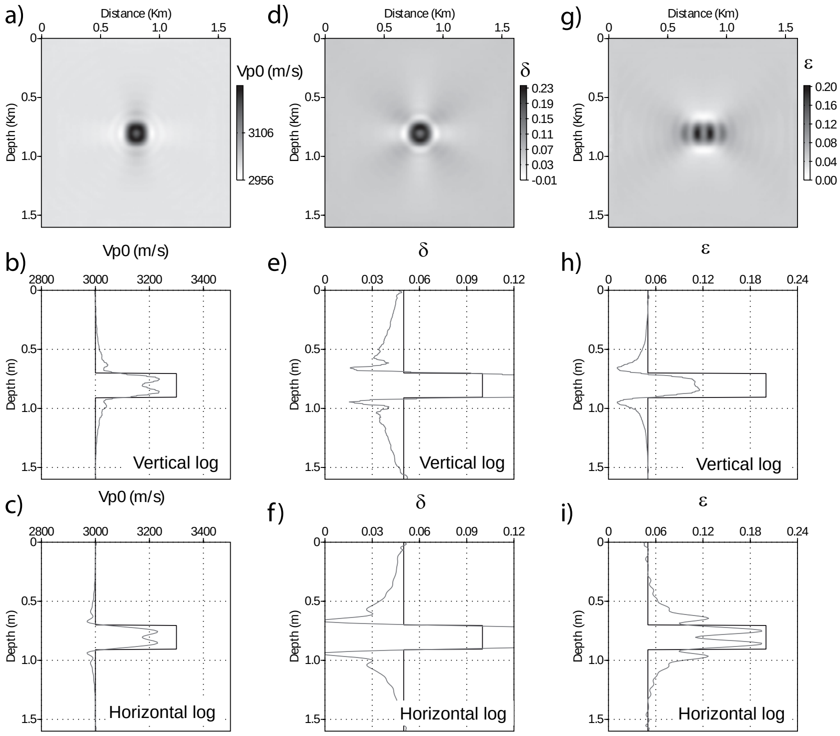


Figure 3.12: Inclusion-model FWI test. Joint up-date of V_{P0} (a-c), δ (d-f) and ϵ (g-i) when the (V_{P0},δ,ϵ) parameterization is used. Same as Figure 3.9 except that only V_{P0} is normalized during the minimization of the misfit function.

is shown in Figure 3.13. We first show that the horizontal and vertical profiles of V_h (Figure 3.13h,i) are close to those of ϵ, when ϵ is built by mono-parameter FWI (Figure 3.10h,i). This is consistent because V_h and ϵ have similar radiation patterns in terms of scattering-angle coverage in the (V_{P0},δ,V_h) and (V_{P0},δ,ϵ) parameterizations (Figure 3.3b,e). However, the radiation pattern of V_h has much higher amplitudes than that of ϵ. This explains why reliable peak-to-peak amplitudes of the V_h perturbations are reconstructed during the simultaneous up-date of the three parameter classes (Figure 3.13h,i), unlike for ϵ (Figure 3.9h,i). The amplitudes of the δ perturbations are significantly underestimated, which is consistent with the weak influence of this parameter in the (V_{P0},δ,V_h) parameterization. The vertical profile of the vertical wavespeed is well reconstructed, although there are some small artifacts near the ends of the inclusion. A broad spectrum of wavenumbers is reconstructed in the vertical profile from small to intermediate scattering angles spanned by the surface acquisition (Figure 3.3e, gray line) and from the transmission regime ($\theta = 180°$) associated with the vertical acquisition (Figure 3.4e, gray line). A possible notch in the wavenumber spectrum that results from the lack of sensitivity of the surface-acquisition data at intermediate scattering angles possibly explains the small-amplitude artifacts in the vertical profile. The horizontal profile of the vertical wavespeed shows more limited bandwidth effects and overestimated peak-to-peak amplitudes, which suggests some possible trade-off between vertical and horizontal wavespeeds. The horizontal profile of the vertical wavespeed should mainly show a deficit of high wavenumbers, because this profile is reconstructed from the large scattering angles of the vertical acquisition (Figure 3.4e, gray line), while the negative velocity perturbations shown in Figure 3.13c

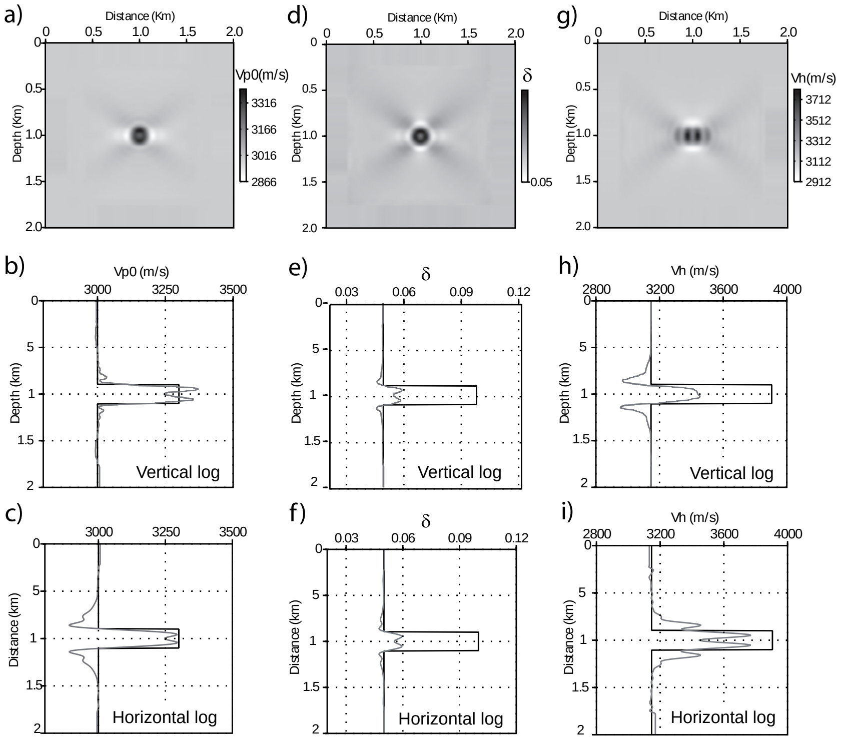


Figure 3.13: Inclusion-model FWI test. Joint update of V_{P0} (a-c), δ (d-f) and V_h (g-i) when the (V_{P0},δ,V_h) parameterization is used. (a,d,g) Final FWI models. (b,e,h) Vertical profiles across the true inclusion (black) and the reconstructed inclusion (gray). (c,f,i) As for (b,e,h) for the horizontal profiles.

suggest that extra high wavenumbers were added to the V_{P0} perturbations. To validate the hypothesis of trade-off effects, we again perform a series of three mono-parameter inversions for V_{P0},δ and V_h (Figure 3.14). Then, we consider a perturbation for the vertical wavespeed only, and jointly update the three parameters (Figure 3.15). The mono-parameter reconstruction of the three parameters shows reliable peak-to-peak amplitude reconstruction. The vertical wavespeed is well reconstructed by mono-parameter FWI along both the horizontal and vertical directions (Figure 3.14a-c). The difference between the horizontal profile of V_{P0} obtained by multiparameter and mono-parameter FWI (Figures 3.13c and 3.14c) is the first evidence that the reconstruction of horizontal wavenumbers of V_{P0} was impacted upon by trade-off effects during the joint up-date of the three parameters. The results of multiparameter FWI for the V_{P0}-inclusion model confirm the trade-off effects between all of the three parameters (Figure 3.15). These trade-off effects result from the overlap of the radiation pattern of V_{P0}, δ and V_h at intermediate aperture angles (typically, between around 80 and 90°) (Figures 3.3e and 3.4e). For completeness, the results of the joint reconstruction of V_{NMO}, δ, and V_h are shown in Figure 3.16. The reconstruction of V_{NMO} and V_h are close to those of V_{P0} and V_h that were obtained with the (V_{P0},δ,V_h) parameterization. However, we note underestimated amplitudes in the vertical profile of V_{NMO} (Figure 3.16b) unlike in the reconstruction of V_{P0} (Figure 3.13b). Moreover, the perturbations of δ have the wrong polarity. These artifacts probably result from the significant trade-off between V_{NMO} and δ (Figures 3.3f and 3.4f, gray and black lines).

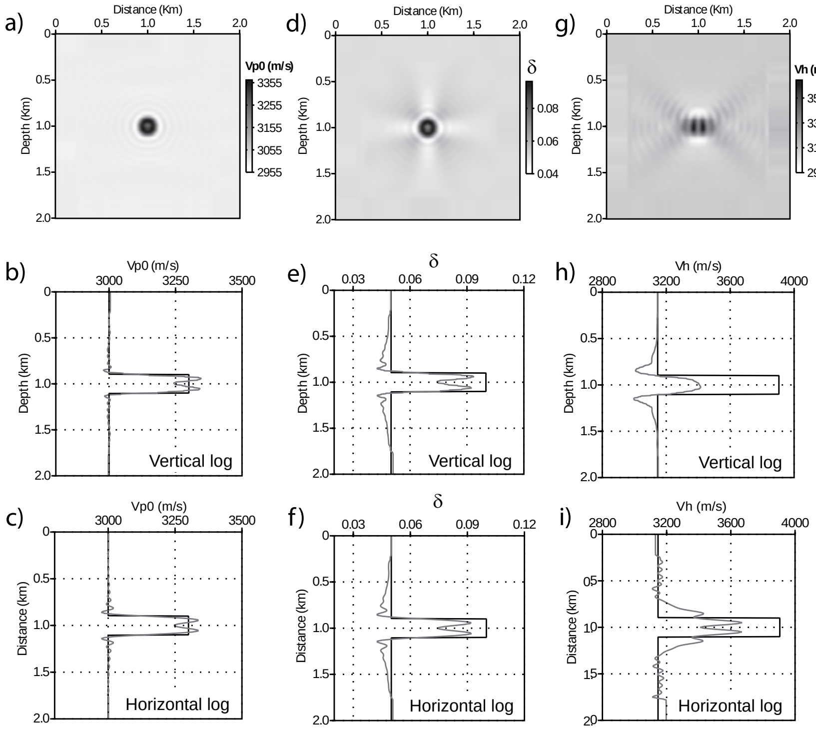


Figure 3.14: Inclusion-model FWI test. Results of mono-parameter FWI for V_{P0} (a-c), δ (d-f) and V_h (g-i) when the (V_{P0},δ,V_h) parameterization is used. (a,d,g) Final FWI models. (b,e,h) Vertical profiles across the true inclusion (black) and the reconstructed inclusion (gray). (c,f,i) As for (b,e,h) for the horizontal profiles. Comparison between these results and those shown in Figure 3.13allows an assessment of the effects of the trade-off between the parameters.

Type 3 parameterization: elastic moduli

The results of the joint up-date of the elastic coefficients are shown in Figure 3.17. We first note that the vertical and horizontal profiles of the c_{13} model are not identical, as they should be according to the symmetric radiation pattern of c_{13} with respect to the axis $\theta = 90\,;\,270°$ (Figure 3.3g). On the other hand, the amplitudes of the c_{13} perturbations are better reconstructed than those of δ for the first two parameterizations. The wavenumber contents of both c_{11} and c_{33} are similar to that of V_h from the (V_{P0},δ,V_h) parameterization, which is consistent because their radiation patterns have identical elliptical shapes (Figures 3.3e,g and 3.4e,g). Of note, the radiation pattern of V_{P0} in the (V_{P0},δ,V_h) parameterization spans over a broader scattering-angle range than c_{33} in the (c_{11},c_{33},c_{13}) parameterization. This defines why the vertical profile of c_{33} is not reconstructed as well as that of V_{P0} in the (V_{P0},δ,V_h) parameterization. The multiparameter FWI reconstructions based upon the (c_{11},c_{33},c_{13}) parameterization are clearly hampered by significant trade-off artifacts that result from the significant influence of c_{13} on the horizontal and vertical wavepaths. This is shown, for example, by the differences between the vertical and horizontal profiles of c_{13}, which should be identical according the symmetry of its radiation pattern with respect to the the axis $\theta = 90\,;\,270°$ (Figures 3.3g and 3.4g).

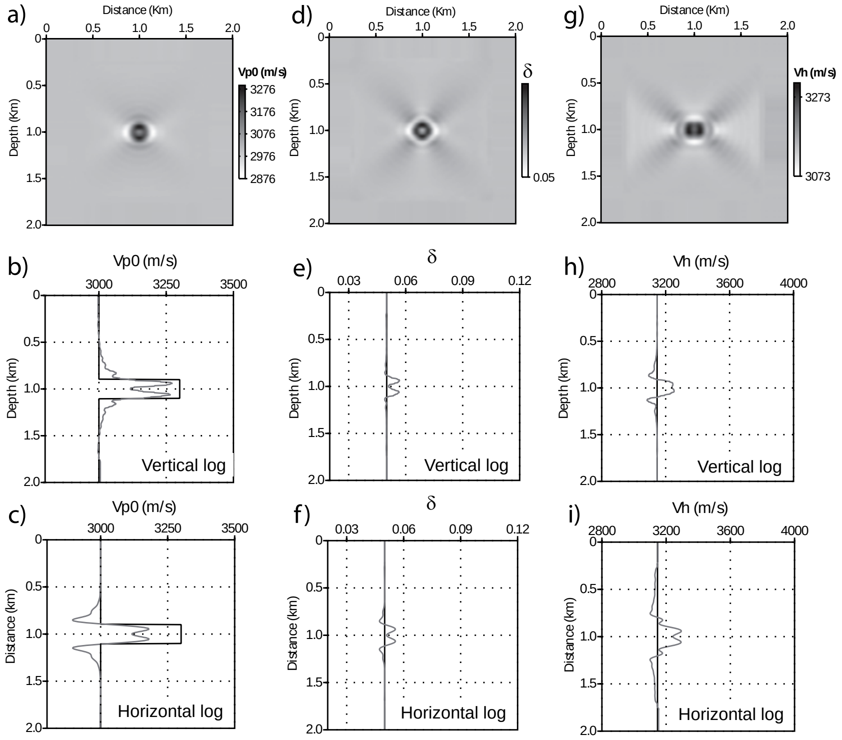


Figure 3.15: Inclusion-model FWI test. Joint update of V_{P0} (a-c), δ (d-f) and V_h (g-i) when the (V_{P0},δ,V_h) parameterization is used. The true model is homogeneous in V_h and δ, and contains an inclusion in V_{P0}. (a,d,g) Final FWI models. (b,e,h) Vertical profiles across the true inclusion (black) and the reconstructed one (gray). (c,f,i) As for (b,e,h) for the horizontal profiles.

3.2.4.4 On the importance of adaptive regularization in multiparameter FWI

We conclude this numerical analysis with some results on the sensitivity of the multiparameter FWI to the damping regularization. We perform a new series of tests with the (V_{P0}, δ, ϵ) parameterization, where we adapt the value of the damping coefficient λ_i to the parameter class, Equation 3.4. Moreover, the regularization is modified such that the prior model is the initial model of each mono-frequency inversion, rather than the initial model of the current inversion iteration, Equation 3.1. In this case, the regularization term in the gradient, Equation 3.4, is zero only at the first iteration of each frequency inversion. For these tests, each parameter classes are normalized (Figure 3.8c). The FWI tests obtained for λ_i = 1e-3, for i=1,3 are shown in Figure 3.18. Compared to the results of the previous tests, the new regularization increases the influence of ϵ in the FWI, as the amplitudes of the ϵ perturbations are higher in Figure 3.18 than in Figure 3.9. As the amplitudes of ϵ are increased, the trade-off artifacts decrease significantly in the V_{P0} reconstruction (compare, for example, the vertical profile of V_{P0} in Figures 3.9b and 3.18b). When λ_δ is set one order of magnitude lower than $\lambda_{V_{P0}}$ and λ_ϵ, the amplitudes of the δ perturbations are overestimated (Figure 3.19), while they were underestimated in the previous tests. When λ_ϵ is in-between the values of $\lambda_{V_{P0}}$ and λ_δ, the amplitudes of the V_{P0} perturbations are damped, while the amplitudes of ϵ are increased (Figure 3.20). Note that not only do the amplitudes of the perturbations change with the values of

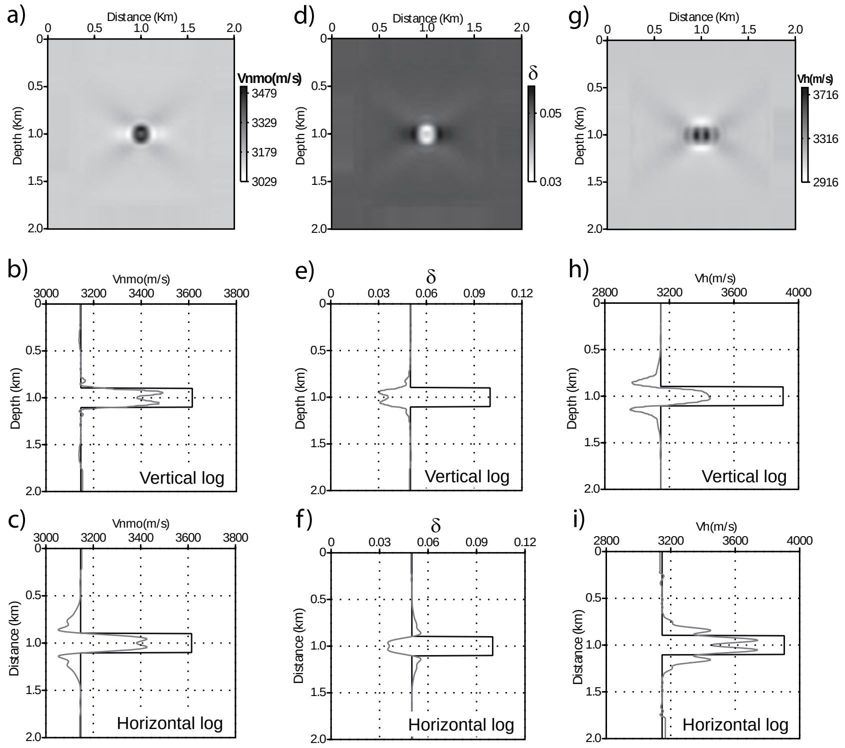


Figure 3.16: Inclusion-model FWI test. Joint update of V_{NMO} (a-c), δ (d-f) and V_h (g-i) when the (V_{NMO},δ,V_h) parameterization is used. (a,d,g) Final FWI models. (b,e,h) Vertical profiles across the true inclusion (black) and the reconstructed inclusion (gray). (c,f,i) As for (b,e,h) for the horizontal profiles.

λ_i but also the strength of the trade-off artifacts between V_{P0} and ϵ. Indeed, the values of λ_i control the relative values of the diagonal blocks of the Hessian, and also the relative values between the diagonal and thee off-diagonal blocks, with these off-diagonal blocks correcting for the trade-off effects in FWI (Figure 3.8). It is worth remembering that these numerical tests are performed without noise. In the presence of noise, the parameters that have a weak influence in the data cannot be reconstructed by FWI whatever the regularization, because the imprint of the parameter in the data will be dominated by the noise. Although the design of a suitable regularization for multi-parameter FWI is beyond the scope of this study, these preliminary results show that careful tuning of the regularization associated with a judicious scaling of each parameter class will be a key for the success of multiparameter FWI.

3.2.5 Discussion

We have presented here a practical approach to choose a suitable parameterization for multi-parameter FWI with applications in the imaging of acoustic VTI media.

The criteria to choose a suitable parameterization for FWI mainly rely on the analysis of the relative influences of the parameter classes on the data, as functions of the scattering angles and propagation directions. This influence controls how the wavenumber spectrum of the subsurface model is filled during FWI, and hence the expected spatial resolution of the

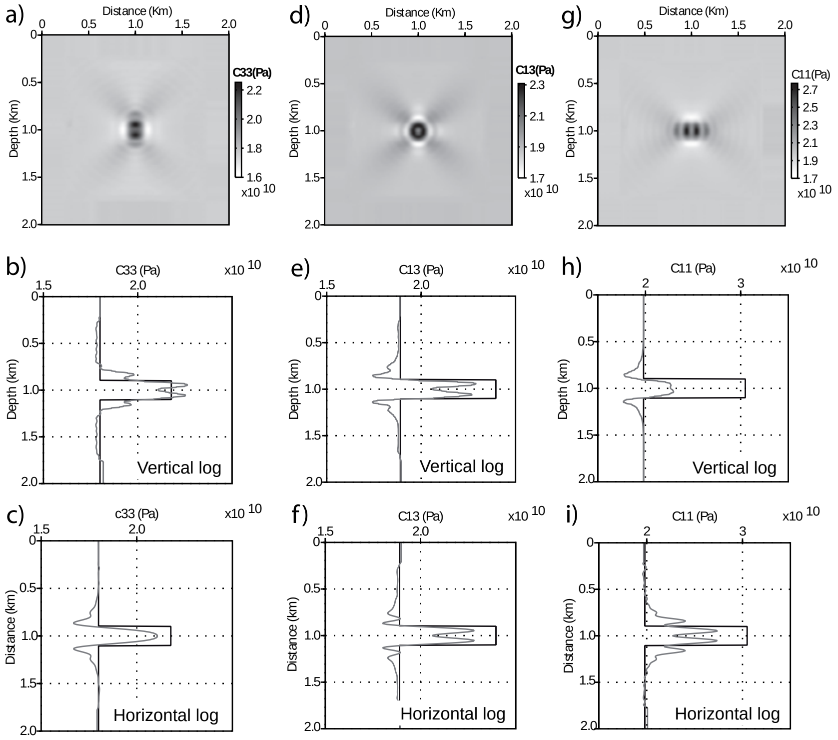


Figure 3.17: Inclusion-model FWI test. Joint update of c_{33} (a-c), c_{13} (d-f) and c_{11} (g-i) when the (c_{33},c_{13},c_{11}) parameterization is used. (a,d,g) Final FWI models. (b,e,h) Vertical profiles across the true inclusion (black) and the reconstructed inclusion (gray). (c,f,i) As for (b,e,h) for the horizontal profiles.

imaging as well as the potential trade-off between the multi-parameter classes. On the one hand, parameterization can be chosen such that as many as possible parameter classes have radiation patterns spanning over a wide range of scattering angles, to guarantee the broad-band-wavenumber reconstruction of these parameters. On the other hand, the radiation patterns of two different parameter classes should not show significant influence of the parameter classes on the data for the same range of scattering angles, to prevent trade-off artifacts during FWI. Indeed, these two criteria cannot be fulfilled simultaneously.

From the optimization viewpoint, the Hessian theoretically corrects for trade-off effects and scale the gradients of the misfit functions associated with the different parameter classes, such that the true values of these model parameters are retrieved, whatever their relative influence on the data and the order of magnitude of the Fréchet derivatives. In practice, the regularization conventionally used in FWI adds a damping term to the Hessian, which might prevent the removal of some trade-off effects and robust reconstruction of the parameters with low-amplitude Fréchet derivatives. These damping terms can however be adapted to each parameter class to account for this variable influence. In addition to regularization tuning, suitable normalization of the parameter classes is another adjustment knob that can be helpful to improve the conditioning of the Hessian. However, this scaling should be chosen carefully to avoid putting to much weight during inversion on parameters that have a small influence on the data. Indeed, it is unlikely that FWI can reconstruct a parameter, which has an influence on the data below the noise level. If too much weight is assigned to these parameters, noisy

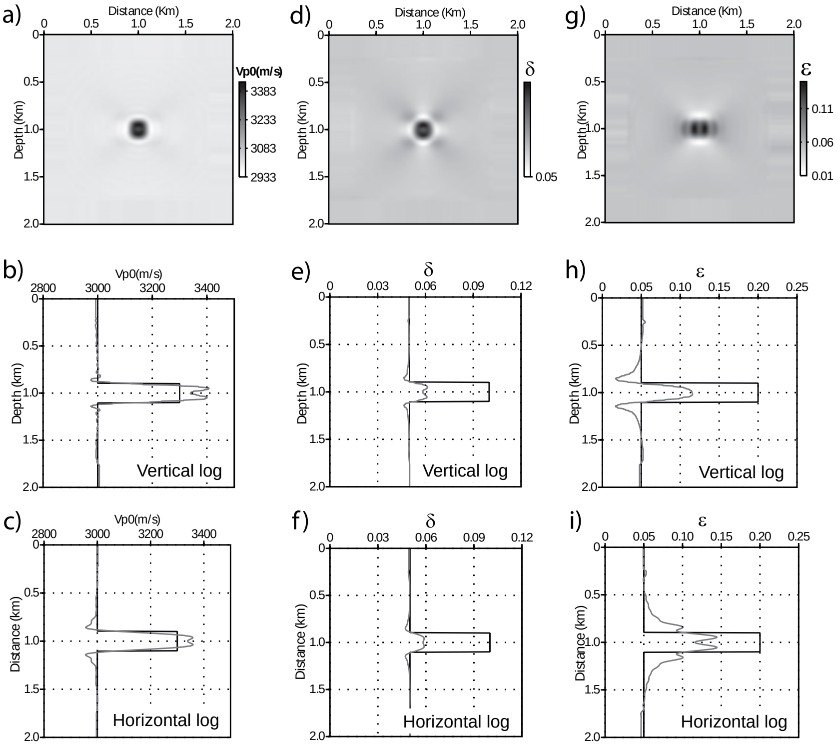


Figure 3.18: Inclusion-model FWI test. Joint update of V_{P0} (a-c), δ (d-f) and ϵ (g-i) when the (V_{P0},δ,ϵ) parameterization is used. (a,d,g) Final FWI models. (b,e,h) Vertical profiles across the true inclusion (black) and the reconstructed inclusion (gray). (c,f,i) As for (b,e,h) for the horizontal profiles. A Tikhonov regularization is used with $\lambda_{V_{P_0}}$=1e-3, λ_δ=1e-3 and λ_ϵ=1e-3 and $\mathbf{m}_{prior} = \mathbf{m}_0$ (see text for details).

reconstruction will be obtained.

Although a detailed analysis of regularization and parameter scaling is beyond the main scope of the present study, we would conclude from the present study that any subsurface parameterization for FWI needs judicious scaling of each parameter class involved in the parameterization as well as a careful parametric analysis of the regularization.

We have identified two main categories of parameterization for acoustic VTI media. For the first, the wavespeed parameter has a dominant influence on the data, and has a radiation pattern that spans over the full range of scattering angles. The two other parameters, the Thomsen parameters, have weaker influence on the data over a narrower range of scattering angles. The Thomsen parameter δ has the weakest influence on the data, with this influence limited to intermediate scattering angles. The parameter ϵ has an influence on the near vertical wavepaths, or on the near horizontal ones, depending whether ϵ is associated with the horizontal or vertical wavespeed in the parameterization. As significant trade-off is expected between the wavespeed and the two Thomsen parameters, one reliable strategy might be to keep the Thomsen parameters fixed during FWI and only up-date the dominant parameter, provided that sufficiently accurate starting models of the Thomsen parameters are available. If the strategy is relevant, we would tend to favor a parameterization that involves the vertical velocity rather than the NMO velocity, because δ has a stronger influence on the data when combined

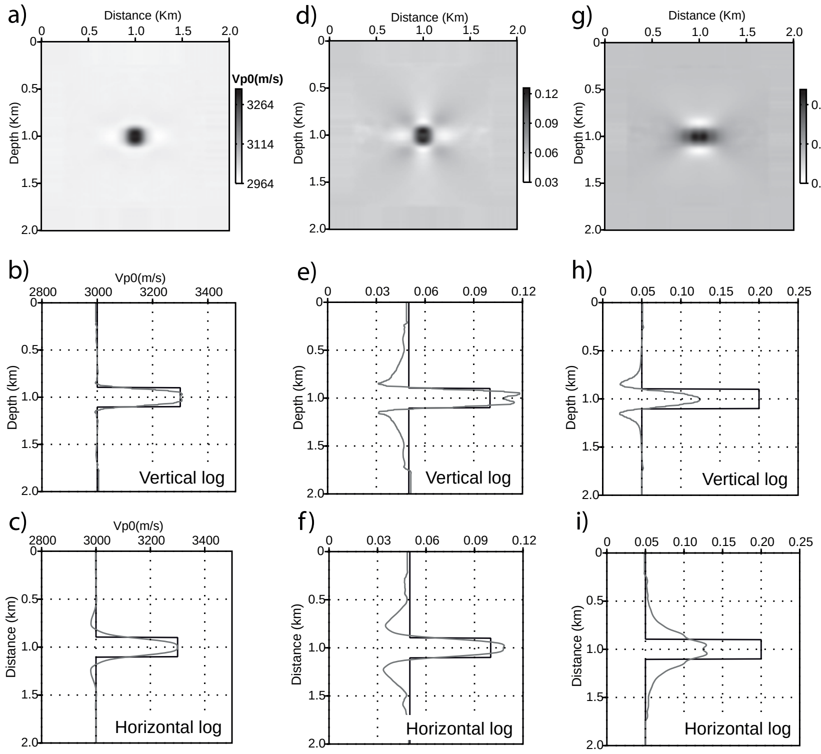


Figure 3.19: As for Figure 3.18 for $\lambda_{V_{P_0}}$=1e-3, λ_δ=1e-4 and λ_ϵ=1e-3. Note how the relative amplitudes of the model perturbations change with the values of the weighting coefficients of the regularization.

with the NMO velocity. In this case, a more accurate initial model of the δ parameter would be required to guarantee a reliable up-date of the velocity model during mono-parameter FWI. We would also recommend parameterization that involves the vertical velocity rather than the horizontal velocity if a smooth background model of ϵ is available. When combined with the vertical velocity in the parameterization, the smooth background model of ϵ should allow sufficiently accurate prediction of the influence of ϵ on the wide-aperture components of the data.

Alternatively, a parameterization that involves two wavespeeds and one Thomsen parameter can be used. This kind of parameterization leads to a limited trade-off between the vertical and horizontal wavespeeds at intermediate scattering angles, while δ has a very minor influence on the data. In this case, joint up-date of the two wavespeeds should be possible with different resolutions. For surface acquisition, the intermediate and short wavelengths of the vertical velocity are reconstructed from the short and intermediate scattering angles, while the intermediate and long wavelengths of the horizontal velocity are reconstructed from the intermediate and wide scattering angles. As the vertical velocity influences the data for the short scattering angles with the (V_{P0},δ,V_h) parameterization, a more accurate background model of the vertical velocity might be necessary, because the large wavelengths of the vertical velocity cannot be up-dated by FWI as they have no influence on the data for this parameterization. If we assume that δ cannot be reliably up-dated by FWI, we would recommend the avoiding of parameterization that involves the NMO velocity combined with the horizontal velocity, because the

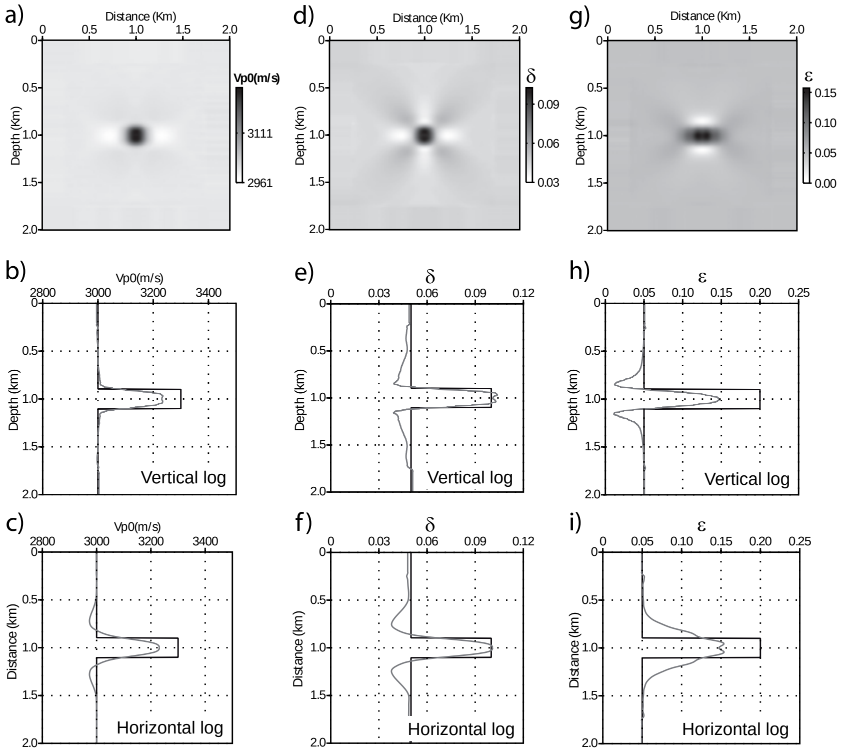


Figure 3.20: As for Figure 3.18 for $\lambda_{V_{P_0}}$=1e-3, λ_δ=1e-4 and λ_ϵ=5e-4.

influence of δ on the data is higher than for parameterization that involves the vertical velocity and shows more significant trade-off. Therefore, a more accurate background model of δ might be required to allow for the reconstruction of the NMO and horizontal velocities. We would also tend to favor the (V_{P0},δ,V_h) parameterization instead of the (c_{11},c_{33},c_{13}) parameterization for similar reasons. The non-overlapping radiation patterns of c_{11} and c_{33} suggest that there is no trade-off between these parameters. However, there is some trade-off between c_{11} and c_{13} on one hand and with c_{33} and c_{13} on the other hand with a significant influence of c_{13} on the data at intermediate scattering angles.

In the following companion report, we validate these statements against the application of anisotropic acoustic FWI to a realistic synthetic example that is representative of the Valhall Field before applying the application to real wide-aperture ocean-bottom-cable data from this field.

3.2.6 Conclusion

Updating multiple parameter classes by FWI is a difficult methodological challenge, because the ill-posedness of the inversion increases when the number of degrees of freedom in the model space increases. In this study, we propose a pragmatical approach to analyze the influence of different subsurface parameterization on wide-aperture data, and some criteria to choose a suitable parameterization for either mono-parameter and multi-parameter FWI. We apply our heuristic approach to the case of VTI acoustic media. Our approach relies on the numerical analysis of the radiation patterns of the different parameter classes, which give some insight into the influence of the parameters on the data as a function of the scattering angle. These radiation patterns represent the directivity of the virtual source that is located at the position of the model parameter, and which generates the partial derivative wavefield. This analysis is validated against canonical synthetic examples that are performed in a simple subsurface model with a perfect illumination. One proposed strategy consists of choosing parameterization that combines one parameter with a dominant influence on the data over a broad range of scattering angles, with two secondary parameters with much smaller influences on the data. In this case, reliable mono-parameter FWI can be performed to up-date the dominant parameter, keeping the secondary parameters fixed during the FWI. Parameterization that combines the vertical wavespeed and the Thomsen parameters δ and ϵ would be the most suitable one for this purpose. Alternatively, we have considered parameterization that combines two wavespeeds and the Thomsen parameter δ. The two wavespeeds have significant influence on the data for different ranges of scattering angles. In this case, the joint up-date of the two wavespeeds should be possible with two different resolutions. We also highlight the role of the Hessian to remove the trade-off between multiple parameter classes. However, the tuning of adaptive dampings in the regularization requires a good trade-off to be found between the needs to improve the conditioning of the Hessian, while preserving the scaling of the gradients of the misfit function with respect to the parameters, the influence of which is weak in the data. The design of suitable regularization for multi-parameter FWI as well as a judicious scaling of the different parameter classes will be the aim of future studies.

3.2.7 Acknowledgments

This study was funded by the SEISCOPE consortium http://seiscope.oca.eu, sponsored by BP, CGG-VERITAS, ENI, EXXON-MOBIL, PETROBRAS, SAUDI ARAMCO, SHELL, STATOIL and TOTAL. The linear systems were solved with the MUMPS package, available on http://graal.ens-lyon.fr/MUMPS/index.html. The mesh generation was performed with the help of TRIANGLE, available on http://www.cs.cmu.edu/∼ quake/triangle.html. This study was granted access to the high-performance computing facilities of the SIGAMM (Observatoire de la Côte d'Azur) and to the HPC resources of [CINES/IDRIS] under the allocation 2010- [project gao2280] made by GENCI (Grand Equipement National de Calcul Intensif). We gratefully acknowledge both of these Facilities and the support of their staff.

3.2.8 Appendix A: Acoustic VTI modeling

In this appendix, we review the VTI acoustic forward modeling operators, that we use, on the one hand to perform seismic wave propagation modeling, and on the other hand to derive the expression of the gradient of the misfit function with the adjoint-state method. These two forward modeling operators are different for reasons explained below and in Appendix B.

To derive the VTI acoustic wave equation, we start from the two-dimensional P-SV velocity-stress equation in VTI media:

$$\begin{aligned}
-\imath\omega v_x &= b\left(\frac{\partial \sigma_{xx}}{\partial x}+\frac{\partial \sigma_{xz}}{\partial z}\right)+bf_x,\\
-\imath\omega v_z &= b\left(\frac{\partial \sigma_{xz}}{\partial x}+\frac{\partial \sigma_{zz}}{\partial z}\right)+bf_z,
\end{aligned}$$

$$\begin{aligned}
-\iota\omega\sigma_{xx} &= c_{11}\frac{\partial v_x}{\partial x} + c_{13}\frac{\partial v_z}{\partial z} - \iota\omega\sigma_{xx_0}, \\
-\iota\omega\sigma_{zz} &= c_{13}\frac{\partial v_x}{\partial x} + c_{33}\frac{\partial v_z}{\partial z} - \iota\omega\sigma_{zz_0}, \\
-\iota\omega\sigma_{xz} &= c_{44}\left(\frac{\partial v_x}{\partial z} + \frac{\partial v_z}{\partial x}\right) - \iota\omega\sigma_{xz_0},
\end{aligned} \tag{3.12}$$

where $(v_x(\omega,\mathbf{x}), v_z(\omega,\mathbf{x}))$ and $(\sigma_{xx}(\omega,\mathbf{x}), \sigma_{zz}(\omega,\mathbf{x}), \sigma_{xz}(\omega,\mathbf{x}))$ denote the particle velocities and stresses, respectively. The VTI medium is parametrized by the buoyancy (i.e., the inverse of density) $b(\mathbf{x})$ and the stiffness coefficients $c_{11}(\mathbf{x})$, $c_{13}(\mathbf{x})$, $c_{33}(\mathbf{x})$ and $c_{44}(\mathbf{x})$, while ω denotes the angular frequency and $\iota = \sqrt{-1}$ is the purely imaginary term. The source is either forces $(f_x(\omega,\mathbf{x}), f_z(\omega,\mathbf{x}))$ or stresses $(\sigma_{xx0}(\omega,\mathbf{x}), \sigma_{zz0(\omega,\mathbf{x})}, \sigma_{zz0}(\omega,\mathbf{x}))$.

A change in the variables can be applied to the stress vector to explicitly introduce the pressure wavefield $\mathbf{p}$ (Brossier et al., 2008; Brossier, 2011a). The new stress components are given by $(p,q,r) = ((\sigma_{xx}+\sigma_{zz})/2, (\sigma_{xx}-\sigma_{zz})/2, \sigma_{xz})$, which can be interpreted as the mean, deviatoric, and shear-stress components.

After the change of the variables, the first-order hyperbolic P-SV system is given by:

$$\begin{aligned}
-\iota\omega v_x &= b\left(\frac{\partial(p+q)}{\partial x} + \frac{\partial r}{\partial z}\right) + b f_x \bar{\delta}(\mathbf{x}-\mathbf{x}_s), \\
-\iota\omega v_z &= b\left(\frac{\partial r}{\partial x} + \frac{\partial(p-q)}{\partial z}\right) + b f_z \bar{\delta}(\mathbf{x}-\mathbf{x}_s), \\
-\iota\omega p &= \frac{c_{11}+c_{13}}{2}\frac{\partial v_x}{\partial x} + \frac{c_{13}+c_{33}}{2}\frac{\partial v_z}{\partial z} - \iota\omega p^0 \bar{\delta}(\mathbf{x}-\mathbf{x}_s), \\
-\iota\omega q &= \frac{c_{11}-c_{13}}{2}\frac{\partial v_x}{\partial x} + \frac{c_{13}-c_{33}}{2}\frac{\partial v_z}{\partial z}, \\
-\iota\omega r &= c_{44}\left(\frac{\partial v_z}{\partial x} + \frac{\partial v_x}{\partial z}\right).
\end{aligned} \tag{3.13}$$

In Equation 3.13, we consider only point sources: vertical and horizontal forces, $f_x\bar{\delta}(\mathbf{x}-\mathbf{x}_s)$ and $f_z\bar{\delta}(\mathbf{x}-\mathbf{x}_s)$, respectively, or explosive source, $p^0\bar{\delta}(\mathbf{x}-\mathbf{x}_s)$, where the Dirac delta function and the source position are denoted by $\bar{\delta}$ and $\mathbf{x}_s$, respectively. We implement the acoustic approximation by setting the shear-wave speed to 0 on the symmetry axis (Alkhalifah, 2000), which cancels out the last equation of the system of linear equations, Equation 3.13.

$$\begin{aligned}
-\iota\omega v_x &= b\frac{\partial(p+q)}{\partial x} + b f_x \bar{\delta}(\mathbf{x}-\mathbf{x}_s) \\
-\iota\omega v_z &= b\frac{\partial(p-q)}{\partial z} + b f_z \bar{\delta}(\mathbf{x}-\mathbf{x}_s) \\
-\iota\omega p &= \frac{c_{11}+c_{13}}{2}\frac{\partial v_x}{\partial x} + \frac{c_{13}+c_{33}}{2}\frac{\partial v_z}{\partial z} - \iota\omega p^0 \bar{\delta}(\mathbf{x}-\mathbf{x}_s) \\
-\iota\omega q &= \frac{c_{11}-c_{13}}{2}\frac{\partial v_x}{\partial x} + \frac{c_{13}-c_{33}}{2}\frac{\partial v_z}{\partial z}
\end{aligned} \tag{3.14}$$

This approach based on the P-SV elastodynamic system is close to that developed by Duveneck et al. (2008). The difference is that we use the first-order system to perform seismic modeling, while Duveneck et al. (2008) eliminate the particle-velocity wavefields from the first-order system to build a second-order wave equation for normal stresses, following a parsimonious approach. Second, we use a change of variables on the normal stresses, which gives a direct access to the pressure.

This system of linear equations, Equation 3.14, can be recast in matrix form as:

$$\mathbf{A}(\mathbf{m}(\mathbf{x}),\omega)\mathbf{w}(\mathbf{x},\omega) = \mathbf{s}_A(\mathbf{x},\omega), \tag{3.15}$$

where $\mathbf{A}$ is referred to as the velocity-stress impedance matrix. The vectors $\mathbf{w} = (\mathbf{v}_x, \mathbf{v}_z, \mathbf{p}, \mathbf{q})$ and $\mathbf{s}_A = (bf_x\bar{\delta}(\mathbf{x}-\mathbf{x}_s), bf_z\bar{\delta}(\mathbf{x}-\mathbf{x}_s), \imath\omega p^0\bar{\delta}(\mathbf{x}-\mathbf{x}_s), \mathbf{0})$ denote the monochromatic velocity-stress wavefields and the source, respectively. We discretize the velocity-stress wave equation, equation 3.14, with a nodal formulation of the discontinuous Galerkin method, based on Lagrange polynomials of order 0, 1, or 2 (referred to as P0, P1, and P2, respectively) (Brossier et al., 2010a; Brossier, 2011a), to perform seismic modeling in VTI acoustic media.

A second-order wave equation for particle velocities that turns out to be useful for FWI implementation can be inferred from the velocity-stress wave equation by eliminating the stress wavefields p and q in the first and second block rows of Equation 3.14, giving the system:

$$\begin{aligned}
\omega^2\rho v_x + \left(\frac{\partial}{\partial x}c_{11}\frac{\partial v_x}{\partial x} + \frac{\partial}{\partial z}c_{13}\frac{\partial v_z}{\partial z}\right) &= \imath\omega p^0\frac{\partial\delta(\mathbf{x}-\mathbf{x}_s)}{\partial x} + \imath\omega f_x\delta(\mathbf{x}-\mathbf{x}_s), \\
\omega^2\rho v_z + \left(\frac{\partial}{\partial x}c_{13}\frac{\partial v_x}{\partial x} + \frac{\partial}{\partial z}c_{33}\frac{\partial v_z}{\partial z}\right) &= \imath\omega p^0\frac{\partial\delta(\mathbf{x}-\mathbf{x}_s)}{\partial z} + \imath\omega f_z\delta(\mathbf{x}-\mathbf{x}_s).
\end{aligned} \tag{3.16}$$

This system of linear equations can be recast in matrix form as:

$$\mathbf{B}(\mathbf{m}(\mathbf{x}),\omega)\mathbf{v}(\mathbf{x},\omega) = \mathbf{s}_B(\mathbf{x},\omega), \tag{3.17}$$

$$\mathbf{B}(\mathbf{m}(\mathbf{x}),\omega)\mathbf{v}(\mathbf{x},\omega) = \mathbf{s}_B(\mathbf{x},\omega), \tag{3.18}$$

where the particle-velocity vector and the source are denoted respectively by $\mathbf{v} = (\mathbf{v}_x, \mathbf{v}_z)$ and $\mathbf{s}_B = (\imath\omega p^0\frac{\partial\delta(\mathbf{x}-\mathbf{x}_s)}{\partial x} + \imath\omega f_x\delta(\mathbf{x}-\mathbf{x}_s), \imath\omega p^0\frac{\partial\delta(\mathbf{x}-\mathbf{x}_s)}{\partial z} + \imath\omega f_z\delta(\mathbf{x}-\mathbf{x}_s))$, respectively. Of note, the second-order system, Equation 3.18, is self-adjoint, as matrix $\mathbf{B}$ is symmetric, an important feature that directs us towards this forward-problem operator for computing the gradient of the misfit function with the adjoint-state method (see next section). Other possible formulations involve either the first-order velocity-stress system, Equation 3.15, or the second-order system for (p,q) where the particle velocities are eliminated from the velocity-stress system, Equation 3.14, for handling pressure directly in relation to the recorded pressure data. However, both of these systems are no more symmetrical, and hence non-self-adjoint. We discretize the forward-problem operator $\mathbf{B}$ with the finite-volume scheme of Brossier et al. (2008), which is equivalent to the P0 discontinuous Galerkin scheme.

