



HAL
open science

Velocity model building by full waveform inversion of early arrivals & reflections and case study with gas cloud effect

Wei Zhou

► **To cite this version:**

Wei Zhou. Velocity model building by full waveform inversion of early arrivals & reflections and case study with gas cloud effect. Earth Sciences. Université Grenoble Alpes, 2016. English. NNT : 2016GREAU024 . tel-01556584

HAL Id: tel-01556584

<https://theses.hal.science/tel-01556584>

Submitted on 5 Jul 2017

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

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

THÈSE

Pour obtenir le grade de

DOCTEUR DE L'UNIVERSITÉ DE GRENOBLE

Spécialité : **Sciences de la Terre, de l'Univers et de l'Environnement**

Arrêté ministériel : 7 Août 2006

Présentée par

Wei ZHOU

Thèse dirigée par **Romain Brossier, Stéphane Operto et Jean Virieux**

préparée au sein de l'**Institut des Sciences de la Terre**
et de l'**École Doctorale Terre Univers Environnement**

Velocity Model Building by Full Waveform Inversion of Early Arrivals & Reflections and Case Study with Gas Cloud Effect

30 Septembre 2016 ,
devant le jury composé de :

Hervé Chauris (Président du jury)

Professeur à MinesParisTech, Fontainebleau, Rapporteur

René-Édouard Plessix

Chercheur Principal (HDR) à Shell, Pays-bas, Rapporteur

Gilles Lambaré

Directeur de Recherche à CGG, Massy, Examineur

François Audebert

Coordinateur R&D à Total, Pau, Examineur

Romain Brossier

Maître de conférences à l'Université Grenoble-Alpes, Directeur de thèse

Stéphane Operto

Directeur de Recherche au CNRS, GéoAzur, Université de Nice-Sophia Antipolis,
Valbonne, Directeur de thèse

Jean Virieux

Professeur à l'Université Grenoble-Alpes, Directeur de thèse



Acknowledgement

“Boom, boom, booom!” The artificial firework has lighted the night of Grenoble. Like every year the ceremony of the French national day ends with its climax, and this is the third time that I enjoy this kind of fire show. It is amazing that I have been living in this foreign city, made new friends, tasted exotic food, enjoyed a different culture and more importantly accomplished a PhD project with the help of my wonderful supervisors.

I am very grateful to Pr. Jean Virieux for giving me opportunities to do an internship and the following PhD in the frame of SEISCOPE. Jean is an erudite scholar and he always likes to share his knowledge with others. I enjoy every discussion with him and have benefited a lot from that. (I still remember our talk related to Fresnel and Fraunhofer optics.) Giving much advice after my presentations and writing numerous comments and annotations in my manuscripts, I can feel his serious attitude and also kind heart to me. I always appreciate his every encouragement.

Also, I appreciate very much the instructions of Stéphane Operto, who is rigorous in my results and writings, but also very patient to explain the ideas for me. Because we use emails to contact, sometimes it is not easy for him to follow the discussions. Nevertheless, he keeps taking care of my advances and asking if I have any difficulties. I feel very sorry to him that I did not send him emails so frequently as he asked. Many thanks to him for providing me the real data set, and more importantly his FWI result such that I can make the assessment of my result.

Special thankfulness is devoted to Romain Brossier, for his useful instructions and inspirations each time I am obstructed by misunderstandings and difficulties. As a supervisor, he is more like a good friend of mine. I can recall the moment that we looked for beers on the streets of Denver, and the relaxing time but became insensitive for me in Amsterdam. I am very grateful to him for helping me contact Total, manage plane tickets, hold the committee of this thesis and improve the manuscript when he is also occupied by running the consortium.

I want to express my sincere gratitude to the reviewers: Hervé Chauris (MinesParis-Tech) and René-Édouard Plessix (Shell), as well as other members of the committee: Gilles Lambaré (CGG) and François Audebert (Total). Thank you for your time. I am looking forward to meeting and discussing with you during my defense!

Many thanks to my dear colleagues and friends that have encouraged and supported me in the past years. I enjoy very much the discussions with Ludovic Métivier, and also his enthusiastic drum shows. I also appreciate his questions and insightful comments on

my work. I always remember the happy time shared with Yang Li, and acknowledge his endless personal assistance. Most importantly, I am very grateful to him for introducing me the special person the second day he knew me. I also want to express my gratitude to Pengliang Yang for the fruitful discussions, which greatly help me accomplish the real-data application. I enjoy very much his handy viscous inversion code as well as his beautiful calligraphy. I also want to acknowledge Guanghui Hu for contacting the accommodation in Grenoble before I arrived. I want to acknowledge my dear colleagues and friends Amir Asnaashari, Boussad Beldjenna, Stéphen Beller, François Bretaudeau, Cécile Brouyant, Clara Castellanos, Laure Combe, John Diaz, Michel Dietrich, Longfei Gao, Stéphane Garambois, Andrzej Górszczyk (Polish academy of sciences), Okba Hamitou, François Lavoué, Philippe Le Bouteiller, Isabella Masoni, Himanshu Midha, Alain Miniussi, Vadim Monteiller, Benjamin Pajot, Hugo Pinard, Alessandra Ribodetti, Josué Tago, Borhan Tavakoli, Phuong-Thu Trinh, Raphaël Valensi (OPERA), Paul Wellington and Huan Zhang. It has been a wonderful period to work with you. I appreciate the assistance and technical support from Magali Gardes, Stéphanie Mandjee and Sandrine Nadau as well as Hafid Bouchafa, Kamil Adoum and Jacque Pellet. I also appreciate the information provided by Christian Bigot and Guy Delrieu. I acknowledge the financial support from the SEISCOPE consortium as well as the HPC resources provided by CIMENT and CINES/IDRIS.

In addition, I want to thank my lovely friends who have supported me in the four-year life. Some of them are also in a PhD project or have finished it, giving me broader views and different ideas. Some of them are fans of football games like me, spending wonderful time together. I give my sincere acknowledgement to Jing He, Gang Li, Zhaohua Li, Bin Ma, Qiuliang Yu, Chao Yuan, Qiujuan Wang, Xinjiletu Zhao, Huaxiang Zhu, Yongliang Chang, Xinteng Gao, Tian Sun and others. I also want to thank Ophélie Passemard, Delphine Tohir, Nicole Coste and all the Ducros' for their understanding and kind assistance these days.

Living at another end of the continent makes me homesick, especially when all family members gather together to celebrate a traditional Chinese festival. I am deeply grateful to my parents and other elder relatives for bringing me up, telling me the truth of life, helping me follow dreams and supporting me to do PhD in a foreign country. These years I cannot stay long with them and their bright faces in my mind have become wrinkly and old. I hope in the near future, when I have a permanent location, I can see them more frequently and do whatever I can do. The gratitude is also for my master degree supervisor, Pr. Zhenxing Yao (IGGCAS), and former colleagues, Jinghai Zhang (IGGCAS) and others, who brought me into the field of geophysics and taught me how to do research. Thank you for your inspirations!

At last, I want to acknowledge my little girlfriend Lingzhu Xu. It is my greatest fortune and proud to meet you, date you, kiss you and finally marry you. Thank you for your consistent and endless care of me. You will be the most important and attractive research in my whole life!

Wei Zhou
July, 2016

The thesis is based on the following work primarily done by W. Zhou¹ during his PhD study from 04.2013 to 10.2016:

Technical reports:

- ZHOU, W. (2013). Toward large-scale imaging by reflected Waves: An analysis of wavenumber sampling of reflection FWI, Technical Report N° 61, SEISCOPE Project;
- ZHOU, W. (2014). Combining diving and reflected waves for velocity model building by waveform inversion, Technical Report N° 65, SEISCOPE Project;

Proceedings:

- ZHOU, W., BROSSIER, R., OPERTO, S. et VIRIEUX, J. (2014). Combining diving and reflected Waves for velocity model building by waveform inversion. *In Expanded Abstracts, 76th EAGE Annual Meeting (Amsterdam)*, We E106 15;
- ZHOU, W., BROSSIER, R., OPERTO, S. et VIRIEUX, J. (2014). Acoustic multiparameter full waveform inversion of diving and reflected waves through a hierarchical scheme. *84th SEG Technical Program Expanded Abstracts (Denver)*, Page 1249–1253;
- ZHOU, W., BROSSIER, R., OPERTO, S. et VIRIEUX, J. (2016). Joint full waveform inversion of early arrivals and reflections: a real OBC case study with gas cloud. *86th SEG Technical Program Expanded Abstracts (Dallas)*;

Publications:

- ZHOU, W., BROSSIER, R., OPERTO, S. et VIRIEUX, J. (2015). Full waveform inversion of diving and reflected waves for velocity model building with impedance inversion based on scale separation. *Geophysical Journal International*, **202**:1535–1554;
- ZHOU, W., BROSSIER, R., OPERTO, S., VIRIEUX, J. et YANG, P. (2016). Joint full waveform inversion of early arrivals and short-spread reflections: a 2D ocean-bottom-cable study including gas cloud effects. *To be submitted*.

¹Contact: wei.zhou@univ-grenoble-alpes.fr

Abstract

Full waveform inversion (FWI) has attracted worldwide interest for its capacity to estimate the physical properties of the subsurface in details. It is often formulated as a least-squares data-fitting procedure and routinely solved by linearized optimization methods. However, FWI is well known to suffer from cycle skipping problem making the final estimations strongly depend on the user-defined initial models. Reflection waveform inversion (RWI) is recently proposed to mitigate such cycle skipping problem by assuming a scale separation between the background velocity and high-wavenumber reflectivity. It explicitly considers reflected waves such that large-wavelength variations of deep zones can be extracted at the early stage of inversion. Yet, the large-wavelength information of the near surface carried by transmitted waves is neglected.

In this thesis, the sensitivity of FWI and RWI to subsurface wavenumbers is revisited in the frame of diffraction tomography and orthogonal decompositions. Based on this analysis, I propose a new method, namely joint full waveform inversion (JFWI), which combines the transmission-oriented FWI and RWI in a unified formulation for a joint sensitivity to low wavenumbers from wide-angle arrivals and short-spread reflections. High-wavenumber components are naturally attenuated during the computation of model updates. To meet the scale separation assumption, I also use a subsurface parameterization based on compressional velocity and acoustic impedance. The temporal complexity of this approach is twice of FWI and the memory requirement is the same.

An integrated workflow is then proposed to build the subsurface velocity and impedance models in an alternate way by JFWI and waveform inversion of the reflection data, respectively. In the synthetic example, JFWI is applied to a streamer seismic data set computed in the synthetic Valhall model, the large-wavelength characteristics of which are missing in the initial 1D model. While FWI converges to a local minimum, JFWI succeeds in building a reliable velocity macromodel. Compared with RWI, the involvement of diving waves in JFWI improves the reconstruction of shallow velocities, which translates into an improved imaging at greater depths. The smooth velocity model built by JFWI can be subsequently taken as the initial model for conventional FWI to inject high-wavenumber content without obvious cycle skipping problems.

The main promises and limitations of the approach are also reviewed in the real-data application on the 2D OBC profile cross-cutting gas cloud. Several initial models and offset-driven strategies are tested with the aim to manage cycle skipping while building subsurface models with sufficient resolution. JFWI can produce an accept-

able velocity model provided that the cycle skipping problem is mitigated and sufficient low-wavenumber content is recovered at the early stage of inversion. Improved scattering-angle illumination provided by 3D acquisitions would allow me to start from cruder initial models.

Résumé

L'inversion des formes d'onde (full waveform inversion, FWI) a suscité un intérêt dans le monde entier pour sa capacité à estimer de manière précise et détaillée les propriétés physiques du sous-sol. La FWI est généralement formulée sous la forme d'un problème d'ajustement des données par moindres carrés et résolu par une approche linéarisée utilisant des méthodes d'optimisation locales. Cependant, la FWI est bien connue de souffrir du problème de saut de phase rendant les résultats fortement dépendant de la qualité des modèles initiaux. L'inversion des formes d'ondes des arrivées réfléchies (reflection waveform inversion, RWI) a récemment été proposée pour atténuer ce problème en supposant une séparation d'échelle entre le modèle de vitesse lisse et le modèle de réflectivité à haut nombre d'onde. La formulation de RWI considère explicitement les ondes réfléchies afin d'extraire de ces ondes une information sur les variations lisses de vitesse des zones profondes. Cependant, la méthode néglige les ondes transmises qui contraignant les informations lisses de vitesse en proche surface.

Dans cette thèse, une étude de la sensibilité en nombre d'ondes des méthodes de FWI et RWI a d'abord été revisitée dans le cadre de la tomographie en diffraction et des décompositions orthogonales. À partir de cette analyse, je propose une nouvelle méthode, à savoir l'inversion jointe des formes d'ondes transmises et réfléchies (joint full waveform inversion, JFWI). La méthode propose une formulation unifiée pour combiner la FWI des transmissions et la RWI pour les réflexions, donnant naturellement une sensibilité commune aux petits nombres d'onde venant des arrivées grand-angle et réfléchies. Les composantes à hauts nombres d'onde sont naturellement atténuées par la formulation. Pour satisfaire l'hypothèse de séparation d'échelle, j'utilise une paramétrisation du sous-sol basée sur la vitesse des ondes de compression et l'impédance acoustique. La complexité temporelle de cette approche est le double de la méthode de FWI classique et la requête mémoire reste la même.

Une procédure d'inversion est ensuite proposée, permettant d'estimer alternativement le modèle de la vitesse du sous-sol par JFWI et l'impédance inversion de formes d'ondes réfléchies. Un exemple synthétique réaliste du modèle de Valhall est d'abord utilisé avec des données de streamer et à partir d'un modèle initial très lisse. Dans ce cadre, alors que la FWI converge vers un minimum local, la JFWI réussit à reconstruire un modèle de vitesse lisse de bonne qualité. La prise en compte des ondes tournante par la JFWI montre un fort intérêt pour la qualité de reconstruction superficielle, comparée à la méthode RWI seule. Cela se traduit ensuite par une reconstruction améliorée en

profondeur. Le modèle de vitesse lisse construit par JFWI peut ensuite être considéré comme modèle initial pour la FWI classique, afin d'injecter le contenu en haut nombres d'onde tout en évitant le problème de saut de phase.

Les avantages et limites de l'approche de JFWI sont ensuite étudiés dans une application sur données réelles, venant d'un profil 2D de données de fond de mer (OBC) recoupant un nuage de gaz au dessus d'un réservoir. Plusieurs modèles initiaux et stratégies d'inversion sont testés afin de minimiser le problème de saut de phase, tout en construisant des modèles de sous-sol avec une résolution suffisante. Sous réserve de mettre en œuvre des stratégies limitant le problème de saut de phase, la JFWI montre qu'elle peut produire un modèle de vitesse acceptable, injectant les bas nombres d'onde dans le modèle de vitesse. L'amélioration de l'éclairage en angles de diffraction fournie par des acquisitions 3D devrait permettre de pouvoir commencer l'inversion par JFWI à partir de modèle encore moins bien définis.

Contents

Acknowledgement	i
Abstract	v
Résumé	vii
1 Introduction	1
1.1 Seismic data and scale separation	2
1.2 Velocity inversion principles	6
1.2.1 Forward problem	6
1.2.2 Inverse problem	7
1.2.3 Summary	15
1.3 Imaging the subsurface velocity field	16
1.3.1 Ray-based traveltime tomography	16
1.3.2 Migration velocity analysis (MVA)	17
1.3.3 Full waveform inversion (FWI)	18
1.3.4 Reflection waveform inversion (RWI)	23
1.3.5 Summary and motivation of this study	25
1.4 Contribution of this work and thesis outline	25
2 From FWI to RWI	29
2.1 Forward modeling	30
2.2 Formulations and the cycle skipping issue	31
2.2.1 FWI as a least-squares linearized optimization	31
2.2.2 Velocity-depth ambiguity – Reflection data-induced cycle skipping	33
2.2.3 RWI based on scale separation and mitigation of cycle skipping .	36
2.3 Sampling analysis	38
2.3.1 FWI: preferentially high wavenumber samplings	38
2.3.2 RWI: preferentially low wavenumber samplings	42
2.3.3 Numerical verification	46
2.4 Discussion	48
2.5 Summary	51
3 Joint Full Waveform Inversion (JFWI)	53

CONTENTS

3.1	Introduction	54
3.2	Review of FWI and RWI	58
3.3	Methodology	61
3.3.1	Formulation	61
3.3.2	Mitigation of high-order isochrones by choosing suitable subsurface parameterization	63
3.3.3	Implementation	65
3.4	Integrated workflow of velocity (V_P) and impedance (I_P) inversion	67
3.4.1	I_P inversion by using short-offset reflection data	67
3.4.2	Cycle workflow of V_P - I_P imaging	68
3.5	Synthetic example: Valhall case study	69
3.5.1	Experimental setup	69
3.5.2	Results and discussions	71
3.5.2.1	Results and comparisons	71
3.5.2.2	Quality control by common image gathers	73
3.5.2.3	Fitting amplitudes	74
3.5.2.4	Broadband imaging of V_P	78
3.5.3	JFWI in presence of multiples	80
3.6	Conclusions and perspectives	84
4	Real Data Application	87
4.1	Introduction	88
4.2	Methodology	90
4.3	Application	92
4.3.1	Preliminary results related to attenuation	96
4.3.2	Inversion setup	99
4.3.3	Results	101
4.3.3.1	Classical FWI	101
4.3.3.2	Joint FWI	105
4.3.3.3	Joint FWI followed by FWI	106
4.3.3.4	Quality control of velocity models	111
4.4	Discussion	113
4.5	Conclusions and perspectives	117
5	Conclusions and Perspectives	119
	Bibliography	122
	Postscript	145
A	Derivation of Gradients through Lagrangian Formulation	149
B	High-order scattering effect	153

Chapter 1

Introduction

If I have seen further, it is by standing on the shoulders of giants.

— Isaac Newton, 1676

The mystery of the Earth's interior has been revealed by earthquake seismology to a much extent. Since the early 1900s, scientists have continuously found that our home planet is rather stratified, with a number of materials under varying pressure and temperature conditions. After a century of development, our knowledge about the underground has been largely enriched, from the physical processes to the chemical compositions, despite an increasing number of questions to be answered.

Since the late 1960s, seismology has also served to prospect the natural resources in the subsurface: a production-oriented branch of seismology that is often termed exploration seismology or simply seismics. Modern oil and gas industry has been and will be heavily relied on seismic methods in the foreseeable future. Researches, innovations, applications have brought a golden age of the seismic technology. Though the ideas are similar, the scales of interest are different. While continents, tectonic plates (hundreds of kilometers) or the whole Earth (thousands of kilometers) is the focus of seismologists, local regions on the crust (0.1 to 10 kilometers) are prospected for industrial application. Other differences can be found in Table 1.1, with a third type of seismology that will not be discussed in this thesis.

This thesis contributes to the community of exploration seismology, in particular, to its sub-field related to velocity inversion in the acoustic approximation. In this preliminary chapter, I shall first introduce some basic principles in seismic imaging followed by the discussion on a group of methods, which are viewed as the starting point of the methodology that I shall describe in Chapter 3.

Table 1.1: Three branches of seismology adopted in different circumstances (adapted from [Yilmaz, 2005](#)). Engineering seismology is oriented to the smallest scale with expected higher resolution by considering very high frequencies.

	Scale	Maximum depth	Source type	Max. frequency
Earthquake seismology	continental global	100 – 1000 km	Passive	10 Hz
Exploration seismology	local regions	10 km	Active	100 Hz
Engineering seismology	local regions	1 km	Active or passive	1000 Hz

1.1 Seismic data and scale separation: a synthetic example of marine case

Let us start from a synthetic model representative of the Valhall oilfield in the North Sea (Figure 1.1a). The seismic waves are recorded at the surface, a typical acquisition system among others in exploration geophysics. In this 2D setting, the x -axis measures the lateral position of the surface and the z -axis measures the depth from the surface. The yellow star represents one seismic shot, emitting transient mechanic energy into the subsurface. The energy of the excitation spreads out in the space with the form of propagating waves, carrying information of the physical properties of the medium, and captured by a set of receivers deployed at the surface. Figure 1.1b gathers from these grouped receivers the recorded phases (or seismic traces, events) as a function of elapsed time t and horizontal axis x . Relatively, each trace can also be associated with the horizontal distance between the source and receiver of that trace, i.e. the offset h . Note that x and h can be deduced from each other by knowing the source position. Due to the transient excitation, the phases have limited time duration. Due to the wave propagation, we can observe lateral continuity of these phases with certain slopes. Among them, several phases have great importance. The colored arrows identify these phases as well as their associated raypaths connecting the source and receiver for the following interpretation (Figure 1.1b):

1. Transmitted waves from shallow zones:

- Direct wave (green arrows). As the superficial part of the model is homogeneous (water column), the direct wave follows a straight path;
- Diving wave (green dashed arrows). According to Fermat's principle, this type of wave follows a curved path that penetrates higher velocity zones at depth so as to arrive at the receiver with shorter time than the direct wave. They are also known as turning wave;

2. First-order scattered waves:

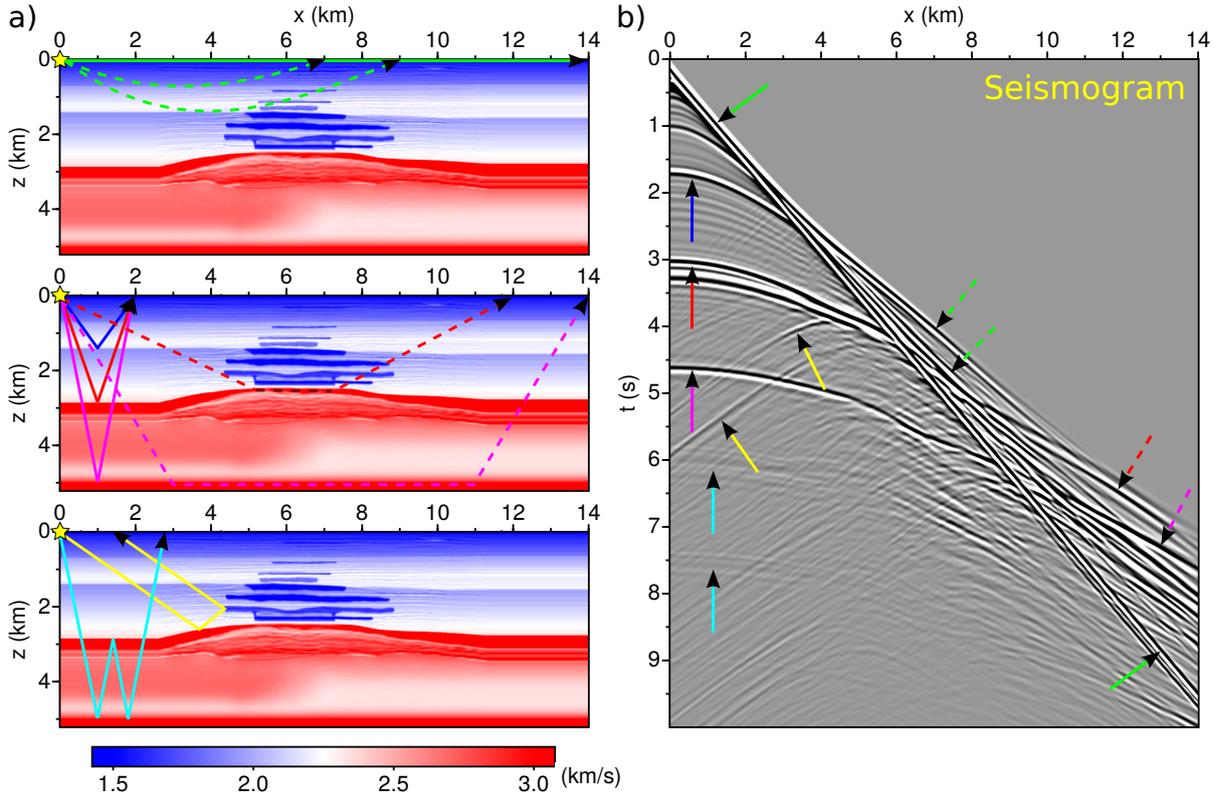


Figure 1.1: An example of seismic data acquired at the surface. (a) Velocity model. (b) Seismogram recorded by an array of receivers. The yellow star denotes the position of the seismic shot. The wavefield is computed under the acoustic approximation to illustrate only the P waves. The colored arrows identify several phases in (b) and the associated raypaths in (a) that are interpreted in the text.

- Reflected wave. When the wave meets an interface, part of its energy is reflected backwards to the surface, leaving a hyperbolic-like shape (normal moveout) in the seismogram. The blue, red and magenta arrows indicate the reflections occurred on interfaces at 1.5 km, 3 km and 5 km depth, respectively. At far distances, these reflected waves indicated by blue and red arrows tend to be tangent with the diving waves due to the inhomogeneity of overlaying medium. Therefore, we can further classify the reflected waves as precritical (or short-spread) reflection and postcritical (or supercritical) reflection based on their different moveouts;

3. Transmitted waves from deep zones:

- Refracted waves. When the wave meets an interface, part of its energy is forward transmitted to deeper layers. As long as the velocity beneath is higher, the wavefront propagating in parallel with the interface serves as secondary sources and generate plane waves propagating to the surface, leaving a nearly straight shape (linear moveout) in the seismogram. This is the case for the

phases indicated by the dashed red arrows, associated with the interface at 2.5 – 3 km depths, and the phases indicated by the dashed magenta arrows, associated with the interface at 5 km depth, respectively. At very long distances, these refracted waves can arrive at the surface earlier than the diving waves.

4. Higher-order scattered waves:

- Multiscattered waves (yellow arrows). When in contact with the sharp edges of the embedded blue layers, the reflected wave (red arrows) serves as secondary sources and generates spherical waves that propagate to the surface, leaving a hyperbolic-like shape in the seismogram. As the incident reflected wave is of first-order scattering, these spherical waves are of higher orders.
- Multiples (cyan arrows). They are high-order reflected waves that are bounced back and forth for multiple times between the surface and interfaces (surface-related multiples), or just between different interfaces (internal multiples). In Figure 1.1b, the absorbing surface condition is used to avoid surface-related multiples. Because of the intrinsic energy partition between reflection and transmission, the internal multiples have very weak amplitudes, therefore, they are usually neglected in reflection seismic imaging.

Generally, we are interested in two types of information included in the seismic data:

1. Kinematic information: the traveltime for one phase to propagate from one source to one receiver, which is a functional of the wavepath and velocity field;
2. Dynamic information: the amplitude of one phase which depends on the source wavelet, the geometrical spreading factor, the dissipation of the medium, and the reflection/transmission coefficients of the interfaces.

In seismic methods, the kinematic and dynamic information is incorporated in different degrees. For example, ray-based tomography methods only use the kinematic property to build a smooth velocity model. Kirchhoff migration relies more on the kinematic than the dynamic to focus the reflector images without properly resolving the reflection/transmission coefficients. Full waveform inversion, on the contrary, simultaneously considers both information for broadband subsurface imaging.

Conventional surface acquisition geometries, such as the streamers in the marine case, often provide insufficient offset and/or azimuth coverage giving small-aperture data. Consequently, the subsurface models we can build lack intermediate wavenumbers in the Fourier domain and provide two parts of the subsurface spectrum. An illustration on the synthetic Valhall model is given in Figure 1.2. This leads to the concept of *scale separation* that distinguishes between large-scale (low wavenumbers) and small-scale (high wavenumbers) models (Claerbout, 1985; Jannane *et al.*, 1989; Wang et Pratt, 1997), assuming that we are unable to resolve the intermediate wavenumbers, and the

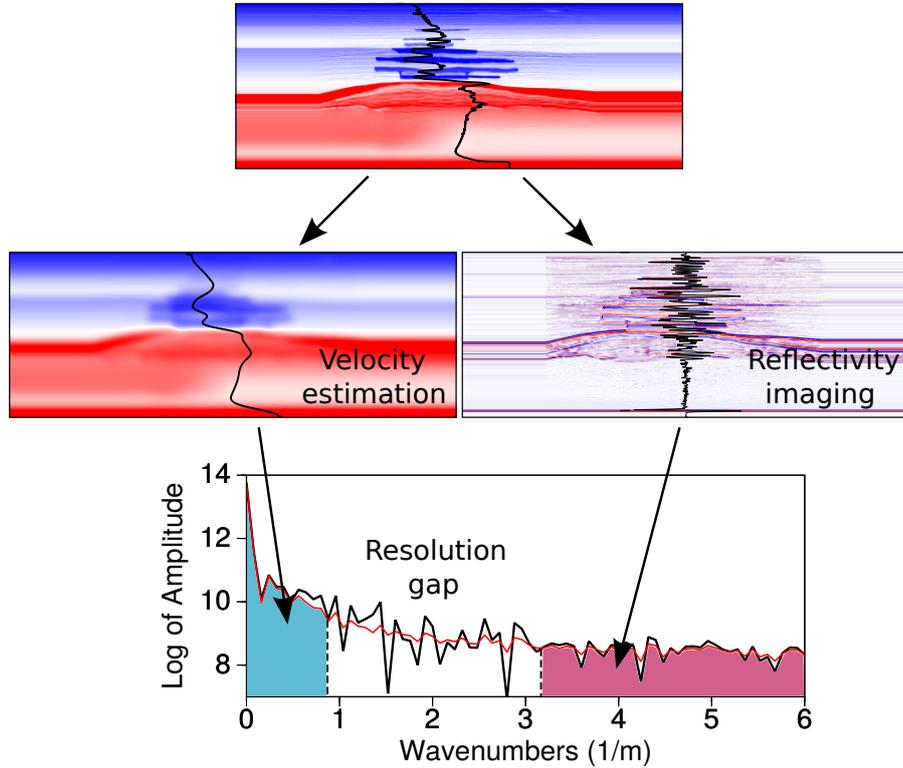


Figure 1.2: Sketch of the scale-separation assumption (inspired by the famous picture of Claerbout (1985), figure 1.4-3). As the variations of the subsurface (top) are sampled by seismic waves, we are able to resolve its low-wavenumber part by velocity estimation methods and high-wavenumber part by reflectivity imaging methods, respectively, and this property is termed scale separation in the literature. However, these two parts often do not overlap, leaving the intermediate wavenumbers unsolved. Fourier analysis (bottom) confirms this point, that the true spectrum (black curve) is matched by the reconstructed spectrum (red curve) in the low-wavenumber part (blue area) and the high-wavenumber part (red area) with a resolution gap in the middle.

seismic model with the resolution gap can reproduce the recorded data. The high-wavenumber part is termed *reflectivity* while the low-wavenumber part is termed (*macro*) *velocity model*. Depending on the objective, seismic imaging methods can be divided into two categories following Wapenaar (1996):

1. Reflectivity imaging (or seismic imaging, migration), which aims to image the high-wavenumber part, or more specifically aims to resolve the reflection/transmission coefficients of the interfaces inferred from the dynamic property of the data;
2. Velocity estimation (or velocity analysis, velocity model building), which aims to reconstruct the low-wavenumber part inferred from the kinematic property. Here, the term “velocity” specifically refers to the smooth velocity background.

Considering this natural separation due to insufficient acquisition coverage, in the conventional seismic processing workflow, the kinematic information is extracted as the initial step to build the large-scale velocity field, followed by seismic imaging of the small-scale reflectivity profile involving the dynamic information. Nevertheless, we should note that such a separation is less true nowadays, because efforts have been taken to recover intermediate wavenumbers by designing wide-azimuth long-offset acquisition geometries, using high-quality broadband data, and applying high-resolution imaging methods such as full waveform inversion (as reported by [Lambaré *et al.*, 2014](#)).

1.2 Velocity inversion principles

Velocity estimation can be formulated as an inverse problem that “looks for a question which can be responded by the proposed answer”. In particular, the recorded seismic data are the proposed “answer” and the subsurface model is the “question” we want to find. However, before solving this inverse problem, we need to know how to solve the forward problem: given a model, how to compute the synthetic seismic data (“how to answer a given question”). These two problems make up velocity inversion that is often formulated in the frame of optimization theory. In this section, I shall discuss some general principles in solving the forward and inverse problems using waveforms. The mathematical formulation is deferred until Chapter 2.

1.2.1 Forward problem

Because of their similar kinematic behavior, seismic waves can be treated as optic rays based on high-frequency approximation ([Červený, 2001](#); [Chapman, 2004](#); [Virieux *et al.*, 2007](#)). The traveltimes of one phase is then computed by the integration of the inverse of local velocities (slowness) along the raypath that connects the source-receiver couple, which is also known as ray-tracing ([Zelt *et al.*, 1992](#); [Bishop *et al.*, 1985](#)). However, according to Fermat’s principle, low-velocity zones tend to be under sampled by raypaths leading to poor resolution in complex media ([Virieux *et al.*, 1991](#)). An alternative to tracing rays consists in solving the eikonal equation, a nonlinear partial differential equation that relates the derivatives of traveltimes to local velocities ([Le Meur *et al.*, 1997](#)). It can be numerically solved either by the fast marching method ([Popovici *et al.*, 1998](#); [Lelièvre *et al.*, 2011](#)) or by the fast sweeping method ([Vidale, 1990](#); [Zhao, 2004](#); [Bretaudeau *et al.*, 2014](#)).

Alternatively, one can consider the wave equation to honor the real physics of seismic waves. The wave equation is a linear partial differential equation and can be discretized by various schemes ([Kelly *et al.*, 1976](#); [Marfurt, 1984](#); [Virieux, 1984](#); [Dablain, 1986](#); [Levander, 1988](#); [Brossier *et al.*, 2008](#)). The finite-difference method ([Virieux, 1984](#); [Levander, 1988](#)) is popular for its simplicity and scalability, but generates artificial boundary reflections that are often mitigated by considering absorbing boundary conditions (e.g. perfectly-matching layers, smart layers, [Bérenger, 1994](#); [Métivier *et al.*,](#)

2014b). On the other hand, the finite-element method (Marfurt, 1984; Seriani et Priolo, 1994; Min *et al.*, 2003; Dumbser et Käser, 2006; Käser et Dumbser, 2008), which is often more computationally intensive, naturally introduce boundary conditions giving more accurate solutions. By using tetrahedral meshes (Frey et George, 2008) the method can almost perfectly fit complex topographies, and the mesh width can be locally adapted to medium properties.

On the other hand, image-domain velocity inversion, as well be described later, considers the reflectivity images inferred from synthetic data as the solution of its forward problem. In other words, the forward problem for image-domain inversion consists of forward modeling (computes synthetic data) followed by migration (computes reflectivity images). For this reason, the computational cost of image-domain inversion is often larger than its data-domain counterpart. For overviews of migration techniques I refer to Sava et Hill (2009) and Etgen *et al.* (2009). Likewise, migration can be also based on the ray approximation (Kirchhoff migration, Schneider, 1978). The computation load is less heavy than wave equation-based migration; however, the latter one is currently more discussed in the literature for complex-structure imaging. A number of approaches have been proposed. Using the approximate one-way wave equation, the wavefield at one depth can be extrapolated to another depth, and images can be built by the coincident time imaging condition (Claerbout, 1985; Wu, 1994; Ristow et Ruhl, 1994; Le Rousseau et de Hoop, 2001). Alternatively, the wavefield can be computed through time marching methods using the true two-way wave equation (*reverse time migration* (RTM), Baysal *et al.*, 1983; McMechan, 1989). Besides, true amplitude or quantitative migration aims to resolve the reflection/transmission coefficients of interfaces, and the second-order information can be considered (Lameloise *et al.*, 2015). Least-squares migration is formulated as a linear inverse problem to explicitly solve for the reflectivity (Nemeth *et al.*, 1999). In Chapter 3, I shall also propose an RTM-like least-squares waveform inversion to image reflectivity in the impedance parameter.

1.2.2 Inverse problem

Without talking about other methodologies, I mention the following two ways to assess the quality of the velocity model:

1. Data domain, which measures the data fitness to assess the velocity model. If the observed data are well matched by the synthetic data computed in the proposed velocity model (and other parameters such as density), then we regard the proposed model as a reliable representation of the subsurface (i.e. true model);
2. Image domain, which measures the “coherency” of reflectivity images respectively resolved from each source-receiver couple. If the positions of these images are independent of the source-receiver offset (or other acquisition parameters that are not a model attribute), then we regard the proposed velocity model as the true model. Other “coherency” criteria have also been proposed.

Therefore, any data mismatch (in the data domain) or image incoherency (in the image domain) are attributed to velocity inaccuracies, and can be used to update the velocity model. Usually, the relationship between the velocity model and the data match or image coherency is nonlinear, but we still apply local optimization schemes to solve the inversion problem due to its large size. In the following, I shall show how full waveform inversion (FWI) and reflection waveform inversion (RWI), representatives of data-domain methods, transform their respective data misfit into model update through the gradient computation flow. After this, image-domain methods will also be described using an example of common image gather (CIG).

Data domain velocity inversion: FWI case

Figure 1.3 illustrates how we build the FWI gradient based on wavefield simulation. Suppose that we have estimated the source wavelet of the observed waves (black wiggle in top left panel). We can simulate the propagation of the incident wavefield (blue-white-red color scale) in the proposed model (left panels). Suppose this model is homogeneous and isotropic, then the wavefront has a circular shape (solid arrows in left panels) with the source position (stars) as its center. We sample the wavefield at the receiver position (triangle) and compare it with the observed wavelet often sample by sample (blue *vs* red wiggles).

This forward modeling process is followed by a back propagation process which computes the adjoint field (middle panels) and the gradient (right panels) to convert the data misfit into model update. For simplicity I take the classical L2 norm-formulated FWI which implies the difference between modeled and observed data (i.e. the data residuals) to be the adjoint source of the adjoint field (black wiggle in bottom middle panel). Based on the formulation which will be present in Section 2.2, the adjoint field should be computed in the reverse time order as indicated by the right-most upward thick arrow. Suppose we have two separated signals in the residuals which are associated to direct and reflected waves, respectively, then we generate two wavefronts in the adjoint field. The inner one comes from the direct wave while the outer one from the reflected wave.

The gradient is formed by pixel-to-pixel zero-lag cross-correlations between the incident and adjoint fields. The direct wave-associated wavefront in the adjoint field interferes with the incident wavefront where their wavenumber vectors (dashed and solid arrows in right panels, respectively) make the open angle (θ^d) nearby 180° . According to the diffraction tomography principle, the magnitude of the imaging wavenumber vector formed by such interference is inversely proportional to the cosine of θ . Consequently, the direct wavefront and its adjoint counterpart generate several isophase elliptic zones in the gradient map with large-wavelength variations along the direct wavepath between the source and receiver, the inner-most of which is known as the first Fresnel zone and others are known as secondary Fresnel zones (red and blue elliptic zones).

On the other hand, the reflected wave-associated wavefront in the adjoint field interferes with the incident wavefront where their wavenumber vectors make the open angle (θ^r) nearby 0° . Consequently, they contribute to small-wavelength components inside

the reflection-associated secondary Fresnel zone, in an elliptic shape with the source and receiver as the focal points, which is also known as the migration isochrone (Tarantola, 1984; Lailly, 1984).

The gradient is stacked over all source-receiver couples (e.g. Figure 1.5a for the Valhall model), and is scaled to update the velocity model. The gradient consists of the large-wavelength component, coming from the transmitted waves (direct, diving and refracted waves), and the small-wavelength component, coming from the reflected waves. As a result, in FWI, we do not assume scale separation and we are looking for a broadband image of the subsurface. However, because of the surface acquisition geometry, we can only record a part of the transmitted waves that travels in a limited depth of the subsurface, and the large-wavelength component of the gradient are concentrated in the shallow zone of the subsurface. The small-wavelength component, on the contrary, is distributed at all depths (they are overlaid by large wavelengths in shallow zone but we do have the sensitivity there).

To mitigate the lack of large-wavelength component in depth, one often looks for larger offsets to record diving waves passing through greater depths. However, this is not efficient: an empirical relation states that the penetration depth of diving waves is only one third to one sixth of the largest offset; we may need an acquisition length more than twenty kilometers in order to record the diving waves that sample the targets at the reservoir level. Using lower frequencies in another solution because it brings lower wavenumber content from the migration isochrone. However, since the source wavelet is band-limited we would be obstructed by the noise in the low-frequency end.

Data domain velocity inversion: RWI case

Large-wavelength updates can be extracted from reflected waves which naturally extends to great depths (Figure 1.4). This is achieved in RWI by generating reflection wavefields during both forward and backward modeling processes using prior reflectors. To clarify which wavefront is mentioned, I shall adopt the words “downgoing” and “upgoing” although the wavefront does not propagate only in the vertical direction.

During forward modeling, the downgoing incident field interacts with the reflector (black line in the model) and generates the upgoing reflection field. Due to the energy partition on the reflector, this upgoing field is weaker than the downgoing field. The upgoing field has a semicircle wavefront that is symmetric to the wavefront of the downgoing field. Considering such symmetry, we can denote the wavenumber vector of the upgoing wavefront (gray arrow) by assuming a virtual source (gray star) at the mirror position of the real source (black star) with regard to the black line.

The whole incident wavefield (both downgoing and upgoing) is sampled at the receiver position (triangle). However, standard RWI only considers the contribution from reflected waves; therefore, the direct wave residual is muted in the adjoint source (black *vs* red wiggles in bottom middle panel). Similar to the incident field, the adjoint field has also the downgoing and upgoing partition due to the presence of reflector, and the wavenumber vector of the upgoing wavefront (dashed gray arrow) can be denoted by as-

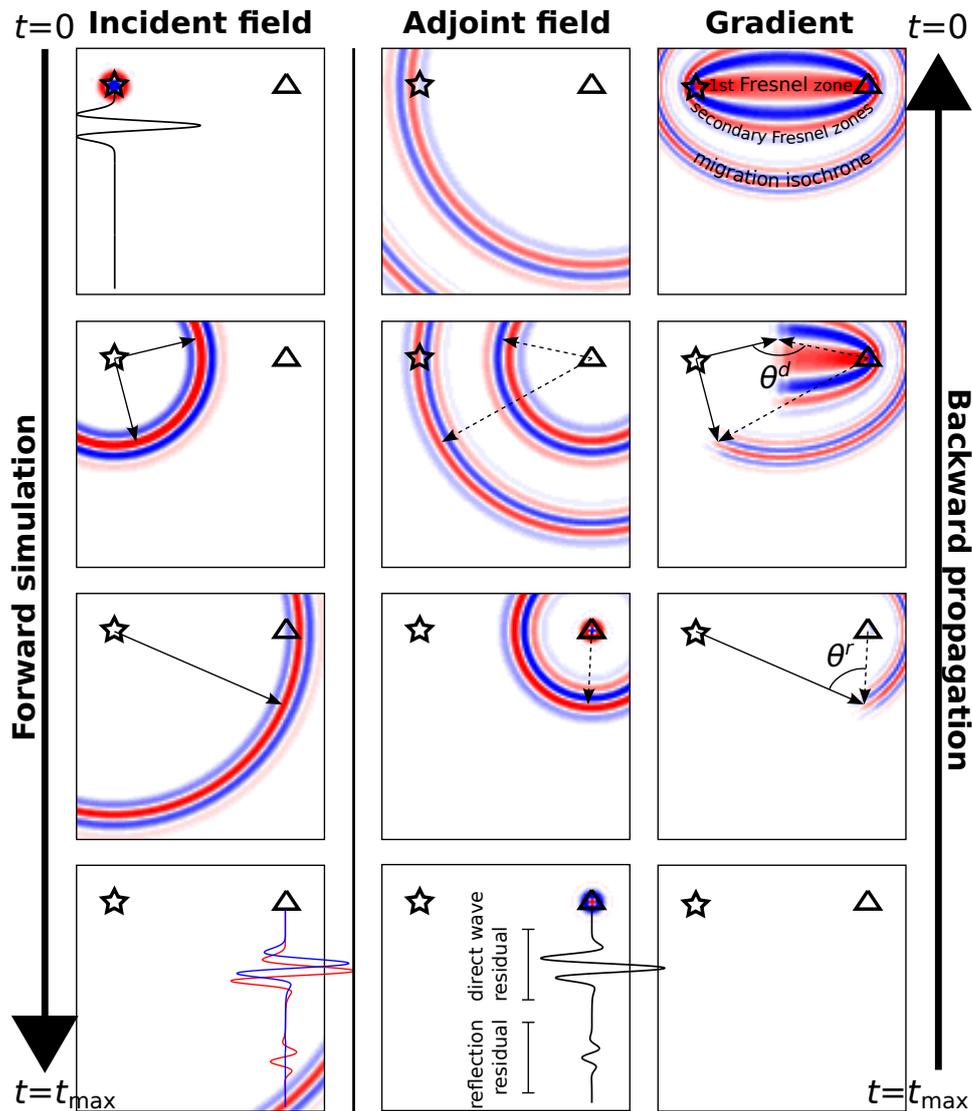


Figure 1.3: Time-domain FWI gradient computation animation. Incident field (left panels) by forward simulation (top to bottom panels), adjoint field (middle panels) and gradient (right panels) by backward propagation (bottom to top panels), plotted in the blue-white-red color scale. Stars and triangles denote the source and receiver positions, respectively. In the homogeneous isotropic model, the wavefronts have a circular shape centered at the source or receiver position. Their wavenumber vectors are denoted by arrows. Two angles are made: 1) θ^d between vectors of the incident wavefront and adjoint wavefront coming from direct wave; 2) θ^r between vectors of the incident wavefront and adjoint wavefront coming from reflected wave. Because of their different angle ranges, large-wavelength Fresnel zones are formed along the direct wavepath between source and receiver, whereas small-wavelength migration isochrone (reflection-associated Fresnel zone) is formed in an elliptic shape with the source and receiver as the focal points. Online animation: <https://drive.google.com/open?id=0Bx0JCm2KZyuebXI5M21YTW8tLTg>

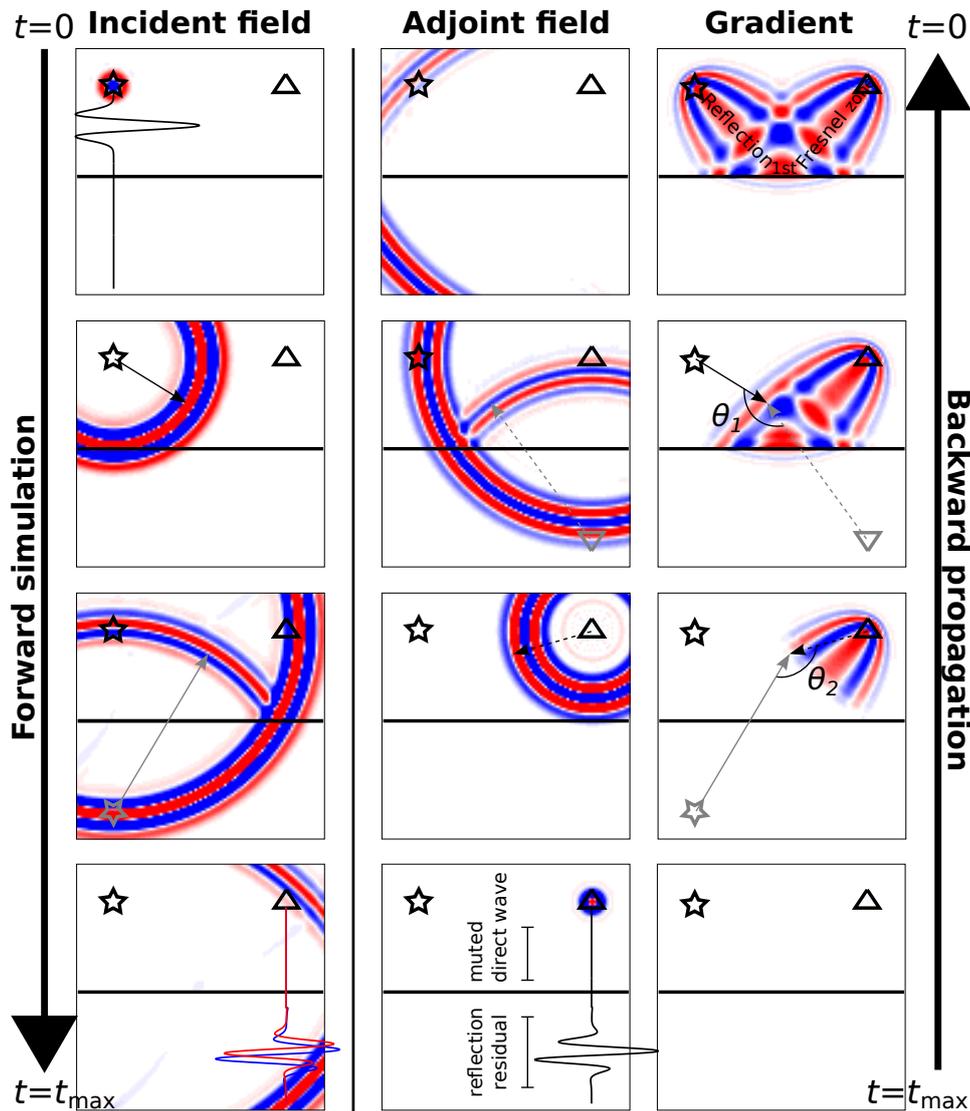


Figure 1.4: Time-domain RWI gradient computation animation. Incident field (left panels), adjoint field (middle panels) and gradient (right panels) same as in Figure 1.3. Black stars and triangles denote the source and receiver positions, and gray stars and triangles denote the mirrored positions, respectively. Direct wave is muted at receiver. Both incident and adjoint fields are scattered due to the existence of reflector. The wavenumber vectors associated to downgoing wavefronts are denoted by black arrows, and the ones associated to upgoing wavefronts are denoted by gray arrows, respectively. Two angles are made: 1) θ_1 between vectors of downgoing incident wavefront and upgoing adjoint wavefront; 2) θ_2 between vectors of upgoing incident wavefront and downgoing adjoint wavefront. Because of their large angles, large-wavelength Fresnel zones are formed along reflection wavepath between surface and reflector. In this animation, I have filtered out the migration isochrone that was generated by the two downgoing wavefronts. Online animation: <https://drive.google.com/open?id=0Bx0Jcm2KZyueM25GMF96STJ3QWs>

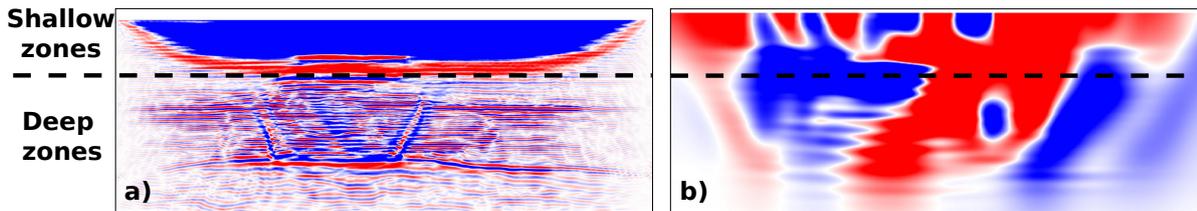


Figure 1.5: FWI (a) and RWI (b) gradients stacked over surface acquisition in the Valhall case. Smooth initial models are used. Migration is implemented before RWI to image reflectivity. For FWI gradient, large-wavelength components coming from transmitted waves are concentrated in shallow zones, whereas small wavelengths coming from reflected waves are distributed in the whole space. In contrast, RWI aims to build a smooth velocity model by using reflected waves, which can remedy the large-wavelength shortage of FWI gradient in deep zones.

suming a virtual receiver (gray triangle) at the mirror position of the real receiver (black triangle) with regard to the black line.

In the right panels, the upgoing wavefront in the incident field interferes with the downgoing wavefront in the adjoint field making an open angle (θ_2) nearby 180° . Similarly, the downgoing wavefront in the incident field interferes with the upgoing wavefront in the adjoint field making another open angle (θ_1) also nearby 180° (notations to be consistent with Section 2.3). Consequently, they contribute to large-wavelength components inside the reflection-associated Fresnel zones along the two-way wavepath between the surface and reflector followed by the reflected wave. Therefore, we can recover large-wavelength variations above the deepest reflector.

On the other hand, the interference of the two downgoing wavefronts in the incident and adjoint fields should have generated the small-wavelength migration isochrone as shown in the FWI gradient. However, we intentionally remove this component because the formalism of the RWI gradient does not include this component (see Chapter 2). We can do this because we can explicitly reproduce the migration isochrone (to the first order at least) through the FWI gradient computation process using only the reflected wave residual and the reflector-free model. In this way, the actual computation of the RWI gradient amounts to compute four wavefields and double time complexity of the FWI gradient. As a result, the RWI gradient is dominated by its large-wavelength components. The stacked gradient for the Valhall model is shown in Figure 1.5b.

Both FWI and RWI rely on data mismatch to update the subsurface model. However, they differ in two main aspects: 1) Unlike FWI, RWI adopts the scale-separation assumption in which the prior reflector is fixed during inversion iterations whereas the smooth velocity is the aim to be updated; 2) RWI relies on reflected waves to image the long wavelengths in deep zones whereas FWI relies on diving waves to do this job but only limited to shallow zones. In Chapter 2, I shall provide a detailed analysis and complete explanation of the FWI and RWI gradients.

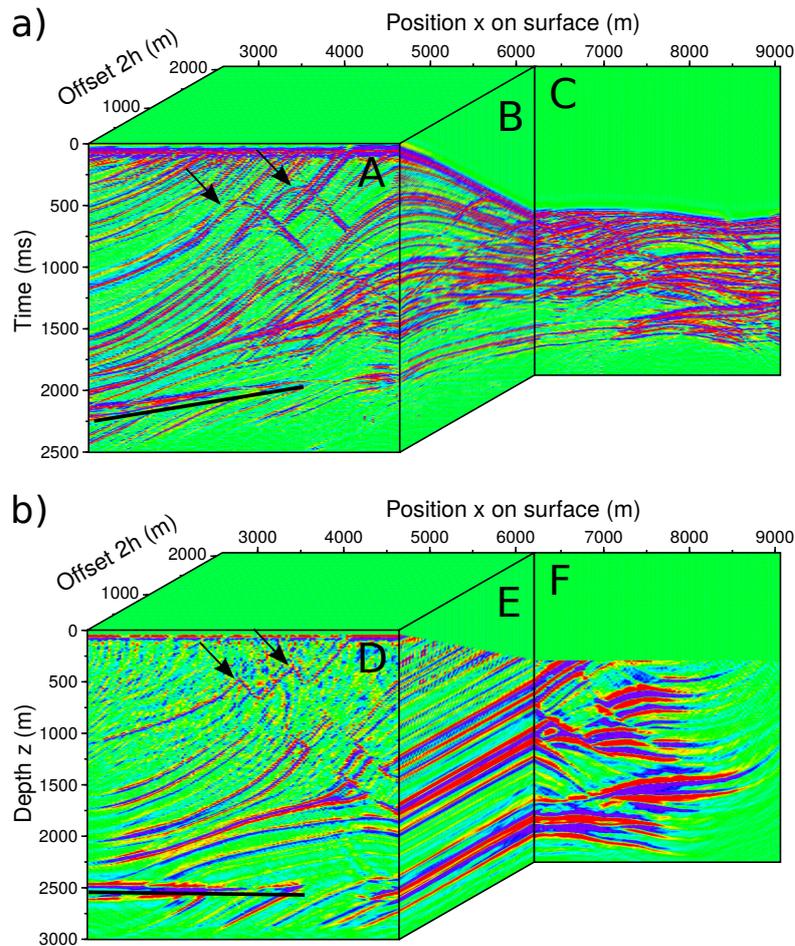


Figure 1.6: Velocity inversion in image domain (adapted from Chauris, 2000). 3D presentation of the time-domain seismic data computed in 2D Marmousi model (a), converted to surface-offset depth-domain common image gathers (b) by prestack depth migration using the true velocity model. A,D: prestack zero-offset common offset gather; B,E: common midpoint gather; C,F: prestack common offset gather collected at largest offset. Accurate velocity model results in flat images across offsets (E). Conversely, residual moveout of migrated images can be used to update the velocity model.

Image domain velocity inversion

Image-domain methods rely on the fact that the data space is always larger than the model space. Take the 2D geometry as an example, we want to build a 2D model (dimensions of z and x) from a 3D data set (dimensions of t and two spatial locations regarding the shot and receiver). Although in the case of multiparameter inversion more 2D models are simultaneously built, we still enjoy an over-determination property of the inverse problem.² Therefore, we can use such data redundancy to quality control the

²Note that this is not to say solving this inverse problem is easy. On the one hand, the relation between the wavefield and the propagation medium is nonlinear therefore linearized optimization methods often converges to a local minimum. On the other hand, deep targets are insufficiently sampled (some-

velocity model and update it.

Among different variants I shall describe the flatness criterion derived from the surface-offset common image gathers. Figure 1.6a shows three profiles of the seismic data computed in the Marmousi model. Profile A shows the prestack common offset gather collected at zero offset. The continuous reflection phases (e.g. black line) indicate a series of reflectors in the subsurface, and the scattered waves with a hyperbolic shape (black arrows) indicate point diffractors located on reflector discontinuities. Profile B shows the common midpoint gather collected at a surface position (fixed x). The normal moveout of the reflected waves, i.e. the behavior of later arrivals at larger offsets, is a typical kinematic information for velocity estimation. In the common offset gather collected at the largest offset (Profile C), the reflection phases are recorded in a rather short time window, hence are hard to be interpreted.

Prestack depth migration converts these time-domain data into depth-domain reflectivity images (Figure 1.6b, note the different vertical axis). If the accurate velocity model is used, the reflection phases in the common offset gathers are reshaped by migration and should be transferred to the associated reflector positions (comparing A and D in Figure 1.6). For example, the phases (black line) arrive later than 2 s in Profile A are reshaped to be a flat reflector image at 2500 m depth in Profile D. The flanks of the hyperbolic scattered waves (arrows in A) are refocused on the associated reflector discontinuities (arrows in D), indicating the fault structure at shallow depth. Similar to the FWI case, larger offsets provides lower wavenumber components, therefore the reflectivity image in Profile F is less sharp as the image in Profile D. By stacking over offsets, a broadband image (including more high wavenumbers) can be built for geological interpretation.

When the proposed velocity model is inaccurate, we cannot stack over offsets due to the destructive stack between low- and high- wavenumber images. Conversely, we can assess the velocity model by measuring their consistency: if the velocity is accurate, the reflectivity images should be positioned at the same depth across offset (e.g. flat images as shown in Profile E); otherwise any depth inconsistency (i.e. residual moveout) indicates an erroneous velocity background and thus can be used to update the velocity model. Note that only one common midpoint gather collected at a fixed position (e.g. Profile E) is not sufficient to update the velocity model in the whole space. Therefore, sufficient amount of common midpoint gathers that are averagely sampled on the surface should be considered, which inevitably causes a large consumption of the memory space.

