



Analyse de vitesse par migration quantitative dans les domaines images et données pour l'imagerie sismique

Yubing Li

► To cite this version:

Yubing Li. Analyse de vitesse par migration quantitative dans les domaines images et données pour l'imagerie sismique. Géophysique [physics.geo-ph]. Université Paris sciences et lettres, 2018. Français. ⟨NNT: 2018PSLEM002⟩. ⟨tel-01814262⟩

HAL Id: tel-01814262

<https://pastel.hal.science/tel-01814262v1>

Submitted on 13 Jun 2018

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.



HAL Authorization

THÈSE DE DOCTORAT

de l'Université de recherche Paris Sciences et Lettres
PSL Research University

Préparée à MINES ParisTech

**Subsurface seismic imaging based on inversion velocity
analysis in both image and data domains**

**Analyse de vitesse par migration quantitative dans les
domaines images et données pour l'imagerie sismique**

École doctorale n°398

GÉOSCIENCES, RESSOURCES NATURELLES ET ENVIRONNEMENT

Spécialité GÉOSCIENCES ET GÉOINGÉNIERIE

Soutenue par **Yubing LI**
le 16 janvier 2018

Dirigée par **Hervé CHAURIS**

COMPOSITION DU JURY :

M Jean-Pierre VILOTTE Président
IPGP

M Gilles LAMBARÉ Rapporteur
CGG

M William SYMES Rapporteur
Rice University

M Romain BROSSIER Membre du jury
Université Grenoble Alpes

Mme Céline GÉLIS Membre du jury
IRSN

M Hervé CHAURIS Membre du jury
MINES ParisTech

M François AUDEBERT Membre du jury
Total



Acknowledgments

I would like to thank all the people who have contributed to this thesis work. There is a saying: "If you want long memory, you need a short pencil." I hope that I could always remember all the great memories for this period of my life.

I would like to express my sincere gratitude to my advisor, Professor Hervé Chauris, for offering me the opportunities to do an internship and to pursue the PhD at MINES ParisTech. I thank him for his patience, support and encouragement. Hervé even would like to help me with the very details of the code and with the basic formulations. He always provided useful instructions and inspirations when I was obstructed by difficulties. He has a healthy attitude towards life and I can always feel his kind heart to me. I am very grateful to him for helping me organize the defense and revise the manuscript line-by-line. His passion for the research encourages me a lot.

I would like to thank the members of my jury. It was my great honor that William Symes and Gilles Lambaré accepted to review my thesis in detail. I thank them for the professional comments and the fruitful recommendations. My sincere thanks also goes to Jean-Pierre Vilotte for chairing the defense and to Romain Brossier, Céline Gélis and François Audbert for examining my work.

I thank Paris Exploration Geophysics Group (GPX) for financially supporting my thesis. All the research described in this thesis was carried out as a part of the GPX project funded by the French National Research Agency, CGG, Total and Schlumberger. I thank the GPX committee, especially Satish Singh, Hervé Chauris and Mark Noble, for accepting my Master application four years ago. It offered me the chance to come to France and to subsequently pursue my PhD. I thank Nobuaki Fuji for kindly helping me to get access to the S-CAPAD platform at IPGP for more computational resources.

Many thanks to the members of the geophysics team at MINES ParisTech. I especially thank Véronique Lachasse for her help to facilitate the administrative affairs since my day one in Fontainebleau. I sincerely thank the permanent researchers, Daniela Donno, Pierre Dublanchet, Alexandrine Gesret and Mark Noble, and the scientific visitors, Nidhal Belayouni and Pierre-François Roux, for their kindness. I also would like to thank all my colleagues during the three and half years, Elise Vi Nu Bas, Charles-Antoine Lameloise, Emmanuel Cocher, Sven Schilke, Yves-Marie Batany, Tiago Barros, Jihane Belhadj, Julien Cotton, Hao Jiang, Keurfon Luu, Alexandre Kazantsev,

Michelle Almakari, Tianyou Zhou, Ahmed Jabrane Maxime Godano, Shanker Krishna, Zufri and Joshua Lartey., for their friendship and support. I express my special thanks to Emmanuel for guiding me to the CROUS residence on the first day and to Hao for selecting the defense gift as the representative of others on the last day. First day, last day, everyday was a nice day.