The first-order and second-order hyperbolic systems, Equations 3.15 and 3.18, give the same solutions for the particle velocities as long as the correct relationships between the sources $\mathbf{s}_A$ and $\mathbf{s}_B$ are used as follows:

$$\begin{aligned}
&\mathbf{s}_A = (bf_x\delta(\mathbf{x}-\mathbf{x}_s), bf_z\delta(\mathbf{x}-\mathbf{x}_s), \mathbf{0}, \mathbf{0}) &&\rightarrow \quad \mathbf{s}_B = (\imath\omega f_x\delta(\mathbf{x}-\mathbf{x}_s), \imath\omega f_z\delta(\mathbf{x}-\mathbf{x}_s)) && for \quad a \quad force. \\
&\mathbf{s}_A = (\mathbf{0}, \mathbf{0}, \imath\omega p^0\delta(\mathbf{x}-\mathbf{x}_s), \mathbf{0}) &&\rightarrow \quad \mathbf{s}_B = (\imath\omega p^0\tfrac{\partial\delta(\mathbf{x}-\mathbf{x}_s)}{\partial x}, \imath\omega p^0\tfrac{\partial\delta(\mathbf{x}-\mathbf{x}_s)}{\partial z}) && for \quad an \quad explosion.
\end{aligned} \tag{3.19}$$

Indeed, the explosive source in the velocity-stress system is transformed in a vertical and horizontal dipoles in the second-order wave equation.

3.2.9 Appendix B: Computing the gradient of the misfit function with the adjoint-state method

In this Appendix, we review the derivation of the gradient of the misfit function with the adjoint-state method based on the two forward modeling operators **A** and **B**, which are introduced in Appendix A. Our motivation behind the use of these two modeling operators is many-fold. On one hand, first-order differential operators in matrix **A** are easier to discretize with the discontinuous Galerkin method for seismic modeling. On the other hand, the second-order forward problem operator **B** allows us to derive the expression of the gradient of the misfit function from the self-adjoint operator. Moreover, it allows us to save memory during the computation of the gradient, as stress wavefields do not need to be kept in memory after seismic modeling (Brossier, 2011b). When pressure data are inverted, the velocity-stress forward problem operator **A** also allows us to easily introduce pressure residuals into the source term of the adjoint equation (Appendix A). Of note, the self-adjointness of the forward modeling operator (and, hence its symmetry) is not very useful in frequency-domain modeling, because one can indifferently multiply the forward modeling operator or its transpose to a vector to compute the adjoint-state variable (Plessix and Cao, 2011b). However, for time-domain modeling, the self-adjointness of the forward modeling operator allows one to use the same numerical scheme to compute the state and adjoint-state wavefields, and hence can greatly facilitate the FWI implementation (Castellanos et al., 2011).

To compute the gradient of the misfit function with the adjoint-state method, we introduce a new functional $\mathcal{L}$

$$\begin{aligned}\mathcal{L}\left(\mathbf{d}_{cal},\mathbf{v},\mathbf{m},\beta_1,\beta_2\right) \;=\;& \frac{1}{2} < \mathbf{d}_{cal}-\mathbf{d}_{obs}\,|\,\mathbf{d}_{cal}-\mathbf{d}_{obs} >_{\mathcal{D}} + \frac{1}{2}\sum_{i=1}^{N_p}\lambda_i\left(\mathbf{m}_i-\mathbf{m}_{prior_i}\right)^{\dagger}\mathbf{W}_{m_i}\left(\mathbf{m}_i-\mathbf{m}_{prior_i}\right)\\ & + \Re < \beta_1 \mid \mathbf{B}\mathbf{v}-\mathbf{s}_B >_{\mathcal{V}}, + \Re < \beta_2 \mid \mathbf{d}_{cal}-\mathbf{R}\mathbf{C}\mathbf{v} >_{\mathcal{D}}\,. \end{aligned} \quad (3.20)$$

where the complex-valued subspaces $\mathcal{D}$ and $\mathcal{V}$ span over the receiver positions and the full computational domain, respectively. The restriction operator **R** samples the pressure wavefield at the receiver positions. The inner product between $\mathbf{x}$ and $\mathbf{y}$ is denoted by $< \mathbf{x}|\mathbf{y} >_{\mathcal{D}} = \int_{\mathcal{D}} x^*\, y$. The Lagrangian function $\mathcal{L}$, Equation 3.12, corresponds to the misfit function $\mathcal{C}$ subject to the constraint that the state equations are satisfied. The state variables are the particle-velocity wavefields $\mathbf{v}$ and the pressure data $\mathbf{d}_{cal}$ at the receiver positions. The Lagrange multipliers β_1 and β_2 are the adjoint state variables. The two state equations allow us to introduce, on the one hand, the self-adjoint forward modeling operator **B** for particle velocities, and on the other hand, the relationships between the modeled particle velocities and the pressure data $\mathbf{d}_{cal}$, through the operator **C**. The operator **C** corresponds the third block row of Equation 3.14:

$$\mathbf{C} = \left[\frac{\iota}{\omega}\frac{c_{11}+c_{13}}{2}\frac{\partial}{\partial x} \quad \frac{\iota}{\omega}\frac{c_{13}+c_{33}}{2}\frac{\partial}{\partial z}\right]. \quad (3.21)$$

At the saddle points of $\mathcal{L}$ with respect to the state and adjoint-state variables, $\nabla\mathcal{C}_{\mathbf{m}} = \nabla\mathcal{L}_{\mathbf{m}}$, this gives for $\nabla\mathcal{C}_{\mathbf{m}}$:

$$\nabla\mathcal{C}_{\mathbf{m}} =< \beta_1 \mid \frac{\partial \mathbf{B}}{\partial \mathbf{m}}\mathbf{v} > + < \beta_2 \mid \mathbf{R}\frac{\partial \mathbf{C}}{\partial \mathbf{m}}\mathbf{v} > + \sum_{i=1}^{N_p}\lambda_i \mathbf{W}_m\left(\mathbf{m}-\mathbf{m}_{prior}\right). \quad (3.22)$$

If we assume that the medium properties at the receiver positions are known, the term $< \beta_2 \mid \mathbf{R}\frac{\partial \mathbf{C}}{\partial \mathbf{m}}\mathbf{v} >$ vanishes. The adjoint-state variables β_1 and β_2 satisfy the adjoint-state equations $\nabla \mathcal{C}_{\mathbf{v}} = \mathbf{0}$ and $\nabla \mathcal{C}_{\mathbf{d}_{cal}} = \mathbf{0}$:

$$\mathbf{B}^T \, \beta_1^* = \mathbf{C}^\dagger \mathbf{R}^T \beta_2^*, \qquad \beta_2 = \Delta \mathbf{d}. \tag{3.23}$$

Exploiting the symmetry of matrix $\mathbf{B}$ and injecting the expression of β_2 into the right-hand-side term of the first state equation gives:

$$\mathbf{B} \, \beta_1^* = \mathbf{C}^\dagger \mathbf{R}^T \Delta \mathbf{d}^*. \tag{3.24}$$

Equation 3.24 indicates that the adjoint wavefield β_1 can be computed by back-propagation of the pressure residuals, after conversion of these latter into particle velocities for interfacing with the second-order wave equation. The adjoint of $\mathbf{C}$ allows us to perform the pressure-to-particle-velocity conversion, and is inferred from the first and second block row of Equation 3.14:

$$\mathbf{C}^\dagger = \frac{\iota b}{\omega} \begin{bmatrix} \frac{\partial}{\partial x} & \frac{\partial}{\partial z} \end{bmatrix}^T, \tag{3.25}$$

where we have assumed that the wavefield q is zero at the receiver positions. Indeed, the aim of the operator $\mathbf{C}^\dagger$ in equation 3.24 is to represent the residual pressure source by two dipoles through the gradient operator. As the building of $\mathbf{B}$ is difficult with the discontinuous Galerkin method, we infer the state variables $\mathbf{v}$ and the adjoint-state variables β_1 from the numerical solutions of the first-order wave equation, Equation 3.15. For this, the correct relationship between an explosive source of the first-order and second-order wave equations, Equation 3.19, should be used to guarantee that the two wave equations give the same solutions. This leads to:

$$\mathbf{v} = \mathbf{P}\mathbf{R}_v \mathbf{w}, \quad \beta_1 = \mathbf{P}\mathbf{R}_v \beta_A, \tag{3.26}$$

where $\mathbf{w}$ and β_A are solutions of the first-order wave equation:

$$\mathbf{A}\,\mathbf{w} = \mathbf{s}_A, \quad \mathbf{A}\,\beta_A^* = \frac{b}{\iota\omega} \mathbf{R}_p^T \mathbf{R}^T \Delta \mathbf{d}^*. \tag{3.27}$$

The operator $\mathbf{R}_v$ and $\mathbf{R}_p$ are the restrictions of the velocity-stress vector to the particle velocities and to the pressure, respectively. The operator $\mathbf{P}$ projects the solutions of the discontinuous Galerkin method computed at the nodes of the P0/P1/P2 mesh onto the barycenter of the triangular elements, for consistency with the finite-volume discretization of the forward-modeling operator $\mathbf{B}$ (Brossier, 2011a). Of note, the pressure residuals on the right-hand side of Equation 3.27 need to be scaled by the factor $\frac{b}{\iota\omega}$ to derive the correct expression of the adjoint wavefield β_1. Plugging the expression of $\mathbf{v}$ and β_1, Equations 3.26 and 3.27, into the expression of the gradient, Equation 3.13, gives:

$$\nabla \mathcal{C}_{\mathbf{m}} = \Re \left\{ \left(\mathbf{P}\mathbf{R}_v \left(\mathbf{A}^{-1} \mathbf{s}_A \right) \right)^T \left(\frac{\partial \mathbf{B}}{\partial \mathbf{m}} \right)^T \mathbf{P}\mathbf{R}_v \left(\mathbf{A}^{-1} \frac{b}{\iota\omega} \mathbf{R}_p^T \mathbf{R}^T \left(\mathbf{R}\mathbf{R}_p \mathbf{A}^{-1} \mathbf{s}_A - \mathbf{d}_{obs} \right)^* \right) \right\}, \tag{3.28}$$

which is identical to equation 3.8 for one frequency and one shot.

3.3 Conclusion of this chapter

A general insight into FWI as a seismic imaging tool tells us that we encounter a problem which is ill-posed in mathematical point of view, and is a local optimisation problem, which depends on the starting conditions. Moreover, some physical (such as acoustic) and mathematical (such as the estimation of Hessian) approximations are considered in the core of numerical problem. Additionally, there are some computer resource limitation for large scale problems. Therefore, the most effort is to provide better conditions for solving this least-square, ill-posed and locally optimised problem. One accessory is the suitable parameterization of the model space. We succeed to provide a practical approach to choose a suitable parameterization for mono-parameter and multi-parameter FWI with an application to the imaging of acoustic VTI media. We showed that this practical approach is reliable because it is based on the sensitivity and trade-off analysis. We showed that the parameterization type has direct effect on the resolution of retrieved parameter by FWI. In the next chapter, we apply our method on large-scale synthetic and real case studies and show that these conclusions are in agreement with the results obtained in reality.

Chapter 4

Application of acoustic anisotropic FWI on Valhall field

Contents

4.1 Introduction

In this chapter, we present the application of acoustic VTI FWI on synthetic and real Valhall dataset. At first, the conclusions obtained from the sensitivity analysis from previous chapter put on application on the synthetic Valhall model. The acquisition geometry represents a wide-aperture acquisition survey. The anisotropic Valhall synthetic model is provided with most similarities to the line 21 of the real Valhall model. This helps to anticipate the application on the real case study. We apply the mono-parameter and multi-parameter acoustic VTI FWI on the synthetic Valhall model. We observe the possibilities of retrieving wavespeeds and Thomsen parameter for a surface acquisition geometry. The wavespeeds can be retrieved with reliable resolutions in terms of reconstructed wavenumbers. The challenge is the reconstruction of Thomsen parameters, which is only possible for long wavelength component of the data. Afterwards, the real Valhall dataset (line 21) is treated. The wavespeeds are reconstructed with good resolution by mono-parameter inversion. The multi-parameter inversion is performed only for two wavespeeds. The FWI models appraisal is performed via the reverse time migration and quality control of the flatness of common image gathers. This chapter contains an article, which has been submitted for publication in *Geophysics* journal.

4.2 Which parameterization for acoustic vertical transverse isotropic full waveform inversion? - Part 2: synthetic and real data case studies from Valhall

Which parameterization for acoustic vertical transverse isotropic full waveform inversion? - Part 2: synthetic and real data case studies from Valhall

Yaser Gholami, Romain Brossier, Stéphane Operto, Vincent Prieux, Alessandra Ribodetti and Jean Virieux

Geophysics, 2012, submitted

4.2.1 Summary

It is necessary to account for anisotropy in full waveform inversion (FWI) of wide azimuth and wide aperture seismic data in most geological environments, for correct depth positioning of reflectors, and for reliable estimations of wavespeeds as a function of the direction of propagation. In this framework, choosing a suitable anisotropic subsurface parameterization is a central issue in mono-parameter and multi-parameter FWI. This is because this parameterization defines the influence of each physical parameter class on the data as a function of the scattering angle, and hence the resolution of the parameter reconstruction, and on the potential trade-off between different parameter classes. We apply mono-parameter and multi-parameter frequency-domain acoustic vertical transverse isotropic FWI to synthetic and real wide-aperture data, representative of the Valhall oil field. We first show that reliable mono-parameter FWI can be performed to build a high-resolution velocity model (for the vertical, the horizontal or normal move-out velocity), provided that the background models of two Thomsen parameters describe the large wavelengths of the subsurface sufficiently accurately. Alternatively, we show the feasibility of the joint reconstruction of two wavespeeds (e.g., the vertical and horizontal wavespeeds), while Thomsen parameter δ is kept fixed during the inversion. The joint update of the two wavespeeds by FWI is feasible, as they have a significant influence on the data with the same order of magnitude, but for a distinct range of scattering angles. Therefore, there is a limited trade-off between the two wavespeeds. The counterpart is that the influence of the wavespeeds on the data for a limited range of scattering angles can significantly hamper the resolution with

which the two wavespeeds are imaged. These conclusions inferred from the application to the real data are fully consistent with those inferred from the theoretical parameterization analysis of acoustic vertical transverse isotropic FWI performed in the companion report.

Keywords

wave propagation, frequency-domain, anisotropy, vertical transverse isotropic anisotropy, seismic imaging, full waveform inversion, multi-parameters inversion

4.2.2 Introduction

Full waveform inversion (FWI) is a data-fitting approach that is theoretically amenable to inversion of the full wavefield recorded by wide-azimuth and wide-aperture acquisition geometries (see Virieux and Operto (2009) for a review). For such acquisitions, it can be critical to account for the intrinsic variations of the wavespeed with the direction of propagation during the inversion: namely, the anisotropy. The footprint of anisotropy in FWI of wide-azimuth and wide-aperture acquisitions was shown by Plessix and Perkins (2010) and Prieux et al. (2011). Plessix and Perkins (2010) showed how accounting for anisotropy in seismic modeling during FWI allows improvement of the FWI velocity model, which is subsequently used as the background model for reverse time migration. The results of the reverse time migration computed in the anisotropic FWI velocity model show flat common image gathers below salt structures. Prieux et al. (2011) compared the results of isotropic and anisotropic acoustic FWI of wide-aperture ocean-bottom cable (OBC) data from the Valhall field. When isotropic FWI is performed, they showed that horizontal velocities are reconstructed in the upper structure, because FWI is mostly driven by the diving waves and the post-critical reflections, which have a direction of propagation close to horizontal. The reconstruction of the horizontal velocities in the upper structure leads to incorrect velocities in the underlying gas layers and mispositioning of the cap rock on the top of the reservoir level. These deep structures are mainly sampled by short-spread reflections, which are mostly sensitive to the normal move-out (NMO) and vertical velocities.

Most of the anisotropic FWI applications aim to update the vertical or NMO velocity, while keeping the Thomsen anisotropic parameters fixed (i.e., δ, ϵ or a combination of the two, as represented by the anellipticity parameter η (Alkhalifah and Tsvankin, 1995)) (Vigh et al., 2010; Plessix and Perkins, 2010; Prieux et al., 2011). Only few attempts have been made to up-date multiple classes of parameters by anisotropic FWI during synthetic experiments in acoustic or elastic vertical transverse isotropic (VTI) approximations (Ji and Singh, 2005; Barnes et al., 2008; Lee et al., 2010; Plessix and Cao, 2011a).

In the initial companion report (Gholami et al., 2012b, this issue), we investigated which classes of parameter need to be selected for FWI, based on scattering pattern analysis with respect to the illumination angle, grid analysis of the misfit function, and simple synthetic FWI examples. In the current study, we continue this investigation through synthetic and real data case studies of two-dimensional mono-parameter and multi-parameter acoustic VTI FWI. For the synthetic and real data case studies, the acquisition geometry represents a long-offset ocean-bottom survey, and the geological target is the Valhall oil field, where significant anisotropy has been reported (Thomsen et al., 1997; Kommedal et al., 1997; Barkved and Heavey, 2003). Our aim is to test whether the previous inferred conclusions that have provided clear insight into the sensitivity of the data to the model parameters as well as into the potential trade-off between different parameter classes can be validated against the application of FWI to a realistic synthetic example and to a real data case study.

We considered three kinds of parameterization: Type 1 parameterization combines one wavespeed (the vertical, the horizontal or the NMO velocity, V_{P0}, V_h, V_{NMO}, respectively) with two Thomsen parameters (Thomsen, 1986) (ϵ, δ or their combination, parameter η [see above] (Alkhalifah and Tsvankin, 1995)). We note that long-spread reflection traveltimes in homogeneous acoustic VTI media are governed by the NMO velocity and η. With type 1 parameterizations, the wavespeed has a dominant influence on the data for the full range of scattering angles, while δ has a minor influence on the data. Thomsen parameter ϵ has a higher influence on the data than δ, and shows significant trade-off with the wavespeed. Alternatively, two wavespeeds, for example the vertical and horizontal velocities, and one Thomsen parameter, for example δ, can be used in the subsurface parameterization (referred to as type 2 parameterization). In this case, the two wavespeeds have significant influence on the data for distinct ranges of scattering angles. For example, the vertical velocity has influence on the small and intermediate scattering angles, while the horizontal velocity has influence on the intermediate and large scattering angles, when they are combined with each other in the parameterization. In this case, the trade-off between the two wavespeeds should be manageable, although it exists at intermediate scattering angles (Gholami et al., 2012b, their Figure 3e,f). However, the resolution with which the two velocity models are reconstructed will be hampered by the narrow range of scattering angles for which they have influence. It is worth noting that together with the frequency of the source, the scattering angle controls the resolution power of the FWI, as has been shown in the general framework of inverse scattering theory (Miller et al., 1987; Wu and Toksöz, 1987; Sirgue and Pratt, 2004). The large scattering angles govern the reconstruction of the large wavelengths of the subsurface, while the small scattering angles govern the reconstruction of the short wavelengths of the subsurface, for a maximum resolution of half of the wavelength in the framework of the single-scattering approximation. The third potential parameterization involves the three elastic coefficients c_{11}, c_{33} and c_{13}, which shows similar behavior to the type 2 parameterization. This parameterization will not be investigated in this study, as it did not show results as good as those obtained with type 2 parameterization, as we have verified.

In the present study, we apply frequency-domain acoustic VTI FWI to synthetic and real wide-aperture data, which are representative of the Valhall target. As the FWI is performed in the acoustic approximation, only the pressure wavefield recorded by the hydrophone component is used. In the first part of this study, we briefly review the FWI algorithm that we use. A more detailed review is presented in Gholami et al. (2012b, this issue). In the second part, we present the application of FWI to the synthetic Valhall model. We first perform isotropic FWI of anisotropic data, to show the footprint of anisotropy on isotropic FWI of wide-aperture data in a similar manner to Prieux et al. (2011) for the real data case study. Then, we apply mono-parameter and multi-parameter FWI to the synthetic data using the type 1 and type 2 parameterizations. We first show that a reliable velocity model can be built by mono-parameter FWI, provided that large-scale background models of the Thomsen parameters are available. The joint update of the vertical or NMO velocity and ϵ or η is possible, although the perturbations of the Thomsen parameter are quite small. This confirms that the data are sensitive only to the large wavelengths of ϵ or η when these parameters are combined with the vertical or NMO velocities in the parameterization. We then show the feasibility of the joint update of two wavespeeds when using type 2 parameterization, while Thomsen parameter δ is kept fixed during the inversion. As expected, the horizontal velocities are reconstructed with low resolution because their influence is confined to the large scattering angles when type 2 parameterization is considered. In the last part of this study, these inversion tests are applied to the real data from Valhall, and they confirm the conclusions inferred from the theoretical analysis shown in Gholami et al. (2012b, this issue) and from the synthetic case study presented in this study.

4.2.3 Full inversion inversion in vertical transverse isotropic acoustic media

In the present study, we perform FWI in the frequency domain with the methodology described by Brossier (2011a). A detailed review of the algorithm is presented in the companion paper (Gholami et al., 2012b, this issue). Only a short review is given here. Seismic modeling in VTI acoustic media is performed with a velocity-stress discontinuous Galerkin finite-element method on unstructured triangular mesh (Brossier et al., 2008, 2010b; Brossier, 2011a). The acoustic approximation is implemented by setting the c_{44} elastic coefficient on the symmetry axis in the P-SV elastodynamic system to zero (Brossier et al., 2010b). The pressure wavefield is taken as the mean of the normal stress components. Inversion is performed following a multiscale approach that proceeds sequentially over increasing frequencies. The local optimization relies on the limited-memory Broyden-Fletcher-Goldfard-Shanno (L-BFGS) quasi-Newton algorithm that provides recursively an approximation of the product of the inverse of the Hessian with the gradient of the misfit function (Nocedal, 1980). The FWI is regularized with a Tikhonov regularization (Tikhonov and Arsenin, 1977).

The least-squares misfit function $\mathcal{C}(\mathbf{m})$ is given by:

$$\mathcal{C}(\mathbf{m}) = \frac{1}{2}\Delta\mathbf{d}^{\dagger}\mathbf{W}_d\Delta\mathbf{d} + \frac{1}{2}\sum_{i=1}^{N_p}\lambda_i\left(\mathbf{m}_i - \mathbf{m}_{prior_i}\right)^{\dagger}\mathbf{W}_{m_i}\left(\mathbf{m}_i - \mathbf{m}_{prior_i}\right), \tag{4.1}$$

where the expression $\Delta\mathbf{d} = \mathbf{d}_{cal}(\mathbf{m}) - \mathbf{d}_{obs}$ denotes the data residual vector, the difference between the modeled data $\mathbf{d}_{cal}(\mathbf{m})$ and the recorded data $\mathbf{d}_{obs}$. In the present study, we consider only the pressure wavefield recorded by the hydrophone component. The multiparameter subsurface model is denoted by $\mathbf{m} = \left(\mathbf{m}_1, ..., \mathbf{m}_{N_p}\right)$, where N_p is the number of parameter classes to be updated during the FWI. The acoustic VTI medium is parametrized by three parameter classes: for example, the vertical velocity and Thomsen parameters δ and ϵ (Thomsen, 1986) and therefore $N_p \leq 3$. The weighting matrices $\mathbf{W}_{m_i}$ seek to penalize the roughness of the difference between the model $\mathbf{m}$ and the prior model $\mathbf{m}_{prior}$. The smoothing operators $\mathbf{W}_{m_i}^{-1}$ are exponential functions given by

$$\mathbf{W}_{m_i}^{-1}(z,x,z',x') = \sigma_i^2(z,x)exp\left(\frac{-|x-x'|}{\tau_x}\right)exp\left(\frac{-|z-z'|}{\tau_z}\right), \tag{4.2}$$

where τ_x and τ_z denote the horizontal and vertical correlation lengths, respectively, defined as a fraction of the local wavelength. The coefficient σ_i represents the standard error, and it is scaled to the physical unit of the parameter class i. An exponential function is used for $\mathbf{W}_{m_i}^{-1}$ because its inverse in the expression of the misfit function, equation 4.1, can be computed analytically (Tarantola, 1987, pages 308-310). Data preconditioning can be applied through the weighting matrix $\mathbf{W}_d$, which weights each component of the data-misfit vector. The scalar hyper-parameters λ_i control the respective weight of the data-space and model-spaced misfit functions in Equation 4.1. Their values can be adapted to each parameter class, although this strategy will not be investigated in the present study because it might require intensive trial-and-error investigations.

The perturbation model that minimizes the misfit function at iteration k is given by:

$$\Delta \mathbf{m}_k = -\gamma_k \left[\frac{\partial^2 C(\mathbf{m}_k)}{\partial \mathbf{m}^2}\right]^{-1} \frac{\partial C(\mathbf{m}_k)}{\partial \mathbf{m}}, \tag{4.3}$$

with $\mathbf{m}_{k+1} = \mathbf{m}_k + \Delta \mathbf{m}_k$ and γ_k as the step length estimated by line search. The estimation of the step length is required by the local quadratic approximation of the misfit function. The first and second derivatives of the misfit function on the right-hand side of Equation 4.3 are the gradient and the Hessian of the misfit function.

The gradient of the data-space misfit function (the first term in Equation 4.1) is computed with the adjoint-state method (Plessix, 2006). We found (Gholami et al., 2012b, their Appendix A):

$$\nabla \mathcal{C}_{\mathbf{m}} = \sum_{\omega} \sum_{s} \Re \left\{ \left(\mathbf{P}\mathbf{R}_v \left(\mathbf{A}^{-1}\mathbf{s}_A\right)\right)^T \left(\frac{\partial \mathbf{B}}{\partial \mathbf{m}}\right)^T \mathbf{P}\mathbf{R}_v \left(\mathbf{A}^{-1}\frac{b}{\imath\omega}\mathbf{R}_p^T\mathbf{R}^T\left(\mathbf{R}\mathbf{R}_p\mathbf{A}^{-1}\mathbf{s}_A - \mathbf{d}_{obs}\right)^*\right)\right\} \tag{4.4}$$

where the transpose is denoted by ^{T} and the conjugate by *. In Equation 4.4,the expression of the gradient is derived from the second-order wave-equation for particle velocities (Gholami et al., 2012b, their equations 5 and 6), with the associated forward-problem operator $\mathbf{B}$ discretized with a simple finite-volume scheme to compute the scattering pattern $\frac{\partial \mathbf{B}}{\partial \mathbf{m}}$ in Equation 4.4 (Brossier et al., 2008; Brossier, 2011a). The second-order operator allows us to efficiently compute the gradient of the misfit function from the self-adjoint operator with significant memory saving as we only store particle velocities. The corresponding incident and adjoint wavefields can be estimated using this operator, but we find it easier to compute these wavefields with the velocity-stress first-order wave equation (Gholami et al., 2012b, their equations 3 and 4). This will give the correct expression of the gradient as long as the correct relationship between the source excitations of the first-order system and the second-order system is used (Gholami et al., 2012b, their equation 7). In Equation 4.4, the operator $\mathbf{A}$ is the discontinuous Galerkin velocity-stress forward problem operator that is used to compute the incident wavefield with the adequate source $\mathbf{s}_A$ and the back-propagated adjoint wavefields with the adequate residual source $\frac{b}{\imath\omega}\mathbf{R}_p^T\mathbf{R}^T\left(\mathbf{R}\mathbf{R}_p\mathbf{A}^{-1}\mathbf{s}_A - \mathbf{d}_{obs}\right)^*$ (back-propagation of the adjoint wavefield is indicated by the conjugate of the residual source). The restriction operators $\mathbf{R}_v$ and $\mathbf{R}_p$ extracts the velocity components and the pressure component from the velocity-stress vector, respectively. The restriction operator $\mathbf{R}$ extracts the values of the pressure wavefield at the receiver positions.

4.2.4 Realistic synthetic Valhall case study

4.2.4.1 Model and data

The two-dimensional synthetic Valhall velocity model (Figure 4.1a) is a realistic synthetic model that is inspired by the geology of the Valhall oil and gas field located in the North Sea (Munns, 1985; Leonard and Munns, 1987). The experimental setup of the synthetic Valhall inversion is performed as close as possible to the real-data case study, which will be discussed later. The vertical velocity model shows a series of horizontal gas layers between 1.5 km and 2.5 km in depth, which hamper the imaging of the underlying reservoir at 2.5 km in depth (Thomsen et al., 1997; Kommedal et al., 1997; Barkved and Heavey, 2003). The δ and ϵ models are show in Figure 4.1b,c. In some areas, the anisotropy is significant, with the horizontal velocity 15% higher than the vertical velocity as revealed by the values of the parameter η (Figure 4.1d). Time-domain synthetic seismograms computed in the true VTI model are shown in Figure

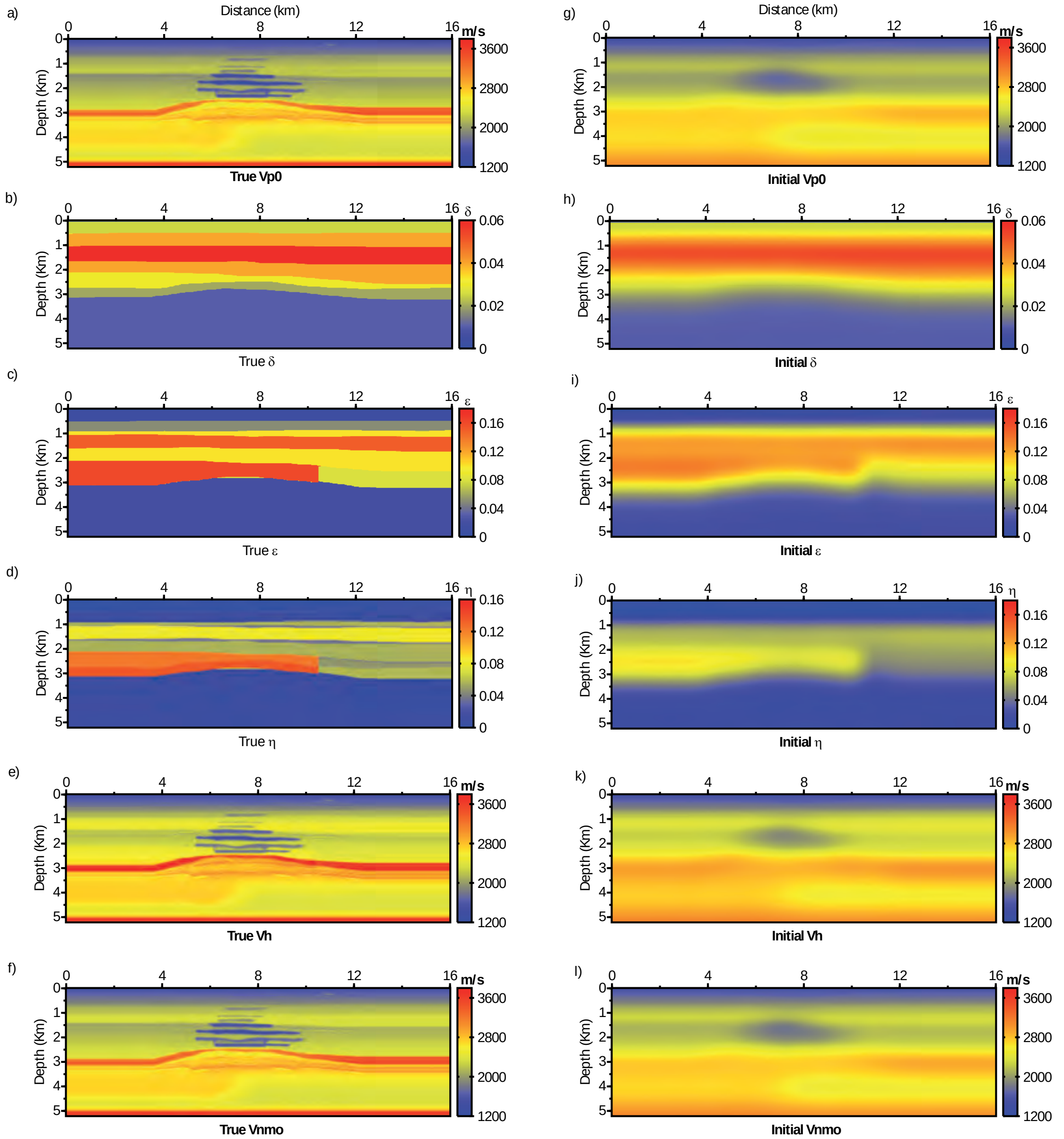


Figure 4.1: Valhall synthetic model. (a) Vertical velocity model. (b,c) Models of Thomsen parameters δ (b) and ϵ (c). (d) Model of η. (e,f) Horizontal (e) and NMO (f) velocity models. (g,l) Same as (a,f) for the smooth initial FWI models.

4.2b for a shot located at 12 km distance. For phase interpretation, we compute isotropic first-arrival ray-tracing in the vertical velocity model (Figure 4.2a). First-arrival rays propagate in the upper structure above the gas layers and at the reservoir level. As the gas layers form on average a low-velocity zone with respect to the surrounding layers, no first-arrival rays pass through them. The main arrivals are diving waves that propagate above the gas layers (Figure 4.2b, D1, D2, D3), a weak-amplitude head wave from the top of the reservoir (Figure 4.2b, Dr), and a series of reflections from the upper sedimentary layers, the top of the gas layers, the top and bottom of the reservoir, and the high velocity interface located at 5000 m in depth (Figure 4.2b, Rs, Rg, Rr, R5). The head wave that propagates along the deep interface at 5 km in depth is also visible as a secondary arrival (Figure 4.2b, D5). The data content suggests that the FWI reconstruction of the target above the gas layers will be driven by both diving waves and pre-critical and post-critical reflections, while the reconstruction at the reservoir level will be dominated by pre-critical reflections. Although post-critical reflections from the reservoir level are shown, complex interactions of these phases with shingling multi-refracted phases in the shallow part probably makes the inversion of the post-critical reflections from the reservoir more challenging. Since the upper target is sampled by waves that propagate with a wide range of directions, the footprint of the anisotropy is expected to be significant in this part of the target (Prieux et al., 2011).

4.2.4.2 Full waveform inversion setup

The acquisition geometry considered for FWI is representative of an OBC survey with receivers on the sea bottom at 71 m in depth (water depth is 70 m) and sources just below the water surface at 6 m in depth. Only the pressure wavefield (hydrophone component) is considered for the acoustic VTI inversion. The maximum offset in the acquisition is 15.7 km, which provides a wide-aperture surface dataset. The shot and receiver spacings are 50 m. Five frequency components between 2 Hz and 6 Hz (2,3,4,5 and 6Hz) are inverted successively with the quasi-Newton L-BFGS optimization method (Nocedal, 1980). The descent direction is preconditioned with a Gaussian smoothing of the misfit-function gradient and of the Hessian. The prior model in Equation 4.1 is set to the current model, $\mathbf{m}_{prior} = \mathbf{m}_k$. The initial guess of the Hessian used during the recursive building performed by L-BFGS contains the diagonal terms of the linear part of the Hessian damped by the pre-whitening factors λ_i, Equation 4.1 (Gholami et al., 2012b, this issue). We use the same value for λ_i, whatever the parameter class i. This value, denoted by λ is set to 1 % of the maximum diagonal coefficients. The starting models for inversion are built by Gaussian smoothing of true models (Figure 4.1g-l). As discussed in Gholami et al. (2012b), the pre-whitening factor λ of the Hessian is scaled to the parameter with the dominant influence on the data. If the influence of the secondary parameters is one order of magnitude lower than that of the dominant parameter, the pre-whitening factor can prevent the correct scaling of the gradient of the secondary parameters. This can lead to underestimated model perturbations. However, the definition of suitable damping adapted to each parameter class is a difficult issue, in particular in the presence of noise, and this will be left for a future study.

4.2.4.3 Isotropic full waveform inversion of anisotropic data

The footprint of anisotropy on the isotropic FWI of the anisotropic wide-aperture data was investigated and demonstrated for then Valhall field dataset by Prieux et al. (2011). They

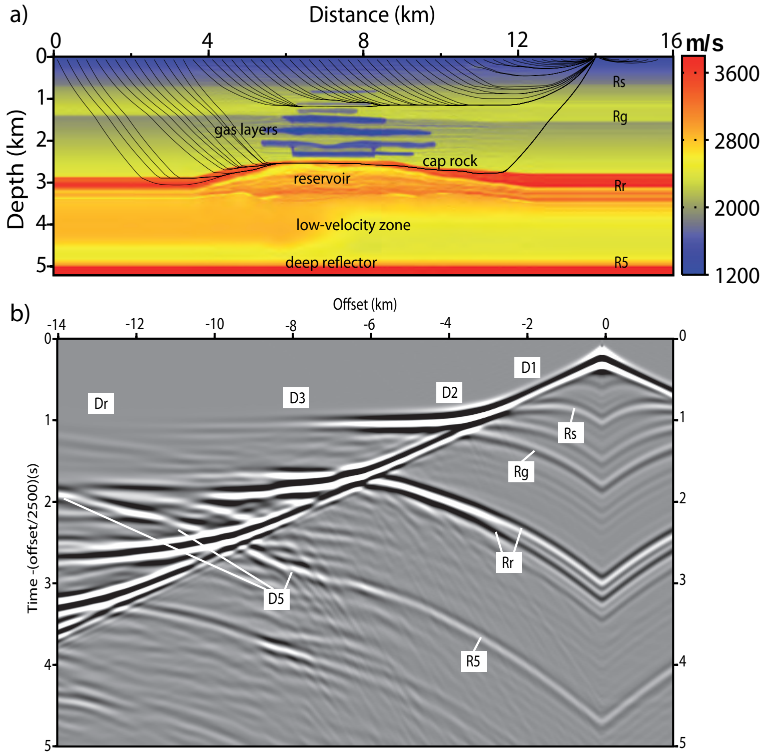


Figure 4.2: Valhall synthetic model. (a) Main structural units: gas layers between 1400 m and 2400 m in depth, the reservoir at 2500 m in depth, a flat reflector at 5000 m in depth. Isotropic first-arrival rays are superimposed on the model for a source at 14 km in distance. (c) Acoustic VTI synthetic seismograms computed in the models shown in Figure 4.1(a-c). The time axis is shown with a reduction velocity of 2.5 km/s. Phase nomenclature: D1, D2, D3: diving waves in the upper structure above the gas layers. Dr: head wave from the top of the reservoir. Rs: shallow reflection from the reflector located at around 700 m in depth. Rg: reflection from the top of the gas layers. Rr: reflections from the top and bottom of the reservoir. D5, R5: head wave and reflection from the interface located at 5000 m in depth.

showed that the isotropic approximation of the modeled data in the FWI produces kinematic inconsistencies between short-aperture and wide-aperture components of the data, due to the difference between the horizontal and the vertical velocities, which is represented by the non-ellipticity parameter η. Prieux et al. (2011) showed that horizontal velocities are mainly reconstructed in the upper part of the target, where the inversion is dominantly driven by diving waves and post-critical reflections. The reconstruction of high horizontal velocities in the upper structure leads to significant mispositioning of deep reflectors, which are mainly sampled by short-aperture reflections due to the limited spread of the acquisition geometry. As the short-aperture reflections are sensitive to the vertical and NMO velocities, deep reflectors are imaged at incorrect depths, to balance the effects of the high velocities reconstructed in the upper target. We perform isotropic FWI of the anisotropic synthetic data computed in the Valhall model to validate this interpretation against a realistic synthetic example. Mono-parameter VTI FWI for vertical velocity performed with the (V_{P0},δ,ϵ) parameterization provides a reference result to assess the footprint of the anisotropy on the isotropic FWI (Figures 4.3a, 4.4a). The FWI vertical-velocity model shows the correct positioning of reflectors and a reliable estimation of the velocity amplitudes down to the reservoir level. Below 2.5 km in depth, the intermediate wavelengths of the velocity structure are not well recovered, which might result from a lack of large scattering-angle coverage at these depths (Figure 4.4a). The reliable vertical velocity model that is built by mono-parameter FWI is a first evidence that the vertical wavespeed has a dominant influence on the data. Moreover, the data are mostly sensitive to the large wavelengths of the Thomsen parameters when the (V_{P0},δ,ϵ) parameterization is used, and these large wavelengths are already present in the background models of the Thomsen parameters. The isotropic FWI models are computed using three different starting models

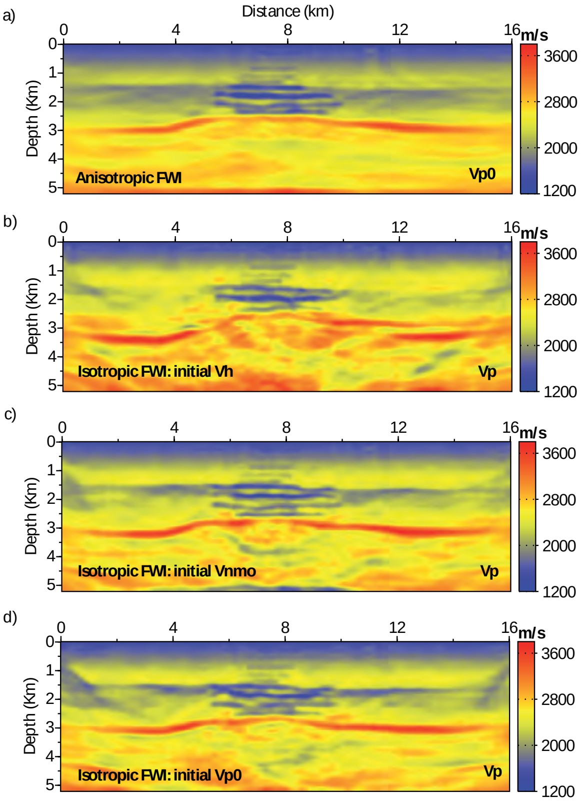


Figure 4.3: Anisotropic *versus* isotropic FWI of anisotropic data. (a) Vertical velocity model built by anisotropic mono-parameter FWI using the (V_{P0},δ,ϵ) parameterization. (b-d) Isotropic FWI model. The initial models are (b) the smooth horizontal velocity model (Figure 4.1k), (c) the smooth NMO velocity model (Figure 4.1l), and (d) the smooth vertical velocity model (Figure 4.1g).

that correspond to the smooth horizontal velocity model, the NMO velocity model, and the vertical velocity model, as shown in Figure 4.1g,k,l. Each FWI model shows that high velocities are reconstructed above the gas layers, with the values even higher than the horizontal velocities at around 1.3 km to1.4 km in depth (Figure 4.4b-d, arrow). These velocities look like those built by the isotropic FWI of the real Valhall data (Prieux et al., 2011, their figure 7c,d). These overestimated velocities above the gas layers lead to erroneous velocities in the gas layers between 1.5 km and 2.5 km in depth, and the deepening of the reservoir reflector at 2.5 km in depth. The reflector at the reservoir level is positioned at the correct depth in the middle of the FWI model built from the initial horizontal velocity model (Figure 4.4b,

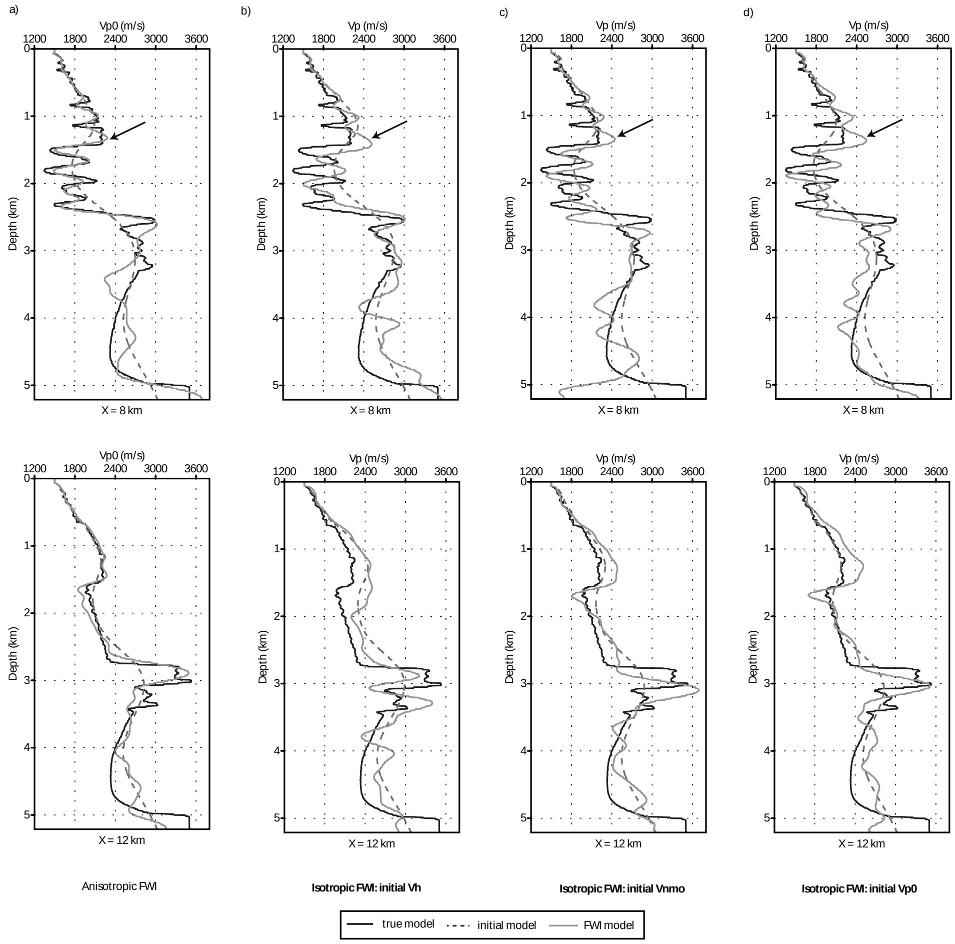


Figure 4.4: Logs of the true model (solid black line), the initial model (dash gray line), and the FWI model (solid gray line). The top panel and the bottom panel correspond to logs, which are located at 8 km and 12 km in distance, respectively. (a) Anisotropic FWI model of Figure 4.3a. (b,d) Isotropic FWI models of Figures 4.3b-d. The initial models are the smooth horizontal velocity model (b), the smooth NMO velocity model (c), and the smooth vertical velocity model (d). The arrow indicates where the isotropic FWI velocities are significantly greater than the true vertical velocity, a clear footprint of anisotropy.

top panel), while this reflector is mispositioned in depth for the FWI models that are built from the initial vertical velocity model and from the NMO velocity model (Figure 4.4c,d). A possible interpretation might be that the horizontal velocity model allows the kinematics of the diving waves and post-critical reflections at long offset (the most nonlinear part of the data) to be explained with sufficient accuracy to prevent cycle-skipping artifacts, unlike the vertical velocities and the NMO velocities (Prieux et al., 2011). Inversion of post-critical reflections from the reservoir might allow the reservoir to be positioned at the correct depths when the initial model represents horizontal velocities. However, this comment does not apply near the ends of the model, where the most significant mispositioning of the reservoir reflector is shown for the initial horizontal velocity model. Direct comparisons between the recorded data and the data modeled in the FWI models show that the isotropic FWI model built from the initial horizontal velocity model leads to the worst match of the reflection from the reservoir, with the modeled reflection phase arriving too early (Figure 4.5b, gray arrow). This probably highlights the footprint of the initial horizontal velocity model, the velocities of which are too high to match the kinematic of the deep reflections. Of note, the seismograms computed in the isotropic FWI models inferred from the initial vertical velocity model and from the NMO velocity model (Figure 4.5c,d) match the seismograms computed in the true model almost as well as the seismograms computed in the anisotropic FWI model (Figure 4.5a). This highlights the ill-posedness of the acoustic VTI FWI in terms of the nonunicity of the solution, as already pointed out by Prieux et al. (2011). We note, however, some degradation of the match of the secondary arrivals, which might correspond to reflections from the gas layers, when the FWI is performed in the isotropic approximation (Figure 4.5b-d, black arrows).