The common midpoint gather mentioned above is one possibility of what is referred to as common image gathers (CIGs) in image-domain methods. Other CIGs, such as the angle-domain CIG and the subsurface-offset CIG are also proposed to reduce the nonuniqueness of the inverse problem, and to facilitate an automatic consistency measurement. A discussion will be devoted to this topic in Section 1.3.2.

Migration is implemented in both RWI and image-domain methods. They both aim to update the smooth velocity based on the scale-separation assumption. However, they

time only by precritical reflected waves) making imaging at this depth very ill-posed. See a real-data example in Chapter 4 that is encountered in my PhD study.

Table 1.2: Classification of velocity inversion methods based on how the forward and inverse problems are solved, arranged in rows and columns, respectively (adapted from Jones, 2010). Examples are provided in each entry with full names given in the text.

	Data domain	Image domain
Ray based (purely kinematic)	FATT, RTT	Prestack migration tomography
Waveform based (kinematic + dynamic)	Finite-frequency tomography, FWI, RWI, JFWI	MVA

differ in two main aspects: 1) Unlike RWI, image-domain methods consider the image inconsistency rather than data misfit to update the velocity model; 2) The migration image for RWI is independent on offset whereas multiple image profiles are built for each offset in image-domain methods. In other words, the dimension of the image for RWI is smaller than that for image-domain methods, therefore requires less memory space and shorter computational time.

1.2.3 Summary

Depending on how the forward and inverse problems are solved, velocity inversion methods can be classified into four categories (Table 1.2). In first arrival traveltime tomography (FATT) and reflection traveltime tomography (RTT), seismic waves are approximated as high-frequency rays and only traveltime information is used to update the model. They cannot provide high-resolution images especially in complex media, and the dynamic information should also be considered. Possible options include the finite-frequency traveltime tomography that measures the time delays by cross-correlation of waveforms, or standard FWI and RWI directly that directly computes the data misfit. On the other hand, migration is also included in the forward problem of image-domain methods, which can be based either on ray approximation or on wave equation. The latter one is also known as migration velocity analysis (MVA), which is currently a hot topic in the literature. In general, because the associated objective functions are more convex than the waveform-based objective function associated to data-domain methods, MVA is more robust in practice (Symes, 2008). Nevertheless, particular data domain methods are also proposed that promise similar stability as MVA (Clément *et al.*, 2001; Chauris *et Plessix*, 2012; Chavent *et al.*, 2014; van Leeuwen, 2010; Gao *et al.*, 2014).

The method proposed in this thesis, namely joint full waveform inversion (JFWI), is a waveform-based data-domain method for velocity inversion. It combines FWI and RWI for large-wavelength imaging whereas small-wavelength contribution is suppressed. In the next section, as a preamble of JFWI, I shall discuss on ray-based traveltime tomography, MVA, FWI and RWI in more details. The introduction of JFWI will be given in Chapter 3 after a theoretical analysis in Chapter 2.

1.3 Imaging the subsurface velocity field

1.3.1 Ray-based traveltimes tomography

Ray-based traveltimes tomography aims to update the velocity model by using the traveltimes of seismic phases. As already mentioned, the synthetic traveltimes can be computed by either ray-tracing or solving the eikonal equation. On the other hand, the observed traveltimes are often manually picked from the recorded data based on a prior phase identification. Picking the first breaks in the seismic data, first arrival traveltimes tomography (FATT) relies on traveltimes of diving and refracted waves (Zelt *et al.*, 1998; Toomey *et al.*, 1994). This method is also widely adopted in global seismology because the very-far acquisition systems can provide significant transmitted waves that have passed through deep zones (Bijwaard *et al.*, 2000; Montelli *et al.*, 2004a,b). However, by surface acquisition deep targets are insufficiently sampled by transmitted waves due to the limited offset range. In contrast, reflection traveltimes tomography (RTT) relies on the traveltimes of reflected waves that better sample deep targets. Similar to reflection waveform inversion, reflectivity images should be built explicitly such that the synthetic traveltimes of reflected waves can be computed (Bishop *et al.*, 1985; Sword, 1987; Farra *et al.*, 1988; Whiting, 1998). The traveltimes residuals are then used to update both the velocity and reflector depths, introducing the dilemma between vertical velocities and depth positions of interfaces. Nevertheless, since the transmitted and reflected waves sample in different depth ranges and contribute to distinguished velocity update, combinations of FATT and RTT have been proposed for a better velocity modeling building (Zelt *et al.*, 1992; Hobro *et al.*, 2003; Huang *et al.*, 2012).

However, picking the reflection traveltimes in the observed data is often difficult. For example, early reflection phases with large slopes could interfere with late ones with small slopes at far offsets, therefore makes the late phase picking troublesome. To overcome this issue, stereotomography is promoted to pick locally coherent seismic events characterized by their slope which is done in a semi-automatic way (Billette *et al.*, 1998; Billette *et al.*, 2003; Lambaré *et al.*, 2004; Lambaré, 2008). Moreover, because the slope is a kind of kinematic information that is independent of traveltimes, it provides additional constraints for velocity model building. At the beginning, the method was applied to reflected waves, but extensions to converted waves (Alerini *et al.*, 2007; Nag *et al.*, 2010) and transmitted waves (Gosselet *et al.*, 2005; Prioux *et al.*, 2013c) have also been proposed.

The development of ray-based traveltimes tomography is reviewed by Woodward *et al.* (2008) and Lambaré *et al.* (2014). Among all efforts, the main is the hunting for high resolution. Waveform inversion could be a substitute of traveltimes tomography for this purpose. However, as the relation between waveform and velocity is highly nonlinear, waveform inversion is generally less stable in practice. Therefore, combining tomography and waveform inversion is more robust to reach the high resolution. One possibility is to build the large wavelengths of the velocity model by traveltimes tomography before imaging fine structures by waveform inversions (e.g. Prioux *et al.*, 2013c; Vigh *et al.*, 2016;

Gorszczyk *et al.*, 2016). Another possibility is to perform tomography and waveform inversion simultaneously (e.g. Zhang *et al.*, 2014; Treister et Haber, 2016). The weights between the two should be tunable by the user to emphasize one method over another.

The principle that local slopes can be used to update the velocity model is not limited to stereotomography; a connection of stereotomography with differential semblance optimization (an MVA method) has been highlighted by Symes (2008); Chauris et Noble (2001). In parallel with tomography, MVA has served as an alternative approach for velocity model building, which is discussed as follows.

1.3.2 Migration velocity analysis (MVA)

Migration velocity analysis relies on the data redundancy to build the velocity model. Prestack migration is performed to generate a certain common image gather (CIG) such that the quality of velocity model can be assessed. As shown in Figure 1.6, surface-offset CIG (SOCIG) is built by gathering migration images from each source-receiver offset. The flatness of images is conventionally quantified by the stacking power condition (or semblance criterion Al-Yahya, 1989; Plessix *et al.*, 2000; Soubaras et Gratacos, 2007), which sums the image across offset. Although it attributes a maximum power to the correct model, the objective function is highly oscillatory due to the constructive interference of unrelated images at different offsets, leading to anomalously high powers for incorrect velocities (Chauris et Noble, 2001). This is a similar case with conventional FWI that the inversion can be easily trapped into local minima of the non-convex L2 norm-based misfit function (See Section 1.3.3). To overcome this issue, Symes et Carazzone (1991) propose to locally measure the derivative of image with respect to the offset in an automated way (differential semblance optimization, DSO), and they are able to show, in the 1D case, the DSO objective function is essentially quadratic, a very good property for local optimization method to find the global minimum (Symes, 1999). Extensions to 2D case have been performed by Chauris et Noble (2001); Mulder et Ten Kroode (2002); Shen (2004); Albertin *et al.* (2006); Shen et Symes (2008). Their successes confirm that this property still holds in realistic applications.

One shortage of SOCIG is that, when the velocity model is inaccurate, the gather may not preserve all the reflection energy after migration, especially when the offset range is limited. One solution is to extend the image space such that unfocused energy can be preserved in the non-physical dimensions of the extended image (the so-called depth-oriented CIG), and then penalize this energy to update the velocity model. Remind that in wave equation-based migration the image is computed by the pixel-to-pixel zero-lag cross-correlation between the source and receiver wavefields (Claerbout, 1985), cross-correlation with temporal lag and/or spatial shift can be used to compute the extended images (Rickett et Sava, 2002; Sava et Biondi, 2004; Sava et Fomel, 2006; Yang et Sava, 2011; Sava et Vasconcelos, 2011). Using the temporal lag consumes a small amount of memory space, whereas using the spatial shift is more consuming as it can be associated to both vertical and lateral directions. However, because it can bring information from

the local slope of events, the lateral shift is often considered in practice leading to the so-called subsurface-offset CIG (SSCIG, Shen *et al.*, 2008; Lameloise *et al.*, 2015).

Alternatively, depth-oriented CIG can be formulated in the angle domain (ADCIG), which relates the images to the open angle between the wavenumber vectors of source and receiver wavefronts (Xu *et al.*, 2001; Chauris *et al.*, 2002; Brandsberg-Dahl *et al.*, 1999; Sava *et al.*, 2003; Biondi *et al.*, 2004; Biondi *et al.*, 2004). Therefore, the velocity update is computed by flattening the images across angles. Although both aim to flatten the images, ADCIG is superior to SOCIG because it can uniquely determine the velocity field (private communication with Pengliang Yang, Stolk *et al.*, 2005). On the other hand, Shen *et al.* (2008) prefers SSCIG to ADCIG from a numerical-condition point of view. Constructing ADCIG by ray-based migration is straightforward, as the angle can be easily deduced from the ray parameter that is evaluated during migration (Chauris *et al.*, 2002). However, in order to tackle complex media the wave equation-based migration should be applied, and the resulting SOCIG or SSCIG can be converted to ADCIG via the Radon transform (Sava *et al.*, 2003; Silvestrov *et al.*, 2016, etc).

While successful applications have been reported, MVA suffers from biases in the objective function and gradient artifacts caused by limited acquisition geometry and complex structures (Fei *et al.*, 2010; Mulder, 2014). Proposed strategies to compensate for this illumination issue include Shen *et al.* (2011); Yang *et al.* (2013) etc. Lameloise *et al.* (2015) propose quantitative migration by introducing Hessian matrix to mitigate the gradient artifacts. Recent advances of MVA deal with combining transmitted waves in MVA for more constraints on near-surface velocities (Lameloise, 2015), using multiples for higher resolution (Berkhout *et al.*, 2015; Cocher *et al.*, 2015), and replacing the migration process by linearized inversion (inversion velocity analysis) which can also mitigate the gradient artifacts for the velocity update (Liu *et al.*, 2014b; Chauris *et al.*, 2015).

Compared with FWI, the main advantage of MVA is its stable convergence to reliable velocity models. On the other hand, FWI has its reputation of high-resolution results. Nonetheless, researchers have investigated the possibility to combine two methods, hoping to reach high resolution in a robust way (e.g. Biondi *et al.*, 2013; Fleury *et al.*, 2012; Allemand *et al.*, 2015; Santos *et al.*, 2016). Among them, Biondi *et al.* (2013, 2014) proposed an objective function which combines the L2 norms of data difference (FWI-related) and image unfocusing in SSCIG (MVA-related), such that the true velocity model can be built even if a poor initial model is used.

1.3.3 Full waveform inversion (FWI)

Full waveform inversion was first introduced by A. Tarantola and P. Lailly in 1980s (Tarantola, 1984; Lailly, 1984; Tarantola, 1986). They showed that the first iteration of FWI resembles a migration process that produces reflectivity images (Devaney, 1984; Wu *et al.*, 1987; Mora, 1989). Due to the restriction to very short offsets at that time,

however, they were unable to update the long wavelengths of the model (Gauthier *et al.*, 1986; Mora, 1988; Jannane *et al.*, 1989; Crase *et al.*, 1990; Pica *et al.*, 1990). In 1990s, the computational load of FWI was reduced by Pratt's group (Pratt et Worthington, 1990; Pratt *et al.*, 1998; Pratt, 1999) by reformulating FWI in the frequency domain. They pointed out that a reduced set of frequencies is already sufficient to provide a non-biased image of the subsurface (Sirgue et Pratt, 2004; Sirgue, 2006). For robustness, they applied FWI to cross-well data to favor transmitted waves (Pratt *et al.*, 1996; Pratt et Shipp, 1999). With the development of wide-aperture/long-offset acquisition geometries, intermediate wavenumbers of the subsurface can be imaged (Neves et Singh, 1996; Shipp et Singh, 2002; Sirgue *et al.*, 2007; Wang *et al.*, 2015), making this method unlike others: in principle, it does not assume scale separation between velocity background and high-wavenumber reflectivity. On the other hand, as the dense arrays of receivers provide much larger data volumes than ever, the automatic misfit evaluation in FWI outstrips the cumbersome picking step in traveltime tomography. Thanks to the development of computational infrastructures, we can implement more large-size FWI jobs on clusters, and since 2006 we have witnessed a rapid growth in the number of citations in the literature (Figure 1.7), from geophysical prospecting (e.g. Ravaut *et al.*, 2004; Malinowski et Operto, 2008; Plessix, 2009; Sears *et al.*, 2010; Vigh *et al.*, 2010; Prioux *et al.*, 2011; Plessix *et al.*, 2012; Etienne *et al.*, 2012; Brossier *et al.*, 2013; Gholami *et al.*, 2013a; Warner *et al.*, 2013; Vigh *et al.*, 2014; Operto *et al.*, 2015; Amestoy *et al.*, 2015; Wellington *et al.*, 2015) to crustal and lithospheric seismology (e.g. Operto *et al.*, 2006; Brenders et Pratt, 2007; Fichtner *et al.*, 2010; Beller *et al.*, 2014). For a nice overview of FWI in exploration geophysics, I refer to Virieux et Operto (2009).

FWI can be implemented either in the time domain or in the frequency domain. In the time domain, the elapsed time is reduced by distributing all shot gathers over the processors of parallel computers. A second level of parallelism can be used for domain decomposition of the computational mesh (Etienne *et al.*, 2010). Nonetheless, such cost can be reduced, at the expense of convergence rate and quality of results, by limiting the number of shots with random selection during FWI iterations (van Leeuwen et Herrmann, 2012; Warner *et al.*, 2013), or deterministic or stochastic source encoding (Vigh et Starr, 2008; Krebs *et al.*, 2009; Ben Hadj Ali *et al.*, 2011; Bansal *et al.*, 2013; Castellanos *et al.*, 2015). Alternatively, FWI can be performed in the frequency domain, in which case the computational cost is proportional to the number of discrete frequencies involved in the inversion. The wave equation transfers to the Helmholtz equation in the frequency domain, which can be solved by direct or iterative solvers (Demmel, 1997). The drawbacks of direct solvers, i.e. the poor scalability and significant in-core memory requirement, can be mitigated by new numerical methods such as block low-rank approximation (Amestoy *et al.*, 2015). However, as we move to 3D geometries, the direct solver becomes very resource demanding and we have to apply iterative solvers (Erlangga et Herrmann, 2008; Plessix, 2009; Li *et al.*, 2014; Hamitou *et al.*, 2015). Hybrid approaches are also promoted to reduce the cost, in which the discrete frequencies are extracted from time-domain seismograms via discrete Fourier transform (Nihei et Li, 2007; Sirgue *et al.*, 2008; Brossier *et al.*, 2014). These three strategies have been recently reviewed by Pajot *et al.* (2014), who point out the scalability issue remains for the direct and iterative

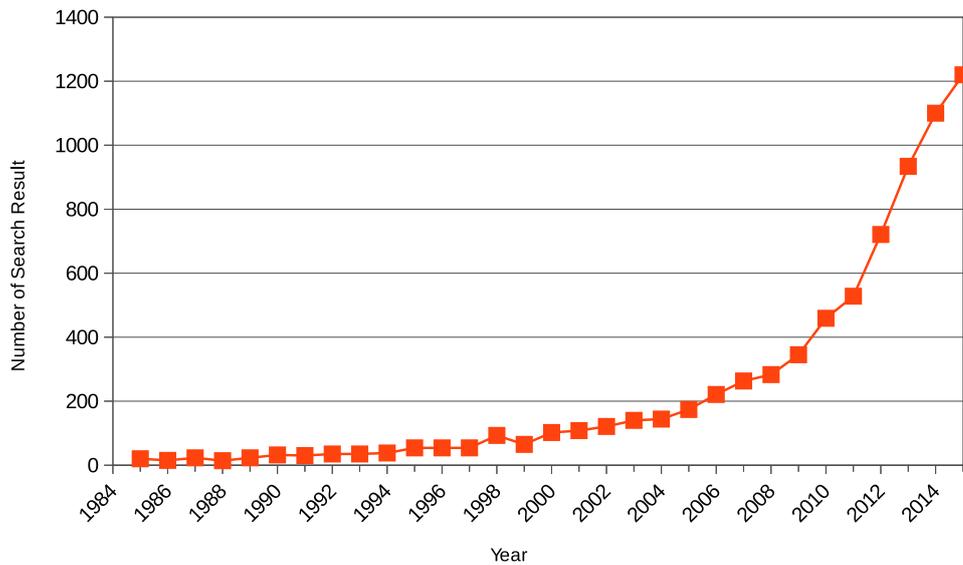


Figure 1.7: Increasing citations of FWI in the literature (from Google Scholar, keyword “full waveform inversion”, 13/05/2016). The numbers include publications, presentations, reports etc., hence give an upper limit of published works in each year. Proposed in 1984, FWI was not widely accepted until 2006, over 200 citations was reached thanks to the upgraded computing hardware. The curve has increased for five times after 9 years. By the end of 2016, a number around 1300 could be anticipated.

solvers. [Brossier *et al.* \(2014\)](#) compare time-domain FWI and frequency-domain FWI based on time-domain modeling, and show a same temporal complexity for the gradient computation embedded in the two methods.

Successful applications rely on how we simulate the wavefield propagation in the medium. Acoustic modeling is firstly tried because the acoustic approximation makes the cost of forward problem affordable and allows to focus the inversion on the estimation of the P-wave velocity (i.e., the first-order parameter in seismic imaging, [Gauthier *et al.*, 1986](#); [Pica *et al.*, 1990](#); [Mulder et Plessix, 2008](#)). However, elastic FWI is more desirable for applications that detect fluids, gasses, high velocity contrasts ([Barnes et Charara, 2009](#)) with the capacity to probably model the AVO effect, converted waves etc. [Plessix et Pérez Solano \(2015\)](#) also highlighted the artifacts generated by the acoustic approximation in the presence of carbonate layers. [Shipp et Singh \(2002\)](#) reconstruct P-wave velocity while S-wave velocity is deduced via an empirical relation, and [Sears *et al.* \(2008, 2010\)](#) used multi-component data to reconstruct both parameters. Based on Born and Rytov approximations, [Gélis *et al.* \(2007\)](#) highlight the dramatic footprint of the surface waves and thus they only involve body waves of short offsets in the early stage of inversion. Another issue for elastic FWI is the cycle skipping problem related to S waves. Because the wavelengths of S waves are shorter than P waves, a good initial model for acoustic FWI may be not accurate enough for elastic FWI to properly match the S waves ([Brossier *et al.*, 2009b](#)). To mitigate this problem, [Masoni *et al.* \(2014\)](#) and

Pérez Solano *et al.* (2014) measure the data misfit in alternative domains (ω - p or ω - k) for more convex misfit functions. Nevertheless, acoustic FWI is more considered in 3D geometries mainly for a lower computational price (Sirgue *et al.*, 2008; Ben Hadj Ali *et al.*, 2008; Vigh et Starr, 2008; Warner *et al.*, 2008; Operto *et al.*, 2015).

Recently, more researchers have taken attenuation into consideration as it can provide invaluable inferences on reservoir properties as porosity and fluid saturation (Ribodetti et Hanyga, 2004; Plessix, 2006a; Malinowski *et al.*, 2011; Prioux *et al.*, 2011; Kunert *et al.*, 2016; Plessix, 2016; Yang *et al.*, 2016). In a synthetic case, Kurzmann *et al.* (2013) show that neglecting attenuation can lead to unreliable reconstructions of the velocity model. Belahi *et al.* (2016) point out that a smooth attenuation background is only sufficient for inverting long-offset refracted waves. The data at intermediate offsets, close to the critical angle, depends very much on the attenuation contrast in the models. In this situation, it is necessary to update both the velocity and attenuation models to properly account for critical reflections.

Multiparameter FWI can significantly improve the data fit; however, this may be achieved by an inappropriate mapping from one parameter update into another: the trade-off issue between different classes of parameters. An example is given in Figure 1.8, which illustrates the trade-offs between P-wave velocity and anisotropic parameter. Before implementing inversion, potential trade-offs can be revealed by analyzing the radiation pattern (or diffraction pattern in some studies). Besides, the singular value decomposition of the Hessian matrix also brings information for trade-off analysis (Operto *et al.*, 2013). From the methodological point of view, an optimal set of parameters should be defined before inversion such that the trade-offs are kept under control at the same time the imaging resolution is preserved. In the acoustic VTI case, a parameterization of NMO velocity and anisotropic parameter η (or vertical and horizontal velocities) are often chosen to have better decoupling due to their distinguished influences on reflected waves and diving waves, respectively. Then, the anisotropic parameter δ is used as a secondary parameter to compensate for the shortcomings that the true earth is represented by the acoustic model (the so-called “garbage parameter”, Plessix et Cao, 2011; Gholami *et al.*, 2013b,a; Alkhalifah et Plessix, 2014; Stopin *et al.*, 2014). This parameterization seems to be still useful for elastic FWI as the S-wave velocity has a minor influence on the P-wave velocity inversion, especially in the marine case (Prioux *et al.*, 2013b; Alkhalifah et Guitton, 2016). In contrast, attenuation has a strong trade-off with the velocity parameters when the dispersion is not taken into account (Hak et Mulder, 2010). In this case, we can simultaneous reconstruct the two parameters (Plessix, 2016) to avoid overfitting of the data caused by monoparameter inversion, and the second-order Hessian matrix may be required to decouple the parameters.

Perhaps the most severe problem in FWI is the cycle skipping of observed and modeled data (Figure 1.9). Many efforts have been taken to attack this problem. Multiscale FWI is investigated (Bunks *et al.*, 1995; Sirgue et Pratt, 2004), the idea of which is to start by inverting lower frequencies and progressively introduce higher frequencies in the inversion (Kolb *et al.*, 1986; Virieux *et al.*, 2011). However, current acquisition techniques still do not allow to access very low frequencies (typically below 2 Hz), and prevent

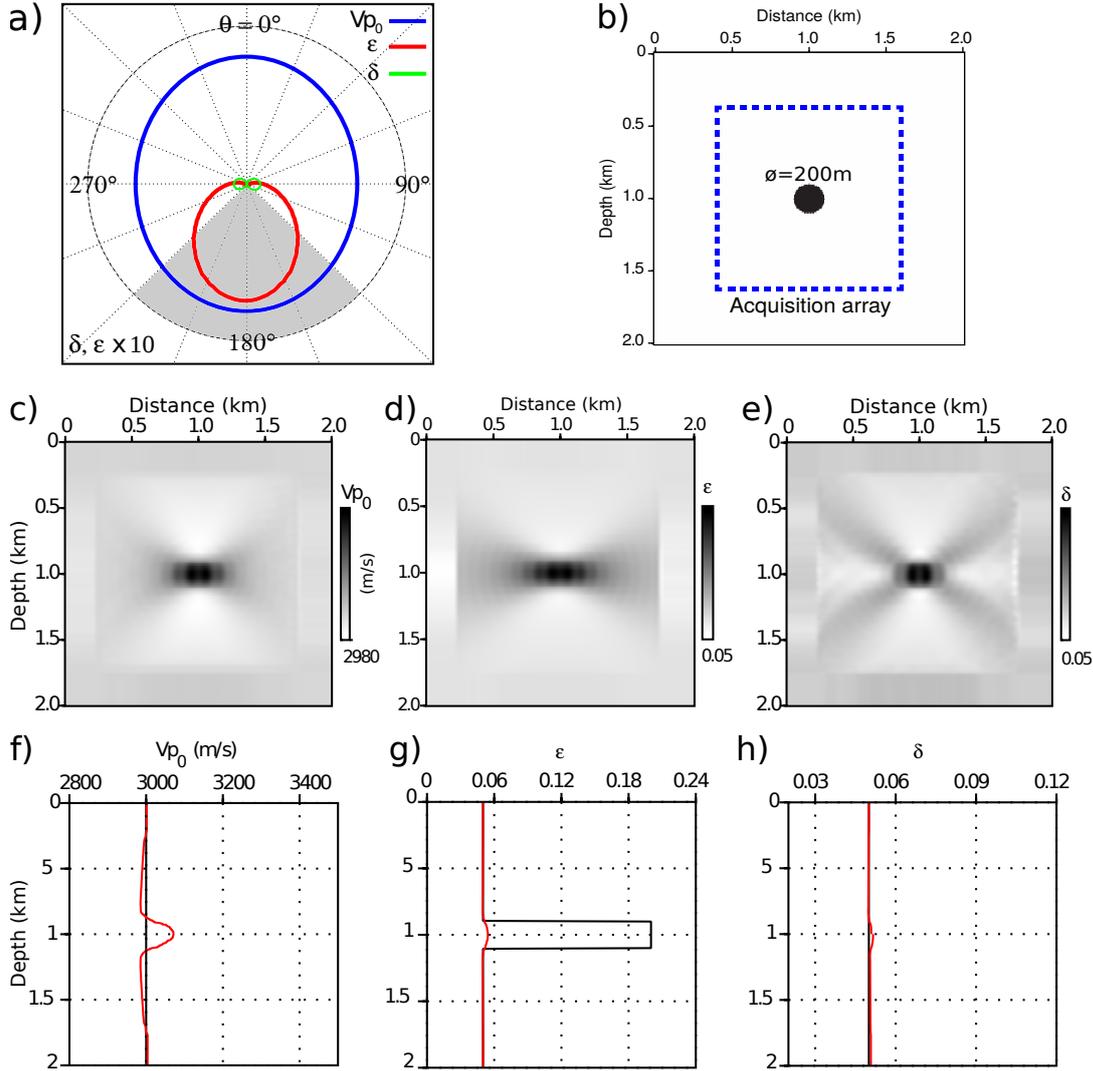


Figure 1.8: Synthetic example of trade-off for acoustic VTI multiparameter FWI (adapted from [Gholami *et al.*, 2013b](#)). (a) Diffraction angle-dependent amplitudes of scattered waves with respect to a single diffractor in the medium caused by inclusions of vertical P-wave velocity V_{P_0} , anisotropic Thomsen parameters ϵ and δ , respectively (scaled amplitudes for ϵ and δ scattering), carried out in the frame of ray+Born approximation ([Forgues et Lambaré, 1997](#)). The pattern reveals the trade-off between V_{P_0} and ϵ for large angles (gray area). (b) Homogeneous true models except a ball-shape anomaly in the ϵ parameter. Starting from homogeneous models, multiparameter FWI creates heterogeneity in all three parameters (c-e). The vertical sections (f-h) of true models (black lines) and recovered models (red lines) at 1 km distance shows the trade-off between ϵ and V_{P_0} . Note how the true ϵ perturbation (g) is translated to V_{P_0} (f), leading to erroneous reconstructions. Due to negligible sensitivity to δ perturbation in this parameterization (a), the δ heterogeneity in (e,h) is very weak.

multiscale inversion from using poor initial models. Alternatively, Laplace-domain and Laplace-Fourier domain inversions are proposed to boost the low frequencies or early arrivals to make the inversion more robust (Shin *et al.*, 2008, 2009; Pyun *et al.*, 2008, 2011). Meanwhile, researchers also try to modify the gradient expression to enhance the low-wavenumber components in the gradient, by different approaches such as wavefield decomposition (Wang *et al.*, 2013a; Tang *et al.*, 2013; Wu *et al.*, 2016) and scattering angle-based filtering (Alkhalifah, 2015). On the other hand, many investigations are also devoted to reducing the nonlinearity in a more automatic way. A number of misfit functionals are proposed, such as those based L1 or Huber norm (Brossier *et al.*, 2009a; Ha *et al.*, 2009; Brossier *et al.*, 2010), cross-correlation (Luo *et al.*, 1991; van Leeuwen *et al.*, 2010; Choi *et al.*, 2016), deconvolution (Luo *et al.*, 2011; Warner *et al.*, 2014; Guasch *et al.*, 2014; Warner *et al.*, 2015), envelope of wavelets (Fichtner *et al.*, 2008; Bözdag *et al.*, 2011; Luo *et al.*, 2015), unwrapped phase (Alkhalifah *et al.*, 2012), dynamic image warping (Ma *et al.*, 2013), registration-guided norm (Baek *et al.*, 2014) and optimal transport (Métivier *et al.*, 2015). Using alternative misfit functionals is still a vivid topic and new ideas are continuously introduced to the literature. Nevertheless, as Plessix (2012) has pointed out, an automatic and also robust FWI strategy relies on a common effort in both acquisition innovation and methodological development.

1.3.4 Reflection waveform inversion (RWI)

As depicted in Figure 1.3, the FWI gradient lacks large-wavelength components in deep zones. To build the smooth velocity model, reflection waveform inversion was proposed to carry out the transmission kernels from reflected waves by considering the migrated images explicitly (Figure 1.4). Moreover, the separation of scales in the model description allows to adjust the depths of migrated images, such that the reflected waves at near-zero offsets are perfectly matched during velocity update (Zhou *et al.*, 2012; Xu *et al.*, 2012)³. Consequently, the associated misfit function has less number of local minima than that of FWI. An example will be given in Section 2.2.

Adjusting the image depths can be achieved by re-performing migration in smooth models, which leads to a cycle workflow that alternates between migration and RWI velocity update. However, the workflow could take a long time to converge, because the temporal complexity of migration is non negligible, especially when true-amplitude migration is used for perfect data match at near offsets. To mitigate this issue, Brossier *et al.* (2015) suggests to adjust the depths based on the vertical time-depth relation (Plessix, 2013), such that the depths are automatically adjusted according to the velocity update without repetition of the migration process.

RWI is inspired by the idea of migration-based traveltimes (MBTT) inversion (Chavent *et al.*, 1994; Chavent, 1996; Plessix *et al.*, 1999; Clément *et al.*, 2001; Tcheverda *et al.*,

³At that time the name of this method was “Reflection FWI”. But later the letter “F” was abandoned as R.-É. Plessix argued that the transmitted waves are not included: it’s not full waveform.

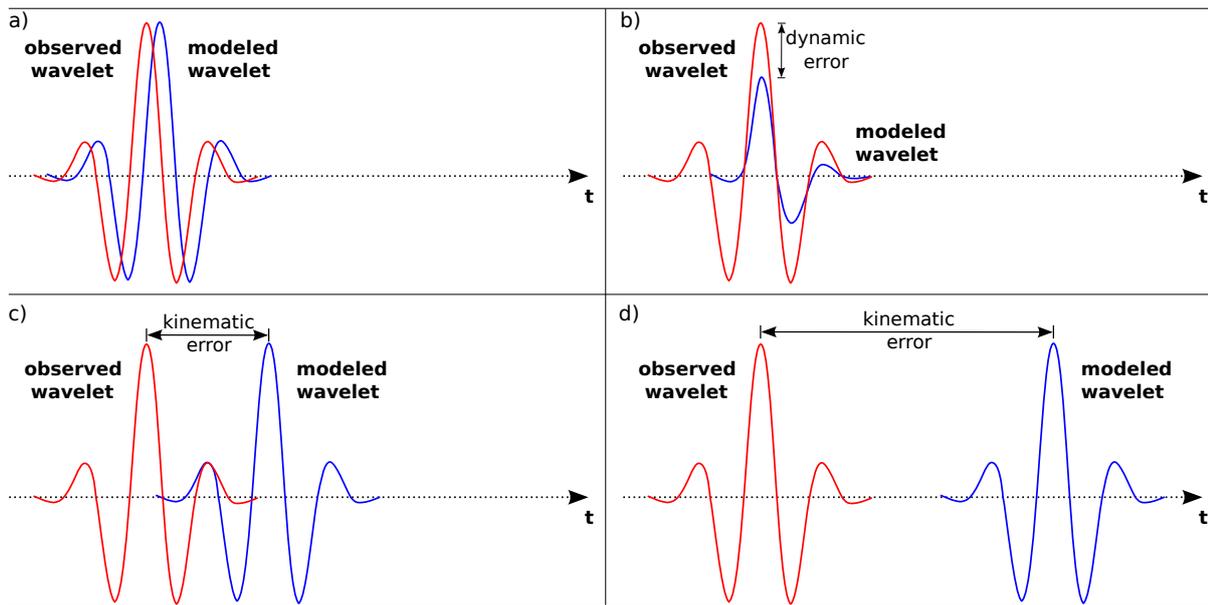


Figure 1.9: Sketch to explain cycle skipping phenomena. (a) No cycle skipping problem: the initial velocity model is accurate enough in terms of kinematics such that the modeled wavelet is within half cycle of the observed wavelet. However, it is often not the case in practice (b)-(d). When the modeled wavelet is slightly out of half cycle of the observed one, the inversion could reshape the wavelet such that the amplitude-related misfit is reduced, which is often interpreted as a dynamic error (b). When the modeled wavelet departs further from the observed one, their side slopes may constructively correlated corresponding to a local minimum of the misfit function (c). In (d), the modeled wavelet has a large time shift from the observed one. The inversion often cannot recognize their relationship and would be lost its direction in the local plateau of the misfit function.

2016) that also uses migrated images for velocity inversion. On the other hand, RWI shares some similarities with MVA (Xu *et al.*, 2012) and promises higher resolution (Alkhalifah *et al.*, 2015). Liu *et al.* (2014a) combine RWI and MVA such that additional velocity update can be gained from extended images. Wang *et al.* (2013b) reformulate RWI in the frequency domain and illustrate the importance of low frequencies – a similar property to FWI. While Staal (2015) and Guo *et al.* (2016) extend RWI to multiple reflections and elastic case, respectively, Wu *et al.* (2015) and Zhou *et al.* (2015) consider diving waves in RWI in order to better constrain the velocity inversion at shallow depths. In specific, Zhou *et al.* (2015) propose to parameterize the subsurface by velocity and impedance for velocity macromodel building and reflectivity imaging, respectively, which forces the scale separation between the two subsurface representations.

1.3.5 Summary and motivation of this study

A summary of aforementioned velocity estimation methods is given in Table 1.3. Unlike the others, FWI does not assume scale separation, which makes it highly nonlinear and often requires sufficiently accurate initial models in a sense that the data can be reproduced within an error of half cycle. However, the interest in FWI grows steadily because its high-resolution imaging power and automated process give the possibility for interpreters to better pinpoint reservoirs, salt body, faults etc. Therefore, investigations are more towards initial model building methods. Without considering other possibilities, initial models could be built by traveltimes tomography, MVA or RWI. While the effectiveness of traveltimes tomography is case-dependent, MVA often requires more computational resources consumed by the migration step and extended imaging conditions (Duveneck, 2013). Although being considered less immune to kinematic errors than MVA (Alkhalifah et Wu, 2015), RWI is a new way to build initial velocity models. The advantage is its higher resolution that FWI may demand to start with (tomography may not satisfy this resolution-related condition), and its cheap implementation than MVA because no extended-domain images are built. Considering both resolution and computational expenses, RWI is the best among the three methods to build the initial model for FWI.

One shortage of RWI is the lack of contribution associated to transmitted waves, or specifically the diving and refracted waves in the case of surface acquisition. With the development of very long-offset/wide-azimuth acquisition geometries,⁴ there is an abundance of transmitted energy recorded in seismic traces, and any high-resolution velocity model building methods should take benefit from that. On the other hand, despite this increasing length of acquisition aperture, FWI remains unqualified for very deep target imaging which makes mandatory an prior implementation of RWI. However, it is also crucial to keep using diving waves to recover the large wavelengths of shallow zones as accurately as possible. Therefore, the motivation of this PhD study is to add these waves into RWI for more robust velocity macromodel building while the computational price is kept same as RWI.

1.4 Contribution of this work and thesis outline

Main contributions:

1. Diving waves (or early arrivals) and reflected waves are combined in a unified formulation for velocity macromodel building (namely JFWI). All low-wavenumber components associated to diving and reflected waves are included in the sensitivity kernel, whereas high-wavenumber migration isochrones are mitigated by an explicit data separation of diving and reflected waves, and by choosing suitable

⁴A recent example is the dual-coil multivessel method of WestenGeco that can potentially provide ultralong offsets up to 29 km! (Vigh *et al.*, 2013)

Table 1.3: A comparison among different velocity estimation methods.

	Separating scales	Human intervention	Linearity	Output resolution
Ray-based tomography	Yes	Manual picking	Quasilinear	Low
MVA	Yes	No	Quasilinear (except stacking-power MVA)	Low
FWI	No	No	Highly nonlinear	High
RWI	Yes	No	Nonlinear	Low
JFWI as a variant RWI	Yes	Prior data separation	Nonlinear	Low

parameterization of the subsurface as well as a stacking of sensitivity kernels over source-receiver couples. The computational time to evaluate the gradient is twice of the conventional FWI gradient while the memory cost is the same.

2. Based on scale separation, an integrated workflow is proposed which alternates an impedance inversion and JFWI, using short-spread reflection data and long-offset diving+reflected waves, respectively. The velocity model resulted from JFWI can be used as the initial model of FWI for the purpose of broadband reconstruction.
3. The synthetic example shows that like RWI, JFWI ensures no cycle skipping at near offsets, and provides sensitivity to low wavenumbers at depths greater than the penetration depth of diving waves. However, the advantage of JFWI over RWI is the improved image quality at shallows depth, due to considering the complementary information carried by diving waves, which translates to better reconstruction in terms of kinematics at great depths.
4. The real case study shows that for successful applications of JFWI a) the diving and reflected waves should be easily separable from each other; b) the cycle skipping problem should be mitigated. Otherwise, the inversion would converge to local minima; c) low-wavenumber content should be injected into the velocity model at the early stage of inversion and before resolving the high-wavenumber content.

Organization of the thesis:

- Chapter 2 will analyze the FWI and RWI gradient by vectorial decompositions in simple circumstances, by which I shall explain why RWI has better sensitivity to deep targets than FWI, and why RWI is able to reduce the cycle skipping problem when reflected waves are used. In addition, the reliance of RWI on frequencies and acquisition aperture will also be studied in a simple test, which can serve as a guideline for realistic applications.

- The formulations of FWI, RWI and JFWI will be systematically studied in Chapter 3, from which the reader will see how diving waves can be combined into RWI through a unified formulation. To further fulfill scale separation, I shall propose to parameterize the subsurface by velocity and impedance, rather than the conventional velocity and density. The benefit of this choice will be explained by analyzing the radiation patterns in the framework of single scattering formulation, and will be illustrated in a two-layer example. Based on this parameterization, a cycle workflow will be proposed, which alternates between high-wavenumber impedance inversion and velocity macromodel building, using near-offset reflected waves and large-offset diving and reflected waves, respectively.

In addition, Chapter 3 will also show a numerical example of JFWI on a synthetic Valhall model. This model is composed of nearly flat interfaces superimposed on the smooth background, which helps separate velocity and impedance in terms of wavenumbers, as is assumed by the approach. Moreover, the flat interfaces generate primary reflections that can be easily separated from the direct/diving waves within the considered acquisition aperture. In the preprocessing stage, I only perform an offset-dependent time window for data separation while during inversion, I apply layer-stripping and data weighting to increase the robustness of the approach and enhance the illumination at depths. Vertical sections of the results will show the realization of the scale separation between velocity and impedance. The reconstructed velocity model can be used as the initial model for a sequential FWI such that the intermediate-to-high wavenumbers are further injected for broadband reconstruction of the subsurface.

- With the confidence gained from the synthetic study, I shall further study the promises and pitfalls of JFWI with a real-data case study (Chapter 4). Three points are noted:
 - a) Because the seismic modeling is performed in the time domain, I omit the attenuation effect for high numerical efficiency. By using an existing initial model, the attenuation effect is illustrated by classical FWI sequentially implemented in the time and frequency domains, which tells us that the attenuation has limited effect for frequencies below 5.6 Hz;
 - b) The second point is the nonlinearity caused by the interference of early arrivals and critical reflections after low-pass filtering. In order to mitigate this, I feed JFWI only short-to-intermediate offsets and then extend to longer offsets during the sequential FWI process;
 - c) By using different initial models with increasing low-wavenumber contents, the importance of low wavenumbers is highlighted to recover the velocity model: either the initial model should contain sufficient low wavenumbers, either longer offsets should be used at the early stage of JFWI. As the cycle-skipping issue prevents us from using long offsets, I need a better initial model in the real-data case than the one in the synthetic case.

- The final chapter will draw conclusions and propose several options for future study.

Chapter 3 is extracted from the publication:

ZHOU, W., BROSSIER, R., OPERTO, S. et VIRIEUX, J. (2015). Full waveform inversion of diving and reflected waves for velocity model building with impedance inversion based on scale separation. *Geophysical Journal International*, **202**:1535–1554

Chapter 4 is extracted from the manuscript:

ZHOU, W., BROSSIER, R., OPERTO, S., VIRIEUX, J. et YANG, P. (2016). Joint full waveform inversion of early arrivals and short-spread reflections: a 2D ocean-bottom-cable study including gas cloud effects. *To be submitted*.

Chapter 2

From FWI to RWI – Mitigation of cycle skipping and approaching to low wavenumbers at great depths

A lack of information cannot be remedied by any mathematical trickery!

— Cornelius Lanczos (Quoted by [Kirsch, 2011](#))

In this chapter, I shall establish the theoretical foundation for the proposed method of the thesis, which will be fully described in the next chapter. Without essential investigations of the methodology, the advantage of RWI over FWI and the additional benefit of the proposed method would not be manifest. Meanwhile, this chapter also serves as a guideline for realistic applications of RWI as well as the new method I propose.

The cycle skipping problem is always encountered in FWI applications, as seismic traces are oscillatory signals. It generates local minima for the misfit function and leads to erroneous subsurface models. One usual source of cycle skipping can be attributed to inaccurate reproductions of the reflection phases. How does RWI encounter this problem? In many surface experiments, the offsets between sources and receivers are limited to a certain range, which forbids a direct implementation of FWI to recover deep targets. Is it still the case for RWI? As will be shown that RWI naturally images low wavenumbers at depth, do we still need long offsets or low-frequency data for this imaging? In this chapter, these questions will be answered by theoretical studies as well as numerical illustrations.

2.1 Forward modeling

To promote the analysis related to FWI and RWI, I shall briefly review the acoustic isotropic wave equation that is used in this thesis. The first-order partial differential equations (PDEs) are written as

$$\partial_t v_x = b \partial_x p + b f_x, \quad (2.1)$$

$$\partial_t v_y = b \partial_y p + b f_y, \quad (2.2)$$

$$\partial_t v_z = b \partial_z p + b f_z, \quad (2.3)$$

$$\partial_t p = \kappa (\partial_x v_x + \partial_y v_y + \partial_z v_z) + s \quad (2.4)$$

where v_x, v_y, v_z denote the three components of the vectorial velocity wavefield \vec{v} , p the scalar pressure wavefield, f_x, f_y, f_z the components of the vectorial velocity sources, and s the stress source, respectively. Two model parameters are considered: the buoyancy b (the inverse of density ρ) and the bulk modulus κ . The compressional wavespeed V_P and acoustic impedance I_P can be deduced by $V_P = \sqrt{\kappa b}$ and $I_P = \sqrt{\kappa/b}$. In this thesis, the temporal partial derivative (∂_t) is computed by an explicit second-order leap-frog scheme, and the spatial derivatives ($\partial_x, \partial_y, \partial_z$) are discretized by staggered-grid fourth-order finite-difference mesh (Levander, 1988; Fornberg, 1988). In this thesis, I consider explosive sources, so $f_x = f_y = f_z = 0$.

In the frequency domain, on the other hand, I only consider the pressure field p . The associated second-order PDE can be obtained by taking the temporal derivatives of Equation (2.4) and merging the four equations together. The Fourier transform with respect to time transfers this PDE into a linear system which is also known as the Helmholtz equation:

$$\omega^2 p + \rho V_P^2 \nabla \cdot \left(\frac{1}{\rho} \nabla p \right) = s \quad (2.5)$$

where ω denotes the circular frequency, $\nabla = (\partial_x, \partial_y, \partial_z)$ the gradient operator.

No matter which domain is considered, the two systems (2.1)-(2.4) and (2.5) can be summarized in a unified formalism by using the modeling operator notation B :

$$B(m)u = W, \quad (2.6)$$

where m denotes the models of all physical parameters (for instance V_P and ρ in the acoustic isotropic case, or many others if we consider anisotropy, elastic or viscous inversion). The operator B , field u and source term W include different quantities depending on which system is considered (Table 2.1). The wavefield u are sampled at sparse receiver locations implying a linear matrix R that projects the wavefield into the data space:

$$d_{cal} = Ru, \quad (2.7)$$

where d_{cal} represents the synthetic data computed in the proposed model m during inversion. Conversely, its transpose R^T , reallocates the seismic traces at positions where they are recorded.

Table 2.1: Quantities that are represented by B , u and W in Equation (2.6)

	Time domain (2.1)-(2.4)	Frequency domain (2.5)
B	spatial discretization stencil, finite-difference approximation of ∂_t	spatial discretization stencil, ω
u	v_x, v_y, v_z and p	p
W	f_x, f_y, f_z and s	s

2.2 Formulations and the cycle skipping issue

2.2.1 FWI as a least-squares linearized optimization

Classical FWI attempts to minimize the following object function:

$$C_{FWI}(m) = \frac{1}{2} \|d_{obs} - d_{cal}(m)\|_2^2, \quad (2.8)$$

where $\|\cdot\|_2$ denotes the L2 norm of the mismatch between the observed data d_{obs} and the synthetic data $d_{cal}(m)$ computed in a proposed subsurface model m . It translates to the integration on the time window in the time-domain implementation, or implies the moduli of the mismatch vectors sampled by different frequencies in the frequency-domain implementation, plus a summation over shots and receivers. Due to its large size (10^4 – 10^6 unknowns in 2D and much more in 3D) and the non negligible time to compute d^{cal} , this inverse problem is generally solved by linearized iterative methods, which updates the subsurface model by using the gradient of C_{FWI} with respect to m . The formulation can be found by using the adjoint-state method (Plessix, 2006b). For sake of simplicity, the frequency-domain formulation is used here (Appendix 5):

$$\nabla C_{FWI}(m) = \Re \left\{ u_0^T \left(\frac{\partial B}{\partial m} \right)^T \lambda_0^* \right\}, \quad (2.9)$$

which is the real part (i.e. \Re) of the inner product of the transposed modeled wavefield u_0^T and the conjugate of the adjoint wavefield λ_0^* , weighted by the transposed partial derivative of the modeling operator B with respect to the model m . The modeled wavefield is the solution of the Helmholtz equation $B(m)u_0 = W$ with W the source function and the adjoint wavefield is the solution of the following adjoint-state equation

$$B(m)^\dagger \lambda_0 = R^T (d_{cal} - d_{obs})^* \quad (2.10)$$

with a virtual source term using the conjugate of the data residuals. The symbol \dagger denotes the adjoint operation (transpose conjugate). However, without consideration of absorbing layers the modeling operator is self-adjoint ($B(m)^\dagger = B(m)$). Therefore, the similar form of (2.10) and $B(m)u_0 = W$ shows that the same modeling engine can be used to solve both u_0 and λ_0 . In addition, the conjugate operation on the source term of

Equation (2.10) translates to the signal reversal in the time domain, implying that the time-domain computation of λ_0 should be performed in the reverse time order.

The gradient is further weighted by the Hessian matrix H in Newton-based optimization:

$$\Delta m = -H^{-1}\nabla C, \quad (2.11)$$

where the Hessian is the second-order information from the data, formulated as (Pratt *et al.*, 1998)

$$H(m) = \nabla^2 C(m) \quad (2.12)$$

$$= H_a + H_b \quad (2.13)$$

$$= \Re \{ J^T J^* \} + \Re \left\{ \left(\frac{\partial J^T}{\partial m^T} \right) (\Delta d^* \ \Delta d^* \ \dots \ \Delta d^*) \right\} \quad (2.14)$$

where $J = \nabla u_0 = \left(\frac{\partial u_{0i}}{\partial m_j} \right)$ denotes the sensitivity kernel (i.e. Jacobian matrix) and Δd the data residuals. The first term H_a is known as the approximate Hessian in Gauss-Newton methods, diagonally dominant and banded due to correlations of first-order scattered fields of a finite frequency band. The diagonal corresponding to spatial autocorrelations, which account for amplitude decay due to geometrical spreading. On the other hand, the off-diagonal corresponding to spatial crosscorrelations, which is nonzero due to the finite frequency content and limited acquisition sampling. Therefore, inverting H_a helps balance the gradient amplitudes at different depths by its diagonal terms (precondition), and sharpen the image by its the off-diagonal terms (deconvolution). In addition, an implicit summation over shots in H_a (not written above) also helps balance the contributions from different shot gathers.

In the second term H_b , the partial derivative matrix $\frac{\partial J^T}{\partial m^T} = \left(\frac{\partial^2 u_{0i}}{\partial m_j \partial m_k} \right)$ is related to the second-order scattering with respect to the parameters m_k and m_j . Considering this information allows to mitigate the cross-talk effect of two parameters of one and/or two classes. However, due to the second-order derivatives, H_b is more difficult to compute than H_a , and is generally neglected with the argument that it would have little influence when the residuals Δd are largely reduced. Nonetheless, investigations have manifest its key role in the discrimination of parameters (e.g. velocity–density, velocity–attenuation. Métivier *et al.*, 2015), especially at the early stage of the inversion. With the upgraded computing hardware the full Hessian has been accounted in the truncated Newton approach (Métivier *et al.*, 2014a).

The answer to the inverse problem (2.8) is not unique due to the lack of acquisition coverage and frequency content; two distinct models could result in a same level of data fitness (see an example in Chapter 4). Besides matching the observed data, model constraints can be added into Equation (2.8) to exclude unwanted models:

$$C_{Reg}(m) = \frac{1}{2} \|d_{obs} - d_{cal}(m)\|_2^2 + \|P(m)\|_2^2, \quad (2.15)$$

where P denotes a certain operator. In Tikhonov regularization, for example, $P(m)$ includes some spatial derivatives of m , such that rapid variations (or high wavenumbers)

are penalized in the model update (Tikhonov et Arsenin, 1977; Guitton, 2012; Esser et al., 2016; Liu et al., 2016). Alternatively, $P(m)$ can involve some prior information of the subsurface (geology, well-logs etc.), such that the result is close to this information and also geologically meaningful (Asnaashari et al., 2013; Vigh et al., 2015).

In this chapter, I shall omit the possibility of using model constraints to mitigate cycle skipping, as it is not the topic of this thesis. Also, the Hessian matrix will not be considered for a straightforward sampling analysis with regard to FWI and RWI gradients.

2.2.2 Velocity-depth ambiguity – Reflection data-induced cycle skipping

For deep-target imaging, reflected waves should be considered. However, the ambiguity arises since both the depths of subsurface interfaces and background velocities can determine the observed traveltimes of reflected waves; the influences of the two parameters may not be separable in the available offset window. Incorrect reflectivity depths, and/or inaccurate velocities result in a mismatch of the traveltimes, which cause the cycle skipping problem that has been widely encountered in many studies. In the following, I shall first point out that FWI tends to resolve interface depths before velocity building, causing cycle skipping problems and inaccurate velocity models. From a simple example, the advantage of RWI over FWI will then be highlighted.

Suppose the model we look for consists of two homogeneous layers. We observe the direct and reflected waves from a surface shot:

$$d_{obs} = d_{obs}^{dir} + d_{obs}^{refl}. \quad (2.16)$$

Starting from a homogeneous model $m^{(0)}$, the gradient expression (2.9) for the first iteration can be decomposed as

$$\nabla C_{FWI}(m^{(0)}) = u_0 \star \lambda_0^{dir} + u_0 \star \lambda_0^{refl} \approx u_0 \star \lambda_0^{refl} \quad (2.17)$$

where the transpose and conjugate operations, the partial derivatives and \Re are represented by \star for simplicity. Without loss of generality, I neglect the first term to highlight the cycle skipping caused by reflected waves (i.e. assuming a good knowledge of the surface). The second term, on the other hand, represents the high-wavenumber migration isochrone due to the small aperture angle made between u_0 and λ_0^{refl} , resulting reflectivity images in $m^{(1)}$ after the first iteration (migration mode of FWI). In the second iteration, these reflectivity images generate the scattered fields δu and $\delta \lambda^{refl}$ in the forward and adjoint modeling processes, respectively, leading to the following updated gradient expression that includes three terms:

$$\nabla C_{FWI}(m^{(1)}) = u_0 \star \lambda_0^{refl} + u_0 \star \delta \lambda^{refl} + \delta u \star \lambda_0^{refl}. \quad (2.18)$$

The first term still represents the high-wavenumber migration isochrone. The last two terms, on the other hand, represent the low-wavenumber update (see the sampling

analysis in Section 2.3, Figure 2.9). To assess the respective contributions to the gradient (2.18) of these three terms, let us assume the amplitude of u_0 is 1. The relative amplitude of each term can be estimated by using the reflection coefficient Z of the imaged reflector:

$u_0 \sim 1$	$\lambda_0^{refl} \sim Z$	$u_0 \star \lambda_0^{refl} \sim Z;$	(2.19)
$\delta u \sim Z$	$\delta \lambda^{refl} \sim Z^2$	$\delta u \star \lambda_0^{refl} \sim Z^2;$	
		$u_0 \star \delta \lambda^{refl} \sim Z^2.$	

For short-spread reflections, Z is very small (~ 0.1), meaning that the total contribution of $\delta u \star \lambda_0^{refl}$ and $u_0 \star \delta \lambda^{refl}$ is negligible compared with $u_0 \star \lambda_0^{refl}$ (10%). In other words, the gradient (2.18) is dominated by its high-wavenumber term $u_0 \star \lambda_0^{refl}$ that contributes to reflectivity imaging. Note that this conclusion does not generally hold for critical and postcritical reflections (Z tends to 1). Nevertheless, the postcritical reflection and refraction fields represented by δu or $\delta \lambda^{refl}$ will be treated as transmission field that is not considered by RWI in principle.

For the moment, let us still assume the dominance of $u_0 \star \lambda_0^{refl}$. Consequently, I can assume that the depth of the imaged reflectivity is fixed through iterations. This is fairly true. Indeed, experiences have shown that in realistic cases shifting the depths of imaged reflectors often requires more iterations than updating the background velocity.

Let us first suppose that the depth of the reflector is correctly determined in $m^{(1)}$. This implies that I have somehow well resolved the ambiguity between reflector depth and velocity, and the remaining job is to determine the velocity by involving the two terms $u_0 \star \delta \lambda^{refl}$ and $\delta u \star \lambda_0^{refl}$. However, it is not easy, as the sinusoidal nature of the seismic waves is translated to the oscillating behavior of the misfit function, and a gradient-based searching method would converge to a local minimum if the initial velocity is far from the true one (Figure 2.1). For example, with a velocity of 2300 m/s (Point a in the figure), the reflection phase is reproduced outside one dominant cycle of the observed phase (the black wiggle disclosing the blue part in Figure 2.1a). Because shifting the synthetic phase to earlier times makes a higher misfit value, the minimization of the misfit function would rather postpone the arrival times using a lower velocity (e.g. 2200 m/s). One may mitigate this issue by decreasing the frequency content of the data (usually by implementing a low-pass filter), such that the observed wavelet is wide enough to accept the synthetic wavelet. As a result, in the figure, Point b is inside the local attraction valley of the global minimum unlike Point a, and the iteration would converge to the true velocity. Note that the lower frequency also gives a less steep misfit function near the global minimum, which may slow down the convergence rate of the iteration. Therefore, higher frequencies can be re-used for the speedup (Bunks *et al.*, 1995). However, the success of this hierarchical scheme is case-dependent. For example, starting from Point c in the figure which is on the local plateau of the misfit function, FWI would fail to converge because the magnitude of the gradient is too small to update the model (zero derivatives at the local plateau).

Taking the velocity-depth ambiguity into consideration, one may image the reflector at incorrect depths in $m^{(1)}$ due to the erroneous initial velocity provided by $m^{(0)}$. The adjustment of the background velocity is thus hampered by these incorrect depths that

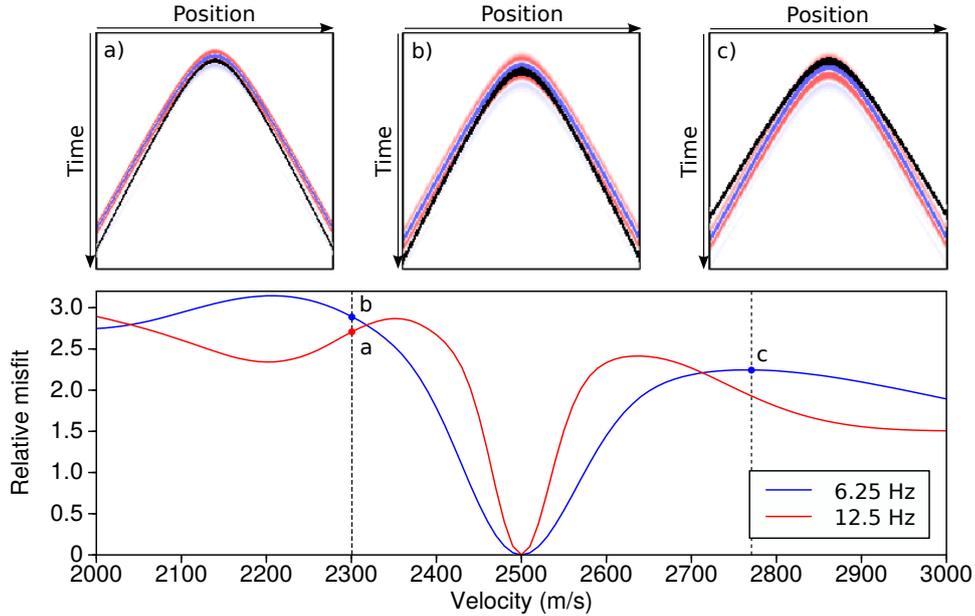


Figure 2.1: Illustration of reflection data-related cycle skipping (top) causing local minima in the L2 norm-defined FWI misfit function (bottom), due to inaccurate velocity background. The observed data are plotted with a blue-white-red color scale, superimposed by the synthetic data plotted with a variable area wiggle display (40% of opacity). The two data are in phase if the black area covers the blue part of the real data (top). The Ricker wavelet is used to generate the data, with central frequencies 6.25 Hz and 12.5 Hz, respectively (bottom). The true velocity equals 2500 m/s. The depth of the reflector is 500 m in the true and tested velocity models. When the tested velocity is inaccurate (Point a), the synthetic data are reproduced outside the dominant cycle of the observed data (note that the black area discloses the blue area); the minimization of the misfit function would propose an erroneous velocity value (e.g. 2200 m/s). Using lower frequencies may mitigate this problem but slows down the convergence rate. See text for more discussions.

are fixed during inversion (Figure 2.2). First of all, the misfit function does not show a global minimum at the true velocity, because the traveltimes deduced from incorrect depths and true velocity are not equal to the observed traveltimes. Second, the global minimum may be located at an erroneous velocity through a relatively good match of the observed phase; however, mismatches do exist at far offsets so that the global minimum is always above zero (Point a). In offset-domain MVA, such ambiguity is tackled by resolving the depths for each offset (roughly speaking computing $u_0 \star \lambda_0^{refl}$ for each offset), such that the deduced traveltimes are equal to the observed ones for all offsets. The consistency of these depths is then used to update the velocity. However, this solution often demands more computing resources.