In addition, I want to thank my lovely friends who have supported me in France. Some of them were also doing research, giving me broader views and different ideas in science. Some of them had different paths of life but we did get along with each other. I give my acknowledgement to Quan Chen, Wei Li, Shuqiao Zou, Fuchen Liu, Shizhu Zheng, Dandan Niu, Shihao Yuan, Yanfang Qin, Weiguang He, Haiyang Wang, Xin Gao, Hanchao Jian, Jialan Wang, Zhengbin Deng, Dongyang Huang, Miao Yu, Hanjun Yin, Shaozhuo Liu, Xiangdong Song, Zezhong Zhang, Shuaitao Wang, Tianyuan Wang, Shun Huang, Mingguang Yang, Feng Yang, Botao Qin, Kurama Ohkubo, Eke-abino Momoh, Pierre Romanet, Mohammad Shahrukh, Fernando Villanueva-Robles, Alvina Kusumadewi and others.

I would express a deep sense of gratitude to my beloved family for their consistent support and unconditional love. My research would not have been possible without the contribution of my dearest parents, who have given me life and have always stood by me like a pillar in times of need. Special thanks are due to my one and only loving sister Ruiyun Li who stays close to our parents and keeps them feel less lonely.

Finally, I want to acknowledge my sweet wife Yanhui He who has been accompanying me during these important but happy years. It is my fortune to meet you, to love you and to have you by my side for the rest of my lifetime. Life is a long journey: you were and always will be the most impressive part of my grand tour, mon amour.

Abstract

Active seismic experiments are widely used to characterize the structure of the subsurface for the oil and gas industry. Based on the assumption of scale separation, numerous approaches split the velocity model into a smooth background model controlling the kinematics of wave propagation and a reflectivity model characterizing the rapid changes of the model parameters. Macro velocity estimation and reflectivity imaging are formulated as two inverse problems. The macro velocity estimation scheme can be derived either in the data-domain, where one seeks an optimal fit between modeled and observed data, or in the image-domain, where one tries to improve the image coherency.

Migration techniques aim at determining the reflectivity in a given macromodel. Classic migration is the adjoint operator of the forward linearized modeling and suffers from migration artifacts. Recent studies recast the asymptotic inversion in the context of reverse time migration. They define a direct way for inverting the Born modeling operator, which automatically compensates for uneven illuminations and geometrical spreading losses, removing in practice migration artifacts.

Migration Velocity Analysis (MVA) techniques assess the quality of the estimated macromodel by observing the migrated images. The analysis is carried out on the panels called common image gathers. These panels can be built in two manners: the surface-oriented methods first perform prestack migration on different subsets of input data, such as the common-shot gathers, and then collect images along the redundant parameter; the depth-oriented methods extend the image volume with an additional parameter, for example the subsurface-offset as a spatial delay, inserted during the construction of reflectivity images. Recent investigations propose to couple the direct inversion to MVA in the subsurface-offset domain, introducing better robustness. This approach is numerically demanding, even in 2D, and cannot be currently extended to 3D. In this thesis, I propose to transpose this strategy to the more conventional common-shot migration based MVA.

I first develop an alternative approach to a recent published work, related to the common-shot true-amplitude reverse time migration. It is a pseudo-inverse of the Born modeling operator in the asymptotic sense. The method allows producing prestack reflectivity images free of migration smiles in a direct way. Then, I propose to couple this operator to velocity analysis. Inversion Velocity Analysis (IVA) is thus an alternative

to MVA consisting of replacing migration by an asymptotic inverse. I analyze how the approach can deal with complex models, in particular with the presence of low velocity anomaly zones or discontinuities, better than the classic MVA. Common-shot IVA benefits from the natural parallel implementation, and requires less numerical cost than its counterpart in the subsurface-offset domain.

I also propose to extend IVA to the data-domain, leading to a more linearized inverse problem than classic full waveform inversion. It simply consists of applying the modeling operator to the images after the application of the annihilator. The new approach is close to Full Waveform Inversion, in the sense that the optimal model is obtained when the norm of the data residual is minimum. On the other hand, the new approach is still based on the coherency criteria for which the inverse problem is known to have a better convexity, at least for simple models. I compare the new approach to other reflection-based waveform inversion to establish formal links between data-fitting principle and image coherency criteria.