4.2.4.4 Mono-parameter anisotropic full waveform inversion

We consider a subsurface VTI parameterization that consists of one wavespeed (vertical velocity, horizontal velocity, and NMO velocity) and two Thomsen parameters δ and ϵ. We showed in Gholami et al. (2012b, this issue) that for this kind of parameterization, the wavespeed has a dominant influence on the wavefield for the full range of scattering angles. The Thomsen parameter δ is the parameter with the most negligible influence on the data at intermediate scattering angles. The Thomsen parameter ϵ has an influence on the data at large-to-intermediate or small-to-intermediate scattering angles, depending whether ϵ is associated with the vertical velocity or the horizontal velocity in the subsurface parameterization. According to the dominant influence of the wavespeed on the wavefield, a possible strategy is to first perform a mono-parameter FWI to reconstruct a high-resolution model of the wavespeed, keeping the δ background model and the ϵ background model fixed. For doing so, we shall assume that the background models that describe the large wavelengths of δ and ϵ are available (Figure 4.1h,i). The final models of the mono-parameter FWI for the vertical velocity, for the horizontal velocity, and for the NMO velocity are shown in Figure 4.6a-c. Comparisons between the logs extracted from the true models and from the FWI models show that reliable velocity models are obtained for each parameterization (Figure 4.7). These results confirm that prior knowledge of the large wavelengths of δ and ϵ can be sufficient to build high-resolution velocity models by mono-parameter anisotropic FWI. For highlighting the sensitivity of the mono-parameter FWI to the subsurface parameterization, we perform mono-parameter FWI for the vertical velocity using the (V_{P0},δ,V_h) parameterization instead of the (V_{P0},δ,ϵ) parameterization (Figure 4.6d). The background model for the horizontal velocity is shown in Figure 4.1k, and it is built from the smooth background model of the vertical velocity and from the smooth background model of δ. Unstable reconstruction of the vertical velocity with overestimated perturbations is shown in this case (Figures 4.6d, 4.7d). The unstable imaging of the vertical velocity results, on the one hand, from the trade-off between the vertical and horizontal velocities at intermediate scattering angles, and on the other hand, from the significant influence of the horizontal velocity on the data (Gholami et al., 2012b, their figures 3 and 14). As the horizontal velocity background model is kept fixed during inversion, the influence of the horizontal velocities on the wavefield is interpreted as the vertical velocity perturbations. This highlights the higher sensitivity of the data to parameter V_h, relative to parameter ϵ.

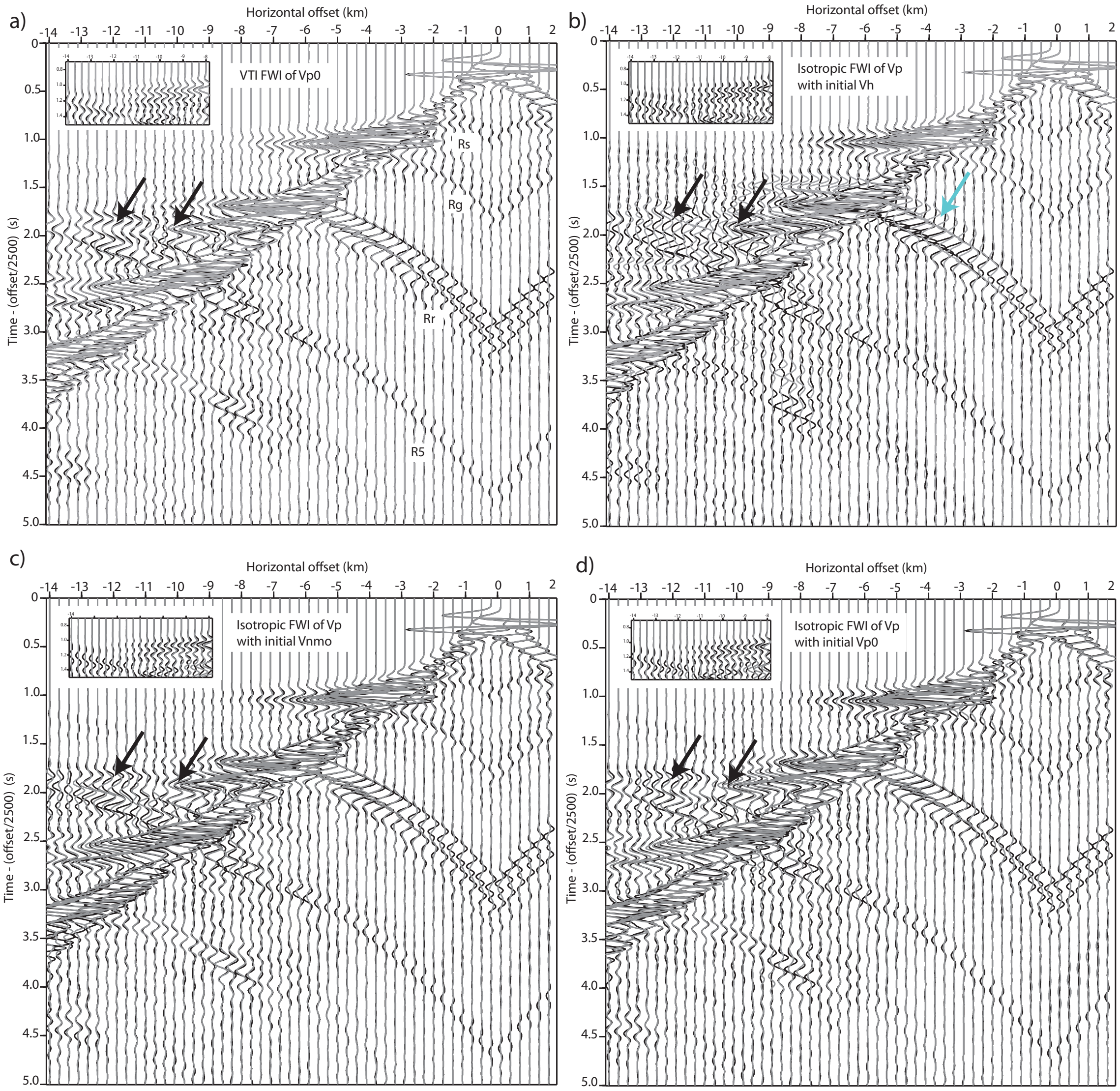


Figure 4.5: Comparison between time-domain seismograms computed in the true anisotropic model (black lines) and in the FWI models (gray lines). (a-d) Synthetic seismograms were computed in the FWI models of Figure 4.3a-d. The inset shows a close-up of the seismograms centered on the first arrival between -14 km and -8 km in offset. Black arrows indicate secondary phases (possible reflections from the gas layers), for which the data match is degraded when the isotropic approximation is used during FWI. The gray arrow indicates reflection from the top of the reservoir, that arrives too early, when the initial model represents horizontal velocities.

4.2.4.5 Joint multi-parameter anisotropic full waveform inversion

We now apply multi-parameter FWI for joint reconstruction of two anisotropic parameters, while the smooth background model of δ is kept fixed during the FWI iterations due to its small influence on the data (Figure 4.8). The initial models are the smooth background models shown in Figure 4.1g-l. We first update the vertical velocity and parameter ϵ using parameterization (V_{P_0}, δ, ϵ). In this case, the vertical velocity model is slightly improved, in particular at the

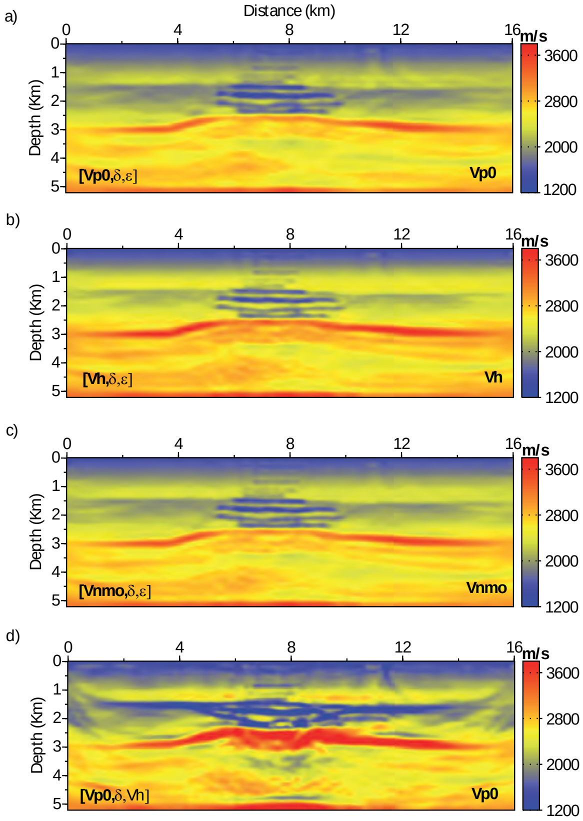


Figure 4.6: Mono-parameter VTI FWI. (a-c) The vertical velocity model (a), the horizontal velocity model (b) and the NMO velocity model (c) are reconstructed by mono-parameter FWI for the type 1 parameterization that combines one wavespeed and two Thomsen parameters δ and ϵ. (d) The vertical velocity model is reconstructed by mono-parameter FWI for the (V_{P_0}, δ, V_h) parameterization.

reservoir level, compared to the model obtained by mono-parameter FWI (compare Figures 4.6a and 4.8a, and Figures 4.7a and 4.9a). The ϵ model is weakly updated by FWI accordingly (Figures 4.8b, 4.9b), which is consistent with the reliable reconstruction of the vertical velocity performed by the mono-parameter FWI. The weak perturbation of ϵ is also justified from a more theoretical viewpoint, because the data are sensitive to the large wavelengths of the parameter ϵ only, when the parameterization (V_{P_0}, δ, ϵ) is used (Gholami et al., 2012b, their figure 3b, black dashed curve). These large wavelengths are already accurately represented in the initial model. We also perform the joint update of the NMO velocity and η using the (V_{P_0}, δ, η) parameterization (Figures 4.8c,d and 4.9c,d). The results show the same trend as for the (V_{P_0}, δ, ϵ) parameterization: the velocity model is updated with high resolution, while the η model is kept almost unmodified. The negligible perturbation of parameter η, which is even smaller than that of parameter ϵ, is consistent with the smaller influence of parameter η relative to parameter ϵ for large scattering angles (Gholami et al., 2012b, their figure 3b,c). This smaller influence of parameter η for large scattering angles is balanced by the greater influence of parameter δ at small scattering angles when the (V_{P_0}, δ, η) parameterization is used. The greater influence of parameter δ for this parameterization, which is kept fixed during the inversion, does not impact on the reconstruction of the NMO velocity in Figure

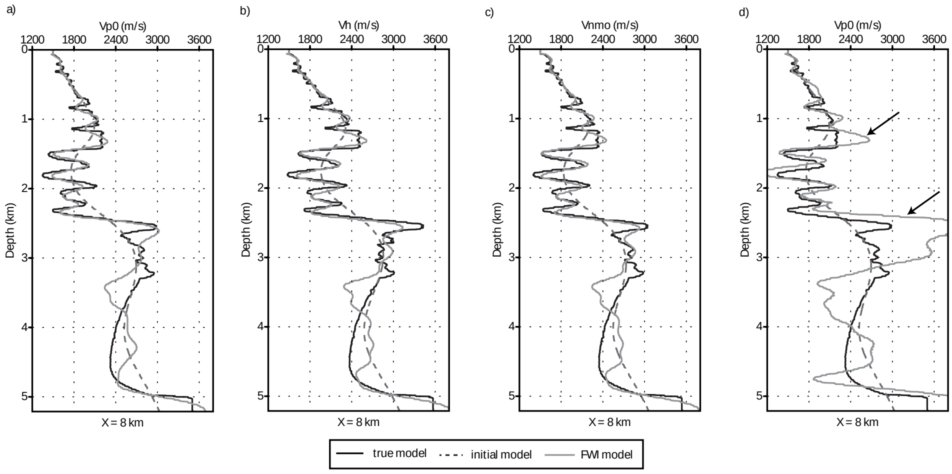


Figure 4.7: Vertical logs of the FWI models shown in Figure 4.6. The log is located at 8 km in distance in the middle of the model. The logs of the true model and of the initial models are plotted with solid black and dash gray lines, respectively. (a) Vertical velocity FWI model for the (V_{P_0}, δ, ϵ) parameterization. (b) Horizontal velocity FWI model for the (V_h, δ, ϵ) parameterization. (c) NMO velocity FWI model for the (V_{NMO}, δ, ϵ) parameterization. (d) Vertical-velocity FWI model for the (V_{P_0}, δ, V_h) parameterization.

4.8c. The results of multi-parameter FWI performed with the (V_{P_0}, δ, V_h) and (V_{NMO}, δ, V_h) parameterizations confirm that two wavespeeds (V_{P0} plus V_h and V_{NMO} plus V_h, respectively) can be jointly updated during FWI (Figure 4.8e-h and Figure 4.9c,d). The feasibility of the joint reconstruction of the two wavespeeds is consistent with the radiation patterns of the two wavespeeds when combined with each other in the subsurface parameterization, which shows a significant influence of the two wavespeeds on the data for a nearly non-overlapping band of scattering angles (Gholami et al., 2012b, their figure 3e, solid gray and black dash curves). The vertical velocity model and the NMO velocity model are, however, less accurate than those inferred from the type 1 parameterization ((V_{P_0}, δ, ϵ) and (V_{NMO}, δ, ϵ)) (compare Figure 4.9a and 4.9c, and Figure 4.9b and 4.9d). Moreover, the FWI horizontal velocity model has a low wavenumber content, and hence lacks resolution (Figure 4.9c,d). This is again consistent with the radiation pattern analysis presented in Gholami et al. (2012b, this issue): the vertical velocity or the NMO velocity have a radiation pattern that spans over the full range of scattering angles when combined with the two Thomsen parameters in the subsurface parameterization (Gholami et al., 2012b, their figure 3b,d, gray curve), which leads to high-resolution reconstruction of the velocity field, while the radiation pattern has a narrower coverage of the short-scattering angles when the vertical velocity or the NMO velocity are associated with the horizontal velocity in the subsurface parameterization (Gholami et al., 2012b, their figure 3(e-f), gray curve). The lack of resolution of the horizontal velocity model is consistent with the radiation pattern of the horizontal velocity in the (V_{P_0}, δ, V_h) and (V_{NMO}, δ, V_h) parameterizations. The influence of the horizontal velocity for large scattering angles governs the large wavelength reconstruction of the subsurface (Gholami et al., 2012b, their figure 3(e-f), black dashed curve). The seismograms computed in the final multi-parameter FWI models obtained with the (V_{P_0}, δ, ϵ) and (V_{P_0}, δ, V_h) parameterizations show similar agreement with the seismograms computed in the true model (Figure 4.10). Up to this point, we have assumed that smooth background models of the Thomsen parameters were available

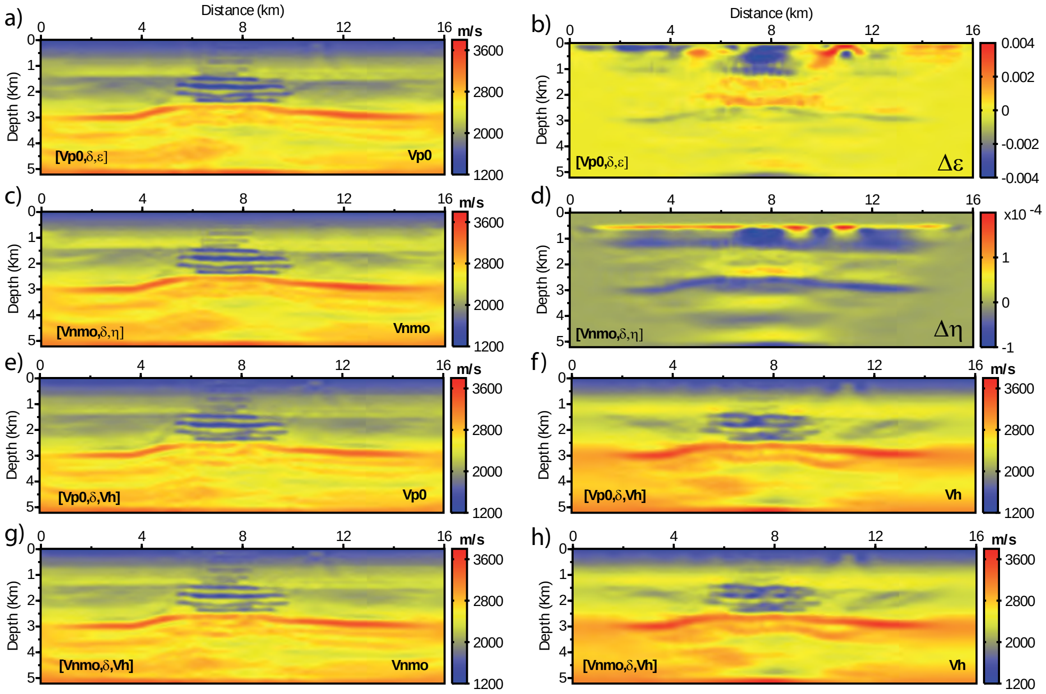


Figure 4.8: Multi-parameter anisotropic FWI. (a,b) Joint update of the vertical velocity (a) and the Thomsen parameter ϵ (b) for the (V_{P_0}, δ, ϵ) parameterization. Note that we show the difference between the final and initial ϵ models rather than the final ϵ model, because the final ϵ model and the initial ϵ model are quite close. (c,d) Joint update of the NMO velocity (c) and η (d) for the (V_{NMO}, δ, η) parameterization. As for ϵ, we show the difference between the final η model and the initial η model. (e,f) Joint update of the vertical velocity (e) and of the horizontal velocity (f) for the (V_{P_0}, δ, V_h) parameterization. (g-h) Joint update of the NMO velocity (g) and of the horizontal velocity (h) for the (V_{NMO}, δ, V_h) parameterization.

to perform mono-parameter FWI and multi-parameter FWI. We now consider the case where the initial models of δ and ϵ are homogeneous, with values $\delta = \epsilon = 0.01$, with the aim to increase the influence of the Thomsen parameters in the data residuals, and hence to create a more suitable configuration for retrieving them. We perform the joint update of the vertical velocity and parameter ϵ with the (V_{P_0}, ϵ, δ) parameterization (Figure 4.11). The ϵ model remains unchanged, while the vertical velocities are strongly underestimated in the gas layer (Figure 4.12). This highlights, on the one hand, the trade-off between the vertical velocity and parameter ϵ, and on the other hand, the dominant influence of the vertical velocity on the data. Indeed, a more suitable tuning of the regularization, the aim of which is to preserve the scaling effect of the Hessian on the gradient of the misfit function, might allow a better scaling of the model perturbations of V_{P0} relative to those of parameter ϵ. This is a difficult issue, which requires extensive trial-and-error and remains beyond the scope of the present study. Of note, both the vertical velocity and parameter ϵ are underestimated in the final FWI models, in particular in the gas layers (Figure 4.12), which leads to a severe underestimation of the horizontal velocities. Therefore, errors in the vertical velocity and parameter ϵ do not appear to compensate each other, but appear to accumulate their effects. This strongly indicates that the inversion is hampered by cycle skipping artifacts and remains locked into a local minimum

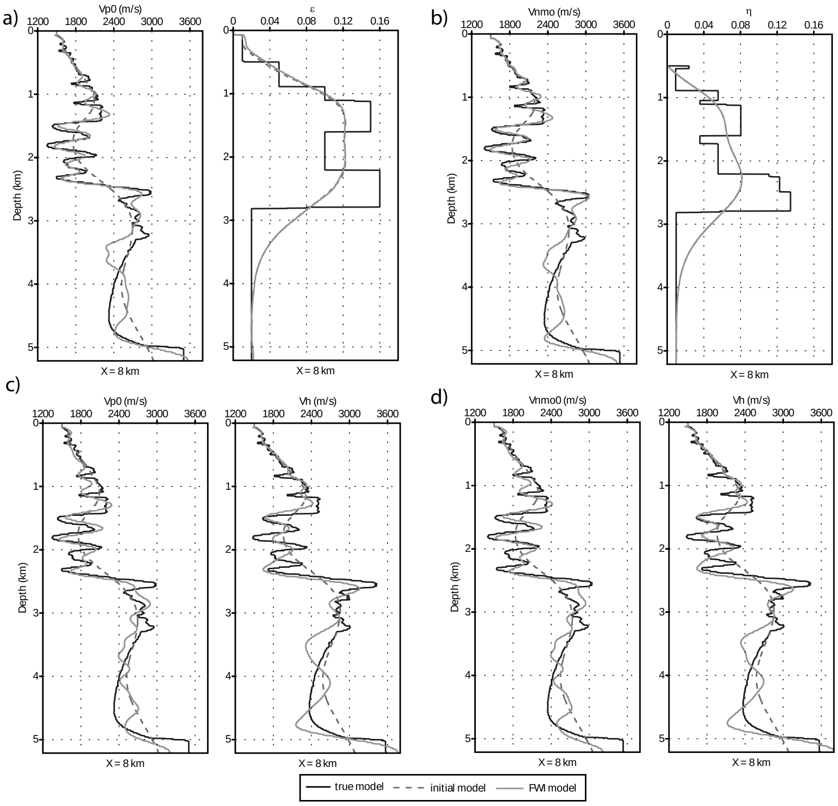


Figure 4.9: Vertical logs of multi-parameter FWI models (solid gray line) shown in Figure 4.8. The true model and the initial model are plotted with solid black and dashed gray lines, respectively. The log is taken at 8 km in distance in the middle of the model. (a) Vertical velocity logs (left) and ϵ logs (right) extracted from the models of Figure 4.8a,b. (b) Vertical velocity logs (left) and η logs (right) extracted from the models of Figure 4.8c,d. (c) Vertical velocity log (left) and horizontal velocity log (right) extracted from the models of Figure 4.8e,f. (d) NMO velocity log (left) and horizontal velocity log (right) extracted from the models of Figure 4.8g,h.

because of the insufficient accuracy of the Thomsen parameter background models.

4.2.4.6 Hierarchical mono-parameter anisotropic full waveform inversion

Other FWI strategies consist of updating the multiple parameter classes hierarchically, rather than simultaneously: the parameters with a dominant influence on the data are updated first, before updating the secondary parameters (Tarantola, 1986; Sears et al., 2008; Brossier et al., 2010c). This hierarchical approach aims to increase the influence of the secondary parameters on the data residuals during the late inversion stages, once the influence of the dominant parameter has been decreased during the early inversion steps. Indeed, the risk is that the influence on the data of the secondary parameters is interpreted as the footprint of the dominant parameter during the early stages of the inversion, if some trade-off exists between the dominant parameters and the secondary parameters.

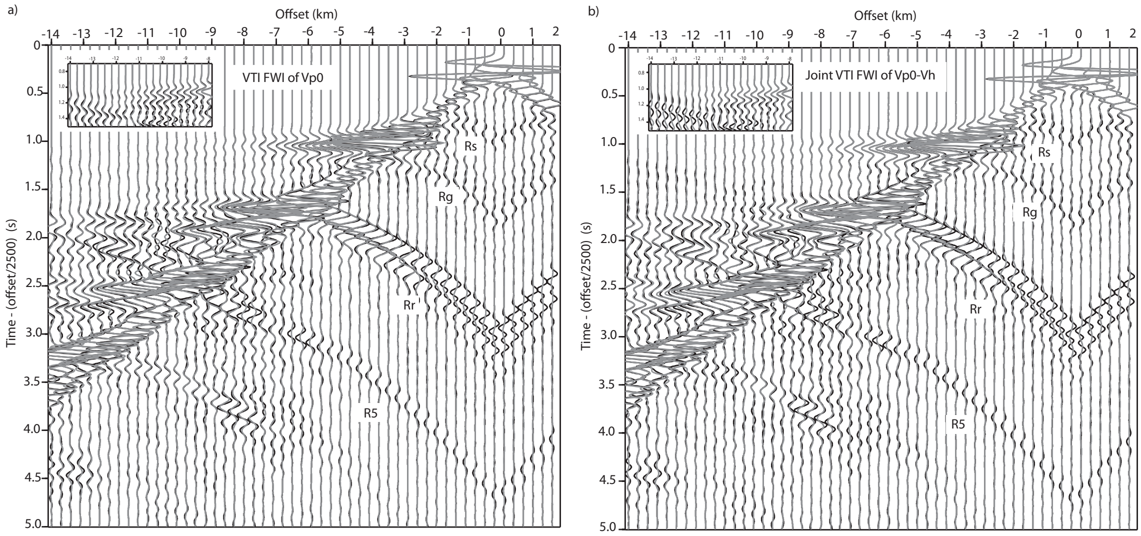


Figure 4.10: Synthetic seismograms computed in the multi-parameter FWI models. (a) Seismograms computed in the FWI models of Figure 4.8a,b ((V_{P_0}, δ, ϵ) parameterization). (b) Seismograms computed in the FWI models of Figure 4.8c,d ((V_{P_0}, δ, V_h) parameterization).

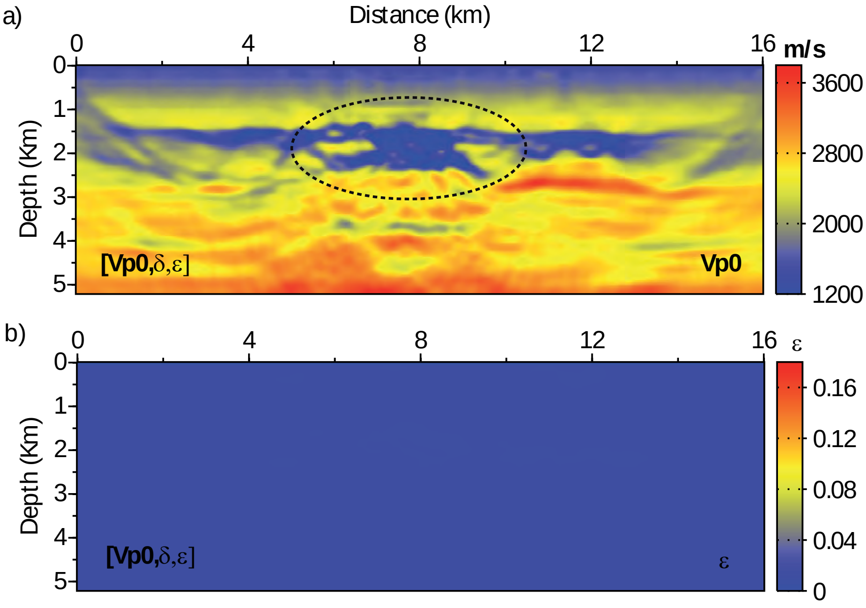


Figure 4.11: Multi-parameter VTI FWI for vertical velocity (a) and ϵ (b) when the initial background models of ϵ and δ are homogeneous ($\epsilon = \delta = 0.01$).

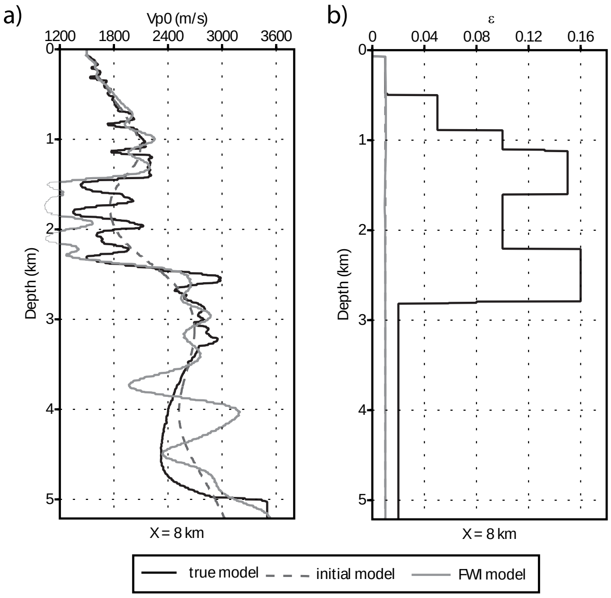


Figure 4.12: Logs of the V_{P0} FWI model (a) and the ϵ FWI model (b) (solid gray lines) shown in Figure 4.11. The true and the initial models are plotted with solid black and dashed gray lines, respectively. The log is taken at 8 km in distance in the middle of the model. Note the strong underestimation of the vertical velocity in the gas layers.

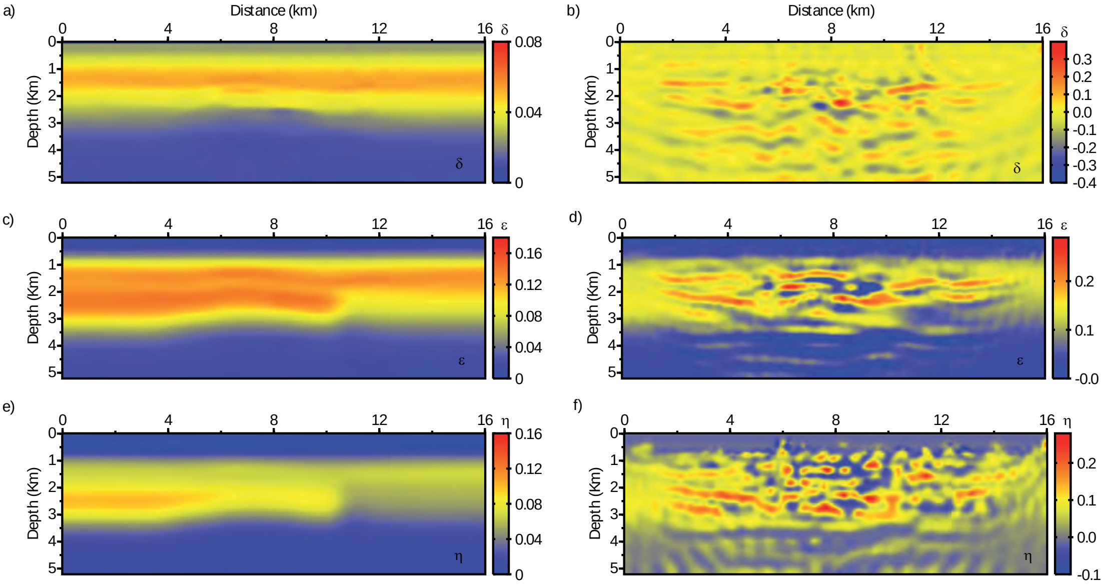


Figure 4.13: Hierarchical VTI FWI for the update of the Thomsen parameters. The background V_{P_0} and V_{NMO} models built by mono-parameter FWI are shown in Figure 4.6a,c. (a,b) Update of δ with the (V_{P_0}, δ, ϵ) parameterization at 2 Hz (a) and 4 Hz (b). The smooth ϵ model is kept fixed during FWI (Figure 4.1i). (c,d) Update of ϵ with the (V_{P_0}, δ, ϵ) parameterization at 2 Hz (a) and 4 Hz (b). The smooth δ model is kept fixed during FWI (Figure 4.1h). (e,f) Update of η with the (V_{NMO}, δ, η) parameterization at 2 Hz (e) and 4 Hz (f). The smooth δ model is kept fixed during FWI (Figure 4.1h).

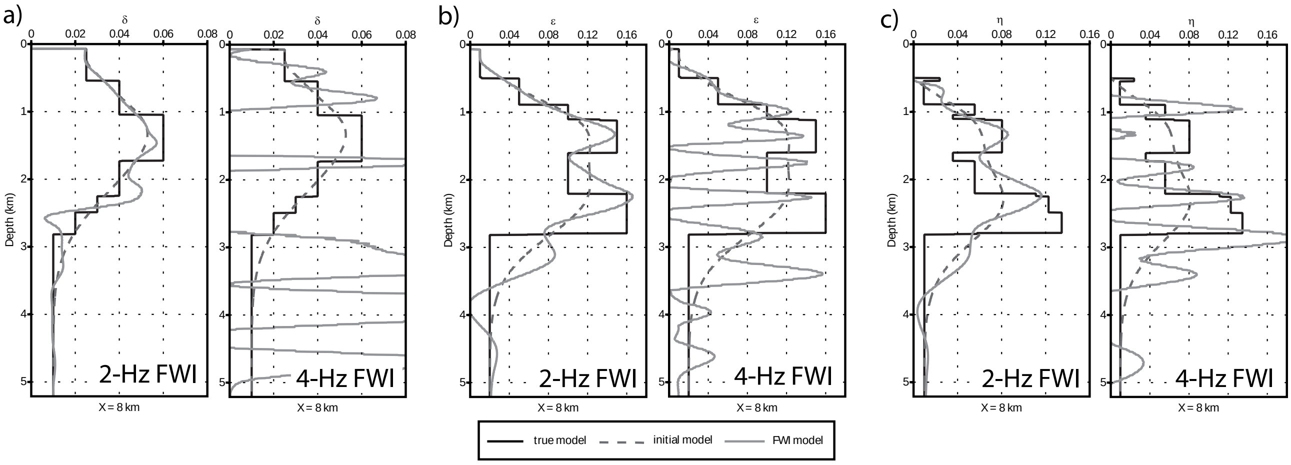


Figure 4.14: Vertical logs of the δ, ϵ, η FWI models (solid gray line) of Figure 4.13. The true and initial models are plotted with solid black and dashed gray lines, respectively. The log is taken at 8 km in distance in the middle of the model. (a) Update of δ at 2 Hz (left) and 4 Hz (right) (Figure 4.13(a-b)). (b) Update of ϵ at 2 Hz (left) and 4 Hz (right) (Figure 4.13(c-d)). (c) Update of η at 2 Hz (left) and 4 Hz (right) (Figure 4.13(e-f)). The update becomes unstable at 4 Hz.

We apply this hierarchical inversion for the update of Thomsen parameters δ, ϵ and η using the velocity models developed by mono-parameter FWI as background models (Figure 4.6a,c). The (V_{P_0}, δ, ϵ) parameterization is used for the update of δ and ϵ, while the (V_{NMO}, δ, η) parameterization is used for the update of η. The updates are performed independently for each Thomsen parameter, keeping the other parameters of the parameterization fixed. The initial models of the Thomsen parameters are the smooth models shown in Figure 4.1g-l. For each mono-parameter inversion, the damping of the Hessian is scaled to the highest diagonal term of the mono-parameter Hessian associated with the updated Thomsen parameter. The updated models are shown after the inversion of the initial frequency (2 Hz) and the third frequency (4 Hz) (Figures 4.13, 4.14). These results show that the Thomsen parameters can be only reliably updated at low frequencies (2 Hz), while at higher frequencies, the inversion becomes unstable. This confirms that the seismic data are mainly sensitive to the large wavelengths of the Thomsen parameters in the (V_{P_0}, δ, ϵ) and (V_{NMO}, δ, η) parameterizations.

4.2.5 Application to real Valhall data

We test the conclusions inferred from the synthetic experiments against the application of two-dimensional visco-acoustic VTI FWI to the hydrophone component of an OBC dataset recorded in the Valhall oil and gas field. The Valhall field is an over-pressured, under-saturated, Upper Cretaceous chalk reservoir located in the North Sea, approximately 290 km off-shore of southern Norway, with a water layer of 69 m in depth. The field is located in the most southwestern corner of the Norwegian continental shelf (Barkved and Heavey, 2003). This shallow-water field is characterized by a gas cloud above the reservoir, which hampers the imaging of reflectors at the oil reservoir level to 2.5 km in depth (Sirgue et al., 2010). The significant intrinsic anisotropy with a vertical velocity that is 15% slower than the horizontal velocity in some areas is another characteristic of this field (Kommedal et al., 2004). In the present study, we consider the cable 21 of the three-dimensional, four-component datasets, which has been processed previously by three-dimensional, isotropic acoustic FWI (Sirgue et al., 2010) and by two-dimensional, isotropic and anisotropic FWI (Prieux et al., 2011). The footprint of anisotropy on the isotropic FWI was shown for this dataset by Prieux et al. (2011). This line is located next to the gas cloud, which is highlighted by the horizontal slice of the three-dimensional FWI model of Sirgue et al. (2010) at 1 km in depth (Prieux et al., 2011, their figure 1). The shot and receiver spacings were 50 m. The maximum offset in the acquisition was 13 km. The two-dimensional section along the position of cable 21 through the anisotropic three-dimensional model of the Valhall field is shown in Figure 4.15. The anisotropic models for V_{P0}, δ and ϵ were built by reflection travel-time tomography (courtesy of BP), and they are used as the initial model for the VTI FWI in this study. The horizontal and NMO velocities are inferred from V_{P_0}, δ and ϵ using the relationships: $V_{NMO} = V_{P_0}\sqrt{1+2\delta}$ and $V_h = V_{P_0}\sqrt{1+2\epsilon}$. A receiver gather for the hydrophone component at position $x = 14.1$ km is shown in Figure 4.16a. The first arrivals (Figure 4.16a, D1, D2), the reflection from the top of the gas layer (Figure 4.16a, Rg), the reflection from the base of the gas layers, and the reflection from the top of the reservoir (Figure 4.16a, Rr) are highlighted in the seismograms. The reflections from the top of the reservoir are disrupted at critical and super-critical distances by shingling dispersive guided waves propagated in the near surface (Figure 4.16a, SW) (Robertson et al., 1996). We compute the first-arrival traveltimes and the reflection traveltimes from the top and the bottom of the gas layers in the NMO and horizontal velocity models using the isotropic eikonal solver of Podvin and Lecomte (1991) (Figure 4.16b,c), to check the kinematic accuracy of these velocities against the direction of propagation. The first-arrival rays do not sample the structure at the reservoir level, as they only travel down to 1.5 km in depth at their farthest offset. Superimposition of the computed traveltime curves on the receiver gather shows that the NMO velocities do not allow the matching of the traveltimes of diving waves at long offsets and the long-spread reflection Rg to be matched (Figure 4.16b). The traveltime delay exceeds half of the period of a starting frequency of 3.5 Hz, which prevents the criterion required to avoid cycle skipping to be satisfied. In contrast, the NMO velocity model is expected to match the short-spread reflection traveltimes of phases Rg and Rr to be matched, which is supported by Figure 4.16b. The horizontal velocity model allows for much better agreement of the first-arrival

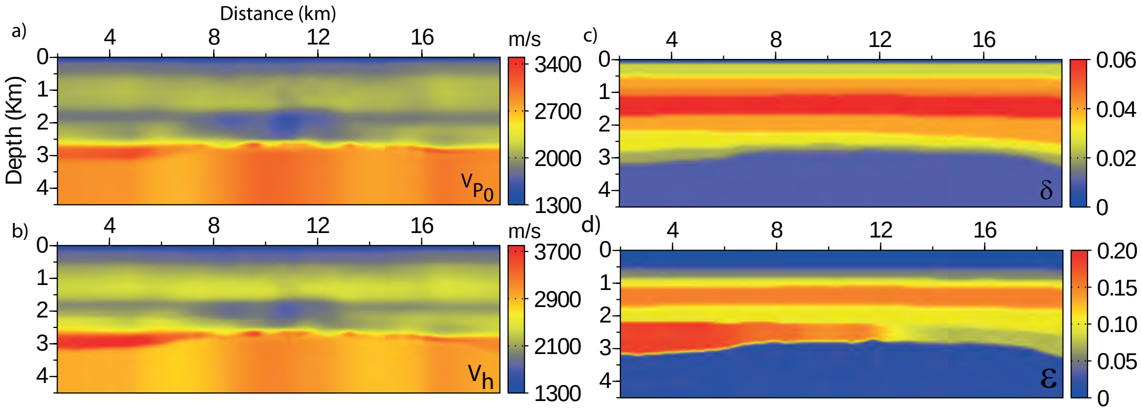


Figure 4.15: Valhall case study: initial FWI models. (a) Vertical velocity (V_{P_0}), (b) horizontal velocity (V_h), (c) Thomsen parameter δ, (d) Thomsen parameter ϵ. The V_{P_0}, δ and ϵ models were built by reflection traveltime tomography (courtesy of BP).

traveltimes (Figure 4.16c). However, the horizontal velocity model less accurately matches the reflection traveltimes at intermediate offsets, which highlights the footprint of the anisotropy in this data set. We apply VTI FWI to this OBC dataset. Following the FWI set-up of Prieux et al. (2011), we successively invert five frequency group: [3.5, 3.78, 4], [4, 4.3, 4.76], [4.76, 5, 5.25], [5.25, 5.6, 6] and [6, 6.35, 6.7] Hz with 25 iterations per frequency groups. The density model is inferred from the initial V_{P_0} model using the Gardner's law (Gardner et al., 1974). A homogeneous attenuation model was defined by trial-and-error, such that the root-mean-square amplitudes of the early-arriving phases computed in the initial model roughly matche those of the recorded data (Prieux et al., 2011). Prieux et al. (2011) reported a value of 150 for the best fitting attenuation factor Q_P. Both the density and the attenuation models are kept fixed during the FWI iterations. The prior model is set as the starting model of each frequency group inversion, equation 4.1, and is kept the same over the iterations. Tikhonov regularization is performed with a Gaussian function of adaptive correlation lengths of 20 % of the local wavelengths, Equation 5.5.