In RWI, the velocity-depth ambiguity is accounted for by two folds: on the one hand, by assuming scale separation, the depth of the reflector can be adapted to the background velocity for a perfect match at short offsets; the mismatch at longer offsets, on the other

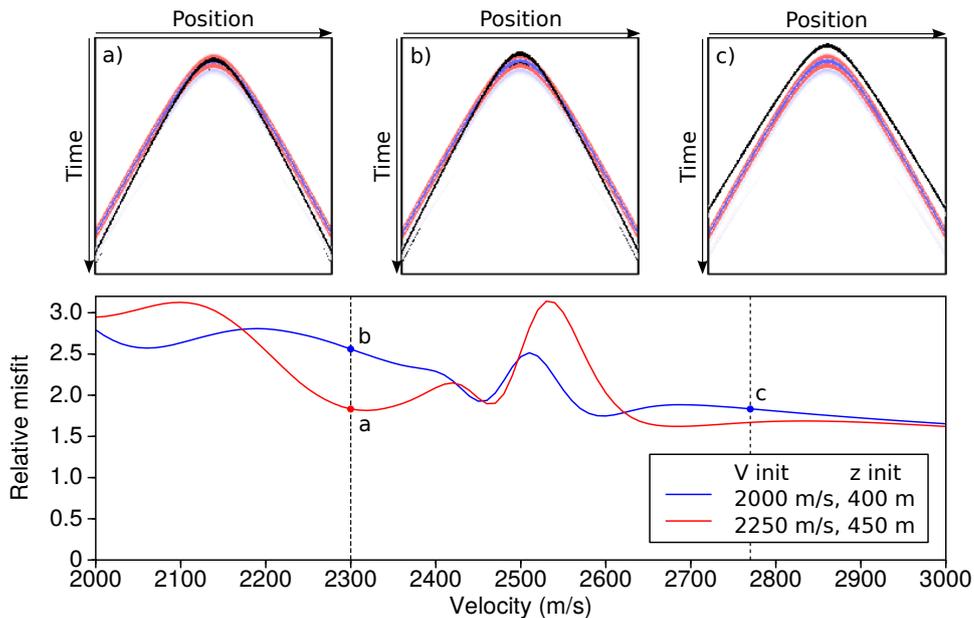


Figure 2.2: Illustration of reflection data-related cycle skipping (top) and non convexity of the L2 norm-defined FWI misfit function (bottom), due to inaccurate velocity background and incorrectly imaged reflector depth. The data are plotted in the same way as Figure 2.1, which are generated by using the Ricker wavelet with central frequency 12.5 Hz (top). The true velocity equals 2500 m/s. The depth of the reflector is 500 m. The depth of the imaged reflector is estimated by $z_{init} = v_{init}/t_0$ where t_0 denotes the vertical two-way traveltime of the reflected waves (bottom). The global minimum can be reached at an erroneous velocity due to a better match in short offsets (Point a). However, the normal moveout of the observed reflected waves is not correctly reproduced so that mismatches exist in longer offsets giving a nonzero global minimum, unlike in Figure 2.1. For the same tested velocity, a different initial velocity leads to a higher misfit value due to the mismatch at short offsets (Point b). This highlights that the initial velocity has a strong impact on the data fitness. Like in Figure 2.1, a local plateau can be found at Point c due to the complete mismatch at all offsets.

hand, can be used to update the velocity without creating additional reflectors, thus no additional ambiguity is introduced. In this way, the number of local minima is largely reduced. The remaining ones are attributed to cycle skipped traces at longer offsets depending on the frequency content. If lower frequencies are used, there should be less number of cycle skipped traces in the intermediate-to-far offsets.

2.2.3 RWI based on scale separation and mitigation of cycle skipping

Let m_0 denote the low-wavenumber velocity model and δm the high-wavenumber reflectivity model. Using δm to reproduce the reflected waves, RWI updates m_0 by minimizing

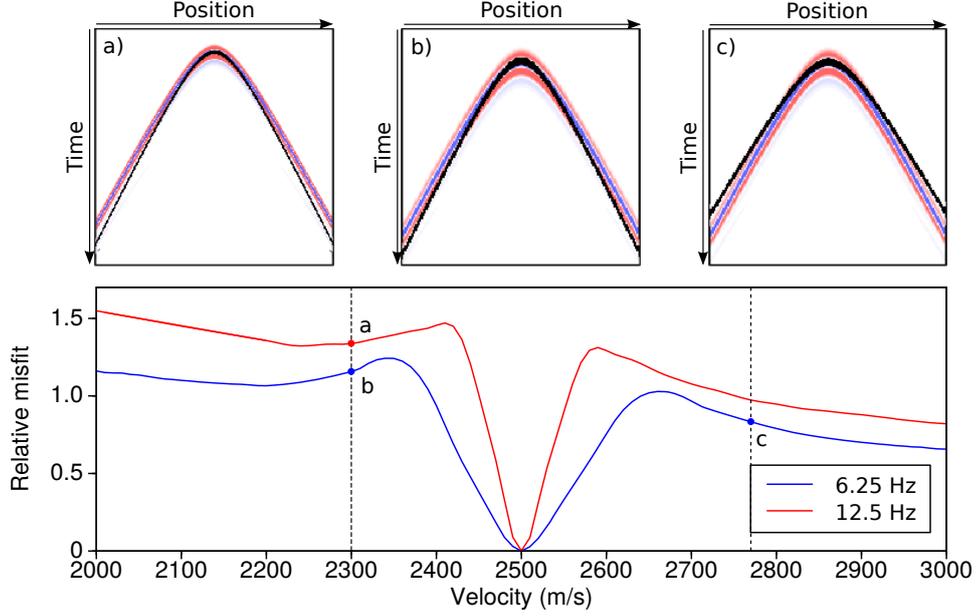


Figure 2.3: Mitigation of cycle skipping (top) and better convexity of the L2 norm-defined FWI misfit function (bottom), by adaptation of reflector depths to background velocity. The data are plotted in the same way as Figure 2.1 (top panel), which are generated by using the Ricker wavelet with central frequencies 6.25 Hz and 12.5 Hz (bottom). The true velocity equals 2500 m/s. The depth of the reflector is 500 m. The depth of the imaged reflector is estimated by $z = v/t_0$ where v denotes the tested velocity and t_0 the vertical two-way traveltime of the reflected waves. Please note the difference between the fixed depth z_{init} in Figure 2.2 and the changeable depth z here. In terms of traveltimes, the reproduction of reflection phase is gradually inaccurate from near to far offsets, and is more accurate for lower frequencies at short offsets (Point a *vs* Point b). Because of this, a good inversion scheme is to consider short offsets (and/or lower frequencies) before longer offsets (higher frequencies). Note that the local plateau shown in Figures 2.1 and 2.2 has been avoided here.

the following misfit function associated to reflected waves only:

$$C_{RWI}(m_0) = \frac{1}{2} \|d_{cal}^{refl}(m_0, \delta m) - d_{obs}^{refl}\|_2^2, \quad (2.20)$$

where the synthetic reflected waves $d_{cal}^{refl}(m_0, \delta m)$ are computed in m_0 (associated to kinematics) and δm (associated to dynamics). Note that δm should be nearly accurate in the sense that the reflector depths should be equal to the one-way vertical traveltime of reflected waves multiplied by the background velocity, as shown in Figure 2.2 that an improper reflector depth would result in a non-minimum point for the true velocity. Note also that RWI aims to update m_0 only; the adaptation of reflector depths (i.e. δm updates) should be implemented by using other methods, such as time-depth conversion

or linearized migration. The RWI gradient can be formulated as

$$\nabla C_{RWI}(m_0^{(0)}) \approx u_0 \star \delta\lambda^{refl} + \delta u \star \lambda_0^{refl}, \quad (2.21)$$

where I have neglected a minor term. Comparing with the FWI gradient (2.18) in the second iteration, the above formulation shows that the RWI gradient excludes the high-wavenumber term $u_0 \star \lambda_0^{refl}$. Therefore, RWI does not create reflectors in m_0 , as requested by the scale separation assumption.

Because the reflectivity model is decoupled from the velocity model, I have the freedom to re-image the reflectors once the background velocity is updated, such that the reflected waves are kept perfectly matched at short offsets throughout the iterations (Figure 2.3). Consequently, the misfit function presents fewer numbers of local minima, which are mainly attributed to long-offset mismatch. As in Figure 2.1, lower frequencies still improves the convexity of the misfit function. Furthermore, we could expect the global convergence to the true velocity, by using alternative misfit definitions such as those based on crosscorrelation (Xu *et al.*, 2012; Brossier *et al.*, 2015) or optimal transport (Métivier *et al.*, 2015), or some offset-driven strategies as those adopted in Section 3.5.

2.3 Sampling analysis in the frame of diffraction tomography and orthogonal decomposition

Another advantage of RWI is the capacity to approach to zero-to-low wavenumbers of subsurface at the early stage of inversion. This will be manifest by analyzing the imaging wavenumber vectors in the frame of generalized diffraction tomography. Huang *et Schuster* (2014) have also performed a comprehensive analysis of the resolution power I shall consider. I further decompose these vectors in the Cartesian system to assess the influence of acquisition offset and frequency content, which will be illustrated by a simple test.

2.3.1 FWI: preferentially high wavenumber samplings

Let us start the discussion from Equation (2.9), which expresses the FWI gradient in the frequency domain as a weighted product of the incident field u_0 and the conjugate of the adjoint field λ_0 . In a homogeneous medium (wavespeed = c_0), analytical expressions for the two fields can be deduced by using the Green function, which are functions of the arbitrary point $A(z, x)$ in space (see the coordinate system in Figure 2.4) and the circular frequency ω :

$$u_0(z, x; \omega) = \frac{W(\omega)}{4\pi|SA|} \exp [i\vec{s} \cdot \overrightarrow{SA}], \quad (2.22)$$

$$\lambda_0(z, x; \omega) = \frac{\Delta d(\omega)}{4\pi|GA|} \exp [i(-\vec{g}) \cdot \overrightarrow{GA}], \quad (2.23)$$

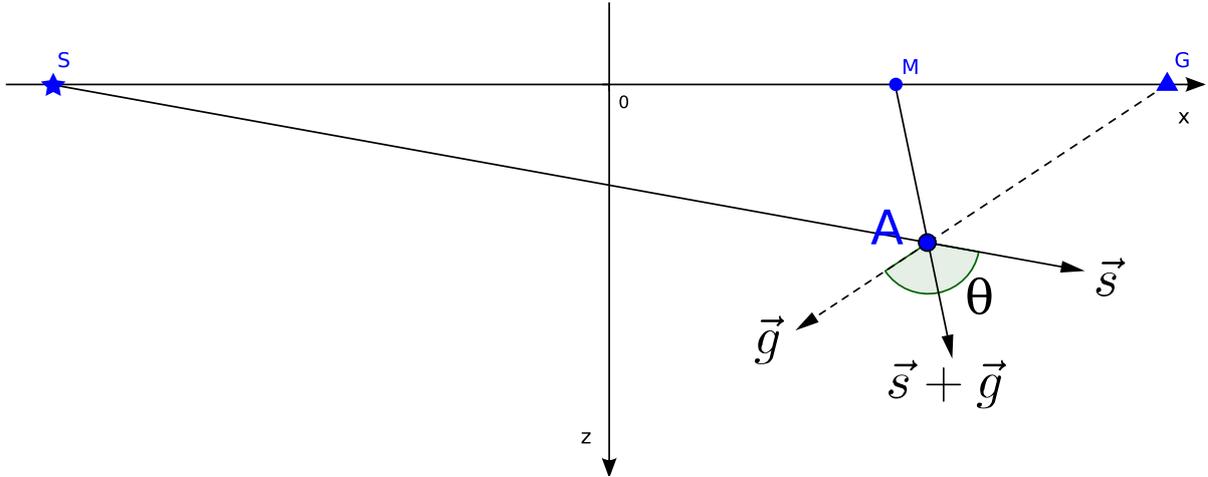


Figure 2.4: Sketch of the wavenumber vectors \vec{s} , \vec{g} and their sum (Equations (2.22)-(2.24)) in the 2D Cartesian system. S denotes the source position, G the receiver position, A an arbitrary point in space and M is the point on surface such that \overrightarrow{MA} is parallel to k . The coordinate system is built by choosing \overrightarrow{SG} as the x-axis and their midpoint as the origin. The positive z-axis is toward depth. The acquisition offset is related to the aperture angle θ which controls the modulus $|\vec{s} + \vec{g}|$, representing the recovered wavenumbers of the subsurface. In shallow zones, θ has a wide range and thus a broadband of wavenumbers can be recovered whereas in deep zones, θ is limited to small ranges and thus only high wavenumbers can be recovered without using low frequencies.

where $W(\omega)$ and $\Delta d(\omega)$ denote the spectra of the source and data residual, \vec{s} and \vec{g} (moduli= ω/c_0) denote the wavenumber vectors of the u_0 and λ_0 , respectively. Due to the spherical wavefront, \vec{s} (respectively \vec{g}) is parallel to the distance vector \overrightarrow{SA} (\overrightarrow{GA}) that starts from S to A (from G to A), the modulus of which is denoted by $|SA|$ ($|GA|$). The minus sign in the second expression derives from the conjugate operation on the adjoint field λ_0 of Equation (2.9), which also means the retropropagation in the time domain. Then, the FWI gradient (2.9) can be written as

$$\nabla C_{FWI}(z, x; \omega) \approx \Re \left\{ \frac{W(\omega)\Delta d(\omega)^*}{16\pi^2|SA||GA|} \exp[i\vec{k} \cdot \overrightarrow{MA}] \right\} \quad (2.24)$$

where $\vec{k} = \vec{s} + \vec{g}$ denotes the imaging wavenumber vector. The approximation is related to the ignorance of the partial derivative matrix $\frac{\partial B}{\partial m}$. I omit this factor to focus on the wavenumber samplings resulted from the vectors \vec{s} and \vec{g} . The influence of the partial derivative matrix will be discussed later.

The imaging wavenumber vector \vec{k} for a given source-receiver couple highlights the sampled spectral component of the subsurface perturbation, weighted by the (adjoint) source spectrum and the distance from the surface. Specifically, considering the aperture

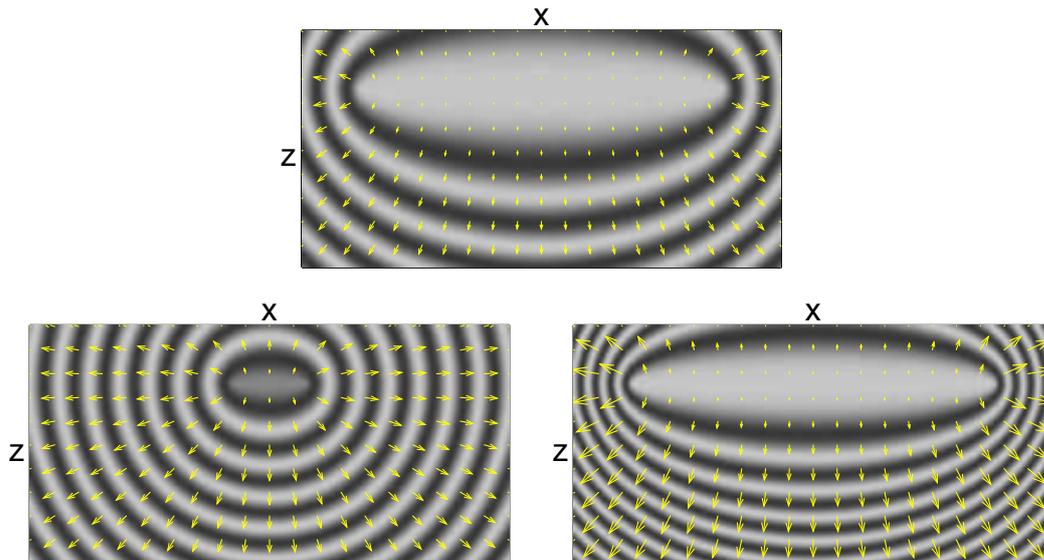


Figure 2.5: Illustration of the exponential kernel of the FWI gradient (Equation (2.24)), superimposed by the vector field \vec{k} deduced from Equation (2.26). Low wavenumbers of shallow zones are sampled while high wavenumbers are sampled at depths (top). By using shorter offsets (bottom left) or higher frequency (bottom right), we have samplings of higher wavenumbers. However, normal frequency bands and offset ranges often lacks low wavenumber samplings in depths, that can be theoretically supplemented by involving (very) low frequencies. In the frequency domain, the kernel is formed by a series of horizontal ellipses with the source and receiver as the focal points. The innermost ellipse is known as the first Fresnel zone while others are called secondary Fresnel zones, which form the migration isochrones when a group of frequencies are summed.

angle θ made between \vec{s} and \vec{g} , the modulus of \vec{k} is given by

$$|\vec{k}| = |\vec{s} + \vec{g}| = \frac{2\omega}{c_0} \cos \frac{\theta}{2}, \quad (2.25)$$

which is also known as the generalized diffraction tomography principle (Devaney, 1982; Wu et Toksöz, 1987; Mora, 1989, and others). It makes a link from the sampled wavenumbers of the subsurface to the aperture angle and data frequency (Figure 2.5). If the seismic data contain the whole 0–360° information, all wavenumbers will be imaged. From surface acquisitions, however, we often face a restriction that the range of θ decreases as A is deeper, meaning that the bandwidth of sampled wavenumbers are relatively narrow at depths. The missing low wavenumbers could be compensated for by decreasing the frequency content of the data (usually by implementing a low-pass filter). However, in practice this solution is often obstructed by the (very) low S/N of the low frequency data, largely because of the limited-band nature of the source spectrum $W(\omega)$.

The above analysis assumes a single source-receiver couple. For multiple couples, the total sampled wavenumbers are the gather of all imaging vectors \vec{k} associated to

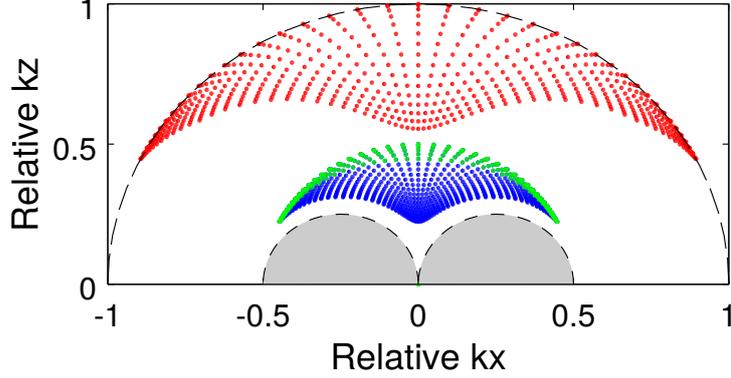


Figure 2.6: Diagram of sampled wavenumbers in the k_z - k_x axes. This diagram is generated by displaying as scattered points the vector heads of those \vec{k} 's in Equation (2.26) with varying x and H (but fixed z). These points depict areas from which vertical and horizontal wavenumbers are sampled. High frequency gives high k_z samplings (Red points) while low frequency gives low k_z samplings (Green and blue points). For one frequency, small offsets lead to the green points while large offsets lead to the blue points. The total sampled area is bounded by the black dashed semicircles. The blank area inside this bound can be filled in by using larger offsets and more distant source-receiver couples. In an asymptotic sense, infinitely long acquisitions allow to sample zero k_z at k_x of values between ± 0.5 and ± 1 , respectively. According to the analysis based on Equation (2.26), the two shaded areas between ± 0.5 cannot be filled unless lower frequencies are used.

each couple. To illustrate this total sampling, I first decompose \vec{k} of one couple into its orthogonal components using the coordinates of Figure 2.4:

$$\vec{k} = \vec{s} + \vec{g} = \frac{\omega}{c_0} \frac{(z, x + H)}{\sqrt{z^2 + (x + H)^2}} + \frac{\omega}{c_0} \frac{(z, x - H)}{\sqrt{z^2 + (x - H)^2}}, \quad (2.26)$$

where $H = \frac{1}{2}|SG|$ denotes the half offset. Then, I can determine the gather by varying the values of x and H (equivalent to moving sources and receivers on surface). At a common depth (fixing z), Figure 2.6 shows this gather in the k_z - k_x plot, where the vertical (respectively horizontal) axis represents which vertical (respectively horizontal) wavenumbers are sampled. This plot allows us to distinguish the respective samplings in the two directions, unlike Equation (2.25). The conclusions, nonetheless, are consistent: **a)** the FWI gradient preferentially images high k_z ; **b)** the sampling area is proportional to the considered frequency; **c)** for one frequency, increasing offset allows to approach to zero k_z samplings but still misses low k_x coverage (see shaded areas in Figure 2.6).

I should mention that the above reasoning is limited to the case of smooth models (i.e. no scattering in the considered frequency range). When the model has reflectivity content, however, the gradient has more terms than Equation (2.24) which gives low k_x samplings. I shall come back to this point in the discussion section.

Unlike FWI, RWI relies on the transmission regimes attributed to reflected waves. In the following, the associated sampling analysis will be carried out in the same manner as above. Orthogonal decompositions of the imaging wavenumber vectors will show that RWI naturally images lower wavenumbers than does FWI at great depths.

2.3.2 RWI: preferentially low wavenumber samplings

Suppose the subsurface comprises a single flat reflector at the depth D in the homogeneous background (wavespeed = c_0). The incident field u_0 encounters the reflector and generates the scattered field δu that includes both transmission and reflection. Because the contribution of the transmission field to the gradient is negligible in this simple setting, I only consider the reflection counterpart. Same consideration also holds for $\delta\lambda^{refl}$. In order to express the reflection fields by using the Green function, I adopt the concept of virtual source (Figure 2.7): above the reflector, the reflection field δu (respectively $\delta\lambda^{refl}$) can be approximated by an upgoing field that is generated from a virtual source located at the mirrored position S' (G') of the real source at S (real receiver at G), with regard to the reflector. Therefore, the four fields can be formulated as (for $z < D$)

$$u_0(z, x; \omega) = \frac{W(\omega)}{4\pi|SA|} \exp[i\vec{s}_1 \cdot \overrightarrow{SA}], \quad (2.27)$$

$$\delta\lambda(z, x; \omega) \approx \frac{Z\Delta d^{refl}(\omega)}{4\pi|G'A|} \exp[i(-\vec{g}_1) \cdot \overrightarrow{G'A}], \quad (2.28)$$

and

$$\delta u(z, x; \omega) \approx \frac{ZW(\omega)}{4\pi|S'A|} \exp[i\vec{s}_2 \cdot \overrightarrow{S'A}], \quad (2.29)$$

$$\lambda_0(z, x; \omega) = \frac{\Delta d^{refl}(\omega)}{4\pi|GA|} \exp[i(-\vec{g}_2) \cdot \overrightarrow{GA}], \quad (2.30)$$

where, again, the reflection coefficient of the reflector is simplified by the ratio Z , and \vec{s}_1 , \vec{s}_2 , \vec{g}_1 and \vec{g}_2 denote the wavenumber vectors of the four fields, respectively. This is a very important feature we add in RWI compared with FWI. Then, the RWI gradient (2.21) can be written as

$$\nabla C_{RWI}(z, x; \omega) \approx \Re \left\{ \frac{ZW(\omega)\Delta d(\omega)^*}{16\pi^2|SA||G'A|} \exp[i\vec{k}_1 \cdot \overrightarrow{M_1A}] \right\} \quad (2.31)$$

$$+ \Re \left\{ \frac{ZW(\omega)\Delta d(\omega)^*}{16\pi^2|S'A||GA|} \exp[i\vec{k}_2 \cdot \overrightarrow{M_2A}] \right\} \quad (2.32)$$

where $\vec{k}_1 = \vec{s}_1 + \vec{g}_1$ and $\vec{k}_2 = \vec{s}_2 + \vec{g}_2$ denote the imaging wavenumber vectors derived from the first and second terms of Expression (2.21). Points M_1 and M_2 are the intersection \vec{k}_1 and $\overrightarrow{SG'}$, \vec{k}_2 and $\overrightarrow{S'G}$, respectively. (For sake of clarity, M_1 and M_2 are not depicted in Figure 2.7. Please see Figure 2.4 for a reference.) The approximation is related to

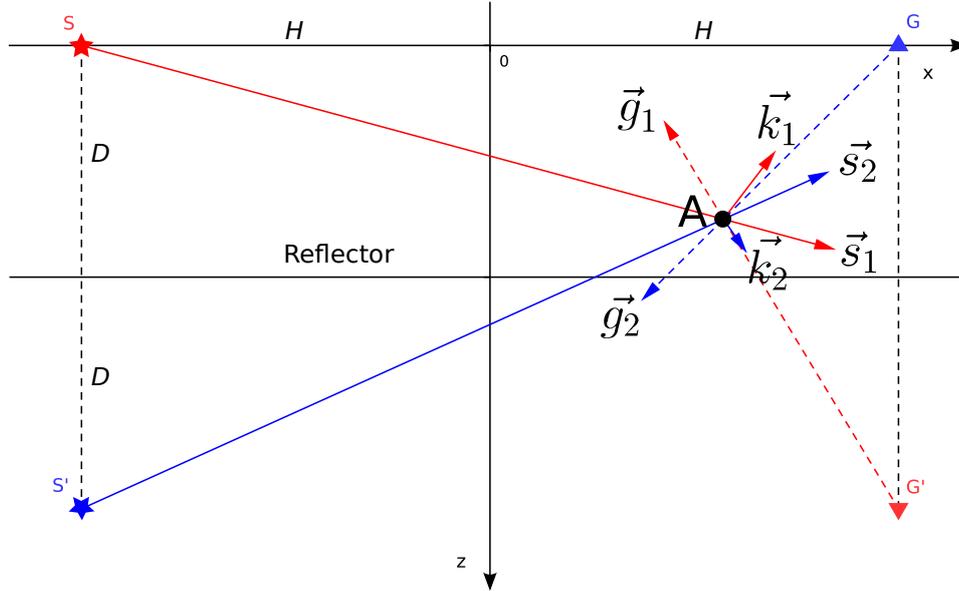


Figure 2.7: Sketch of wavenumber vectors \vec{s}_i , \vec{g}_i and their sums \vec{k}_i ($i = 1, 2$, Equations (2.27)-(2.24)). The coordinate system is built by the same manner in Figure 2.4. S denotes the source position, S' the mirrored position of S with regard to the reflector, G the receiver position and G' the related mirrored position. Virtual sources are proposed at S' and G' for analytical expressions of the scattered fields δu and $\delta \lambda$, respectively.

the ignorance of the partial derivatives of forward modeling operators with respect to model parameters, as I am interested in the geometrical configuration in this discussion. Figure 2.8 illustrates the two exponential kernels of the gradient, superimposed by the vector fields k_1 and k_2 , respectively. Unlike the FWI gradient case (2.25), the aperture angles associated to \vec{k}_1 and \vec{k}_2 are relatively large, resulting low wavenumber samplings in deep zones.

I decompose \vec{k}_1 and \vec{k}_2 in the orthogonal coordinates to account for multiple source-receiver couples:

$$\vec{k}_1 = \vec{s}_1 + \vec{g}_1 = \frac{\omega}{c_0} \frac{(z, x + H)}{\sqrt{z^2 + (x + H)^2}} + \frac{\omega}{c_0} \frac{(z - 2D, x - H)}{\sqrt{(z - 2D)^2 + (x - H)^2}}, \quad (2.33)$$

$$\vec{k}_2 = \vec{s}_2 + \vec{g}_2 = \frac{\omega}{c_0} \frac{(z - 2D, x + H)}{\sqrt{(z - 2D)^2 + (x + H)^2}} + \frac{\omega}{c_0} \frac{(z, x - H)}{\sqrt{z^2 + (x - H)^2}}. \quad (2.34)$$

As in Figure 2.6, I plot the sampled wavenumbers of vertical and horizontal directions in Figure 2.9, from which I can infer that

- a) The RWI gradient preferentially images zero-to-low k_z , which is complementary to the FWI gradient in terms of samplings. This issue will be detailed in the next paragraph;

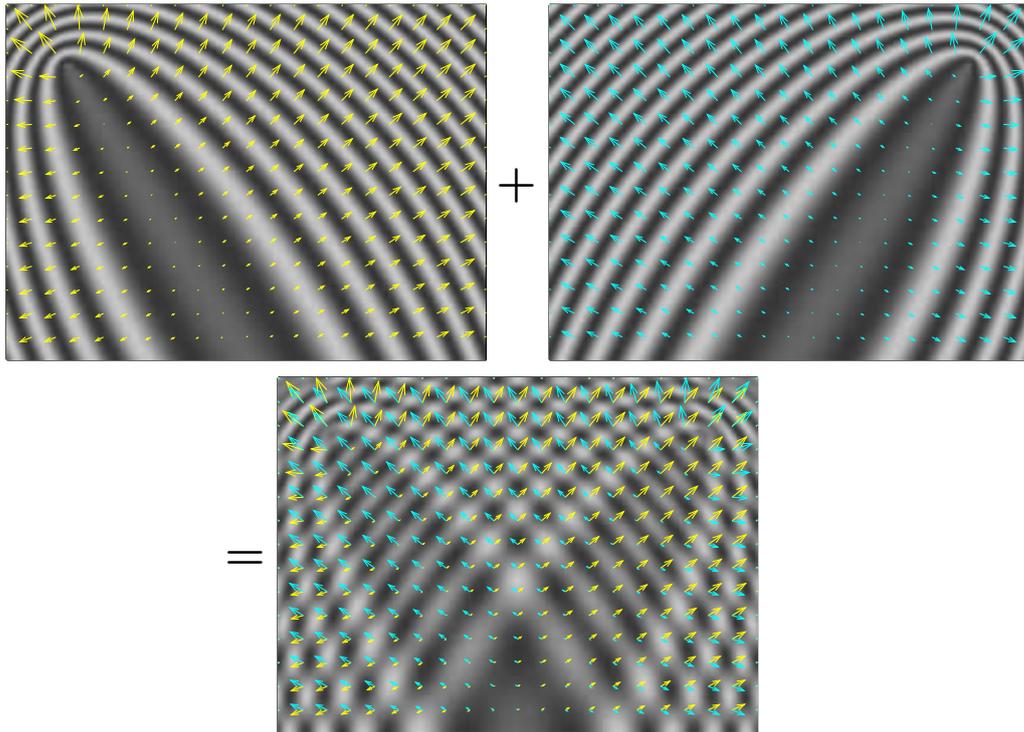


Figure 2.8: Illustration of the first exponential kernel (top left), second kernel (top right) and their sum (bottom) of the RWI gradient (Equations (2.31) and (2.32)) for one source-receiver couple and one frequency, superimposed by the vector fields \vec{k}_1 and/or \vec{k}_2 deduced from Equations (2.33) and (2.34). The flat reflector is positioned at the bottom of each panel. The kernels are formed by series of inclined ellipses. Their common focal points are at the real source position and the virtual receiver position (first kernel), or at the virtual source position and the real receiver position (second term). The total sensitivity kernel is a superimposition of these two groups of ellipses, which are related to the transmission regimes along the reflection path. This kernel gives much lower wavenumber samplings than does FWI gradient at great depths (see Figure 2.10).

- b) As for FWI, the sampling area of RWI is proportional to the considered frequency. The difference is that for RWI the coverage of low wavenumbers is not enlarged by decreasing the frequency content; the high frequency provides the same low-wavenumber content as the low frequency. Therefore, the benefit of using low frequencies is not related to a wider range of wavenumbers, but a denser sampling of low wavenumbers;
- c) For one frequency, increasing offset allows to sample higher k_z , thus results in higher resolution. Again, this is different from FWI for which larger offsets lead to lower wavenumbers.

In the next section, these behaviors are partially verified by a simple numerical example.

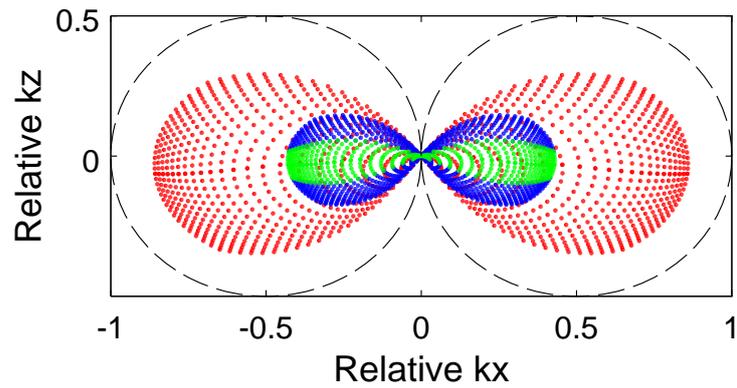


Figure 2.9: Diagram of sampled wavenumbers in the k_z - k_x axes. This diagram is plotted in a similar way of Figure 2.6. Zero-to-low wavenumbers are sampled. Compared with the low frequency, the high frequency (Red points) gives higher k_z and k_x samplings with larger intervals (sparse sampling). Large offsets (Blue points) increase the bandwidth of k_z samplings, with the same interval of small offsets (Green points). The total sampled area is bounded by the black dashed circles. The blank area inside this bound can be filled in by using larger offsets and more distant source-receiver couples.

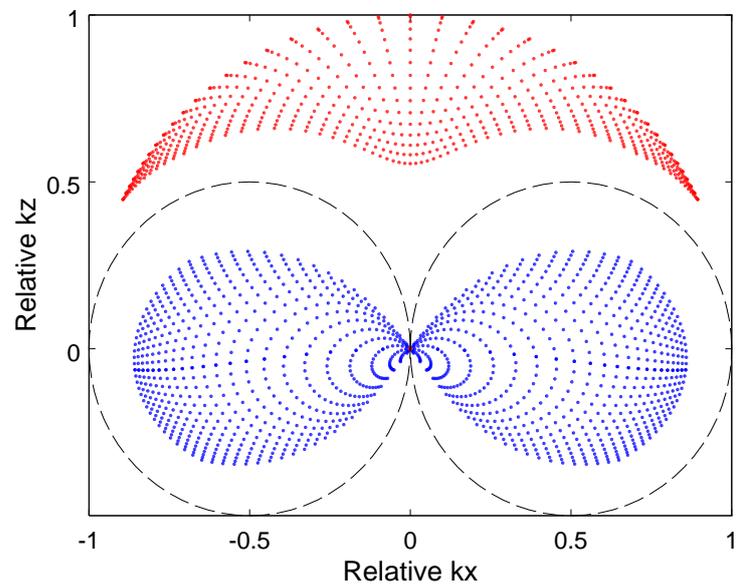


Figure 2.10: Diagram of sampled wavenumbers by FWI (red points) and RWI (blue points) gradients, respectively, separated by the blank area that is also known as the scale gap in seismic imaging. The black dashed circles denote the boundary of the two samplings that could be reached in the limit case (infinite long acquisitions).

Figure 2.10 compares the wavenumber samplings respectively from FWI (Figure 2.6) and RWI (Figure 2.9) for multiple offsets and one frequency. While FWI preferentially samples in a narrow band of high k_z , RWI preferentially samples zero-to-small k_z components of the subsurface. The two disjoint areas highlights the scale gap in terms of k_z samplings that is resulted from the limited length of the acquisition geometry (remind Figure 1.2). With larger offsets/wider azimuth coverage, the gap can be narrowed, and the two areas asymptotically reach the theoretical boundary depicted by the dashed circles. A broad frequency band can also be used to fill this gap efficiently.

2.3.3 Numerical verification

I verify our former analysis by a simple experiment. The true velocity and density models are shown in Figure 2.11. A ball-shape anomaly with a higher velocity value is present in the velocity model, which is excluded in the homogeneous initial model (1000 m/s). Because the box-car function corresponds to the sinc function in the Fourier domain, the sharp boundary of the anomaly provides a broadband spectrum, from which I can determine which wavenumbers are sampled by FWI and RWI.

The true density model comprises a flat interface to generate the reflected wave. I take this model as the initial density model and keep it unchanged. Therefore, the misfit function includes the residuals associated to the direct and reflected waves, and combines the FWI misfit (2.8) and RWI misfit (2.20) in a single L2 norm:

$$C(m_0) = \frac{1}{2} \left\| \underbrace{\left(d_{cal}^{dir}(m_0) - d_{obs}^{dir} \right)}_{\text{FWI}} + \underbrace{\left(d_{cal}^{refl}(m_0, \delta m) - d_{obs}^{refl} \right)}_{\text{RWI}} \right\|_2^2, \quad (2.35)$$

where m_0 and δm denotes the velocity and density models, respectively (let us omit the influence of δm to d_{cal}^{dir}). Then, the associated gradient joints the FWI gradient (2.17) and RWI gradient (2.21) as⁵

$$\nabla C(m_0) = \underbrace{u_0 \star \lambda_0^{dir} + u_0 \star \lambda_0^{refl}}_{\text{FWI}} + \underbrace{u_0 \star \delta \lambda^{refl} + \delta u \star \lambda_0^{refl}}_{\text{RWI}}. \quad (2.36)$$

I consider two offset ranges represented by the red line (large offsets) and blue line (short offsets) in Figure 2.11, respectively. Four frequencies are simultaneously used: [2, 4, 6, 10] Hz. The range of sampled wavenumbers can be assessed by analyzing the gradient in the Fourier domain (Figure 2.12). I identify five fundamental features (Table 2.2). Features 1–4 confirm the analysis in Figure 2.6 that FWI preferentially images high k_z and this imaging is proportional to the considered frequency. Comparing Feature 5 with Feature 4, we see that RWI preferentially samples zero-to-small k_z that is complementary to the FWI high k_z sampling, as pointed out in Figure 2.10. Here, we do not see RWI samplings from other frequencies. I attribute this nearly-zero imprint to the

⁵Another equivalent phrasing is that the gradient (2.36) is the joint of the FWI gradient (2.18) in the second iteration and RWI gradient (2.21) in the first iteration with shared terms $u_0 \star \delta \lambda^{refl} + \delta u \star \lambda_0^{refl}$.

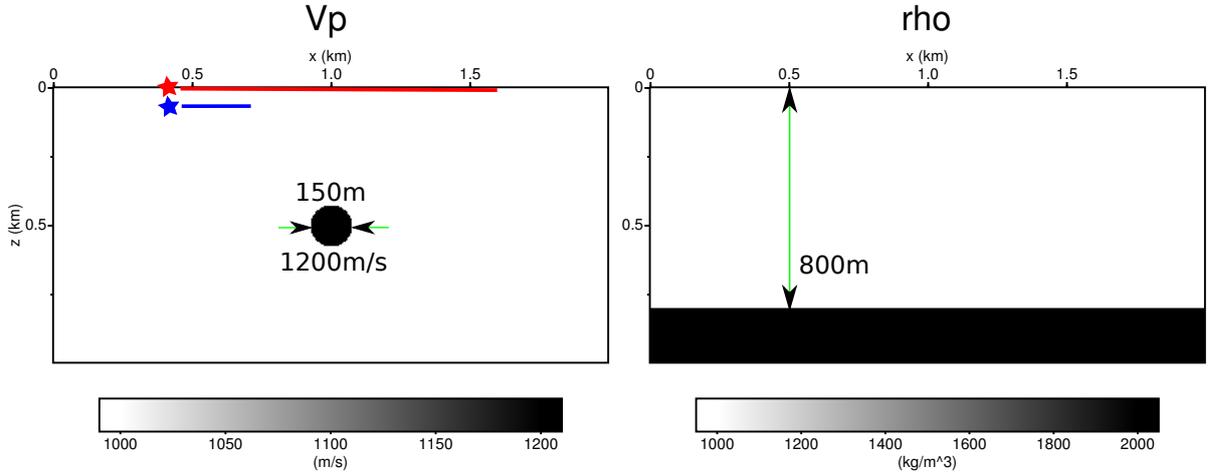


Figure 2.11: True velocity and density models. The initial velocity model excludes the ball-shape anomaly. The model is discretized by $N_x \times N_z = 400 \times 200$ with a 5 m interval. The sources and receivers are deployed on the model surface with a 50 m spacing. Red line: 0–1200 m offsets. Blue line: 0–300 m offsets (deployed also on the surface).

weak energy of RWI terms compared with $u_0 \star \lambda_0^{refl}$ (remind the amplitude comparison in (2.19) that the RWI terms are around 10% of it). On the other hand, large offsets help enlarge the respective sampling areas: lower k_z is sampled by FWI (white arrows in Figure 2.12) and relatively higher k_z is sampled by RWI (note the higher amplitudes for Feature 5 in the left panel comparing with the right one). In other words, by using large offsets, the scale gap between FWI and RWI samplings is narrowed, and the intermediate wavenumbers are more continuously sampled (remind the gap in Figure 2.10).

As a result, we can see a good reconstruction from large offsets in Figure 2.13. Higher velocities are imaged at the correct position of the ball. The periodic shadows around the ball are attributed to the sparse sampling of frequencies, the content of which is not sufficient to cancel out the side lobes of the signals. If larger offsets and more frequencies are used, these shadows should be further attenuated. On the contrary, the short-offset result can be considered as the low-wavenumber component imaged by the RWI kernels (the grey zones along the $x = 1$ km profile), superimposed by the high-wavenumber perturbations imaged by the FWI kernels that are similar to those in the large-offset result. Due to the narrow offset range, we lack more high wavenumbers. Therefore, these shadows are more visible in the result (Figure 2.12). The ball is also not fully recovered because intermediate wavenumbers are insufficiently sampled.

In the large-offset result, a high-velocity perturbation with a linear shape is observed below the ball. It is a footprint of the migration isochrone ($u_0 \star \lambda_0^{refl}$ term) that should have been used to image the reflectivity in the density model. Such a leakage derives from the possibility to explain the density-associated reflected waves by velocity perturbations,

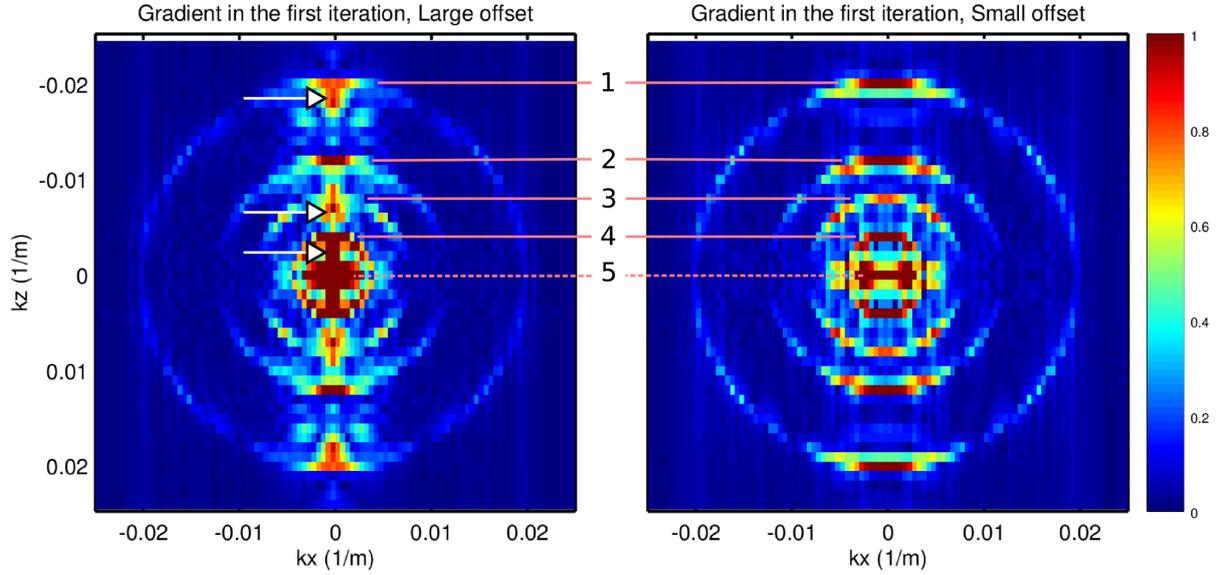


Figure 2.12: Spectra of gradients with large offsets (left) and short offsets (right). Five fundamental features are identified. Features 1–4 are associated to FWI high wavenumber samplings and Feature 5 is associated to RWI low wavenumber samplings. The scale gap can be narrowed by using large offsets (white arrows).

Table 2.2: Five features identified in the gradient spectra for analysis

#	Feature	Sampling area	Source of contribution	Freq.
1	Outermost circle	highest k_z	$u_0 \star \lambda_0^{refl}$ (FWI)	10 Hz
2	Sub-outer circle	higher k_z	$u_0 \star \lambda_0^{refl}$ (FWI)	6 Hz
3	Middle circle	middle k_z	$u_0 \star \lambda_0^{refl}$ (FWI)	4 Hz
4	Innermost circle	low k_z	$u_0 \star \lambda_0^{refl}$ (FWI)	2 Hz
5	Red-filled oval	low k_x	the two RWI terms	2 Hz

a trade-off between two parameters. Other trade-offs are also possible (e.g. velocity–anisotropic parameters), and should be mitigated for reliable subsurface imaging. I shall go back to this point in Section 2.5 and shall suggest a decoupling scheme in Chapter 3.

2.4 Discussion

According to the classification of Table 1.2, FWI and RWI belong to data-domain waveform inversion methods. Though their target functions are different, both seek the nearby minimum by computing the gradients without building extended images. Their key difference, however, resides in the scale separation assumption. In FWI, no separation is

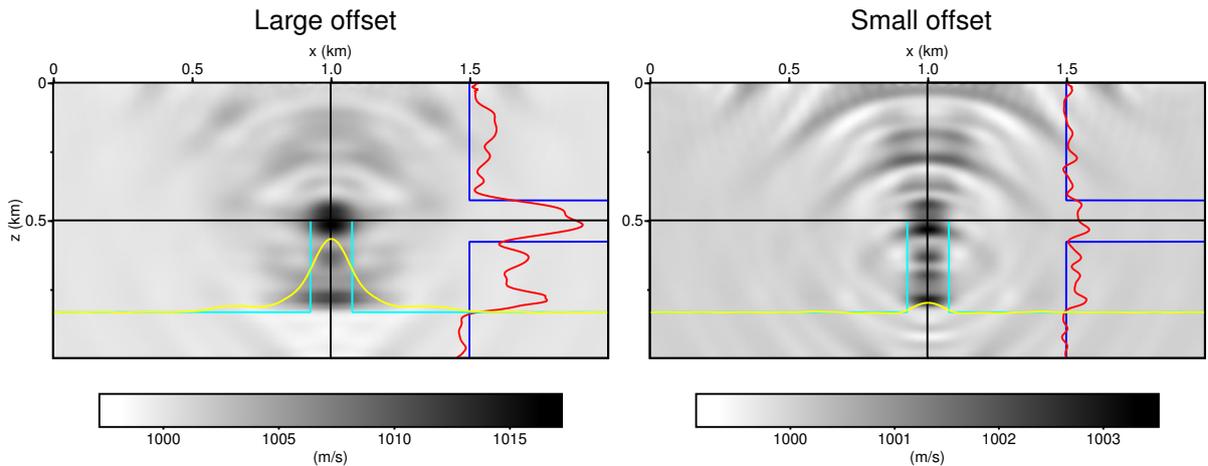


Figure 2.13: Results by large (left) and small (right) offset data. Four frequencies are simultaneously used: [2, 4, 6, 10] Hz. The gradient is smoothed to avoid too rapid changes. Eleven iterations are implemented on the long-offset data and nine iterations on the short-offset data, both reduced the associated misfit function by 25%. I superimpose on the results the vertical profiles of the true velocity model (blue curves) and the result (red curves) at $x = 1$ km position, and the horizontal profiles of the true velocity model (cyan curves) and the result (yellow curves) at 0.5 km depth, respectively. The samplings of intermediate wavenumbers from large offsets translates to an improved reconstruction of the ball anomaly. See text for details.

assumed. Although low wavenumbers are often missing at depths due to insufficient coverage, there is no intrinsic limitation to recover the full wavenumber spectrum. In contrast, RWI assumes the separation between low and high wavenumbers, as represented by the background velocity model m_0 and reflectivity model δm , respectively. A cycle workflow that alternates true-amplitude migration and RWI can be used to simultaneously reconstruct m_0 and δm . This is considered to be more robust than classical FWI in two aspects:

- Mitigation of cycle skipping. The ambiguity of reflector depth and background velocity actually reflects the ill-posedness of the problem: we want to resolve two quantities from mainly one data attribute (i.e. the traveltimes). Using longer offsets help resolve the ambiguity but the problem becomes more nonlinear. In addition, the L2 norm defined misfit function is too sensitive to the waveform, such that many local minima are generated. On the other hand, by scale separation, RWI actually only solve for velocity from the traveltimes, making the problem better posed. The remaining parameter, i.e. the depth is allowed to be re-defined such that the near-offset data are kept matched after velocity updates, producing fewer local minima than FWI. Conversely, this match can stabilize the data fitting procedure at longer offsets if we incorporate them in later iterations. Note that, since the migration

isochrone has been filtered out, fitting near-offset data does not generate reflectivity in the velocity model, unlike FWI.

- Low wavenumber samplings at depths. Because of the ambiguity mentioned above, FWI mainly relies on diving waves to image low wavenumbers, at least in the early stage of iteration. However, the diving wave is well known to have the limitation on depth penetration, therefore may not sample deeper targets. Using lower frequencies is sometimes ineffective as it is always a question of noise. As the ambiguity is mitigated by RWI, we can enjoy the low-wavenumber information carried by the reflected waves. The way that RWI image these low wavenumbers is similar to the ideology of reflection tomography in the data domain, but as the waveform is involved RWI can provide higher resolution. In addition, as no extended images are built, it requires less computing resources than MVA methods.

In practice, FWI also use reflected waves to recover deep targets, assuming the ambiguity is somehow under control. As shown in Section 2.2, the use of reflected waves generates reflectivity images in the velocity model, which can produce in the second iteration the two scattering quantities δu and $\delta\lambda^{refl}$ giving RWI-like sensitivity kernels. This is the case of the misfit function (2.35) and gradient expression (2.36). In other words, FWI includes the RWI functionality that samples low wavenumbers at depths. However, I need to emphasize that, because the depth of reflectors tend to be unchanged through FWI iterations, as discussed in Section 2.2, the cycle skipping problem or ambiguity still exists, and therefore the reconstructions may not be reliable at depth. Therefore, the main advantage of RWI over FWI in terms of reflection data processing is related to the enhancement of low wavenumber samplings at the early stage of the iteration, reducing the number of local minima as many as possible.

The last point to discuss is related to the use of frequency. Like FWI, RWI would need lower frequencies to avoid potential cycle skipping at long offsets, and enjoy faster convergence rate and higher resolution from higher frequencies. This is same as the multiscale strategy proposed by [Bunks *et al.* \(1995\)](#). However, this does not mean that we could expect more low wavenumber reconstructions from low frequencies. As shown in Section 2.3, high frequencies are able to provide the same content of low wavenumber samplings as low frequencies. As soon as cycle skipping is avoided, we can start inversion from higher frequencies.

In the cycle workflow, one may consider to use higher frequency for true-amplitude migration and lower frequency for RWI, which helps satisfy the scale separation assumption. However, experiences have suggested use a same frequency band for both processes. This is mainly because if two bands are respectively used, the reflection coefficients inferred from the high-frequency reflectivity may not be sufficiently accurate to reproduce the low-frequency reflection waveform for RWI; the same frequency band ensures that the waveform residuals at longer offsets are totally attributed to velocity errors, in the frame of acoustic assumption.

2.5 Summary

In this chapter, I review essential ingredients of FWI and RWI, particularly in the case of surface acquisitions. The ambiguity derived from reflection data is often inappropriately handled in FWI, leading to cycle skipping problems and erroneous velocity models. In contrast, the scale separation assumed in RWI allows the adaptation of reflector depths, which makes perfect match at short offsets and reduces the number of local minima for the misfit function. As a result, RWI is more robust than FWI. Classical inversion schemes that work for FWI, such as travelttime-oriented misfit functions, multiscale inversion, offset-driven strategies etc., can also be used for RWI for global convergence.

The analysis on resolution power is carried out in the frame of generalized diffraction tomography. The orthogonal decomposition of the imaging wavenumber vectors allows to further assess the wavenumber samplings respectively in the vertical and horizontal directions. While both FWI and RWI have samplings of all horizontal wavenumbers, a scale gap of vertical wavenumber samplings is observed for FWI and RWI: FWI preferentially samples high wavenumbers whereas RWI preferentially samples low wavenumbers. This gap can be filled in by using larger acquisition coverage; in the limit case this gap would disappear. However, the distinction between FWI and RWI samplings has been lined out in this theoretical study, and also verified by a numerical example. Because of this, at the depths beyond diving-wave penetration, FWI results in small-scale perturbations (large-scale updates are case-dependent and less reliable) whereas RWI gives rise to large-scale reconstructions.

The frequency bandwidth plays different roles in FWI and RWI. While FWI needs low frequencies for low wavenumber samplings, RWI can extract the low wavenumbers from high frequency data as long as cycle skipping problem is avoided. The benefit of low frequencies is related to the dense samplings of low wavenumbers.

Although the conclusions are deduced from simple cases, they can be generalized to more sophisticated models and serve as guidelines for real data applications. Here, for a preparation of the next chapter, I put forward the discussion related to two aspects:

- Subsurface parameterization. The influence of the parameterization has been omitted in Equations (2.24), (2.31) and (2.32). Specifically, they do not change the wavenumber samplings provided by the imaging wavenumber vectors, but affects the amplitudes of the sensitivity kernels. Depending on the parameterization, a diffractor may generate stronger transmitted waves than reflected waves (or the opposite situation), favoring the illumination coming from large aperture angles ($\sim 180^\circ$) to small ones ($\sim 0^\circ$). These angle-dependent illuminations, in turn, act as a weighting function on the sensitivity kernels (damping high wavenumbers) and thus results in low wavenumber reconstruction of that diffractor. This wavenumber-related property can be considered to further meet the scale separation assumption.

- Using diving waves in RWI. From Figure 2.9 or 2.10, I concluded that RWI samples low k_z . However, the bandwidth of this sampling reduces as k_x approaches to zero. This reflects that RWI misses the low k_z information carried by the diving waves, and this shortage cannot be remedied by using only short-spread reflected waves. Therefore, diving waves should be considered. However, this is not straightforward, as I still want to preserve the scale separation property of RWI. For example, Equation (2.35) provides a brute-force way that directly inserts diving wave-related residuals into the misfit function, but the resulting gradient formulation (2.36) presents the migration isochrone term which leads to high-wavenumber velocity updates. The mitigation of this contribution may be realized by choosing a particular parameterization, as mentioned above, but this may be less effective as expected due to its high amplitude. Therefore, I should follow an alternative way which enables a natural suppressing of this term, and the diving way can also be included through a unified formulation. This new methodology, named Joint Full Waveform Inversion, is the topic of the next chapter.

Chapter 3

Joint Full Waveform Inversion

The world under heaven, after a long period of division, tends to unite; after a long period of union, tends to divide. This has been so since antiquity.

— Guanzhong Luo, *Romance of Three Kingdoms*

In the previous chapter a simple example is used to illustrate the cycle skipping problem of FWI and how it is mitigated by RWI assuming scale separation. Moreover, I review the wavenumber samplings provided by FWI and RWI. In shallow zones, we often enjoy the abundance of transmitted and reflected waves for broadband reconstructions whereas in deep zones, the limitation of depth penetration of diving waves leads to high-wavenumber samplings only. In contrast, RWI emphasizes the transmission regimes of reflected waves to recover low wavenumbers at depths. In this chapter, I shall prove this point more rigorously by analyzing the gradient formulations, and based on this I shall propose an extended RWI approach, namely Joint FWI, which also includes diving waves for more complete samplings of low wavenumbers. How is it formulated, such that the associated gradient does not include the migration isochrones? Which subsurface parameterization should be used, such that we can naturally separate large scales from small ones? In this chapter, I shall answer these two questions and provide a synthetic example highlighting the advantages of JFWI over RWI and FWI.

**Full Waveform Inversion of Diving & Reflected Waves
for Velocity Model Building with Impedance Inversion
Based on Scale Separation**ZHOU, W., BROSSIER, R., OPERTO, S. & VIRIEUX, J.
2015, *Geophysical Journal International*, **202**:1535–1554.**Abstract**

Full waveform inversion (FWI) aims to reconstruct high-resolution subsurface models from the full wavefield which includes diving waves, post-critical reflections and short-spread reflections. Most successful applications of FWI are driven by the information carried by diving waves and post-critical reflections to build the long-to-intermediate wavelengths of the velocity structure. Alternative approaches, referred to as reflection waveform inversion, have been recently revisited to retrieve these long-to-intermediate wavelengths from short-spread reflections by using some prior knowledge of the reflectivity and a scale separation between the velocity macromodel and the reflectivity. This study presents a unified formalism of FWI, named as Joint FWI, whose aim is to efficiently combine the diving and reflected waves for velocity model building. The two key ingredients of Joint FWI are, on the data side, the explicit separation between the short-spread reflections and the wide-angle arrivals and, on the model side, the scale separation between the velocity macromodel and the short-scale impedance model. The velocity model and the impedance model are updated in an alternate way by Joint FWI and waveform inversion of the reflection data (least-squares migration), respectively. Starting from a crude velocity model, Joint FWI is applied to the streamer seismic data computed in the synthetic Valhall model. While the conventional FWI is stuck into a local minimum due to cycle skipping, Joint FWI succeeds in building a reliable velocity macromodel. Compared with reflection waveform inversion, the use of diving waves in Joint FWI improves the reconstruction of shallow velocities, which translates into an improved imaging at deeper depths. The smooth velocity model built by Joint FWI can be subsequently used as a reliable initial model for conventional FWI to increase the high-wavenumber content of the velocity model.