The methodologies are analyzed on 2D synthetic data sets from a series of velocity models, in particular models with the presence of a low-anomaly zone for which common-shot migration is not necessarily appropriate, and the Marmousi model, to justify the robustness. The main contribution of this work is (1) the development of common-shot true-amplitude reverse time migration and, more importantly, the coupling with velocity analysis; (2) the extension of common-shot IVA to the data-domain and, along this line, the analysis of the links between image-domain and data-domain methods.

Résumé

Dans le domaine de la prospection pétrolière, les expériences sismiques actives sont largement utilisées pour caractériser la structure de la subsurface. Avec l'hypothèse de la séparation des échelles, de nombreuses approches divisent le modèle de vitesse entre un modèle de grande longueur d'onde qui contrôle la cinématique de propagation des ondes, et un modèle de réflectivité qui caractérise les changements rapides. Les estimations du macro-modèle de vitesse et de la réflectivité sont formulées comme deux problèmes inverses imbriqués. La détermination du macro-modèle peut être obtenue soit dans le domaine des données, où est mesurée l'écart entre les données modélisées et les données observées, ou dans le domaine image, où l'objectif est d'avoir des images cohérentes.

Les techniques de migration visent à déterminer le modèle de réflectivité dans un macro-modèle donné. La migration classique est seulement l'adjoint de l'opérateur de modélisation linéarisé. La méthode est connue pour causer des artefacts de migration. Récemment, une formule d'inversion au sens asymptotique a été développée pour remplacer la migration. C'est une méthode directe sans itération. Elle compense pour l'illumination irrégulière, pour le facteur d'atténuation géométrique et donne des images beaucoup plus propres en pratique.

L'analyse de vitesse par migration est une technique qui juge de la qualité d'un macro-modèle de vitesse en comparant différentes images issues de sous-ensembles des données, comme par exemple un point de tir. Des panneaux sont construits en modifiant la condition d'imagerie soit avec un paramètre de surface, soit avec un paramètre lié à la profondeur, comme un délai en espace ou en temps. Des résultats récents proposent de coupler l'inversion asymptotique avec l'analyse de vitesse pour la version extension en profondeur. L'analyse de vitesse est rendue beaucoup plus robuste. Cette approche cependant demande des capacités de calcul et de mémoire importantes, même en 2D, et ne peut actuellement être étendue en 3D. Dans ce travail, je propose de développer le couplage entre l'analyse de vitesse et la migration plus conventionnelle par point de tir.

Je développe dans un premier temps une alternative à un travail récent autour de la migration quantitative par point de tir. La formule est un pseudo-inverse de l'opérateur de Born au sens asymptotique. Elle permet d'obtenir des images migrées propres sans recourir à des itérations. Ensuite, je propose de coupler cet opérateur inverse avec

l'analyse de vitesse dite par inversion et non plus par migration. La nouvelle approche permet de prendre en compte des modèles de vitesse complexes, comme par exemple en présence d'anomalies de vitesses plus lentes ou de réflectivités discontinues. C'est une alternative avantageuse en termes d'implémentation et de coût numérique par rapport à la version profondeur.

Je propose aussi d'étendre l'analyse de vitesse par inversion au domaine des données. Ceci conduit à une approche du problème inverse plus linéarisée que celle de l'inversion des formes d'onde. Il suffit d'appliquer l'opérateur de modélisation aux images après la multiplication par l'annihilateur. Cette nouvelle approche est proche de l'inversion des formes d'onde dans le sens que le modèle optimal est obtenu lorsque la norme des résidus est minimale. D'un autre côté, l'approche est toujours basée sur le critère de cohérence. Le problème inverse est connu pour être plus convexe, au moins pour des modèles simples. Je compare la nouvelle approche avec d'autres méthodes pour établir des liens formels entre des méthodes dans le domaine des données et dans le domaine des images.

Les méthodologies sont analysées sur les jeux de données 2D et au travers de toute une série de méthodes de vitesse, en particulier des modèles avec la présence de zones de vitesses plus faibles à l'origine de triplications. Ces modèles ne sont pas nécessairement appropriés pour la migration par point de tir. J'applique aussi les méthodes au modèle Marmousi pour tester la robustesse. Les principales contributions de ce travail sont (1) le développement de la migration par point de tir avec amplitude préservée, et surtout le couplage avec l'analyse de vitesse ; et (2) l'extension de l'analyse de vitesse au domaine des données, et le lien entre les domaines données et images.