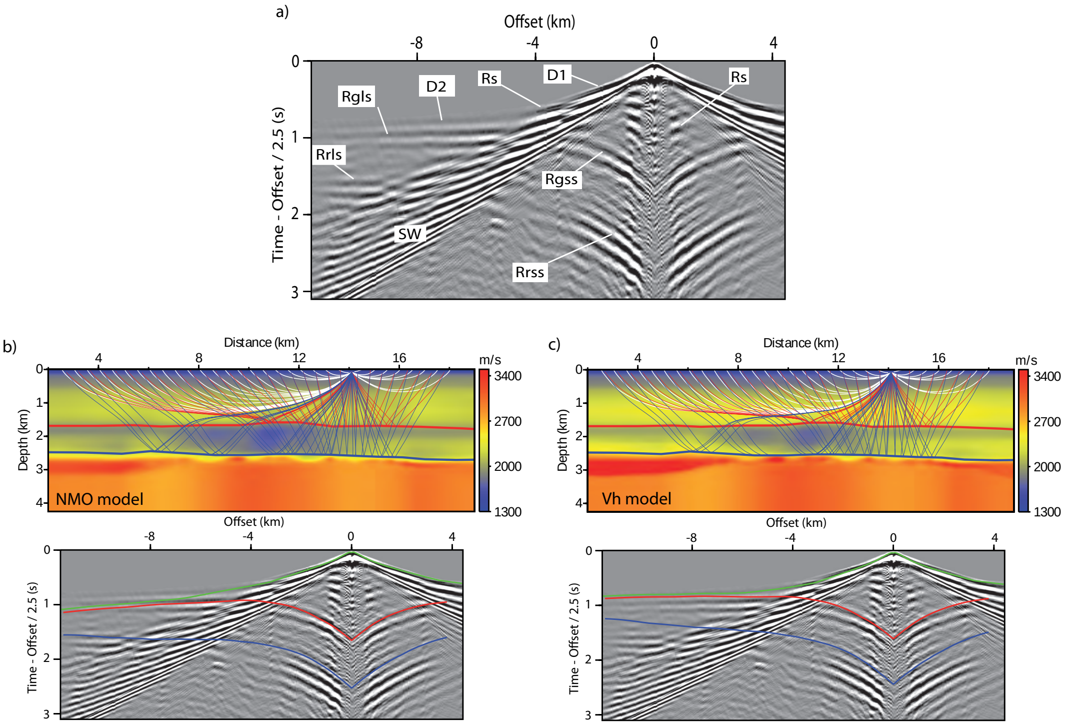


Figure 4.16: Valhall case study: data anatomy. (a) Example of preprocessed recorded receiver gather, at position $x = 14100\ m$. The vertical axis is plotted with a reduction velocity of 2.5 km/s. Phase nomenclature: D1, D2: diving waves. Rs: shallow reflection. Rgss / Rgls: short-spread and long-spread reflections from the top of the gas. Rrss / Rrls: short-spread and long-spread reflections from the top of the reservoir. SW: shingling waves. (b) Top: ray tracing in the NMO velocity model for the first arrival (white rays), and the reflections from the top of the gas (red) and the reservoir (blue). The top of the gas and the reservoirs are delineated by red and blue solid lines, respectively. Bottom: receiver gather shown in (a) with superimposed traveltime curves computed in the NMO model for these three phases. (c) As for (b), for the horizontal velocity model (from Prieux et al. (2011)).

4.2.5.1 Mono-parameter vertical transverse isotropic FWI

We first perform mono-parameter VTI FWI to update the vertical velocity, horizontal velocity, and NMO velocity, following the same approach as for the synthetic test. The initial models for FWI are shown in Figure 4.15. The three FWI velocity models (V_{P_0}, V_h, and V_{NMO}) are of comparable quality and resolution (Figure 4.17). This is consistent with their similar radiation patterns when type 1 parameterization is used (Gholami et al., 2012b, their figure 3(b-d)). Confrontation of the FWI vertical velocity model against the well log shows good agreement between the measured and reconstructed vertical velocity (Figure 4.18).

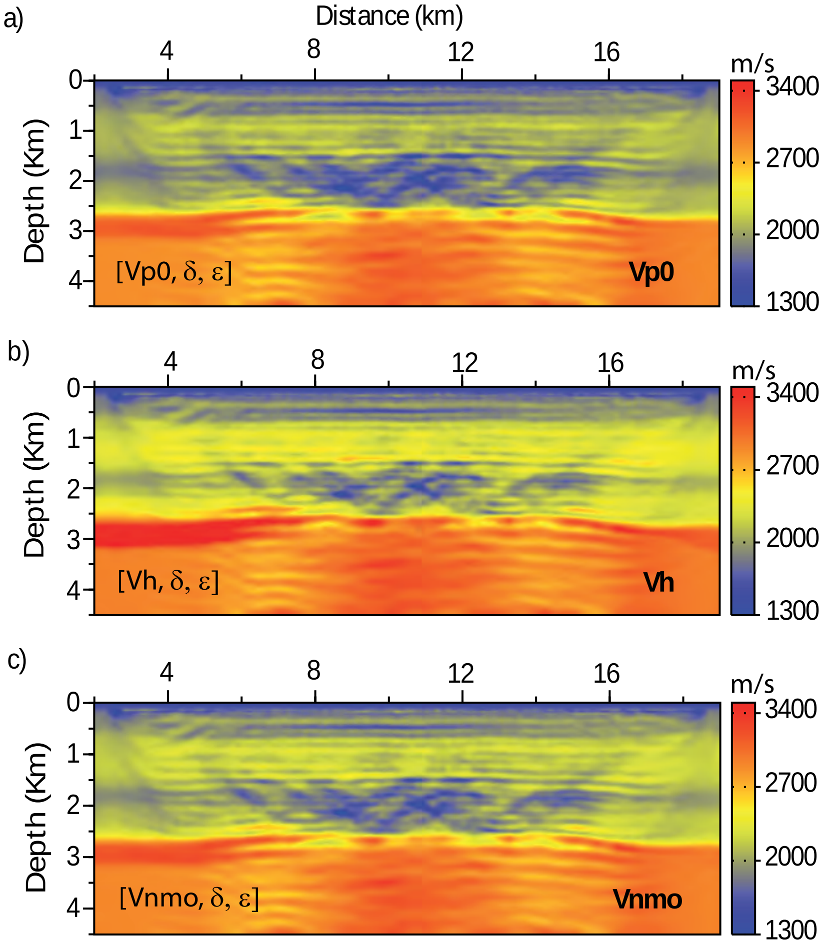


Figure 4.17: Valhall case study: final models of mono-parameter FWI. (a) Vertical velocity model. (b) Horizontal velocity model. (c) NMO velocity model. The model space is parametrized by one wavespeed (a) V_{P_0}, (b) V_h, (c) V_{NMO}, and Thomsen's parameters δ and ϵ, these latter are kept fixed during inversion.

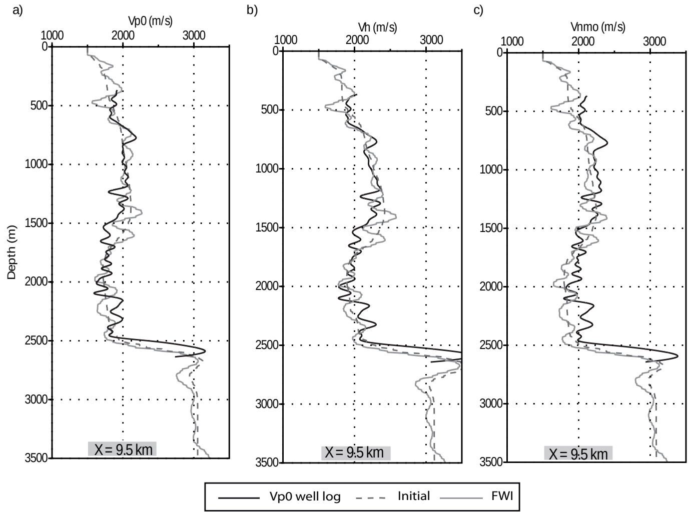


Figure 4.18: Valhall case study: Comparison between well log (black line) and FWI velocities (solid gray line). Log of the initial FWI models is plotted with dashed gray line. The sonic log has been smoothed according to the theoretical resolution of FWI (Prieux et al., 2011)). (a) Vertical velocity. (b) Horizontal velocity. (c) NMO velocity. In (b) and (c), the well log for horizontal and NMO velocity were built from the original well log for vertical velocity (a) and the background δ and ϵ models.

4.2.5.2 Multi-parameter vertical transverse isotropic FWI

We now perform the joint update of two wavespeeds by multi-parameter VTI FWI. Two parameterizations are tested: (V_{P_0},δ,V_h) and (V_{NMO},δ,V_h). The smooth background model of δ is kept fixed during the inversion, when the same value of the damping factor λ_i is also used for the two wavespeeds, equation 4.1, as they have radiation pattern of similar amplitudes (Gholami et al., 2012b, their figure 3(e-f)). The final FWI models for the two parameterizations are shown in Figure 4.19. The vertical velocity model and the NMO velocity model inferred from the two parameterizations are of comparable quality and resolution (Figure 4.19a,c), which is consistent because their radiation patterns are similar. These velocity models are also similar to those inferred from mono-parameter FWI (Figure 4.17). Nevertheless, we note some vertical oscillations in the upper part of the vertical and NMO velocity models inferred from multi-parameter FWI, and the footprint of these oscillations appears less significant in the velocity models inferred from mono-parameter FWI (compare Figure 4.17a,c and Figure 4.19a,c). These vertical oscillations might result from the limited vertical-wavenumber bandwidth of the velocity models. If this interpretation is correct, the higher footprint of the oscillations in multi-parameter FWI models is consistent: indeed, the velocity models inferred from type 1 parameterization should have a broader wavenumber content in the upper structure (i.e., where the aperture illumination is broad) than the velocity models inferred from type 2 parameterization, because their radiation patterns span over the full range of scattering angles when type 1 parameterization is used (Gholami et al., 2012b, their figure 3(b,e)).

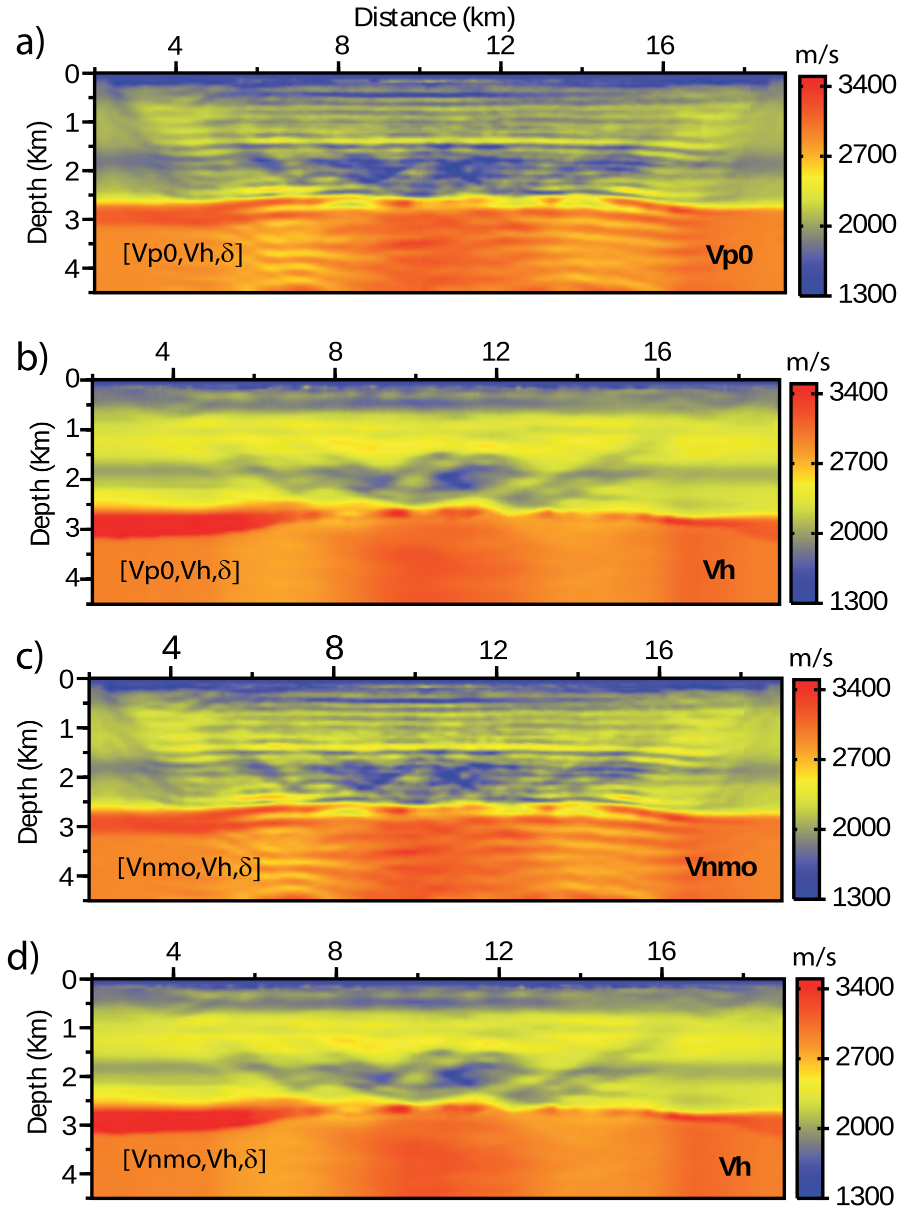


Figure 4.19: Valhall case study: Final models of multi-parameter FWI. (a,b) Joint update of the vertical velocity (a) and the horizontal velocity (b) with the (V_{P_0}, δ, V_h) parameterization. (c,d) Joint update of the NMO velocity (c) and of the horizontal velocity (d) with the (V_{NMO}, δ, V_h) parameterization. The smooth δ background model is kept fixed during FWI in both cases (Figure 4.15c. Note that the vertical oscillations in the upper part of the vertical and NMO velocity models is more significant than in the corresponding velocity models built by mono-parameter FWI with the type 1 parameterization (Figure 4.17a,c). See text for additional discussion.

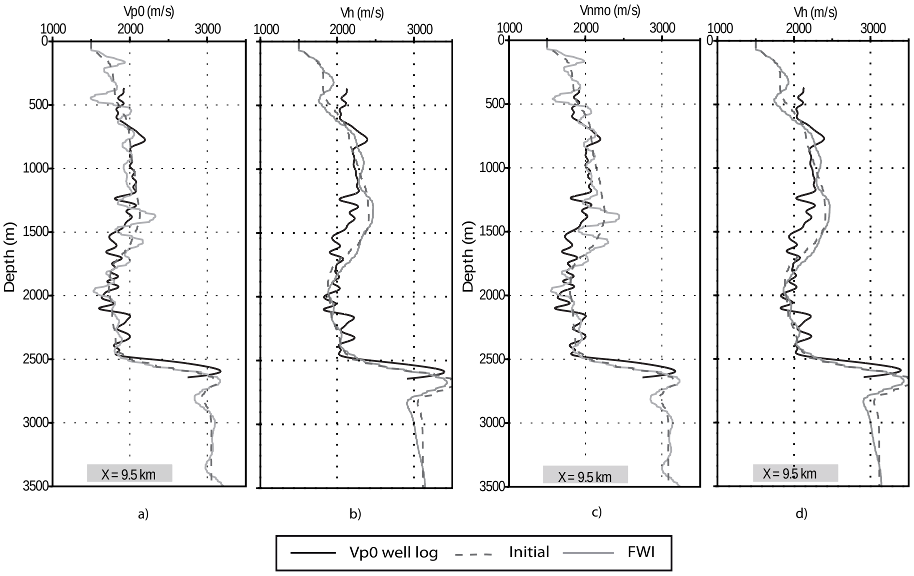


Figure 4.20: Valhall case study: comparison between the well log (black line) and FWI velocities from Figure 4.19 (solid gray line). (a,b) Vertical velocity (a) and horizontal velocity (b) (Figure 4.19(a-b)). (c,d) NMO velocity (c) and horizontal velocity (d) (Figure 4.19c,d). The well log for horizontal velocity and NMO velocity were built with the same procedure as in Figure 4.18. The logs of the initial models are plotted with dashed gray line.

The horizontal velocity models obtained with the (V_{P_0},δ,V_h) and (V_{NMO},δ,V_h) parameterizations are very similar (Figures 4.19b,d, 4.20b,d). They show a more limited resolution than that obtained by mono-parameter FWI (Figure 4.17b). This is a consistent result because the horizontal velocity has an influence on the wide aperture components of the data only with type 2 parameterization (Gholami et al., 2012b, their figure 3e,f). The spatial resolution of the horizontal velocity models (Figure 4.19b,d) is, however, improved relative to that of the initial V_h model (Figure 4.15b). Time-domain seismograms computed in the final FWI models computed with the (V_{P_0},δ,ϵ) and (V_{P_0},δ,V_h) parameterizations show good and similar agreement with the recorded data for both diving waves and reflected phases (Figure 4.21). We compute anisotropic reverse time migration and common image gathers (CIGs) in the initial models and in the final FWI models obtained with the (V_{P_0},δ,ϵ) and (V_{P_0},δ,V_h) parameterizations using the same approach as Prieux et al. (2011). The migration computed in the initial anisotropic model provides a highly accurate image of the subsurface at all depths (Figure 4.22a), which is supported by fairly flat reflectors in the CIGs (Figure 4.22d). As for Prieux et al. (2011), the background migration models built by FWI do not allow any improvement of the migrated image inferred from the reflection-traveltime tomography background model, except in the first kilometer in depth of the subsurface, where the CIGs computed in the FWI models show flatter events (Figure 4.22d-f). This is expected because the reflection traveltime tomography is optimally designed to focus reflection energy in depth, unlike FWI, and should perform well in structural environments like Valhall, where no significant dips are shown. However, it is worth noting that the reflectors are positioned at similar depths in each migrated section of Figure 4.22, while severe mispositioning of the reflectors was shown in the migrated images computed in isotropic FWI background models (Prieux et al., 2011, their figures 8 and 9).

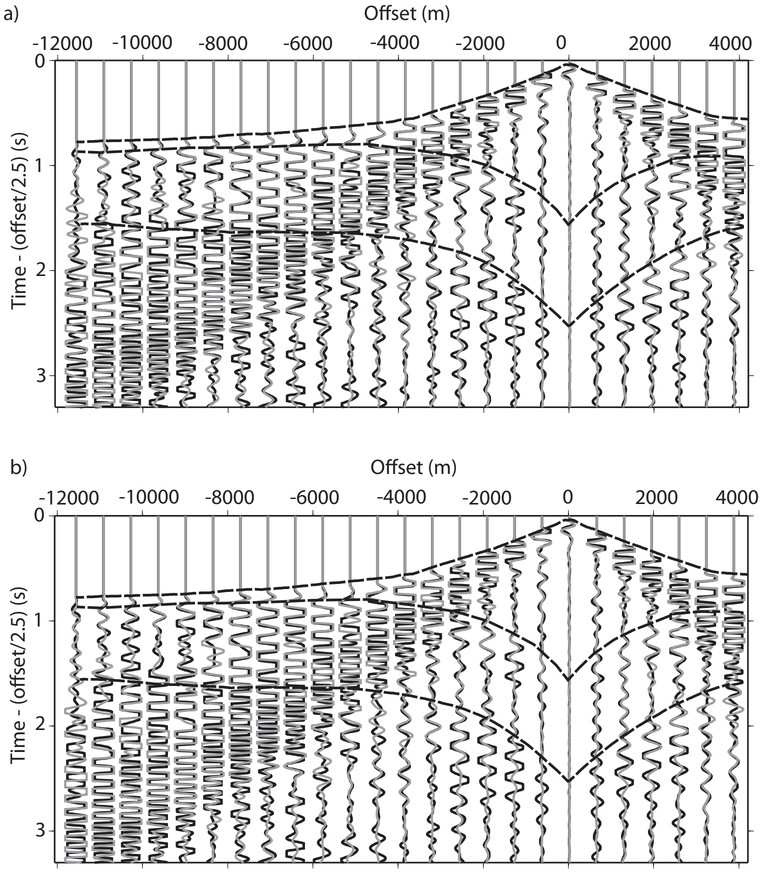


Figure 4.21: Valhall case study: direct comparison between recorded seismograms (black line) and seismograms computed in the final anisotropic FWI models (gray line). (a) Seismograms are computed in the vertical velocity model updated by mono-parameter FWI (Figure 4.17a) and in the smooth background δ and ϵ models (Figure 4.15c,d). (b) Seismograms are computed in the V_{P_0} and V_h models updated by multiparameter FWI (Figure 4.19a,b) and the smooth background δ model (Figure 4.15c). The dashed black curves are picked first-arrival traveltimes and picked reflection traveltimes from the top and from the bottom of the gas layers.

4.2.6 Conclusion

The FWI case studies presented in the present study support the conclusions of the theoretical parameterization analysis of acoustic VTI FWI presented in the companion report. As long as accurate long-wavelength models of the Thomsen parameters δ and ϵ are developed during the preliminary tomographic step, a high-resolution velocity model can be reconstructed by mono-parameter FWI, when the subsurface parameterization combines the wavespeed with the two Thomsen parameters. There were no significant differences in the reconstruction quality when the updated wavespeed is the vertical velocity, horizontal velocity, or NMO velocity. The joint update of the vertical velocity and the parameter ϵ (or, of the NMO velocity and the parameter η) has been shown to be feasible, although this led to a marginal update of the Thomsen parameter, as most of the influence of parameters ϵ or η on the data is predicted by the initial models. Alternatively, the joint update of two wavespeeds (the vertical velocity and horizontal velocity, or the NMO velocity and horizontal velocity) shows robust results, because

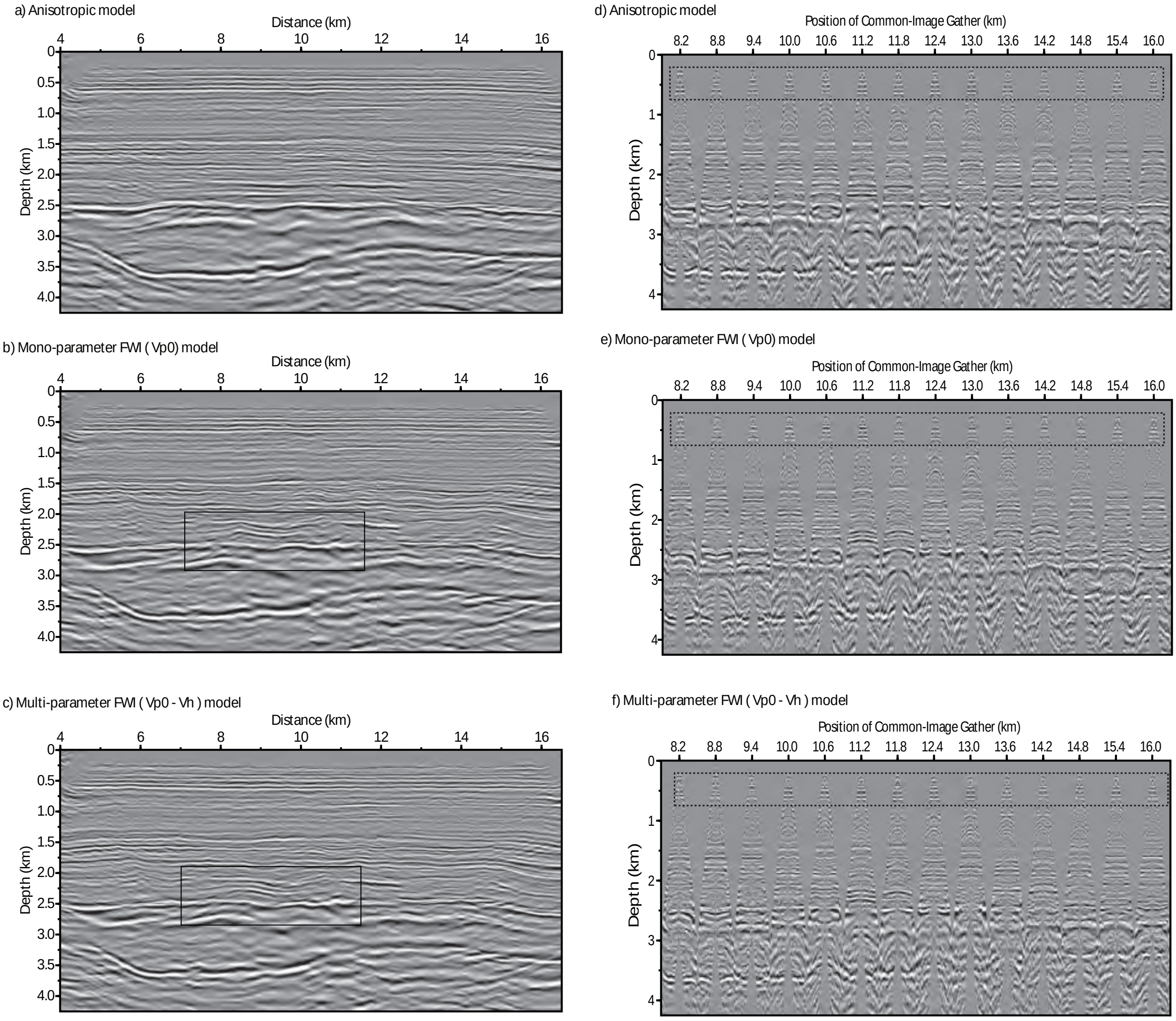


Figure 4.22: Valhall case study: VTI reverse time migration. (a-c) Migrated images computed in the initial models (a) (Figure 4.15) and in the anisotropic models updated by mono-parameter FWI (b) (Figure 4.17a) and by multi-parameter FWI (c) (Figure 4.19(a-b). The frame delineates the area in the migrated images computed in the FWI models, where reflectors show suspicious undulations, similar to those shown in Prieux et al. (2011), their figure 8(e-h). (d-f) CIGs computed in the offset-depth domain. The frame delineates where CIGs inferred from the FWI models show flatter reflectors.

the two wavespeeds have significant influences on the data that are of similar magnitude, except for their distinct ranges of scattering angles. Hence, trade-off artifacts should not significantly impact on the inversion. The horizontal velocity model has a limited resolution because it has influence on the large scattering angles only, when the horizontal velocity is combined with the vertical velocity or the NMO velocity in the parameterization. The vertical velocity model or the NMO velocity model should have a narrower wavenumber content when they are combined with the horizontal velocity rather than with parameter ϵ, because they have influence on the data for small and intermediate scattering angles in the former case, while they have influence on the data over the full range of scattering angles in the latter case. However, for surface acquisition and given the accuracy of the available initial velocity models, we do not show significant differences in resolution between the vertical velocity models and the NMO velocity models built with the two parameterizations. In all of the tests, the background model of Thomsen parameter δ is kept fixed due to its limited influence in the data. We would conclude that the choice of a suitable parameterization for acoustic VTI FWI should be driven by the accuracy of the initial models as well as any prior information which can be used during FWI. If sufficiently accurate large-scale models of the Thomsen parameters are available, to build a high-resolution velocity model of the subsurface, we would tend to favor a parameterization that involves only one wavespeed.

Multiparameter anisotropic FWI can involve parameters that have contrasting influences on the data. In this case, a key issue is to account for the Hessian in the optimization to correctly scale the gradients of the misfit function associated with each parameter class. The L-BFGS quasi-Newton method provides a computationally efficient framework, to take into account the effects of the Hessian in the FWI. The regularization damping term, conventionally added on the diagonal of the Hessian, can have a large influence on the reconstruction of the parameters. In the present study, we scale this damping to the maximum coefficient of the Hessian. In this case, the reconstruction of the parameter with the dominant influence in the data is favored at the expense of the reconstruction of the secondary parameters. Future work will aim to more carefully design the regularization of the FWI for an assessment of the feasibility of the reconstruction of multiple classes of parameters with variable imprints in the data, keeping in mind that the reconstruction of parameters that have an influence below the noise level in the data is unlikely. This improved regularization might allow the joint updating of the vertical velocity and the large wavelengths of parameter ϵ (or the NMO velocity and parameter η), if the available background model of the Thomsen parameter is not accurate enough.

4.2.7 Acknowledgments

This study was funded by the SEISCOPE consortium http://seiscope.oca.eu, and sponsored by BP, CGG-VERITAS, ENI, EXXON-MOBIL, PETROBRAS, SAUDI ARAMCO, SHELL, STATOIL and TOTAL. The linear systems were solved with MUMPS package, available on http://graal.ens-lyon.fr/MUMPS/index.html. The mesh generation was performed with help of TRIANGLE, available on http://www.cs.cmu.edu/∼ quake/triangle.html. Access to the high-performance computing facilities of the SIGAMM (Observatoire de la Côte d'Azur) is acknowledged gratefully. This study was granted access to the HPC resources of [CINES/IDRIS] under the allocation 2010- [project gao2280] of GENCI (Grand Equipement National de Calcul Intensif), and we gratefully acknowledge both of these facilities and the support of their staff. We also thank BP Norge AS and their Valhall partner Hess Norge AS for allowing access to the Valhall dataset, as well as the dataset preprocessed by PGS for migration, and the well-log velocities.

Chapter 5

Elastic anisotropic FWI

Contents

5.1 Introduction

The multi-parameter (joint) FWI suffers from the non-linearity of the problem. This is more distinguishable for multi-parameter elastic FWI with respect to acoustic inversion. The presence of the shear wave components and their converted modes add more complexity to the problem. Additionally, the elastic FWI has some difficulties related to the computational cost. Many approaches are proposed in order to perform elastic FWI such as, the hierarchical construction of elastic model parameters based on their influence on the data (Tarantola, 1986), and implementation of multi-scale FWI algorithm by building a loop over frequency groups and time dampings (Brossier, 2011a). Tarantola (1986) showed that the parameterization of FWI is a highly hierarchical problem. This means that the inversion of physical parameters should be performed by hierarchical steps, based on their weight on the wavefield residual. Djikpéssé and Tarantola (1999) used a L_1 norm, time domain, 2.5-D elastic waveform inversion method to extract the variations of acoustic impedance and Poisson's ratio (2D map of these models) from marine multi-offset reflection seismograms collected in the Gulf of Mexico area. They showed that the estimation of physical P-impedance and Poisson's ratio models by full waveform fitting allows for lithology characterization and, therefore, the delineation of a shale-over-gas sand reservoir in Gulf of Mexico.

Another offshore application of time domain elastic FWI was shown by Sears et al. (2008) on synthetic examples. They exploited both the hydrophone and geophone components to reconstruct the P-wave and S-wave velocities by using wide-angle, multi-component OBC seismic data. Later, Sears et al. (2010) showed the application of the time domain elastic FWI on the real OBC data for Alba field in North sea. Brossier et al. (2009b) performed 2-D elastic frequency-domain FWI for imaging isotropic and complex onshore structures such as the SEG/EAGE overthrust model. They showed that the non-linearity of elastic FWI arises both from the presence of converted and surface waves, and from the limited accuracy of the V_S starting model. In order to ease this problem they proposed a data preconditioning by time damping plus the successive inversions of overlapping frequency groups for removing the near-surface instabilities of the FWI models. Recently, Chung and Shin (2011) proposed a new algorithm for elastic FWI, which uses the Gauss-Newton optimisation algorithm and the partial derivative wavefield is computed directly. But, due to the computational cost, their applications are restricted to small-sized case studies. Manukyan et al. (2012) explored the exploitable data-information in isotropic elastic-waveform inversion. They demonstrated that several source-receiver configurations provide sufficient information for imaging small-scale features with elastic cross-hole waveform tomography. They showed also, for realistic cross-hole case study, the satisfactory V_P and V_S tomograms can be derived, but the density structure remains unresolved. Unfortunately, there has been no investigation on sensitivity analysis (such as radiation pattern analysis) of the elastic anisotropic FWI.

In this chapter, we discuss the application of frequency-domain elastic FWI for VTI medium. The sensitivity analysis mostly includes the elastic VTI FWI on the so-called *inclusion* model. The three considered types of parameterizations for acoustic VTI FWI are valid for elastic case. The shear wave velocity V_{SV} is included in the parameterizations and is the common parameter class for all parameterization types. For more details about the parameterization types of 2D elastic medium refer to chapter 1.2.5 (table 1.2). The 2D elastic VTI FWI is applied to the Valhall field data set. The inversion is performed for vertical and horizontal geophone components of the 4C data set.

5.2 Synthetic examples of elastic FWI: *inclusion* model

In this section, we apply VTI elastic FWI to the inclusion model to perform a sensitivity and a trade-off analysis of the inversion. We seek to image an inclusion in a homogeneous background model. The acquisition geometry is perfect, in a way that it surrounds the target. The reason is to analyze the filtering effects performed by the parameterization in the full scattering angle. The initial FWI model is the homogeneous background model. Minimization of the misfit function is performed for parameters that are normalized by their value in the background model. This scaling allows us to manipulate different class of parameter which have the same range of values. This scaling also steers the inversion toward reconstruction of the parameters that have the dominant influence in the data (see chapter 3). We follow the similar configuration used for acoustic VTI FWI (chapter 3). Nine frequencies between 4.8 and 19.5 Hz are sequentially inverted. We use again the l-BFGS optimization with $\mathbf{m}_{prior} = \mathbf{m}^{(k)}$ in equation 1.76. The diagonal terms of Hessian matrix with $\lambda_i = 0.10\%$, $i = 1, N_c$, of the highest diagonal coefficient is used in the role of initial guess of the Hessian for the l-BFGS optimization (as for acoustic VTI FWI), where N_c is the number of parameters involved in inversion. In case of multi-parameter inversion, the same value of the damping λ_i is used for all parameters, whatever is the parameter class i. In other words, we have the following equation:

$$\Delta \mathbf{m}_i = \frac{\mathcal{G}_i}{\mathbf{H}_i + \lambda}, \qquad where \quad \lambda = 0.1\% MAX(H_i;\ \forall i). \tag{5.1}$$

$\mathcal{G}$ denotes the gradient and $\mathbf{H}$ denotes the Hessian. This implies that, if the diagonal coefficients of the Hessian associated with one parameter class are much smaller than the selected damping (here 0.10 % of the maximum coefficient), the inversion, unlikely, will succeed in scaling (by $\mathbf{H}_i + \lambda$) properly the model perturbation associated with this parameter class because the high damping term will annihilate the scaling action of the Hessian.

5.2.1 Mono-parameter and multi-parameter elastic anisotropic FWI

We apply the mono-parameter and multi-parameter elastic VTI FWI on inclusion model. The mono-parameter FWI means that only one parameter class is estimated by FWI, and multi-parameter FWI means that more than one parameters are estimated jointly by FWI. The categorized parameterization for acoustic VTI inversion is valid here, where three types of parameterizations are recognized. The parameterization by (V_{P0}, δ, ϵ, V_{SV}) is considered for parameterization type one. For parameterization type two the inversion with (V_{P0}, δ, V_h, V_{SV}) is evaluated. And the elastic moduli are the parameter classes for parameterization type three.

5.2.1.1 Parameterizations of type one: two wavespeeds + two Thomsen parameters

The joint reconstruction of V_{P0}, δ, ϵ, and V_{SV} is performed when the true model contains an inclusion perturbation for each model parameter class (figure 5.1). Extraction of vertical and horizontal profiles across the centre of the inclusion allows us to assess the resolution of the reconstructed model parameter. The amplitudes of the V_{P0} wavespeed perturbations are over-estimated (figure 5.1a-c), while the perturbations of Thomsen parameters are underestimated, dominantly for δ (figure 5.1d-i). The perturbations of the shear velocity V_{SV}, is reconstructed, except some oscillations around the inclusion, which may come from the trade-off with δ parameter. The Thomsen parameter ϵ is better reconstructed than its acoustic case, relatively. The over-estimation of V_{P0} and the under-estimation of Thomsen parameters δ and ϵ, plus the oscillations around the V_{SV} reconstructed model, suggest that there is a potential trade-off between these parameters. Moreover, these results show the dominant influence of V_{P0} and V_{SV} on the data compared to the Thomsen parameters. The diagonal coefficients of the Hessian (H_{diag}) computed in the initial models are shown in figure 5.2(a-d). We have $H^{(\delta)}_{max} = 0.02\% H^{(V_{P0})}_{max}$ and $H^{(\epsilon)}_{max} = 0.04\% H^{(V_{P0})}_{max}$ and $H^{(V_{SV})}_{max} = 25\% H^{(V_{P0})}_{max}$, where $H^{(V_{P0})}_{max}$, $H^{(\delta)}_{max}$, $H^{(\epsilon)}_{max}$ and $H^{(V_{SV})}_{max}$ denote the maximum diagonal coefficient of the Hessian for the V_{P0}, δ, ϵ and V_{SV} parameters, respectively. This implies that, the chosen damping term is higher by one order of magnitude than the maximum coefficient of the Hessian associated with Thomsen parameters. In this case, the action of the Hessian on the gradient, consisting in proper scaling the units of the model perturbation, is strongly reduced, leading to underestimated perturbations of δ and ϵ. On the other hand, this damping term does not affect the scaling of V_{SV} model perturbation.

Afterwards, we perform four mono-parameter FWI for V_{P0}, δ, ϵ, and V_{SV}, where the true

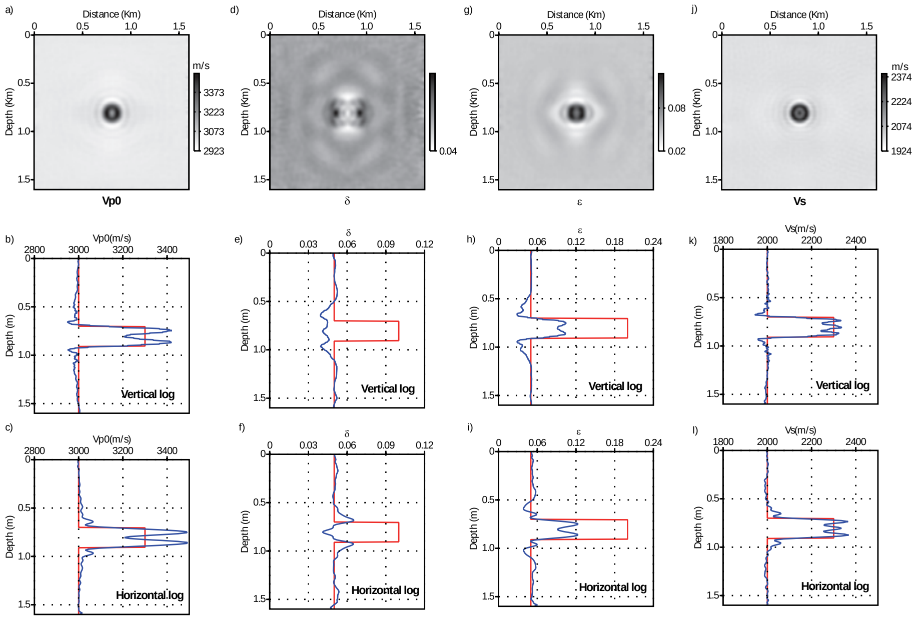


Figure 5.1: Inclusion-model FWI test. Joint update of V_{P0} (a-c), δ (d-f), ϵ (g-i), and V_{SV} (j-l) when (V_{P0},δ,ϵ,V_{SV}) parameterization is used. (a,d,g,j) Final FWI models. (b,e,h,k) Vertical profiles across the true inclusion (red) and the reconstructed one (blue). (c,f,i,l) Same as (b,e,h,k) for horizontal profiles.

models contain an inclusion associated with the parameter to be reconstructed and the true models of the other three parameters are kept fixed during inversion and are homogeneous. In this case, the damping term of the Hessian is scaled to the maximum coefficient of the mono-parameter Hessian. This means:

$$\Delta \mathbf{m}_i = \frac{\mathcal{G}_i}{\mathbf{H}_i + \lambda_i}, \qquad where \quad \lambda_i = 0.1\% MAX(H_i) \tag{5.2}$$

All the four parameter classes are perfectly reconstructed. This experiment verifies the suggested potential trade-off by their multi-parameter inversion. Because, they are not well reconstructed by multi-parameter inversion, even though the inversion configuration is same for both mono-parameter and multi-parameter experiments. Let us compare the mono-parameter elastic FWI with the one for acoustic FWI (chapter 3, figure 3.10). The mono-parameter acoustic VTI FWI for reconstruction of ϵ (figure 3.10g-i) parameter showed the filtering effects in the vertical and horizontal profiles that induced by its radiation pattern. On the other hand, its reconstruction by mono-parameter elastic FWI does not show this filtering effect and is well reconstructed. For the moment, we do not have enough information about radiation pattern of this parameter in elastic VTI medium, to reach to a solid conclusion of this result. The same scenario is for mono-parameter reconstruction of δ by elastic FWI comparing with its acoustic FWI.

Based on this experiment, the under-estimation of Thomsen parameters in multi-parameter inversion can be related (again) to their weak influence on the data and the potential trade-off between parameters, as V_{P0} and V_{SV} wavespeeds absorb the amplitudes of reconstructed Thomsen parameters.

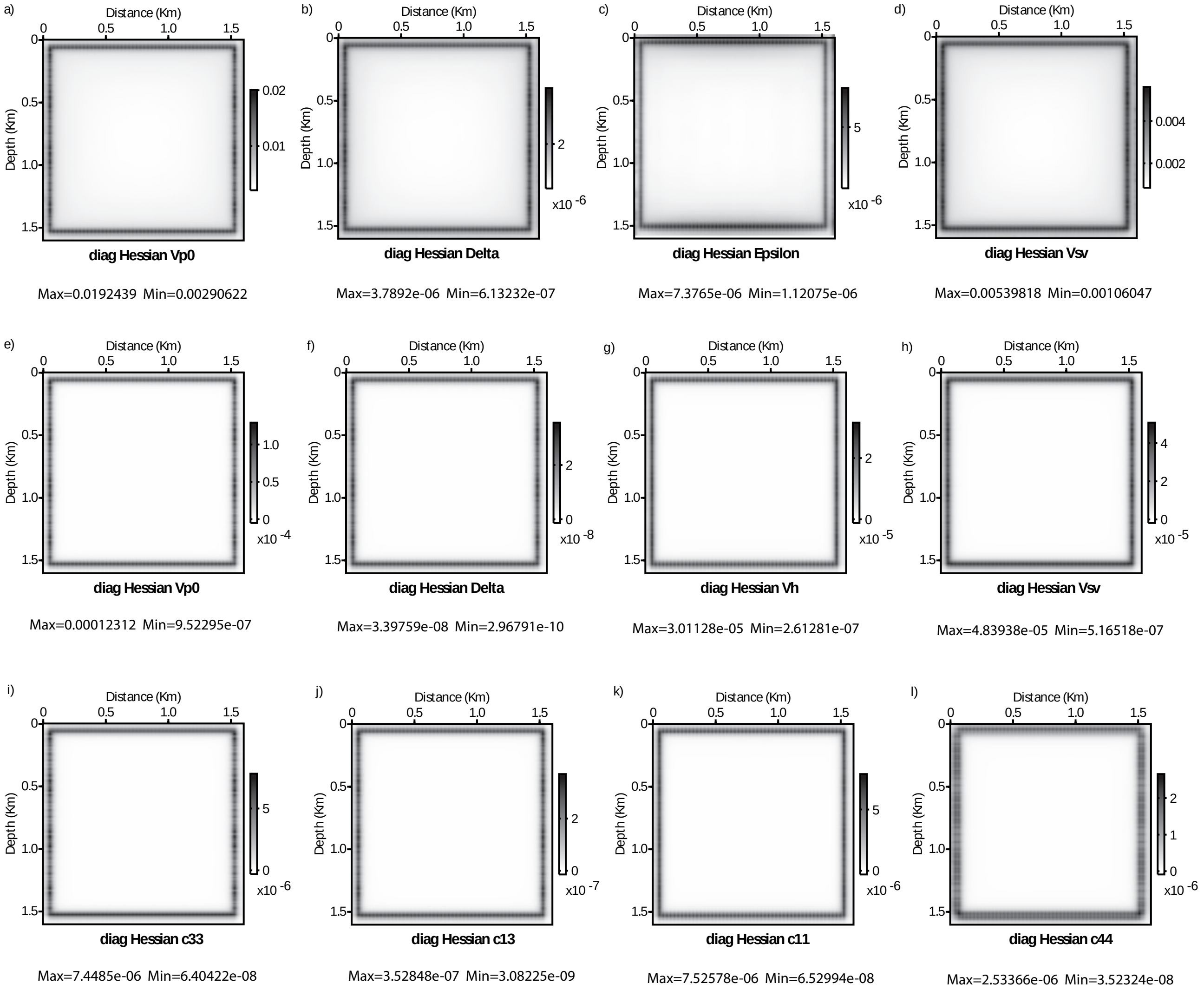


Figure 5.2: Diagonal terms of pseudo-Hessian associated with (a-d) parameterization V_{P_0}, δ, ϵ, V_{SV}, (e-h) parameterization V_{P_0}, δ, V_h, V_{SV}, and (i-l) parameterization c_{33}, c_{13}, c_{11}, c_{44} plotted in depth-distance domain. The footprint of the shot positions located around the target is clearly visible. The minimum and maximum values of the diagonal terms, provided by the fields Max and Min, represent an approximation of the square amplitude of the partial derivative wavefield and hence, give some insight on the relative influence of the three parameter classes on the data.

In order to analyze the trade-off each parameter class with other parameter classes, we perform four joint inversion experiment for V_{P0}, δ, ϵ, and V_{SV}, where the true models contain an inclusion associated with only one parameter at a time and all four parameters are reconstructed jointly. The purpose of this experiment is to find out, step-by-step, the potential trade-off between parameter classes. The experiment for V_{P0} (figure 5.4) illustrates the estimation of V_{P0} and some amplitude absorbed by δ and ϵ parameters. There is not any absorbed amplitude into the reconstructed V_{SV} model. This result suggest that, in multi-parameter elastic FWI,

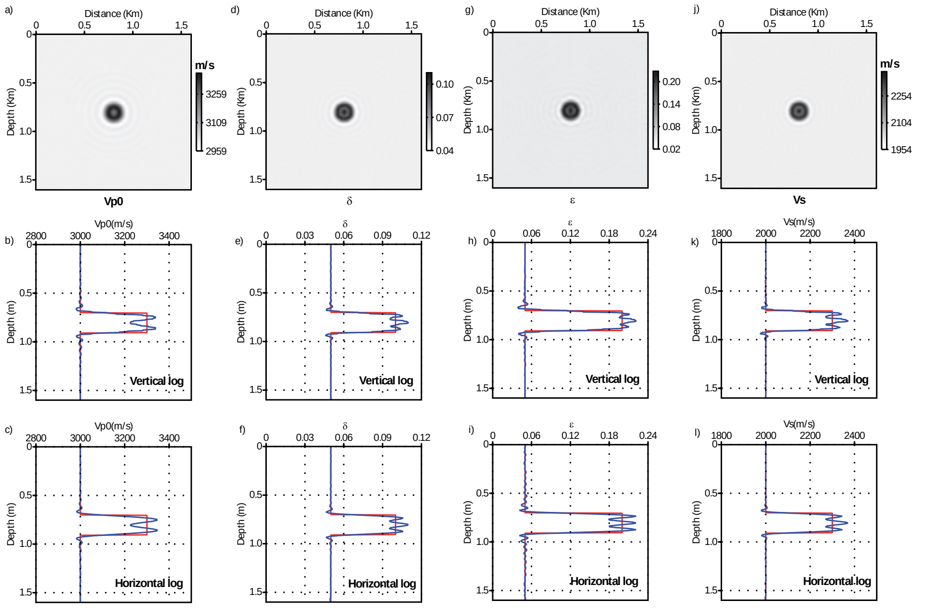


Figure 5.3: Inclusion-model FWI test. Results of mono-parameter FWI for V_{P0} (a-c), δ (d-f), ϵ (g-i), and V_{SV} (j-l) when (V_{P0},δ,ϵ,V_{SV}) parameterization is used. (a,d,g,j) Final FWI models. (b,e,h,k) Vertical profiles across the true inclusion (red) and the reconstructed one (blue). (c,f,i,l) Same as (b,e,h,k) for horizontal profiles.