3.1 Introduction

With the emergence of long-offset wide-azimuth acquisitions and broadband sources, full waveform inversion (FWI) has been recognized as an efficient tool for velocity model building (Virieux et Operto (2009) for a review). In these long-offset experiments, FWI is mainly driven by the information carried by diving waves and post-critical reflections to build the long-to-intermediate wavelengths of the velocity structure. The connection between the acquisition geometry and the spatial resolution of FWI has been clearly established in the theoretical framework of the generalized diffraction tomography (e.g. Devaney, 1982; Miller *et al.*, 1987). It has been shown that the wavenumber component \mathbf{k} ,

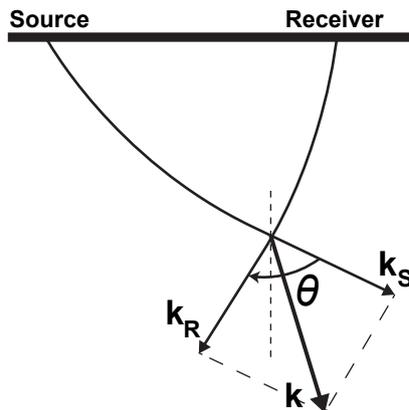


Figure 3.1: Spatial resolution of diffraction tomography and its connection with acquisition geometry. The wavenumber vectors associated with the rays connecting the source and the receiver to the diffractor are denoted by \mathbf{k}_S and \mathbf{k}_R , respectively. The scattering angle is denoted by θ . The wavenumber vector $\mathbf{k} = \mathbf{k}_S + \mathbf{k}_R$ is the spectral component mapped at the diffractor point by the source-receiver pair during FWI.

injected at a diffractor point in the subsurface is related to the local wavelength λ and the scattering angle θ by the relationship

$$\mathbf{k} = \frac{2}{\lambda} \cos\left(\frac{\theta}{2}\right) \mathbf{n}, \quad (3.1)$$

where \mathbf{n} is the normalization of the vector \mathbf{k} (Figure 3.1). This relationship shows that the wide-scattering angles associated with diving waves and post-critical reflections contribute low-to-intermediate wavenumber updates to the subsurface. Conversely, the small scattering angles associated with short-spread reflections contribute high wavenumber updates to the subsurface. One key issue in the classical formulation of FWI (for short, referred to as FWI in the following) is that the penetration depths of diving waves are often insufficient to reach the deepest targeted structures, even from modern wide-azimuth surveys. At these depths, FWI behaves as a least-squares migration of the short-spread reflections rather than as a tool for velocity model building, and would fail to update the low-to-intermediate wavenumbers of deep targets.

Alternatively, migration-based velocity analysis has been developed in the image domain to build the velocity macromodel by using the reflection data (Diaz *et al.*, 2013; Liu *et al.*, 2013; Allemand et Lambaré, 2014). These approaches focus on the flattening of the common image gathers generated by migration (Symes et Carazzone, 1991; Sava et Biondi, 2004). Extended-domain approaches have also been proposed, which attempt to minimize the energy left in the non-physical dimensions added to the model space (Sava et Fomel, 2006; Yang et Sava, 2011; Almomin et Biondi, 2012; Biondi et Almomin, 2012; Sun et Symes, 2012; Lameloise *et al.*, 2015). The main issue of these approaches is their high computational cost that is made, on the one hand, by the migration step performed during each velocity update and, on the other hand, by the extended-domain

imaging condition. Although 2D applications have shown promising results, the cost of these approaches seems to prevent a direct extension to 3D cases, in particular when the forward modeling is performed with the two-way wave equation.

Inspired by the pioneering work of [Chavent *et al.* \(1994\)](#), [Chavent \(1996\)](#) and [Clément *et al.* \(2001\)](#) on migration-based traveltime inversion (MBTT), recent data-domain FWI strategies, referred to as reflection waveform inversion (RWI) in this paper, have been proposed as a new alternative to build the velocity macromodel from the reflection data (e.g., [Xu *et al.*, 2012](#); [Ma *et al.*, 2013](#); [Brossier *et al.*, 2015](#)). As most of the seismic reflection processing workflows, RWI relies on the explicit scale separation between a smooth velocity macromodel and a rough reflectivity. This scale separation results from the gap between wavenumber contents of the velocity macromodel built by reflection tomography or migration velocity analysis and the reflectivity built by migration ([Claerbout, 1985](#); [Jannane *et al.*, 1989](#); [Mora, 1989](#); [Wu *et al.*, 1987](#)). Such a separation leads to a two-step imaging workflow in which one repeatedly alternates the velocity model building assuming a known reflectivity and the reflectivity update by migration using the previous velocity update as the background model. More sophisticated approaches can be viewed to mitigate the computational burden of this workflow, for example by building the reflectivity in the pseudo-time domain ([Plessix, 2013](#)) to avoid performing migration at each iteration of the velocity update ([Brossier *et al.*, 2015](#); [Wang *et al.*, 2015](#)). In RWI, the governing idea behind the velocity model building task is to assume the reflectivity is known in priori, by which the reflected waves are predicted and the residuals are minimized for the velocity macromodel update. Under this assumption, high-wavenumber contributions such as migration isochrones are not present in the sensitivity kernels of RWI. Moreover, this prior reflectivity is used by RWI as the secondary sources to highlight in the sensitivity kernel the contribution of the transmission paths followed by the reflected waves. Indeed, the wide-scattering angles associated with these transmission regimes are amenable to update the low-to-intermediate wavenumbers of the subsurface located between the reflectors and the surface (Equation 3.1).

Brute-force approaches might be performed without explicit scale separation in the FWI formalism (e.g. the one proposed by Equation (2.35) in Section 2.3.3). They consist of applying the conventional FWI to build reflectivity during early iterations by a migration-like processing (e.g. the associated gradient expression (2.36) in Section 2.3.3) before updating the low-to-intermediate wavenumbers of the subsurface from the transmission paths of the reflected waves ([AlTheyab *et al.*, 2013](#)). These approaches could be further developed for the velocity macromodel building or for the reflectivity imaging by a wavenumber-driven filtering of the gradient of the FWI misfit function ([Alkhalifah *et al.*, 2014](#); [Alkhalifah, 2014](#); [Wu *et al.*, 2014](#); [Alkhalifah, 2015](#)). Note that RWI can be implemented either in the time domain or in the frequency domain as shown by [Wang *et al.* \(2013c\)](#). Other data-domain approaches for velocity model building rely on a wavefield decomposition into upgoing and downgoing waves to separate the contribution of forward-scattering and backward-scattering in the sensitivity kernel of waveform inversion ([Wang *et al.*, 2013a](#); [Tang *et al.*, 2013](#)).

One key limitation of RWI, that will be overcome in this study, is the exclusive re-

liance on the use of reflected waves, discarding the low-wavenumber information on the shallow targets that are carried by the diving waves. In this study, we propose a new FWI method, referred to as joint full waveform inversion (JFWI), which integrates the ingredients of the conventional FWI and RWI into a unified formalism. Like RWI, JFWI still rests on the scale separation between the velocity macromodel and the reflectivity. The added value of JFWI compared with RWI, is the combination of the diving waves and the reflected waves such that the low-wavenumber information carried by these two wave modes, as described above, are simultaneously used in the velocity model building task. In JFWI, we regenerate the reflectivity by conventional FWI using only short-offset reflected waves (i.e., non-linear least-squares migration) at each iteration of the scale-separation workflow. A key feature of JFWI is to require the explicit separation of the early arriving phases (diving waves and post-critical reflections) and the pre-critical reflections in the data. Our approach shares some similarities with Wang *et al.* (2015), who also use refracted and reflected waves to update the low-to-intermediate wavenumbers of the subsurface. The main difference is that our workflow relies on the waveform-difference misfit function to update the velocity macromodel and the reflectivity, whereas the approach of Wang *et al.* (2015) relies on the cross-correlation based misfit function, the so-called wave-equation tomography (Luo et Schuster, 1991), to perform the velocity macromodel update. Although our approach relies on the explicit separation between reflected waves and refracted waves, it does not require assigning a time window to isolate phases as it does in the workflow of Wang *et al.* (2015). This phase identification might be quite cumbersome for the reflection recordings in the complex environments. In addition, we use a velocity-impedance parameterization to alternate the update of the velocity macromodel and the reflectivity, unlike Wang *et al.* (2015) who choose the velocity-density parameterization.

This paper is organized as follows. First, we shall review the intrinsic limitation of FWI and RWI as well as the benefit that can be expected from JFWI with a simple synthetic model. Second, we shall discuss about the choice of a suitable subsurface parameterization to perform the velocity model and reflectivity updates, such that the scale separation between these two models is satisfied. This naturally directs us toward a subsurface parameterization in terms of wavespeed (V_P) and impedance (I_P). Third, we shall review the ingredients of the JFWI workflow that alternates the updates of V_P and I_P before showing its application to the synthetic Valhall case. The experimental setup of this case study is designed in a way such that the diving waves only sample the shallow part of the subsurface. We shall firstly show how the use of diving waves in JFWI improves the reconstruction of the shallow velocities compared with RWI, and how this shallow improvement translates into an improved imaging at greater depths. Then, we shall show that the smooth velocity model built by JFWI can be subsequently used as a reliable initial model for conventional FWI, resulting a broadband velocity model that can be taken for the purpose of structural interpretation.

3.2 Review of FWI and RWI

FWI is a data-fitting procedure during which the subsurface model m is iteratively updated in order to match the synthetic data $d = d(m)$ with the recorded data d_{obs} . The misfit function is conventionally defined as the least-squares norm of the data residuals weighted by a linear operator W , that is

$$C_{FWI}(m) = \frac{1}{2} \|W(d_{obs} - d(m))\|_2^2, \quad (3.2)$$

with an implicit summation over sources, receivers and time. As the wavefield does not linearly depend on the subsurface parameters, this optimization problem is highly non-linear. Many local optimization approaches have been proposed to mitigate this non-linearity and make the misfit function as convex as possible: they might differ in the misfit definition and/or in the domain within which the minimization is performed, but all of them aim to update the model by a linear-search method (Shin et Cha, 2008; van Leeuwen et Mulder, 2010; Luo et Sava, 2011; van Leeuwen et Herrmann, 2013). The local descent direction relies on the gradient of the misfit function with respect to the model parameters, which can be efficiently computed by the adjoint-state method (for a review see Plessix (2006b)). For one source-receiver pair, the gradient can be written in a compact form as

$$\nabla C_{FWI} = u_0 \star \lambda_0, \quad (3.3)$$

where the symbols $u_0 = u_0(m)$ and $\lambda_0 = \lambda_0(m)$ denote the incident wavefield and the back-propagated adjoint wavefield, respectively. For multiple sources and receivers, the gradient is a summation of $u_0 \star \lambda_0$ over all sources and receivers. The adjoint wavefield is computed with a source term that gathers the data residuals associated with all kinds of waves (diving waves, reflected waves, scattered waves etc.). The gradient of the misfit function is computed by a zero-lag cross-correlation between the incident and adjoint fields. This correlation operation, denoted by a single star (\star) in Equation (3.3), embeds, for the sake of compactness, the partial derivative of the forward modeling operator with respect to model parameters (the so-called diffraction pattern) that is cumbersome to be expressed in the time domain. Hiding this kind of complexity in the star symbol will not obscure the governing idea underlying the following derivation of the misfit function gradients. All details about the incident and adjoint fields as well as the gradient expressions are provided in the Appendix A in the framework of frequency-domain FWI allowing for more compact notations. Note that all our implementations are performed in the time domain.

We first illustrate the FWI gradient (Equation 3.3) in the case of a homogeneous subsurface model in which a flat reflector is embedded (Figure 3.2a). The background model is homogeneous with a wrong velocity and does not contain the reflector, so that the source of the adjoint equation contains the residuals of the direct wave and reflected wave (Figure 3.2b). These residuals give rise to two components in the adjoint field denoted by λ_0^d and λ_0^r , respectively. Correlations of the incident wavefield u with these two components of the adjoint wavefield respectively build a wide first Fresnel

zone and a secondary Fresnel zone (the so-called migration isochrone, Figure 3.3a), over which the direct-wave residual and the reflection residual are respectively back-projected (Woodward, 1992). The width of these isophase surfaces gives the spatial resolution with which a point diffractor located in these surfaces is imaged by the current source-receiver pair, according to Equation (3.1). A seismic acquisition generally provides the samplings of the shallow subsurface with both diving and reflected waves, allowing for a broadband imaging of the shallow targets. In contrast, only short-spread reflections sample the subsurface at greater depths, i.e. beyond the penetration depths of the diving waves, leading to a high-wavenumber imaging of deep targets.

RWI focuses on the reflection data to build the subsurface model. The method relies on the prior knowledge of the reflectivity to predict the reflected waves. The misfit function is given by

$$C_{RWI}(m_0) = \frac{1}{2} \left\| W^r \left(d_{obs}^{refl} - d_{pred}^{refl}(m_0, \delta m) \right) \right\|_2^2, \quad (3.4)$$

where the symbols d_{obs}^{refl} and d_{pred}^{refl} denote the observed and predicted reflected waves, respectively, weighted by the linear operator W^r . The low-wavenumber background m_0 and the high-wavenumber reflectivity δm are separated in scales, and RWI seeks to reconstruct the background m_0 only. Following Brossier *et al.* (2015), the gradient with respect to the background model is (for the counterpart in the frequency domain, see Appendix A),

$$\nabla C_{RWI} \approx u_0 \star \delta \lambda^r + \delta u \star \lambda_0^r + \delta u \star \delta \lambda^r, \quad (3.5)$$

where the symbols $u_0 = u_0(m_0)$ and $\delta u = \delta u(m_0, \delta m)$ denote the incident wavefield computed in m_0 and the wavefield scattered by δm , respectively. Similarly, the adjoint wavefield can be decomposed as a component λ_0^r propagating in the background model and a scattered component $\delta \lambda^r$. The scattered wavefield includes forward and backward scatterings (transmissions versus reflections) of any order. Note that the three correlation operations in Equation (3.5) embed two different diffraction pattern operators computed in m_0 and $m_0 + \delta m$, respectively. The detailed expressions are provided in Appendix A.

The RWI gradient corresponding to the one-reflector synthetic model is shown in Figure 3.3(b). The source of the adjoint-state equation contains only the residual of the reflected wave (Figure 3.2c). The first term $u_0 \star \delta \lambda^r$ represents the correlation between the down-going field u_0 and the up-going scattered field $\delta \lambda^r$ (ray path indicated by yellow arrows). This correlation builds a wide first Fresnel zone connecting the reflector and the source position. Similarly, the second term $\delta u \star \lambda_0^r$ gives rise to a mirror Fresnel zone generated by the up-going scattered field δu and the down-going field λ_0^r (ray path indicated by cyan arrows). Known as ‘‘rabbit ears’’, this pair of Fresnel zones allow for a long-wavelength reconstruction of the deep targets where FWI encounters difficulties. The third term represents higher-order migration isochrones. Due to the weak amplitude of the scattered fields δu and $\delta \lambda^r$, these higher-order isochrones are generally of small amplitudes and thus negligible. In Figure 3.3(b), two mirror high-order migration

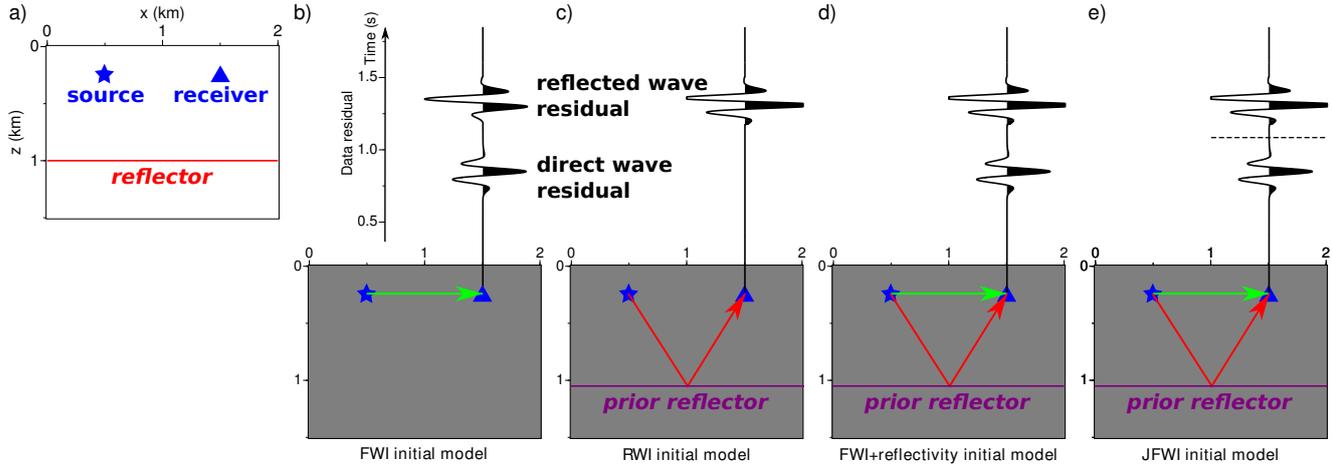


Figure 3.2: Illustration of the initial models and the data residuals used in different FWI approaches. (a) True velocity model including a reflector at 1 km depth. One source and receiver couple is indicated. (b) Homogeneous initial model (m) without reflector and erroneous background velocity for FWI. Only the direct wave (green arrow) is generated. Residuals include the direct wave and the reflected wave residuals. (c) Homogeneous background (m_0) and prior reflectivity (δm) for RWI. Only the reflected wave (red arrow) is modeled. (d) Initial model with prior reflector for FWI. Both direct and reflected waves are modeled. (e) Homogeneous background (m_0) and prior reflectivity (δm) for JFWI. Compared with RWI (c), diving waves are modeled; compared with FWI (d), direct and reflected wave residuals are explicitly separated (see text for details).

isochrones are shown near the reflector position. They are built, on the one hand by the zero-lag correlation between the incident and adjoint fields transmitted across the reflector and, on the other hand by the zero-lag correlation between the incident and adjoint fields reflected from the reflector. The reader is referred to Appendix B for a more detailed description of these higher-order contributions.

The limitation of RWI is the reliance of the exclusive use of the reflected waves, discarding the low-wavenumber information carried by the diving waves (i.e. $u_0 \star \lambda_0^d$ term, see analysis of conventional FWI gradient). One may simply insert the diving waves into the RWI misfit function, augmenting the reflection data residuals in Equation (3.4) with the diving wave residuals (i.e. $d_{obs} - d_{pred}^{div}(m_0) - d_{pred}^{refl}(m_0, \delta m)$). This is equivalent to performing FWI with a prior reflectivity in the initial model (Figure 3.2d). Although all low-wavenumber contributions are gathered in the gradient (Figure 3.3c), the dominant imprint of the high-wavenumber information carried by the migration of the reflection residuals (i.e. $u_0 \star \lambda_0^r$, Figure 3.3(c), red arrows) makes challenging the extraction of the low-wavenumber information carried by the first Fresnel zones associated with RWI (Figure 3.3b). This highlights the necessity to force a scale separation between the low-wavenumber and high-wavenumber components in the FWI formalism, and this is the aim of the following section where we propose an alternative FWI formulation that

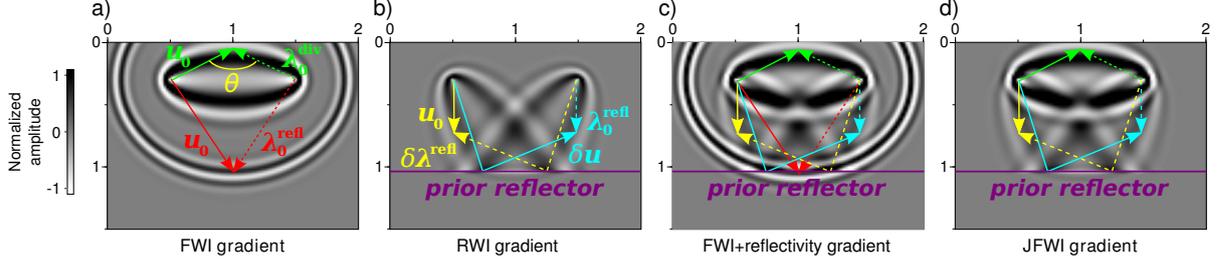


Figure 3.3: Illustration of gradients generated by different FWI approaches. Solid and dashed arrows denote the ray paths respectively followed by the incident and adjoint fields that interfere constructively. (a) FWI gradient combines a low-wavenumber first Fresnel zone (represented by $u_0 \star \lambda_0^d$) between source and receiver, and a high-wavenumber migration isochrone (represented by $u_0 \star \delta \lambda^r$). Note that the first Fresnel zone has a limited penetration in depth. (b) RWI gradient shows two wide first Fresnel zones centered on the two-way paths followed by the reflected wave between the reflector (behaving as a secondary source) and the source/receiver positions. (c) FWI gradient, with a prior reflectivity in the initial model, combines FWI (a) and RWI (b) gradients. Low-wavenumber and high-wavenumber information enter into the gradient, hence breaking down the scale-separation prerequisite. (d) JFWI gradient combines the first Fresnel zone generated by direct-wave residuals and RWI gradient. Compared with (c) the migration isochrone was removed to honor the scale separation. Online animations: <https://drive.google.com/open?id=0Bx0JCM2KZyueU1JqMFBNYkNWWWM>

combines the diving waves and the reflected waves for the velocity macromodel building without generating the high-wavenumber isochrones.

3.3 Methodology

We shall first review the main idea behind JFWI that will allows us to mitigate the high-wavenumber contributions during the velocity model building before discussing the implementation of the JFWI gradient.

3.3.1 Formulation

The governing idea of JFWI is to explicitly separate the contributions of the diving and reflected waves (Figure 3.2e), which means that the misfit function is decomposed as the sum of two terms,

$$C_{JFWI}(m_0) = \frac{1}{2} \|W^d (d_{obs}^{div} - d_{pred}^{div}(m_0))\|_2^2 + \frac{1}{2} \|W^r (d_{obs}^{refl} - d_{pred}^{refl}(m_0, \delta m))\|_2^2, \quad (3.6)$$

where the symbols d_{obs}^{div} and $d_{pred}^{div}(m_0)$ denote the observed and predicted diving waves, and W^d and W^r denote the weighting operator that are applied to the diving and re-

flected waves, respectively. The role of the two weighting operators is to balance the respective contributions of diving and reflected residuals in the misfit function, considering that reflected wavefields have generally weaker amplitudes than the diving waves. The operator W^r can also embed a time-dependent weighting of reflection residuals to enhance the contribution of late reflected arrivals at the expense of early ones in the misfit function. This inversion preconditioning can contribute to balance the amplitudes of the shallow and deep perturbations in the gradient of the misfit function, and hence improve the convergence rate. Other weightings can also be added in the W^d and W^r operators such as offset-dependent weighting, which can be useful to design layer-stripping strategies. These issues will be illustrated in the synthetic example presented in the sequel of this study.

Both the observed data and predicted data should be decomposed into the diving part and the reflection part during the data preprocessing stage. Many preprocessing tools, commonly used to preprocess the reflection data before migration, can be applied here, such as the dynamic time windowing or F-K filtering. While the observed data are preprocessed one and for all before the inversion starts, the decomposition of the predicted data, however, has to be performed at each iteration of JFWI. This might be cumbersome if the decomposition should be refined according to the velocity model update. For the synthetic example shown in the following of this study, we did not need to perform this refinement: we applied to the modeled data, at each JFWI iteration, the same decomposition rule based on the offset-dependent time window as is applied to the recorded data. We could afford to keep the same windowing law at each JFWI iteration because the initial model for JFWI is already accurate enough to predict the first-arrival traveltimes within half the dominant period. If the separation between the diving waves and the reflected waves is awkward, a systematic separation procedure would consist in performing two forward modelings, one in m_0 and one in $m_0 + \delta m$. The first simulation provides the diving wavefield, while the subtraction between the two simulated wavefields provides the reflection wavefield.

The gradient of the misfit function with respect to the background model m_0 is given by (for the counterpart in the frequency domain see Appendix A)

$$\nabla C_{JFWI} \approx u_0 \star \lambda_0^d + u_0 \star \delta\lambda^r + \delta u \star \lambda_0^r + \delta u \star \delta\lambda^r, \quad (3.7)$$

where the symbols λ_0^d and λ_0^r denote the background components of the adjoint field generated by the diving-wave residuals and the reflection residuals, and $\delta\lambda^r$ denotes the scattered component of the adjoint wavefield generated by the reflection residuals, respectively. The first term in Equation (3.7) builds the first Fresnel zone associated with the diving waves, while the second and third terms are those generated during RWI. The key point is that the gradient in Equation (3.7) does not include the $u_0 \star \lambda_0^r$ term associated with the high-wavenumber migration isochrone (Figure 3.3d).

The last three terms in Equation (3.7) encapsulate all of the high-order scattering propagation. Among them, higher-order migration isochrones can be generated and hence, inject undesired high-wavenumber components into the gradient. A two-reflector model, in which internal multiples are generated, is used to illustrate the imprint of these

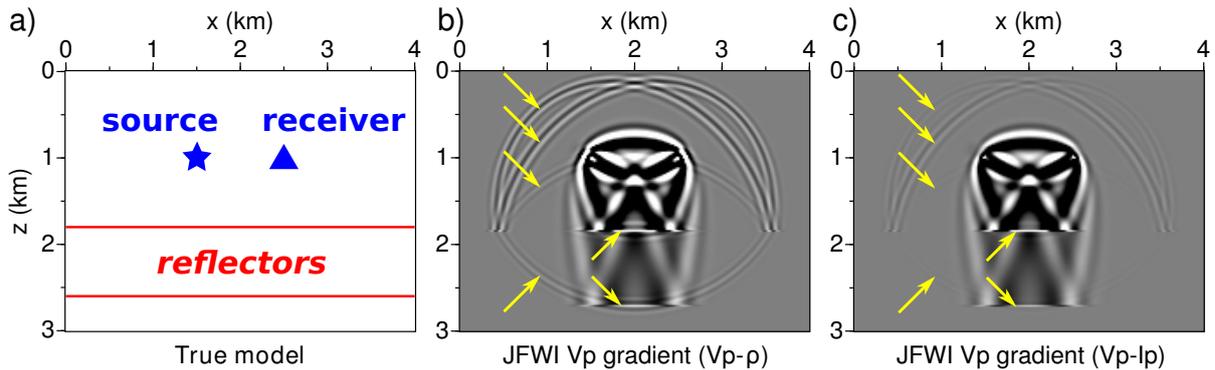


Figure 3.4: Two-reflector model and the corresponding JFWI gradients. (a) Two-reflector model to generate multi-scattered fields. (b) JFWI gradient for the V_P - ρ parameterization. Higher-order isochrones are produced due to constructive interference of the multi-scattered fields (Appendix B for a detailed review of all of these isochrones). (c) Same as (b) except that the V_P gradient was built for the V_P - I_P parameterization. The low-wavenumber components of the gradient are kept unchanged, while the undesired high-wavenumber components are filtered out. Online animation for gradient with V_P - ρ parameterization: <https://drive.google.com/open?id=0Bx0JCM2KZyueQTc1dUhFVE1SQWc>

high-order migration isochrones on the gradient (Figures 3.4a,b). A detailed review of the different high-order contributions is provided in Appendix B. Due to their inconsistent spatial locations, most of these high-order migration isochrones interfere with each other in a destructive way when the contributions from multiple shot gathers are stacked, and thus the gradient tends to be free from these high-wavenumber components. Moreover, in the next section, we shall show how to further reduce their footprint by choosing a proper subsurface parameterization for JFWI.

3.3.2 Mitigation of high-order isochrones by choosing suitable subsurface parameterization

In most acoustic FWI approaches, the subsurface is parameterized by the velocity (V_P) and the density (ρ) (e.g., Wang *et al.*, 2015). The diffraction patterns associated with these parameters are shown in Figure 3.5(a) and (b), respectively. They show two wavefields scattered by a single V_P and ρ point perturbation in a homogeneous background. These wavefields represent the partial derivatives of the incident wavefield with respect to the model parameters that undergo a perturbation. Zero-lag correlations of this kind of partial derivative wavefields at the receiver positions with the data residuals form the FWI gradients. The amplitude variation of the partial derivative wavefield with the scattering angle θ (i.e., the diffraction pattern) controls the effective range of scattering angles that will confer the spatial resolution to the FWI gradient for one parameter class (see Operto *et al.* (2013) for a recent tutorial on multi-parameter FWI). The wavefield scat-

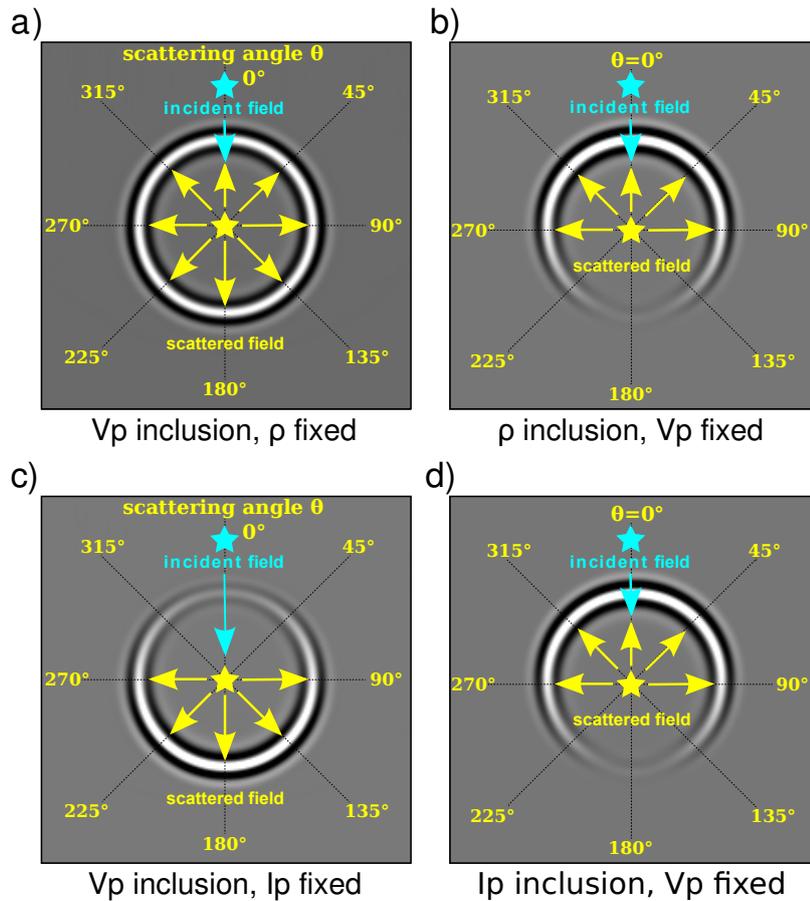


Figure 3.5: Diffraction patterns of V_P - ρ (a) and V_P - I_P (b) parameterizations. Single diffractors are located in the center of the homogeneous background. A source at the vertex of the diffractor generates the scattered field, whose amplitude variation with the scattering angle is representative of the diffraction pattern of the parameter. Note how the diffraction pattern of V_P changes depending on the other parameter involved in the parameterization (ρ versus I_P). See text for details.

tered by the V_P diffractor shows an isotropic radiation pattern, conferring a broadband wavenumber content to the V_P gradient. On the other hand, the wavefield scattered by the ρ diffractor has significant amplitudes at small-to-intermediate θ , conferring a narrower (high) wavenumber content to the ρ gradient. This V_P - ρ parameterization is a natural choice for FWI to reconstruct a broadband V_P model, keeping in mind that cross-talks or leakage between V_P and ρ are necessarily present in the high-wavenumber part of the V_P and ρ gradients. A careful accounting for the Hessian should help to remove this leakage in the FWI subsurface models.

In contrast, the V_P - I_P parameterization leads to a natural scale separation between the two parameter classes. The corresponding diffraction patterns are shown in Figures

3.5(c) and (d). The wavefield scattered by the V_P diffractor has significant amplitudes for large θ , leading to a V_P gradient with a low-wavenumber content. Conversely, the wavefield scattered by the I_P diffractor has significant amplitudes for small θ , leading to an I_P gradient with a high-wavenumber content. Compared with the V_P - ρ parameterization, the V_P - I_P parameterization is more suitable to satisfy the scale-separation condition underlying JFWI, although it will prevent the reconstruction of the high wavenumbers in the V_P model.

According to this diffraction-pattern analysis, we choose the V_P - I_P parameterization to perform JFWI, where the low-wavenumber part of the subsurface m_0 is parameterized by V_P and the high-wavenumber part δm is parameterized by I_P (Equation 3.6). The effect of the subsurface parameterization on the JFWI gradient is illustrated in Figure 3.4(c) by the effective attenuation of the high-order isochrones when the V_P - I_P parameterization is used. (Another advantage of the V_P - I_P parameterization over the V_P - ρ parameterization is described in Snieder *et al.* (1989), their figures 2 and 3.)

3.3.3 Implementation

In order to perform a computationally efficient implementation of JFWI, we rewrite Equation (3.7) in a more compact form by regrouping scattered and background wavefields, which is

$$\begin{aligned} \nabla C_{JFWI} &= u_0 \star \lambda_0^r + \delta u \star \lambda_0^r + u_0 \star \delta \lambda^r + \delta u \star \delta \lambda^r + u_0 \star \lambda_0^d - u_0 \star \lambda_0^r & (3.8) \\ &= \underbrace{u^r \star \lambda^r}_{G_1} + \underbrace{u_0 \star (\lambda_0^d - \lambda_0^r)}_{G_2} & (3.9) \end{aligned}$$

where we have (re-)introduced the total reflection field $u^r = u_0^r + \delta u^r$ and the total reflection adjoint wavefield $\lambda^r = \lambda_0^r + \delta \lambda^r$. Note that Equation (3.8) is a true identity even though we have concealed the partial derivatives of the modeling operator inside the symbol \star . Strict demonstration is provided by the developments starting from Equation (A.10) to Equation (A.14) in Appendix A.

This alternative expression of the gradient leads to a workflow that sequentially computes two quantities G_1 and G_2 (Algorithm 3.1). The quantity represented by G_1 is computed in $m_0 + \delta m$ (Figure 3.3c) by conventional FWI of the reflection data (hence, the first Fresnel zone associated with the diving waves represented by $u_0 \star \lambda_0^r$ are not embedded in this expression), while the quantity represented by G_2 is computed in m_0 by conventional FWI of the diving waves and reflection data (Figure 3.3a), in which the sign of the reflection data residuals is reversed. This reversal, indicated by the minus sign in front of λ_0^r in Equation (3.9), cancels by subtraction the first-order migration isochrone involved in G_1 , making the total gradient $G_1 + G_2$ dominated by low-wavenumber components. Moreover, computing G_1 or G_2 requires the same tasks as those performed during FWI to evaluate the FWI gradient in the time domain (the boundary-saving strategy is used to alleviate the memory load, Clapp (2008)). These tasks can be summarized as follows: simulate the modeled wavefield and save the values on the model boundaries,

compute the data residuals at receiver positions, compute the adjoint field by taking the residuals as the source function, and re-simulate the modeled wavefield in reverse time by using the boundary values as the Dirichlet condition, cross-correlate on the fly the two fields. Therefore, for one evaluation of the FWI gradient, we need to do three modelings and save two fields at maximum. As a combination of G_1 and G_2 , Algorithm 3.1 requires to perform six modelings and to store two fields for the cross-correlation. Consequently, the workflow to evaluate the JFWI gradient has a time complexity two times higher than the one of conventional FWI but the memory requirement is the same (see Table 3.1). In addition, the cost of JFWI is the same as the cost of RWI.

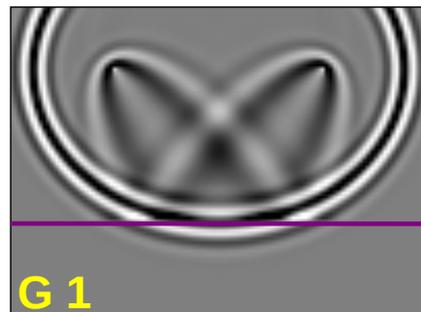
Table 3.1: Cost comparison

Cost	FWI	RWI	JFWI
Modelings	3	6	6
Storage	2	2	2

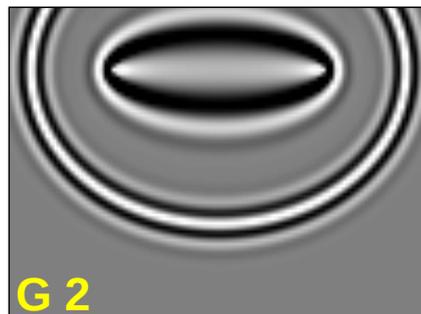
Algorithm 3.1 Efficient evaluation of ∇C_{JFWI}

Step 1

USE prior reflectivity
DO forward modeling: $u = u_0 + \delta u$
USE reflected wave residuals
DO backpropagation and correlation
GET $G_1 = \text{rabbit ears} + 1^{st}\text{-order isochrone}$


Step 2

USE smooth models (NO reflectivity)
DO forward modeling: u_0
USE diving wave residuals
USE same reflected wave residuals as in Step 1
 but with a minus sign
DO backpropagation and correlation
GET $G_2 = \text{diving-wave first Fresnel zone}$
 – $1^{st}\text{-order isochrone}$


Step 3

DO $G_1 + G_2 = \nabla C_{JFWI}$

If more subtle data separations are used, i.e. involving the decomposition of the modeled wavefield on the fly by subtraction of the full field and the background field, the implementation of G_1 in equation 3.9 is impossible. This is because G_1 is only related to the predicted reflection data, which would not be readily available at the time of the adjoint simulation when the data separation is performed on the fly with the aforementioned subtraction procedure. In this case, we would recommend computing an approximation of ∇C_{JFWI} with the following workflow: (1) Compute the conventional FWI gradient in the model $m_0 + \delta m$ using both the diving and reflected wave residuals (Figure 3.3c). (2) Build the first-order migration isochrones by conventional FWI performed in model m_0 using the reflection residuals as the source of the adjoint equation. At this stage the predicted reflection data becomes available since the subtraction of the modeled full field and the background field can be readily performed. (3) Subtraction of the two quantities gives a low-wavenumber quantity, which is a good approximation of ∇C_{JFWI} . A relative error of 2% is found in the synthetic Valhall case study presented later. This implementation requires the same resources as the one discussed above (Table 3.1, third column).

3.4 Integrated workflow of velocity (V_P) and impedance (I_P) inversion

So far we have derived the principles of JFWI for low-wavenumber velocity building, and choose the V_P - I_P parameterization for high-wavenumber isochrones mitigation. In the following we shall design a complementary imaging tool that provides a high-wavenumber model of the subsurface, and combine it with JFWI to have an integrated inversion workflow.

3.4.1 I_P inversion by using short-offset reflection data

We perform the high-wavenumber imaging by waveform inversion using only the short-offset reflection data (referred to as IpWI in the following). According to the previous diffraction-pattern analysis, the subsurface is parameterized by V_P - I_P and only the I_P parameter is updated using the V_P model as the background model. The corresponding misfit function is given by

$$C(I_P) = \frac{1}{2} \left\| W^r \left(d_{obs}^{refl} - d_{pred}^{refl}(V_P, I_P) \right) \right\|_2^2. \quad (3.10)$$

The corresponding gradient (Figure 3.6) contains the 1st-order isochrones similar to the ones which would have been imaged by a migration processing

$$\nabla C = u_0 \star \lambda_0^r = u_0 \star B^{\dagger-1} (d_{obs}^{refl} - d_{pred}^{refl}), \quad (3.11)$$

where the forward modeling operator is denoted by the symbol B .

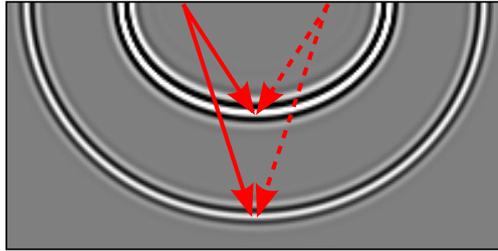


Figure 3.6: Impedance gradient resulting from the migration of one residual seismogram in a homogeneous background. The true model contains two reflectors. The gradient only contains 1st-order isochrones. Ray paths of the incident and adjoint fields for the source-receiver pair are illustrated.

The only noticeable difference from a least-squares migration is that the observed reflection data is used as the source function to compute the back-propagated field in the migration process, whereas the source function of the adjoint-state equation in this IpWI is the reflection data residual.

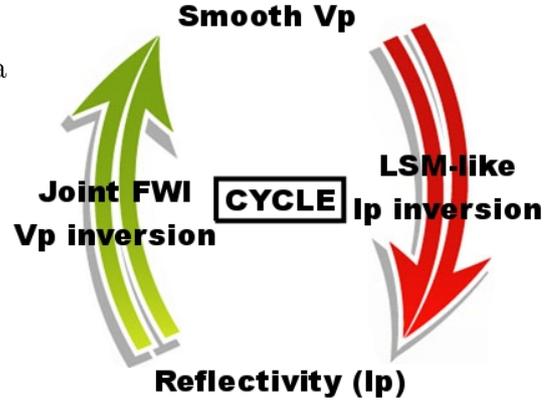
3.4.2 Cycle workflow of V_P - I_P imaging

Recall that the V_P - I_P parameterization naturally leads to the scale separation between the two parameter classes. The low wavenumbers of V_P are reconstructed by JFWI, which requires a prior high-wavenumber I_P model built by IpWI. Conversely, the impedance imaging requires a background V_P that can be provided by JFWI. Therefore, it is natural to combine JFWI and IpWI to reconstruct these two parameters (a similar strategy can be found in [Ma et Hale, 2013](#)). A critical issue is that, once the V_P model has been modified by JFWI to some extent, a new I_P reflectivity model should be generated accordingly in order to match the reflection data before continuing the V_P update, and the initial smooth impedance model should be used as the starting model for this new IpWI implementation (i.e. removing the imaged reflectivity) ([Brossier et al., 2015](#)). This directs us toward a cycle workflow in which the V_P model and the I_P model are repeatedly updated in an alternate way by several non-linear iterations of JFWI and IpWI (Algorithm 4.1), and during each JFWI and IpWI step the two parameters are considered independently from each other. As the velocity model is not expected to be accurate during the cycle workflow, we build the impedance model from very short-offset reflections to enhance the focusing of the reflectivity image, even if mispositioned in depth, in order to avoid any biases associated to the residual move-out in the depth-migrated domain.

Algorithm 3.2 Cycle Workflow of V_P - I_P ImagingSmooth initial $V_P^{(0)}$ and $I_P^{(0)}$, $k = 1$ **repeat**

USE $V_P^{(k-1)}$, $I_P^{(0)}$ (NO reflectivity)
USE short offsets
DO I_P inversion by IpWI of reflection data
GET high-wavenumber $I_P^{(k)}$
 (building reflectivity)

USE $V_P^{(k-1)}$, $I_P^{(k)}$
USE full offsets
DO V_P inversion by JFWI
GET low-wavenumber $V_P^{(k)}$
 (updating background)

Cycle number $k = k + 1$ **until** convergence

3.5 Synthetic example: Valhall case study

3.5.1 Experimental setup

We apply our cycle workflow on a synthetic case representative of the Valhall oil field. The true I_P and V_P models are shown in Figures 3.7(a) and (c), respectively. The sea floor is at 62.5 m depth, above the sediment layers that overlay several low-velocity gas zones depicted in red. From 2.5 km to 3.3 km depths, the high-velocity oil reservoir is separated from the low-velocity gas zones by a cap rock of anticline structure. The sand is laid below the reservoir with smooth V_P , I_P variations, supported by the bedrock at 5 km depth. The model, which is 8.8 km in width and 5.2 km in depth, is discretized by 418×704 grid points with a grid interval of 12.5 m.

We use a Gaussian filter to smooth the true velocity model (excluding the water layer), and then extract one vertical profile to build the 1D initial model for JFWI and IpWI (Figures 3.7b and d). This initial velocity model captures the large-scale variation of the true model and discards all features of the gas zones. Therefore, the main task of JFWI is to reconstruct the gas zones. Based on a former analysis (Prioux *et al.*, 2011, their Figure 2b), more than 14 km of offset would be needed to record diving waves that propagate at reservoir depths. In this study, the maximum offset is set to 6 km and the diving waves reach a maximum penetration-depth of 1.5 km, which is shallower than most of the gas layers. Therefore, only the reflected waves can contribute to the reconstruction of the deep targets. The initial impedance model is shown in Figure 3.7(b), which was built by taking a two-layer density model (setting 1000 kg/m^3 for the water and 2000 kg/cm^3 below).

The forward problem is solved by a classical $\mathcal{O}(\Delta t^2, \Delta x^4)$ staggered-grid finite-difference method. The absorbing boundary condition implemented with perfectly-matched layers (PMLs, e.g. [Bérenger, 1994](#); [Komatitsch et Martin, 2007](#)) is applied along each edge of the model. Therefore, no surface-related multiples are generated in the data. We exclude this kind of multiples for a purpose to avoid potential cycle skipping among primary reflections and secondary reflections. More realistic case involving the surface-related multiples will be discussed in Section 3.5.3. The source function is a Ricker wavelet with a peak frequency of 6.25 Hz. We consider a streamer acquisition of 80 shots, with the first shot gather being shown in Figure 3.7(e). The reflection phases with zero-offset two-way traveltimes ranging from 0.4 s to 2.8 s are reflections from the gas layers. The reflections from the cap rock and the reservoir have zero-offset two-way traveltimes ranging from 2.8 s to 3.4 s, followed by the multi-scattered phases (indicated by the yellow arrows) especially the one arriving at 3.8 s (by the yellow ellipse). These multi-scattered waves are generated from the highly-reflective cap rock and from the edges of the gas layers. As the direct/diving waves do not significantly overlap the reflected waves in time, we simply apply a linear time-offset window to the data set for the purpose of data separation, defined by the following formula:

$$t_{\text{sep}} \text{ (s)} = \text{offset (km)} / 1.5 \text{ (km/s)} + 0.33 \text{ (s)}, \quad (3.12)$$

where t_{sep} is the time of the window boundary as a function of offset (indicated by the red dashed line in Figure 3.7e), the slope of the linear boundary is given by the water wavespeed (1.5 km/s) and the intercept is equal to 0.33 s according to the dominant period of the Ricker wavelet. Furthermore, we estimate that the direct/diving waves have amplitudes ten times higher than those of the reflected waves. This prompts us to leave the amplitudes of the direct/diving waves unchanged in the misfit function (i.e. W^d equals to the identity matrix), while the reflection residuals are multiplied by a factor of ten in the misfit function.

Having been observed in several tests, JFWI tends to match the recorded multi-scattered phases (indicated by the yellow arrows and ellipse in Figure 3.7e) with the modeled primary reflections if the full reflection wavefield is processed in one go, which implies that JFWI also suffers from non-linearity as FWI does. Therefore, we apply a progressively increasing time window procedure ([Kolb et al., 1986](#)) to the data to enhance the robustness of JFWI. This is implemented by muting the full-offset reflection data after 3.5 s during first few cycles to remove the contribution of the multi-scattered waves in the misfit function, then gradually restore their amplitude to their original level during later cycles.

The velocity gradients that are computed in the initial smooth model by FWI, RWI and JFWI are shown in Figures 3.7(f-h), respectively. As expected, the FWI gradient shows the limited penetration depth (~ 0.8 km) of the first Fresnel zones associated with the diving waves and a high-wavenumber content at greater depths generated by the stack of the migration isochrones (Figure 3.7f). The RWI gradient shows how the migration isochrones were avoided by assuming a known reflectivity in the waveform inversion formalism, while low-wavenumber components were injected at all depths along the transmission wavepaths of the reflected waves that are predicted by using the prior reflectivity

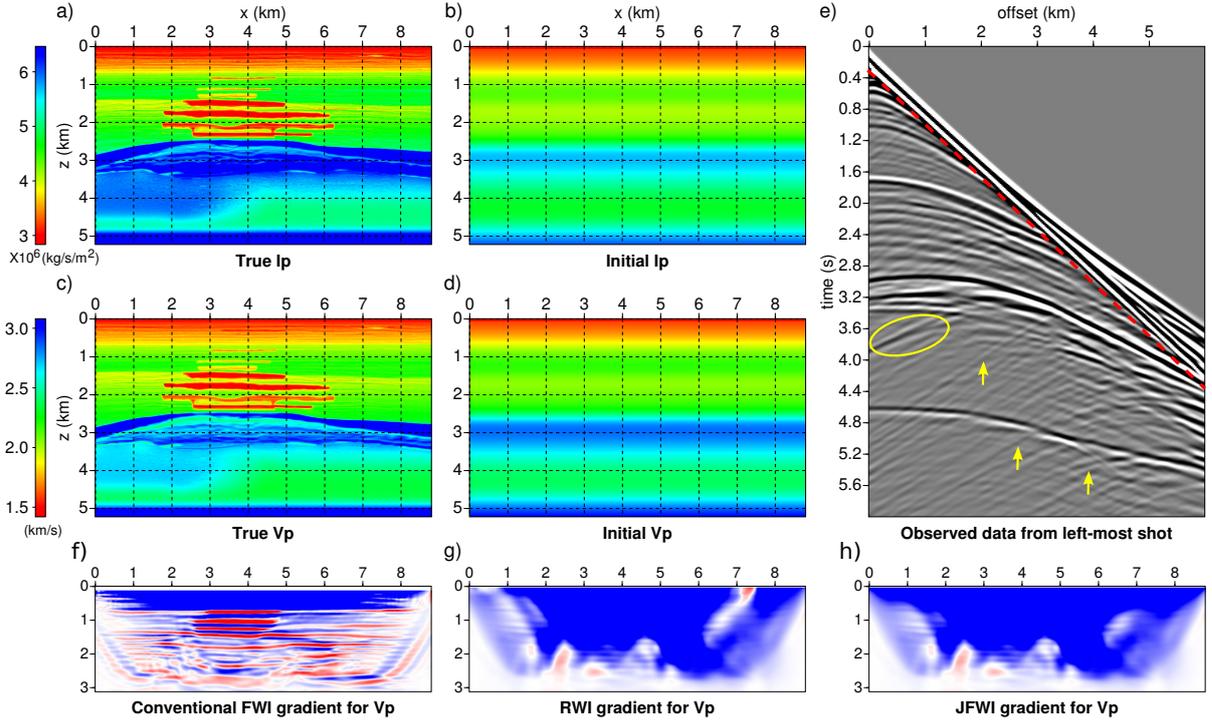


Figure 3.7: Synthetic Valhall case study. (a) True I_P model. (b) Initial I_P model. (c) True V_P model. (d) Initial V_P model. (e) First shot gather. Diving and reflected waves are recorded (separated by the time-offset boundary delineated by the red dashed line), as well as several multi-scattered waves indicated by the yellow ellipse and arrows. (f) Conventional FWI gradient, (g) RWI gradient and (h) JFWI gradient for V_P .

(Figure 3.7h). The JFWI gradient yet highlights how a deficit of low-wavenumber coverage in the shallow part of the RWI gradient (Figure 3.7g, see at 0.5 km and 7 km in distance) can be filled through the diving wave contribution (Figure 3.7h). The optimization method relies on the l -BFGS quasi-Newton approach (Nocedal, 1980). We perform 10 non-linear iterations of IpWI and 20 non-linear iterations of JFWI during each cycle of the workflow. We use only offsets smaller than 200 m to perform IpWI in order to increase the linearity of the problem (see Section 4).

3.5.2 Results and discussions

3.5.2.1 Results and comparisons

Figure 3.8 displays a selection of I_P perturbations (namely I_P updates) and V_P models obtained at different cycles of the workflow (Algorithm 4.1). The initial I_P perturbation is zero and the initial V_P is smooth (Cycle 0). After the first cycle, the I_P perturbation, computed using the initial V_P model as the background model, is not well focused especially for the sand-bedrock interface. The image of the gas-zone reflectors in the first I_P

perturbation are mispositioned at excessive depths due to the overestimated velocities in the initial model. Large-scale V_P variations of the gas zones start to show up at about 1.8 km depth in the first model built by JFWI. This velocity update contributes to move the gas-zone reflectors at their correct position during IpWI in the following cycles. We perform 22 cycles of the IpWI+JFWI workflow. The final V_P and I_P models are shown in Figure 3.9(b) and (e), respectively. The reflectors have been moved at their correct positions, as JFWI has injected long-to-intermediate wavelengths into the velocity model. In particular, the zone of influence of the low-velocity gas-zone between 2.5 km and 5 km in distance has been well delineated. The long-wavelengths of V_P between the reservoir and bedrock are rarely imaged due to two facts: most of the incident energy is reflected backwards by the hard cap rock and the smooth variations in the sand zone provides few reflection information.

We also apply FWI and RWI on this data set to highlight the add-value provided by JFWI. Results are shown in Figures 3.9(a), (c) and (d). FWI is implemented with the V_P - ρ parameterization and only the V_P model is reconstructed. RWI is performed with the V_P - I_P parameterization. The same cycle workflow as was used for JFWI is used again to perform RWI, except that the contribution of diving waves is discarded by RWI.

The final V_P model obtained by FWI is roughly the superimposition of short-wavelength velocity perturbations on the smooth initial velocity model. The inversion clearly fails to update the long-to-intermediate wavelengths of the velocity model because of the inaccuracy of the initial model and the lack of long offsets. The short-wavelength components of the reconstructed V_P are poorly focused due to the inaccurate long-wavelength components.

In contrast, both RWI and JFWI have reconstructed the long wavelengths of the gas zones to some extent. However, the final V_P model built by JFWI is significantly more accurate than the one built by RWI, especially in the shallow part where diving waves penetrate. An inspection of the vertical profiles extracted from the true model, RWI model and JFWI model supports this statement (Figure 3.10, to be discussed later). The final I_P models obtained by IpWI, from either JFWI or RWI V_P models, are purely superimpositions of the short-wavelength components imaged by IpWI on the smooth initial I_P model, according to the diffraction pattern of the I_P parameter (Figure 3.5d). In summary, the scale-separation condition has been fulfilled through this proposed cycle inversion.

Figures 3.10(a) and (b) show the logs of the I_P perturbations obtained by IpWI and the V_P models inferred from RWI and JFWI at $x = 3.75$ km, respectively. Above $z = 1$ km, the I_P perturbations computed from the RWI and JFWI V_P models are merely equivalent. However, the RWI V_P model is clearly less accurate than the JFWI V_P model, particularly in the shallow part where the diving waves penetrate. From 1.5 km to 2.5 km depth, due to the accumulation of inaccuracies from the near surface, the velocities of the RWI model are overestimated and prevent the correct positioning in depth of the I_P perturbations. This highlights that, indeed, an accurate near-surface reconstruction is also critical to properly image deeper zones.

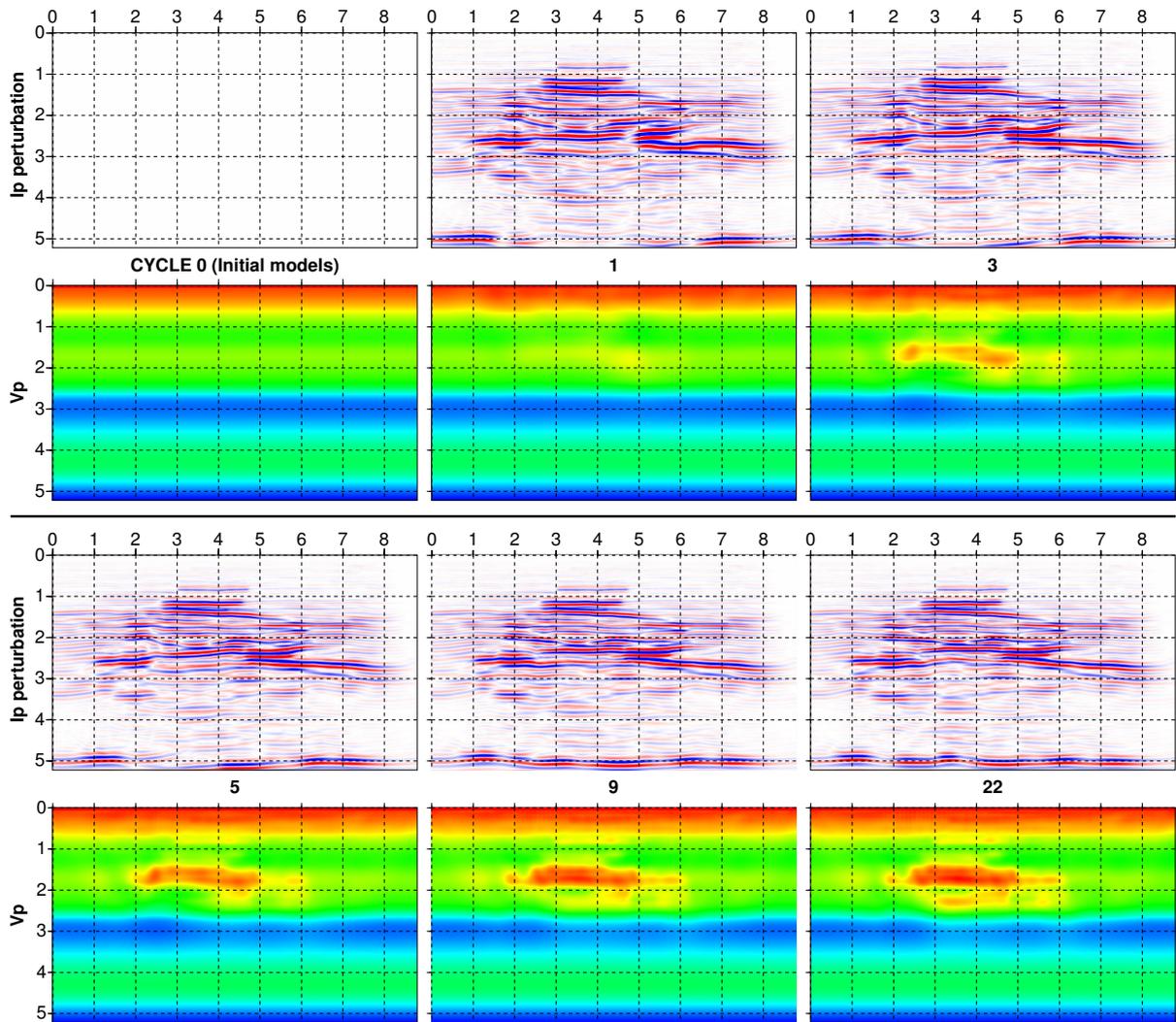


Figure 3.8: Selection of I_P perturbations and V_P models from the cycle IpWI+JFWI workflow. Long wavelengths are built in V_P while short wavelengths are imaged in the I_P perturbation model. Online animation of all intermediate updates: <https://drive.google.com/open?id=0Bx0JcM2KZyueQThsakJHbW9odm8>

3.5.2.2 Quality control by common image gathers

We further assess the quality of the velocity models inferred from RWI and JFWI by generating common image gathers (CIGs) in the offset-depth domain (Figure 3.11). These CIGs are computed in the initial V_P , RWI V_P and JFWI V_P models by reverse-time migration using the same modeling engine as was used during JFWI and IpWI.

Compared with the initial V_P model, both the RWI and JFWI V_P models can significantly improve the flatness of the events in the CIGs. The improvement provided by JFWI compared with RWI is more subtle as the CIGs generated by RWI are already

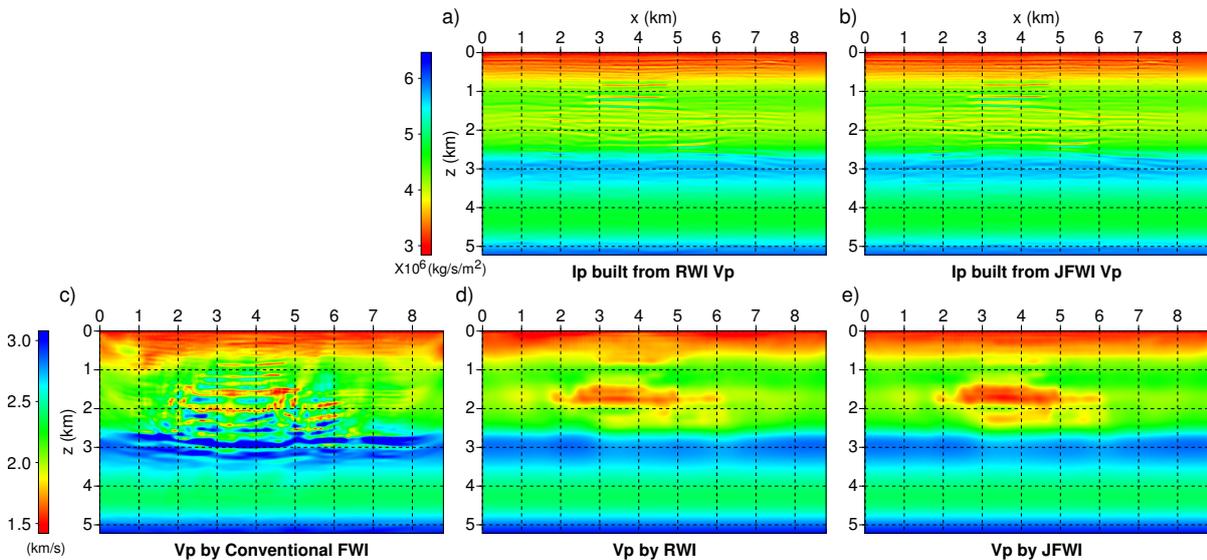


Figure 3.9: Final subsurface models of different waveform inversion approaches. (a-b) Impedance models obtained by RWI (a) and JFWI (b). (c-e) V_P model obtained by FWI (c), RWI (d) and JFWI (e). See text for explanations. Online animation of all intermediate updates: <https://drive.google.com/open?id=0Bx0Jcm2KZyueS2JjMW01TTA0NEE>

quite flat. Nonetheless, we still show that shallow events are flatter in the CIGs inferred from the JFWI V_P model relative to those inferred from the RWI V_P model (Figure 3.11, horizontal yellow arrows). This manifests the diving wave contribution used in JFWI as well as the difficulty of the reflection-based imaging methods to update the shallow part of the subsurface. This point was illustrated with a real-data case study from Valhall by [Prioux *et al.* \(2011\)](#), who showed that FWI of diving waves and reflected waves improved the flatness of the CIGs in the first kilometers in depth of the subsurface compared with the CIGs inferred from a reflection traveltime tomography velocity model. Some improvements achieved by JFWI compared with RWI are also shown at the cap rock level (Figure 3.11, red arrows), which highlight how the more accurate shallow velocity reconstruction impacts on the focusing of the deep reflector images. The horizontal reflector at 5 km depth is also better imaged in the JFWI model than in the RWI model (Figure 3.11, vertical green arrow).

3.5.2.3 Fitting amplitudes

Due to the geometrical spreading effects that are incompletely removed during the FWI process, the amplitude of the impedance perturbations decreases with depth. Without an accurate reconstruction of the impedance contrasts, the amplitude of the modeled reflected waves can be significantly smaller than the observed amplitude. In order to

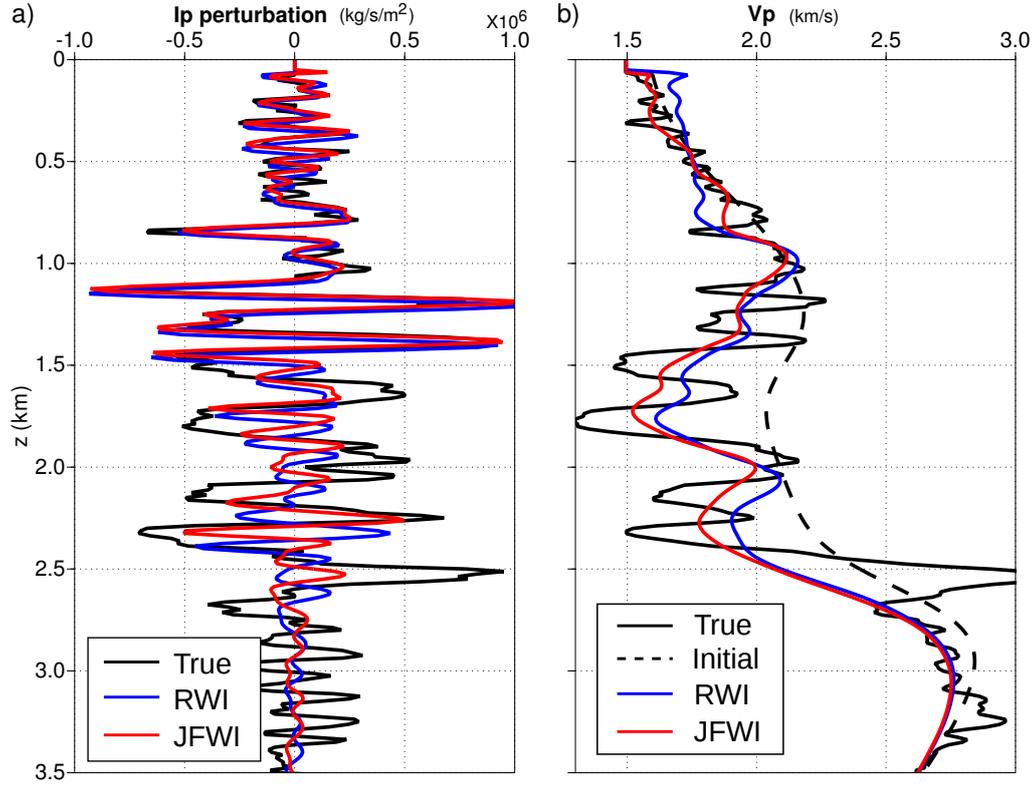


Figure 3.10: (a-b) Logs of I_P perturbations (a) and V_P models (b) at $x = 3.75$ km for IpWI+RWI and IpWI+JFWI. Accounting for diving waves in JFWI improves the velocity reconstruction in the shallow part, which translates into an improved imaging of the deep velocity and the impedance structures.

improve the amplitude fit, we precondition the gradient to strengthen the deep perturbations at the expense of shallow ones and perform 40 IpWI iterations with the preconditioned l -BFGS optimization scheme, starting from the final V_P model of JFWI (Figure 3.9e). We still use offsets smaller than 200 m.

The number of iterations is taken twice as before such that tiny differences in the data amplitude can influence the model update. The refined impedance perturbations computed in the RWI and JFWI velocity models are shown in Figure 3.12. Compared with the previous I_P models (Figures 3.9 a and b), the image of the deep reflectors, such as the cap rock-reservoir interface, has been significantly enhanced. Note also how the geometry of the reservoir is much more accurately delineated in the I_P model computed by using the JFWI velocity model compared with the one using the RWI velocity model.

For the sake of completeness, we plot the vertical profile of the new I_P perturbation models at $x = 3.75$ km in Figure 3.13. The amplitude of the impedance perturbations is now much better estimated in the deep part of the model.

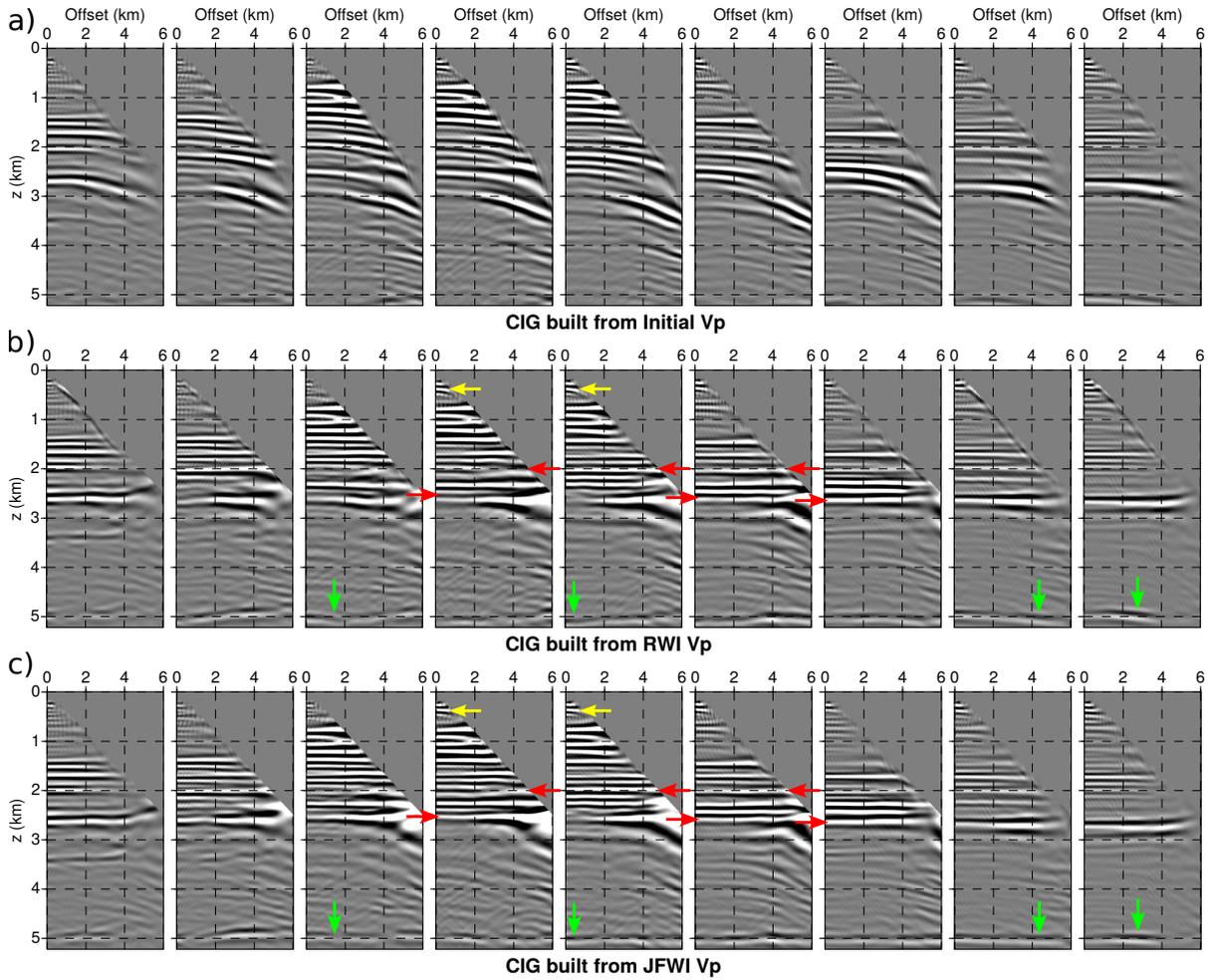


Figure 3.11: Common image gathers at positions $x = \{1.88, 2.50, 3.13, 3.75, 4.38, 5.00, 5.63, 6.25, 6.88\}$ km from the initial V_P (a), RWI V_P (b) and JFWI V_P (c) models. Improvements to the event flatness provided by JFWI compared with RWI are pointed by horizontal yellow arrows in the shallow part, red arrows at the cap rock level and vertical green arrows at 5 km depth.

The first shot gather computed in the true models and in the final RWI and JFWI models are compared in Figures 3.14 and 3.15. Phases and amplitudes in the JFWI-calculated data agree quite well with those of the observed data, except for the multi-scattered waves (e.g. time = 4.4 s to 5.2 s at offset = 6 km). Nevertheless, the absence of the multi-scattered waves in the calculated data helps us avoid somehow the cycle-skipping issues: the phase ended in time = 4.6 s at offset = 6 km seems not to be cycle-skipped. As the sand zone is hard to be recovered (due to few reflection information from the seismogram), the move-out of the latest reflection coming from the sand-bedrock interface is less properly matched especially at offset = 4 km. More convex

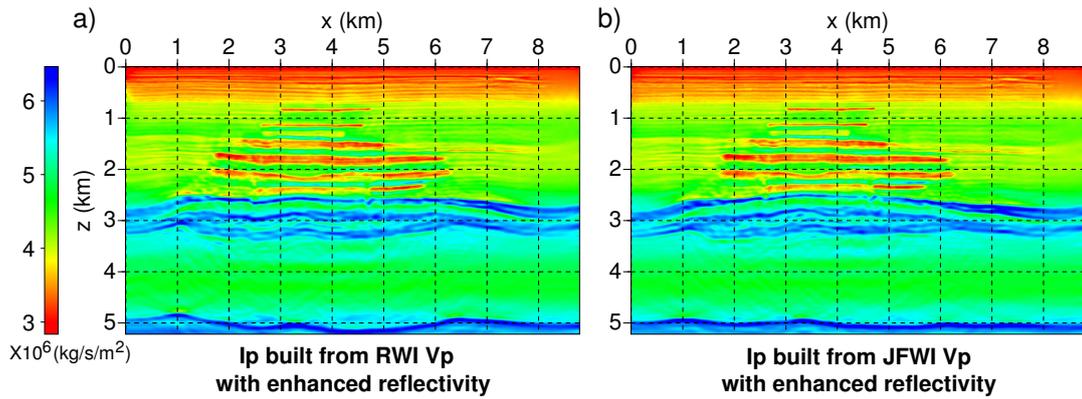


Figure 3.12: I_P models after 40 iterations of IpWI performed with preconditioned l -BFGS optimization, using RWI (a) and JFWI (b) V_P models as background models (Figures 3.9 (d) and (e), respectively).

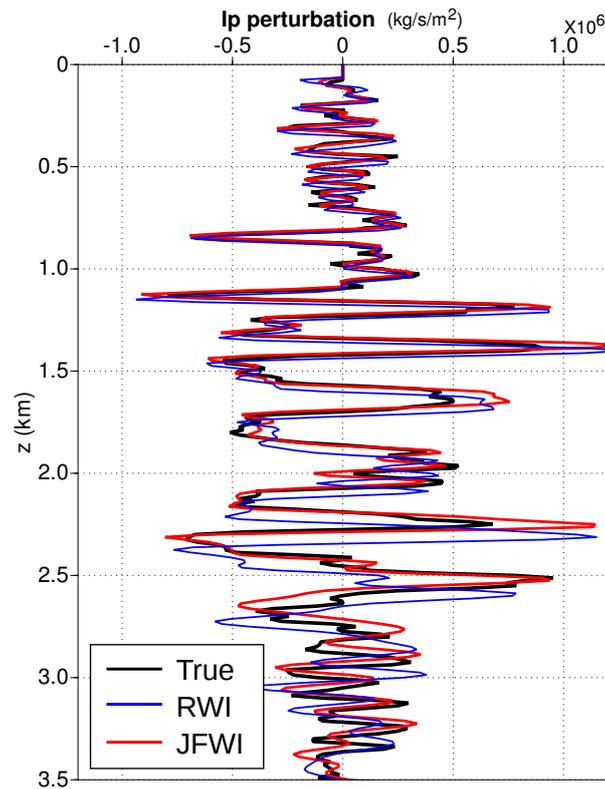


Figure 3.13: Logs of enhanced reflectivity (I_P update) at $x = 3.75$ km. Compared with previous I_P logs (Figure 3.10a), deeper reflectivity can be matched by preconditioned l -BFGS optimization starting from the RWI or JFWI V_P models. Still, JFWI performs better than RWI in terms of depth positioning.

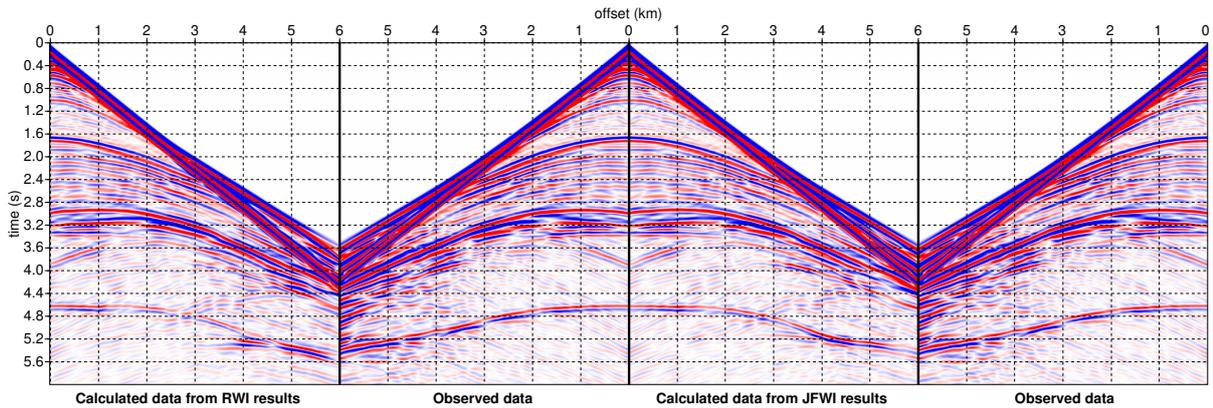


Figure 3.14: Comparison of the first shot gather computed in the true V_P model (second and fourth panels) and in the final RWI (first panel) and JFWI (third panel) models. See text for details.

misfit functions (Luo et Schuster, 1991; Luo et Sava, 2011; Brossier *et al.*, 2009a, 2010, 2015; Ma et Hale, 2013; Warner et Guasch, 2014) would be helpful to relax the cycle-skipping issue raised by waveform-difference misfit functions, and allow large traveltimes shifts for the inversion procedure to match the full-offset reflection phases. On the other hand, the data calculated in the RWI final models match the observed data at short offsets, but fail at long offsets (e.g. time = 3.6 s, 5.4 s at offset = 6 km). This means that the diving-wave information is also critical for the matching of the reflection data (phases, move-outs etc.), and should not be discarded in the high-resolution imaging techniques.

3.5.2.4 Broadband imaging of V_P

The impedance model in Figure 3.12 could be used for geophysical interpretation. Alternatively, a broadband V_P model is also very helpful to understand the structure of the subsurface and the rock properties, but conventional approaches like FWI may fail in building such a V_P model from a crude initial model due to the lack of low frequencies (Figure 3.9a). Joint FWI can be used as a robust tool for initial model building since it can build the long wavelengths that are required to perform reliable FWI (Figure 3.9e).

Figure 3.16(a) shows the result of FWI starting from the JFWI V_P model (Figure 3.9e) under the V_P - ρ parameterization. The two-layer density model that was used as the initial model for JFWI is used here as the background model, and is kept fixed during this FWI implementation. The final FWI velocity model matches quite closely to the true velocity model, except in the deep part due to the lack of illumination. This result confirms the relevance of JFWI as a robust tool to build an initial velocity model for FWI, leading to a two-step velocity model building workflow of successive JFWI and

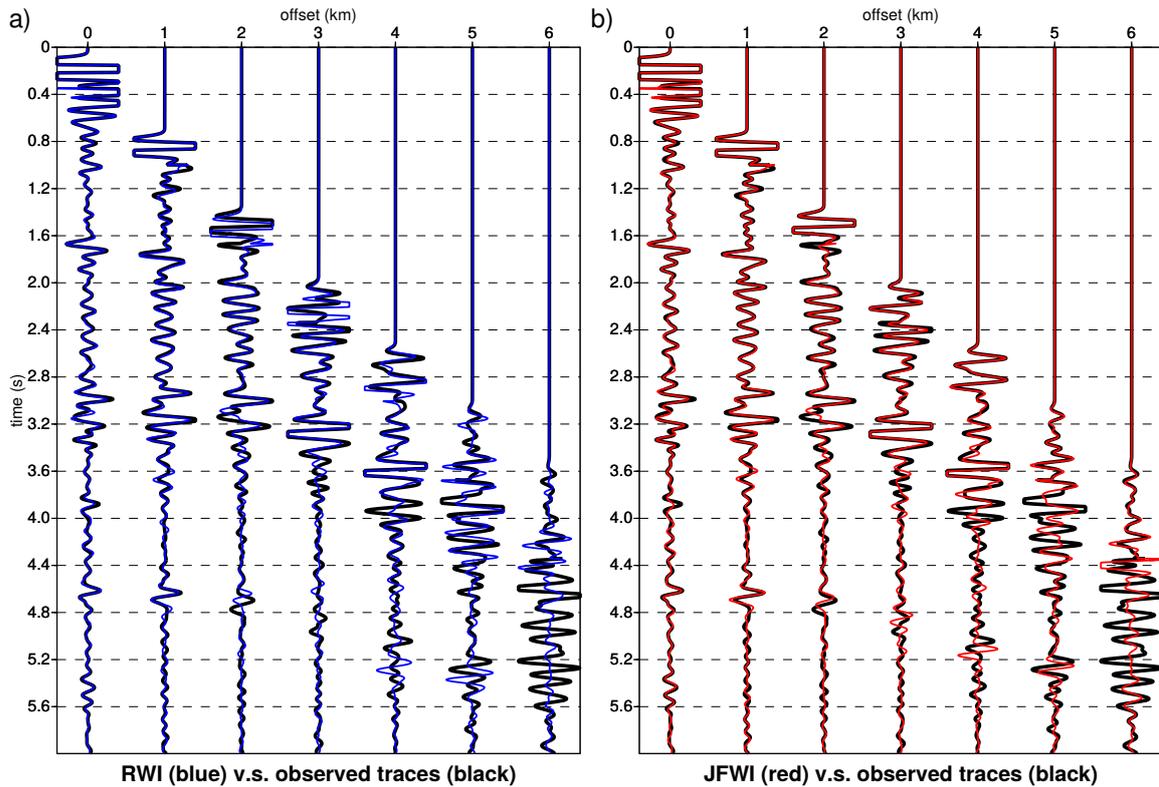


Figure 3.15: (a) Direct comparison between seismograms computed in the true V_P model (black lines) and the final RWI model (blue lines) for the first shot gather (Figure 3.14). (b) Same as (a) for seismograms computed in the JFWI model (red lines). See text for details.

FWI. However, this V_P model is prone to contain imprints from V_P - ρ cross-talks: the short-scale heterogeneities that are missed in the background density model might be interpreted as the short-scale velocity perturbations by FWI, leading to an erroneous velocity estimation. This might explain some amplitude mismatches between the true and the FWI velocities in the log profile (Figure 3.16b), for example at 2 km to 2.5 km depths.

The data fit of the first shot gather computed in the true V_P model and in the final JFWI followed by FWI V_P model is shown in Figure 3.17. The multi-scattered waves are generated (e.g. time = 4.4 s to 5.2 s at offset = 6 km), due to the injection of the intermediate wavenumbers into the final V_P model, which allows us to model these higher-order scattering effects from the top of the cap rock and the vertical edges of the gas layers. Except at zero offset, the match of the phase and amplitude is further improved comparing with Figure 3.15, especially for the multi-scattered waves and the late reflection at long offsets. In summary, this experiment manifests that JFWI has a great potentiality to be used as a robust tool for initial velocity model building for

conventional FWI.

The match at zero offset has been degraded (within the two-way traveltimes ranging from 1.8 s to 3.4 s) due to the offset weighting that has been applied to enhance the reflection data at long offsets. Moreover, this degradation might be caused by the fact that we have kept the smooth density model to its initial values, suggesting that the model space is not large enough to account for the amplitude effects at short offsets. In other words, these amplitudes cannot be matched by a velocity-only inversion procedure. Therefore, a multi-parameter FWI for V_P and ρ should be considered to improve the data fit.

3.5.3 JFWI in presence of multiples

In this section, we discuss the robustness of JFWI in the presence of surface multiples. The interest is that if JFWI allows the multiples to be present in the data, the preprocessing workflow could be simplified.

We still consider the synthetic Valhall model (Figures 3.7a and c) and generate a data set with surface-related multiples (Figure 3.18). This data set is processed by JFWI and IpWI by considering a free-surface boundary condition on the surface during the seismic modeling. Compared with the data set computed without free-surface multiples (Figure 3.7a), the diving waves have weaker amplitudes in the 1 km to 2 km offset range (using the same clip), and more complex reflection wavefields are recorded. This prompts us to reduce the scaling factor applied to the reflected waves through the operator W^T from 10 to 5, such that the contributions from the diving and reflected waves are re-balanced in the JFWI misfit function. Moreover, as free surface effects generate more multi scattering, we progressively feed the inversion with late-arriving reflections at a slower rate than in the former case performed without surface multiples. Here, we mute the reflection data after 3.3 s over the full offset range during first few cycles, then gradually restore their amplitude to their original level during later cycles (also weighted by 5). We apply the same time-offset window (Equation 3.12) to separate the data.

Figure 3.19(a) shows the JFWI result after 14 cycles. Although the final JFWI V_P model shows the long wavelengths of the true V_P model as in Figure 3.9(e), we have witnessed some degradations of the quality from this velocity result:

1. The thin reflector at the water bottom between $x = 3$ km and 4 km is an artifact coming from the very early reflections in the diving-wave time window. Because we have applied a simple linear time-offset window, the migration isochrones generated from the reflections left in the diving-wave time window are not totally canceled during the summation of G_1 and G_2 . When multiples are present, these reflections may be enhanced, making the migration isochrones apparent in the gradient. This highlights that a successful application of JFWI heavily relies on the accuracy of data separation.

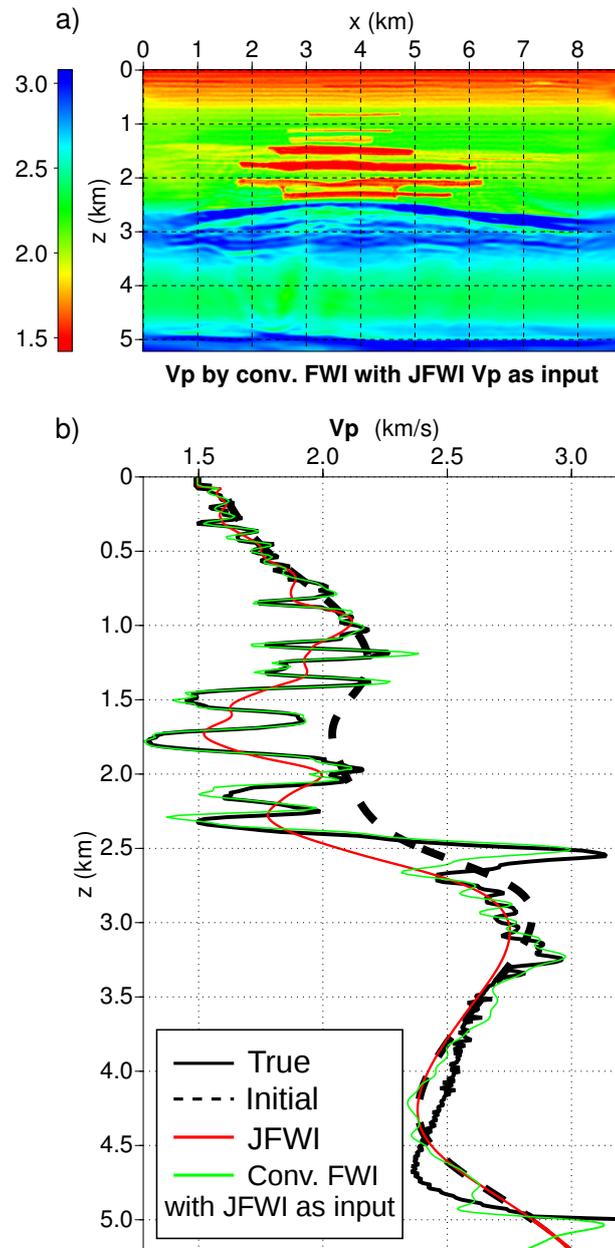


Figure 3.16: (a) Broadband reconstruction of the velocity model by conventional FWI using JFWI V_P model as the starting model. The V_P - ρ parameterization is used for inversion keeping the density model fixed). (b) Comparison between velocity profiles ($x=3.75$ km) extracted from the true model (solid black line), the initial V_P model (dashed line), the JFWI V_P model (red line) and the broadband V_P model shown in (a) (green line). Leakage of ρ reflectivity is expected (e.g. at 2 km to 2.5 km depths). Online animation for all intermediate updates: <https://drive.google.com/open?id=0Bx0JCm2KZyueZWZDXzN6VFQ5W1E>

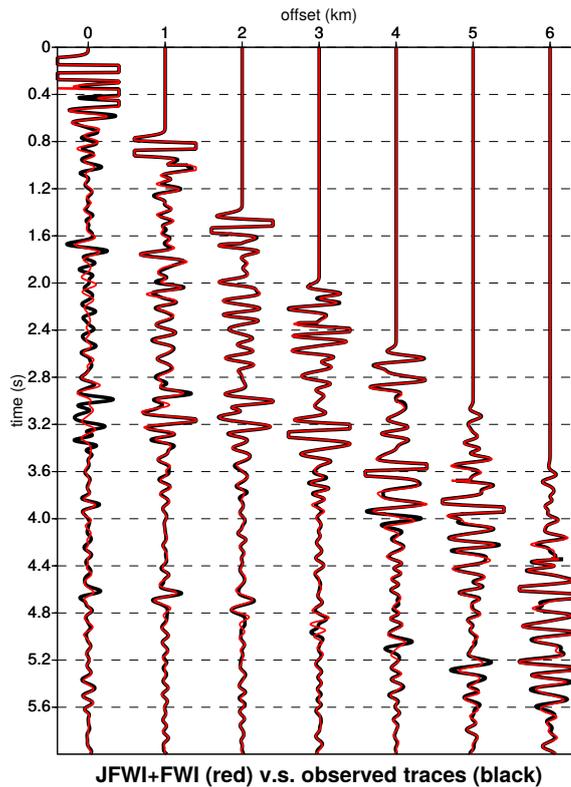


Figure 3.17: Data fit of the first shot gather computed in the true V_P model (black) and in the final JFWI+FWI V_P model (red), which is comparable with the one obtained with the JFWI V_P model (Figure 3.15). Here, the fit at long offsets of multi-scattered waves has been nicely improved at the expense of the fit at short offsets. See text for interpretation.

2. Above 2 km depth, the resolution seems to be higher in Figure 3.19(a) than in Figure 3.9(e), even though we have applied the same Gaussian smoothing regularization to the gradient (vertical and horizontal correlation lengths equal to twice of the dominant wavelength). This improved resolution might result from the improved subsurface coverage provided by the surface-related multiples. This statement deserves however further investigations.
3. In contrast, the low velocity zone at 2.3 km depth is reconstructed less accurately in Figure 3.19(a) than in Figure 3.9(e). We consider that imaging at this depth, covered by only reflection kernels, is a difficult task especially when multiples are present.

As was in Section 3.5.2.4, we launch the conventional FWI starting from the JFWI V_P macromodel (Figure 3.19a) using the V_P - ρ parameterization. The two-layer density model that was used as the initial model for JFWI is also used here as the background

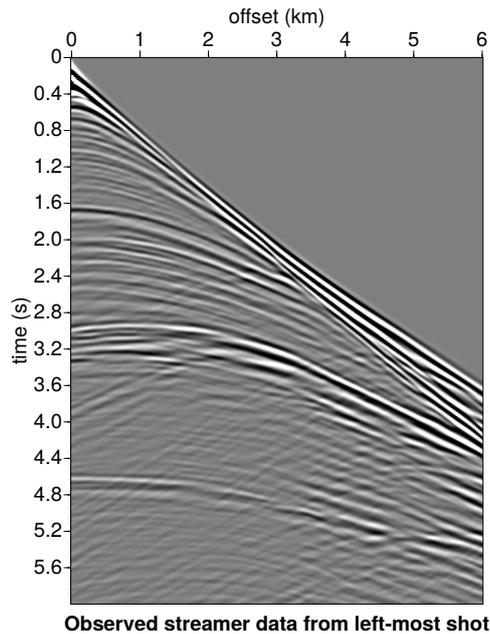


Figure 3.18: First short gather of synthetic Valhall case study with free surface multiples.

model, and is kept fixed during this FWI implementation. The result is shown in Figure 3.19(b). The gas zone is properly reconstructed, with some noise due to the multiples. The image of the cap rock, however, is not as accurate as the one built from multiple-free data. We believe that this degradation results from the surface multiples and the multi-scattered waves generated by the hard interfaces and the edge of the gas layers. Remembering the former application without multiples, these multi-scattered waves had already made the velocity model building quite non-linear, which prompted us to design a time-windowing approach to mitigate this non-linearity. Here, as expected, the presence of multiples has significantly increased the non-linearity, which forces us to apply a more prudent time-windowing approach. With such an effort, the degradation of the resolution is limited.

Regardless the difficulties raised by these two wave modes, we still obtain a reasonable velocity model by FWI using the JFWI model as the starting model. The log extracted at $x = 3.75$ km is shown in Figure 3.19(c). Without considering possible cross-talk effects between parameters, the final V_P values (green curve) fit quite well the true V_P values (black curve) above 3.2 km depth, and is much comparable with the one inferred from multiple-free data (Figure 3.16).

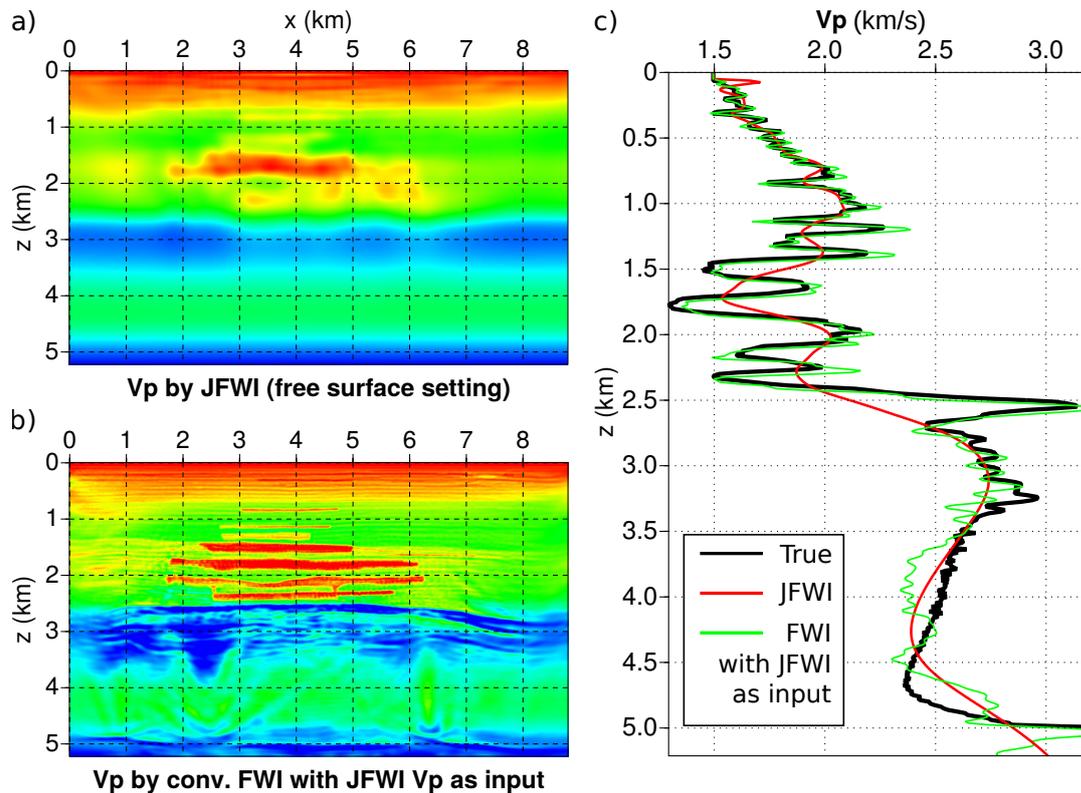


Figure 3.19: Synthetic Valhall case study with free surface multiples. (a) Long-wavelength V_P model by JFWI. (b) Broadband reconstruction of the V_P model (conventional FWI implementation starting from (a), V_P - ρ parameterization and density fixed). (c) Logs of V_P models at $x = 3.75$ km for (a) and (b). JFWI suffers from the high-order scattering effects from the surface-related multiples and the multi-scattered waves, therefore the V_P results are worse than the previous results (Figure 3.9e and Figure 3.16a). However, from the log we see that the gas layers are still well imaged. See text for details.

3.6 Conclusions and perspectives

Most applications of the conventional FWI are driven by diving waves and post-critical reflections. However, imaging the long-to-intermediate wavelengths in deep regions remains challenging if the transmitted waves do not penetrate these regions due to insufficient offsets. In contrast, RWI succeeds to some extent in imaging the long-to-intermediate wavelengths of the deep regions focusing on the two transmission wavepaths followed by the reflected waves from the surface to the reflector positions. This kind of approaches relies on a scale separation between a smooth velocity model and a known reflectivity, that allows one to suppress the contribution of the unwanted (high-wavenumber) migration isochrones during the velocity model building. The limitation of RWI is that diving waves and post-critical reflections are not used, although they carry the essen-

tial long-wavelength information of the shallow subsurface. We have presented in this paper an extension of RWI, namely Joint FWI, the aim of which is to account for the contribution of the diving waves and post-critical reflections during the velocity model building. The relevance of the method is demonstrated with the synthetic Valhall case study, which has shown that how the improvement of the near surface imaging provided by the diving waves translates into an improved imaging of the deep targets performed by RWI. Although modern wide-azimuth long-offset seismic acquisition may still not allow to record the diving waves with a sufficient penetration depth to sample the deepest targeted structures, diving waves and post-critical reflections can have an increasing contribution in the seismic wavefield. In this study, we have proposed a seismic workflow which makes an optimal use of the information carried by all kinds of waves to build a reliable velocity macromodel. Our approach requires however the explicit separation of the diving waves (or post-critical reflections) and the pre-critical reflections.

The conventional V_P - ρ parameterization of the subsurface favors a broadband reconstruction of the V_P parameter, but the mitigation of the cross-talks between V_P and ρ is challenging from reflected waves. In contrast, the V_P - I_P parameterization leads to a more natural uncoupling between the two parameter classes, which is consistent with the scale separation between the velocity model and the reflectivity underlying RWI and JFWI. It becomes therefore natural to combine within an iterative workflow the velocity model building performed by JFWI and the impedance imaging performed by conventional FWI of reflected waves (IpWI). As the impedance model needs to be updated according to the velocity updates, JFWI and IpWI are performed in an alternate way leading to the cycle workflow.

For more efficient implementations of the cycle workflow, one possibility is to reduce the iteration number of the I_P inversion. Although the amplitudes of the imaged reflectivity will be incompletely estimated, the kinematics of the data can be retrieved to reconstruct a reliable V_P background model. As soon as the kinematic attributes of the data are matched, amplitudes can be further used to build more accurate reflectivity images.

Having respectively built the low-part and the high-part of the V_P and I_P spectra, it is natural to wonder whether imaging a broad spectrum of the subsurface is possible. Starting from the velocity macromodel built by JFWI, a broadband velocity model can be tentatively imaged by conventional FWI using the V_P - ρ parameterization. The quality of the reconstructed velocity model shows that the low-to-intermediate wavenumber components of the JFWI model are accurate enough to successfully image the subsurface from reflected waves by FWI. In this experiment, only the V_P parameter is updated keeping the density fixed to its original value. Therefore, this V_P model is prone to contain imprints from V_P - ρ cross-talks. These cross-talks probably manifest by the overestimation and/or underestimation of the reconstructed velocities. To tackle this kind of cross-talks, multi-parameter inversions which involves the Hessian operator should be considered. Depending on the computational facilities, l -BFGS (Nocedal, 1980) or Truncated Newton (Métivier *et al.*, 2013) methods could be considered to introduce the Hessian operator.

On the other hand, imaging low-wavenumber part of the I_P spectrum could be more difficult, as neither V_P - ρ nor V_P - I_P parameterization can provide low-wavenumber sensitivity to the I_P parameter (cross-talks arise in the transmission regimes for the ρ - I_P parameterization). In this case, we have to rely on empirical parametric relations such as the Gardner relation. From this kind of relation, we could rebuild the initial I_P in each cycle to translate low wavenumbers from V_P to I_P , or set a trust region to bound the searching area of the I_P inversion.

We also look forward to the applications of JFWI to real data sets such as the real Valhall one. The cycle-skipping issue would appear and could hinder the local search method to reach the global minimum. Therefore, more robust misfit definitions would be required. Besides, how the cycle inversion scheme is affected by the elastic information of the data should be studied. For example, we foresee that the misinterpretation of elastic wavefields as acoustic ones would generate reflection residuals due to inaccurate amplitudes. The inversion of these residuals by JFWI would generate some artificial velocity perturbations even if the initial V_P model is correct. Accounting for density to absorb the elastic effects is an option, although this inevitably questions the use of the V_P - I_P parameterization during JFWI (Borisov *et al.*, 2014; Plessix *et al.*, 2014; Plessix et Pérez Solano, 2015). Another possible strategy is to correct the acoustic wavefields for the elastic effects by using artificial source terms (Chapman *et al.*, 2014; Hobro *et al.*, 2014). A last-but-not-least possibility could be to rely on the kinematic-associated misfit functions such as the ones based on cross-correlation (van Leeuwen et Mulder, 2010), deconvolution (Luo et Sava, 2011; Warner et Guasch, 2014), instantaneous phase (Fichtner *et al.*, 2008; Bözdog *et al.*, 2011), or dynamic warping (Hale, 2013). Extensions to 3D geometry could be performed, but the repetition of the I_P inversion inside the cycle workflow would be a computational obstacle. One possible solution is to build the reflectivity in the pseudo-time domain instead of the depth domain, by which reflection phases are always matched in short offsets (Plessix, 2013; Wang *et al.*, 2015).

Acknowledgements

This study was partially funded by the SEISCOPE consortium (<http://seiscope2.osug.fr>), sponsored by BP, CGG, CHEVRON, EXXON-MOBIL, JGI, PETROBRAS, SAUDI ARAMCO, SCHLUMBERGER, SHELL, SINOPEC, STATOIL, TOTAL and WOODSIDE. This study was granted access to the HPC resources of the Froggy platform of the CIMENT infrastructure (<https://ciment.ujf-grenoble.fr>), which is supported by the Rhône-Alpes region (GRANT CPER07_13 CIRA), the OSUG@2020 labex (reference ANR10 LABX56) and the Equip@Meso project (reference ANR-10-EQPX-29-01) of the program Investissements d’Avenir supervised by the Agence Nationale pour la Recherche, and the HPC resources of CINES/IDRIS under the allocation 046091 made by GENCI. Authors appreciate fruitful discussions with Ludovic Métivier (LJK-CNRS, Université Grenoble Alpes) and Yang Li (Harbin Institute of Technology). We thank the editor R.-É. Plessix, Partha Routh and an anonymous reviewer for their constructive comments which helped us improve the manuscript.

Chapter 4

Real Data Application

When in doubt, smooth.

— Sir Harold Jeffreys (Quoted by Moritz, 1980 and [Tarantola, 2005](#))

Avec les rêves aussi on peut faire des confitures. Il suffit d'ajouter des fruits et du sucre.

— Stanislaw Jerzy Lec, *Nouvelles pensées échevelées*

Encouraged by the promising result of JFWI in the synthetic Valhall case, I move forward to real case studies to assess its promises and pitfalls, particularly in initial velocity model building for classical FWI. Can JFWI recover sufficient low wavenumbers such that the subsurface model can be located inside the attraction valley of the global minimum of FWI? What are the main obstacles for JFWI to build such desirable model?

A 2D section of OBC data set cross-cutting a gas cloud is adopted to answer these questions. The first obstacle is the inseparability of the postcritical branch of the reflected waves from the refracted and direct/diving waves at far offsets (hence, they are uniformly called “early arrivals” in this chapter). However, due to their weak amplitudes the high-wavenumber updates from the reflected waves can be mitigated by using the V_P – I_P parameterization and model regularization. Therefore, I shall directly augment the diving-wave term of JFWI misfit function with all early arrivals, and I shall only separate between early arrivals and precritical reflected waves which can be implemented simply by time windowing.

Previous study has shown reasonable velocity models by 3D monoparameter FWI in the frequency domain ([Operto *et al.*, 2015](#)). They believe that the elastic effect can be ignored due to mild velocity contrasts in the field. However, their data misfits at far offsets and degraded resolution at high frequencies are attributed to the underestimation of attenuation, an issue that was not considered in the previous synthetic study of JFWI. In this chapter, I shall also address this via the implementation of classical FWI using different attenuation models in order to well manage the data for JFWI applications.

Joint Full Waveform Inversion of Early Arrivals and Short-spread Reflections: a 2D Ocean-Bottom-Cable Study Including Gas Cloud EffectsZHOU, W., BROSSIER, R., OPERTO, S., VIRIEUX, J. & YANG, P.
*To be submitted.***Abstract**

Joint full waveform inversion (JFWI) combines reflection (RWI) and early-arrival (EWI) waveform inversions to build a large-scale velocity model of the subsurface. The misfit function of JFWI requires an explicit separation between the short-spread reflections and early arrivals, the feasibility of which is illustrated with a real long-offset data set. JFWI is alternated with a waveform inversion/migration of short-spread reflections to provide a short-scale impedance model that is used as an input to build the sensitivity kernel of RWI along the two-way reflection paths. The velocity macromodel built by JFWI can be taken as the initial model for classical FWI to enrich the high wavenumber content of the subsurface model. We present an application of this workflow to a real 2D OBC profile cross-cutting a gas cloud in the North Sea to review its main promises and pitfalls. Several initial models and offset-driven strategies are assessed with the aim to manage the cycle-skipping issue while producing subsurface models of sufficient vertical and horizontal resolution. Our workflow produces an acceptable FWI velocity model that fits the main early arrivals and reflections when JFWI starts from a smoothed version of an existing traveltimes tomographic model. We expect that considering the 3D geometry would relax the strict planar propagation assumed in 2D. The improved scattering angle illustration would also reduce the requirement on the quality of the initial model.

4.1 Introduction

Full waveform inversion (FWI) is a promising tool to image the subsurface (Virieux et Operto, 2009). Successful applications often require a good initial guess of the subsurface: the proposed velocity model must match the recorded phases within a half cycle; otherwise the inversion is easily trapped by a local minimum associated to an unphysical model. Ray-based tomographic approaches, such as first-arrival traveltimes tomography (Sei et Symes, 1994; Zelt et Barton, 1998; Leung et Qian, 2006; Taillandier et al., 2009), reflection traveltimes tomography (Bishop et al., 1985; Sword, 1987; Whiting, 1998; Farra et Madariaga, 1988), stereotomography (Billette et Lambaré, 1998; Lambaré, 2008; Prioux et al., 2013c) are conventionally applied for the purpose of initial model building for FWI. However, the underlying high-frequency approximation may limit the applicability of these approaches to complex media. Alternatively, wave equation-based inversion can overcome such limitation. Waveform inversion using early arrivals (EWI, Shipp et Singh, 2002; Sheng et al., 2006; Sirgue, 2006; Shen, 2014; Wang et al., 2015), such as diving waves and super-critical reflections is usually performed to image the large scales of the subsurface before short ones. Sufficiently wide apertures and/or long

offsets are required to reach the deep target, which in turn intensify the risk of cycle skipping due to long propagation distances. To mitigate this difficulty, we can rely on low frequency data to reduce the nonlinearity (Bunks *et al.*, 1995; Sirgue *et Pratt*, 2004). Formulations in the Laplace domain can also boost the low frequencies and early arrivals in the data (Shin *et Cha*, 2008, 2009; Pyun *et al.*, 2008, 2011). However, the problem is that current acquisition techniques still cannot provide very low frequencies (typically < 2 Hz); the reliability of the results can be questioned due to the low S/N of the data. On the other hand, misfit functions that present less local minima than the L2-norm misfit are also considered. Numerous investigations are promoted in this direction, such as those based on correlation (Luo *et Schuster*, 1991; van Leeuwen *et Mulder*, 2010), deconvolution (Luo *et Sava*, 2011; Warner *et Guasch*, 2014), envelope (Bözdag *et al.*, 2011; Luo *et Wu*, 2015), unwrapped phase (Alkhalifah *et Choi*, 2012), dynamic warping (Ma *et Hale* (2013) and optimal transport (Métivier *et al.*, 2015).

Short-spread reflections have the sensitivity to deep targets without the need of very far offsets, and the redundancy of the reflection data enables image domain approaches (Diaz *et al.*, 2013; Liu *et al.*, 2013; Allemand *et Lambaré*, 2014) to update the velocity model. By migrating the reflection energy to common-depth-point images, these approaches focus on the flatness in the image gathers (Symes *et Carazzone*, 1991; Sava *et Biondi*, 2004), or attempt to minimize the energy left in the non-physical dimensions added to the model space (Sava *et Fomel*, 2006; Yang *et Sava*, 2011; Almomin *et Biondi*, 2012; Biondi *et Almomin*, 2012; Sun *et Symes*, 2012; Lameloise *et al.*, 2015). The main issue is related to their high computational cost due to the migration step and image evaluations in extended domains performed during each velocity update.

In general, data-domain inversion requires less computational resources than image-domain inversion. Inspired by migration-based traveltime inversion (Chavent *et al.*, 1994; Chavent, 1996; Clément *et al.*, 2001), reflection waveform inversion (RWI) has been proposed to build the velocity macromodel by restricting the sensitivity kernel of FWI along the two-way transmission paths of short-spread reflections (Zhou *et al.*, 2012; Xu *et al.*, 2012). It is alternated with the migration or impedance waveform inversion that produces a short-scale reflectivity model to be used as a prior information to build the RWI sensitivity kernel. A key property of RWI is the significant reduction of cycle skipping at short offsets. The need for regenerating the impedance model can be relaxed by building the reflectivity in the pseudo-time domain (Plessix, 2013; Brossier *et al.*, 2015; Wang *et al.*, 2015).

Due to the different wavepaths followed by early arrivals and short-spread reflections, EWI and RWI tend to preferentially sample the vertical and horizontal components of the wavenumbers, respectively. This prompts us to combine EWI and RWI into a joint inversion workflow (JFWI) to enrich the wavenumber content of the velocity macromodel and to improve the reconstruction at shallow depths (Zhou *et al.*, 2015). The JFWI misfit function relies on an explicit separation between early arrivals and short-spread reflections, which might require a careful data preprocessing. Nevertheless, as argued in Wang (2015), sufficiently broadband data rather than low frequencies should be considered in the early stage of RWI to extract the low-wavenumber information from

the normal moveout of precritical reflections. The focused wavelets, resulting from this broadband signal, facilitate the separation of the data by a simple time windowing as will be illustrated in this study. In [Zhou *et al.* \(2015\)](#), we also show with a synthetic experiment that JFWI is immune to cycle skipping as long as a time-driven scheme ([Kolb *et al.*, 1986](#); [Bunks *et al.*, 1995](#)) is used to image shallow targets before deeper ones.

This study aims to further assess the promises and pitfalls of JFWI with an application to a 2D line of the OBC data set that is collected across a gas cloud in the North Sea. This field has been studied by [Prioux *et al.* \(2011, 2013a\)](#) and [Operto *et al.* \(2015\)](#), who have reported the attenuation effect of the wavefield in the whole available offset range due to the dissipative medium in the near surface and in the gas. As it can affect both phases and amplitudes of the wavefield, considering attenuation in waveform inversion gives more reliable results. This is recently confirmed by the synthetic study of [Kurzman *et al.* \(2013\)](#), who have shown that an acoustic inversion scheme based on viscoacoustic modeling can significantly improve the velocity models, requiring only an appropriate representation of attenuation in the near surface. In view of the numerical simulation of the attenuated wavefield, the frequency-domain implementation is cheaper than in the time domain ([Carcione *et al.*, 1988](#); [Robertsson *et al.*, 1994](#); [Bai *et al.*, 2014](#); [Dutta *et al.*, 2014](#); [Plessix, 2016](#); [Yang *et al.*, 2016](#), among others). However, JFWI is implemented in the time domain for the purpose of explicit data separation.

For the sake of computational efficiency, we shall perform time-domain modeling without considering attenuation, and a lower frequency range will be looked for such that disregarding attenuation would not significantly affect the inversion. With this assurance, we shall test different source-receiver offset-driven strategies and initial models of increasing accuracies to assess the sensitivity of the inversion to cycle skipping. The velocity macromodel built by JFWI will be further assessed in terms of kinematic accuracy and spatial resolution in order to be used as an initial model for FWI. Starting from a smooth model with lateral variations, we succeed in building a velocity model from JFWI followed by FWI that compares reasonably well with an existing 3D FWI model from a former study ([Operto *et al.*, 2015](#)). On the contrary, when starting from a crude 1D model, JFWI fails to update enough the low-wavenumber content of the gas cloud, leading to mispositioned reflectors after FWI. However, an equally good data fit is achieved after inversions respectively starting from the smooth and 1D initial models, which highlights the ill-posedness of the inverse problem due to the insufficient azimuthal and offset coverage. We hope this issue would be overcome by extension of JFWI to 3D geometries.

4.2 Methodology

Classical FWI tries to recover the subsurface model m by minimizing the following L2-norm misfit function

$$C_{FWI}(m) = \frac{1}{2} \|W(d - Ru(m))\|^2, \quad (4.1)$$

where d denotes the recorded data and $u(m)$ denotes the modeled wavefield computed in the model m (same misfit function as Equation (3.2)). This wavefield is then sampled at receiver positions by the operator R such that the residual can be evaluated. All waves, including early arrivals and reflections are included in d and Ru . To highlight or mute certain wave modes, the weighting operator denoted by W is introduced. In EWI, for instance, W has nonzero values for early arrivals whereas it equals zero for short-spread reflections.

Alternatively, JFWI assumes that, on the one hand, the early arrivals d^e can be a priori separated from the short-spread reflected data d^r , and, on the other hand, the short-scale variations of the subsurface δm can be separated from the background model m_0 (i.e. scale separation). In this way, the misfit function of JFWI is defined as follows (Zhou *et al.*, 2015)

$$C_{JFWI}(m_0)|_{\delta m} = \frac{1}{2} \|W^e (d^e - Ru_0(m_0))\|^2 + \frac{1}{2} \|W^r (d^r - R\delta u(m_0, \delta m))\|^2 \quad (4.2)$$

(same misfit function as Equation (3.6)). Note that the model δm is fixed during inversion and only the model m_0 is updated iteratively. At each iteration, the wavefield u_0 is computed in the smooth velocity model m_0 whereas the perturbation wavefield δu is computed in the whole model $m_0 + \delta m$. The two terms of the misfit function, associated with early arrivals and reflections, preferentially contribute to updating the vertical and horizontal components of the low wavenumber vectors, respectively, and hence are complementary. Thanks to the explicit data separation, the associated JFWI gradient (i.e. Equation (3.7)) excludes the high-wavenumber migration isochrones that were introduced into the FWI gradient. The remaining high-wavenumber update in the JFWI gradient, coming from higher-order scattering, can be mitigated by stacking the gradient from each source-receiver couple due to destructive interference. Furthermore, based on the analysis of diffraction patterns with certain parametrizations, we propose to use $V_P - I_P$ to effectively force the scale separation, in which V_P is chosen as m_0 and I_P is chosen as δm in Equation (4.2) (Jannane *et al.*, 1989; Operto *et al.*, 2013).

The purpose of introducing scale separation and thus the short-scale impedance model δm is to enhance the low-wavenumber update coming from reflected waves, and make the velocity inversion immune to cycle skipping at least in short offsets (see discussion in Chapter 2). For this reason, the impedance model cannot be given arbitrarily. For example, the depth of one reflector, Z , in the impedance model should be equal to

$$\int_0^Z V_P(z) t_0(z) dz \quad (4.3)$$

where t_0 is the one-way vertical traveltime of the reflected wave associated to that reflector, and V_P is the background velocity along the vertical reflection wavepath. Therefore, to build an effective impedance model, we proposed impedance waveform inversion using short offset reflection data (referred to as IpWI for short). We use the same regular mesh to discretize the two parameters (velocity and impedance) for convenience, although blocky models could also be used especially for impedance. I refer the reader

to Section 3.3 or the publication Zhou *et al.* (2015) for a complete explanation of the approach.

The overall workflow is listed in Algorithm 4.1. Note that in Step 2, $V_P^{(k-1)}$ and $I_P^{(0)}$ are used as the initial models whereas in Step 3, $V_P^{(k-1)}$ and $I_P^{(k-1)}$ are used as the initial models. In each cycle, the computational cost of IpWI is equivalent to FWI. The temporal complexity of JFWI is twice of FWI but their memory requirement is the same.

Algorithm 4.1 Cycle workflow of V_P - I_P inversion

Initial condition: smooth models $V_P^{(0)}$ and $I_P^{(0)}$, $k = 1$.