V_{P0} shows the trade-off with δ and ϵ parameters, but not with V_{SV} parameters.
The experiment for V_{SV} parameter (figure 5.5) illustrates that, in multi-parameter elastic FWI V_{SV} shows a trade-off with all other parameter classes, but this trade-off is not strong. The experiment for δ parameter (figure 5.6) demonstrated the under-estimation of this parameter, and its amplitude is absorbed by reconstructed V_{P0} and V_{SV} wavespeeds. The reconstructed ϵ model remains unchanged. We observe the trade-off between δ parameter and V_{P0} and V_{SV} parameters in multi-parameter inversion. The reconstructed ϵ model (figure 5.7g-i) in its corresponding experiment is under-estimated (as for δ parameter experiment), and shows the amplitudes absorbed in the retrieved V_{P0} and V_{SV} models. This result suggest the trade-off between ϵ parameter and V_{P0} and V_{SV} velocities, and no significant trade-off with δ parameter. Note that ϵ has stronger trade-off with V_{P0} parameter than with V_{SV} parameter.
In overall, these mono-parameter and multi-parameter elastic FWI experiments suggest that the V_{P0} and V_{SV} wavespeeds are the dominant parameter classes for this parameterization. The Thomsen parameters are the second-order parameters in term of their influence on the data. Therefore, one should be careful about their reciprocal action when the multi-parameter inversion reconstructs the wavespeeds and Thomsen parameters together. The δ parameter shows same degree of trade-off with V_{P0} and V_{SV} parameters. The ϵ parameter also have trade-off with V_{P0} and V_{SV} parameters, more significantly with V_{P0} than with V_{SV} parameter. In comparison with acoustic VTI FWI, some remarks can be highlighted. The elastic mono-parameter inversion shows better reconstruction of individual parameter classes with respect to acoustic approximation, especially for ϵ (figure 5.7), suggesting that elastic VTI FWI can improve the reconstruction of this parameter. The elastic multi-parameter inversion shows same features as acoustic one, such as the under-estimation of Thomsen parameters and over-estimation of

V_{P0} wavespeed. The δ parameter is the weakest parameter class in both elastic and acoustic cases. Also the trade-off between parameters (in high level between wavespeeds and Thomsen parameters), seems to be a characteristic of anisotropic FWI, both in elastic and acoustic states.

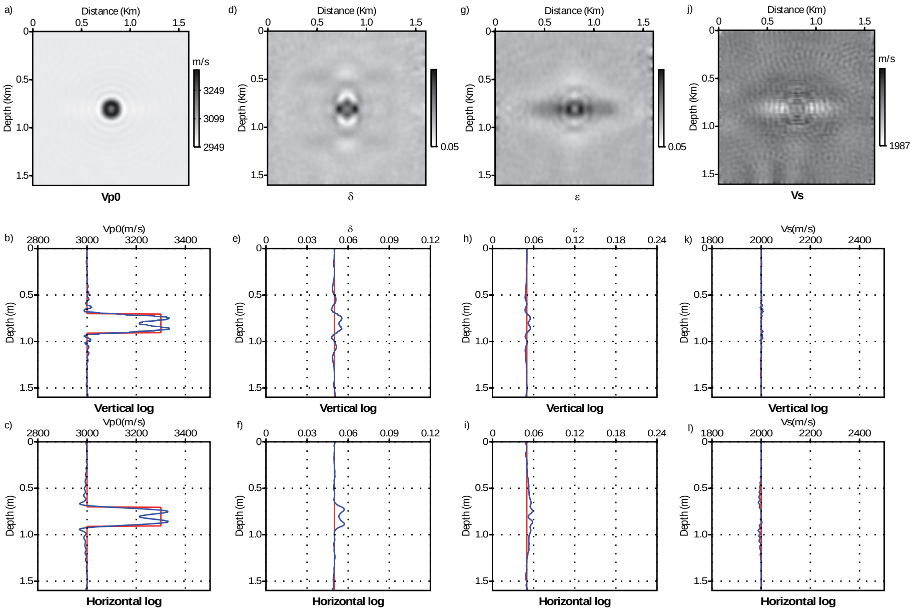


Figure 5.4: Inclusion-model FWI test. Joint update of V_{P0} (a-c), δ (d-f), ϵ (g-i), and V_{SV} (j-i) when (V_{P0},δ,ϵ,V_{SV}) parameterization is used. The true model is homogeneous in δ, ϵ, and V_{SV} and contains an inclusion in V_{P_0}. (a,d,g,j) Final FWI models. (b,e,h,k) Vertical profiles across the true inclusion (red) and the reconstructed one (blue). (c,f,i,l) Same as (b,e,h,k) for horizontal profiles.

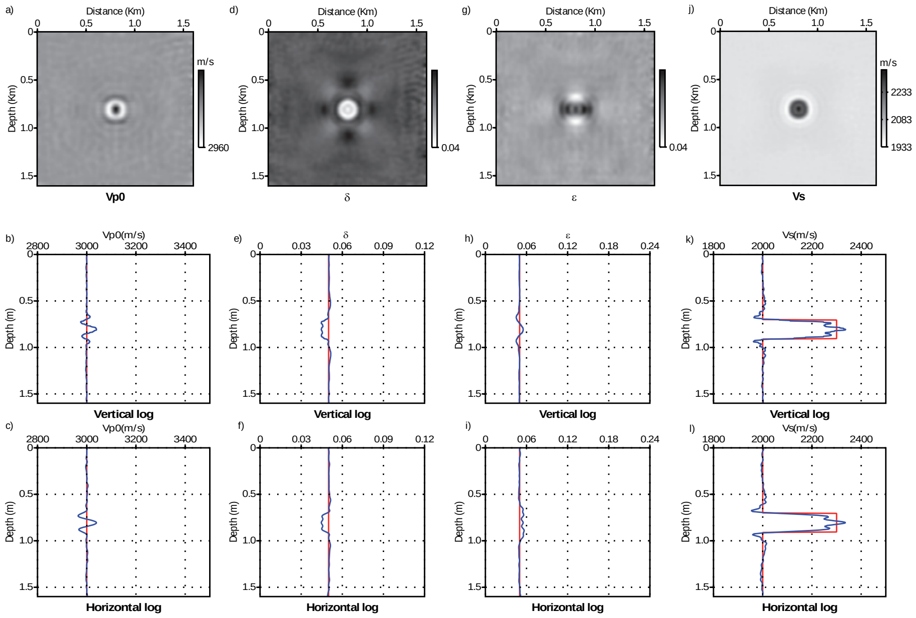


Figure 5.5: Same configuration as in figure 5.4 but the true model is homogeneous for V_{P0}, δ, and ϵ and contains an inclusion for V_{SV}.

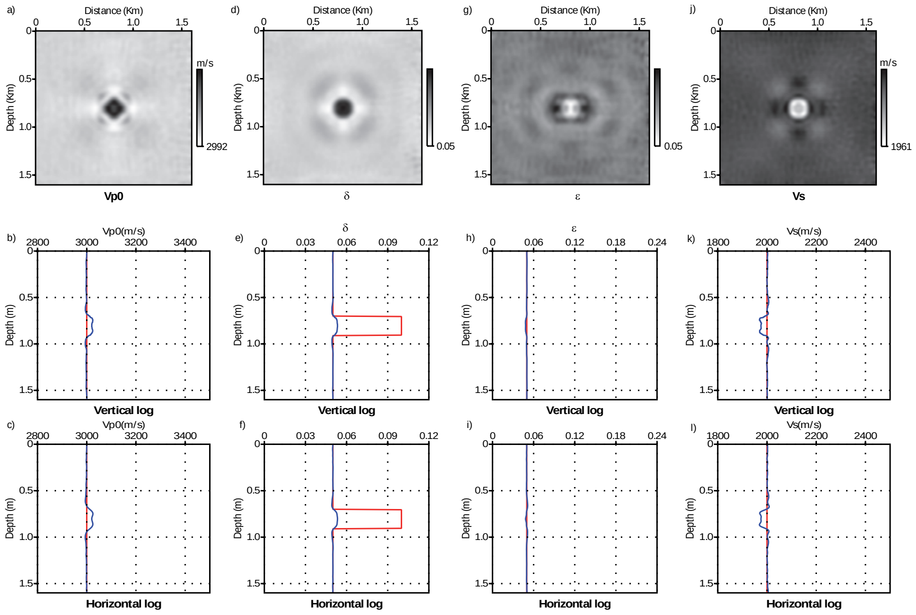


Figure 5.6: Same configuration as in figure 5.4 and 5.5 but the true model is homogeneous for V_{P0}, ϵ, and V_{SV} and contains an inclusion for δ.

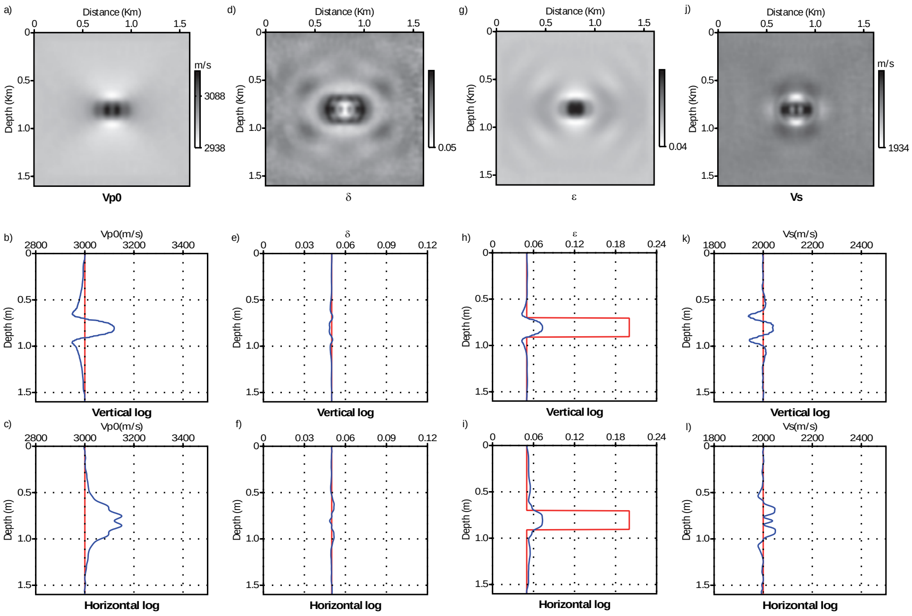


Figure 5.7: Same configuration as in figures 5.4 and 5.5 and 5.6, but the true model is homogeneous for V_{P0}, δ, and V_{SV} and and contains an inclusion for ϵ.

5.2.1.2 Parameterizations of type two: three wavespeeds + one Thomsen parameter

In this part, the joint reconstruction of an example of parameterization of type 2 is considered, when the inversion is parametrized with three wavespeeds and Thomsen's parameter δ. The multi-parameter (joint) reconstruction of V_{P0}, δ, V_h, and V_{SV} is performed, when the true medium contains an inclusion perturbation for each model parameter class (figure 5.8). The amplitudes of the V_{P0} wavespeed perturbations are estimated (figure 5.8(a-c)). The estimated V_{P0} model in this parameterization does not show the over-estimation in amplitude like the one estimated by parameterization type one (figure 5.1a-c). while the perturbations of the Thomsen parameter δ is underestimated, (figure 5.8(d-i)). The perturbations of the horizontal velocity V_h, is reconstructed, more accurate than for ϵ in figure 5.1. Therefore, this result suggests that V_h has higher influence in the data than ϵ parameter. The perturbations of the shear velocity V_{SV} is also reconstructed. The retrieved δ is under-estimated, but less significant than the one for parameterization type one (figure 5.1d-f). These results show good reconstruction of V_{P0}, V_h, and V_{SV} and equally dominant influence of wavespeeds on the data compared to the one of the Thomsen parameter δ. The diagonal coefficients of the Hessian computed in the initial models are shown in figure 5.2(e-h). For this parameterization we have $H^{(\delta)}_{max} = 0.04\% H^{(V_{P0})}_{max}$ and $H^{(V_h}_{max}) = 30\% H^{(V_{P0})}_{max}$ and $H^{(V_{SV})}_{max} = 50\% H^{(V_{P0})}_{max}$, where $H^{(V_{P0})}_{max}$, $H^{(\delta)}_{max}$, $H^{(V_h)}_{max}$ and $H^{(V_{SV})}_{max}$ denote the maximum diagonal coefficient of the Hessian for the V_{P0}, δ, V_h, and V_{SV} parameters, respectively. Note that, the maximum values of diagonal Hessian for V_h is much higher than the one for ϵ in parameterization 1, which explains better reconstruction of horizontal velocity by proper scaling the model perturbations, the same scenario is for V_{SV}. The four

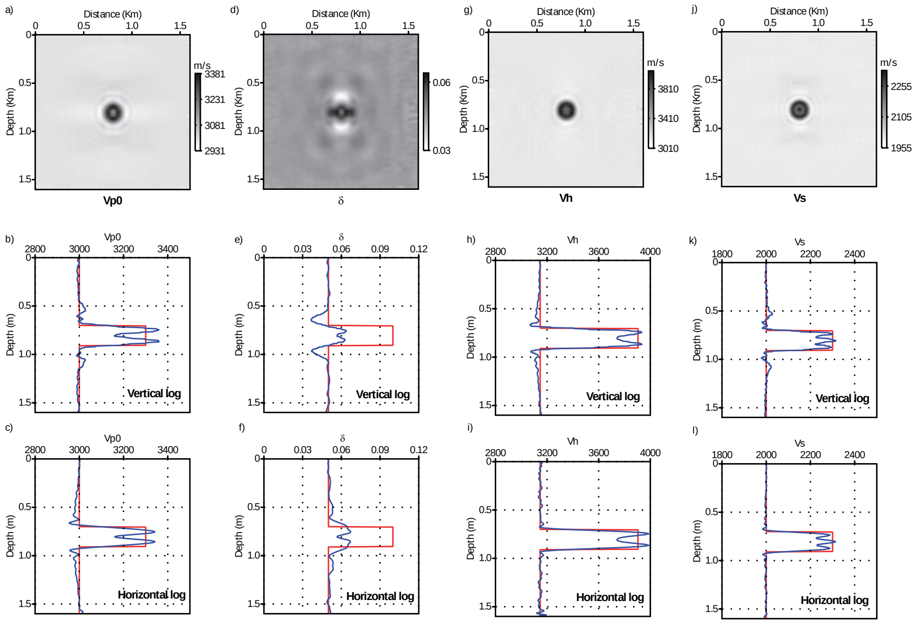


Figure 5.8: Inclusion-model FWI test. Joint update of V_{P0} (a-c), δ (d-f), V_h (g-i), and V_{SV} (j-l) when (V_{P0},δ,V_h,V_{SV}) parameterization is used. (a,d,g,j) Final FWI models. (b,e,h,k) Vertical profiles across the true inclusion (red) and the reconstructed one (blue). (c,f,i,l) Same as (b,e,h,k) for horizontal profiles.

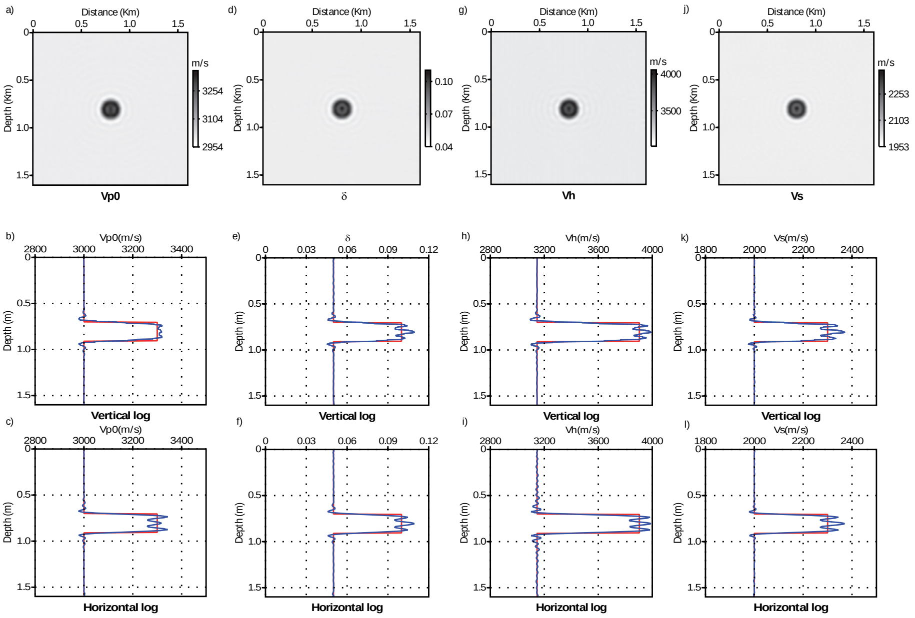


Figure 5.9: Inclusion-model FWI test. Results of mono-parameter FWI for V_{P0} (a-c), δ (d-f), V_h (g-i), and V_{SV} (j-l) when (V_{P0},δ,V_h,V_{SV}) parameterization is used. (a,d,g,j) Final FWI models. (b,e,h,k) Vertical profiles across the true inclusion (red) and the reconstructed one (blue). (c,f,i,l) Same as (b,e,h,k) for horizontal profiles. Comparison between the results shown in figure 5.8 and this figure allows one to understand that there is no distinguishable trade-off between parameters as for the one for acoustic FWI experiment.

mono-parameter FWI are performed for V_{P0}, δ, V_h, and V_{SV}, with same strategy as for parameterization type one. Again, the damping term of the Hessian is scaled to the maximum coefficient of the each corresponding mono-parameter diagonal Hessian values (equation 5.2). Repeatedly, the perfect reconstruction of all four parameters is shown in figure 5.9. Referring to the multi-parameter inversion, some trade-off can be observed. Specially for reconstructed V_{P0} and δ parameter classes. Comparing with same experimental setup applied for acoustic approximation there is a dominant improvement in reconstruction of V_h parameter class. There is not any deficit of small or large wavenumbers, suggesting (again) that elastic inversion can improve the reconstruction of parameter classes for this parameterization. Another event is the perfect reconstruction of δ parameter. Same for the parameterization 1, based on this experiment we can conclude that the under-estimation of Thomsen parameters in multi-parameter inversion is related to trade-off between this parameter and wavespeeds.
In purpose of evaluating the potential trade-off between parameter classes of this parameterization the same strategy of trade-off experiment of parameterization one is implemented here. We perform four joint inversion experiments for V_{P0}, δ, V_h, and V_{SV}, where the true models contain an inclusion associated with only one parameter at a time and all four parameters are reconstructed jointly. The experiments for V_{P0}, V_h, and V_{SV} are shown in figure 5.10, 5.12, and 5.11, respectively. These experiments illustrate that there are trade-off between δ and V_{P0} and V_{SV} wavespeeds (figures 5.10 and 5.11). The trade-off between δ and V_h (figure 5.12) is not significant. For the wavespeeds, some of amplitudes of reconstructed V_{P0} are absorbed by the retrieved V_{SV}, when the true models have perturbation in V_{P0} and other true models are homo-

geneous and kept fixed during FWI iterations (figure 5.10). The trade-off between wavespeeds is not significant, as the absorbed amplitudes are not big with respect to their background values. On the other hand, when the perturbation in true models is for δ parameter (figure 5.13), a noteworthy under-estimation of reconstructed parameter is recognized, supplemented with trade-off with V_{P0} and V_{SV} wavespeeds.
In overall, the mono-parameter and multi-parameter experiments for parameterization of type two suggest that three wavespeeds are the principal parameter classes and can be reconstructed with acceptable accuracy. The trade-off between parameters is less than the one for parameterization one. Three wavespeeds are the dominant parameter classes and the only weak parameter is δ. Hence, a good initial model of δ can improve the reconstruction of anisotropic parameters. The mono-parameter inversion demonstrates better reconstruction of individual parameter classes with respect to acoustic approximation (see chapter 3, figure 3.14), especially for V_h, suggesting again that elastic inversion improves the inversion condition (same conclusion inferred from multi-parameter inversion).

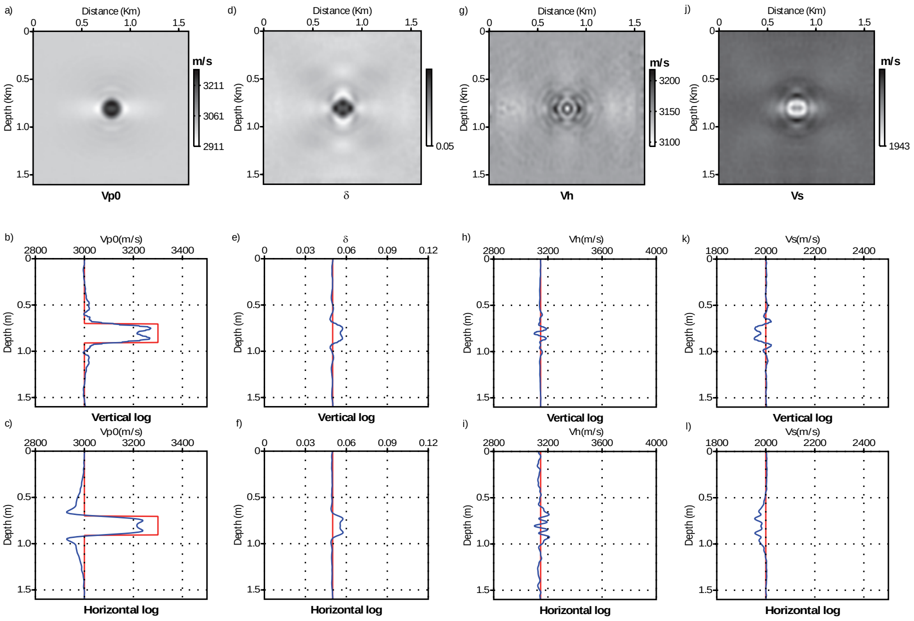


Figure 5.10: Inclusion-model FWI test. Joint update of V_{P0} (a-c), δ (d-f), V_h (g-i) and V_{SV} (j-i) when (V_{P0},δ,V_h,V_{SV}) parameterization is used. The true model is homogeneous in δ, V_h and V_{SV} and contains an inclusion in V_{P_0}. (a,d,g,j) Final FWI models. (b,e,h,k) Vertical profiles across the true inclusion (red) and the reconstructed one (blue). (c,f,i,l) Same as (b,e,h,k) for horizontal profiles.

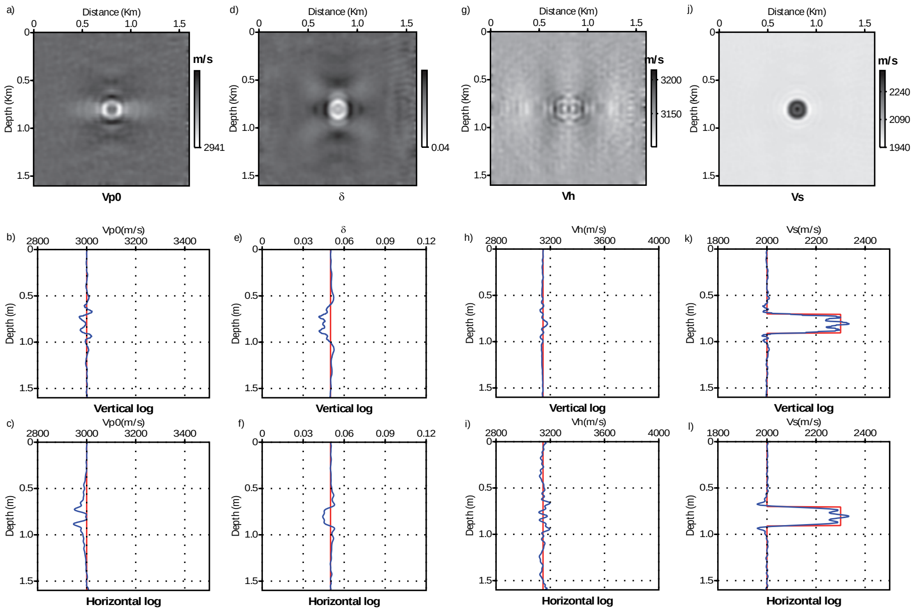


Figure 5.11: Same configuration as in figures 5.10 and 5.12 but the true model is homogeneous for V_{P0}, δ, and V_h and contains an inclusion in V_{SV}.

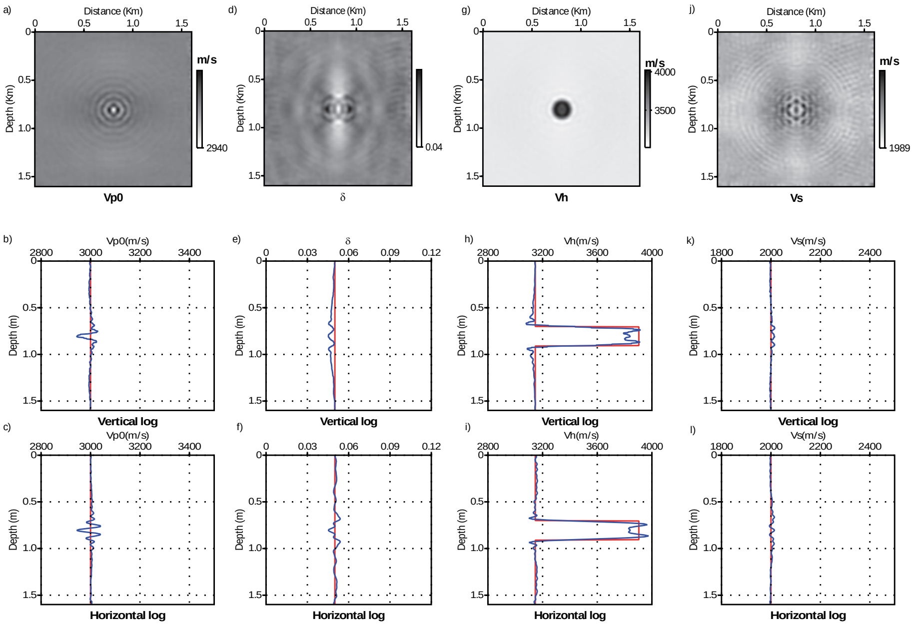


Figure 5.12: Same configuration as in figure 5.10 but the true model is homogeneous for V_{P0}, δ, and V_{SV} and contains an inclusion in V_h.

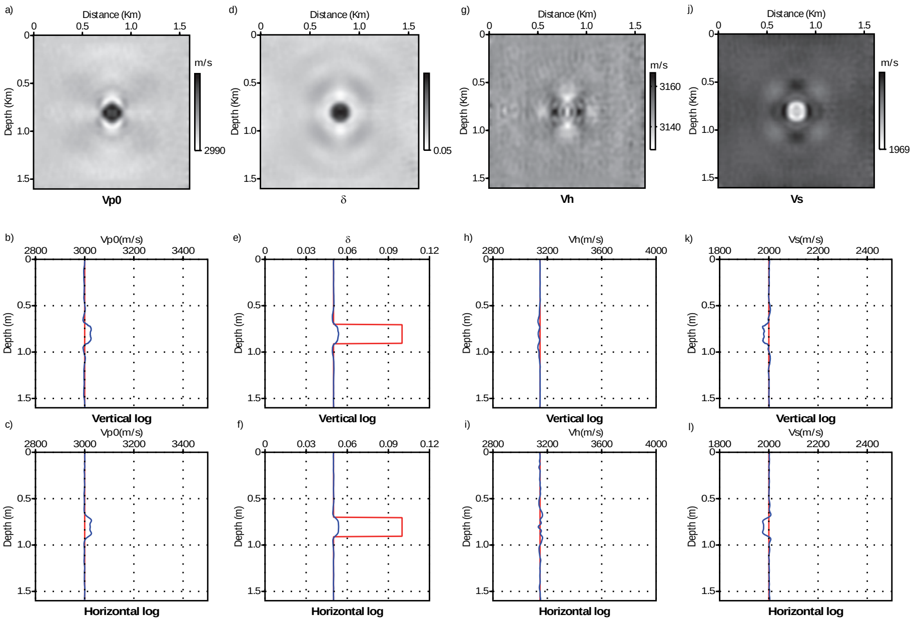


Figure 5.13: Same configuration as in figures 5.10, 5.12, and 5.11 but the true model is homogeneous for V_{P0}, V_h, and V_{SV} contains an inclusion in δ.

5.2.1.3 Parameterization type 3: elastic moduli

Results of the joint update of the elastic moduli are shown in figure 5.14. The first remarkable point is the poor reconstruction of the c_{13} model perturbations, which almost looks like a random noise without any physical sense (unlike the acoustic case in chapter 3, figure 3.17). On the other hand, the amplitudes of the c_{11} and c_{33} perturbations are better reconstructed than the ones of acoustic case. The wavenumbers content for c_{11} and c_{33} elastic moduli in both vertical and horizontal profile do not show deficit of wavenumbers. However, there are some oscillations, which can be related to weak footprint of c_{13}. The reconstruction of c_{44} in companion with well resolved images of c_{11} and c_{33} gives the hope that elastic VTI FWI can obtain better resolution with respect to acoustic VTI for elastic moduli. But the drawback is the c_{13} parameter, which the suggestion is to keep it fixed during FWI iterations as it has weak influence in the data. The reconstructions of c_{33} and c_{11} by elastic FWI are positively more reliable than the one obtained by acoustic approximation, whereas the reconstruction of c_{13} is significantly degraded, and almost nothing is recovered except artifacts. The diagonal coefficients of the Hessian computed in the initial models for elastic moduli are shown in figure 5.2(i-l). The maximum value of diagonal approximated Hessian for three parameters c_{33}, c_{11}, and c_{44} are relatively in same order ($H^{(c_{33})}_{max} = H^{(c_{11})}_{max} \simeq H^{(c_{44})}_{max}$) but $H^{(c_{13})}_{max}$=5%H_{max}. Again,

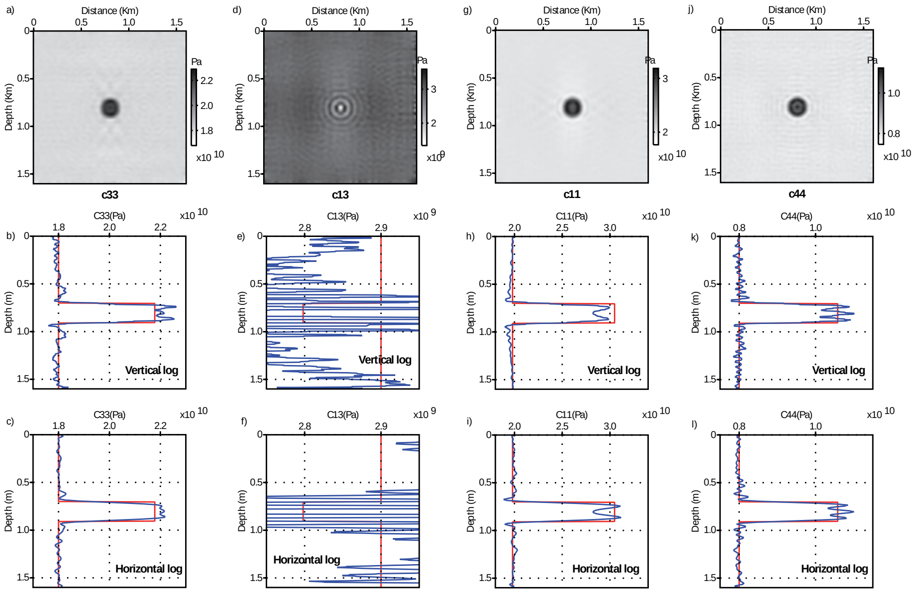


Figure 5.14: Inclusion-model FWI test. Joint update of c_{33} (a-c), c_{13} (d-f), c_{11} (g-i) and c_{44} (j-l) when (c_{33},c_{13},c_{11},c_{44}) parameterization is used. (a,d,g,j) Final FWI models. (b,e,h,k) Vertical profiles across the true inclusion (red) and the reconstructed one (blue). (c,f,i,k) Same as (b,e,h,l) for horizontal profiles. Note the random and highly noisy reconstruction of c_{13}.

four mono-parameter FWI are performed for c_{33}, c_{13}, c_{11}, and c_{44}, with same strategy as for parameterization type one and two. Besides, the damping term of the Hessian is scaled to the maximum coefficient of the mono-parameter Hessian, which is close to one for c_{33}, c_{11}, and c_{44} and far from the one for c_{13} (refer to equation 5.2 and figure 5.2(i-l)). We forecast the reliable and good reconstruction of c_{33}, c_{11}, and c_{44} parameters, which is verified by figure 5.15. Surprisingly, the c_{13} parameter class is well retrieved, which is far from the one obtained

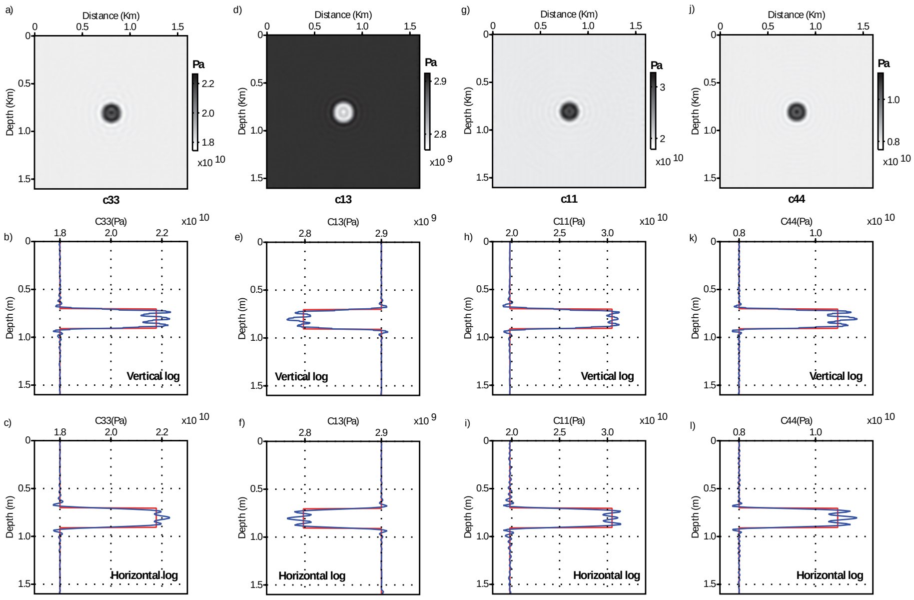


Figure 5.15: Inclusion-model FWI test. Results of mono-parameter FWI for c_{33} (a-c), c_{13} (d-f), c_{11} (g-i), and c_{44} (j-l) when (c_{33},c_{13},c_{11},c_{44}) parameterization is used. (a,d,g,j) Final FWI models. (b,e,h,k) Vertical profiles across the true inclusion (red) and the reconstructed one (blue). (c,f,i,l) Same as (b,e,h,k) for horizontal profiles. Note the good reconstruction of c_{13} with respect to one in figure 5.14.

from multi-parameter inversion. This result suggests a strong trade-off between c_{13} with other parameter class(es). Comparing with same experimental setup applied for acoustic approximation there is an improvement in reconstruction of c_{11} parameter class by both mono-parameter and multi-parameter inversion (there is not any deficit of small or large wavenumbers).

The experiment to seek the potential trade-off between c_{13} and one or more of the other elastic moduli seems necessary. Once again, we perform four joint inversion experiments for c_{33}, c_{13}, c_{11}, and c_{44}, where the true models contain an inclusion associated with only one parameter at a time and all four parameters are reconstructed jointly. The experiment for c_{33}, c_{11}, and c_{44} (figures 5.16, 5.17 and 5.18, respectively) demonstrates the strong trade-off between all these parameter classes and c_{13}. On the other hand, the experiment for c_{13} (figure 5.19) does not shows footprint of this parameter on other reconstructed parameters, suggesting that c_{13} is the most inferior and defective parameter within elastic moduli in 2D VTI FWI. Besides, the reconstructed c_{11} shows its footprint in reconstructed c_{33} for long wavelength (mostly in vertical profile), and the footprint with reconstructed c_{44} for short wavelength (mostly in horizontal profile).

One of the main conclusions from elastic inversion of elastic moduli is that the parameter class c_{13} is the weakest parameter class in term of data sensitivity and the inversion can not update this parameter by multi-parameter inversion. We perform an experiment to verify this conclusion. The experiment has same setup as for the one in figure 5.14, but only three parameter classes c_{33}, c_{11}, and c_{44} are updated during inversion iterations, while the c_{13} is kept fixed during inversion iterations (even though the true model contains perturbation related to this parameter class). As is shown in figure 5.20 the inversion succeeds to retrieve the three parameters c_{33}, c_{11} and c_{44} without any significant footprint or diffused amplitude from c_{13}. The surprising phenomenon is that their recovered perturbations is close to the one obtained by experiment shown in figure 5.14. This illustrates again the weak sensitivity of data with respect to c_{13} parameter.

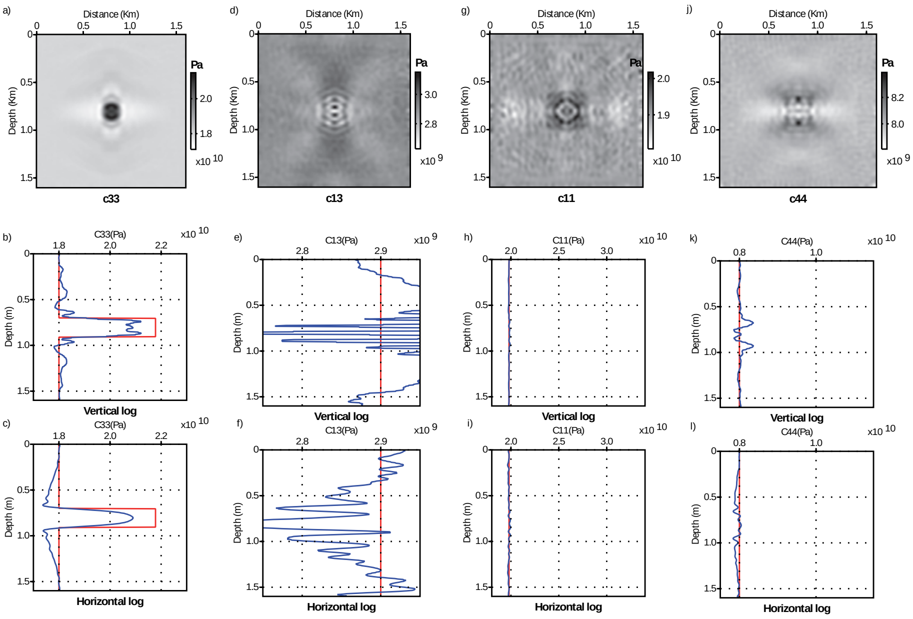


Figure 5.16: Inclusion-model FWI test. Joint update of c_{33} (a-c), c_{13} (d-f), c_{11} (g-i), and c_{44} (j-l) when (c_{33},c_{13},c_{11},c_{44}) parameterization is used. The true model is homogeneous in c_{13}, c_{11}, and c_{44} and contains an inclusion in c_{33}. (a,d,g,j) Final FWI models. (b,e,h,k) Vertical profiles across the true inclusion (red) and the reconstructed one (blue). (c,f,i,l) Same as (b,e,h,k) for horizontal profiles.

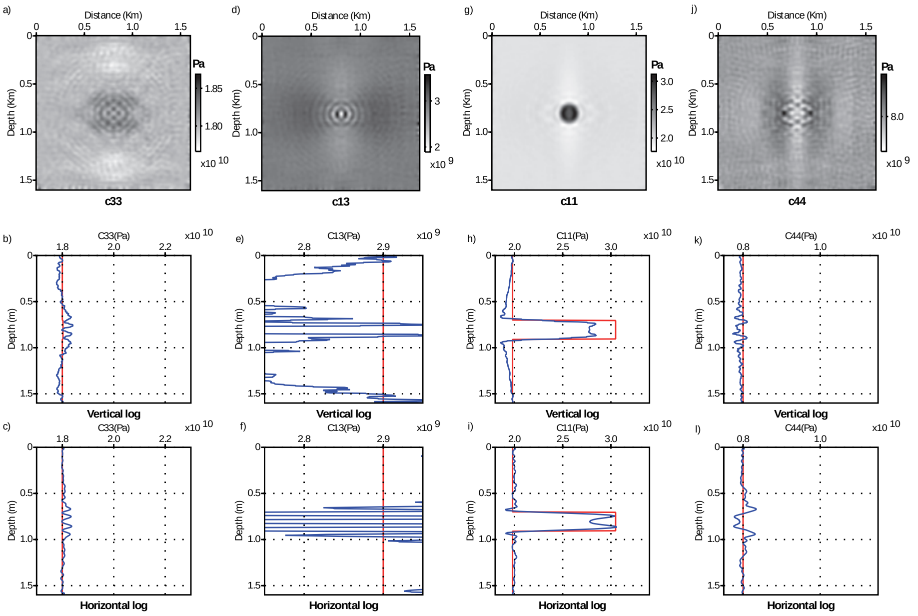


Figure 5.17: Same configuration as in figure 5.16, but the true model is homogeneous for c_{33}, c_{13}, and c_{44}, and contains an inclusion in c_{11}.

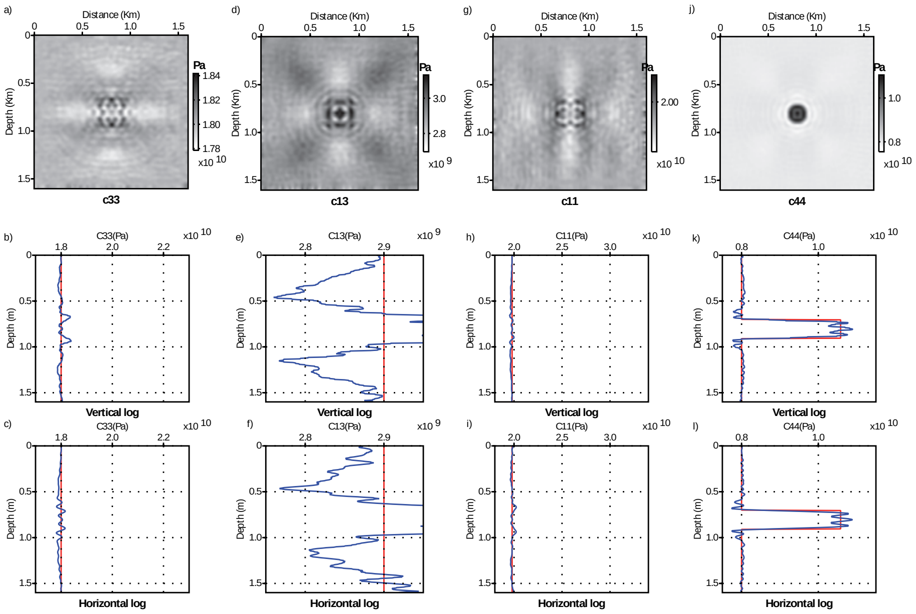


Figure 5.18: Same configuration as in figures 5.16 and 5.17, but the true model is homogeneous for c_{33}, c_{13}, and c_{11}, and contains an inclusion in c_{44}.

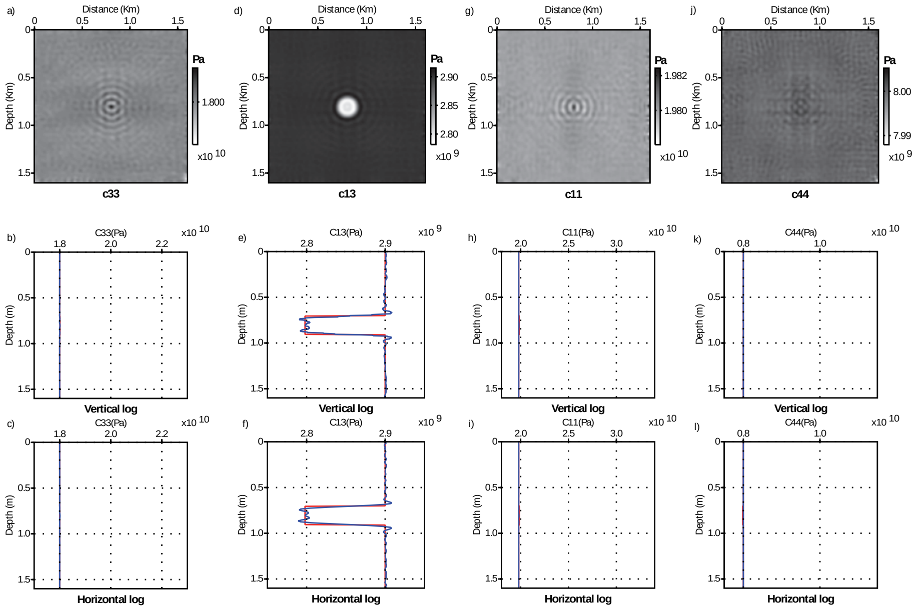


Figure 5.19: Same configuration as in figures 5.16, 5.17, and 5.18, but the true model is homogeneous in c_{33}, c_{11}, and c_{44}, and contains an inclusion in c_{13}. Surprisingly, the multi-parameter inversion succeeds to reconstruct c_{13} parameter, when the true model has perturbation only in c_{13}.