WHILE

1. Source wavelet $s(t)$ estimation by the approach proposed in (Pratt, 1999):

$$s(t) = \mathcal{F}^{-1} \left(\frac{\mathcal{F}(R\hat{u})\mathcal{F}(d)^*}{\mathcal{F}(R\hat{u})\mathcal{F}(R\hat{u})^* + \varepsilon} \right), \quad (4.4)$$

where \mathcal{F} and \mathcal{F}^{-1} denote the Fourier transform and its inverse, respectively. $\hat{u} = \hat{u}(t)$ denotes the modeled data computed in $V_P^{(k-1)}$ and $I_P^{(k-1)}$ using the delta function as source wavelet. The complex-valued signals $\mathcal{F}(d)$ and $\mathcal{F}(R\hat{u})$ are respectively conjugated (denoted by symbol $*$) to evaluate the cross- and auto-correlations in the Fourier domain with a small number ε for numerical stability;

2. IpWI to build reflectivity image: $I_P^{(k)} = I_P^{(0)} + \Delta I_P$;
3. JFWI to update the macromodel: $V_P^{(k)} = V_P^{(k-1)} + \Delta V_P$;
4. $k = k + 1$

UNTIL convergence in Step 3.

4.3 Application

The area under exploration is the Valhall oilfield in the North Sea (Barkved *et al.*, 2010a,b; Sirgue *et al.*, 2010). The data set is acquired at the sea bottom (70 m depth) with ocean bottom cables. For this 2D study we use the inline gathers collected over the gas cloud from the 3D acquisition system (Cable 13, Operto *et al.*, 2015, their figure 2a). The maximal offset range is 13.44 km (Figure 4.1a, white line and dots). The two white stars represents two data gathers that will be discussed below. A velocity field has been built by 3D reflection tomography, and the corresponding 2D section (Figure 4.1a, courtesy of BP) shows the gas area at 1.5 km depth (depicted in blue) embedded in the overburden (depicted in yellow). The presence of gas forms a local low-velocity region where strong attenuation of wavefields has been reported (Prioux *et al.*, 2011, 2013a; Operto *et al.*,

2015), making the seismic imaging methods challenging at the reservoir depths (red). To assess the quality of this tomographic model, we generate the associated migrated image (Figure 4.1b) through source estimation (Step 1 of Algorithm 4.1) followed by waveform inversion of near-offset reflected waves (Step 2) filtered in a narrow frequency band (3–5.6 Hz). The detailed preprocessing flow will be listed in Section 4.3.2 and the reason why using low frequencies will be explained in Section 4.3.1. A lateral discontinuity is observed (ellipse) for the caprock reflector, which is attributed to the insufficiently low velocities in the gas cloud provided from the velocity model (Figure 4.1a). In the 3D study performed by *Operto et al. (2015)*, FWI was implemented in the frequency domain, where attenuation was considered by using complex-valued velocities during the modeling process. They used a maximum frequency of 10 Hz and a homogeneous quality factor Q of 200 below the sea bed to account for attenuation. The associated 2D section is shown in (Figure 4.1c). Starting from the tomographic model, FWI has increased the velocities in the overburden while decreased the velocities in the gas cloud (below 2 km), therefore leads to a relatively flat caprock reflector with good lateral continuity (Figure 4.1d). The reference reflectivity model is extracted from the velocity model (c) by filtering out the low wavenumber components using Gaussian smoothing. We do not perform migration as in Figure 4.1b because we favor high-resolution images for comparison (remind the low-frequency content used in migration). In the following, we shall use the two models (c,d) to assess the models developed in this study.

The anisotropic parameters are also reconstructed by reflection tomography (Figures 4.2a,c). Mild anisotropy is present in the soft sediments above the reservoir. We also smooth these two models below the sea bed using the Gaussian function, such that they do not generate apparent scattered fields (Figures 4.2b,d) in our application. In this study, we do not consider anisotropic inversion.

For the sake of data separation, we shall implement JFWI in the time domain. The modeling code is parallelized over shot gathers, but the original OBC acquisition provides more airguns than hydrophones. Therefore, we apply source-receiver reciprocity to the data meaning that we treat shots as receivers and vice versa. Therefore, we shall refer to a receiver as a reciprocal shot and the word “reciprocal” will not be mentioned hereafter. By filtering out high-frequency noise, two shot gathers are shown in Figure 4.3. The shots are positioned far away from the gas cloud (a) and over the gas cloud (b), respectively (stars in Figure 4.1a representing the two source positions). Main body-wave arrivals are indicated by arrows. At short offsets, the direct waves (yellow arrows) propagate with an apparent velocity of around 2 km/s followed by a series of reflections (blue arrows) generated by the sediment layers (Figure 4.1c). Their constructive interference results in a higher amplitude observed at 2.5 s. The strong reflection (red arrows) comes from the sediment-caprock interface, which is followed by multiples and deep reflections after around 3.2 s. Due to the low-velocity gas cloud, the right-hand-side branches of the reflected waves in Figure 4.3b travel slower than their left-hand-side counterparts, giving non hyperbolic shapes unlike the gather in Figure 4.3a (blue and red arrows) whose source is far from the gas. The green arrows indicate the Stoneley surface waves that are to be filtered out during the preprocessing stage (discussed later). At far offsets, the

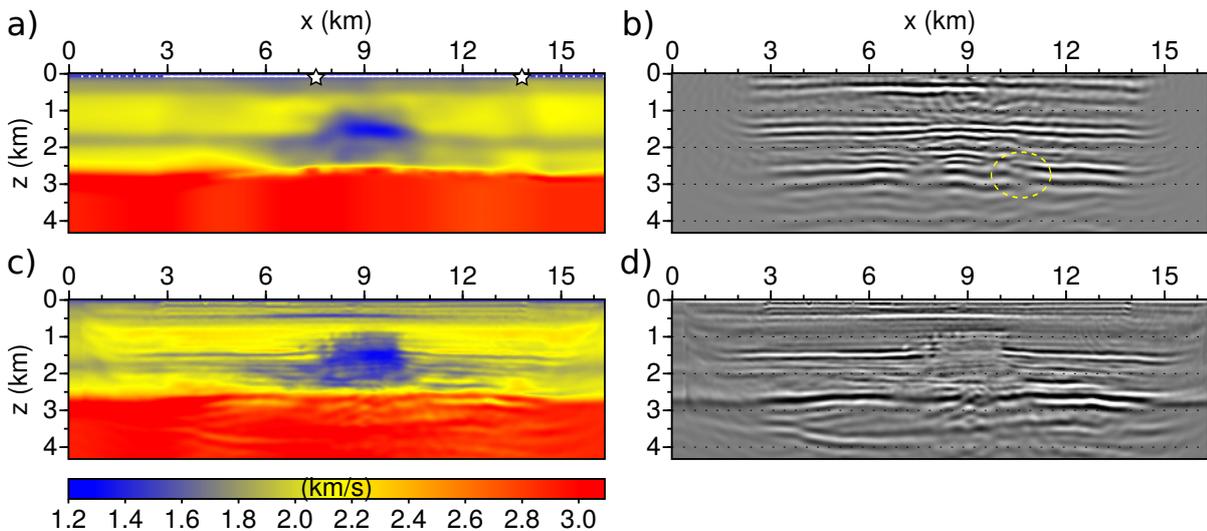


Figure 4.1: Prior knowledge of the subsurface under investigation. Associated 2D section of reflection tomographic velocity model (a) (courtesy of BP). The white line represents the ocean bottom cable ($x = 2.88$ to 13.81 km, $z = 0.07$ km) and the white line and the dots denote the shot positions ($x = 0.37$ to 16.32 km, $z = 0.07$ km). We apply source-receiver reciprocity due to the parallelism of our time modeling code. Two reciprocal shots are indicated by the starts ($x = 7.59, 13.81$ km), and the corresponding gathers are shown in Figures 4.3b,a, respectively. Low-velocity gas cloud (blue) is embed in the sediment layers (yellow) making strong lateral inhomogeneity overriding the hard interface at around 2.5 km depth formed by the sediment layers and the caprock (red). The migrated image (b) shows a discontinuous caprock reflector (ellipse) due to the overestimated velocities in the gas cloud (a). Associated 2D section of the 3D FWI velocity model (c), whose large scales are extracted by Gaussian smoothing and then are subtracted from the reference model to have the small-scale reflectivity model (d) for reference. Unlike in (b), the caprock reflector is relatively flat due to the velocity update in the gas cloud. The color bar of (c) also depicts the color scale of Figure 4.4, Figures 4.5a,b, Figures 4.8a,c, Figures 4.9a-c, Figures 4.10a,c,e and Figures 4.12a,c.

diving waves (dashed yellow arrows) have lower amplitudes than the direct waves due to long propagation distances and attenuation. The (post)critical branch (dashed red arrows) of the sediment-caprock reflection arrives later than the direct and diving waves but propagates with a faster apparent velocity. Due to the low-velocity gas cloud, the refraction branch (magenta arrow) is not recorded as the first arrival in the available offset range, but shows a tendency to surpasses the diving waves at further offsets.

Let us remind that, for the sake of computational efficiency we shall implement JFWI without taking attenuation into account, which causes inversion artifacts that will be shown later. Therefore, we must first address this issue and aim to find a suitable frequency range to mitigate such artifacts. This range will be found as follows. Starting from the tomographic models (Figure 4.1a and Figures 4.2b,d), we shall perform classical

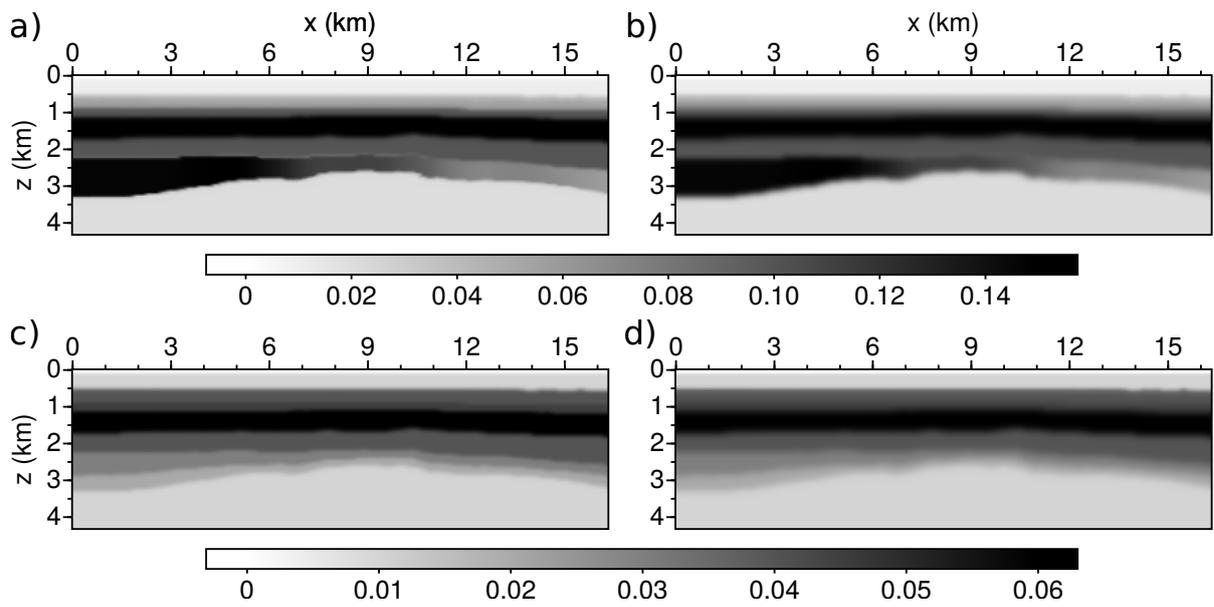


Figure 4.2: 2D sections of reflection tomographic Thomsen parameters ϵ (a) and δ (c) (courtesy of BP) and their smoothed versions for this study (b,d).

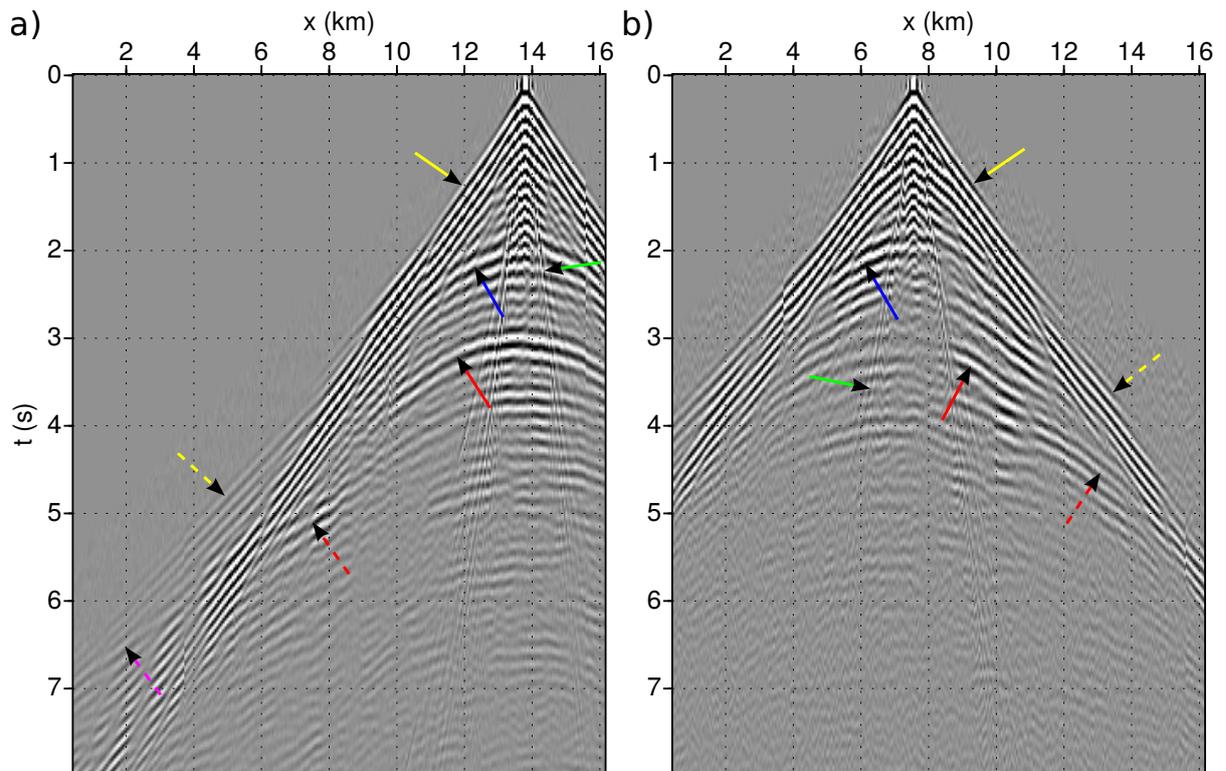


Figure 4.3: Real data after high-frequency noise filtering (high cutoff at 9 Hz). Main body waves are identified by colored arrows. See text for interpretations.

FWI in the frequency domain to assess the results by using different Q models, and then we shall perform time-domain FWI assuming an infinite Q (no attenuation) to assess the results with increasing frequency ranges. With these findings, the desirable frequency range will be determined for JFWI with disregard of attenuation, which is a reasonably wide range of intermediate-to-low frequencies with regard to the data spectrum.

4.3.1 Preliminary results related to attenuation

Seven frequency groups are chosen for the frequency-domain FWI test (Table 4.1). The inversion result from the previous group with lower frequencies is used as the initial model for the inversion of the next group with higher frequencies. In each group, three or four frequencies are simultaneously used. We do not use frequencies lower than 3.6 Hz due to low S/N. We keep the intermediate frequencies (5.5 and 5.9 Hz) in Groups 6 and 7 to mitigate potential cycle skipping at higher frequencies. Figure 4.4 shows the results from Groups 6 (a,c) and 7 (b,d) using $Q = 200$ (a,b) and $Q = 1000$ (c,d) below the seabed, respectively. We obtain an acceptable velocity model with $Q = 200$ whereas high-velocity artifacts are shown in the shallow part of the velocity model with $Q = 1000$ (black ellipses in c and arrows in d). Rather than true anomalies, we explain these anomalies as the trade-off from the inadequate Q parameter to the V_P parameter. Specifically, with a fixed Q model, the inversion has to increase V_P for a better match of the data. When the value of Q is sufficiently low as in (b), this trade-off phenomenon disappears. Note that these anomalies are less outstanding in the low-frequency model (c), implying that we could expect to not generate these anomalies at rather low frequencies even with a very large Q .

For the time-domain FWI test assuming an infinite Q , we apply the Butterworth filter to the data with increasing passbands (Table 4.2). These bands include rather low frequencies (<3.6 Hz) to avoid the ringing effect (i.e. Gibbs phenomenon) to take place, unlike in the frequency domain where low frequencies are discarded for computational efficiency (Table 4.1). The inversion result from the previous band is used as the initial model for the inversion of the next band which has a larger frequency range. The inversion results from Bands C and E are shown in Figures 4.5a,b, respectively. The high-velocity anomalies, presented in Figure 4.4c, are also created in the Band E result (enclosed by ellipses), issued from the trade-off caused by the infinite Q assumed in the modeling process. In contrast, they are not generated in the Band C model which considers lower frequencies. Consequently, the image of the underlying reflector (at 0.5 km depth) is reasonably flat in (a) as indicated by the line segment, whereas it is bended in (b) as indicated by the curve line due to the overestimation of the above velocity field.

To choose a desirable frequency range for our time-domain inversion, we compare the lateral profiles of these time-domain FWI results at the depth of the high-velocity anomaly (245 m) in Figure 4.5c. For reference, we also show the profiles of the 3D FWI and tomographic velocity models at the same depth. The plot illustrates the tendency of increasing velocities at $x = 14$ km position at higher frequencies, and reach a value of 2100 m/s by the Band E result. In contrast, the references models show a value

Table 4.1: Frequency groups sequentially used in the frequency-domain FWI test. Frequencies in each group are simultaneously used.

#	Frequency group (Hz)
1	3.6 3.9 4.3
2	4.1 4.4 4.8
3	4.6 4.9 5.3
4	5.0 5.4 5.8
5	5.5 5.9 6.3
6	5.5 5.9 6.4 7.0
7	5.5 5.9 8.0 9.0

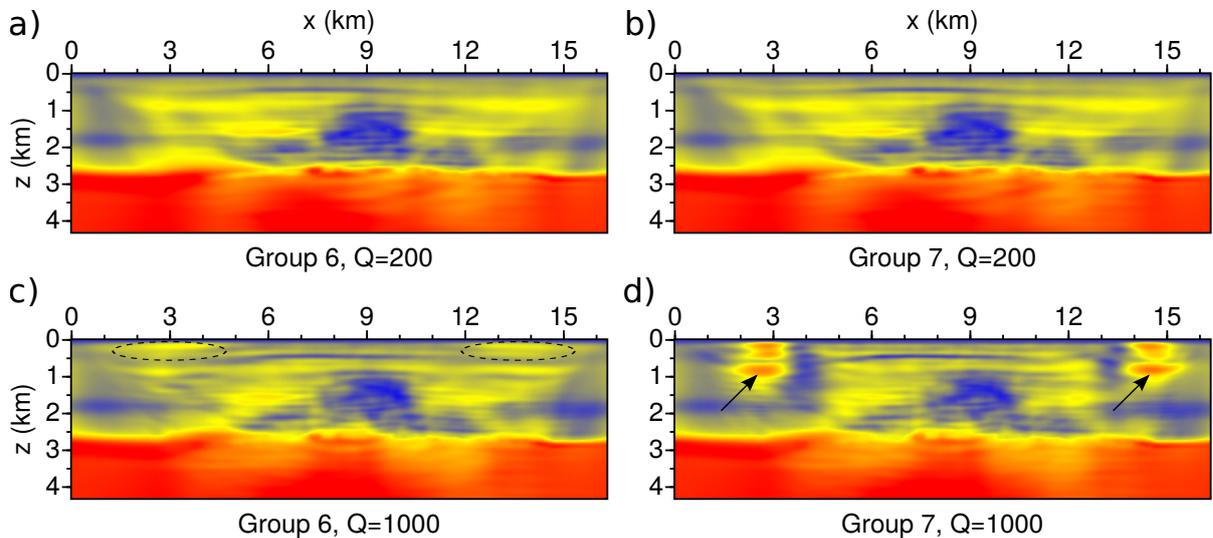


Figure 4.4: Preliminary frequency-domain FWI results with $Q = 200$ (a,b) and $Q = 1000$ (c,d), using frequency groups 6 (a,c) and 7 (b,d) that are listed in Table 4.1 (see labels below the results). High-velocity anomalies are generated in (c) and (d), respectively indicated by ellipses and arrows, highlighting the trade-off issue when an inadequate Q model is used without updates during the inversion. These anomalies are less outstanding in (c) because Group 6 does not include the high frequencies 8 and 9 Hz. We could expect to further mitigate these anomalies by using lower frequencies even with a very large Q .

lower than 1900 m/s. This discrepancy highlights the attenuation effect on velocity model reconstruction, and also the influence of the 3D-to-2D conversion used in this test. In this study, we only consider 2D applications of JFWI, and we need to balance two opposite behaviors. In order to exclude the attenuation footprints, we would need to use frequency content as low as possible. However, a broad frequency band improves the S/N of data (particularly at low frequencies) and also facilitates the explicit separation between early arrivals and reflections: if no sufficient high frequencies are considered, the two types of body waves mix with each other and hence obstructs a straightforward

Table 4.2: Characteristic frequencies defining five passbands for the Butterworth filter, sequentially applied to the data for the time-domain FWI test. Through the comparison made in Figure 4.5, we choose Band D for JFWI applications.

Band	Characteristic frequencies (Hz)
A	2.9 3 3.6 3.9
B	2.9 3 4.1 4.4
C	2.9 3 4.6 4.9
D	2.9 3 5.1 5.4
E	2.9 3 5.6 5.9

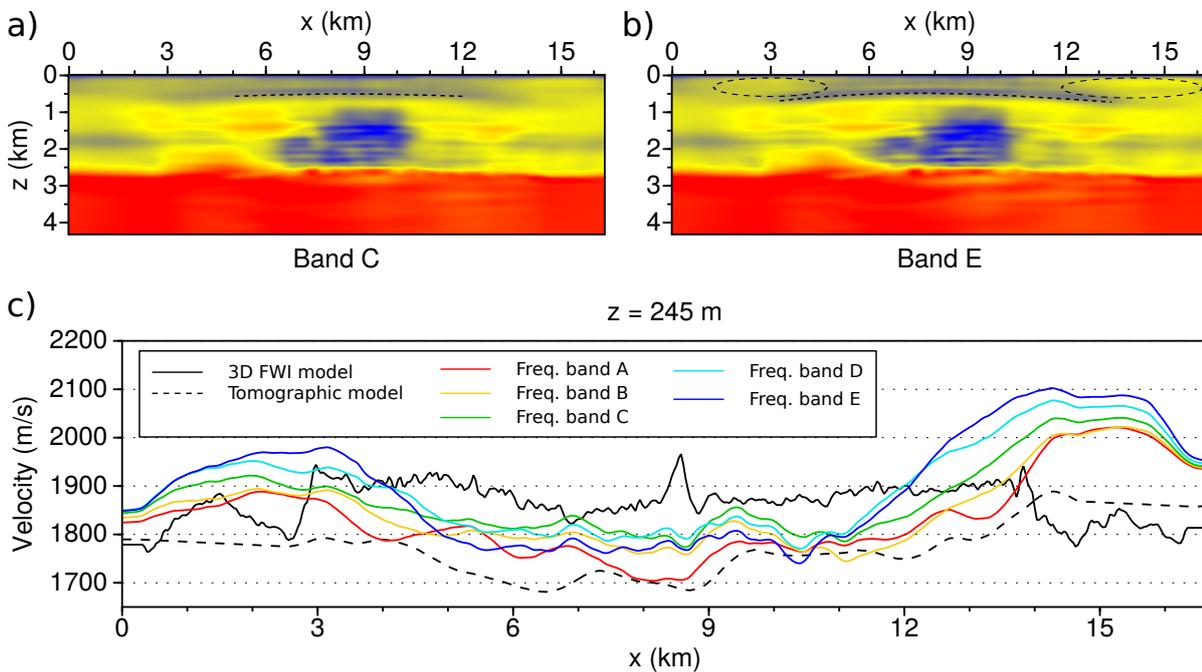


Figure 4.5: Preliminary time-domain FWI results using frequency bands C (a) and E (b) that are listed in Table 4.2 (see labels below the results). The modeling is based on acoustic wave equation therefore no attenuation is considered (assuming an infinite Q). High-velocity anomalies that are present in Figure 4.4c,d are also generated in (b), indicated by ellipses, which is a trade-off effect coming from disregard of attenuation. As a result, the reflector image below the anomalies is bent (dashed curve) that should be flat (line segment in b). We mitigate this trade-off issue by choosing a frequency range excluding high frequencies. Comparing the lateral profiles of all intermediate results from the five frequency bands and the referencing profiles of the 3D FWI and tomographic model (c), we pick Band D (Table 4.2) for our JFWI applications which includes sufficient low-to-intermediate frequencies to handle the low frequency noise as well as for the explicit data separation between early arrivals and reflections (Figure 4.6).

separation (Wang, 2015). Considering these two effects, we choose to limit the frequency range below 5.6 Hz (critical frequency) and choose Band D for our JFWI applications.

4.3.2 Inversion setup

We design the data preprocessing workflow as follows:

1. Normalize shot gathers with respect to the maximum absolute value to balance the contribution from each gather;
2. Remove Stoneley waves by f - k filtering;
3. Amplitude gain by a factor of $(t - h/1480)^{0.7}$ where t denotes the time (in seconds) and h the offset (in meters). The reduced velocity 1480 m/s is used to approximately follow the first breaks through offsets in the data. Empirically, we choose 0.7 as the power to roughly convert the amplitude from 3D to 2D geometric spreading and to compensate the energy absorption due to attenuation. However, such compensation does not overcome the dispersion;
4. Minimum-phase Butterworth filtering from 2.9 Hz (much noise at lower frequencies) to 5.6 Hz (i.e. Band D of Table 4.2);
5. Predictive deconvolution to whiten the data spectrum for better separability of early arrivals and short-spread reflections.

After preprocessing, the gathers in Figures 4.3a,b change into the gathers in Figures 4.6a,b, respectively. We identify the phases with the same colored arrows. In general, the minimum-phase filter in Step 4 results in later traveltimes for all body waves, and provides the desired low-frequency window on the data spectrum (gray lines in Figures 4.6c,d). The wraparound behavior at near-zero offsets does not cause a failure in source estimation, yet it would dominate the misfit function among other offsets and thus should be muted. The separability of body waves are recovered in the deconvolution step. We apply a time windowing to the data to separate the early waves from short-spread reflections. The time window starts at $t = 0$ s and finishes at $t = t_{\text{bound}}$ defined by

$$t_{\text{bound}}(h) = \begin{cases} (1.59 + h/1.608) \text{ s}, & \text{for } h \leq 5 \text{ km;} \\ 9 \text{ s}, & \text{for } h > 5 \text{ km.} \end{cases} \quad (4.5)$$

where h denotes the offset (in km). Please refer to the yellow lines in Figures 4.6a,b. Outside this window is regarded as reflections.

In order to assign an accurate water bottom (70 m depth) during inversion, we use a 143×477 grid with 35 m interval for discretization (16.7 km in width and 5 km in depth). The forward problem is solved by a classical $\mathcal{O}(\Delta t^2, \Delta x^4)$ staggered-grid finite-difference method (Levander, 1988), and the associated CFL condition leads to a sampling rate of 5 ms. The absorbing boundary condition implemented with convolutional perfectly

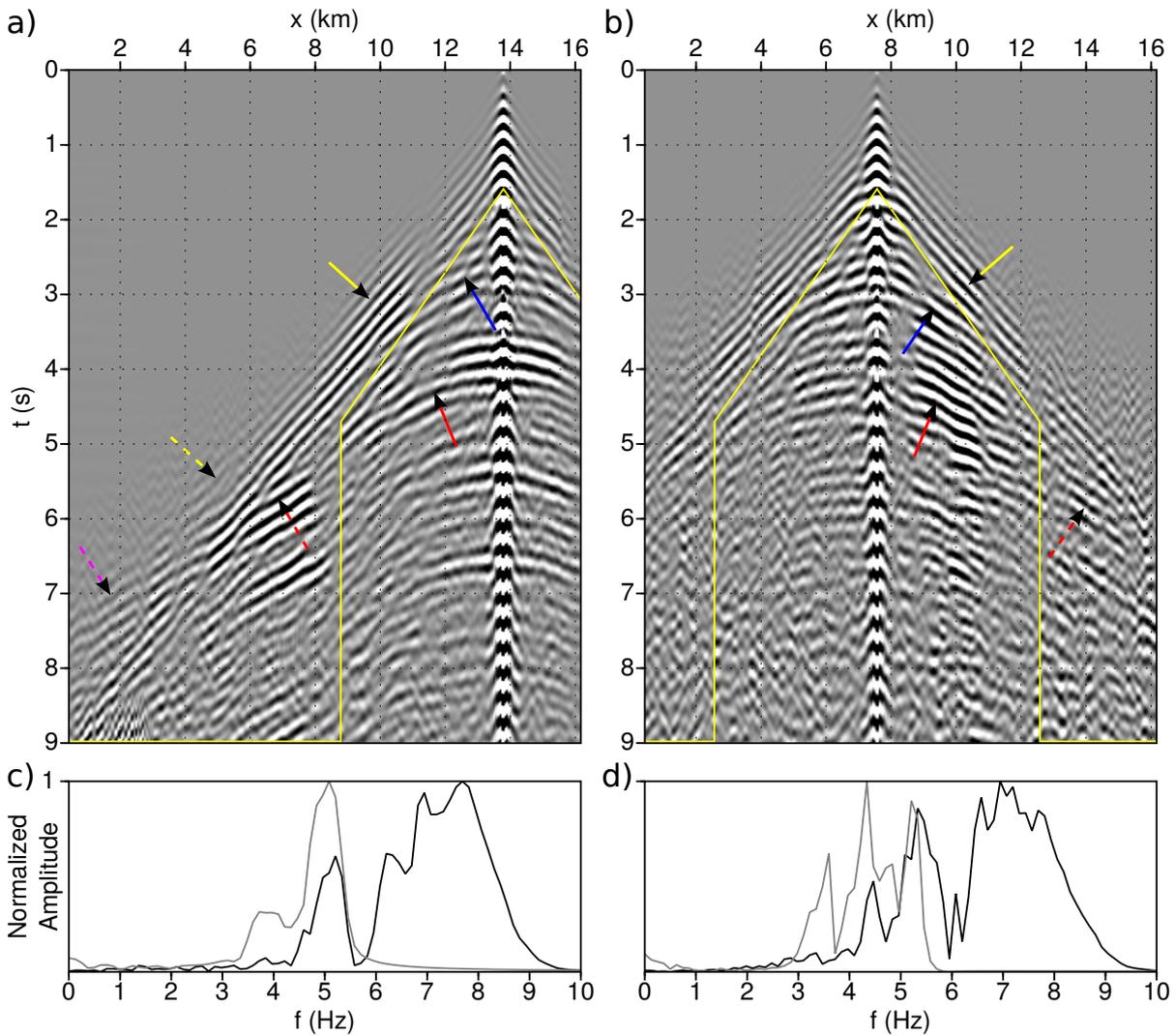


Figure 4.6: Preprocessed data (a,b) respectively from the two shot gathers shown in Figure 4.3a,b. Due to the low-pass filtering, the wavelets are less focused than in Figure 4.3 (missing high frequencies), but we still have a relatively good separability between different body waves thanks to the afterward deconvolution operation. The yellow line indicates the boundary of the time window used to separate the early arrivals from the short-spread reflections, and we use the same colored arrows to identify the typical phases as in Figure 4.3. The amplitude spectra of one trace recording at $x = 7$ km are shown in (c,d). The black lines depict the real data spectrum and the gray lines depict the spectrum after preprocessing, respectively. We can observe two frequency bands in the real data spectrum with a gap at 5.5 to 6 Hz. The low-pass filtering extracts the desired low frequency band (3–5.6 Hz), and mutes the high frequencies that are related to the attenuation issue. The extracted spectrum is not exactly proportional to the real data spectrum due to the whitening effect of the deconvolution. We preserve around half of the total energy after preprocessing.

matched layers (PMLs, e.g. [Bérenger, 1994](#); [Komatitsch et Martin, 2007](#)) is applied along the edges of the model other than the free surface.

In the chosen frequency band 3–5.6 Hz, the tomographic model (Figure 4.1a) allows for a reasonable match of the diving waves up to 8 km offsets but is not accurate enough for data fitting at longer offsets (Figure 4.7a). On the contrary, the reference 3D FWI model (Figure 4.1c) improves the match at long offsets (Figure 4.7b). To build an initial model for JFWI, we smooth the tomographic model with the 2D Gaussian function (Figure 4.8a) such that no reflected waves are generated (Figure 4.7c). However, the smoothing also increases the background velocity in the gas cloud. Therefore, we can observe a large mismatch at $x = 5$ km in Figure 4.7c. Like in Figure 4.1b, the migrated image for the caprock reflector is still discontinuous (ellipse in Figure 4.8b) due to the shortage of low wavenumbers in the proposed and tomographic velocity models. At last, we build a 1D model (Figure 4.8c) from the $x = 8$ km profile of the smoothed model to further assess the sensitivity of JFWI to the accuracy of the initial model. Note that the velocities in the gas cloud are too high to properly match the critical reflections (mismatch at $x = 6$ to 8 km in Figure 4.7d). These high velocities also lead to a deepening of the migrated images below 2.5 km depth (ellipse in Figure 4.8d). Initial density models are inferred from their respective velocity models by using the Gardner relation ($\rho = 0.23V_p^{0.25}$, [Gardner et al. \(1974\)](#)) except that the water column is set to 1 g/cm³. The smooth anisotropic models (Figures 4.1b,d) are used as background modeling parameters during inversion.

In the following, we shall successively present the classical FWI and JFWI results. The implementation of classical FWI is recast into cycles. One cycle consists of the source wavelet estimation (i.e. Step 1 in Algorithm 4.1, [Pratt et Shipp, 1999](#)) using the final velocity model from the previous cycles, followed by several nonlinear iterations of the velocity model update. Considering that the initial models produce an inaccurate wavelet, we perform only one iteration of velocity inversion in the first cycle and then increase the number of iteration in later cycles. The optimization method relies on the l -BFGS quasi-Newton approach ([Nocedal, 1980](#)). Offsets smaller than 1200 m are muted for the velocity inversion (i.e. FWI and JFWI) whereas 700 to 1200 m offsets are used in the impedance waveform inversion. No complicated weights on data are incorporated because the amplitudes of early arrivals and reflections are close after data preprocessing.

4.3.3 Results

4.3.3.1 Classical FWI

We first apply FWI to the whole offset range using the smooth initial model (Figure 4.8a). The mismatch at far offset (e.g. $x = 5$ km in Figure 4.7c) turns out to be a cycle skipping issue that guides the inversion to preferentially create high velocities in shallow zones (ellipses in Figures 4.9a), rather than decrease the velocity in the gas cloud. This wrong update can make earlier traveltimes for the synthetic early arrivals such that the misfit value is reduced. As a result, in the data match plot (Figures 4.9d), we only

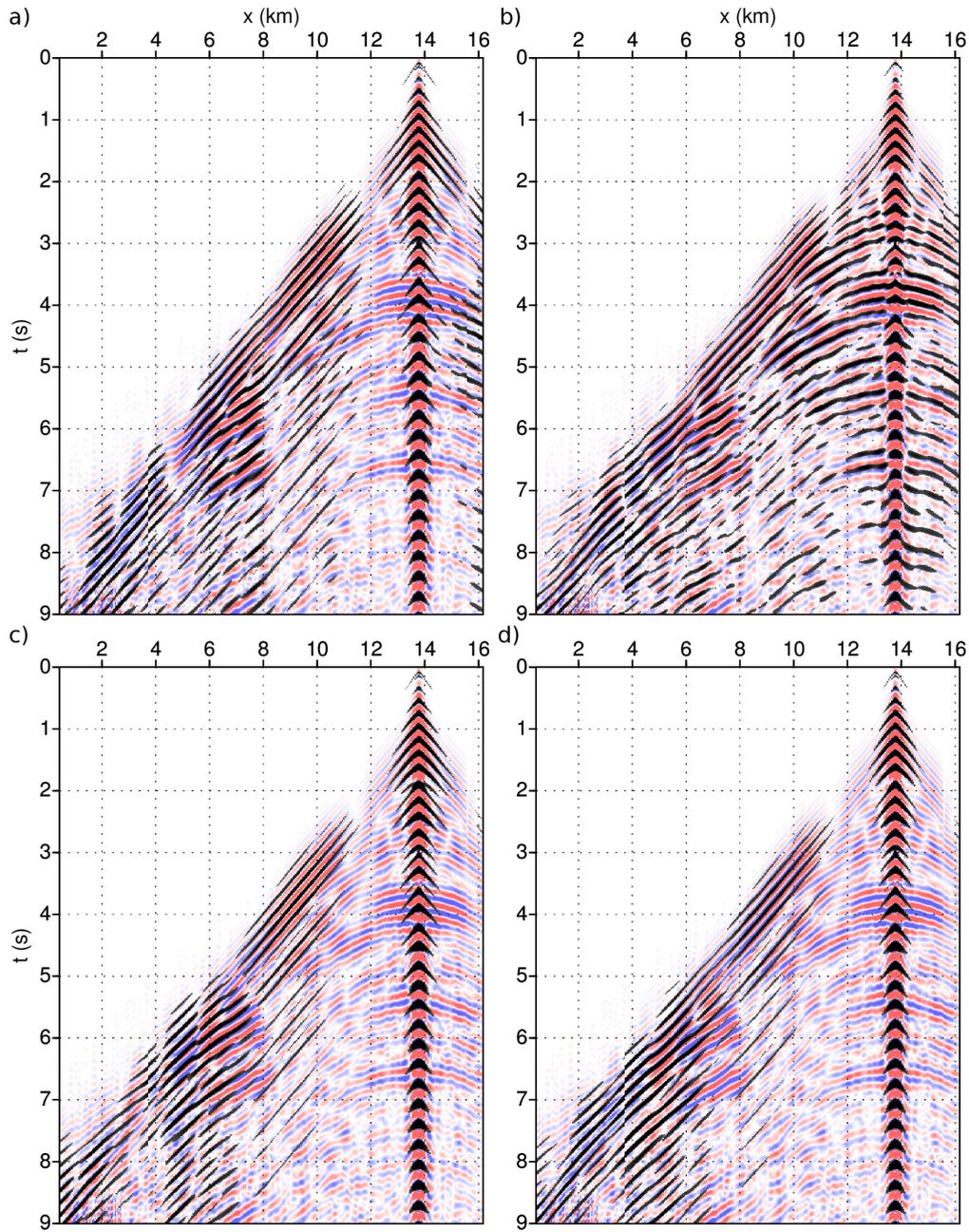


Figure 4.7: Direct comparison between real and modeled data computed in the tomographic model (a), reference model (b), smoothed model (c) and 1D model (d). The real data plotted with a blue-white-red color scale are superimposed by the modeled data plotted with a variable area wiggle display (40% of opacity). The modeled data are computed in the velocity models of Figures 4.1a,c and Figures 4.8a,c, respectively. The two datasets are in phase if the black area covers the blue part of the real data. See text for details. The color scale and the opacity are kept same for Figures 4.9d-f, Figures 4.10g-i and Figures 4.12c,d.

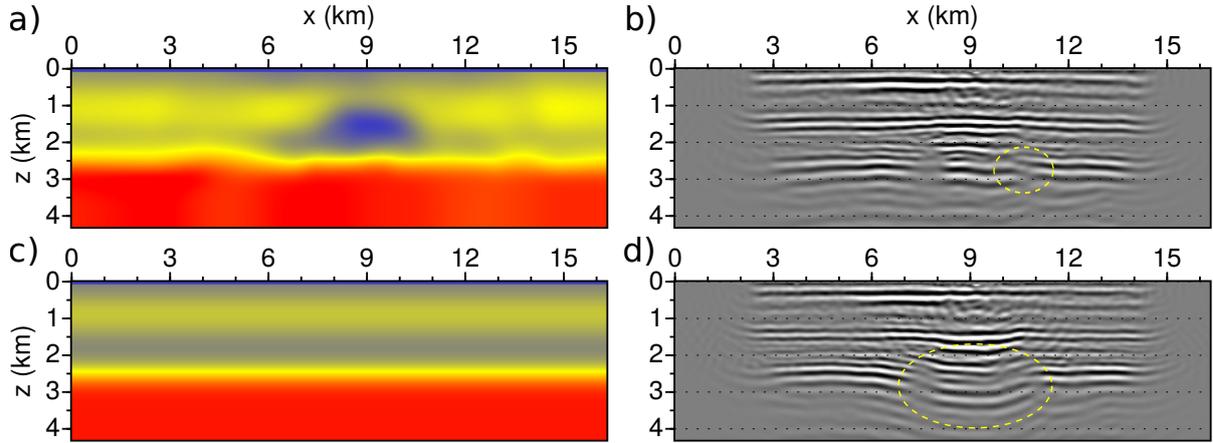


Figure 4.8: Two proposed initial models for FWI and JFWI applications (a,c), respectively built by (a) Gaussian smoothing of the tomographic model (Figure 4.1a), and (c) 1D profile at the left edge of the gas cloud without the low-velocity zone (blue region in a). No wavefield reflections take place in the considered frequency band (3–5.6 Hz). The migrated images (b,d) are generated in the same way as in Figure 4.1b using velocity models of (a,c), respectively. Because the smoothing retains some lateral variations associated to the gas cloud, the images are relatively flat in (b). Similarly, the discontinuity of reflector image present in Figure 4.1b is still present here (ellipse in b), indicating a shortage of low wavenumbers in the velocity model (a). On the other hand, the velocity field in the 1D model is too high such that the inferred image shows deepening of reflector images below 2.5 km depth (ellipse) with weaker lateral continuity.

observe the synthetic data in very short offsets which are mainly related to the source signature (black wiggles); the amplitudes at other offsets are too weak to be comparable to the observed data amplitudes, and cannot be seen within the chosen clip of this data plot: the minimization of the misfit function strongly penalizes the synthetic amplitudes.

To mitigate such cycle skipping issue, we propose an offset-driven layer-stripping strategy as follows. We first mute far offsets (larger than 5 km) to avoid cycle skipping (Figure 4.9e). This leads to a reasonable velocity reconstruction (Figure 4.9b) and a good data match within the 5 km offset range. In addition, some arrivals at longer offsets are matched although they were not involved in the inversion. The amplitudes of short-spread reflections are not so well reproduced since we have smoothed the model update for inversion stability. Then, we gradually increase the offset range to involve earlier arrivals, and relax the smoothing to match short-spread reflections. As a result, an equally match of data can be achieved at all offsets (Figure 4.9f) as well as a reasonable velocity model is reconstructed (Figure 4.9e).

Comparing with the reference model (Figure 4.1c), the degraded resolution of this FWI model can be attributed to many reasons. For example, the initial model for the 3D FWI reference model is the tomographic model that contains more high-wavenumber

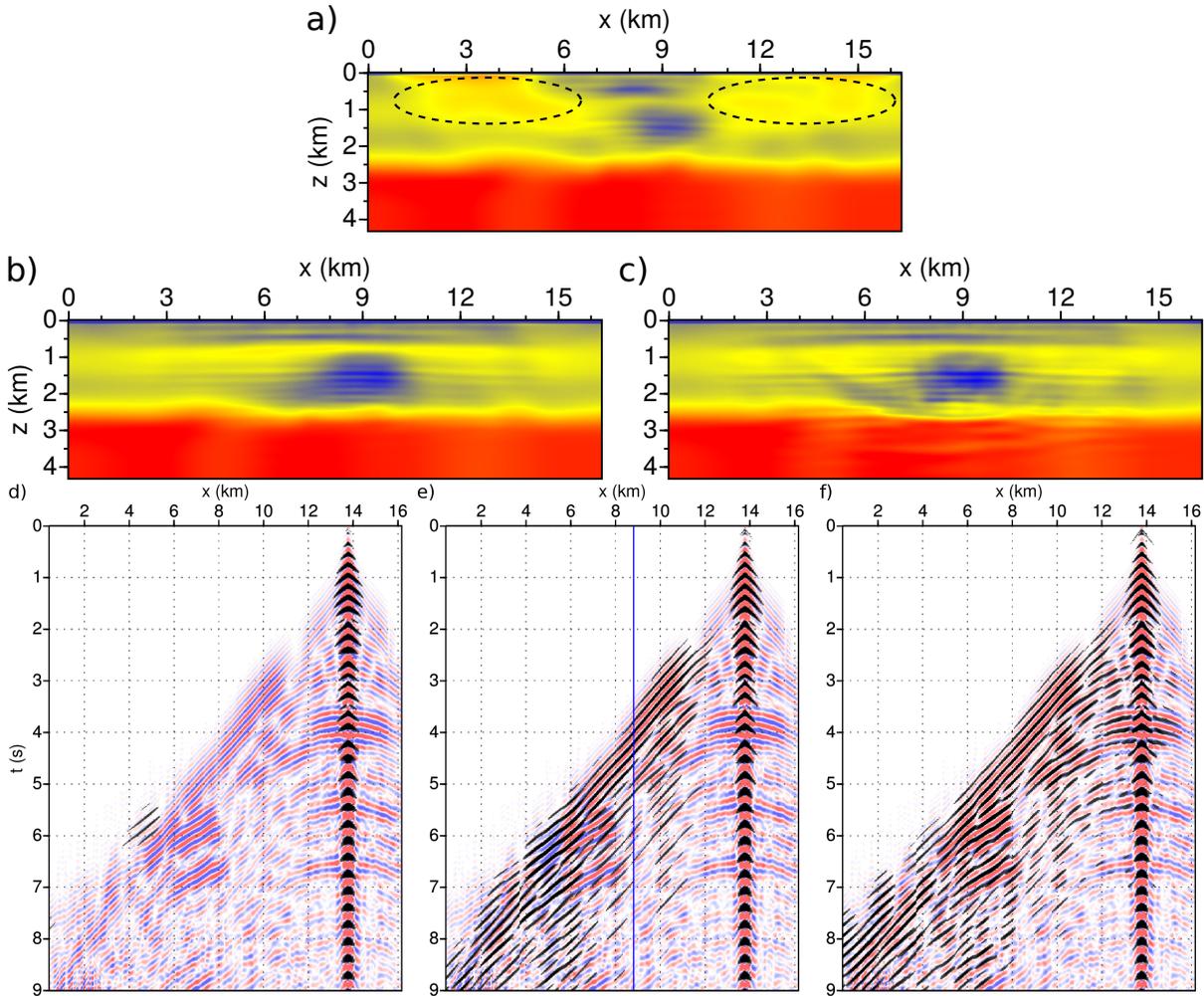


Figure 4.9: FWI models (a-c) and associated data comparisons (d-f). The cycle skipping at far offsets prevents a direct FWI application (artifacts enclosed by ellipses in a and rather weak amplitudes of the synthetic data in d). Therefore, we first use offsets smaller than 5 km (blue line in e) to build a stable intermediate model (b) and then gradually feed inversion with longer offsets. This results in a reasonable velocity reconstruction (c) and a good data match at all offsets (f). See text for details.

components than our initial model (Figure 4.8a). As a result, we may lack sufficient reflectivity information at the beginning of the inversion to highlight the low wavenumber update associated to reflected waves. We may also suffer from a shortage of data content because we use only a 2D line of the 3D data set. The deficit of low wavenumbers in the gas cloud is probably due to the narrower offset provided from the 2D geometry. The off-plane propagation may also have an impact on the results. Besides, we have excluded high frequencies to mitigate the attenuation effect, which also make the inversion more ill-posed. Nonetheless, we accept this result as a reasonable velocity model.

4.3.3.2 Joint FWI

The cycle skipping problem at far offsets also prevents JFWI from building a reasonable velocity model, if no specific strategies are used. Figures 4.10a,b,g show the JFWI velocity model using the whole offsets simultaneously, the inferred impedance model and the associated data match, respectively. Similar to the FWI result (Figures 4.9a), the cycle skipping leads to high velocities in the near surface (arrows in Figure 4.10a), which make earlier traveltimes and weaker amplitudes for the synthetic direct and diving waves. Unlike FWI, JFWI requires prior reflectivity images, therefore, reflected waves can be generated in the synthetic data. Comparing with the data match associated to the smooth model (Figure 4.7c), an improved match can be observed at $x = 5$ km. However, this is a cycle skipping phenomenon that happens to postcritical reflections. In the velocity model (Figure 4.10a), a pair of high-velocity tubes is produced (dashed lines) near the low-velocity bulb in the center of the gas cloud (blue region). These tubes are attributed to the transmission regimes generated by the two-way reflected waves, and, due to cycle skipping, they cannot be canceled out by other transmission regimes associated to reflected and/or diving waves. Due to their different wavepaths, the reflected waves preferentially sample the horizontal components of the low wavenumber vectors whereas the transmitted waves preferentially sample vertical components of the low wavenumber vectors (see the sampling analysis in Section 2.3). The cycle skipping, which mainly affect reflected waves, leads to strong lateral variations above the cap rock in the velocity model. These lateral variations in the velocity model translate into undulations in the impedance model (2.5 km depth in Figure 4.10g).

To mitigate cycle skipping, we apply the same offset-continuation strategy as proposed in the classical FWI application. Figures 4.10c,d,h show the velocity model using 5 km offset range, the inferred impedance model and the associated data match, respectively. As expected, the high velocity anomalies are avoided as well as the direct waves in a short offset range are matched, indicating no cycle skipping issue in the considered offset. Comparing with the smooth model (Figure 4.8a), the inversion has decreased the velocity in the gas cloud, therefore can improve the match on early arrivals from $x = 5.5$ to 8 km (Figure 4.10e) even though they are not included in the inversion. The inferred impedance model also shows a flatter image for the caprock reflector than the one in Figure 4.10b). Comparing with the image inferred from the smooth model (Figure 4.8b), there is a slight shift of the caprock reflector (e.g. $x = 6$ km, $z = 3$ km) coming from the low wavenumber update in the velocity model. However, the discontinuity enclosed by ellipse still exists, implying a shortage of low wavenumbers in the velocity field at this depth due to the offset constraint.

Using the same offset range (5 km), we implement JFWI starting from the 1D velocity model (Figure 4.8c). Unlike others, this model gives higher velocities for the gas cloud, therefore lacks more low wavenumber components especially in the horizontal direction. Figures 4.10e,f,i show the velocity model, the inferred impedance model and the associated data match, respectively. The inversion fails to recover sufficient low wavenumber components in the gas cloud leading to overestimated velocities. As a result,

Table 4.3: Misfit comparisons associated to JFWI models. For last two models the misfit is computed within 5 km offset range. Note their close values indicating the same level of data fitness in Figures 4.10h,i.

Models	Figure #	Misfit	Relative misfit (%)
JFWI of all offsets	4.10a	1.56E-5	100
JFWI of short offsets	4.10c	9.85E-7	6.31
JFWI of short offsets with 1D initial	4.10e	9.05E-7	5.80

the impedance model shows an apparent deepening of the caprock image (ellipse), accompanied with the unfocused energy shown as the residual images in a semicircle shape. Compared with the image inferred from the 1D model Figures 4.8d, the erroneous reflector depth is largely corrected, though incompletely, resulted from the velocity update by JFWI. However, the data match in the considered offset range (i) is reasonably as good as the data match of the previous model (h). Indeed, the related misfit values are close to each other (Table 4.3). This highlights that the desired low wavenumbers, missing in the model of Figure 4.10e, actually belong to the null space of the inverse problem due to the insufficient offset coverage. At further offsets, however, the critical reflections are less well matched at $x = 6$ km in (i) than in (h), suggesting that the null space may be reduced by involving larger offsets provided no cycle skipping problems exist.

These missing low wavenumbers are more evident by looking at the profiles of the models (Figure 4.11). The JFWI model built from the 1D initial model (blue lines in a,b,c) fails to capture the long-wavelength variations of the reference model (black lines) at 1 to 2 km depths, unlike the JFWI model with the smooth initial input (red lines). Consequently, the reflector images inferred from the poor JFWI model (blue lines in d) tend to be mispositioned comparing with those inferred from the reference model (black lines) below 0.5 km depth, whereas the reflectors inferred from the effective JFWI model (red lines) shows good agreement with those reference reflectors above 2.2 km depth. The discrepancy below this depth can be explained by the insufficient low wavenumber updates due to the limited offset range.

In summary, we have confronted with the cycle skipping problem at far offsets in both FWI and JFWI implementations. Consequently, we limit the offset range to mitigate cycle skipping at the expense of updates of low wavenumber components in the gas cloud carried in the far offsets. These low wavenumbers are believed to reside in the null space of the inverse problem. Therefore, a sufficiently accurate initial model is expected.

4.3.3.3 Joint FWI followed by FWI (JFWI+FWI)

The JFWI models Figure 4.10b and c can be naturally taken as the initial models of monoparameter FWI for broadband reconstruction of the velocity parameter. We perform the inversion in the time domain and the same frequency band is used. To mitigate

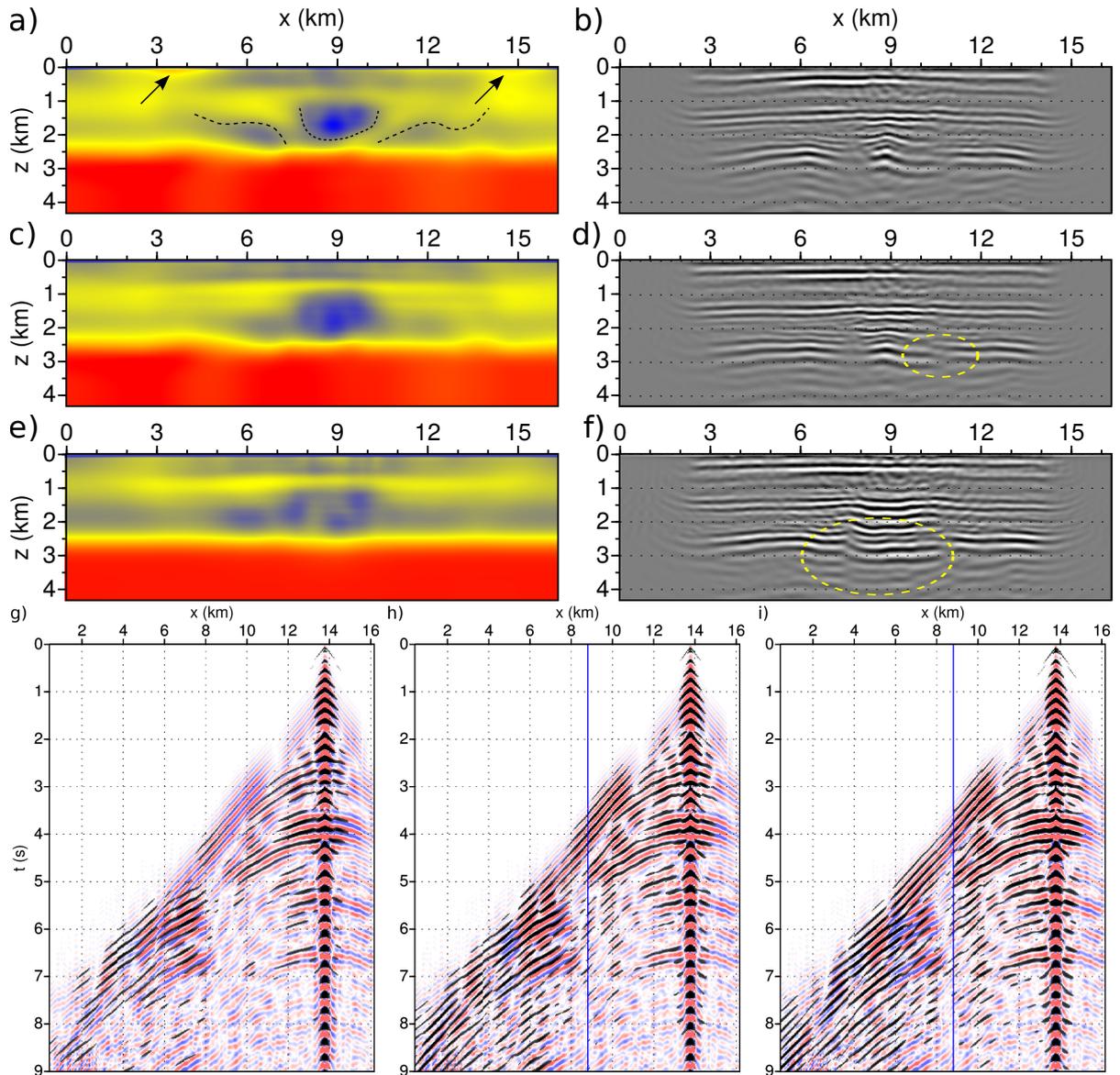


Figure 4.10: JFWI models (a,c,e), inferred reflectivity (b,d,f) and associated data comparisons (g-i). The cycle skipping at far offsets prevents a direct application of JFWI (a,b,g). Therefore, we limit the offset within 5 km (blue line in h) as in the FWI application (Figure 4.9e). This leads to a stable velocity model (c) with flatter reflector image (d) as well as a good data match at short offsets (h). However, the discontinuous caprock image (ellipse in d) indicates the insufficient low wavenumber components in (c). To assess its sensitivity to the accuracy of initial model, we apply JFWI to the 1D model (Figure 4.8c) using 5 km offset range. The inversion fails to recover the low velocities in the gas cloud (e), making discontinuous caprock images (ellipse in f) despite a good match at short offsets (i). The equally match in (h) and (i) highlights the nonuniqueness of the inverse problem due to the small offset range. See text for details.

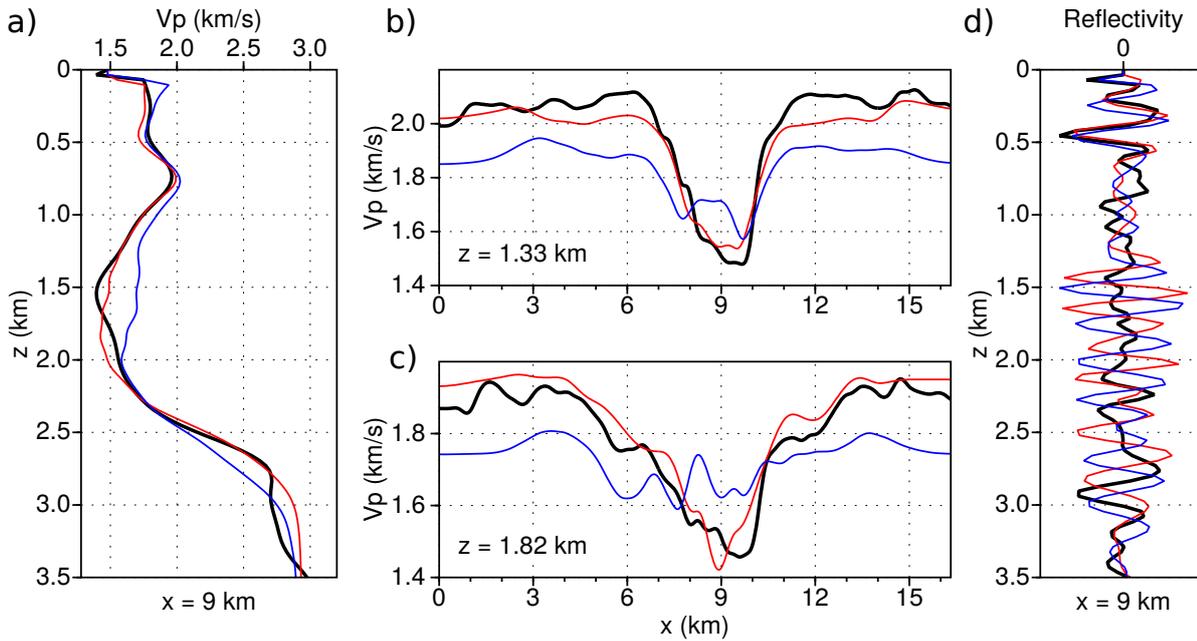


Figure 4.11: Vertical (a) and horizontal (b,c) profiles of JFWI models of Figures 4.10c (red) and Figure 4.10e (blue), as well as reference velocity model of Figure 4.1c (black) in (b,c) and the difference between Figures 4.1c,d as the reference background model (black in a). Vertical (d) profiles of reflectivity of Figure 4.10d (red) and Figure 4.10f (blue) inferred from the two JFWI models, respectively, as well as the reference reflectivity model of Figure 4.1d (black). Because the low wavenumbers are better recovered in Figure 4.10c, the velocity profile and inferred reflectivity depicted in red are closer to the black referencing profiles, especially at 1 to 2 km depths.

the cycle skipping issue, we gradually increase the offset range from 5 km to the maximum available offset (13.5 km). Figures 4.12a,b,e show the FWI result using the JFWI model of Figure 4.10c, the extracted reflectivity model and the data comparison, respectively. The reflectivity models are built by filtering out large-scale components from the velocity models using Gaussian smoothing. Again, we do not perform migration as in Figure 4.10 because we favor high-resolution images for comparison (remind the low-frequency band D used in migration). Except the small-scale updates by FWI, we also see the large-scale updates at 2.5 km depth in the velocity model coming from far offset data. Because of this, the discontinuity of the caprock reflector image that is present in Figure 4.10d (ellipse) is mitigated in (b) to some extent. With higher frequencies, we can show more focused images to better assess this improvement. Nonetheless, the associated data plot (e) shows a high level of match at all offsets.

The JFWI+FWI result using the poor JFWI model (Figure 4.10e), the extracted reflectivity model and the data comparison are shown in Figures 4.12c,d,f, respectively. The high velocities built by JFWI in the gas cloud (Figure 4.10e) are not fully corrected in (c), leading to a curved reflector image at 1.5 km depth (arrow in d). Consequently, in or-

Table 4.4: Misfit comparisons associated to JFWI and JFWI+FWI models. For fair comparisons, the misfits are computed from all available offsets without source signals. Therefore, the two JFWI misfits are different from the values in Table 4.3. Note how FWI has reduced the misfit in both cases, and built two models with close misfit values indicating the same level of data fitness in Figures 4.12e,f.

Models	Figure #	Misfit	Relative misfit (%)
Good JFWI	4.10c	2.50E-5	100
Good JFWI + FWI	4.12a	2.02E-5	80.91
Poor JFWI	4.10e	2.45E-5	98.15
Poor JFWI + FWI	4.12c	1.98E-5	79.19

der to match the vertical two-way traveltimes of reflected waves, a relatively low-velocity zone (arrow in c) is generated to compensate for these high velocities. In addition, compared with the caprock reflector image shown in Figure 4.10f, the image in (d) is more continuous inside the ellipse due to the low wavenumber updates derived from far offsets.

Interestingly, the associated data comparison (f) shows a same level of match as (f) despite the incorrect velocities in the gas cloud. Indeed, the related misfit values are close to each other (Table 4.4), which highlights the nonuniqueness of the inverse problem. From JFWI to JFWI+FWI, although we have used far offsets, this nonuniqueness means that we still suffer from insufficient data content to uniquely image the subsurface model (ill-posedness). Using the whole 3D data set, or including higher frequencies should provide a supplementary constraint to mitigate such nonuniqueness. Nonetheless, judging from the reference model, we accept the velocity model of Figure 4.12a as a reasonable result.

The accuracy of each JFWI+FWI models are more evident by comparing the profiles. Figure 4.13 shows the profiles of the two initial models (Figures 4.8a,c), the inferred JFWI models (Figures 4.10c,e) and JFWI+FWI models (Figures 4.12a,c) as well as the reference model, respectively. The vertical profiles (a) of the accepted JFWI+FWI model (green lines) depicts the injection of short wavelengths into the JFWI models (red lines) mapped from reflectivity. More importantly, we see an update of long wavelengths in the lateral profile (c) between $x = 3$ to 6 km and $x = 12$ km when comparing with the reference model (black lines). This is attributed to the use of far-offset critical reflections that can contribute to horizontal low-wavenumber components in the velocity update. In contrast, starting from the poor JFWI model (blue line in b), the subsequent FWI of all offsets (magenta line) fails to correct the long wavelengths at 1 to 2 km depths. Therefore, the magenta line is far from the black reference line especially at depths between 1 and 1.7 km, unlike the accepted JFWI+FWI model (green line in a). Below 2 km, velocity uncertainty increases due to insufficient data content and angle illumination, and thus we see a degraded match of the reference by both JFWI+FWI models. Nonetheless, the caprock reflector inferred from the accepted model (green arrow in a) is positioned at a similar depth with the black reference reflector (black arrow), whereas the reflector

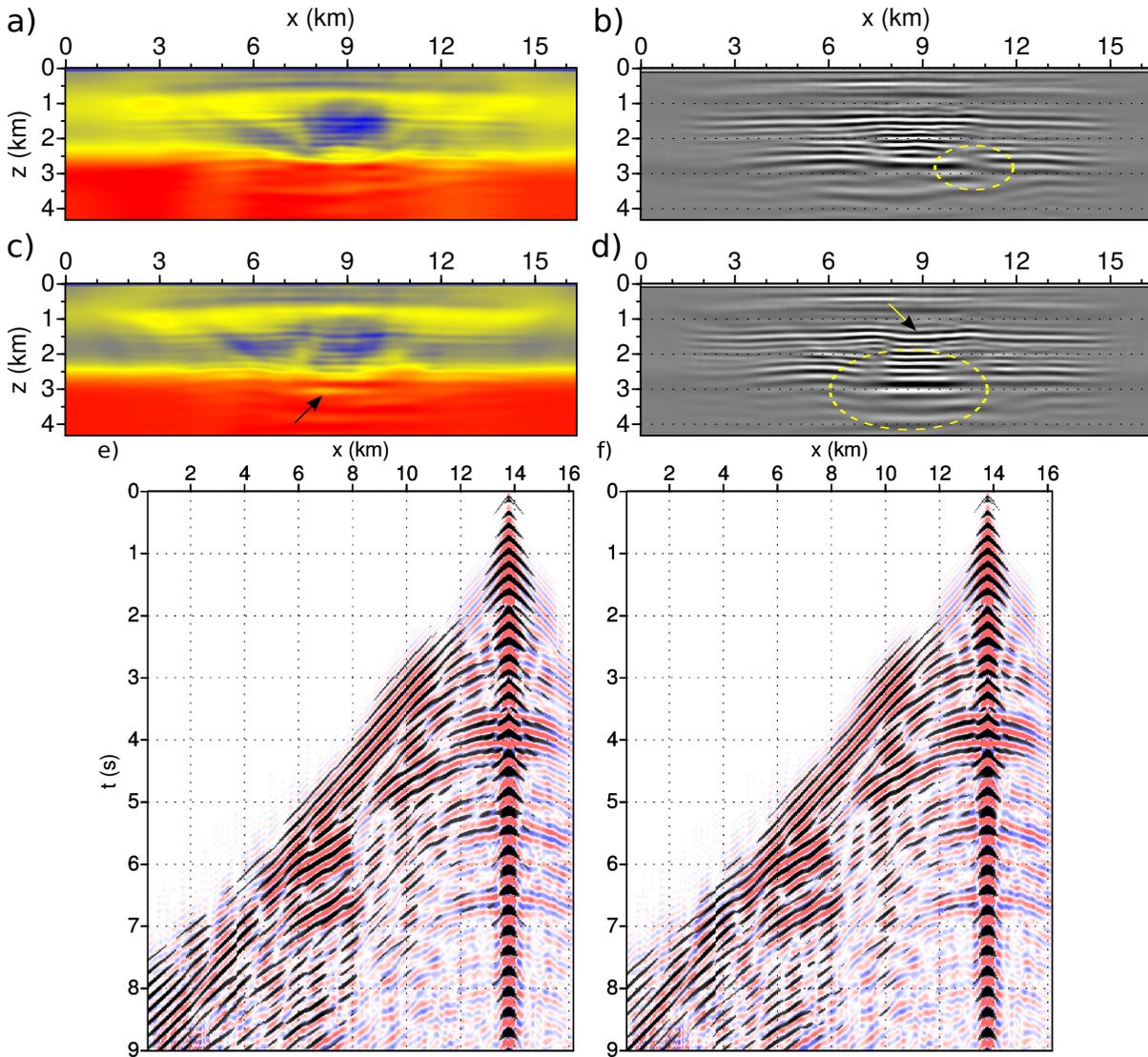


Figure 4.12: FWI models with JFWI models as initial input (a,c), extracted reflectivity models (b,d) and data comparisons (e,f). (a,b,e) starting from JFWI model of Figure 4.10c. (b,d,f) starting from JFWI model of Figure 4.10e. The reflectivity models are extracted from the velocity models by filtering out large-scale components from the velocity models using Gaussian smoothing. A broadband velocity model is built by FWI (a), with the improved reflector images (ellipse in b) thanks to the low-to-intermediate wavenumber updates given by far offsets. The use of far offsets also flattens the caprock reflector image (ellipse in d) that was curved in Figure 4.10f (ellipse). However, the reflector crossing the gas cloud is still non flat (arrow in d) due to the high velocity background (c). The low-velocity zone pointed by the arrow in (c) is a consequence of such high velocity in order to match the data. Despite its high level of data match (f) the velocity model proposed in (c) should not be trusted due to the dissimilar feature with the reference model of Figure 4.1c.

inferred from the poor model (magenta arrow in b) is about 200 m deeper than the reference reflector: this difference reflects a better recovery of velocities in the gas cloud in the accepted JFWI+FWI model.

Comparing with the direct implementation of FWI (Figure 4.9e), we find an improved resolution at 6 km location between 1.6 and 2.1 km depths resulted from JFWI followed by FWI (blue ellipse in Figure 4.14). The JFWI+FWI model is closer to the reference model than the FWI model due to the prior implementation of JFWI, who has successfully recovered more low-to-intermediate wavenumber components than FWI. Such improvement could be found at other locations when the 3D geometry is considered in our further investigations.

4.3.3.4 Quality control of velocity models by common image gathers

We assess the quality of our inversion results by generating the common image gathers (CIGs) in the surface-offset–depth domain (Figures 4.15 and 4.16). Six models are respectively assessed: the smooth initial model (Figure 4.8a), the 1D initial model (Figure 4.8c) and JFWI models starting from the two initial models (Figures 4.10c,e), as well as the JFWI+FWI models starting from the JFWI models (Figures 4.12a,c), at six positions above the gas cloud ($x = 6$ to 11 km, from top to bottom rows). The images are built by reverse time migration. For high-resolution images we use the broadband data (Figure 4.3), including high frequencies that have attenuation footprints (Section 4.3.1). Therefore, we solve a viscoacoustic wave equation based on the rheology of generalized Maxwell body (Emmerich et Korn, 1987; Moczo et Kristek, 2005). We use three relaxation mechanisms to approximate the frequency-independent Q (Emmerich et Korn, 1987), and estimate the anelastic coefficients via least-squares fitting (Yang et al., 2016). We use a homogeneous $Q = 200$ during modeling to account for the attenuation below the sea bed, since it has led to reasonable reconstructions as shown in Section 4.3.1. Due to the source-receiver reciprocity, the images in the negative offset gathers (receiver on left side of source) should be same with the images in the positive offset gathers (receiver on right side of source) that share the same midpoints. Therefore, we sum the “twin” gathers to increase the S/N of the images, and display only absolute offsets in 0–10 km.

Because it is difficult to separate direct and diving waves from the reflected waves from shallow interfaces, we generate two sets of CIGs using different portions of the data, which allow to assess the quality of velocity models at different depths. In the first set (Figures 4.15), we incorporate all waves and no preprocessing is applied. Therefore, images associated to diving waves are present below 0.5 km depth with linear residual moveouts, for example, the dashed curves in the top left panel. The images above this depth level are mainly associated to reflected waves from shallow interfaces. They are rather discontinuous at offsets beyond 2 km in the CIGs inferred from the smooth model (a) and 1D model (b). In contrast, we can see the significant improvements of image flatness and focusing provided by JFWI (c,d) due to the low-wavenumber updates at shallow depths in both cases. Moreover, the JFWI model inferred from the smooth model (c) produces flatter images than the one inferred from the 1D initial model (d),

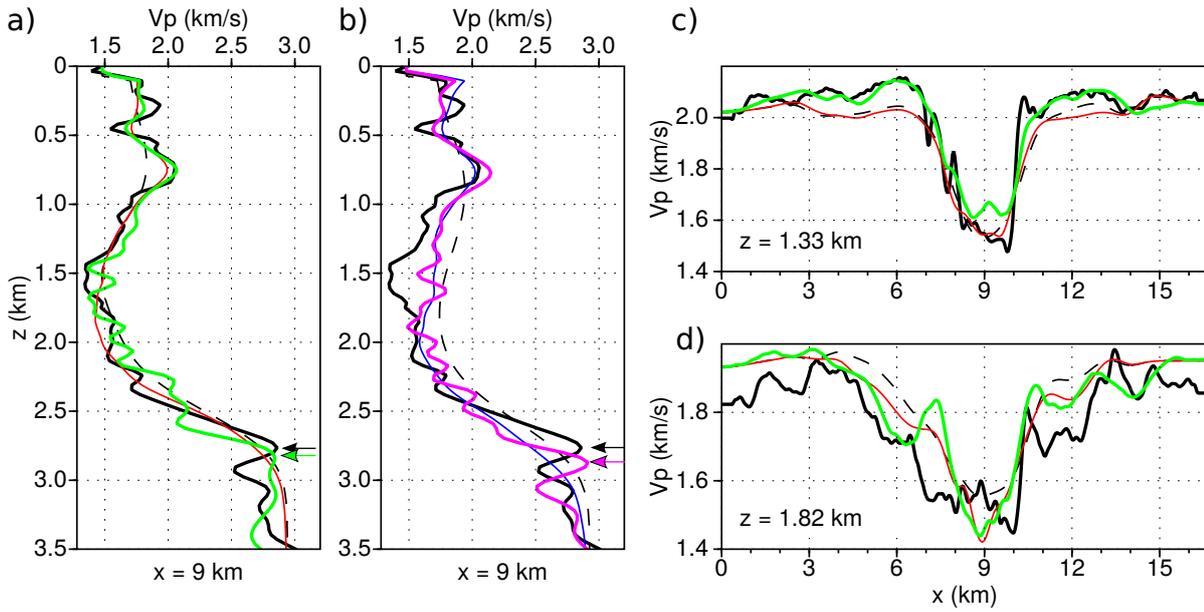


Figure 4.13: Vertical (a,b) and horizontal (c,d) profiles of FWI models (green, magenta) with JFWI inputs of Figure 4.10b (red) and of Figure 4.10c (blue), respectively, as well as profiles of proposed initial models of Figures 4.8a,c (dashed black) and reference model (Figure 4.1c, black). Arrows indicate the caprock reflector inferred from the two JFWI+FWI models (green and magenta) and the reference model (black), respectively. See text for details.

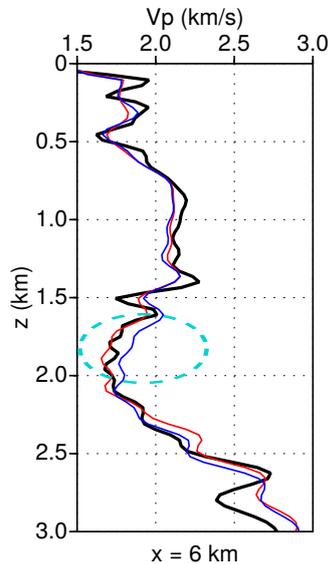


Figure 4.14: Vertical profile of JFWI+FWI model (red), FWI model (Figure 4.9c, blue) and reference model (Figure 4.1c, black). Note that the red line is closer to the black line than the blue one at 1.6 to 2.1 km depths (blue ellipse) due to the improved low-to-intermediate wavenumber updates by JFWI.

as delineated by yellow curves in middle panels, which is consistent with our former analysis that the former JFWI model has better kinematic accuracy. On the other hand, the images inferred from JFWI+FWI models (e,f) are difficult to follow laterally because the panels are noisy. We attribute this decreased S/N to multiple reflections generated from shallow interfaces that are imaged by FWI, and an inadequate Q for viscoacoustic modeling in the near surface. Nonetheless, improvements of image flatness can be seen just below the sea bed (e.g. positions pointed by arrows), which means a better near surface reconstruction achieved by FWI using larger offsets.

The second set of CIGs (Figures 4.16) are generated by using reflected waves inside a time window, which is defined from an offset-dependent instant t to the maximum instant (8 s), where $t = 1.59 + h/1.563$ s and h denotes the offset (in km). The value 1.563 km/s is chosen to exclude direct waves without losing too many reflected waves at short-to-intermediate offsets. The disagreement of reflector depths between (a) and (b) at zero offsets is resulted from the different velocity fields provided from the smooth model (a) and 1D model (b). The velocities in the gas cloud are more accurate in the smooth model as indicated by the flat images (yellow line in a) compared with the curved images (yellow line in b). Such superiority is preserved after the implementation of JFWI (yellow lines in c,d). Indeed, starting from the 1D model JFWI did not fully recover the low wavenumber components of the gas cloud due to the limited offset range. On the other hand, starting from the smooth model with more low wavenumber content, JFWI improves the resolution in the gas cloud, therefore leads to flatter images comparing (d) with (c), cyan lines. The low wavenumber information carried by far offsets are further injected to the velocity models by the ultimate FWI implementation. We see the improvements of image flatness from (c) to (d), green lines for the smooth initial case, as well as the image flatness from (d) to (f), magenta lines for the 1D initial case. Note that the magenta line is not as flat as the green line in (e,f), which implies that we could use high frequencies to mitigate the nonuniqueness of inversion models that are encountered in the Figure 4.12. The non flat images in (e) can be attributed to inaccurate velocities proposed by the JFWI+FWI model, or other issues that are not considered in this study, such as the off-plane effect due to 3D propagations.

4.4 Discussion

In this study, the successful application of JFWI was driven by the need to manage the attenuation footprint in the data and the cycle skipping issue. As argued in many other investigations, waveform inversion requires an accurate forward modeling algorithm that should be able to mimic the real physics of the recorded data. However, for the sake of computational efficiency we do not take attenuation into account, and inevitably generate inversion artifacts when using high frequencies. Nevertheless, we emphasize that this is not a limitation of JFWI. In the future we believe high frequencies can be involved in JFWI with the development of computing infrastructures and viscous modeling techniques.

On the contrary, like classical FWI, the cycle skipping issue is a problem of JFWI in

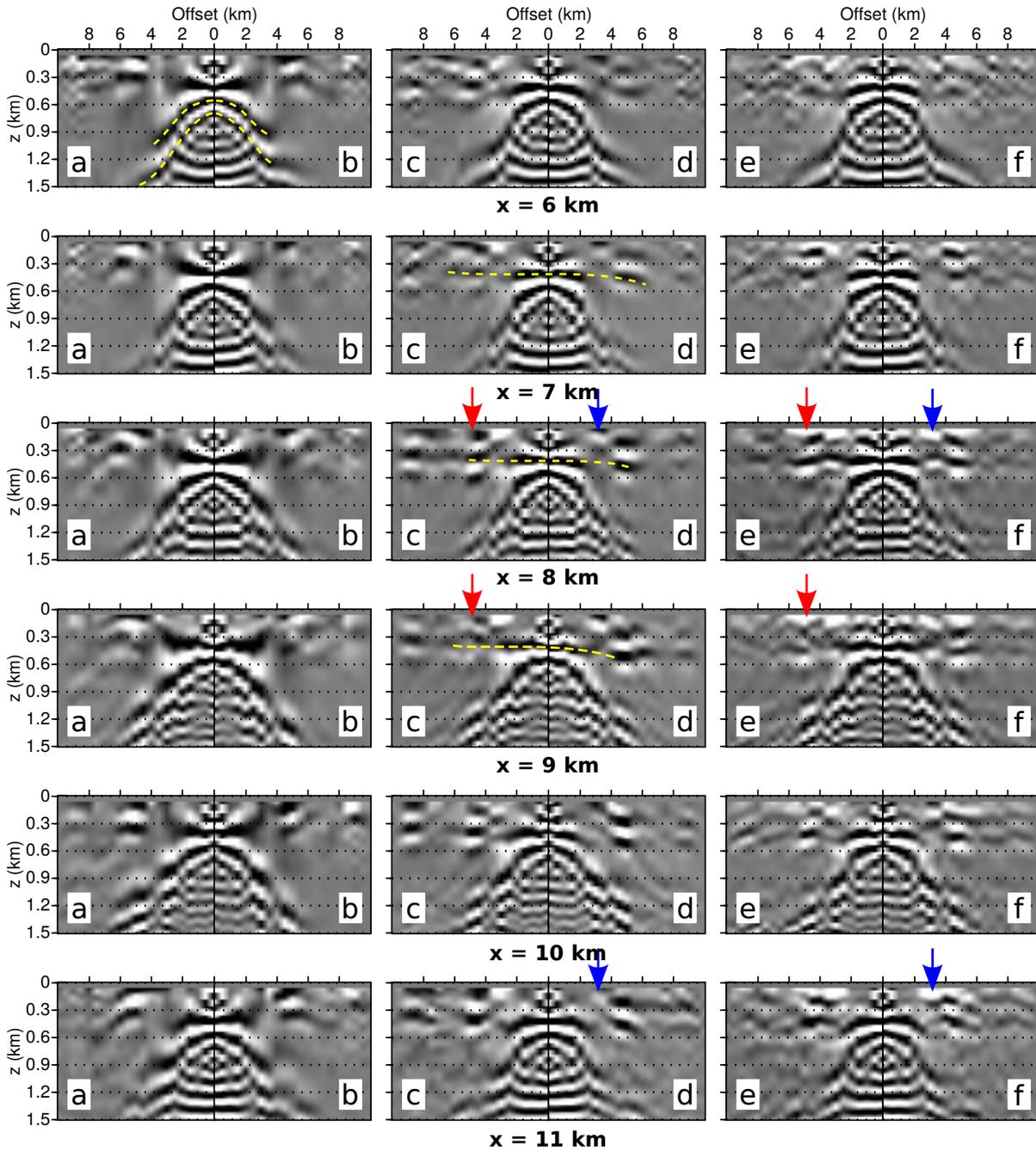


Figure 4.15: Surface-offset domain common image gathers at $x = [6\ 7\ 8\ 9\ 10\ 11]$ km positions, using (a) smooth initial model (Figure 4.8a), (b) 1D initial model (Figure 4.8c), (c,d) JFWI models (Figure 4.10c and Figure 4.10e) of short offsets respectively using the smooth and 1D initial model, and (e,f) JFWI+FWI models (Figure 4.12a and Figure 4.12c). Broadband data (Figure 4.3) are used without muting direct and diving waves, which generate images with large slopes (e.g. yellow curves in the top left panel). Improvements to image flatness are pointed. See text for discussions.

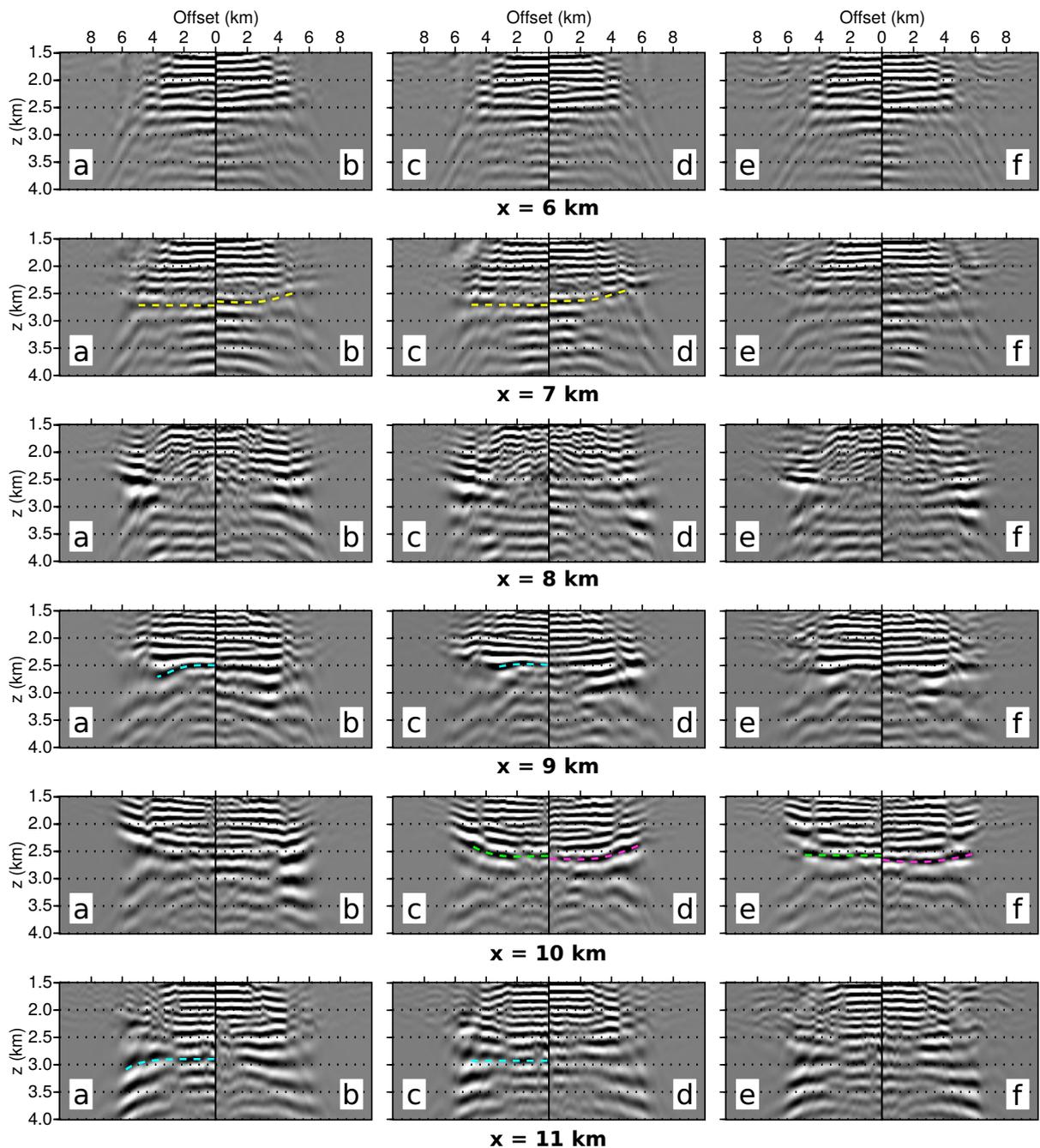


Figure 4.16: Surface-offset domain common image gathers at $x = [6\ 7\ 8\ 9\ 10\ 11]$ km positions, using six models (a-f) as explained in Figure 4.15. Broadband data (Figure 4.3) are used, and direct and diving waves are muted by time windowing to assess the image flatness in deep zones. Improvements to image flatness are pointed. See text for discussions.

the sense that a brute-force application to the real data always outputs unphysical models corresponding to local minima. In this study, we have identified the cycle skipping problem occurred on early arrivals at far offsets. This results from the low-pass filtering implemented in the data preprocessing which mixes up the diving waves with the (pre)critical reflections. Besides, this also makes difficult the data separation driven by time windowing. Because of the cycle skipping and rough data separation, the inversion falls into a local minimum giving high-velocity anomalies. Possible remedies are the use of more robust misfit functions (e.g. Warner *et al.*, 2014; Métivier *et al.*, 2015) and phase identification techniques (e.g. Ma *et al.*, 2013).

To tackle the cycle skipping at far offsets, in this study, we limit the offset range during inversion, which inevitably limits our ability to update low-to-intermediate wavenumbers due to the insufficient wide angle coverage. Comparing with the 3D FWI model, the JFWI model with the 1D initial fails to properly recover the gas cloud, unlike in the synthetic case that a successful reconstruction was achieved using a crude 1D initial model (Section 3.5). Therefore, a better initial model is required such that the needed low wavenumbers are already available, such as the 2D model derived from the existing tomographic models. Otherwise, an unphysical model could be built with good data match due to the insufficient data content making the inverse problem very ill-posed.

In this study, a key question needs to answer is whether or not we can separate the scales of the subsurface by using this data set. The concept of scale separation concept has long served as the fundamental assumption in seismic imaging. However, the real Earth is a medium of continuous spectrum and intermediate wavenumbers have been exploited in various FWI applications. Indeed, the continuous mapping from angle apertures to imaged wavenumbers (Devaney, 1982) allows to retrieve the intermediate wavenumbers from intermediate-to-long reflection angles/offsets. At postcritical reflection angles, intermediate and low wavenumbers simultaneously define the kinematics of the postcritical reflections and refracted waves (e.g. critical offsets, traveltimes and their slopes, etc.). Considering this, we suspect that the aforementioned cycle-skipping problem occurred on early arrivals is actually a matter of missing intermediate wavenumbers in the initial model: methods based on scale separation, like JFWI cannot be readily applied to these offsets. As soon as this assumption is valid, for example, by using short offsets JFWI starts to recover low wavenumbers from the data.

Therefore, we consider that for this data set it is more suitable to keep intermediate wavenumbers during the inversion, and aim to build a broadband velocity model. As in Operto *et al.* (2015) the tomographic model is used as their initial model which contain sufficient intermediate wavenumbers to reproduce the critical reflections, and they were able to reconstruct a reasonably good model via a direct implementation of classical FWI. However, we emphasize the advantage of JFWI over classical FWI, the improved low wavenumber samplings in difficult area such as the low-velocity gas cloud by reflected waves without generating cycle skipping at short offsets (Chapter 2). Injecting intermediate-to-high wavenumbers, without much effort, can be achieved by performing a subsequent classical FWI starting from the JFWI model.

4.5 Conclusions and perspectives

In this study, we have combined the sensitivity kernels associated to early arrivals and reflections for a robust velocity macromodel building that is suitable for a subsequent classical FWI implementation. We have applied the approach to the 2D long-offset data set collected across a gas cloud. We have found that the attenuation imprint in the gas cloud has a strong impact on the time-domain inversion, when attenuation is not taken into account in the modeling process, therefore only low frequencies are considered. Critical or postcritical reflections and refracted waves from below the gas cloud (i.e. early arrivals at far offsets) are crucial to update the low wavenumbers in the gas cloud that significantly affect the kinematic accuracy, but the cycle-skipping issue prevents a direct use of these waves. Therefore, we use an offset-continuation strategy which progressively incorporates longer offset in the inversion combined with the use of a more accurate initial model. Narrowing the offset range and the aperture illumination of the subsurface limits the ability of the inversion to update the low part of the wavenumber spectrum. Therefore, it is still required to design an accurate initial model for the proposed least-squares misfit function. A reasonably good velocity model is built with high level of data fit, through the workflow that alternates between JFWI and impedance waveform inversion followed by classical FWI.

Further investigations could deal with the 3D extension of JFWI to assess whether the improved coverage could bring more low wavenumber samplings of the gas, whether the full data cube may be sufficient to eliminate nonuniqueness of the solution, and whether off-plane effects/propagation have influence in the 2D experiment. For these goals, we may require more robust misfit functions to mitigate the related cycle skipping problem and relax the need of an accurate velocity model.

Acknowledgments

This study is funded by the SEISCOPE consortium (<http://seiscope2.osug.fr>) sponsored by BP, CGG, CHEVRON, EXXON-MOBIL, JGI, PETROBRAS, SAUDI ARAMCO, SCHLUMBERGER, SHELL, SINOPEC, STATOIL, TOTAL and WOODSIDE. The authors thank BP Norge AS and Hess Norge AS for providing the data set. They appreciate fruitful discussions with François Audebert (TOTAL), Andrzej Górszczyk (Polish academy of sciences), Ludovic Métivier (LJK-CNRS, UGA), Paul Wellington (UGA) and Yang Li (Harbin Institute of Technology). This study is granted access to the HPC resources of the Froggy platform of the CIMENT infrastructure (<https://ciment.ujf-grenoble.fr>), which is supported by the Rhône-Alpes region (GRANT CPER07_13 CIRA), the OSUG@2020 labex (reference ANR10 LABX56) and the Equip@Meso project (reference ANR-10-EQPX-29-01) of the programme Investissements d’Avenir supervised by the Agence Nationale pour la Recherche, and the HPC resources of CINES/IDRIS under the allocation 046091 made by GENCI.

Chapter 5

Conclusions and Perspectives

Full waveform inversion (FWI) is a powerful tool for broadband subsurface imaging. However, it often requires an accurate initial velocity model and/or sufficient low-frequency content that are not easily available. Although modern acquisition geometries can provide very long-offset/wide-azimuth coverage, the angle illumination is still insufficient for FWI to properly recover deep targets. Artifacts cannot be totally avoided due to the shortage of data constraint, which requires complementary information to quality control the results. In contrast, reflection waveform inversion (RWI) is carried out in the frame of scale separation between large and small wavelengths of subsurface. The ambiguity of velocity and reflector depths are decoupled, such that the misfit function associated to RWI presents less local minima than that of FWI, therefore the line searching can target the true model more easily. Furthermore, it has been shown that, by highlighting the transmission regimes associated to reflected waves, RWI enhances small-wavelength reconstructions at greater depths than FWI, therefore is more suitable for the surface acquisition geometry. However, RWI omits the diving waves that are quite useful for near-surface imaging.

In this thesis, I propose joint full waveform inversion (JFWI) to combine diving waves and reflected waves for velocity macromodel building. Similar to RWI, this new method is formulated under the scale separation assumption and benefits from reflected waves to image large wavelengths of deep targets. However, it also includes the diving waves in its formulation. Moreover, by analyzing the radiation patterns, I choose the V_P - I_P parameterization of the subsurface to emphasize during inversion the scale separation between the velocity model and reflectivity. Consequently, a cycle workflow is proposed to alternate the velocity model building performed by JFWI and impedance imaging performed by classical FWI of reflected waves (IpWI).

The synthetic Valhall experiment has shown the importance of using diving waves for velocity inversion, as they can bring large-wavelength information which is not carried by reflected waves. Compared with RWI, JFWI improves the velocity images at shallow depths, which are further translated to a better recovery of deep targets. This implies that, when the initial velocity model is inaccurate and the offset range is limited, RWI is

unable to fit all reflected waves even for those generated by shallow interfaces; the additional information provided by diving waves assists in better constraining the inversion at shallow depths and improve the fitting of all reflected waves. The promising velocity model built by JFWI followed by FWI numerically confirms that 1) from a crude initial model, JFWI can build a velocity model that is located inside the attraction valley of the global minimum of the FWI misfit function; 2) the large-wavelength components in the smooth JFWI velocity model are sufficient for FWI to build a broadband model.

The proposed hierarchical workflow (JFWI followed by FWI) is further assessed in the real case study of the 2D data set including effect coming from a gas cloud above a reservoir. For high numerical efficiency, I do not model the attenuation effect, and the attenuation footprint in the results is mitigated by considering enough low frequencies below 5.6 Hz. On the other hand, the cycle-skipping resulted from postcritical reflections and refracted waves prevents a direct use of long offsets, and I use an offset-continuation strategy starting from a more accurate initial model to stabilize the inversion. The acceptable broadband velocity model built by JFWI followed by FWI manifests the ability of JFWI to build a reasonably good initial model for classical FWI, if 1) the cycle skipping problem is mitigated, and 2) sufficiently long offsets are available or alternatively, a better initial model is provided for JFWI that is not so accurate as expected for FWI.

Further necessary work should consider more convex misfit function rather than the L2-norm difference-based misfit function (Ma et Hale, 2013; Warner et Guasch, 2014; Métivier *et al.*, 2015, etc). The candidates should allow an explicit operation to separate the early arrivals and reflected waves, and for this reason, the ones formulated in the time domain are preferred. In perspectives, such a good misfit function can allow a direct use of the whole data, unlike in the synthetic and real-data case where I should select a proper time window or offset range to mitigate cycle skipping. Undoubtedly, the benefit is to build the velocity model in a more robust way.

I am also encountered with the cross-talks issue during the broadband reconstruction process performed by classical FWI starting from the JFWI model. It is generated because I use the V_P - ρ parameterization which introduces the ambiguity between V_P and ρ for small scattering angles or short offsets. The mitigation of such cross-talk effects has been well identified and studied in the literature (Operto *et al.*, 2013). Using the second-order information is considered to tackle this problem and successful examples have been published (Métivier *et al.*, 2014a).

JFWI is also oriented to 3D applications in the future. Nonunique solutions presented in the real-case study reflect the insufficiency of data coverage as only a 2D line is considered. Longer offsets provided from the diagonal direction of the 3D geometry may relax the need of accurate initial models for JFWI. The increased elapse time from 2D to 3D can be offset by using the time-depth conversion (Plessix, 2013; Brossier *et al.*, 2015). On the other hand, the attenuation effect has prevented me from using higher frequencies and recovering high-resolution models, therefore viscous modeling should be

embedded in the future to honor the name “full waveform”. Besides, extensions to elastic inversion could be considered: how to introduce the shear-wave velocity into the frame of scale separation between compressional velocity and acoustic impedance is still an open question in this study. For this purpose, I am also interested in other datasets to further assess the promises and pitfalls of the proposed hierarchical workflow in different circumstances.

Bibliography

- AL-YAHYA, K. (1989). Velocity analysis by iterative profile migration. *Geophysics*, 54(6):718–729.
- ALBERTIN, U., SAVA, P., ETGEN, J. et MAHARRAMOV, M. (2006). Adjoint wave-equation velocity analysis. In *SEG Technical Program Expanded Abstracts, 76th SEG Annual Meeting*, pages 3345–3349.
- ALERINI, M., LAMBARÉ, G., BAINA, R., PODVIN, P. et BÉGAT, S. L. (2007). Two-dimensional PP/PS-stereotomography: P- and S-waves velocities estimation from OBC data. *Geophysical Journal International*, 170:725–736.
- ALKHALIFAH, T. (2014). Monofrequency waveform acquisition and inversion: a new paradigm. In *Expanded Abstracts, 84th Annual SEG Meeting (Denver)*.
- ALKHALIFAH, T. (2015). Scattering-angle based filtering of the waveform inversion gradients. *Geophysical Journal International*, 200(1):363–373.
- ALKHALIFAH, T. et CHOI, Y. (2012). Taming waveform inversion non-linearity through phase unwrapping of the model and objective functions. *Geophysical Journal International*, 191:1171–1178.
- ALKHALIFAH, T. et GUITTON, A. (2016). An optimal parameterization for full waveform inversion in anisotropic media. In *78th EAGE Conference and Exhibition 2016*.
- ALKHALIFAH, T. et PLESSIX, R. (2014). A recipe for practical full-waveform inversion in anisotropic media: An analytical parameter resolution study. *Geophysics*, 79(3):R91–R101.
- ALKHALIFAH, T. et WU, Z. (2014). FWI and MVA the natural way. In *EAGE, 76th EAGE Conference and Exhibition incorporating SPE EUROPEC 2014*.
- ALKHALIFAH, T. et WU, Z. (2015). Migration velocity analysis using prestack wavefields. In *77th EAGE Conference and Exhibition 2015*.
- ALLEMAND, T. et LAMBARÉ, G. (2014). Full waveform inversion guided migration velocity analysis. In *Expanded Abstracts*, pages 4712–4717. Society of Exploration Geophysics.

BIBLIOGRAPHY

- ALLEMAND, T. et LAMBARÉ, G. (2015). Combining full waveform inversion and tomography: Full waveform inversion-guided tomography. *In 77th EAGE Conference and Exhibition 2015*. EAGE.
- ALMOMIN, A. et BIONDI, B. (2012). Tomographic full waveform inversion: Practical and computationally feasible approach. *SEG Technical Program Expanded Abstracts 2012*, pages 1–5.
- ALTHEYAB, A., WANG, X. et SCHUSTER, G. T. (2013). Time-domain incomplete gauss-newton full-waveform inversion of gulf of mexico data. *In SEG Technical Program Expanded Abstracts 2013*, pages 5175–5179.
- AMESTOY, P., BROSSIER, R., BUTTARI, A., L'EXCELLENT, J.-Y., MARY, T., MÉTIVIER, L., MINIUSI, A., OPERTO, S., RIBODETTI, A., VIRIEUX, J. et WEISBECKER, C. (2015). Efficient 3D frequency-domain full-waveform inversion of ocean-bottom cable data with sparse block low-rank direct solver: a real data case study from the North Sea. *In Expanded Abstracts, 85th Annual SEG Meeting (New Orleans)*, pages 1303–1308. SEG.
- ASNAASHARI, A., BROSSIER, R., GARAMBOIS, S., AUDEBERT, F., THORE, P. et VIRIEUX, J. (2013). Regularized seismic full waveform inversion with prior model information. *Geophysics*, 78(2):R25–R36.
- BAEK, H., CALANDRA, H. et DEMANET, L. (2014). Velocity estimation via registration-guided least-squares inversion. *Geophysics*, 79(2):R79–R89.
- BAI, J., YINGST, D., BLOOR, R. et LEVEILLE, J. (2014). Viscoacoustic waveform inversion of velocity structures in the time domain. *Geophysics*, 79(3):R103–R119.
- BANSAL, R., KREBS, J., ROUTH, P., LEE, S., ANDERSON, J., BAUMSTEIN, A., MULLUR, A., LAZARATOS, S., CHIKICHEV, I. et MCADOW, D. (2013). Simultaneous-source full-wavefield inversion. *The Leading Edge*, 32(9):1100–1108.
- BARKVED, O., ALBERTIN, U., HEAVEY, P., KOMMEDAL, J., van GESTEL, J., SYNNOVE, R., PETERSEN, H. et KENT, C. (2010a). Business impact of full waveform inversion at Valhall. *In Expanded Abstracts, 91 Annual SEG Meeting and Exposition (October 17-22, Denver)*, pages 925–929. Society of Exploration Geophysics.
- BARKVED, O., HEAVEY, P., KOMMEDAL, J. H., van GESTEL, J.-P., VE, R. S., PETERSEN, H., KENT, C. et ALBERTIN, U. (2010b). Business impact of full waveform inversion at valhall. *SEG Technical Program Expanded Abstracts*, 29(1):925–929.
- BARNES, C. et CHARARA, M. (2009). The domain of applicability of acoustic full-waveform inversion for marine seismic data. *Geophysics*, 74(6):WCC91–WCC103.
- BAYSAL, E., KOSLOFF, D. et SHERWOOD, J. (1983). Reverse time migration. *Geophysics*, 48:1514–1524.

- BELAHI, T., SINGH, S. et FUJI, N. (2016). Viscoelastic full waveform inversion of sea bottom long offset seismic data in presence of attenuation. *In 78th EAGE Conference and Exhibition 2016*.
- BELLER, S., MONTEILLER, V., OPERTO, S., NOLET, G., COMBE, L., MÉTIVIER, L., VIRIEUX, J., NISSEN-MEYER, T. et PAUL, A. (2014). 3d lithospheric imaging by time-domain full-waveform inversion of teleseismic body-waves. *In Abstracts*, page A2388.
- BEN HADJ ALI, H., OPERTO, S. et VIRIEUX, J. (2008). Velocity model building by 3D frequency-domain, full-waveform inversion of wide-aperture seismic data. *Geophysics*, 73(5):VE101–VE117.
- BEN HADJ ALI, H., OPERTO, S. et VIRIEUX, J. (2011). An efficient frequency-domain full waveform inversion method using simultaneous encoded sources. *Geophysics*, 76(4):R109.
- BÉRENGER, J.-P. (1994). A perfectly matched layer for absorption of electromagnetic waves. *Journal of Computational Physics*, 114:185–200.
- BERKHOUT, A., VERSCHUUR, D., STAAL, X. et DAVYDENKO, M. (2015). Utilization of multiples in simultaneous image building and velocity estimation. *In Proceedings of the 77th EAGE Conference & Exhibition, Workshops*. EAGE.
- BIJWAARD, H. et SPAKMAN, W. (2000). Non-linear global p-wave tomography by iterated linearized inversion. *Geophysical Journal International*, 141:71–82.
- BILLETTE, F. et LAMBARÉ, G. (1998). Velocity macro-model estimation from seismic reflection data by stereotomography. *Geophysical Journal International*, 135(2):671–680.
- BILLETTE, F., LE BÉGAT, S., PODVIN, P. et LAMBARÉ, G. (2003). Practical aspects and applications of 2D stereotomography. *Geophysics*, 68:1008–1021.
- BIONDI, B. et ALMOMIN, A. (2012). Tomographic full waveform inversion (tfwi) by combining full waveform inversion with wave-equation migration velocity analysis. *SEG Technical Program Expanded Abstracts 2012*, pages 1–5.
- BIONDI, B. et ALMOMIN, A. (2013). Tomographic full waveform inversion (TFWI) by combining FWI and wave-equation migration velocity analysis. *The Leading Edge*, September, special section: full waveform inversion:1074–1080.
- BIONDI, B. et ALMOMIN, A. (2014). Simultaneous inversion of full data bandwidth by tomographic full-waveform inversion. *Geophysics*, 79(3):WA129–WA140.
- BIONDI, B. et SYMES, W. (2004). Angle-domain common-image gathers for migration velocity analysis by wavefield-continuation imaging. *Geophysics*, 69(5):1283–1298.

BIBLIOGRAPHY

- BIONDI, B. et TISSERANT, T. (2004). 3d angle-domain common-image gathers for migration velocity analysis. *Geophysical Prospecting*, 52(6):575–591.
- BISHOP, T. N., BUBE, K. P., CUTLER, R. T., LANGAN, R. T., LOVE, P. L., RESNICK, J. R., SHUEY, R. T. et SPINDER, D. A. (1985). Tomographic determination of velocity and depth in laterally varying media. *Geophysics*, 50:903–923.
- BORISOV, D., STOPIN, A. et PLESSIX, R.-E. (2014). Acoustic pseudo-density full waveform inversion in the presence of hard thin beds. In *Expanded Abstracts, 76th Annual EAGE Meeting (Amsterdam)*.
- BRANDSBERG-DAHL, S., DE HOOP, M. et URSIN, B. (1999). Velocity analysis in the common scattering-angleazimuth domain. In *SEG Technical Program Expanded Abstracts, 69th Annual Meeting (Houston)*, pages 1715–1718. SEG.
- BRENDERS, A. J. et PRATT, R. G. (2007). Efficient waveform tomography for lithospheric imaging: implications for realistic 2D acquisition geometries and low frequency data. *Geophysical Journal International*, 168:152–170.
- BRETAUDEAU, F., BROSSIER, R., MÉTIVIER, L. et VIRIEUX, J. (2014). First-arrival delayed tomography using 1st and 2nd order adjoint-state method. In *Expanded Abstracts*, pages 4757–4762. Society of Exploration Geophysics.
- BROSSIER, R., ETIENNE, V., HU, G., OPERTO, S. et VIRIEUX, J. (2013). Performances of 3D frequency-domain full-waveform inversion based on frequency-domain direct-solver and time-domain modeling: application to 3D OBC data from the Valhall field. *International Petroleum Technology Conference, Beijing, China*, page IPTC 16881.
- BROSSIER, R., OPERTO, S. et VIRIEUX, J. (2009a). Robust frequency-domain full-waveform inversion using the l_1 norm. *Geophysical Research Letters*, 36:L20310.
- BROSSIER, R., OPERTO, S. et VIRIEUX, J. (2009b). Seismic imaging of complex onshore structures by 2D elastic frequency-domain full-waveform inversion. *Geophysics*, 74(6):WCC105–WCC118.
- BROSSIER, R., OPERTO, S. et VIRIEUX, J. (2010). Which data residual norm for robust elastic frequency-domain full waveform inversion? *Geophysics*, 75(3):R37–R46.
- BROSSIER, R., OPERTO, S. et VIRIEUX, J. (2015). Velocity model building from seismic reflection data by full waveform inversion. *Geophysical Prospecting*, 63:354–367.
- BROSSIER, R., PAJOT, B., COMBE, L., OPERTO, S., MÉTIVIER, L. et VIRIEUX, J. (2014). Time and frequency-domain FWI implementations based on time solver: analysis of computational complexities. In *Expanded Abstracts, 76th Annual EAGE Meeting (Amsterdam)*.
- BROSSIER, R., VIRIEUX, J. et OPERTO, S. (2008). Parsimonious finite-volume frequency-domain method for 2-D P-SV-wave modelling. *Geophysical Journal International*, 175(2):541–559.

- BUNKS, C., SALEK, F. M., ZALESKI, S. et CHAVENT, G. (1995). Multiscale seismic waveform inversion. *Geophysics*, 60(5):1457–1473.
- BÖZDAG, E., TRAMPERT, J. et TROMP, J. (2011). Misfit functions for full waveform inversion based on instantaneous phase and envelope measurements. *Geophysical Journal International*, 185(2):845–870.
- CARCIONE, J. M., KOSLOFF, D. et KOSLOFF, R. (1988). Wave propagation simulation in a linear viscoacoustic medium. *Geophysical Journal International*, 93(2):393–401.
- CASTELLANOS, C., MÉTIVIER, L., OPERTO, S., BROSSIER, R. et VIRIEUX, J. (2015). Fast full waveform inversion with source encoding and second-order optimization methods. *Geophysical Journal International*, 200(2):720–744.
- ČERVENÝ, V. (2001). *Seismic Ray Theory*. Cambridge University Press, Cambridge.
- CHAPMAN, C. (2004). *Fundamentals of seismic waves propagation*. Cambridge University Press, Cambridge, England.
- CHAPMAN, C. H., HOBRO, J. W. D. et ROBERTSSON, J. O. A. (2014). Correcting an acoustic wavefield for elastic effects. *Geophysical Journal International*, 197:1196–1214.
- CHAURIS, H. (2000). *Analyse de vitesse par migration pour l'imagerie des structures complexes en sismique réflexion*. Thèse de doctorat, Ecole des Mines de Paris.
- CHAURIS, H., LAMELOISE, C.-A. et COCHER, E. (2015). Inversion velocity analysis - the importance of regularisation. In *Proceedings of the 77th EAGE Conference & Exhibition*. EAGE.
- CHAURIS, H. et NOBLE, M. (2001). 2D velocity macro model estimation from seismic reflection data by local Differential Semblance Optimization: applications on synthetic and real data. *Geophysical Journal International*, 144:14–26.
- CHAURIS, H., NOBLE, M., LAMBARÉ, G. et PODVIN, P. (2002). Migration velocity analysis from locally coherent events in 2-D laterally heterogeneous media, Part I: Theoretical aspects. *Geophysics*, 67(4):1202–1212.
- CHAURIS, H. et PLESSIX, R.-E. (2012). Investigating the differential waveform inversion. In *EAGE Technical Program Expanded Abstracts 2012*, pages "presented in Workshop: From Kinematic to Waveform Inversion – Where Are we and where Do we Want to Go? A Tribute to Patrick Lailly".
- CHAVENT, G. (1996). Duality methods for waveform inversion. Rapport technique, Duality methods for waveform inversion, Research report, RR-2975, inria-00073723.
- CHAVENT, G., CLÉMENT, F. et GÒMEZ, S. (1994). Automatic determination of velocities via migration-based traveltime waveform inversion: A synthetic data example. *SEG Technical Program Expanded Abstracts 1994*, pages 1179–1182.

BIBLIOGRAPHY

- CHAVENT, G., GADYLSHIN, K. et TCHEVERDA, V. (2014). Full waveform inversion in migration based travel-time formulation. *In 6th EAGE Saint Petersburg International Conference and Exhibition.*
- CHOI, Y. et ALKHALIFAH, T. (2016). An optimized correlation-based full waveform inversion. *In 78th EAGE Conference and Exhibition 2016.*
- CLAERBOUT, J. (1985). *Imaging the Earth's interior.* Blackwell Scientific Publication.
- CLAPP, R. (2008). Reverse time migration: Saving the boundaries. Rapport technique SEP-136, Stanford Exploration Project.
- CLÉMENT, F., CHAVENT, G. et GÓMEZ, S. (2001). Migration-based travelttime waveform inversion of 2-D simple structures: A synthetic example. *Geophysics*, 66:845–860.
- COCHER, E., CHAURIS, H. et LAMELOISE, C.-A. (2015). Imaging with surface-related multiples in the subsurface-offset domain. *In Proceedings of the 77th EAGE Conference & Exhibition.* EAGE.
- CRASE, E., PICA, A., NOBLE, M., McDONALD, J. et TARANTOLA, A. (1990). Robust elastic non-linear waveform inversion: application to real data. *Geophysics*, 55:527–538.
- DABLAIN, M. (1986). The application of high order differencing for the scalar wave equation. *Geophysics*, 51:54–66.
- DEMME, J. W. (1997). *Applied numerical linear algebra.* SIAM, Philadelphia.
- DEVANEY, A. (1984). Geophysical diffraction tomography. *Geoscience and Remote Sensing, IEEE Transactions on*, GE-22(1):3–13.
- DEVANEY, A. J. (1982). A filtered backprojection algorithm for diffraction tomography. *Ultrasonic Imaging*, 4:336–350.
- DIAZ, E., SAVA, P. et YANG, T. (2013). Data-domain and image-domain wavefield tomography. *The Leading Edge*, September:1064–1072.
- DUMBSER, M. et KÄSER, M. (2006). An Arbitrary High Order Discontinuous Galerkin Method for Elastic Waves on Unstructured Meshes II: The Three-Dimensional Isotropic Case. *Geophysical Journal International*, 167(1):319–336.
- DUTTA, G. et SCHUSTER, G. T. (2014). Attenuation compensation for least-squares reverse time migration using the viscoacoustic-wave equation. *Geophysics*, 79(6):S251–S262.
- DUVENECK, E. (2013). A pragmatic approach for computing full-volume rtm reflection angle/azimuth gathers. *In 75th EAGE Conference & Exhibition incorporating SPE EUROPEC 2013.* EAGE.

- EMMERICH, H. et KORN, M. (1987). Incorporation of attenuation into time-domain computation of seismic wavefield. *Geophysics*, 52:1252–1264.
- ERLANGGA, Y. A. et HERRMANN, F. J. (2008). An iterative multilevel method for computing wavefields in frequency-domain seismic inversion. *In Expanded Abstracts*, pages 1956–1960. Soc. Expl. Geophys.
- ESSER, E., GUASCH, L., HERRMANN, F. J. et WARNER, M. (2016). Constrained waveform inversion for automatic salt flooding. *The Leading Edge*, 35(3):235–239.
- ETGEN, J., GRAY, S. H. et ZHANG, Y. (2009). An overview of depth imaging in exploration geophysics. *Geophysics*, 74(6):WCA5–WCA17.
- ETIENNE, V., HU, G., OPERTO, S., VIRIEUX, J., BARKVED, O. et KOMMEDAL, J. (2012). Three-dimensional acoustic full waveform inversion: algorithm and application to Valhall. *In Expanded Abstracts, 74th Annual EAGE Conference & Exhibition, Copenhagen*. EAGE.
- ETIENNE, V., OPERTO, S., VIRIEUX, J. et JIA, Y. (2010). Computational issues and strategies related to full waveform inversion in 3d elastic media: methodological developments. *SEG Technical Program Expanded Abstracts*, 29(1):1050–1054.
- FARRA, V. et MADARIAGA, R. (1988). Non-linear reflection tomography. *Geophysical Journal*, 95:135–147.
- FEI, W. et WILLIAMSON, P. (2010). On the gradient artifacts in migration velocity analysis based on differential semblance optimization. *In SEG Technical Program Expanded Abstracts*. SEG.
- FICHTNER, A., KENNETT, B. L. N., IGEL, H. et BUNGE, H. P. (2008). Theoretical background for continental- and global-scale full-waveform inversion in the time-frequency domain. *Geophysical Journal International*, 175:665–685.
- FICHTNER, A., KENNETT, B. L. N., IGEL, H. et BUNGE, H. P. (2010). Full waveform tomography for radially anisotropic structure: New insights into present and past states of the Australasian upper mantle. *Earth and Planetary Science Letters*, 290(3–4):270–280.
- FLEURY, C. et PERRONE, F. (2012). Bi-objective optimization for the inversion of seismic reflection data: combined FWI and MVA. *In Expanded Abstracts*, volume 548, pages 1–6. Society of Exploration Geophysics.
- FORGUES, E. et LAMBARÉ, G. (1997). Parameterization study for acoustic and elastic ray+born inversion. *Journal of Seismic Exploration*, 6:253–278.
- FORNBERG, B. (1988). Generation of finite difference formulas on arbitrarily spaced grids. *Mathematics of Computation*, 51:699–706.