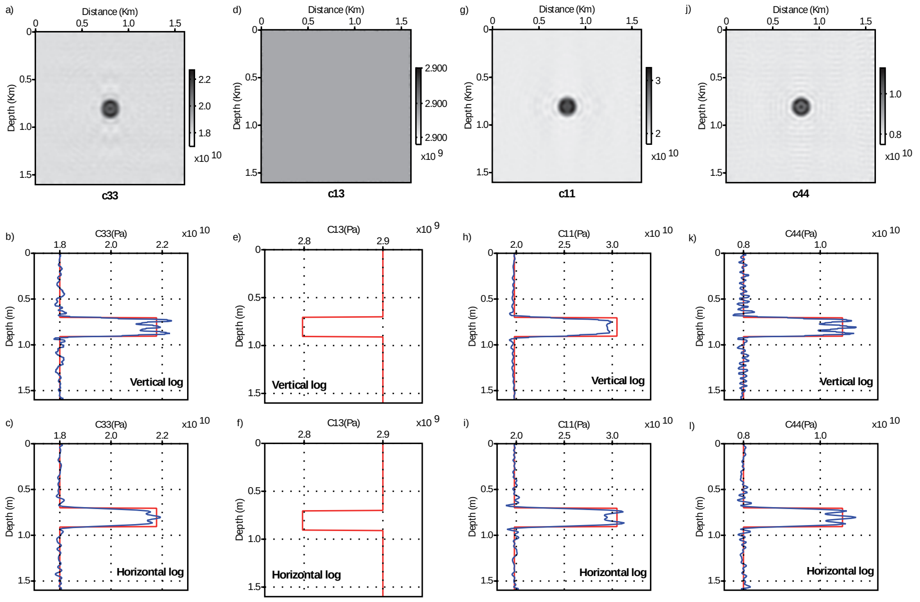


Figure 5.20: Inclusion-model FWI test. same configuration as in figure 5.14, but only the joint update of c_{33} (a-c), c_{11} (g-i), and c_{44} (j-l) is considered. (a,d,j) Final FWI models of c_{33}, c_{11}, c_{44}, and (d) the homogeneous c_{13}, which is kept fixed during inversion iterations.

5.3 Application to Valhall field

The bottleneck of FWI is the non-linearity of the inverse problem. The presence of wide-aperture components in the data increases the degree of non-linearity. The number of parameter classes involved in the inversion, also, increases the non-linearity of the problem. The provider of data residuals is the perturbations originating from variety of parameter classes with their specific order of sensitivity and influence on the dataset. Therefore, the hierarchical inversion strategy is usually suggested to ease the non-linearity of inverse problem. For example, Tarantola (1986) showed that the parameterization of FWI is a highly hierarchical problem. He proposed a hierarchical approach for elastic FWI, where a model of the long and short wavelengths of the compressional wavespeed are reconstructed, then the V_{SV} is reconstructed on the second step. The choice of hierarchical approach is highly dependant on the knowledge of the sensitivity of data with respect to parameter classes, which conducts us toward the interpretation of the resources of residuals.

Later, Prieux (2012) proposed different hierarchical approaches for reconstruction of shear wavespeed parameter class from hydrophone and geophone components of the shallow water marine data set of Valhall field. Prieux et al. (2011) applied the acoustic FWI on hydrophone component of the Valhall marine dataset. They reconstructed the pressure wavespeed. Then, the V_{P_0} model obtained from hydrophone component is inserted as the initial model for hierarchical elastic FWI. The joint reconstruction of V_{P_0} and V_{SV} parameters for hydrophone component of the data showed that the long wavelength components of the shear velocity can be reconstructed accurately. For reconstruction of higher resolution of shear wave speed model, the joint elastic FWI of the geophone components of data are proposed to estimate V_{P_0} and V_{SV} parameter classes.

The sensitivity analysis by the inclusion experiment showed that V_{P_0} and V_{SV} velocities have the higher influence on the data residual, than the δ and ϵ parameter classes. This interpretation is valid for parameterization (V_{P_0}, V_{SV}, δ, ϵ). This synthetic elastic experiment is an ideal case where all of the scattering modes are included in the data (such as SV-SV, SV-P and P-SV). But, the Valhall data set is from marine environment and there is no guaranty that the shear wavespeed components have distinguishable footprint on the dataset.

However, we follow another hierarchical approach than Tarantola (1986) and Prieux (2012). The inversion is performed only for vertical and horizontal geophone components of the data. The joint elastic inversion estimates V_{P_0} and V_{SV} parameter classes as two major parameters. The parameterization type 1 (V_{P_0}, δ, ϵ, V_{SV}) is used. The starting models for V_{P_0}, δ, ϵ, and V_{SV} are smooth and no information from inversion of hydrophone component are inserted for this parameter classes (in Prieux (2012), the initial V_{P_0} model is not smooth).

When the final model of V_{P_0} and V_{SV} wavespeeds are retrieved, their corresponding anisotropic velocities (i.e. NMO and horizontal velocities) models are computed from reconstructed V_{P_0} model and initial δ and ϵ models. These NMO and horizontal velocities models are reconstructed (updated) with parameterizations type one and two at the second step. The Thomsen parameters δ and ϵ are not reconstructed in our hierarchical procedure. We aim to reconstruct only the most resolvable elastic anisotropic parameters with an optimised computational cost. In the end, the model appraisal is performed for FWI reconstructed models of cable 29.

5.3.1 Elastic VTI Valhall models

The elastic VTI FWI is performed for two cables out of the 3D data sets, the cable 21 and 29 (courtesy of BP). As mentioned before these models were obtained by elastic anisotropic travel-time tomography. The starting models of both lines are shown in figures 5.21 and 5.23. A same experiment setup is designed for both lines. For both experiment the models of density and the quality factor of the rocks, obtained by acoustic VTI FWI from the hydrophone component, are inserted as starting models and kept fixed during inversion iterations (figures 5.22 and 5.24). The reason is to diminish their effects on the data residuals.

The starting models of cable 21

The starting models of elastic anisotropic models corresponding to cable 21 of 3D dataset are shown in figure 5.21. The P-wave velocity is strongly decreased in the gas layers, specially for some local areas. The S-wave decreases also for this formation, but in a more bigger area, and less significant than for P-wave velocity, that is consistent because the shear wavespeed influenced by the presence of gas. The density and quality factor models as starting models for elastic anisotropic FWI are shown in figure 5.22. The density and Q_P models are obtained by acoustic VTI FWI from hydrophone component of 3D 4C dataset corresponding to cable 21 (Prieux, 2012). As an assumption, the values for shear wave quality factor Q_s, are considered the same as for P-wave quality factor values. We use them as starting models in order to reduce the data residuals originating from density and quality factor physical parameters.

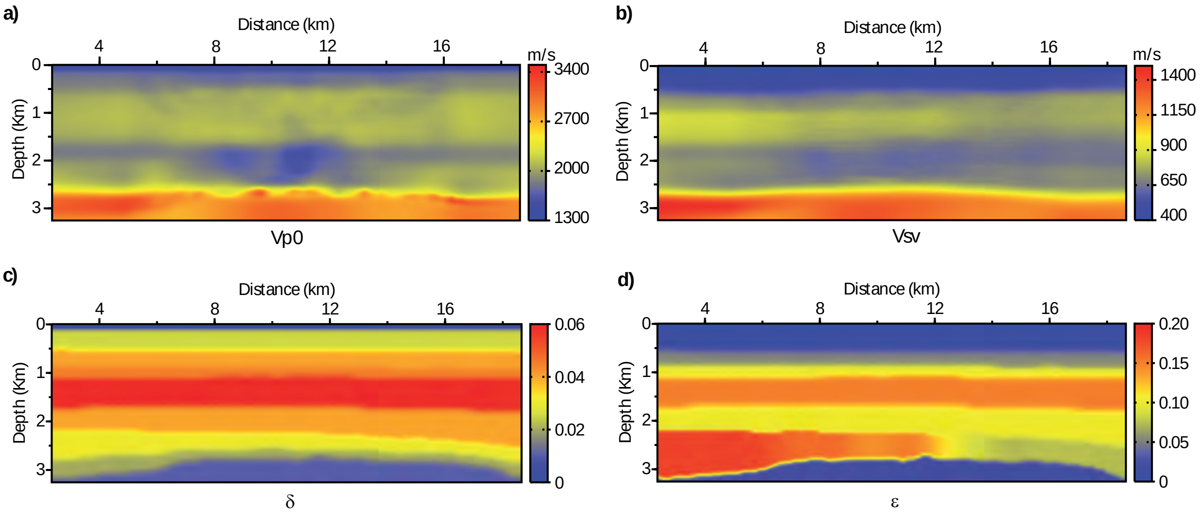


Figure 5.21: **Cable 21-** The 2D cross-section of the 3D velocity model dataset along cable 21. The elastic model (a) V_{P_0}, (b) V_{SV} (c) δ and (e) ϵ are obtained by travel-time tomography. These models are used as starting models for elastic VTI FWI in companion with density and quality factor models shown in figure 5.22.

The starting models of cable 29

The starting models of elastic anisotropic models corresponding to cable 29 of 3D dataset are shown in figure 5.23. The δ and ϵ models (5.23c-d) share many features with those of cable 21. However, the low velocity gas zone in this cable is less significant than for cable 21, because

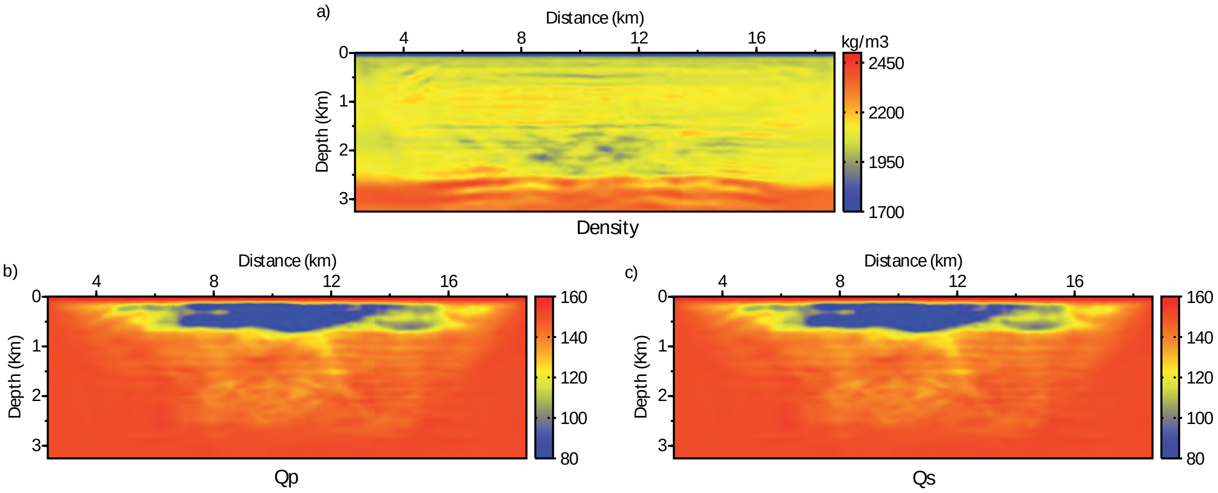


Figure 5.22: **Cable 21-** The 2D cross-section of the cable 21 for (a) the density ρ, (b) pressure wavespeed quality factor Q_P models obtained by acoustic VTI FWI from hydrophone component, and (c) shear wavespeed quality factor Q_s model (Prieux, 2012). They are used as background models for elastic VTI FWI and kept fixed during inversion iterations.

cable 29 is more distant from gas cloud. As like for cable 21, the starting models for density and

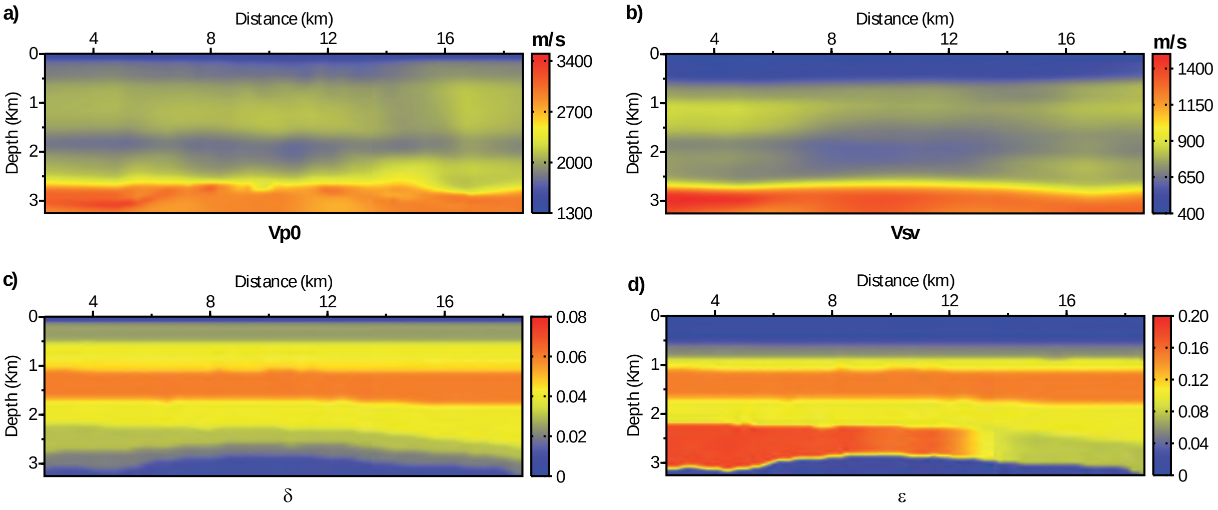


Figure 5.23: **Cable 29-** The 2D cross-section of the 3D velocity model dataset along cable 29. The elastic model (a) V_{P_0}, (b) V_{SV} (c) δ and (e) ϵ are obtained by travel-time tomography. These models are used as starting models for elastic VTI FWI in companion with density and quality factor models shown in figure 5.24.

quality factor are obtained by acoustic VTI FWI from hydrophone component (figure 5.24). Again, the values for Q_s are considered the same as for P-wave quality factor values. For the same purpose as for cable 21, we insert density and quality factor models as starting models in order to reduce the data residuals originating from these parameters.

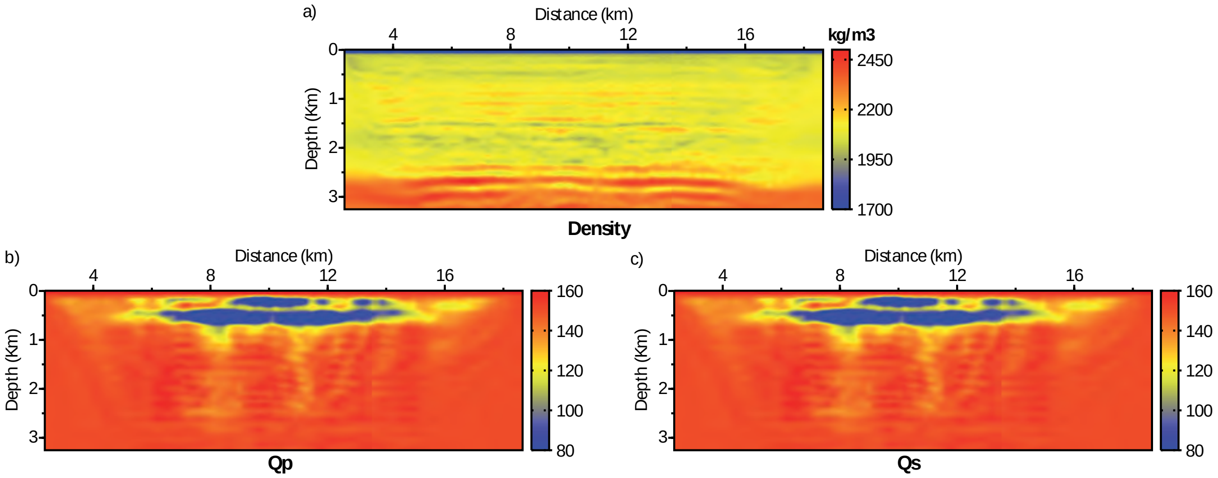


Figure 5.24: **Cable 29-** The 2D cross-section of the 29 for (a) the density ρ (b) pressure wavespeed quality factor Q_P models obtained by acoustic VTI FWI from hydrophone component and (c) shear wavespeed quality factor Q_s model (Prieux, 2012). As for the experiment of cable 21, these models are inserted as starting models for elastic VTI FWI and kept fixed during inversion iterations.

5.3.2 Elastic VTI FWI: the application on a real case study

The experimental setup

The experimental setup of elastic VTI FWI for Valhall field is similar to the acoustic VTI FWI to some extent. For elastic case we consider the vertical and horizontal geophone components of the data. The reciprocity of the source-receiver is not applied for inversion of geophone components, as the receivers have two components and the number of sources (320 shots) is less than two times the number of receivers (220 receivers). The maximum offset is 13 km. The inversion starts from 4 Hz frequency up to 6.7 Hz. The pre-processing of the geophone data consists of a minimum-phase whitening, followed by Butter-worth filtering with cut-off frequencies of 3 and 20 Hz, and a mute before the first arrivals. The 3D to 2D amplitude correction implemented by multiplying the data by $\sqrt{t}$. No time damping is applied to the geophone data during the inversion, therefore the full information content of the data components are used at each iteration. We invert four frequency groups: [4, 4.3, 4.76], [4.76, 5, 5.25], [5.25, 5.6, 6] and [6, 6.35, 6.7] Hz in two partitions (table 5.1). At first partition, the two frequency groups of [4, 4.3, 4.76] and [4.76, 5, 5.25] Hz, are involved in the inversion for a coarse mesh with a grid interval of mesh $h = 7.5m$ (2.8×10^6 cells). Then, the results obtained at first part are inserted as starting models for the second partition where two frequency groups [5.25, 5.6, 6] and [6, 6.35, 6.7] Hz, are involved in inversion with a finer mesh of $h = 6m$ (4.3×10^6 cells). The maximum number of iterations is 25 per frequency group. The density model is inferred from the acoustic FWI of V_{P_0} model (chapter 4, figure 4.17a) using the Gardner law (Gardner et al., 1974), and is kept fixed during inversion iterations for both partitions.

The starting pressure wavespeed quality factor Q_P model is the recovered model obtained by acoustic FWI (figures 5.22b and 5.24b). As mentioned before, we assumed the same quality factor values for both pressure and shear wavespeeds. Therefore, the shear weave speed quality factor Q_{SV} model is same as the Q_P model (figures 5.22c and 5.24c). Both of the quality factor models are kept fixed during FWI iterations as for the density.

	frequency (Hz)	mesh size (m)
Partition 1	[4, 4.3, 4.76] [4.76, 5, 5.25]	7.5
Partition 2	[5.25, 5.6, 6] [6, 6.35, 6.7]	6.0

Table 5.1: The four frequency groups are used during two partitions to perform the elastic VTI FWI.

The misfit function $\mathcal{C}(m)$ of the elastic inversion of Valhall is considered as:

$$\mathcal{C}(m) = \frac{1}{2}\Delta d^{\dagger} W_d \Delta d + \frac{1}{2}\sum_{i=1}^{N_p} \lambda_i \left(m_i - m_{i_{prior}}\right)^{\dagger} W_{m_i} \left(m_i - m_{i_{prior}}\right), \tag{5.3}$$

where no gain with offset is applied on the data (W_d=I). The Tikhonov regularization is applied and *a priori* model is set as starting model of each frequency group during inversion. The parameter classes are updated by

$$\Delta m_i = -\Re[W_{m_i}^{-1} J^{\dagger} W_d J + \lambda]^{-1} \Re[W_{m_i}^{-1} J^{\dagger} W_d \Delta d + \lambda(m_i - m_{i_{prior}})]. \tag{5.4}$$

The quasi-Newton l-BFGS optimization algorithm is used for updating the model parameter. An approximation of the product of the inverse of the Hessian with the gradient, from five gradients and solution vectors from previous iterations is recursively computed by l-BFGS algorithm. The Hessian matrix is estimated recursively by the l-BFGS optimization method per frequency group. The diagonal approximation of the approximate Hessian (equation 3.7) is used as initial guess of the l-BFGS algorithm per frequency group. The weighting matrices W_{m_i} seek to penalize the roughness of the difference between the model m and the prior model m_{prior}. The smoothing operators $W_{m_i}^{-1}$ are exponential functions given by

$$W_{m_i}^{-1}(z, x, z', x') = \sigma_i^2(z, x) exp\left(\frac{-|x - x'|}{\tau_x}\right) exp\left(\frac{-|z - z'|}{\tau_z}\right), \tag{5.5}$$

where τ_x and τ_z denote the horizontal and vertical correlation lengths, respectively, defined as a fraction of the local wavelength. The damping factor λ_i is defined according to the order of sensitivity of the data with respect to the parameter classes, where for vertical, horizontal, and NMO velocities (λ =4e10^8) is considered 4 times higher than for shear-wave velocity. A source wavelet per shot gather is estimated at each iteration of the geophone data by solving a linear inverse problem.

The elastic VTI FWI workflow

Generally, the cost of elastic FWI is higher in terms of both time and computing resources with respect to the acoustic FWI. Therefore, the elastic FWI needs to be performed carefully with most optimised time and cost. In this part, we reconstruct anisotropic elastic parameter classes (V_{P_0}, V_{SV}, V_{NMO} and V_h), with joint and hierarchical FWI strategies in order to retrieve all these parameter classes with optimised number of FWI experiments. This hierarchical approach is in agreement with the sensitivity of the data with respect to each parameter class.

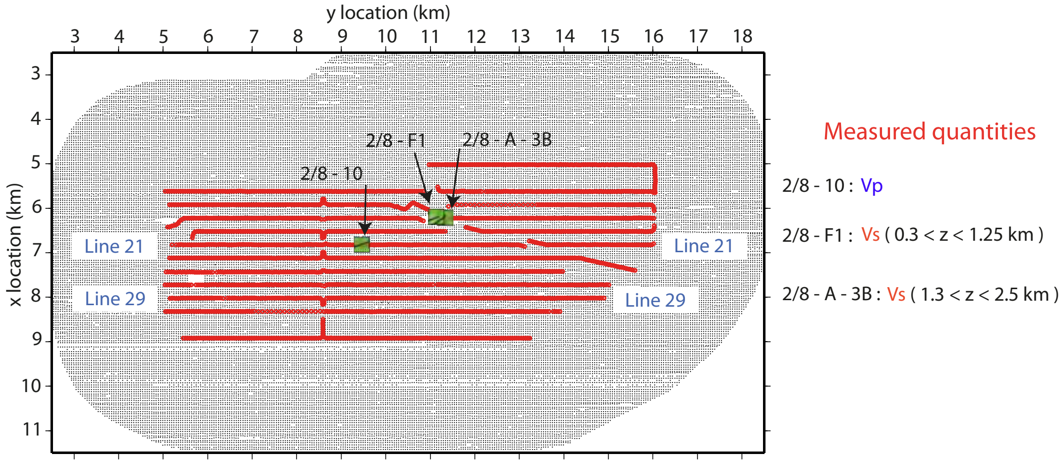


Figure 5.25: Valhall acquisition layout with the location of the wells, which are shown with green squares. The measured sonic log values of V_{P_0} and V_{SV} on the wells are used for FWI model appraisal of both 21 and 29 lines. Note that the V_{SV} well logs are measured in two wells, for two different depth intervals. But the wells positions are very close.

The same hierarchical strategy is performed for the geophone data of both cable 21 and 29, but only the results from cable 21 is shown here, because the final inferred results are the same. From the starting models shown in figures 5.21 and 5.22 (corresponding to cable 21) and figures 5.23 and 5.24 (corresponding to cable 29), the joint reconstruction of the two wavespeeds V_{P_0} and V_{SV} is performed when the parameterization type is (V_{P_0}, δ, ϵ, V_{SV}). The δ and ϵ parameters are fixed during inversion iterations due to their weak influence on the data.
Afterwards, the retrieved models of V_{P_0} and V_{SV} wavespeeds with 6.7 Hz resolution is inserted as starting model in order to apply joint inversion of V_{NMO} and V_h, when the V_{SV} and δ parameters are kept fixed during inversion iteration. The parameterization type is (V_{NMO}, δ, V_h, V_{SV}).
This strategy allows to save the computational cost for inversion up to 5.25 Hz frequency where the inversion is performed only for [5.25, 5.6, 6] and [6, 6.35, 6.7] Hz frequency groups. On the other hand, this strategy allows to compensate the data residuals originating from V_{P_0} and V_{SV} wavespeeds. All steps of this hierarchical inversion procedure are shown by a flowchart in figure 5.26.

The elastic FWI results: cable 21

The joint update of V_{P_0} and V_{SV} wavespeeds (figure 5.26, step 2) is performed by elastic VTI FWI for cable 21 (figure 5.27). The retrieved models are shown for two partitions of inversion. Figure 5.27a shows the retrieved V_{P_0} model obtained by the partition 1 (up to 5.25 Hz frequency) and its corresponding vertical profile at 9.5 km distance is shown by figure 5.27e. The final model of V_{P_0} at 6.7 Hz has a good resolution (figure 5.27b). The low velocity interfaces near 0.5 km depth are distinguishable. The two low velocity areas of gas cloud zone are clearly recognized and separated and show slight deformation of horizontal continuous layers. This

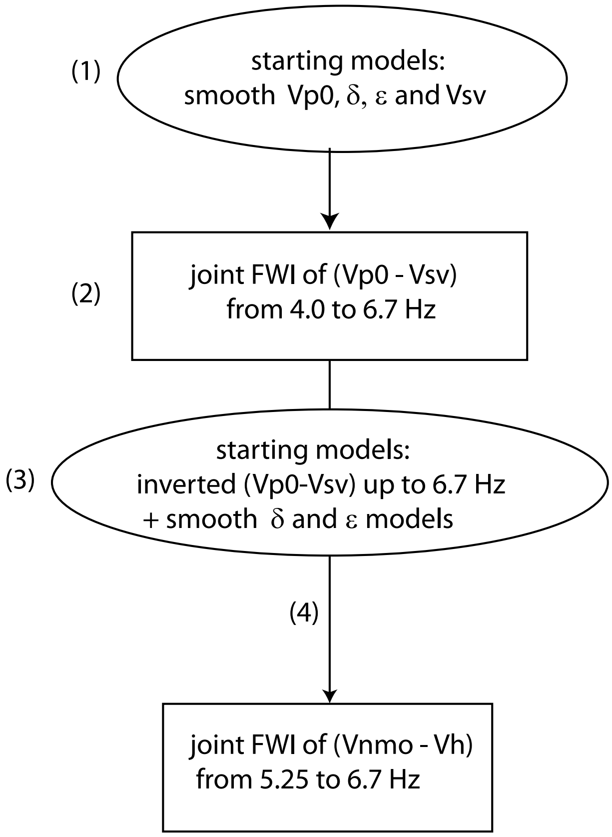


Figure 5.26: The flowchart of the performed hierarchical elastic VTI FWI. Each step of the hierarchical experiment is shown by a numbers from 1 to 6. The density and quality factor models (obtained by acoustic FWI) are kept fixed during elastic FWI iterations.

can be related to a recent local compressional tectonic stresses (as these deformation is on the top of the crest). The corresponding profile at 9.5 km distance shows a good fit with up-scaled sonic well log of V_{P_0}, dominantly for 1.5-2.0 km depth (the gas cloud zone).

The retrieved V_{SV} model obtained at both partitions are show in figures 5.27c-d. The final model obtained at 6.7 Hz is well resolved in term of positioning and focusing the reflectors. The vertical profile of FWI reconstructed model fits with up-scaled sonic well log of V_{SV} (figure 5.27g-h) at depth $\simeq$0.7-1.0 km and before 1.5 km. The retrieved shear velocity by FWI shows higher velocity than well log at depth between 1.5 km to 2.4 km. There is a gap between values of velocity for this part. The means that FWI did not succeed to update the long wavelength components of the shear wavespeeds. Only the short wavelength components are retrieved by FWI. Therefore, the final model is close to its initial model. This is maybe due to deficit of some wavelengths in the FWI estimated model (reconstructed at low frequencies, up to 6.7 Hz). Another hypothesis is that this gap is maybe related to difference in their geological properties, as the wells are far from cable 21 (figure 5.25). However, the positive point is that the variation of reconstructed image of V_{SV} with respect to the depth is close enough to the sonic well logs of 2/8-F1 and 2/8-A-3B wells.

The Poisson ratio ν, computed from the initial and retrieved V_{P_0} and V_{SV} wavespeeds are shown in figures 5.28a-b. The Poisson ratio is the indicator of presence of gas where the pressure wavespeed is decreased. The shear wave speed does not propagate through the fluid,

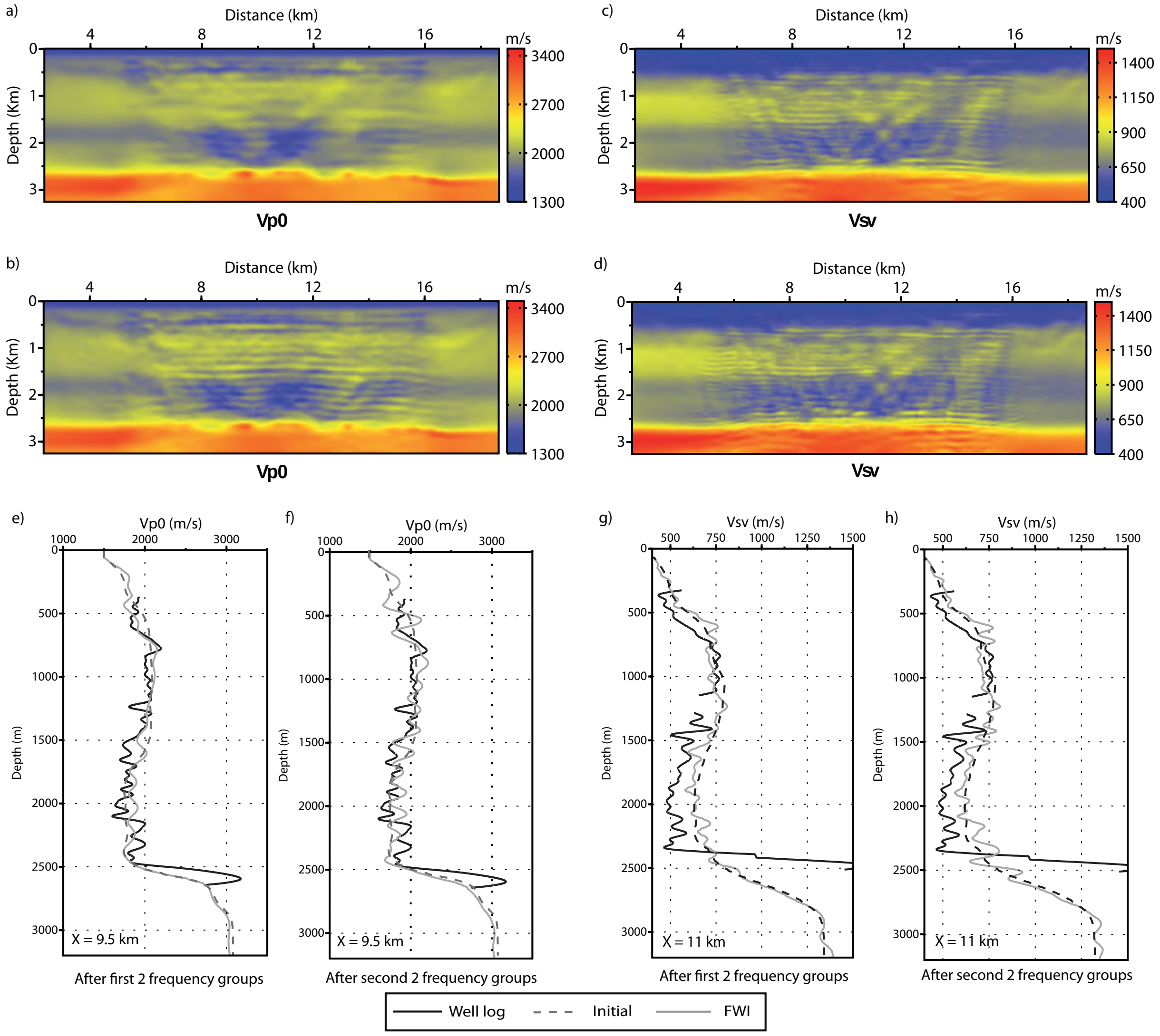


Figure 5.27: **Cable 21-** The joint elastic inversion of V_{P_0} and V_{SV} wavespeeds for cable 21 geophone data. (a-b) the reconstruction of V_{P_0} parameter after inversion of (a) the second group and (b) after fourth group of frequencies. (c-d) same as (a-b) for V_{SV} parameter (e-f) the vertical profiles at 9.5 km distance corresponding to (a) and (b), respectively. (g-h) the vertical profiles 11 km corresponding to (c) and (d), respectively.

therefore, it is not affected by gas area. The local areas where the compressional wavespeed decreases, due to gas filled layers, are highlighted with low values of Poisson ratio. In addition to Poisson ratio the V_{P_0}/V_{SV} and $V_{P_0}\times V_{SV}$ models are computed and shown in figures 5.28c and 5.28d, respectively. The V_{P_0}/V_{SV} ratio allows to discriminate the fluid saturated areas. As for the Poisson ratio model, the (local) fluid saturated areas are discriminated in the V_{P_0}/V_{SV} ratio model. The $V_{P_0}\times V_{SV}$ quantity is representative of the lithological variations in term of porosity and cracks density (Vanorio and Virieux, 2005) i.e. the areas that can be saturated by fluid. In gas filled zone the $V_{P_0}\times V_{SV}$ values decrease.

In the next step, the hierarchical elastic FWI approach for reconstruction of V_{NMO} and V_h wavespeeds is performed. The final FWI models of V_{P_0} and V_{SV} (figures 5.27b and 5.27d)

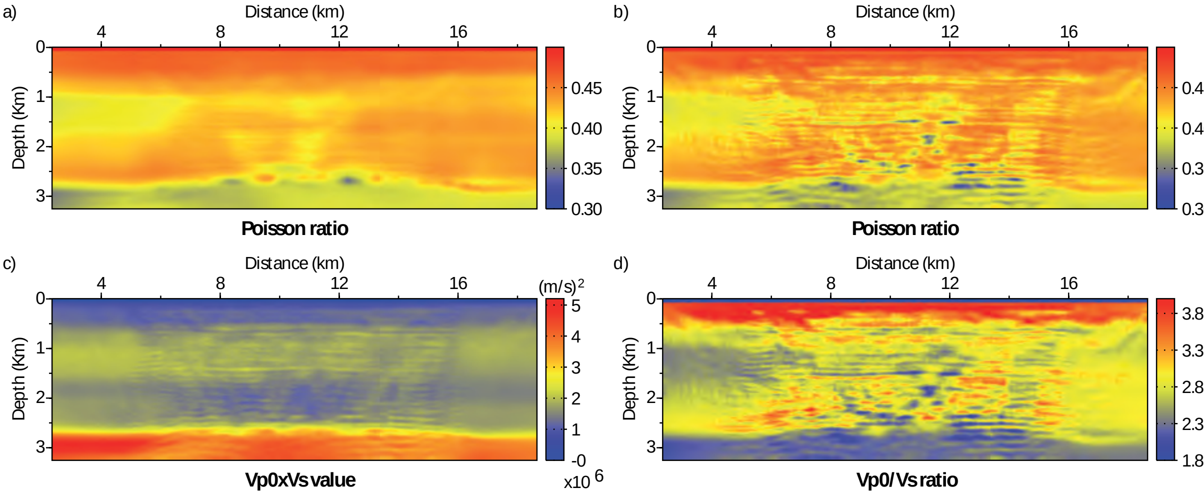


Figure 5.28: **Cable 21-** The Poisson ratio (ν) computed from (a) starting and (b) FWI final models of V_{P_0} and V_{SV} wavespeeds corresponding to cable 21. Two combinations of P-wave and S-wave velocities, (c) the $V_{P_0} \times V_{SV}$ values and (d) the V_{P_0}/V_{SV} ratio for their FWI final models. For water depth (acoustic zone) the V_{P_0}/V_{SV} ratio is set to zero.

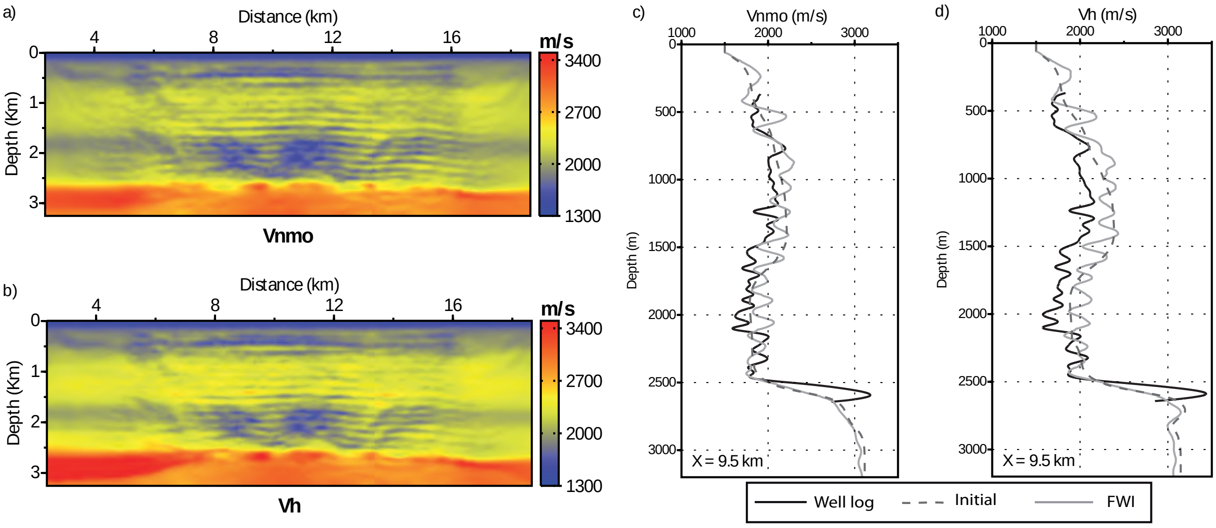


Figure 5.29: **Cable 21-** The joint elastic inversion of (a) V_{NMO} and (b) V_h wavespeeds for cable 21 geophone data. (c-d) Their corresponding vertical profiles at 9.5 km of distance. The (V_{NMO}, δ, V_h, V_{SV}) parameterization type is used.

plus the initial models of δ and ϵ (figure 5.21b-c) are inserted in order to provide the starting models for V_{NMO} and V_h wavespeeds (figure 5.26, step 3) are used as starting models. Then, the joint FWI for reconstruction V_{NMO} and V_h wavespeeds ($FWI_{V_{NMO}-V_h}$) is performed for third and fourth groups of frequencies (figure 5.26, step 4). The parameterization type is (V_{NMO}, δ, V_h, V_{SV}). The final models of the $FWI_{V_{NMO}-V_h}$ are shown in figure 5.29. Their corresponding vertical profiles at 9.5 km in distance (figure 5.29c-d) show a good fit with the sonic well log at some depth (their sonic logs are computed from the V_{P_0} well log and values of δ and ϵ parameter vertical profile at this distance). However, the fit for V_h model

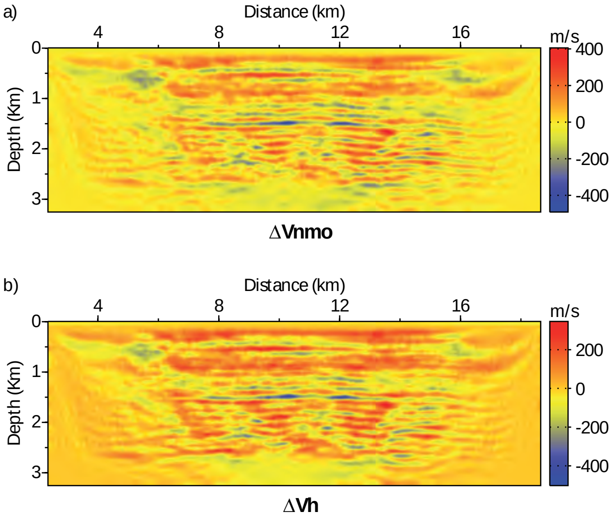


Figure 5.30: **Cable 21-** The difference between the final models of (a) V_{NMO} and (b) V_h obtained by joint elastic inversion (figure 5.29) and their initial model inserted at step 3 of the figure 5.26.

is not remarkable. The amount of update values of V_{NMO} and V_h wavespeeds by the joint inversion is displayed in figure 5.30. Their update values at 0.5 km and at 1.5 km depth are distinguishable, which show a continuous line corresponding a reflector. However, this figure shows that this hierarchical inversion approach succeed to update the value of anisotropic velocities corresponding to Thomsen parameters.

The elastic FWI results: cable 29

The joint elastic VTI FWI of V_{P_0} and V_{SV} wavespeeds is performed for cable 29 with same workflow as for cable 21 (figure 5.31). Again, both wavespeeds are retrieved with a good resolution. The vertical profile of the reconstructed compressional wavespeed shows the agreement with well log at corresponding depths. The cable 29 is not enough close to the gas filled area and as for cable 21. The V_{P_0} sonic log for model appraisal is from the well located on cable 21. However, the retrieved V_{P_0} model (figure 5.31b) shows the agreement with the sonic log (figure 5.31e-f). On the other hand, the vertical profile of the reconstructed shear wavespeed (figure 5.31d) shows a good agreement with its 2/8-F1 and 2/8-A-3B well logs, specially the retrieved model at 6.7 Hz (figure 5.31g-h).

The computed Poisson ratio from starting and final FWI V_{P_0} and V_{SV} models (figure 5.32a-b) show minor low values as trace of gas zone, except for reservoir's crest. The V_{P_0}/V_{SV} model (figure 5.32d) also highlight the fact that the P-wave velocity does not decreases very much for this cable in the gas zone area. This gas area is not saturated as much as for cable 21. The V_{P_0}/V_{SV} ration decreases in the reservoir zone.

The same hierarchical approach applied for cable 21 for reconstruction of V_{NMO} and V_h can be developed for the cable 29. The results of the next steps of hierarchical approach for cable 29 are not shown here because the final conclusion is the same.

The model appraisal of elastic VTI FWI: cable 29

The model appraisal of FWI reconstructed model permits to validate the data fit between

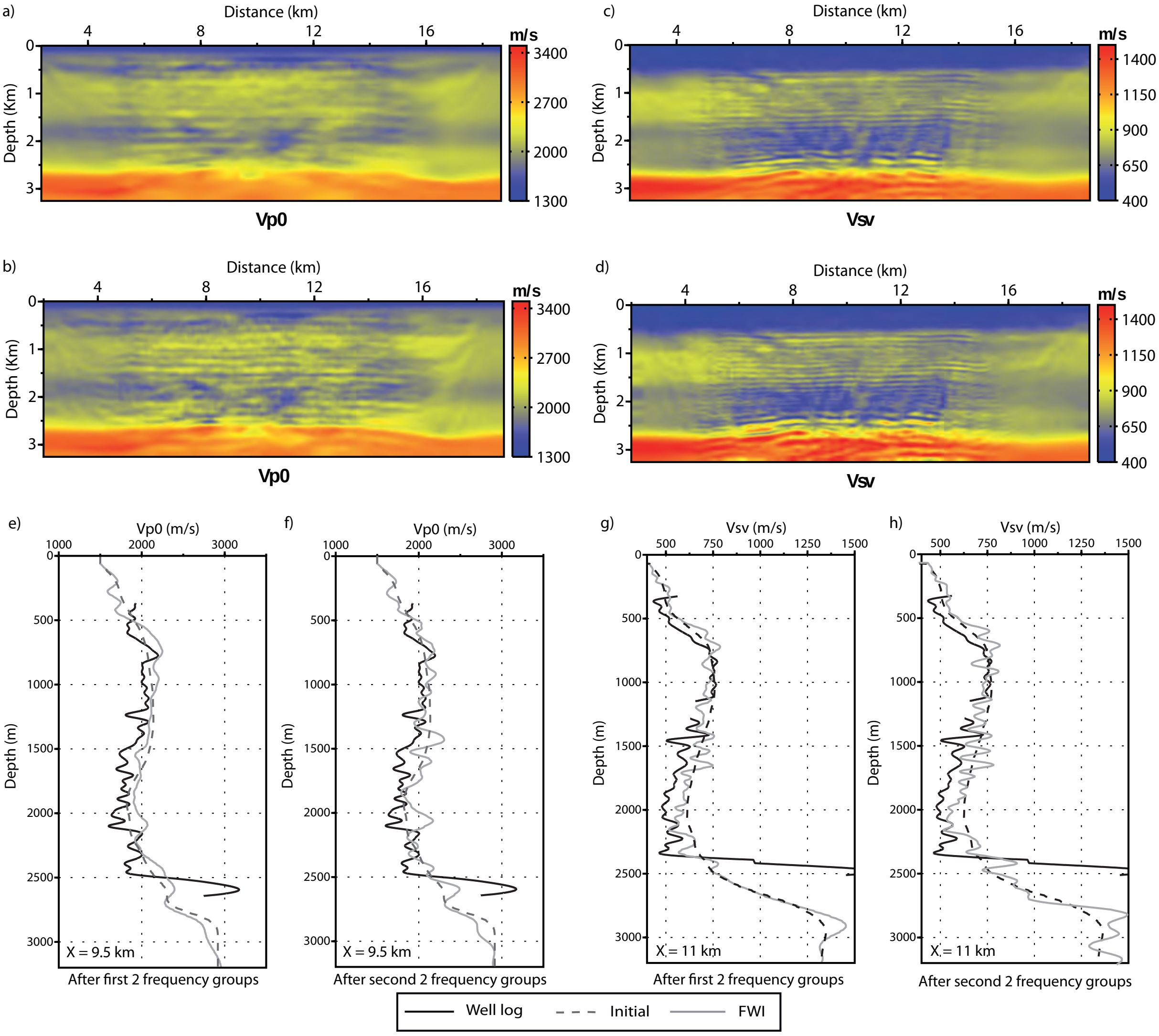


Figure 5.31: **Cable 29-** The joint elastic inversion of V_{P_0} and V_{SV} wavespeeds for cable 29 geophone data. (a-b) the reconstruction of V_{P_0} parameter after inversion of (a) two first group and (b) after two second group of frequencies. (c-d) the V_{SV} parameter after (c) the first two frequency-group and (d) the second two frequency-group. (e-f) the vertical profiles at 9.5 km distance corresponding to (a) and (b), respectively. (g-h) the vertical profiles at 11 km distance corresponding to (c) and (d), respectively.

the computed data and recorded data. In this part, we perform the model appraisal for cable 29. The overall conclusion is the same for cable 21. Time-domain shot gather computed in the final $FWI_{V_{P_0}-V_{SV}}$ models (figures 5.31b and 5.31d) computed with the (V_{P_0},V_{SV},δ,ϵ) parameterization is performed for cable 29. The smooth δ and ϵ models are used.
The recorded horizontal and vertical components shot gathers of cable 29 for a shot position at 16 km distance are shown in figures 5.33a and 5.34a, respectively. And, their corresponding computed shot gathers of the same shot for reconstructed V_{P_0} and V_{SV} models by FWI is shown in figures 5.33b and 5.34b, respectively. Some reflections are picked in computed data are localized in the recorded shot gathers. For x-component shot gather, there is relevant fit

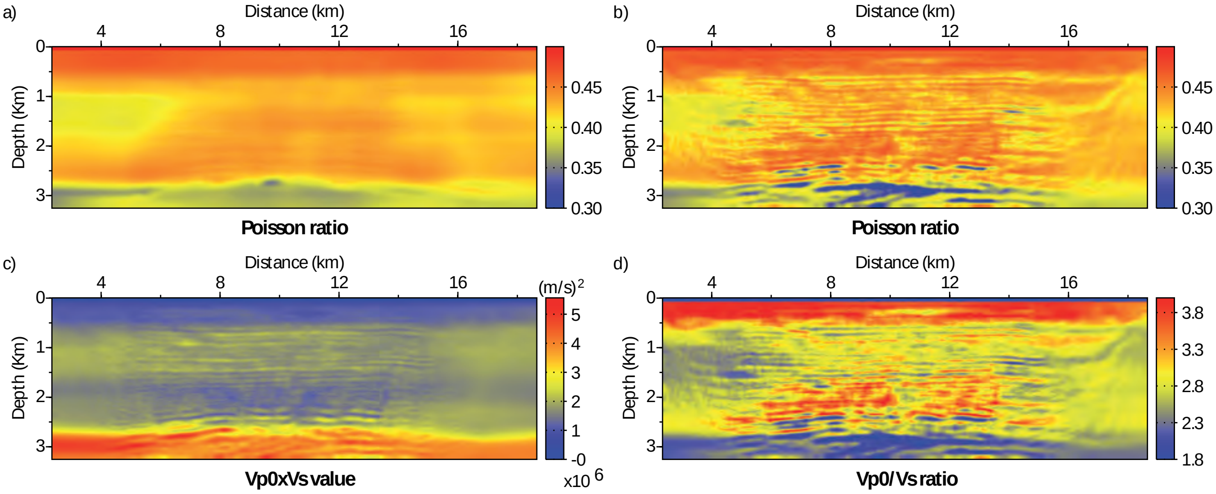
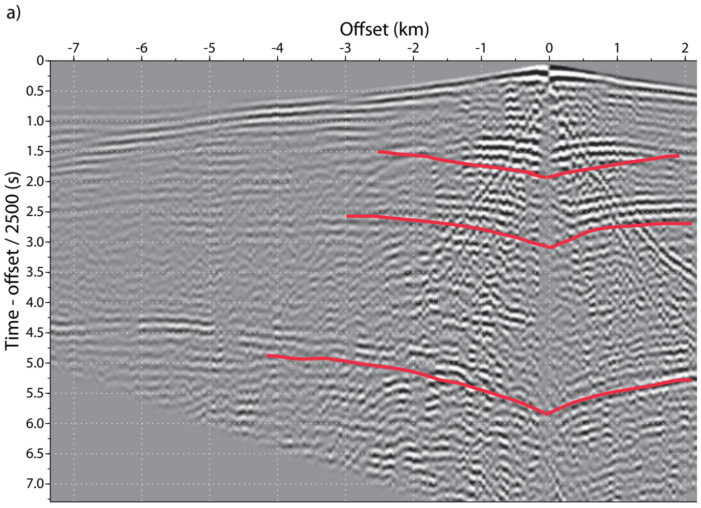


Figure 5.32: **Cable 29-** The Poisson ratio (ν) computed by (a) starting and (b) FWI final models of V_{P_0} and V_{SV} wavespeed corresponding to cable 29. Two combinations of P-wave and S-wave velocities, (c) the $V_{P_0} \times V_{SV}$ values and (d) the V_{P_0}/V_{SV} ratio for their FWI final models. For water depth (acoustic zone) the V_{P_0}/V_{SV} ratio is put to zero.

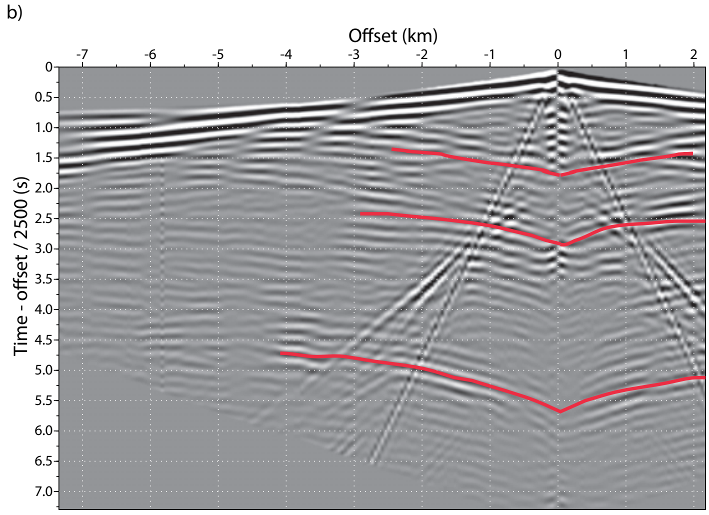
for some traces of the reflection arrivals (figure 5.33a). The fit between z-component recorded and computed data in more visible for this cable. The four picked reflections of z-component computed shot gather are clearly visible on their corresponding recorded data. For more precise demonstration, the mean of estimated sources for the FWI models is shown in figure 5.35a. The mean-estimated source is close to zero-phase wavelet. The more detailed comparison between recorded and computed data is demonstrated for x-component of geophone data by figure 5.35b. The agreement between two seismograms shows that the retrieved models by elastic FWI explain accurately the kinematics of wave propagation.

5.4 Conclusion

We demonstrated the application of elastic VTI FWI on synthetic and real dataset. The experiment on the synthetic inclusion model was performed for a full acquisition geometry and for the real data experiment with a wide-aperture acquisition survey.
The synthetic example is performed for three parameterization types. The joint and mono-parameter reconstruction of parameter classes of parameterization type one for (V_{P_0}, δ, ϵ, V_{SV}) gave the following conclusions. The potential trade-off between ϵ and wavespeeds and in higher order the trade-off between δ and wavespeeds, prevents the reconstruction of Thomsen parameters. This is similar to the acoustic VTI FWI but the reconstruction of ϵ parameter shows a slight improvement. The shear wave velocity is estimated by mono-parameter and multi-parameter inversions and show negligible trade-off with V_{P_0}. The δ is the weakest parameter class for this parameterization in terms of ability of joint FWI to recover this parameter in a stable condition. Because, this parameter has strong trade-off with two wavespeeds, V_{P_0} a,s V_{SV}.
The mono-parameter and multi-parameter elastic inversion of parameterization type 2 (V_{P_0}, δ,



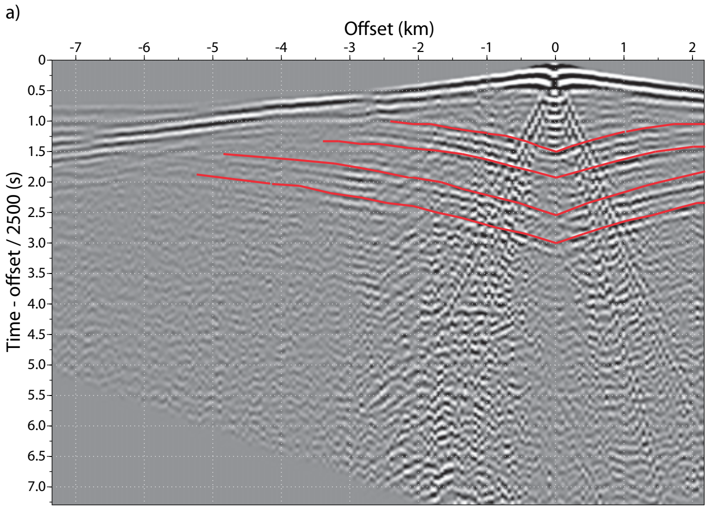
X-component observed data



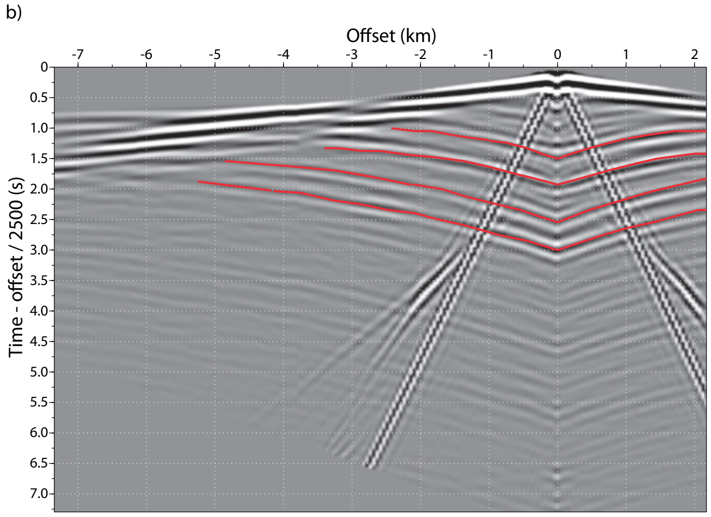
X-component computed data from FWI-reconstructed models

Figure 5.33: **Cable 29-** (a) The shot gather of recorded data for x-component of geophone. (b) The x-component shot gather computed for FWI final models of V_{P_0} and V_{SV} parameters. The starting models of δ and ϵ are used. The reflection arrivals picked in recorded seismograms (red solid lines) can be detected in the seismograms computed for FWI final models of V_{P_0} and V_{SV} parameters.

V_h, V_{SV}) show better reconstruction of these parameter classes with respect to acoustic one. The good point is the less trade-off between parameter classes and slight improvement with respect to the same acoustic experiment. The three wavespeeds are well reconstructed, while δ parameter is still the weakest parameter and shows the trade-off with V_{P_0} and V_{SV} wavespeeds. The parameterization with elastic moduli shows that except for c_{13}, the other parameter classes are well recovered by elastic inversion. The c_{13} parameter is the weakest parameter class and suffers from the highly potential trade-off with other three parameter classes (c_{11}, c_{33}, c_{44}). There is no significant trade-off between c_{11}, c_{33} and c_{44}. In general, there is an improvement of recovering anisotropic parameters by elastic VTI FWI with respect to acoustic VTI FWI for



The z-component observed data



The z-component computed data from FWI-reconstructed models

Figure 5.34: **Cable 29-** Same as in figure 5.33, but for z-components of geophone data.

c_{11} and c_{33}.
The feasibility of the elastic VTI FWI is verified against the Valhall real data with wide-aperture surface dataset. The elastic inversion of real wide-aperture data of Valhall field is performed for vertical and horizontal geophone components of the data. Firstly the V_{P_0} and V_{SV} models are reconstructed with good agreement with their corresponding up-scaled sonic well logs. This is the case for both cables 21 and 29. This results demonstrate the ability of our elastic VTI FWI for joint reconstruction of P-wave and S-wave velocities with a reliable accuracy. Finally, I proposed a hierarchical approach that permits to retrieve (jointly) V_{NMO} and V_h anisotropic wavespeeds via the parametrisation type two. The FWI model appraisal for cable 29 showed the fit between recorded and computed data in term of kinematics (in term of amplitude is not ideal). The first-arrivals and reflections showed the fit with to the data. We succeed to retrieve the model of V_{P_0} and V_{SV} wavespeeds for an anisotropic shallow water marine field, where the gas zone is challenging.

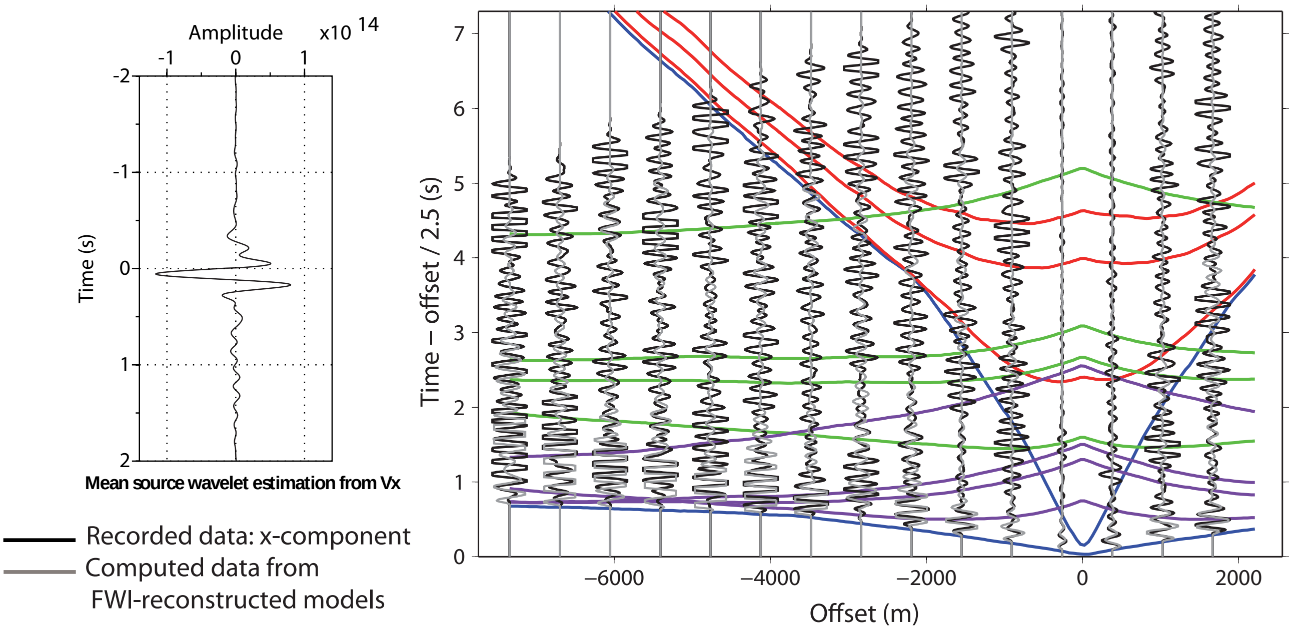


Figure 5.35: **Cable 29-** (a) The mean source wavelet estimated (Prieux, 2012) from FWI final models of V_{P_0} and V_{SV} parameters. The starting models of δ and ϵ are used. (b) The x-component seismograms of this computed data are compared with the correspond observed seismograms. The good agreement between these two seismograms validates the joint elastic inversion results.

Conclusions and perspectives

I implemented, validated, and applied the frequency-domain full waveform inversion method for imaging two dimensional anisotropic visco-acoustic and visco-elastic media. I discussed the choice of suitable parameterizations for multi-parameter VTI FWI. The four main objectives of this PhD project are: (a) the implementation of the anisotropic forward problem with a frequency-domain finite-element method, followed by validation of the forward modeling method (chapter 1), (b) the implementation of anisotopic full waveform inversion in the frequency-domain (chapters 1 and 3), (c) a sensitivity study of the reconstruction of the anisotropic parameters (anisotropic wavespeeds and Thomsen parameters) by partial derivative wavefield analysis and grid analysis of misfit function of least-squares problem. The sensitivity analysis was the support for evaluating the FWI results (chapter 3), (d) and finally, the application of the method to synthetic and real dataset, acquired in seismic exploration frameworks (chapters 4 and 5).

VTI FWI: a seismic imaging method

I presented the theory of applied seismic imaging technique in this study. The visco-acoustic and visco-elastic 2D VTI wave equations are introduced, which are used for full waveform modeling of anisotropic wave propagation as the core of the inverse problem (Brossier et al., 2010b). The linear first-order elasto-dynamic system of equations is applied in frequency-space domain and is discretized by finite-element Discontinuous Galerkin method with unstructured triangular meshes (Brossier et al., 2008; Brossier, 2011a). The system of linear equations is solved with the sparse MUMPS direct solver (Amestoy et al., 2001; MUMPS-team, 2009). The FWM is applied for different VTI and TTI synthetic examples. I validated the frequency-domain discontinuous Galerkin VTI FWM method against the time-domain finite-difference VTI FWM modeling method, and demonstrated the feasibility and robustness of the forward problem. The VTI FWM faces to some instabilities for VTI complex and highly anisotropic media in the absorbing boundaries (PML) area (Berenger, 1994). Also, the TTI FWM is unstable in case of sharp heterogeneities in the model of the angle of deviation from vertical axis (Zhang and Zhang, 2008; Fletcher et al., 2009; Liu et al., 2009; Duveneck and Bakker, 2011; Bakker and Duveneck, 2011).

The frequency-domain full waveform inversion of anisotropic medium is represented as a least-squares optimisation problem. In spite of the suitable multi-scale approaches (Pratt and Worthington, 1990; Bunks et al., 1995; Shipp and Singh, 2002; Brossier et al., 2009b; Shin and Cha, 2009), that are designed to reduce the non-linearity of FWI, the parameter classes with low influence on the data are not likely to be reconstructed by multi-parameter FWI. An alternative solution is designed. The inverse problem is parametrized in order to reconstruct anisotropic parameter by multi-parameter FWI. I proposed seven possible parameterization modes for VTI model space, and categorized them in three parameterization types (chapter 1.2.5). These parameterization types are applied inside the core of the inverse problem in order to provide the possibility to retrieve the favorite anisotropic parameters.

The sensitivity analysis of the VTI acoustic FWI is designed by a heuristic approach. The sensitivity analysis approach relies on the numerical analysis of the diffraction pattern of the different parameter classes. This analysis gave some insight on the influence of the parameters on the data as a function of the scattering angle. The diffraction pattern of scattering point represents the directivity of the virtual source located at the position of the model parameter, which generates the partial derivative wavefield. A normalization consist in scaling each parameter class by its value in the background model is applied for sensitivity analysis and VTI FWI. This normalization respects the relative amplitudes of the wavefield perturbations associated with each parameter class, and represent the real (i.e., physical) influence of the parameters in the data.

The radiation pattern analysis for parameterization type one showed that the wavespeeds (V_{P_0}, V_h or V_{NMO}) are the principal parameter classes. They contain a broad range of scattering angles (isotropic radiation pattern). The grid analysis of misfit function for parameterization type one showed that wavespeed have higher influence in the data than the Thomsen parameters. The Thomsen parameters (δ and ϵ or η) are the secondary parameters and have narrow range of scattering angles. Thomsen parameters ϵ and η have narrow banded wide-aperture scattering pattern. They are the secondary parameters but with higher influence in the data than δ parameter. In parameterization type two the wavespeeds (V_{P_0} and V_h, or V_{NMO} and V_h) are the principal parameters. But their radiation patterns are not broad banded anymore. The V_{P_0} and V_{NMO} wavespeeds have narrow short-aperture scattering pattern in parameterization type two, while their radiation patterns in parameterization type one are isotropic. The V_h wavespeed has narrow wide-aperture scattering pattern in parameterization type two. The advantage in parameterization type two is that V_{P_0} and V_h show less trade-off with respect to the potential trade-off between V_{P_0} and ϵ in parameterization type one. The radiation patterns of V_{P_0} and V_h do not over-lap. But, the radiation patterns of V_{P_0} and ϵ show over-lap for wide-aperture scattering angles. Moreover, the V_h has higher influence in the data than ϵ parameter, this is verified by grid analysis of misfit function. Thomsen parameter δ showed to be the weakest parameter class in term of its influence on the data in both parameterization types one and two. The radiation patterns of elastic moduli (c_{33}, c_{13}, and c_{11}) are similar to the one for V_{P_0}, δ, and V_h in parametrization type two. The grid analysis of misfit function demonstrate that the elastic moduli have same influence on the data.

These analysis are validated against canonical synthetic examples performed in simple subsurface model with a perfect illumination of incident angles. Then, the parameter classes that can be well recovered with reliable resolution are recognized. In case of parametrization type one, a reliable mono-parameter FWI can be performed to update the dominant parameter, while the secondary parameters are kept fixed during FWI. In this case, the wavespeeds can be reconstructed by mono-parameter FWI with good resolution, with a condition that Thomsen parameters are sufficiently accurate, and describe the long wavelength components of the medium. Among the parameterizations of parameterization type one, the parameterization combining the vertical wavespeed and the Thomsen parameters δ and ϵ should be the most suitable one for this purpose. In case of parametrization type two, the multi-parameter FWI succeed to reconstruct two wavespeeds. The starting V_{P_0} model should be precise enough to describe the long wavelength components of the data (because V_{P_0} has short-aperture radiation pattern). The δ parameter as the weakest parameter class is not reconstructed in a stable way. The sensitivity analysis is used to investigate the application of different subsurface parameterization for FWI of wide-aperture data. This investigation lead to acquire the knowledge about the sensitivity of the data to the parameter with respect to the aperture angle and the potential trade-off between parameter classes. The sensitivity of the data to the parameter for a range of aperture angle should be sufficiently high enough such that information content of the data can be extracted from noise during inversion. The trade-off between parameter classes must be low enough such that information content of one parameter can not be inserted into other parameter class(es) during inversion. The acoustic VTI FWI was applied on synthetic and real Valhall dataset for wide-aperture acquisition geometry. We showed that as long as accurate long-wavelength model of the Thomsen parameters δ and ϵ have been developed during a preliminary tomographic step, a high-resolution velocity model (either V_{P_0}, V_h or V_{NMO}), can be reconstructed by mono-parameter FWI when the subsurface parameterization combines the wavespeed with the two Thomsen parameters. We also demonstrated that the joint update of two wavespeeds (V_{P_0}/V_{NMO}, V_h), provides robust results, because the two wavespeeds have a significant influence on the data of similar magnitude but for distinct range of scattering angles. On the other hand, the trade-off artifacts do not impact significantly the joint inversion of two wavespeeds. The retrieved horizontal velocity model retrieved multi-parameter inversion (when is combined with the vertical or the normal moveout velocities in the parameterization), has a limited resolution than its mono-parameter inversion, because it has influence on the wide scattering angles only.

Due to the limited influence of the Thomsen parameter δ in the data, the suggestion is to keep its background model fixed for wide-aperture surface acquisition geometry. On the other hand, the data has higher sensitivity to ϵ and η than to δ, where I succeed to retrieve long wavelength components of these parameter classes by a hierarchical acoustic inversion approach. The accuracy of the initial models as well as any prior information, which can be used during FWI, have direct influence on the choice of suitable parameterization for acoustic VTI FWI. For example, when sufficiently accurate large-scale models of the Thomsen parameters are available, the favourable parameterization involves only in reconstruction of a high-resolution velocity model of the subsurface.

The application of elastic VTI FWI was shown on the synthetic inclusion model with full acquisition geometry, and on the real data experiment for a wide-aperture acquisition survey. The vertical and horizontal component of geophone data are involved in elastic FWI. A slight improvement of recovering anisotropic parameters were observed by elastic VTI FWI with respect to acoustic VTI FWI in inclusion synthetic experiment. I succeed to retrieve the P-wave and S-wave velocities from wide-aperture geophone component of Valhall dataset by elastic VTI FWI. Only the geophone data are used and no information from hydrophone data are inserted into the inversion. The reason is to apply the inversion in a fully elastic framework. The reconstructed compressional and shear wavespeeds shows the well resolved image of these parameters when they are verified against their up-scaled sonic well logs for cables 21 and 29. I showed that the proposed hierarchical approach permits to retrieve all of the anisotropic wavespeeds by a parametrisation type two.

The key issue in multi-parameter anisotropic FWI is accounting for the Hessian in the optimization in order to scale properly the gradients of the misfit function associated with each parameter class. We used an estimation of the inverse of the Hessian provided by l-BFGS algorithm. Our studies show that for a suitable multi-scaled inversion algorithm, the choices of regularization parameter, parametrization type and the normalization (scaling) of parameter classes has direct influence on the final results of FWI. The regularization damping term λ, conventionally added on the diagonal of the Hessian, has shown a large influence on the reconstruction of the parameters. We scaled this damping to the maximum coefficient of the Hessian. In consequence, the reconstruction of the parameter with the dominant influence on the data is favored at the expense of the reconstruction of the secondary parameter. Keeping in mind that, for multi-parameter inversion of two parameters with different sensitivity, the choice of the regularization parameter λ is a difficult task. The good choice needs trial-and-error tests to realise the best trade-off and weighting between data and model space terms of misfit function. As mentioned before, the choice of parameterization lead to change in the sensitivity of data with respect to parameter classes. Therefore, this change has direct influence on the final results of FWI. The choice of parametrization should be based on the accuracy of staring models and the information contents of the dataset. On should note that, in any parametrization type, FWI retrieves a better image of the parameter class with high influence on the data. The choice of normalization of parameter classes also has direct influence on the final results of FWI. When the normalization respects the physical influence of the parameters in the data, the inversion is likely to reconstruct the parameter classes with higher influence on the data. When this normalization is not respected the multi-parameter FWI tends to retrieve weak parameter classes. However, this issue is under investigation and needs more studies.

Perspectives

The forward problem

The forward problem for VTI and TTI media suffers from some instabilities, which are related to absorbing boundary conditions and sharp discontinuities of the medium, respectively. In the first place, the application of a more robust absorbing boundary condition for VTI FWM is suggested. In the second place, the instabilities of TTI FWM due to sharp discontinuities specially in salt dome structure should be investigated and resolved.

The extension of the method from two dimension to three dimensional anisotropic forward modeling permits to simulate the anisotropic wave propagation in nature. Reminding that, the direct solver encounters the computational memory limitation in frequency domain, therefore the time domain anisotropic 3D FWM is suggested.

The inverse problem

The sensitivity analysis is performed for the acoustic VTI medium, and the results obtained in this part fed into development and application of acoustic VTI FWI for synthetic and real dataset. The sensitivity analysis of the elastic VTI medium via the radiation pattern analysis of the partial derivative wavefields and grid analysis of the misfit function is suggested. I predict that the results obtained by acoustic VTI sensitivity analysis will have minor differences with the elastic VTI medium.

The acoustic and elastic VTI FWI are performed for data without any noise. The ability of the inversion to reconstruct anisotropic parameters in the presence of noise needs to be investigated. We predict that the weak parameter classes can not be recovered in the presence of the noise, even for a simple synthetic model with complete acquisition geometry.

In terms of optimisation algorithm, the quasi-Newton l-BFGS algorithm estimates the inverse of Hessian, which improves the multi-parameter inversion. Our applications show that the multi-parameter inversion can not jointly reconstruct two parameters with extremely different sensitivity. A robust algorithm such as Gauss-Newton method is suggested which takes into account the proper weighting of variable parameter classes during inversion. Still, there is a doubt about this perspective, but I suggest the experiment.

The acoustic and elastic VTI FWI are applied for shallow water Valhall marine dataset. In term of applications, the method needs to be verified for an on-shore dataset. The Valhall field consist of mostly horizontal layers, which represents an ideal VTI medium in reality. The validation of the method against a more complex structure (maybe) with irregular topography should be considered.

The extension of the anisotropic FWI to three dimensions provides a tool to investigate the real dimensional case studies. The sensitivity analysis as like as same procedure performed in this study is suggested. In three dimension anisotropic medium the shear waves anisotropy parameter (γ), related to vertical and horizontal shear wavespeeds variations, plus δ and ϵ are the dimensionless anisotropic parameter classes. The Valhall 3D dataset is a very suited case study for application of the 3D VTI FWI, even though this field is characterized by some azimuthal anisotropy due to fracturing in variable scales, specially on the crest of the anticlinal.

In order to move beyond the VTI anisotropy, the forward modeling for more complex anisotropic medium such as orthorhombic is suggested. The orthorhombic forward modeling permits to apply the orthorhombic FWI on the dataset from the fields with fracturing and azimuthal anisotropy.

Bibliography

Abubakar, A., Li, M., Pan, G., Liu, J., and Habashy, T. M. (2011). Joint mt and csem data inversion using a multiplicative cost function approach. *Geophysics*, 76(3):F203–F214.

Aki, K. and Richards, P. (1980). *Quantitative Seismology: Theory and Methods.* W. H. Freeman & Co, San Francisco.

Alford, R. M., Kelly, K. R., and Boore, D. M. (1974). Accuracy of finite-difference modeling of the acoustic wave equation. *Geophysics*, 39:834–842.

Alkhalifah, T. (1996). *Seismic processing in transversely isotropic media.* PhD thesis, Center of Wave Phenomena, Colorado School of Mines.

Alkhalifah, T. (1998). Acoustic approximations for processing in transversely isotropic media. *Geophysics*, 63:623–631.

Alkhalifah, T. (2000). An acoustic wave equation for anisotropic media. *Geophysics*, 65:1239–1250.

Alkhalifah, T. and Tsvankin, I. (1995). Velocity analysis for transversely isotropic media. *Geophysics*, 60:1550–1566.

Alterman, Z. and Karal, F. C. (1968). Propagation of elastic waves in layared media by finite-difference methods. *Bulletin of the Seismological Society of America*, 58:367–398.

Amestoy, P. R., Duff, I. S., Koster, J., and L'Excellent, J. Y. (2001). A fully asynchronous multifrontal solver using distributed dynamic scheduling. *SIAM Journal of Matrix Analysis and Applications*, 23(1):15–41.

Amestoy, P. R., Duff, I. S., and L'Excellent, J. Y. (2000). Multifrontal parallel distributed symmetric and unsymmetric solvers. *Computer Methods in Applied Mechanics and Engineering*, 184:501–520.

Amestoy, P. R., Duff, I. S., L'Excellent, J. Y., and Koster, J. (2003). *Multifrontal massively parallel solver (MUMPS version 4.3): Users'guide.* http://enseeiht.fr/apo/MUMPS/.

Anagaw, A. and Sacchi, M. (2011). Regularized 2d acoustic full waveform inversion. In *Expanded Abstracts.* EAGE.

Audebert, F., Nichols, D. E., Rekdal, T., Biondi, B., Lumley, D. E., and Urdaneta, H. (1997). Imaging complex geologic structure with single-arrival Kirchhoff prestack depth migration. *Geophysics*, 62(5):1533–1543.

Backus, G. E. (1962). Long-wave elastic anisotropy produced by horizontal layering. *Journal Geophysical Research*, 67:4427–4440.

Bakker, P. M. and Duveneck, E. (2011). Stability analysis for acoustic wave propagation in tilted TI media by finite differences. *Geophysical Journal International*, 185:911–921.

Bakulin, A., Grechka, V., and Tsvankin, I. (2000a). Estimation of fracture parameters from reflection seismic data—Part II: fractured models with orthorhombic symmetry. *Geophysics*, 65(6):1803–1817.

Bakulin, A., Grechka, V., and Tsvankin, I. (2000b). Estimation of fracture parameters from reflection seismic data—Part III: fractured models with monoclinic symmetry. *Geophysics*, 65(6):1818–1830.

Barkved, O. and Heavey, P. (2003). Society of petroleum engineers inc. Valhall field - still on plateau after 20 years of production. *This paper was prepared for presentation at Offshore Europe 2003 held in Aberdeen, UK.*

Barnes, C., Charara, M., and Tsuchiya, T. (2008). Feasibility study for an anisotropic full waveform inversion of cross-well data. *Geophysical Prospecting*, 56:897–906.

Baysal, E., Kosloff, D., and Sherwood, J. (1983). Reverse time migration. *Geophysics*, 48:1514–1524.

Bécache, E., Fauqueux, S., and Joly, P. (2003). Stability of perfectly matched layers, group velocities and anisotropic waves. *Journal of Computational Physics*, 188(2):399 – 433.

Bécache, E., Petropoulos, P. G., and Gedney, S. G. (2004). On the long-time behavior of unsplit perfectly matched layers. *IEEE Transactions on Antennas and Propagation*, 52:1335–1342.

Ben Hadj Ali, H. (2009). *Three dimensional visco-acoustic frequency domain full waveform inversion.* PhD thesis, Université de Nice-Sophia-Antipolis.

Ben Jemaa, M., GLinsky-Olivier, N., Cruz-Atienza, V., Virieux, J., and Piperno, S. (2007). Dynamic non-planar crack rupture by a finite volume method. *Geophysical Journal International*, 172(1):271–285.

Ben Menahem, A. and Gibson, R. L. (1990). Scattering of elastic waves by localized anisotropic inclusions. *Journal of Acoustical Society of America*, 87:2300–2309.

Berenger, J.-P. (1994). A perfectly matched layer for absorption of electromagnetic waves. *Journal of Computational Physics*, 114:185–200.

Berryman, J. G., Grechka, V., and Berge, P. A. (1999). Analysis of Thomsen parameters for finely layered VTI media. *Geophysical Prospecting*, 47:959–978.

Bevc, D. (1997). Imaging complex structures with semirecursive Kirchhoff migration. *Geophysics*, 62(2):577–588.

Beydoun, W. B. and Mendes, M. (1989). Elastic ray-Born ℓ^2 - migration/inversion. *Geophysical Journal*, 97:151–160.

Beydoun, W. B. and Tarantola, A. (1988). First Born and Rytov approximation: Modeling and inversion conditions in a canonical example. *Journal of the Acoustical Society of America*, 83:1045–1055.

Beylkin, G. (1985). Imaging of discontinuities in the inverse scaterring problem by inversion of a causal generalized Radon transform. *Journal of Mathematical Physics*, 26:99–108.

Boore, D. M. (1972). Finite-difference methods for seismic wave propagation in heterogeneous materials. In Ed., B. B. A., editor, *Methods in computational physics*, volume 11. Academic Press, Inc.

Born, M. and Wolf, E. (1980). *Principles of optics.* Pergamon Press, 6^{th} edition edition.

Born, M. and Wolf, E. (1993). *Principles of optics : electromagnetic theory of propagation, interference and diffraction of light; sixth edition with corrections.* Pergamon Press.

Bostock, M. (2003). Linearized inverse scattering of teleseismic waves for anisotropic crust and mantle structure: I. theory. *Journal of Geophysical Research*, 108-B5:2258.

Bracewell, R. N. (1986). *The Fourier transform and its applications, second edition, revised.* McGraw-Hill International editions.

Brenders, A. J. and Pratt, R. G. (2007a). Efficient waveform tomography for lithospheric imaging: implications for realistic 2D acquisition geometries and low frequency data. *Geophysical Journal International*, 168:152–170.

Brenders, A. J. and Pratt, R. G. (2007b). Full waveform tomography for lithospheric imaging: results from a blind test in a realistic crustal model. *Geophysical Journal International*, 168:133–151.

Brenner, S. and Ridgway Scott, L. (2008). *The Finite Element Method for Solid and Structural Mechanics.* McGraw Hill, New York.

Brossier, R. (2009). *Imagerie sismique à deux dimensions des milieux visco-élastiques par inversion des formes d'onde: développements méthodologiques et applications.* PhD thesis, Université de Nice-Sophia-Antipolis.

Brossier, R. (2011a). Two-dimensional frequency-domain visco-elastic full waveform inversion: Parallel algorithms, optimization and performance. *Computers & Geosciences*, 37(4):444 – 455.

Brossier, R. (2011b). Two-dimensional frequency-domain visco-elastic full waveform inversion: Parallel algorithms, optimization and performance. *Computers & Geosciences*, 37(4):444 – 455.

Brossier, R., Etienne, V., Operto, S., and Virieux, J. (2010a). Frequency-domain numerical modelling of visco-acoustic waves based on finite-difference and finite-element discontinuous galerkin methods. In Dissanayake, D. W., editor, *Acoustic Waves*, pages 125–158. SCIYO.

Brossier, R., Gholami, Y., Virieux, J., and Operto, S. (2010b). 2D frequency-domain seismic wave modeling in VTI media based on a hp-adaptive discontinuous galerkin method. In *Expanded Abstracts, 72^{th} Annual International Meeting, EAGE.*

Brossier, R., Operto, S., and Virieux, J. (2009a). 2D elastic frequency-domain full-waveform inversion for imaging complex onshore structures. In *Expanded Abstracts*, page U019. EAGE.

Brossier, R., Operto, S., and Virieux, J. (2009b). Seismic imaging of complex onshore structures by 2D elastic frequency-domain full-waveform inversion. *Geophysics*, 74(6):WCC63–WCC76.

Brossier, R., Operto, S., and Virieux, J. (2010c). Which data residual norm for robust elastic frequency-domain full waveform inversion? *Geophysics*, 75(3):R37–R46.

Brossier, R., Virieux, J., and Operto, S. (2008). Parsimonious finite-volume frequency-domain method for 2-D P-SV-wave modelling. *Geophysical Journal International*, 175(2):541–559.

Broto, K. and Ehinger, A. (1998). How to access 3D prestack kinematic information in case of complex geologies? In *Expanded Abstracts*, pages 1401–1404. Soc. Expl. Geophys.

Bunks, C., Salek, F. M., Zaleski, S., and Chavent, G. (1995). Multiscale seismic waveform inversion. *Geophysics*, 60(5):1457–1473.

Burridge, R., de Hoop, M., Miller, D., and Spencer, C. (1998). Multiparameter inversion in anisotropic elastic media. *Geophysical Journal International*, 134:757–777.

Cagniard, L. (1962). *Reflection and refraction of progressive seismic waves.* McGraw-Hill (translated from Cagniard 1932).

Calvet, M., Chevrot, S., and Souriau, A. (2006). P-wave propagation in transversely isotropic media i. finite-frequency theory. *Physics of the Earth and Planetary Interiors*, 56:12–20.

Cara, M. (2002). Seismic Anisotropy, in International Handbook of Earthquake and Engineering Seismology, Part A.

Carcione, J. (1996). Wave propagation in anisotropic, saturated porous media: Plane-wave theory and numerical simulation. *Journal of Acoustical Society of America*, 99(5):2655–2666.

Carcione, J., Kosloff, D., Behle, A., and Seriani, G. (1992). A spectral scheme for wave propagation simutation in 3D elastic-anisotropic media. *Geophysics*, 57:1593–1607.

Carcione, J. M. (1998). Radiation patterns for 2-D GPR forward modeling. *Geophysics*, 63(2):424–430.

Carcione, J. M. and Helle, H. (2004). The physics and simulation of wave propagation at the ocean bottom. *Geophysics*, 69(3):825–839.

Casadei, F., Gabellini, E., Fotia, G., Maggio, F., and Quarteroni, A. (2002). A mortar spectral/finite element method for complex 2D and 3D elastodynamic problems. *Computer Methods in Applied Mechanics and Engineering*, 191:5119–5148.

Casarotti, E., Stupazzini, M., Lee, S.-J., Komatitsch, D., Piersanti, A., and Tromp, J. (2007). CUBIT and seismic wave propagation based upon the spectral-element method: An advanced unstructured mesher for complex 3D geological media. pages 579–597. Springer.

Castellanos, C., Etienne, V., Hu, G., Operto, S., Brossier, R., and Virieux, J. (2011). Algorithmic and methodological developments towards full waveform inversion in 3d elastic media. *SEG Technical Program Expanded Abstracts*, 30:2793–2798.

Červený, V. (1985). The application of ray tracing to the numerical modeling of seismic wavefields in complex structures. In Helbig, K. and Treitel, S., editors, *Handbook of geophysical exploration, section I: Seismic exploration*, volume 15, pages 1–124. G. Dohr, Geophysical Press, London.

Červený, V. (1999). Ray tracing in factorized anisotropic inhomogeneous media. *Geophysical Journal International*, pages 91–100.

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

Červený, V., Molotkov, I. A., and Pšenčik, I. (1977). *Ray Method in Seismology.* Charles University Press, Praha.

Červený, V. and Pšenčík, I. (2005). Plane waves in viscoelastic anisotropic media - I. Theory. *Geophysical Journal International*, 161:197–212.

Chapman, C. H. and Coates, R. T. (1994). Generalized Born scattering in anisotropic media. *Wave Motion*, 19:309–341.

Chavent, G. (1974). Identification of parameter distributed systems. In Goodson, R. and Polis, M., editors, *Identification of function parameters in partial differential equations*, pages 31–48. American Society of Mechanical Engineers, New York.

Chavent, G. (2009). *Nonlinear least squares for inverse problems.* Springer Dordrecht Heidelberg London New York.

Chavent, G. and Jacewitz, C. A. (1995). Determination of background velocities by multiple migration fitting. *Geophysics*, 60(2):476–490.

Chen, K. (1984). Numerical modeling of elastic wave propagation in anisotropic inhomogeneous media: a finite element approach. In *Expanded Abstracts, 54^{th} Annual meeting*, page 631–632. Soc. Expl. Geophys.

Chew, W. C. and Liu, Q. H. (1996). Perfectly matched layers for elastodynamics: a new absorbing boundary condition. *Journal of Computational Acoustics*, 4:341–359.

Chung, W. and Shin, C. (2011). Elastic full waveform inversion using the time-stacked shot gathers. volume 30, pages 2809–2813. SEG.

Claerbout, J. (1971). Towards a unified theory of reflector mapping. *Geophysics*, 36:467–481.

Claerbout, J. F. (1976). *Fundamentals of Geophysical Data Processing.* McGraw-Hill Book Co.

Clough, R. J. (1960). The finite element method in plane stress analysis. In *ASCE Conf Electron Computat, Pittsburg, PA.*

Coates, R. T. and Chapman, C. H. (1990). Ray perturbation theory and the Born approximation. *Geophysical Journal International*, 100:379–392.

Cockburn, B. (2003). Discontinuous galerkin methods. *Journal of Applied Mathematics and Mechanics*, 83:731–754.

Cockburn, B., Karniadakis, G., and Shu, C.-W. (2000). The development of discontinuous galerkin methods. In B. Cockburn, G. K. and Shu, C.-W., editors, *Discontinuous Galerkin methodsM theory; computation and application*, Lecture notes in comutational Science and Engineering, 11. Springer.

de Hoop, M. V., Burridge, R., Spencer, C., and Miller, D. (1994). *GRT/AVA migration/inversion in anisotropic media*, volume 2301, pages 15–27. SPIE-The International Society for Optical Engineering.

de la Puente, J., Dumbser, M., Käser, M., and Igel, H. (2008). Discontinuous Galerkin methods for wave propagation in poroelastic media. *Geophysics*, 73(5):77–97.

Djikpéssé, H. A. and Tarantola, A. (1999). Multiparameter l_1 norm waveform fitting: Interpretation of gulf of mexico reflection seismograms. *Geophysics*, 64(4):1023–1035.

Dormy, E. and Tarantola, A. (1995). Numerical simulation of elastic wave propagation using a finite volume method. *Journal of Geophysical Research*, 100:2123–2133.

Duff, I. S. and Reid, J. K. (1983). The multifrontal solution of indefinite sparse symmetric linear systems. *ACM Transactions on Mathematical Software*, 9:302–325.

Duveneck, E. and Bakker, P. M. (2011). Stable P-wave modeling for reverse-time migration in tilted TI media. *Geophysics*, 76(2):S65–S75.

Duveneck, E., Milcik, P., Bakker, P. M., and Perkins, C. (2008). Acoustic VTI wave equations and their application for anisotropic reverse-time migration. *SEG Technical Program Expanded Abstracts*, 27(1):2186–2190.

Eaton, D. and Stewart, R. (1994). 2.5d elastic Ray-Born migration/inversion theory for transversely isotropic media. *Geophysical Journal of the Royal Astronomical Society*, 119:issue2, 667–683.