BIBLIOGRAPHY

- FREY, P. et GEORGE, P. (2008). *Mesh Generation*. ISTE Ltd & John Wiley Sons Inc, London (UK) & Hoboken (USA).
- GAO, F., WILLIAMSON, P. et PRATT, R. G. (2014). A new objective function for full waveform inversion: Differential semblance in the data domain. *In SEG Technical Program Expanded Abstracts 2014*, pages 1178–1183.
- GARDNER, G. H. F., GARDNER, L. W. et GREGORY, A. R. (1974). Formation velocity and density—the diagnostic basics for stratigraphic traps. *Geophysics*, 39:770–780.
- GAUTHIER, O., VIRIEUX, J. et TARANTOLA, A. (1986). Two-dimensional nonlinear inversion of seismic waveform : numerical results. *Geophysics*, 51:1387–1403.
- GÉLIS, C., VIRIEUX, J. et GRANDJEAN, G. (2007). 2D elastic waveform inversion using Born and Rytov approximations in the frequency domain. *Geophysical Journal International*, 168:605–633.
- GHOLAMI, Y., BROSSIER, R., OPERTO, S., PRIEUX, V., RIBODETTI, A. et VIRIEUX, J. (2013a). Which parametrization is suitable for acoustic VTI full waveform inversion? - Part 2: application to Valhall. *Geophysics*, 78(2):R107–R124.
- GHOLAMI, Y., BROSSIER, R., OPERTO, S., RIBODETTI, A. et VIRIEUX, J. (2013b). Which parametrization is suitable for acoustic VTI full waveform inversion? - Part 1: sensitivity and trade-off analysis. *Geophysics*, 78(2):R81–R105.
- GORSZCZYK, A., MALINOWSKI, M. et OPERTO, S. (2016). Crustal-scale imaging from ultra-long offset node data by full waveform inversion-how to do it right? *In 78th EAGE Conference and Exhibition 2016*.
- GOSSELET, A., BÉGAT, S. L. et PETERSEN, S. (2005). Joint slope tomography of borehole transmitted and surface seismic data. *In Expanded Abstracts*, pages 2577–2580.
- GUASCH, L. et WARNER, M. (2014). Adaptive Waveform Inversion - FWI Without Cycle Skipping - Applications. *In 76th EAGE Conference and Exhibition 2014*.
- GUITTON, A. (2012). Blocky regularization schemes for full waveform inversion. *Geophysical Prospecting*, 60(5):870–884.
- GUO, Q. et ALKHALIFAH, T. (2016). Elastic reflection based waveform inversion in isotropic media. *In 78th EAGE Conference and Exhibition 2016*. EAGE.
- HA, T., CHUNG, W. et SHIN, C. (2009). Waveform inversion using a back-propagation algorithm and a Huber function norm. *Geophysics*, 74(3):R15–R24.
- HAK, B. et MULDER, W. A. (2010). Migration for velocity and attenuation perturbations. *Geophysical Prospecting*, 58:939–951.
- HALE, D. (2013). Dynamic warping of seismic images. *Geophysics*, 78(2):S105–S115.

- HAMITOU, O., MÉTIVIER, L., LABBÉ, S., BROSSIER, R., VIRIEUX, J. *et al.* (2015). Pre-conditioning and multiple-right hand sides strategies for the solution of the frequency-domain wave propagation problem using the cgmn method. *In 2015 SEG Annual Meeting*. SEG.
- HOBRO, J. W. D., CHAPMAN, C. H. et ROBERTSSON, J. O. A. (2014). A method for correcting acoustic finite-difference amplitudes for elastic effects. *Geophysics*, 79:T243–T255.
- HOBRO, J. W. D., SINGH, S. C. et MINSHULL, T. A. (2003). Three-dimensional tomographic inversion of combined reflection and refraction seismic traveltimes data. *Geophysical Journal International*, 152:79–93.
- HUANG, J. et BELLEFLEUR, G. (2012). Joint transmission and reflection traveltimes tomography using the fast sweeping method and the adjoint-state technique. *Geophysical Journal International*, 188(2):570–582.
- HUANG, Y. et SCHUSTER, G. (2014). Resolution limits for wave equation imaging. *Journal of Applied Geophysics*, 107:137–148.
- JANNANE, M., BEYDOUN, W., CRASE, E., CAO, D., KOREN, Z., LANDA, E., MENDES, M., PICA, A., NOBLE, M., ROETH, G., SINGH, S., SNIEDER, R., TARANTOLA, A. et TREZEGUET, D. (1989). Wavelengths of Earth structures that can be resolved from seismic reflection data. *Geophysics*, 54(7):906–910.
- JONES, I. (2010). Tutorial: Velocity estimation via ray-based tomography. *First Break*, 28(2):45–52.
- KÄSER, M. et DUMBSER, M. (2008). A highly accurate discontinuous Galerkin method for complex interfaces between solids and moving fluids. *Geophysics*, 73(3):23–35.
- KELLY, K., WARD, R., TREITEL, S. et ALFORD, R. (1976). Synthetic seismograms - a finite-difference approach. *Geophysics*, 41:2–27.
- KIRSCH, A. (2011). *An introduction to the mathematical theory of inverse problems*, volume 120. Springer Science & Business Media.
- KOLB, P., COLLINO, F. et LAILLY, P. (1986). Prestack inversion of 1-D medium,. *In Extended Abstracts*, volume 74, pages 498–508.
- KOMATITSCH, D. et MARTIN, R. (2007). An unsplit convolutional perfectly matched layer improved at grazing incidence for the seismic wave equation. *Geophysics*, 72(5):SM155–SM167.
- KREBS, J., ANDERSON, J., HINKLEY, D., NEELAMANI, R., LEE, S., BAUMSTEIN, A. et LACASSE, M. D. (2009). Fast full-wavefield seismic inversion using encoded sources. *Geophysics*, 74(6):WCC105–WCC116.

BIBLIOGRAPHY

- KUNERT, M., KURZMANN, A. et BOHLEN, T. (2016). Application of 2d acoustic full waveform inversion to obc-data in shallow water. *In 78th EAGE Conference and Exhibition 2016*.
- KURZMANN, A., PRZEBINDOWSKA, A., KOHN, D. et BOHLEN, T. (2013). Acoustic full waveform tomography in the presence of attenuation: a sensitivity analysis. *Geophysical Journal International*, 195(2):985–1000.
- LAILLY, P. (1984). The seismic inverse problem as a sequence of before stack migrations. *In* BEDNAR, R. et WEGLEIN, éditeurs : *Conference on Inverse Scattering, SIAM, Philadelphia*, pages 206–220. Soc. Ind. appl. Math.
- LAMBARÉ, G. (2008). Stereotomography. *Geophysics*, 73(5):VE25–VE34.
- LAMBARÉ, G., ALERINI, M., BAINA, R. et PODVIN, P. (2004). Stereotomography: a semi-automatic approach for velocity macromodel estimation. *Geophysical Prospecting*, 52:671–681.
- LAMBARÉ, G., GUILLAUME, P. et MONTEL, J. (2014). Recent advances in ray-based tomography. *In 76th EAGE Conference & Exhibition incorporating SPE EUROPEC 2014*. EAGE.
- LAMELOISE, C.-A. (2015). *Analyse de vitesse par migration quantitative et introduction des ondes transmises*. Thèse de doctorat, l'École nationale supérieure des mines de Paris.
- LAMELOISE, C.-A., CHAURIS, H. et NOBLE, M. (2015). Improving the gradient of the image-domain objective function using quantitative migration for a more robust migration velocity analysis. *Geophysical Prospecting*, 63:391–404.
- LE MEUR, H., VIRIEUX, J. et PODVIN, P. (1997). Seismic tomography of the gulf of Corinth: a comparison of methods. *Annali Di Geofisica*, XL(1):1–24.
- LE ROUSSEAU, J. et de HOOP, M. (2001). Modelling and imaging with the scalar generalized-screen algorithms in isotropic media. *Geophysics*, 66:1551–1568.
- LELIÈVRE, P. G., FARQUHARSON, C. G. et HURICH, C. A. (2011). Inversion of first-arrival seismic traveltimes without rays, implemented on unstructured grids. *Geophysical Journal International*, 185:749–763.
- LEUNG, S. et QIAN, J. (2006). An adjoint state method for three-dimensional transmission traveltome tomography using first-arrivals. *Communications in Mathematical Sciences*, 4(1):249–266.
- LEVANDER, A. R. (1988). Fourth-order finite-difference P-SV seismograms. *Geophysics*, 53(11):1425–1436.

- LI, Y., MÉTIVIER, L., BROSSIER, R., HAN, B. et VIRIEUX, J. (2014). 2D and 3D frequency-domain elastic wave modeling in complex media with a parallel iterative solver. *Geophysics*, page Accepted.
- LIU, Y., SYMES, W., HUANG, Y. et LI, Z. (2013). Linearized extended waveform inversion via differential semblance optimization in the depth-oriented extension. In *Expanded Abstracts*, volume 171, pages 4869–4873. Society of Exploration Geophysics.
- LIU, Y., SYMES, W. W. et LI, Z. (2014b). Inversion velocity analysis via differential semblance optimization. In *Proceedings of the 76th EAGE Conference & Exhibition*. EAGE.
- LIU, Y., SYMES, W. W., LI, Z. *et al.* (2014a). Extended reflection waveform inversion via differential semblance optimization. In *2014 SEG Annual Meeting*. SEG.
- LIU, Y., WANG, S., WANG, B., YIN, Y. et GUO, T. (2016). Simultaneous multisource full waveform inversion with total variation regularization in time domain. In *78th EAGE Conference and Exhibition 2016*.
- LUO, J. et WU, R.-S. (2015). Seismic envelope inversion: reduction of local minima and noise resistance. *Geophysical Prospecting*, 63(3):597–614.
- LUO, S. et SAVA, P. (2011). A deconvolution-based objective function for wave-equation inversion. *SEG Technical Program Expanded Abstracts*, 30(1):2788–2792.
- LUO, Y. et SCHUSTER, G. T. (1991). Wave-equation travelttime inversion. *Geophysics*, 56(5):645–653.
- MA, Y. et HALE, D. (2013). Wave-equation reflection travelttime inversion with dynamic warping and full waveform inversion. *Geophysics*, 78(6):R223–R233.
- MALINOWSKI, M. et 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.
- MALINOWSKI, M., OPERTO, S. et RIBODETTI, A. (2011). High-resolution seismic attenuation imaging from wide-aperture onshore data by visco-acoustic frequency-domain full waveform inversion. *Geophysical Journal International*, 186(3):1179–1204.
- MARFURT, K. (1984). Accuracy of finite-difference and finite-element modeling of the scalar and elastic wave equations. *Geophysics*, 49:533–549.
- MASONI, I., BROSSIER, R., BOELLE, J.-L., VIRIEUX, J. *et al.* (2014). Robust full waveform inversion of surface waves. In *2014 SEG Annual Meeting*. SEG.
- MCMECHAN, G. (1989). A review of seismic acoustic imaging by reverse time migration. *International Journal of Imaging Systems and Technology*, 1:18–21.

BIBLIOGRAPHY

- MÉTIVIER, L., BRETAUDEAU, F., BROSSIER, R., OPERTO, S. et VIRIEUX, J. (2014a). Full waveform inversion and the truncated Newton method: quantitative imaging of complex subsurface structures. *Geophysical Prospecting*, 62:1353–1375.
- MÉTIVIER, L., BROSSIER, R., MÉRIGOT, Q., OUDET, E. et VIRIEUX, J. (2015). Measuring the misfit between seismograms using an optimal transport distance: Application to full waveform inversion. *Geophysical Journal International (Submitted)*.
- MÉTIVIER, L., BROSSIER, R., OPERTO, S. et VIRIEUX, J. (2014b). Smart: Robust absorbing layer and s-waves filtering for acoustic anisotropic wave simulation. *In Expanded Abstracts, 84th Annual SEG Meeting (Denver)*.
- MÉTIVIER, L., BROSSIER, R., OPERTO, S. et VIRIEUX, J. (2015). Acoustic multi-parameter FWI for the reconstruction of P-wave velocity, density and attenuation: preconditioned truncated newton approach. *In Expanded Abstracts, 85th Annual Meeting*, pages 1198–1203. SEG.
- MÉTIVIER, L., BROSSIER, R., VIRIEUX, J. et OPERTO, S. (2013). Full Waveform Inversion and the truncated Newton method. *SIAM Journal On Scientific Computing*, 35(2):B401–B437.
- MILLER, D., ORISTAGLIO, M. et BEYLKIN, G. (1987). A new slant on seismic imaging: Migration and integral geometry. *Geophysics*, 52(7):943–964.
- MIN, D.-J., SHIN, C., PRATT, R. G. et YOO, H. S. (2003). Weighted-averaging finite-element method for 2D elastic wave equations in the frequency domain. *Bulletin of the Seismological Society of America*, 93(2):904–921.
- MOCZO, P. et KRISTEK, J. (2005). On the rheological models used for time-domain methods of seismic wave propagation. *Geophysical Research Letters*, 32(1).
- MONTELLI, R., NOLET, G., DAHLEN, F. A., MASTERS, G., ENGDahl, E. R. et HUNG, S. H. (2004a). Finite-frequency tomography reveals a variety of plumes in the mantle. *Science*, 303:338–343.
- MONTELLI, R., NOLET, G., MASTERS, G., DAHLEN, F. A. et HUNG, S. H. (2004b). P and PP global travel time tomography. *Geophysical Journal International*, 158:637–654.
- MORA, P. R. (1988). Elastic wavefield inversion of reflection and transmission data. *Geophysics*, 53:750–759.
- MORA, P. R. (1989). Inversion = migration + tomography. *Geophysics*, 54(12):1575–1586.
- MULDER, W. (2014). Subsurface offset behaviour in velocity analysis with extended reflectivity images. *Geophysical Prospecting*, 62(1):17–33.

- MULDER, W. et PLESSIX, R. E. (2008). Exploring some issues in acoustic full waveform inversion. *Geophysical Prospecting*, 56(6):827–841.
- MULDER, W. et TEN KROODE, A. (2002). Automatic velocity analysis by differential semblance optimization. *Geophysics*, 67(4):1184–1191.
- NAG, S., ALERINI, M. et URSIN, B. (2010). PP/PS anisotropic stereotomography. *Geophysical Journal International*, 181(1):427–452.
- NEMETH, T., WU, C. et SCHUSTER, G. T. (1999). Least-squares migration of incomplete reflection data. *Geophysics*, 64(1):208–221.
- NEVES, F. A. et SINGH, S. C. (1996). Sensitivity study of seismic reflection/refraction data. *Geophysical Journal International*, 126(2):470–476.
- NIHEI, K. T. et LI, X. (2007). Frequency response modelling of seismic waves using finite difference time domain with phase sensitive detection (TD-PSD). *Geophysical Journal International*, 169:1069–1078.
- NOCEDAL, J. (1980). Updating Quasi-Newton Matrices With Limited Storage. *Mathematics of Computation*, 35(151):773–782.
- OPERTO, S., BROSSIER, R., GHOLAMI, Y., MÉTIVIER, L., PRIEUX, V., RIBODETTI, A. et VIRIEUX, J. (2013). A guided tour of multiparameter full waveform inversion for multicomponent data: from theory to practice. *The Leading Edge*, Special section Full Waveform Inversion(September):1040–1054.
- OPERTO, S., MINIUSI, A., BROSSIER, R., COMBE, L., MÉTIVIER, L., MONTEILLER, V., RIBODETTI, A. et VIRIEUX, J. (2015). Efficient 3-D frequency-domain monoparameter full-waveform inversion of ocean-bottom cable data: application to Valhall in the visco-acoustic vertical transverse isotropic approximation. *Geophysical Journal International*, 202(2):1362–1391.
- OPERTO, S., VIRIEUX, J., DESSA, J. X. et PASCAL, G. (2006). Crustal imaging from multifold ocean bottom seismometers data by frequency-domain full-waveform tomography: application to the eastern Nankai trough. *Journal of Geophysical Research*, 111(B09306):doi:10.1029/2005JB003835.
- PAJOT, B., LI, Y., BERTHOUMIEUX, V., WEISBECKER, C., BROSSIER, R., MÉTIVIER, L., THIERRY, P., OPERTO, S. et VIRIEUX, J. (2014). A review of recent forward problem developments used for frequency-domain fwi. In *Expanded Abstracts, 76th Annual EAGE Meeting (Amsterdam)*.
- PÉREZ SOLANO, C., DONNO, D. et CHAURIS, H. (2014). Alternative waveform inversion for surface wave analysis in 2-d media. *Geophysical Journal international*, 198:1359–1372.

BIBLIOGRAPHY

- PICA, A., DIET, J. P. et TARANTOLA, A. (1990). Nonlinear inversion of seismic reflection data in laterally invariant medium. *Geophysics*, 55(3):284–292.
- PLESSIX, R. (2016). Visco-acoustic full waveform inversion. *In Expanded Abstracts, 78th Annual EAGE Meeting (Vienna)*.
- PLESSIX, R.-E. (2006a). Estimation of velocity and attenuation coefficient maps from crosswell seismic data. *Geophysics*, 71(6):S235–S240.
- PLESSIX, R. E. (2006b). 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:WCC149–WCC157.
- PLESSIX, R. E. (2012). Waveform inversion overview: Where are we? and what are the challenges? *In Expanded Abstracts*. EAGE.
- PLESSIX, R. E. (2013). A pseudo-time formulation for acoustic full waveform inversion. *Geophysical Journal International*, 192:613–630.
- PLESSIX, R.-E., BAETEN, G., de MAAG, J. W. et 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. et CAO, Q. (2011). A parametrization study for surface seismic full waveform inversion in an acoustic vertical transversely isotropic medium. *Geophysical Journal International*, 185:539–556.
- PLESSIX, R. E., CHAVENT, G. et ROECK, Y.-H. D. (1999). Waveform inversion of reflection seismic data for kinematic parameters by local inversion. *SIAM Journal of Scientific Computing*, 20:1033–1052.
- PLESSIX, R.-E., MULDER, W. et TEN KROODE, A. (2000). Automatic cross-well tomography by semblance and differential semblance optimization: theory and gradient computation. *Geophysical Prospecting*, 48(5):913–935.
- PLESSIX, R.-E. et PÉREZ SOLANO, C. (2015). Modified surface boundary conditions for elastic waveform inversion of low-frequency wide-angle active land seismic data. *Geophysical Journal International*, 201:1324–1334.
- PLESSIX, R. E., STOPIN, A., MILCIK, P. et MATSON, K. (2014). Acoustic and anisotropic multiparameter seismic full waveform inversion case studies. *In Expanded Abstracts, 84th Annual SEG Meeting (Denver)*.
- POPOVICI, M. et SETHIAN, J. (1998). Three dimensional traveltimes using the fast marching method. *In Expanded Abstracts*, pages 1–22. Eur. Ass. Expl. Geophys.

- PRATT, R. G. (1999). Seismic waveform inversion in the frequency domain, part I : theory and verification in a physical scale model. *Geophysics*, 64:888–901.
- PRATT, R. G., SHIN, C. et 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. et 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., SONG, Z. M., WILLIAMSON, P. R. et WARNER, M. (1996). Two-dimensional velocity models from wide-angle seismic data by wavefield inversion. *Geophysical Journal International*, 124:323–340.
- PRATT, R. G. et WORTHINGTON, M. H. (1990). Inverse theory applied to multi-source cross-hole tomography. Part I: acoustic wave-equation method. *Geophysical Prospecting*, 38:287–310.
- PRIEUX, V., BROSSIER, R., GHOLAMI, Y., OPERTO, S., VIRIEUX, J., BARKVED, O. et KOMMEDAL, J. (2011). On the footprint of anisotropy on isotropic full waveform inversion: the Valhall case study. *Geophysical Journal International*, 187:1495–1515.
- PRIEUX, V., BROSSIER, R., OPERTO, S. et VIRIEUX, J. (2013a). Multiparameter full waveform inversion of multicomponent OBC data from valhall. Part 1: imaging compressional wavespeed, density and attenuation. *Geophysical Journal International*, 194(3):1640–1664.
- PRIEUX, V., BROSSIER, R., OPERTO, S. et VIRIEUX, J. (2013b). Multiparameter full waveform inversion of multicomponent OBC data from valhall. Part 2: imaging compressional and shear-wave velocities. *Geophysical Journal International*, 194(3):1665–1681.
- PRIEUX, V., LAMBARÉ, G., OPERTO, S. et VIRIEUX, J. (2013c). Building starting model for full waveform inversion from wide-aperture data by stereotomography. *Geophysical Prospecting*, 61(Issue supplement: 60 year anniversary issue):109–137.
- PYUN, S., SHIN, C., LEE, H. et YANG, D. (2008). 3D elastic full waveform inversion in the Laplace domain. *SEG Technical Program Expanded Abstracts*, 27(1):1976–1980.
- PYUN, S., SON, W. et SHIN, C. (2011). 3d acoustic waveform inversion in the laplace domain using an iterative solver. *Geophysical Prospecting*, 59:386–399.
- RAVAUT, C., OPERTO, S., IMPROTA, L., VIRIEUX, J., HERRERO, A. et 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.

BIBLIOGRAPHY

- RIBODETTI, A. et HANYGA, A. (2004). Inversion in viscoelastic power-law media. some effects of the memory kernel singularity on wave propagation and inversion in viscoelastic media, II: inversion. *Geophysical Journal International*, 158:426–442.
- RICKETT, J. et SAVA, P. (2002). Offset and angle-domain common image-point gathers for shot-profile migration. *Geophysics*, 67(3):883–889.
- RISTOW, D. et RUHL, T. (1994). Fourier finite difference migration. *Geophysics*, 59:1882–1893.
- ROBERTSSON, J., BLANCH, J. et SYMES, W. (1994). Viscoelastic finite-difference modeling. *Geophysics*, 59:1444–1456.
- SANTOS, H., SCHLEICHER, J., NOVAIS, A., KURZMANN, A. et BOHLEN, T. (2016). Robust time-domain migration velocity analysis methods for initial-model building in a full waveform tomography workflow. In *Proceedings of the 78th EAGE Conference and Exhibition 2016*. EAGE.
- SAVA, P. et BIONDI, B. (2004). Wave-equation migration velocity analysis. i. theory. *Geophysical Prospecting*, 52(6):593–606.
- SAVA, P. et FOMEL, S. (2006). Time-shift imaging condition in seismic migration. *Geophysics*, 71(6):S209–S217.
- SAVA, P. et HILL, S. J. (2009). Overview and classification of wavefield seismic imaging methods. *The Leading Edge*, 28(2):170–183.
- SAVA, P. et VASCONCELOS, I. (2011). Extended imaging conditions for wave-equation migration. *Geophysical Prospecting*, 59(1):35–55.
- SAVA, P. C. et FOMEL, S. (2003). Angle-domain common-image gathers by wavefield continuation methods. *Geophysics*, 68(3):1065–1074.
- SCHNEIDER, W. A. (1978). Integral formulation for migration in two and three dimensions. *Geophysics*, 43:49–76.
- SEARS, T., SINGH, S. et 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. et 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.
- SEI, A. et SYMES, W. W. (1994). Gradient calculation of the travelttime cost function without ray tracing. In *SEG Technical Program Expanded Abstracts*, pages 1351–1354. SEG.
- SERIANI, G. et PRIOLO, E. (1994). Spectral element method for acoustic wave simulation in heterogeneous media. *Finite elements in analysis and design*, 16:337–348.

- SHEN, H., MOTH, S., ALBERTIN, U. *et al.* (2011). Improving subsalt imaging with illumination-based weighting of rtm 3d angle gathers. *In SEG Technical Program Expanded Abstracts, 81th Annual Meeting*. SEG.
- SHEN, P. (2004). *Wave equation migration velocity analysis by differential semblance optimization*. Thèse de doctorat, Rice University.
- SHEN, P. et SYMES, W. W. (2008). Automatic velocity analysis via shot profile migration. *Geophysics*, 73(5):VE49–VE59.
- SHEN, X. (2014). *Early-arrival waveform inversion for near-surface velocity estimation*. Thèse de doctorat, Stanford University.
- SHENG, J., LEEDS, A., BUDDENSIEK, M. et SCHUSTER, G. T. (2006). Early arrival waveform tomography on near-surface refraction data. *Geophysics*, 71(4):U47–U57.
- SHIN, C. et CHA, Y. H. (2008). Waveform inversion in the Laplace domain. *Geophysical Journal International*, 173(3):922–931.
- SHIN, C. et CHA, Y. H. (2009). Waveform inversion in the Laplace-Fourier domain. *Geophysical Journal International*, 177:1067–1079.
- SHIPP, R. M. et SINGH, S. C. (2002). Two-dimensional full wavefield inversion of wide-aperture marine seismic streamer data. *Geophysical Journal International*, 151:325–344.
- SILVESTROV, I., BAINA, R. et LANDA, E. (2016). Poststack diffraction imaging using reverse-time migration. *Geophysical Prospecting*, 64(1):129–142.
- SIRGUE, L. (2006). The importance of low frequency and large offset in waveform inversion. *In 68th EAGE Conference & Exhibition, Vienna, EAGE*, page A037.
- SIRGUE, L., BARKVED, O. I., DELLINGER, J., ETGEN, J., ALBERTIN, U. et KOMMEDAL, J. H. (2010). Full waveform inversion: the next leap forward in imaging at Valhall. *First Break*, 28:65–70.
- SIRGUE, L., ETGEN, J. et ALBERTIN, U. (2007). 3D full-waveform inversion: Wide-versus narrow-azimuth acquisitions. *SEG Technical Program Expanded Abstracts*, 26(1):1760–1764.
- SIRGUE, L., ETGEN, J. T. et ALBERTIN, U. (2008). 3D Frequency Domain Waveform Inversion using Time Domain Finite Difference Methods. *In Proceedings 70th EAGE, Conference and Exhibition, Roma, Italy*, page F022.
- SIRGUE, L. et PRATT, R. G. (2004). Efficient waveform inversion and imaging : a strategy for selecting temporal frequencies. *Geophysics*, 69(1):231–248.

BIBLIOGRAPHY

- SNIEDER, R., XIE, M., PICA, A. et TARANTOLA, A. (1989). Retrieving both the impedance contrast and background velocity : a global strategy for the seismic reflection problem. *Geophysics*, 54:991–1000.
- SOUBARAS, R. et GRATACOS, B. (2007). Velocity model building by semblance maximization of modulated-shot gathers. *Geophysics*, 72(5):U67–U73.
- STAAL, X. (2015). *Combined imaging and velocity estimation by Joint Migration Inversion*. Thèse de doctorat, TU Delft, Delft University of Technology.
- STOLK, C. C. et DE HOOP, M. V. (2005). Modeling of seismic data in the downward continuation approach. *SIAM journal on applied mathematics*, 65(4):1388–1406.
- STOPIN, A., PLESSIX, R.-E. et AL ABRI, S. (2014). Multiparameter waveform inversion of a large wide-azimuth low-frequency land data set in Oman. *Geophysics*, 79(3):WA69–WA77.
- SUN, D. et SYMES, W. W. (2012). Waveform inversion via nonlinear differential semblance optimization. *SEG Technical Program Expanded Abstracts 2012*, pages 1–7.
- SWORD, C. H. (1987). *Tomographic determination of interval velocities from reflection seismic data: The method of controlled directional reception*. Thèse de doctorat, Stanford University.
- SYMES, W. (1999). All stationary points of differential semblance are asymptotic global minimizers: Layered acoustics. Rapport technique, Stanford University.
- SYMES, W. W. (2008). Migration velocity analysis and waveform inversion. *Geophysical Prospecting*, 56:765–790.
- SYMES, W. W. et CARAZZONE, J. J. (1991). Velocity inversion by differential semblance optimization. *Geophysics*, 56:654–663.
- TAILLANDIER, C., NOBLE, M., CHAURIS, H. et CALANDRA, H. (2009). First-arrival travel time tomography based on the adjoint state method. *Geophysics*, 74(6):WCB1–WCB10.
- TANG, Y., LEE, S., BAUMSTEIN, A., HINKLEY, D. *et al.* (2013). Tomographically enhanced full wavefield inversion. *In 2013 SEG Annual Meeting*. SEG.
- TARANTOLA, A. (1984). Linearized inversion of seismic reflection data. *Geophysical Prospecting*, 32:998–1015.
- TARANTOLA, A. (1986). A strategy for non linear inversion of seismic reflection data. *Geophysics*, 51(10):1893–1903.
- TARANTOLA, A. (2005). *Inverse Problem theory and methods for model parameter estimation*. Society for Industrial and Applied Mathematics, Philadelphia.

- TCHEVERDA, V., CHAVENT, G. et GADYLSHIN, K. (2016). Macrovelocity reconstruction by reflection fwi. *In 78th EAGE Conference and Exhibition 2016*. EAGE.
- TIKHONOV, A. et ARSENIN, V. (1977). *Solution of ill-posed problems*. Winston, Washington, DC.
- TOOMEY, D. R., SOLOMON, S. C. et PURDY, G. M. (1994). Tomographic imaging of the shallow crustal structure of the east pacific rise at 9deg.30'N. *Journal of Geophysical Research*, 99:24135–24157.
- TREISTER, E. et HABER, E. (2016). Joint full waveform inversion and travel time tomography. *In 78th EAGE Conference & Exhibition 2016*. EAGE.
- VAN LEEUWEN, T. (2010). *Correlation-based seismic velocity inversion*. Thèse de doctorat, Technische Universiteit Delft.
- VAN LEEUWEN, T. et HERRMANN, F. (2012). Fast waveform inversion without source-encoding. *Geophysical Prospecting*, 61(s1):10–19.
- VAN LEEUWEN, T. et HERRMANN, F. J. (2013). Mitigating local minima in full-waveform inversion by expanding the search space. *Geophysical Journal International*.
- VAN LEEUWEN, T. et MULDER, W. A. (2010). A correlation-based misfit criterion for wave-equation traveltimes tomography. *Geophysical Journal International*, 182(3): 1383–1394.
- VIDALE, D. (1990). Finite-difference calculation of traveltimes in three dimensions. *Geophysics*, 55:521–526.
- VIGH, D., JIAO, K., CHENG, X., SUN, D., KAPOOR, J. et LEWIS, O. (2016). Efficiently using full-waveform inversion in the velocity model building flow. *In 78th EAGE Conference & Exhibition 2016*. EAGE.
- VIGH, D., JIAO, K., WATTS, D. et SUN, D. (2014). Elastic full-waveform inversion application using multicomponent measurements of seismic data collection. *Geophysics*, 79(2):R63–R77.
- VIGH, D., LEWIS, W., PAREKH, C., JIAO, K., KAPOOR, J. *et al.* (2015). Introducing well constraints in full waveform inversion and its applications in time-lapse seismic measurements. *In 2015 SEG Annual Meeting*. SEG.
- VIGH, D., MOLDOVEANU, N., JIAO, K., HUANG, W. et KAPOOR, J. (2013). Ultralong-offset data acquisition can complement full-waveform inversion and lead to improved subsalt imaging. *The Leading Edge*, 32(9):1116–1122.
- VIGH, D., STARR, B., KAPOOR, J. et LI, H. (2010). 3d full waveform inversion on a gulf of mexico waz data set. *SEG Technical Program Expanded Abstracts*, 29(1):957–961.

BIBLIOGRAPHY

- VIGH, D. et STARR, E. W. (2008). 3D prestack plane-wave, full waveform inversion. *Geophysics*, 73:VE135–VE144.
- VIRIEUX, J. (1984). SH wave propagation in heterogeneous media: Velocity-stress finite difference method. *Geophysics*, 49:1259–1266.
- VIRIEUX, J. et FARRA, V. (1991). Ray tracing in 3D complex isotropic media: an analysis of the problem. *Geophysics*, 16:2057–2069.
- VIRIEUX, J. et LAMBARÉ, G. (2007). Theory and observations - body waves: ray methods and finite frequency effects. In ROMANOWICZ, B. et DIEWONSKI, A., éditeurs : *Treatise of Geophysics, volume 1: Seismology and structure of the Earth*. Elsevier.
- VIRIEUX, J. et OPERTO, S. (2009). An overview of full waveform inversion in exploration geophysics. *Geophysics*, 74(6):WCC1–WCC26.
- VIRIEUX, J., OPERTO, S., ASNAASHARI, A., BROSSIER, R., CASTELLANOS, C., ETIENNE, V., GHOLAMI, Y., HU, G., PAGEOT, D., PRIEUX, V., RIBODETTI, A. et ROQUES, A. (2011). A multi-scale strategy for handling broadband seismic data. *SEG Technical Program Expanded Abstracts*, 30(1):4360–4373.
- WANG, F. (2015). *Waveform inversion based on wavefield decomposition*. Thèse de doctorat, l'École nationale supérieure des mines de Paris.
- WANG, F., CHAURIS, H., DONNO, D. et CALANDRA, H. (2013a). Taking advantage of wave field decomposition in full waveform inversion. In *EAGE Technical Program Expanded Abstracts 2013*, page Tu0708.
- WANG, H., SINGH, S., AUDEBERT, F. et CALANDRA, H. (2015). Inversion of seismic refraction and reflection data for building long-wavelengths velocity models. *Geophysics*, 80(2):R81–R93.
- WANG, S., CHEN, F., ZHANG, H. et SHEN, Y. (2013b). Reflection-based full waveform inversion (rfwi) in the frequency domain. In *SEG Technical Program Expanded Abstracts*, pages 877–881. SEG.
- WANG, S., CHEN, F., ZHANG, H. et SHEN, Y. (2013c). Reflection-based full waveform inversion (RFWI) in the frequency domain. In *Expanded Abstracts*, volume 171, pages 877–881. Society of Exploration Geophysics.
- WANG, Y. et PRATT, G. (1997). Sensitivities of seismic traveltimes and amplitudes in reflection tomography. *Geophysical Journal International*, 131:618–642.
- WAPENAAR, C. (1996). Inversion versus migration: A new perspective to an old discussion. *Geophysics*, 61(3):804–814.
- WARNER, M. et GUASCH, L. (2014). Adaptive waveform inversion - FWI without cycle skipping - theory. In *76th EAGE Conference and Exhibition 2014*, page We E106 13.

- WARNER, M. et GUASCH, L. (2015). Robust adaptive waveform inversion. *In SEG Technical Program Expanded Abstracts 2015*, pages 1059–1063.
- WARNER, M., RATCLIFFE, A., NANGOO, T., MORGAN, J., UMPLEBY, A., SHAH, N., VINJE, V., STEKL, I., GUASCH, L., WIN, C., CONROY, G. et BERTRAND, A. (2013). Anisotropic 3D full-waveform inversion. *Geophysics*, 78(2):R59–R80.
- WARNER, M., STEKL, I. et UMPLEBY, A. (2008). 3D wavefield tomography: synthetic and field data examples. *SEG Technical Program Expanded Abstracts*, 27(1):3330–3334.
- WELLINGTON, P., BROSSIER, R., GARAMBOIS, S. et VIRIEUX, J. (2015). 2d acoustic time domain full waveform inversion - a broadband application in the carnarvon basin, australia. *In 77th EAGE Conference and Exhibition*, page We P1 16. EAGE.
- WHITING, P. (1998). Reflection tomography without picking. *In Expanded Abstracts*, pages 1226–1229. Society of Exploration Geophysics.
- WOODWARD, M. J. (1992). Wave-equation tomography. *Geophysics*, 57:15–26.
- WOODWARD, M. J., NICHOLS, D., ZDRAVEVA, O., WHITFIELD, P. et JOHNS, T. (2008). A decade of tomography. *Geophysics*, 73(5):VE5–VE11.
- WU, R.-S. (1994). Wide-angle elastic one-way propagation in heterogeneous media and an elastic wave complex-screen method. *Journal of Geophysical Research*, 99:751–766.
- WU, R. S. et TOKSÖZ, M. N. (1987). Diffraction tomography and multisource holography applied to seismic imaging. *Geophysics*, 52:11–25.
- WU, Z. et ALKHALIFAH, T. (2014). Full waveform inversion based on the optimized gradient and its spectral implementation. *In Expanded Abstracts*, page Tu P01 11. EAGE.
- WU, Z. et ALKHALIFAH, T. (2015). Simultaneous inversion of the background velocity and the perturbation in full-waveform inversion. *Geophysics*, 80(6):R317–R329.
- WU, Z. et ALKHALIFAH, T. (2016). The optimized gradient method for full waveform inversion and its spectral implementation. *Geophysical Journal International*, 205(3): 1823–1831.
- XU, S., CHAURIS, H., LAMBARÉ, G. et NOBLE, M. (2001). Common angle image gather : a strategy for imaging complex media. *Geophysics*, 66(6):1877–1894.
- XU, S., WANG, D., CHEN, F., LAMBARÉ, G. et ZHANG, Y. (2012). Inversion on reflected seismic wave. *SEG Technical Program Expanded Abstracts 2012*, pages 1–7.
- YANG, P., BROSSIER, R., MÉTIVIER, L. et VIRIEUX, J. (2016). A review on the systematic formulation of 3D multiparameter full waveform inversion in viscoelastic medium. *Geophysical Journal International*, 207(1):129–149.

BIBLIOGRAPHY

- YANG, T. et SAVA, P. (2011). Wave-equation migration velocity analysis with time-shift imaging. *Geophysical Prospecting*, 59(4):635–650.
- YANG, T., SHRAGGE, J. et SAVA, P. (2013). Illumination compensation for image-domain wavefield tomography. *Geophysics*, 78(5):U65–U76.
- YILMAZ, O. (2005). Earthquake seismology, exploration seismology and engineering seismology: How sweet it is—listening to the earth. *In 4th Congress of the Balkan Geophysical Society*.
- ZELT, C. et BARTON, P. J. (1998). Three-dimensional seismic refraction tomography: a comparison of two methods applied to data from the Faeroe basin. *Journal of Geophysical Research*, 103(B4):7187–7210.
- ZELT, C. et SMITH, R. B. (1992). Seismic travelt ime inversion for 2-D crustal velocity structure. *Geophysical Journal International*, 108:16–34.
- ZHANG, J., CHEN, J. *et al.* (2014). Joint seismic travelt ime and waveform inversion for near surface imaging. *In 2014 SEG Annual Meeting*. SEG.
- ZHAO, H. (2004). A fast sweeping method for eikonal equations. *Mathematics of computation*, 74(250):603–627.
- ZHOU, H., AMUNDSEN, L. et ZHANG, G. (2012). Fundamental issues in full waveform inversion. *In SEG Technical Program Expanded Abstracts*, pages 1–5. SEG.
- ZHOU, W., BROSSIER, R., OPERTO, S. et VIRIEUX, J. (2015). Full waveform inversion of diving & reflected waves for velocity model building with impedance inversion based on scale separation. *Geophysical Journal International*, 202(3):1535–1554.

Postscript

I started my PhD thesis in April of 2013. After the EAGE annual meeting of that year, I told one of my supervisor, Romain Brossier, that I would like to attend EAGE the next year. He was astonished, because at that time I even had not touched RWI yet, but if I want to go I should submit an abstract to EAGE. The deadline was some day at the beginning of February of 2014, giving me only about three months to make some new work. So, I kept interrupting Romain in the first week, and gradually understood the concept of sensitivity kernels associated to diving waves and reflected waves, in smooth models or the ones with perturbations, by doing small tests. One day, he told me what he and Stéphane Operto, another supervisor, had “brewed” in their mind. Some other day, Stéphane returned from the 2013 SEG annual meeting and sent us a report on what he had discussed with some one during the meeting (a mystery guy...). Though many equations, I was able to understand what their motivation was behind and what the difficulty they had encountered: it is straightforward to include both diving waves and reflected waves in FWI to build velocity macromodel, but how to formulate the misfit function without generating the high-wavenumber model updates?

The answer was found out based on clearly knowing the contribution from each body wave, and how it is taken out via the Lagrangian formulation of the adjoint-state method. Staring at my drafts, I realized that every hint leads to a data separation in the misfit function: if I could separate the data, the correlation between the incident wave, generated from the seismic source, and the adjoint wave, generated from the adjoint source that uses the reflected wave residual as the wavelet function, disappears in the gradient expression! This is what we need for JFWI to suppress the migration isochrone from the gradient.

However, the separation only works for the suppression of the first-order migration isochrones. When I test with the two-layer case, the inter-bed multiples could generate migration isochrones of higher orders (more exactly to two and three orders, see Appendix B), thus brings new high-wavenumber contributions. Since the data separation between reflected waves from different layers is more difficult and even impossible (though the high-order isochrones can be suppressed if we could do so), Romain told me to use velocity–impedance parameterization based on former studies of SEISCOPE (Yaser Gholami, Vincent Prieux etc). This problem was then nicely solved.

I was happy to attend EAGE in 2014 with my new approach and some results from the synthetic Valhall application. At that time, the cycle skipping problem coming from

the multiscattered waves generated by adjacent sharp interfaces could not be solved, and I was only able to reconstruct shallow regions including gas layers at 1.5 to 2.5 km depths. Nonetheless, as it is said *Rome was not built in a day*, I was satisfied with my preliminary results. What was really a challenge for me was to present something in front of many high-level expertises (especially they are in dark), for the first time! Needless to say, I was nervous. However, as having practiced many times I was still able to speak fluently without organizing the sentences. Coincidentally, Jean was the chairman of that FWI session and he helped me finish it smoothly. So finally, it was good.

The deeper part of the synthetic Valhall model was recovered after I managed to handle the multiscattered waves, simply by ignorance their contribution in the early stage and re-take them at the end. With the improved results I did a poster presentation in 2014's SEG meeting and published a paper on *Geophysical Journal International*. My writing skills, although still defective now, had improved a lot many thanks to my supervisors' instructions, especially thanks to Stéphane who left many comments in my latex file such that I could compare them with mine.

In that year, there was a release of benchmark test provided by Chevron to hold a "competition" among worldwide FWI teams. The winner was CGG, although they actually did not perform FWI. The data were very suitable to JFWI: abundant reflection phases which were well separable from direct waves. The result showed a high-velocity body on the left side in the shallow zone. However, Romain did not believe it, because his multiscale FWI did not produce the same thing. It was funny that resolving this disagreement had to wait for the "magical" moment when the answer was revealed. It turned out that there was no such high-velocity body, which I think would be a translation of the diving wave-associated kernel due to insufficient offset range. However, this data were not touched again after the SEG meeting, and I directly moved forward to the real case study following the established research plan (three years are short!).

The real data application was not as successful as the synthetic study. Many problems arose that I was not aware of at the beginning. Before Andrzej Górszczyk visited Grenoble in the winter of 2015, the result was a mess. He told me how to whiten the data spectrum to reduce potential cycle skipping due to ringing effect after low-pass filtering. Then, following Stéphane and Romain's suggestion, I tested the data by implementing classical FWI respectively in the time and frequency domains, and quickly identified the problem of attenuation (JFWI is implemented in the time domain without considering attenuation for the sake of computational cost). This prompted the discussion with Peng-Liang Yang, a postdoc of SEISCOPE working on cheap implementations of viscous FWI in the time domain. With his help, I could find a band of lower frequencies that is safe for me to perform seismic modeling without considering attenuation. (Therefore, a perspective of JFWI could be to use higher frequencies when a specific modeling tool is available.)

After one year's effort (or say struggle), the finalization of the real case study was set to the deadline of SEG submission, 1st of April, 2016. However, which message to deliver to the audiences was not clear for me. After several telephone discussions with Stéphane, I learned that I should report the promises and limitations of JFWI, rather

than good results that show JFWI was working. (In fact, it must not work due to many reasons!) This change of mind led me to point out some critical issues in JFWI.

I have learned many during the four-year stay in Grenoble, and benefited a lot from my supervisors and colleagues to reach the finishing line. At the end, I want to share a Chinese proverb with the reader and all the people I work with: ⁶

書山有路勤為徑
學海無涯苦做舟

⁶“Diligence is the path to the mountain of knowledge; hard-work is the boat to the endless sea of learning.”

Appendix A

Derivation of Gradients through Lagrangian Formulation

Although we perform full waveform inversion in the time domain, we shall derive the gradient formulations in the frequency domain for the sake of compactness. The Lagrangian quantities of FWI, RWI, FWI with reflectivity (denoted by FWI+reflectivity in Figure 3.2(d), here FWI2 for short), and JFWI can be expressed with common notations (see Plessix (2006b) for a review of adjoint-state method):

$$\mathcal{L} = \text{Misfit function} + \Re\langle a | \text{First constraint} \rangle + \Re\langle b | \text{Second constraint} \rangle, \quad (\text{A.1})$$

where a and b denote the adjoint-state variables and $\Re\langle \cdot | \cdot \rangle$ denotes the real part of the inner product. In particular,

	Misfit function	State variable(s)
FWI	$0.5 \ W(d_{obs} - Ru) \ ^2$	u_0
RWI	$0.5 \ W^r(d_{obs}^{refl} - R\delta u) \ ^2$	u_0 and δu
FWI2	$0.5 \ W^d(d_{obs}^{div} - Ru_0) + W^r(d_{obs}^{refl} - R\delta u) \ ^2$	u_0 and δu
JFWI	$0.5 \ W^d(d_{obs}^{div} - Ru_0) \ ^2 + 0.5 \ W^r(d_{obs}^{refl} - R\delta u) \ ^2$	u_0 and δu

Continued:

	First constraint	Second constraint
FWI	$B(m)u_0 - s$	–
RWI	$B(m_0)u_0 - s$	$B(m_0 + \delta m)\delta u + \Delta Bu_0$
FWI2	$B(m_0)u_0 - s$	$B(m_0 + \delta m)\delta u + \Delta Bu_0$
JFWI	$B(m_0)u_0 - s$	$B(m_0 + \delta m)\delta u + \Delta Bu_0$

where the physical meaning of u_0 and δu are the background and scattered components of the modelled wavefield, respectively. Starting from the smooth initial model, the

scattered field is not generated during the first iteration of FWI, therefore $\delta u = 0$ and only the first constraint is required. The operator B denotes the modeling operator, s the source term and R the real-valued sampling operator that extracts the calculated data from the modelled field.

The expression of the first constraint is derived from the forward problem equation (i.e. $B(m)u = s$) in m or m_0 to constrain u_0 . Similarly, the expression of the second constraint is from the equation $B(m + \Delta m)(u + \delta u) = s = B(m)u$ to constrain δu . The source term $-\Delta B u_0$ with $\Delta B = B(m_0 + \delta m) - B(m_0)$ emits the scattered field when the background field u_0 hits local diffractors. Since the modeling operator $B(m_0 + \delta m)$ depends on δm , high-order scattering effects are accounted for in δu .

Setting the derivatives of Equation (A.1) with respect to the state variables to zero gives the adjoint-state equations

	$\partial_{\delta u} \mathcal{L} = 0$ gives	$\partial_{u_0} \mathcal{L} = 0$ gives
FWI	–	$B(m)^\dagger a = -R^\top W^\top W \Delta d^*$
RWI	$B(m_0 + \delta m)^\dagger b =$ $-R^\top W^{r\top} W^r \Delta d^{refl*}$	$B(m_0)^\dagger a = -\Delta B^\dagger b$
FWI2	$B(m_0 + \delta m)^\dagger b =$ $-R^\top W^{r\top} (W^d \Delta d^{div*} + W^r \Delta d^{refl*})$	$B(m_0)^\dagger a = -\Delta B^\dagger b$ $-R^\top W^{d\top} (W^d \Delta d^{div*} + W^r \Delta d^{refl*})$
JFWI	$B(m_0 + \delta m)^\dagger b =$ $-R^\top W^{r\top} W^r \Delta d^{refl*}$	$B(m_0)^\dagger a =$ $-\Delta B^\dagger b - R^\top W^{d\top} W^d \Delta d^{div*}$

where Δd^* denotes the conjugate of the data residual, \top the transpose operation and \dagger the adjoint operation. In the first column, Δd^* represents the data residuals at the receiver positions, reversed in time, augmented with zeroes in the subsurface model by the prolongation operator R^\top , and used as virtual sources to produce the adjoint field b . In the second column, the other adjoint quantity a is emitted by the residual source at receiver positions or by the secondary sources located at diffractor positions δm . The adjoint propagator B^\dagger indicates that the computation of the adjoint quantities can be implemented by modifying the forward modeling code without much effort, and the cost to compute one adjoint quantity is the same as the cost of one forward modeling computation.

Analysis of these equations allows us to give a more physical interpretation of the adjoint fields a and b :

	b	a
FWI	–	$W^2(\lambda_0^d + \lambda_0^r)$
RWI	$W^{r2}(\lambda_0^r + \delta\lambda^r)$	$W^{r2}\delta\lambda^r$
FWI2	$W^{rd}(\lambda_0^d + \delta\lambda^d)$ $+W^{r2}(\lambda_0^r + \delta\lambda^r)$	$W^{d2}\lambda_0^d + W^{rd}\delta\lambda^d$ $+W^{dr}\lambda_0^r + W^{r2}\delta\lambda^r$
JFWI	$W^{r2}(\lambda_0^r + \delta\lambda^r)$	$W^{d2}\lambda_0^d + W^{r2}\delta\lambda^r$

where λ_0^d denotes the background adjoint field from the diving wave residuals, $\delta\lambda^d$ the scattered adjoint field from the diving wave residuals, λ_0^r the background adjoint field

from the reflected wave residuals, and $\delta\lambda^r$ the scattered adjoint field from the reflected wave residuals. W^2 is the short-hand for $W^T W$, and similarly W^{d2} for $W^{dT} W^d$, W^{r2} for $W^{rT} W^r$, W^{dr} for $W^{dT} W^r$ and W^{rd} for $W^{rT} W^d$. Instead of computing the adjoint quantities a and b , we actually compute the adjoint fields λ 's because their source terms are easier to be built: the evaluation of the operator ΔB is not required and only the data residuals are computed.

In the adjoint-state method, the gradient is found by taking the derivative of Equation (A.1) with respect to m or m_0 . Inserting the respective expressions of a , b , we find that

$$\nabla C_{FWI} = (\lambda_0^d + \lambda_0^r)^\dagger W^2 \frac{\partial B(m)}{\partial m} u_0, \quad (\text{A.2})$$

$$\nabla C_{RWI} = \delta\lambda^{r\dagger} W^{r2} \frac{\partial B(m_0 + \delta m)}{\partial m_0} u_0 + \lambda_0^{r\dagger} W^{r2} \frac{\partial B(m_0 + \delta m)}{\partial m_0} \delta u \quad (\text{A.3})$$

$$+ \delta\lambda^{r\dagger} W^{r2} \frac{\partial B(m_0 + \delta m)}{\partial m_0} \delta u \quad (\text{A.4})$$

$$+ \lambda_0^{r\dagger} W^{r2} \frac{\partial \Delta B}{\partial m_0} u_0 \quad (\approx 0), \quad (\text{A.5})$$

$$\nabla C_{FWI2} = \left[\lambda_0^{d\dagger} W^{d2} + \lambda_0^{r\dagger} W^{r2} \right] \frac{\partial B(m_0)}{\partial m_0} u_0 \quad (\text{A.6})$$

$$+ \left[\delta\lambda^{d\dagger} W^{dr} + \delta\lambda^{r\dagger} W^{r2} \right] \frac{\partial B(m_0 + \delta m)}{\partial m_0} u_0 \quad (\text{A.7})$$

$$+ \left[(\lambda_0^d + \delta\lambda^d)^\dagger W^{dr} + (\lambda_0^r + \delta\lambda^r)^\dagger W^{r2} \right] \frac{\partial B(m_0 + \delta m)}{\partial m_0} \delta u \quad (\text{A.8})$$

$$+ \left[\lambda_0^{d\dagger} W^{dr} + \lambda_0^{r\dagger} W^{r2} \right] \frac{\partial \Delta B}{\partial m_0} u_0 \quad (\approx 0), \quad (\text{A.9})$$

$$\nabla C_{JFWI} = \lambda_0^{d\dagger} W^{d2} \frac{\partial B(m_0)}{\partial m_0} u_0 + \delta\lambda^{r\dagger} W^{r2} \frac{\partial B(m_0 + \delta m)}{\partial m_0} u_0 \quad (\text{A.10})$$

$$+ \lambda_0^{r\dagger} W^{r2} \frac{\partial B(m_0 + \delta m)}{\partial m_0} \delta u + \delta\lambda^{r\dagger} W^{r2} \frac{\partial B(m_0 + \delta m)}{\partial m_0} \delta u \quad (\text{A.11})$$

$$+ \lambda_0^{r\dagger} W^{r2} \frac{\partial \Delta B}{\partial m_0} u_0 \quad (\approx 0). \quad (\text{A.12})$$

Apart from the fact that they are formulated in the frequency domain, these gradients expressions are the exact forms of the compact ones provided in Equations (3.3), (3.5) and (3.7), respectively. For FWI2, the gradient expression is not given but illustrated in Figure 3.3(d).

The terms associated with $\partial \Delta B / \partial m_0$ have non-zero values only at reflector positions (where δm differs from 0). Regularization of the inverse problem helps suppress these unwanted contributions because they have high-wavenumber contents. Therefore, the last terms of RWI, FWI2 and JFWI gradients (A.5), (A.9) and (A.12) can be neglected. Nonetheless, the workflow proposed to evaluate the JFWI gradient (Algorithm

3.1) provides the exact expressions. In the frequency domain, G_1 and G_2 are given by:

$$G_1 = (\lambda_0^r + \delta\lambda^r)^\dagger W^{r2} \frac{\partial B(m_0 + \delta m)}{\partial m_0} (u_0 + \delta u), \quad (\text{A.13})$$

$$G_2 = \lambda_0^{d\dagger} W^{d2} \frac{\partial B(m_0)}{\partial m_0} u_0 - \lambda_0^{r\dagger} W^{r2} \frac{\partial B(m_0)}{\partial m_0} u_0. \quad (\text{A.14})$$

One can verify that adding G_2 to G_1 results in the same expression as (A.10) to (A.12).

Another approximation can be made on Equation (A.6) to (A.9). Terms representing interferences of $\delta\lambda^d$ and u_0 , λ_0^d and δu as well as $\delta\lambda^d$ and δu can be discarded simply because these interferences happen to high-order scattered fields, giving negligible contribution to the gradient (e.g. around 2% of the total energy in the synthetic Valhall case).

Appendix B

High-order scattering effect

Formulations of the high-order migration isochrones can be theoretically deduced from expressions (A.10) to (A.12). Taking the two-reflector case (Figure 3.4a) as an example, the scattered component of the adjoint field can be decomposed as infinite series in an increasing order:

$$\delta\lambda^{r1} = \delta\lambda_1^{r1} + \delta\lambda_2^{r1} + \delta\lambda_{1,2}^{r1} + \dots, \quad (\text{B.1})$$

$$\delta\lambda^{r2} = \delta\lambda_1^{r2} + \delta\lambda_2^{r2} + \delta\lambda_{1,2}^{r2} + \dots, \quad (\text{B.2})$$

where $r1$ and $r2$ respectively indicate the early and late reflection phases that provide the adjoint sources. Numbers in the subscript indicate the reflectors at which the scattering takes place. 1: scattering at the shallow reflector; 2: scattering at the deep reflector; 1, 2: successive scatterings at the shallow and deep reflectors. Similarly, the scattered component of the incident field can be decomposed using the expression $\delta u = \delta u_1 + \delta u_2 + \delta u_{1,2} + \dots$. Inserting these decompositions into expressions (A.10) to (A.12), omitting small-amplitude terms and switching to the time domain, the gradient turns out to be

$$\nabla C_{JFWI} = u_0 \star \lambda_0^d(A) \quad (\text{B.3})$$

$$+ u_0 \star \delta\lambda_1^{r1}(B) + \delta u_1 \star \lambda_0^{r1}(C) + \delta u_1 \star \delta\lambda_1^{r1}(D) \quad (\text{B.4})$$

$$+ u_0 \star \delta\lambda_2^{r2}(E) + u_0 \star \delta\lambda_{1,2}^{r2}(E) + \delta u_2 \star \lambda_0^{r2}(F) + \delta u_{1,2} \star \lambda_0^{r2}(F) \quad (\text{B.5})$$

$$+ u_0 \star \delta\lambda_1^{r2}(G) + \delta u_1 \star \lambda_0^{r2}(H) + \delta u_1 \star \delta\lambda_1^{r2}(I, J) \quad (\text{B.6})$$

$$+ \delta u_2 \star \delta\lambda_2^{r2}(K) + \delta u_{1,2} \star \delta\lambda_2^{r2}(K) + \delta u_2 \star \delta\lambda_{1,2}^{r2}(K) + \delta u_{1,2} \star \delta\lambda_{1,2}^{r2}(K). \quad (\text{B.7})$$

Each term represents one or two Fresnel zones labelled by the capital letters in parentheses. They are shown in Figure B1. Zone A is the classical first-Fresnel zone wavepath generated by direct/diving waves (Figure B1a). Zones B and C are the RWI wavepaths associated with the first reflection, whereas Zone D is a high wavenumber isochrone located near the shallow reflector (Figure B1b). The size of Zone D decreases as higher frequencies are used. The late reflection phase gives rise to other zones E to K (Figure B1c). Zones E and F are the RWI wavepaths associated with the second reflector. Zones I, J, G and H are higher-order migration isochrones. In migration imaging, these

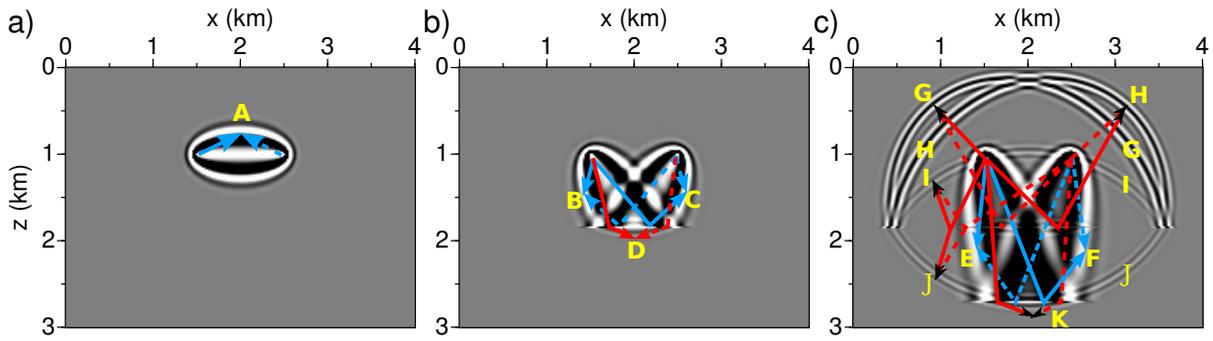


Figure B1: Decomposed JFWI gradient in the two-reflector case, components associated to (a) direct/diving wave, (b) early reflection phase denoted by r_1 , and (c) late reflection phase denoted by r_2 . Fresnel zones A–K are represented by correlation terms in Equation (B.3) to (B.7). Solid and dashed arrows denote the ray paths followed by the modelled and adjoint fields, respectively. Blue paths are useful for low-wavenumber imaging unlike the red ones.

isochrones are conventionally considered as migration artifacts, suppressed by destructive interference from different source-receiver couples. Like Zone D, Zone K is located near the deep reflector position and shrinks with higher frequencies.