Emmerich, H. (1992). PSV-wave propagation in a medium with local heterogeneities: a hybrid formulation and its application. *Geophysical Journal International*, 109:54–64.

Etgen, J., Gray, S. H., and Zhang, Y. (2009). An overview of depth imaging in exploration geophysics. *Geophysics*, 74(6):WCA5–WCA17.

Etienne, V., Chaljub, E., Virieux, J., and Glinsky, N. (2010a). An hp-adaptive discontinuous Galerkin finite-element method for 3D elastic wave modelling. *Geophysical Journal International*, 183(2):941–962.

Etienne, V., Virieux, J., and Operto, S. (2010b). Seismic modelling with the discontinuous Galerkin finite-element method: application to complex 3D elastic media. In *Expanded Abstracts, 72th Annual EAGE Conference & Exhibition, Barcelona*. EAGE.

Farra, V. (1989). Ray perturbation theory for heterogeneous hexagonal anisotropic media. *Geophysical Journal International*, 99:723–737.

Fedorov, F. I. (1968). *Theory of elastic waves in crystals.* Plenum press, New York.

Festa, G. and Nielsen, S. (2003). PML absorbing boundaries. *Bulletin of the Seismological Society of America*, 93:891–903.

Fletcher, R. P., Du, X., and Fowler, P. J. (2009). Reverse time migration in tilted transversely isotropic TTI media. *Geophysics*, 74(6):WCA179–WCA187.

Forgues, E. and Lambaré, G. (1997). Parameterization study for acoustic and elastic ray+born inversion. *Journal of Seismic Exploration*, 6:253–278.

Foss, S.-K., de Hoop, M., and Ursin, B. (2005). Linearized 2.5-dimensional parameter imaging inversion in anisotropic elastic media. *Geophysical Journal International*, 161:722–738.

French, W. S. (1975). Computer migration of oblique seismic reflection profiles. *Geophysics*, 40:961–980.

Gardner, G. H. F., Gardner, L. W., and Gregory, A. R. (1974). Formation velocity and density–the diagnostic basics for stratigraphic traps. *Geophysics*, 39:770–780.

Gauthier, O., Virieux, J., and Tarantola, A. (1986). Two-dimensional nonlinear inversion of seismic waveforms: numerical results. *Geophysics*, 51(7):1387–1403.

Gelis, C., Virieux, J., Grandjean, G., Leparoux, D., and Operto, S. (2004). Full waveform elastic inversion in a space frequency domain formulation: a powerful geotechnical tool for superficial reconstruction. In *EOS Trans. AGU*, volume 85. American Geophysical Union.

Gholami, Y., Brossier, R., Operto, S., Prieux, V., and Virieux, J. (2012a). Which parametrization for acoustic VTI full waveform inversion? - part 2: application to Valhall. *Geophysics*, Submitted.

Gholami, Y., Brossier, R., Operto, S., Ribodetti, A., and Virieux, J. (2012b). Which parametrization for acoustic VTI full waveform inversion? - part 1: sensitivity and trade-off analysis. *Geophysics*, Submitted.

Grechka, V., Zhang, L., and III, J. W. R. (2004). Shear waves in acoustic anisotropic media. *Geophysics*, 69:576–582.

Gung, Y. and Romanowicz, B. (2004). Q tomography of the upper mantle using three component long-period waveforms. *Geophysical Journal International*, 157:813–830.

Hall, S. and Kendall, J. (2003). Fracture characterization at Valhall: Application of P-wave amplitude variation with offset and azimuth (AVOA) analysis to a 3D ocean-bottom data set. *Geophysics*, pages 1150–1160.

Hall, S., Kendall, J., and Barkved, O. (2002). Fractured reservoir characterization using P-wave AVOA analysis of 3D OBC data. *The Leading Edge*, pages 777–781.

Hansen, C. (1998). *Rank-defficient and discrete ill-posed problems - Numerical aspects of linear inversion.* Society for Industrial and Applied Mathematics - Mathematical modeling and computation.

Hanyga, A., Lenartowicz, E., and Pajchel, J. (1984). *Seismic Wave Propagation in the Earth.* (Elsevier), Amsterdam.

Hastings, F. D., Schneider, J. B., and Broschat, S. L. (1996). Application of the perfectly matched layer (PML) absorbing boundary condition to elastic wave propagation. *Journal of Acoustical Society of America*, 100:3061–3069.

Hatchell, P., Kawar, R., and Savitski, A. (2005). Integrating 4d seismic, geomechanics and reservoir simulation in the valhall oil field. In *Expanded Abstracts*.

Helbig, K. (1994). *Foundations of anisotropy for exploration seismic.* Pergamon.

Hudson, J. A. and Heritage, J. R. (1981). *Use of the Born approximation in seismic scattering problems.* Cambridge University Press.

Hughes, T. J. R. (2003). *The finite element method: linear static and dynamic finite element analysis.* Dover Publications Inc.

Ji, Y. and Singh, S. (2005). Anisotropy from full waveform inversion of multicomponent seismic data using a hyn. e. brid optimization method. *Geophysical Prospecting*, 53:435–445.

Kelly, K., Ward, R., Treitel, S., and Alford, R. (1976). Synthetic seismograms - a finite-difference approach. *Geophysics*, 41:2–27.

Kiyashchenko, D., B., K., and Plessix, R.-E. (2004). Anisotropic migration weight for amplitude-preserving migration and sensitivity analysis. *Geophysical Journal International*, 157:753–763.

Komatitsch, D., Barnes, C., and Tromp, J. (2000). Simulation of anisotropic wave propagation based upon a spectral element method. *Geophysics*, 65(4):1251–1260.

Kommedal, J. H., Barkved, O. I., and Howe, D. J. (2004). Initial experience operating a permanent 4C seabed array for reservoir monitoring at Valhall. *SEG Technical Program Expanded Abstracts*, 23(1):2239–2242.

Kommedal, J. H., Barkved, O. I., and Thomsen, L. A. (1997). Acquisition of 4 component obs data - a case study from the valhall field. *Presented at the 59th EAGE conference & Exhibition, Geneva, EAGE.*

Kosloff, D. and Baysal, E. (1982). Forward modeling by a Fourier method. *Geophysics*, 47:1402–1412.

Kuzuoglu, M. and Mittra, R. (1996). Frequency dependence of the constitutive parameters of causal perfectly matched anisotropic absorbers. *IEEE Microwave and Guided Wave Letters*, 6:447–449.

Lailly, P. (1983). The seismic problem as a sequence of before-stack migrations. In Bednar, J., editor, *Conference on Inverse Scattering: Theory and Applications.* SIAM, Philadelphia.

Lee, H.-Y., Koo, J. M., Min, D.-J., Kwon, B.-D., and Yoo, H. S. (2010). Frequency-domain elastic full-waveform inversion for VTI media. *Geophysical Journal International*, 183(2):884–904.

Leonard, R. and Munns, J. (1987). *Valhall Field in Geology of Norwegian Oil and Gas Fields.* Graham and Trotman.

Liu, W., Bube, K., Zhang, L., and Nihei, K. (2009). Stable reverse-time migration in variable-tilt TI media. In *Expanded Abstracts, 71^{th} Annual meeting.* EAGE.

Loris, I., Nolet, G., Daubechies, I., and Dahlen, F. A. (2007). Tomographic inversion using $l1$-norm regularization of wavelet coefficients. *Geophysical Journal International*, 170(1):359–370.

Luo, Y. and Schuster, G. T. (1990). Parsimonious staggered grid finite-differencing of the wave equation. *Geophysical Research Letters*, 17(2):155–158.

Lysmer, J. and Drake, L. A. (1972). A finite element method for seismology. In *Methods in computational physics*, volume 11. Academic Press, New Yort, USA.

Malinowski, M. and Operto, S. (2008). Quantitative imaging of the Permo-Mesozoic complex and its basement by frequency domain waveform tomography of wide-aperture seismic data from the Polish basin. *Geophysical Prospecting*, 56:805–825.

Manukyan, E., Latzel, S., Maurer, H., Marelli, S., and Greenhalgh, S. (2012). Exploitation of data-information content in elastic-waveform inversions. *Geophysics*, 77:105–115.

Marfurt, K. (1984). Accuracy of finite-difference anf finite-element modeling of the scalar and elastic wave equations. *Geophysics*, 49:533–549.

McMechan, G. (1983). Migration by extrapolation of time-dependent boundary values. *Geophysical Prospecting*, 31:413–420.

McMechan, G. (1989). A review of seismic acoustic imaging by reverse time migration. *International Journal of Imaging Systems and Technology*, 1:18–21.

McMechan, G. A. and Fuis, G. S. (1987). Ray equation migration of wide-angle reflections from southern alaska. *Journal of Geophysical Research*, 92(1):407–420.

Menke, W. (1984). *Geophysical Data Analysis: Discrete Inverse Theory*. Academic Press, Inc., Orlando, USA.

Miller, D., Oristaglio, M., and Beylkin, G. (1987). A new slant on seismic imaging: Migration and integral geometry. *Geophysics*, 52(7):943–964.

Moczo, P. (1989). Finite-difference technique for SH waves in 2D media using irregular grids: application to the seismic response problem. *Geophysical Journal International*, 99:321–329.

Moczo, P., Kristek, J., and Bystricky, E. (2001). Efficiency and optimization of the 3D finite-difference modeling of seismic ground motion. *Journal of Computational Acoustics*, 9:593–609.

Moczo, P., Kristek, J., Galis, M., Chaljub, E., and Etienne, V. (2010a). Comparison of Accuracy of the FDM, FEM, SEM and DGM. In *Projet QUEST*. Workshop à Capo Caccia, Italie.

Moczo, P., Kristek, J., Galis, M., and Pazak, P. (2010b). On accuracy of finite-difference and finite-element schemes with respect to P-wave and S-wave speed ratio. *Geophysical Journal International*, 182:493–510.

Moczo, P., Kristek, J., Galis, M., Pazak, P., and Balazovjech, M. (2007). The finite-difference and finite-element modeling of seismic wave propagation and earthquake motion. *Acta Physica Slovaca*, 52(2):177–406.

Montelli, R., Nolet, G., Dahlen, F. A., Masters, G., Engdahl, E. R., and Hung, S. H. (2004). Finite-frequency tomography reveals a variety of plumes in the mantle. *Science*, 303:338–343.

Mora, P. R. (1987). Nonlinear two-dimensional elastic inversion of multi-offset seismic data. *Geophysics*, 52:1211–1228.

Mora, P. R. (1989). Inversion = migration + tomography. *Geophysics*, 54(12):1575–1586.

Muller, G. (1985). The reflectivity method: a tutorial. *Journal of Geophysics*, 58:153–174.

MUMPS-team (2009). *MUMPS - MUltifrontal Massively Parallel Solver users' guide - version 4.9.2 (November 5, 2009).* ENSEEIHT-ENS Lyon, http://www.enseeiht.fr/apo/MUMPS/ or http://graal.ens-lyon.fr/MUMPS.

Munns, J. W. (1985). The Valhall field: a geological overview. *Marine and Petroleum Geology*, 2:23–43.

Nocedal, J. (1980). Updating Quasi-Newton Matrices With Limited Storage. *Mathematics of Computation*, 35(151):773–782.

Nocedal, J. and Wright, S. J. (1999). *Numerical Optimization.* New York, US : Springer.

Nocedal, J. and Wright, S. J. (2006). *Numerical Optimization.* Springer, 2nd edition.

Nye, J. (1985). *Physical properties of Crystals.* Oxford University Press, London.

O'Brien, M., Whitmore, N., Brandsberg-Dahl, S., Edgen, J., and Murphy, G. (1999). Multicomponent modelling of the Valhall field. In *Expanded Abstracts.* EAGE.

Olofsson, B., Probert, T., Kommedal, J., and Barkved, O. (2003). Azimuthal anisotropy from the valhall 4C 3D survey. *The Leading Edge*, pages 1228–1235.

Operto, S. and Virieux, J. (2006). SEISCOPE consortium: seismic imaging of complex structures from multicomponent global offset data by full waveform inversion. Technical report, UMR Géosciences Azur - CNRS. Consortium SEISCOPE.

Operto, S., Virieux, J., Ribodetti, A., and Anderson, J. E. (2009). Finite-difference frequency-domain modeling of visco-acoustic wave propagation in two-dimensional TTI media. *Geophysics*, 74 (5):T75–T95.

Pao, Y.-H. and Varatharajulu, V. (1976). Huygen's principle, radiation conditions, and integral formulas for the scattering of elastic waves. *Journal of Acoustical Society of America*, 59(6):1361–1371.

Plessix, R. E. (2006). A review of the adjoint-state method for computing the gradient of a functional with geophysical applications. *Geophysical Journal International*, 167(2):495–503.

Plessix, R. E. (2009). Three-dimensional frequency-domain full-waveform inversion with an iterative solver. *Geophysics*, 74(6):WCC53–WCC61.

Plessix, R.-E., Baeten, G., de Maag, J. W., and ten Kroode, F. (2012). Full waveform inversion and distance separated simultaneous sweeping: a study with a land seismic data set. *Geophysical Prospecting*, 60:733 – 747.

Plessix, R. E. and Bork, J. (2000). Quantitative estimate of vti parameters from ava responses. *Geophysical Prospecting*, 48(1):87–108.

Plessix, R. E. and Cao, Q. (2011a). A parametrization study for surface seismic full waveform inversion in an acoustic vertical transversely isotropic medium. *Geophysical Journal International*, 185:539–556.

Plessix, R. E. and Cao, Q. (2011b). A parametrization study for surface seismic full waveform inversion in an acoustic vertical transversely isotropic medium. *Geophysical Journal International*.

Plessix, R. E. and Perkins, C. (2010). Full waveform inversion of a deep water ocean bottom seismometer dataset. *First Break*, 28:71–78.

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

Polak, E. and Ribière, G. (1969). Note sur la convergence de méthodes de directions conjuguées. *Revue Française d'Informatique et de Recherche Opérationnelle*, 16:35–43.

Pratt, R. G. (1990). Frequency-domain elastic modeling by finite differences: a tool for crosshole seismic imaging. *Geophysics*, 55(5):626–632.

Pratt, R. G. (1999). Seismic waveform inversion in the frequency domain, part I : theory and verification in a physic scale model. *Geophysics*, 64:888–901.

Pratt, R. G. and Chapman, C. H. (1992). Traveltime tomography in anisotropic media - i. theory. *Geophysical Journal International*, 109:1–19.

Pratt, R. G., Plessix, R. E., and Mulder, W. A. (2001). Seismic waveform tomography: the effect of layering and anisotropy. In *Presented at the 63th EAGE Conference & Exhibition, Amsterdam, EAGE.*, page P092.

Pratt, R. G., Shin, C., and Hicks, G. J. (1998). Gauss-Newton and full Newton methods in frequency-space seismic waveform inversion. *Geophysical Journal International*, 133:341–362.

Pratt, R. G. and Shipp, R. M. (1999). Seismic waveform inversion in the frequency domain, part II: Fault delineation in sediments using crosshole data. *Geophysics*, 64:902–914.

Pratt, R. G. and Symes, W. (2002). Semblance and differential semblance optimisation for waveform tomography: a frequency domain implementation. In *Journal of Conference Abstracts*, volume 7(2), pages 183–184. Cambridge publications.

Pratt, R. G. and Worthington, M. H. (1990). Inverse theory applied to multi-source cross-hole tomography. Part I: acoustic wave-equation method. *Geophysical Prospecting*, 38:287–310.

Press, W. H., Teukolsky, S. A., Vetterling, W. T., and Flannery, B. P. (2007). *Numerical Recipes 3rd Edition: The Art of Scientific Computing.* Cambridge University Press, 3 edition.

Prieux, V. (2012). *Imagerie sismique des milieux visco-acoustiques et visco-élastiques à deux dimensions par stéréotomographie et inversion des formes d'ondes: applications au champ pétrolier de Valhall.* PhD thesis, Université de Nice-Sophia-Antipolis.

Prieux, V., Brossier, R., Gholami, Y., Operto, S., Virieux, J., Barkved, O., and Kommedal, J. (2011). On the footprint of anisotropy on isotropic full waveform inversion: the Valhall case study. *Geophysical Journal International*, 187:1495–1515.

Ravaut, C., Operto, S., Improta, L., Virieux, J., Herrero, A., and dell'Aversana, P. (2004). Multi-scale imaging of complex structures from multi-fold wide-aperture seismic data by frequency-domain full-wavefield inversions: application to a thrust belt. *Geophysical Journal International*, 159:1032–1056.

Reed, W. and Hill, T. (1973). Triangular mesh methods for the neuron transport equation. Technical Report LA-UR-73-479, Los Alamos Scientific Laboratory.

Remaki, M. (2000). A new finite volume scheme for solving Maxwell's system. *Computation and Mathematics in Electrical and Electronic Engineering*, 19(3):913–931.

Remaki, M. and Fézoui, L. (1998). Une méthode de galerkin discontinu pour la résolution des équations de maxwell en milieu hétérogène. Technical Report 3501, INRIA Sophia-Antipolis.

Ribodetti, A. and Virieux, J. (1996). Asymptotic theory for imaging the attenuation factors Q_p and Q_s. In *Inverse Problems of Wave Propagation and Diffraction, Proceedings, Aix-les-Bains, France 1996*, pages 334–353. Springer-Verlag.

Rivière, S. (2008). *Discontinuous Galerkin methods for solving elliptic and parabolic equations: theory and implementation.* SIAM.

Robertson, J. O. A., Holliger, K., Green, A. G., Pugin, A., and Iaco, R. D. (1996). Effects of near-surface waveguides on shallow high-resolution seismic refraction and reflection data. *Geophysical Research Letters*, 23(5):495–498.

Roden, J. A. and Gedney, S. D. (2000). Convolution PML (CPML): An efficient FDTD implementation of the CFS-PML for arbitrary media. *Microwave and Optical Technology Letters*, 27(5):334–339.

Roever, W. E., Vining, T. F., and Strick, E. (1959). Propagation of elastic wave motion from an impulsive source along a fluid/solid interface. Part II: Theoretical pressure response. part III: The pseudo-Rayleigh wave. *Philosophical Transactions of the Royal Society of London*, 251:465–523.

Rudin, L., Osher, S., and Fatemi, E. (1992). Nonlinear total variation based noise removal algorithms. *Physica D*, 60:259–268.

S. G. Nash, J. N. (1991). A numerical study of the limited memory bfgs method and truncated newton method for large scale optimization. *Siam Journal on Optimization*, 1:358–372.

Sands, D. (1982). *Vectors and Tensors in Crystallography.* Addison-Wesley.

Sato, H. (1984). Attenuation and envelope formation of three-component seismograms of small local earthquakes in randomly inhomogeneous lithosphere. *Journal Geophysical Research*, 89(B2):1221–1241.

Sato, H. and Fehler, M. C. (1997). *Seismic wave propagation and scattering in the heterogeneous Earth.* Springer.

Scales, J. A., Docherty, P., and Gersztenkorn, A. (1990). Regularization of nonlinear inverse problems: imaging the near-surface weathering layer. *Inverse Problems*, 6:115–131.

Schneider, W. A. (1978). Integral formulation for migration in two and three dimensions. *Geophysics*, 43:49–76.

Sears, T., Singh, S., and Barton, P. (2008). Elastic full waveform inversion of multi-component OBC seismic data. *Geophysical Prospecting*, 56(6):843–862.

Sears, T. J., Barton, P. J., and Singh, S. C. (2010). Elastic full waveform inversion of multicomponent ocean-bottom cable seismic data: Application to alba field, u. k. north sea. *Geophysics*, 75(6):R109–R119.

Shin, C. and Cha, Y. H. (2009). Waveform inversion in the Laplace-Fourier domain. *Geophysical Journal International*, 177:1067–1079.

Shin, C., Jang, S., and Min, D. J. (2001). Improved amplitude preservation for prestack depth migration by inverse scattering theory. *Geophysical Prospecting*, 49:592–606.

Shin, C., Min, D.-J., Marfurt, K. J., Lim, H. Y., Yang, D., Cha, Y., Ko, S., Yoon, K., Ha, T., and Hong, S. (2002). Traveltime and amplitude calculations using the damped wave solution. *Geophysics*, 67:1637–1647.

Shipp, R. M. and Singh, S. C. (2002). Two-dimensional full wavefield inversion of wide-aperture marine seismic streamer data. *Geophysical Journal International*, 151:325–344.

Sieminski, A., Trampert, J., and Tromp, J. (2009). Principal component analysis of anisotropic finite-frequency sensitivity kernels. *Geophysical Journal International*, 179(2):1186–1198.

Sirgue, L. (2003). *Inversion de la forme d'onde dans le domaine fréquentiel de données sismiques grand offset*. PhD thesis, Université Paris 11, France - Queen's University, Canada.

Sirgue, L. (2006). The importance of low frequency and large offset in waveform inversion. In *Presented at the 68th EAGE Conference & Exhibition, Vienna, EAGE*, page A037.

Sirgue, L., Barkved, O. I., Dellinger, J., Etgen, J., Albertin, U., and Kommedal, J. H. (2010). Full waveform inversion: the next leap forward in imaging at Valhall. *First Break*, 28:65–70.

Sirgue, L., Etgen, T. J., Albertin, U., and Brandsberg-Dahl, S. (2007). System and method for 3D frequency-domain waveform inversion based on 3D time-domain forward modeling. *US Patent Application Publication*, US2007/0282535 A1.

Sirgue, L. and Pratt, R. G. (2004). Efficient waveform inversion and imaging : a strategy for selecting temporal frequencies. *Geophysics*, 69(1):231–248.

Smith, T. M., Collis, S. S., Ober, C. C., Overfelt, J. R., and Schwaiger, H. F. (2010). Elastic Wave Propafation in Variable Media using a Discontinuous Galerkin Method. In *Expanded Abstracts, 80th Annual SEG Conference & Exhibition, Denver*. Society of Exploration Geophysics.

Sourbier, F., Operto, S., Virieux, J., Amestoy, P., and L'Excellent, J.-Y. (2009). FWT2D: A massively parallel program for frequency-domain full-waveform tomography of wide-aperture seismic data–Part 1: Algorithm. *Computers & Geosciences*, 35(3):487 – 495.

Strong, D. and Chan, T. (2003). Edge-preserving and scale-dependent properties of total variation regularization. *Inverse Problems*, 19:S165–S187.

Sun, Z., Brown, R. J., Lawton, D. C., and Wang, Z. (1991). Seismic anisotropy and salt detection: A physical modeling study. volume 10, pages 713–716. SEG.

Tago, J., Cruz-Atienza, V. M., Etienne, V., Virieux, J., Benjemaa, M., and Sanchez-Sesma, F. J. (2010). 3D dynamic rupture with anelastic wave propagation using an hp-adaptive Discontinuous Galerkin method. In *EOS Trans. AGU, abstract S51A-1915*. American Geophysical Union, San Francisco, USA.

Talagrand, O. and Courtier, P. (1987). Variational assimilation of meteorological observations with the adjoint vorticity equation. i: theory. *Quarterly Journal of the Royal Meteorological Society*, 113:1311–1328.

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

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

Tarantola, A. (1987). *Inverse problem theory: methods for data fitting and model parameter estimation*. Elsevier, New York.

Thierry, P., Operto, S., and Lambaré, G. (1999). Fast 2D ray-Born inversion/migration in complex media. *Geophysics*, 64(1):162–181.

Thomsen, L. A. (1986). Weak elastic anisotropy. *Geophysics*, 51:1954–1966.

Thomsen, L. A., Barkved, O., Haggard, B., Kommedal, J. H., and Rosland, B. (1997). Converted wave imaging of valhall reservoir. *Presented at the 59th EAGE conference & Exhibition, Geneva, EAGE.*

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

Tikhonov, A. N. (1963). Resolution of ill-posed problems and the regularization method (in russian, french translation, mir, moscow, 1976). *Dokl. Akad. Nauk SSSR*, 151:501–504.

Tromp, J., Tape, C., and Liu, Q. (2005). Seismic tomography, adjoint methods, time reversal and banana-doughnut kernels. *Geophysical Journal International*, 160:195–216.

Tsvankin, I. (1995). Normal moveout from dipping reflectors in anisotropic media. *Geophysics*, 60:268–284.

Tsvankin, I. (2001a). Seismic signature and analysis of reflection data in anisotropic media. In Helbig, K. and Treitel, S., editors, *Seismic Exploration*, volume 29 of *Handbook of Geophysical Exploration*. Elsevier, 2 edition.

Tsvankin, I. (2001b). Seismic signature and analysis of reflection data in anisotropic media. In Helbig, K. and Treitel, S., editors, *Seismic Exploration*, volume 29 of *Handbook of Geophysical Exploration*. Pergamon.

van den Berg, P., Abubakar, A., and Fokkema, J. (2003). Multiplicative regularization for contrast profile inversion. *Radio Science*, 38:23.1–23.10.

van den Berg, P. M. and Abubakar, A. (2001). Contrast source inversion method: state of the art. *Progress in Electromagnetics research*, 34:189–218.

van Gestel, J., Kommedal, J., Barkved, O., Mundal, I., Bakke, R., and Best, K. (2008). Continuous seismic surveillance of Valhall field. *The Leading Edge*, pages 1616–1621.

Vanorio, T. and Virieux, J. (2005). Three-dimensional seismic tomography from p wave and s wave microearthquake travel times and rock physics characterization of the campi flegrei caldera. *Journal of Geophysical Research*, 110.

Vigh, D., Starr, B., Kapoor, J., and Li, H. (2010). 3d full waveform inversion on a gulf of mexico waz data set. *SEG Technical Program Expanded Abstracts*, 29(1):957–961.

Virieux, J. (1986). P-SV wave propagation in heterogeneous media: velocity stress finite difference method. *Geophysics*, 51:889–901.

Virieux, J. and Operto, S. (2009). An overview of full waveform inversion in exploration geophysics. *Geophysics*, 74(6):WCC127–WCC152.

Wang, Y. and Rao, Y. (2009). Reflection seismic waveform tomography. *Journal of Geophysical Research*, 114(B03304):doi:10.1029/2008JB005916.

Weglein, A. B. and Gray, S. H. (1983). The sensitivity of Born inversion to the choice of reference velocity-a simple example. *Geophysics*, 48:36–38.

Weglein, A. B., Violette, P. B., and Keho, T. H. (1986). Using a constant background multi-parameter Born theory to obtain exact inversion goals. *Geophysics*, 51.

Wenk, S., Käser, M., and Igel, H. (2010). The discontinuous Galerkin finite element method and its application to global wave propagation. In *EOS Trans. AGU, abstract S53F-02*. American Geophysical Union, San Francisco, USA.

Whitmore, N. (1983). Iterative depth migration by backward time propagation. In *53th Annual Meeting*. SEG.

Winterstein, D. F. and Meadows, M. A. (1991). Shear-wave polarizations and subsurface stress directions at Lost Hills field. *Geophysics*, 56(9):1331–1348.

Woodhouse, J. (2007). Normal modes and surface waves: theory. In Romanovitz, B. and Diewonski, A., editors, *Treatise of Geophysics, volume 1: Seismology and structure of the Earth, Treatise of Geophysics*, chapter 3, pages 67–125. Elsevier.

Woodhouse, J. and Dziewonski, A. (1984). Mapping the upper mantle: Three dimensional modelling of earth structure by inversion of seismic waveforms. *Journal of Geophysical Research*, 89:5953–5986.

Wu, R. S. (1989). The perturbation method in elastic wave scattering. *Pure and Applied Geophysics*, 131(4):605–638.

Wu, R.-S. and Aki, K. (1985a). Elastic wave scattering by a random medium and the small scale inhomogeneities in the lithosphere. *Journal Geophysical Research*, 90:10261–10237.

Wu, R. S. and Aki, K. (1985b). Scattering characteristics of elastic waves by an elastic heterogeneity. *Geophysics*, 50(4):582–595.

Wu, R. S. and Toksöz, M. N. (1987). Diffraction tomography and multisource holography applied to seismic imaging. *Geophysics*, 52:11–25.

Zhang, H. J. and Zhang, Y. (2008). Reverse time migration in 3D heterogeneous TTI media. *SEG Technical Program Expanded Abstracts*, 27(1):2196–2200.

Zhang, L., III, J. W. R., and Hoversten, G. M. (2003). An acoustic wave equation for modeling in tilted ti media. *SEG Technical Program Expanded Abstracts*, 22(1):153–156.

Zhang, L., Rector III, J., and Hoversten, G. (2005). Finite-difference modelling of wave propagation in acoustic tilted ti media. *Geophysical Prospecting*, 53:843–852.

Zhang, Y., Zhang, H., and Zhang, G. (2011). A stable TTI reverse time migration and its implementation. *Geophysics*, 76(3):WA3–WA11.

Zhang Y., Z. H. (2011). A stable TTI reverse time migration. In *Expanded Abstracts*.

Zhou, B. and Greenhalgh, S. A. (2009). On the computation of the fréchet derivatives for seismic waveform inversion in 3d general anisotropic, heterogeneous media. *Geophysics*, 74:WB153–WB163.

Zhou, H., Zhang, G., and Bloor, R. (2006). An anisotropic acoustic wave equation for modeling and migration in 2D TTI media. In *Expanded Abstracts - New Orleans*, volume 25, pages 194–198. Society of Exploration Geophysics.

Zhou H., Zhang G., B. R. (2006). An anisotropic acoustic wave equation for VTI media. In *Expanded Abstracts, 68th Annual meeting*. EAGE.

Zhu, C., Byrd, R. H., and Nocedal, J. (1997). L-bfgs-b: Algorithm 778: L-bfgs-b, fortran routines for large scale bound constrained optimization. *ACM Transactions on Mathematical Software*, 23(4):550–560.

Zhu, J. and Dorman, J. (2000). Two-dimensional, three-component wave propagation in a transversely isotropic medium with arbitrary-orientation—finite-element modeling. *Geophysics*, 65(3):934–942.

Zienkiewicz, O. and Taylor, R. (1967). *The Finite Element Method for Solid and Structural Mechanics*. McGraw Hill, New York.

Zienkiewicz, O. C., Taylor, R. L., and Zhu, J. Z. (2005). *The Finite Element Method: Its Basis and Fundamentals*. Elsevier, London. 6th edition.

Appendix A

The 2D TTI wave equation

Contents

A.1 Two dimensional P-SV waves equations in TTI media

The generalized Hook's law gives the constitutive relation between stress and strain, for anisotropic and elastic solid media (see chapter 1.1.1). A general anisotropic medium with symmetry axis oriented arbitrarly (tilted symmetry of axis), is called Tilted Transverse Isotropic (TTI) model. For such a model, the calculation of the anisotropic response to material tilt for elastic seismic wave propagation requires rotation of the stiffness tensors describing the crustal material. In other words, to propagate waves within an anisotropic material possessing internal tilt, the stiffness tensor within the elastic wave equation needs to be oriented via tensor rotation (Sands, 1982; Nye, 1985).

A.1.1 Rotation of the matrix of elastic constants

To discuss the wave motion in the observation system x,y,z, we must rotate the system x',y',z' (corresponding to the orientation of the general TI medium, with $x' - y'$ in the isotropy plane and z' the symmetry axis) to the system x,y,z with tilted symmetry axis (figure A.1).

Zhu and Dorman (2000) showed that the stiffness coefficient matrix for rotated general TI medium is rewritten as :

$$D = \begin{bmatrix} d_{11} & d_{12} & d_{13} & d_{14} & d_{15} & d_{16} \\ d_{21} & d_{22} & d_{23} & d_{24} & d_{25} & d_{26} \\ d_{31} & d_{32} & d_{33} & d_{34} & d_{35} & d_{36} \\ d_{41} & d_{42} & d_{43} & d_{44} & d_{45} & d_{46} \\ d_{51} & d_{52} & d_{53} & d_{54} & d_{55} & d_{56} \\ d_{61} & d_{62} & d_{63} & d_{64} & d_{65} & d_{66} \end{bmatrix} . \tag{A.1}$$

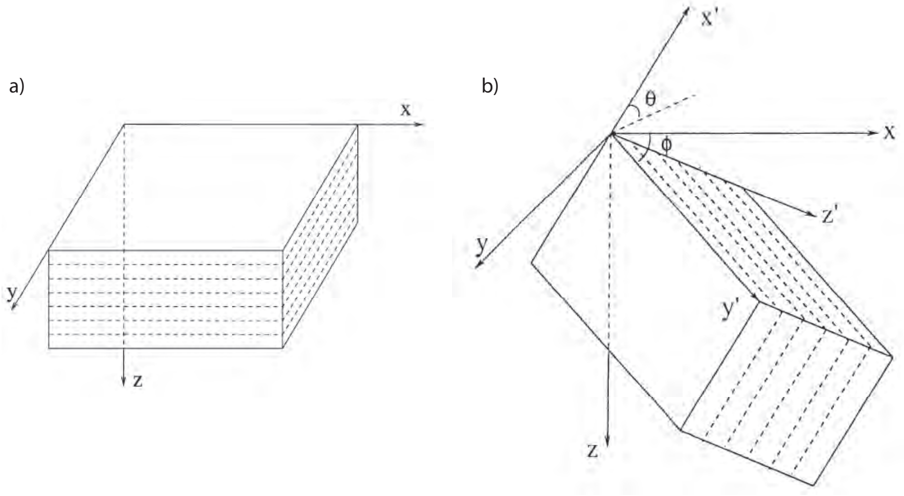


Figure A.1: (a) The TI medium with vertical symmetry axis (z-direction), and (b) the general TI medium with arbitrary rotated symmetry axis, extracted from Zhu and Dorman (2000, figures 1 and 2). The observation coordinate system are the axes x, y, z, with $x - y$ in the horizontal plane and z in the vertical (depth) direction. The axes x', y', z' correspond to a coordinate system where z' is the symmetry axis of the TI medium and $x' - y'$ is in the isotropy plane, with y' in the strike direction of the isotropy plane. The orientation of the general TI medium is defined by the dip (i.e. tilted) angle θ and strike angle ϕ of the isotropy plane. The dip angle is the angle between the $x'-$axis and the horizontal plane. The strike angle is measured from x to y' in the clockwise direction.

This matrix is also a 6×6 symmetric matrix with all the elements generally not equal to zero. The value of entries of the matrix **D** are the elastic constants for the rotated general TI medium in the observation system. In the special case when the axis of symmetry of the TI medium is in the $x - z$ plane ($\phi = 90^0$, see figure A.1b), then the matrix **D** reduces to (Zhu and Dorman, 2000) :

$$D = \begin{bmatrix} d_{11} & d_{12} & d_{13} & 0 & d_{15} & 0 \\ d_{21} & d_{22} & d_{23} & 0 & d_{25} & 0 \\ d_{31} & d_{22} & d_{33} & 0 & d_{35} & 0 \\ 0 & 0 & 0 & d_{44} & 0 & d_{46} \\ d_{51} & d_{22} & d_{53} & 0 & d_{55} & 0 \\ 0 & 0 & 0 & d_{64} & 0 & d_{66} \end{bmatrix} \tag{A.2}$$

A.1.2 Two dimensional wave equation for TTI media

By rotation between the above two coordinate systems, the matrix of elastic constants in the observation system is calculated for a tilted TI medium. Then, the wave propagation is modeled in the observation system, assuming that the medium, the source, and the wavefields are two dimensional, varying only in the $x - z$ domain. All partial derivatives with respect to y are equal to zero, but since $u_y \neq 0$, we calculate the three components of wave motion. Therefore, from the strain-displacement relationship we have:

$$\begin{gathered}\epsilon_{xx} = \frac{\partial u_x}{\partial x}, \quad \epsilon_{yy} = \frac{\partial u_y}{\partial y} = 0, \quad \epsilon_{zz} = \frac{\partial u_z}{\partial z} \\ 2\epsilon_{yz} = \frac{\partial u_y}{\partial z}, \quad 2\epsilon_{zx} = \frac{\partial u_x}{\partial z} + \frac{\partial u_z}{\partial x}, \quad 2\epsilon_{xy} = \frac{\partial u_y}{\partial x}, \end{gathered} \tag{A.3}$$

where u_x, u_y, u_z are the displacements along x, y, z respectively.
If the symmetry axis of the TI medium is in the $x - z$ plane or the strike direction of the isotropy plane is perpendicular to the $x - axis$ $(\phi = 90^0)$, then :

$$\begin{aligned} \sigma_{xx} &= d_{11}\epsilon_{xx} + d_{13}\epsilon_{zz} + 2d_{15}\epsilon_{zx}, \\ \sigma_{zz} &= d_{31}\epsilon_{xx} + d_{33}\epsilon_{zz} + 2d_{35}\epsilon_{zx}, \\ \sigma_{xz} &= d_{51}\epsilon_{xx} + d_{53}\epsilon_{zz} + 2d_{55}\epsilon_{zx}. \end{aligned} \tag{A.4}$$

Inserting the equation A.3 into equation A.4 and taking the temporal derivative, we have :

$$\begin{aligned} \frac{\partial \sigma_{xx}}{\partial t} &= d_{11}\frac{\partial v_x}{\partial x} + d_{13}\frac{\partial v_z}{\partial z} + d_{15}\left\{\frac{\partial v_x}{\partial z} + \frac{\partial v_z}{\partial x}\right\}, \\ \frac{\partial \sigma_{zz}}{\partial t} &= d_{31}\frac{\partial v_x}{\partial x} + d_{33}\frac{\partial v_z}{\partial z} + d_{35}\left\{\frac{\partial v_x}{\partial z} + \frac{\partial v_z}{\partial x}\right\}, \\ \frac{\partial \sigma_{xz}}{\partial t} &= d_{51}\frac{\partial v_x}{\partial x} + d_{53}\frac{\partial v_z}{\partial z} + d_{55}\left\{\frac{\partial v_x}{\partial z} + \frac{\partial v_z}{\partial x}\right\}. \end{aligned} \tag{A.5}$$

The equation A.5 in frequency domain is :

$$\begin{aligned} -\iota\omega\sigma_{xx} &= d_{11}\frac{\partial v_x}{\partial x} + d_{13}\frac{\partial v_z}{\partial z} + d_{15}\left\{\frac{\partial v_x}{\partial z} + \frac{\partial v_z}{\partial x}\right\}, \\ -\iota\omega\sigma_{zz} &= d_{31}\frac{\partial v_x}{\partial x} + d_{33}\frac{\partial v_z}{\partial z} + d_{35}\left\{\frac{\partial v_x}{\partial z} + \frac{\partial v_z}{\partial x}\right\}, \\ -\iota\omega\sigma_{xz} &= d_{51}\frac{\partial v_x}{\partial x} + d_{53}\frac{\partial v_z}{\partial z} + d_{55}\left\{\frac{\partial v_x}{\partial z} + \frac{\partial v_z}{\partial x}\right\}. \end{aligned} \tag{A.6}$$

From the general equations of motion,

$$\rho\frac{\partial^2 u_i}{\partial t^2} = \frac{\partial \sigma_{ij}}{\partial x_j}, \tag{A.7}$$

we have :

$$\begin{aligned} \rho\frac{\partial^2 u_x}{\partial t^2} &= \frac{\partial \sigma_{xx}}{\partial x} + \frac{\partial \sigma_{xz}}{\partial z} \\ \rho\frac{\partial^2 u_z}{\partial t^2} &= \frac{\partial \sigma_{zx}}{\partial x} + \frac{\partial \sigma_{zz}}{\partial z}, \end{aligned} \tag{A.8}$$

which can be rewritten as :

$$\begin{aligned} \rho\frac{\partial v_x}{\partial t} &= \frac{\partial \sigma_{xx}}{\partial x} + \frac{\partial \sigma_{xz}}{\partial z}, \\ \rho\frac{\partial v_z}{\partial t} &= \frac{\partial \sigma_{xz}}{\partial x} + \frac{\partial \sigma_{zz}}{\partial z}, \end{aligned} \tag{A.9}$$

where v_x and v_z are the horizontal and vertical velocities. In frequency domain we obtain :

$$\begin{aligned}
-\iota\omega v_x &= \frac{\partial\sigma_{xx}}{\partial x} + \frac{\partial\sigma_{xz}}{\partial z}, \\
-\iota\omega v_z &= \frac{\partial\sigma_{xz}}{\partial x} + \frac{\partial\sigma_{zz}}{\partial z}.
\end{aligned} \tag{A.10}$$

The first-order two dimensional frequency domain P-SV wave equation for tilted TI media is obtained by summing up the equations A.6 and A.10 :

$$\begin{aligned}
-\iota\omega v_x &= \frac{1}{\rho}\left\{\frac{\partial\sigma_{xx}}{\partial x} + \frac{\partial\sigma_{xz}}{\partial z}\right\} + f_x, \\
-\iota\omega v_z &= \frac{1}{\rho}\left\{\frac{\partial\sigma_{xz}}{\partial x} + \frac{\partial\sigma_{zz}}{\partial z}\right\} + f_z, \\
-\iota\omega\sigma_{xx} &= d_{11}\frac{\partial v_x}{\partial x} + d_{13}\frac{\partial v_z}{\partial z} + d_{15}\left\{\frac{\partial v_x}{\partial z} + \frac{\partial v_z}{\partial x}\right\} - \iota\omega\sigma_{xx_0}, \\
-\iota\omega\sigma_{zz} &= d_{31}\frac{\partial v_x}{\partial x} + d_{33}\frac{\partial v_z}{\partial z} + d_{35}\left\{\frac{\partial v_x}{\partial z} + \frac{\partial v_z}{\partial x}\right\} - \iota\omega\sigma_{zz_0}, \\
-\iota\omega\sigma_{xz} &= d_{51}\frac{\partial v_x}{\partial x} + d_{53}\frac{\partial v_z}{\partial z} + d_{55}\left\{\frac{\partial v_x}{\partial z} + \frac{\partial v_z}{\partial x}\right\} - \iota\omega\sigma_{xz_0},
\end{aligned} \tag{A.11}$$

where $(\sigma_{xx_0}, \sigma_{zz_0}, \sigma_{xz_0})$ are source terms, and :

$$\begin{aligned}
d_{11} &= \cos^2\theta(c_{11}\cos^2\theta + c_{13}\sin^2\theta) + \sin^2\theta(c_{13}\cos^2\theta + c_{33}\sin^2\theta) + 4c_{44}\cos^2\theta\sin^2\theta, \\
d_{13} &= c_{11}\cos^2\theta\sin^2\theta + c_{13}\sin^4\theta + c_{13}\cos^4\theta + c_{33}\sin^2\theta\cos^2\theta - 4c_{44}\sin^2\theta\cos^2\theta, \\
d_{15} &= \cos^2\theta(c_{13} - c_{11})\cos\theta\sin\theta + \sin^2\theta(c_{33} - c_{13})\cos\theta\sin\theta + 2\cos\theta\sin\theta c_{44}(\cos^2\theta - \sin^2\theta), \\
d_{33} &= \sin^2\theta(c_{11}\sin^2\theta + c_{13}\cos^2\theta) + \cos^2\theta(c_{33}\cos^2\theta + c_{13}\sin^2\theta) + 4c_{44}\cos^2\theta\sin^2\theta, \\
d_{35} &= \sin^2\theta(c_{13} - c_{11})\cos\theta\sin\theta + \cos^2\theta(c_{33} - c_{13})\cos\theta\sin\theta - 2\cos\theta\sin\theta c_{44}(\cos^2\theta - \sin^2\theta), \\
d_{55} &= c_{44}(1 - 2\sin^2\theta)^2 + (c_{33} - c_{13})\cos^2\theta\sin^2\theta - (c_{13} - c_{11})\cos^2\theta\sin^2\theta.
\end{aligned} \tag{A.12}$$

Note that $d_{13} = d_{31}$, $d_{15} = d_{51}$ and $d_{35} = d_{53}$ for symmetry reason.
The development of formulation for integration over a surface in 2D TTI can be applied same as one applied for VTI wave equation :

$$\begin{aligned}
-\iota\omega\rho v_x &= \frac{\partial(T_1 + T_2)}{\partial x} + \frac{\partial(T_3)}{\partial z} + \rho f_x, \\
-\iota\omega\rho v_z &= \frac{\partial T_3}{\partial x} + \frac{\partial(T_1 - T_2)}{\partial z} + \rho f_z, \\
-\iota\omega T_1 &= \frac{d_{11} + d_{13}}{2}\frac{\partial v_x}{\partial x} + \frac{d_{13} + d_{33}}{2}\frac{\partial v_z}{\partial z} + \frac{d_{15} + d_{35}}{2}\frac{\partial v_x}{\partial z} + \frac{d_{15} + d_{35}}{2}s_x\frac{\partial v_z}{\partial x} - \iota\omega T_1^0, \\
-\iota\omega T_2 &= \frac{d_{11} - d_{13}}{2}\frac{\partial v_x}{\partial x} + \frac{d_{13} - d_{33}}{2}\frac{\partial v_z}{\partial z} + \frac{d_{15} - d_{35}}{2}\frac{\partial v_x}{\partial z} + \frac{d_{15} - d_{35}}{2}\frac{\partial v_z}{\partial x} - \iota\omega T_2^0, \\
-\iota\omega T_3 &= d_{55}\frac{\partial v_z}{\partial x} + d_{55}\frac{\partial v_x}{\partial z} + d_{51}\frac{\partial v_x}{\partial x} + d_{53}\frac{\partial v_z}{\partial z} - \iota\omega T_3^0,
\end{aligned} \tag{A.13}$$

where $\vec{T}^t = (T_1, T_2, T_3) = ((\sigma_{xx} + \sigma_{zz})/2, (\sigma_{xx} - \sigma_{zz})/2, \sigma_{xz})$.