Contents

Acknowledgments	i
Abstract	iii
List of Figures	xi
List of Tables	xxiii
1 Introduction	1
1.1 Seismic imaging principles	3
1.1.1 Seismic data	4
1.1.2 Scale separation	6
1.1.3 Forward modeling	9
1.1.4 Inverse problem	10
1.2 Data-domain methods	12
1.2.1 Full Waveform Inversion	13
1.2.2 Linearized waveform inversion	16
1.2.3 Alternative methods	21
1.3 Image-domain methods	24
1.3.1 Surface-oriented MVA	26
1.3.2 Depth-oriented MVA	28
1.4 Motivations and thesis organization	29
1.4.1 Motivation I: towards common-shot inversion velocity analysis – a robust approach	30
1.4.2 Motivation II: investigating the link between data-domain and image-domain methods	31
1.4.3 Thesis organization	32
1.4.4 Contributions	34
2 Methodology	37
2.1 Introduction	39
2.2 Forward problem	39

2.2.1	Wave equation	39
2.2.2	Numerical solution	40
2.2.3	Stability and dispersion	41
2.2.4	Boundary condition	42
2.3	Data fitting principle	43
2.3.1	Full Waveform Inversion and alternatives	43
2.3.2	Reflection Waveform Inversion	56
2.3.3	Differential Waveform Inversion	58
2.4	Image coherency criteria	60
2.4.1	Surface-oriented MVA	62
2.4.2	Depth-oriented MVA	63
2.4.3	Limitations of MVA	65
2.5	Summary	70
3	Common-shot Inversion Velocity Analysis in the image-domain	73
3.1	Introduction	75
3.2	From migration to inversion	79
3.3	From MVA to IVA	82
3.4	Numerical examples	84
3.4.1	Homogeneous model	84
3.4.2	Low-velocity anomaly	86
3.4.3	Marmousi model	97
3.5	Discussions	108
3.6	Conclusions	111
3.7	Appendix I: Common-shot inversion scheme	112
3.8	Appendix II: Gradient derivation for IVA	117
4	Investigating the links between image and data domains	119
4.1	Introduction	121
4.2	Common-shot IVA in the data-domain	125
4.2.1	Objective function	125
4.2.2	Gradient of the objective function	127
4.2.3	Numerical examples	129
4.3	Comparisons between image and data domains	149
4.3.1	Equivalence between data-domain IVA and DWI	149
4.3.2	Data fitting versus image coherency	150
4.3.3	Numerical comparisons	154
4.4	Conclusions	164
4.5	Appendix I: Gradient derivation for IVA in the data-domain	166
5	Conclusions and Perspectives	169

5.1	Conclusions	171
5.1.1	Common-shot Inversion Velocity Analysis	171
5.1.2	Links between image-domain and data-domain methods	172
5.2	Perspectives	173
5.2.1	Edge effects	173
5.2.2	Introduction of more physics	173
5.2.3	Extension to 3D	174
5.2.4	Application to real data	175
References		177

List of Figures

1.1	Acquisition geometries for land and marine environments.	5
1.2	A synthetic example of marine seismic data acquired at the surface: (a) scattered paths in the model; (b) transmitted paths in the model; (c) shot gather recorded at the surface. The red star denotes the position of the source. Solid, dot dashed and dashed white lines in (a) denote first-reflected, multi-reflected and multi-diffracted waves, respectively. Solid, dot dashed and dashed white lines in (b) denote direct, diving and refracted waves, respectively.	7
1.3	In black, the classic sketch by Claerbout (1985) illustrating the spatial frequencies that can be resolved from seismic data. The gap is now filled by the improved resolution of tomography (red curve) and by the impact of broadband acquisition on imaging (blue curve) (from Lambaré et al., 2014).	8
1.4	A synthetic example of the scale separation for the Marmousi model. The full velocity model (left) is decomposed into a smooth background velocity model (middle) and a velocity perturbation model (right). . . .	8
1.5	Schematic for local optimization methods.	11
1.6	Illustration of the relationship between the wavenumber $\mathbf{k} = \mathbf{k}_s + \mathbf{k}_r$ and the opening angle θ (adapted from Alkhalifah and Plessix, 2014). . . .	14
1.7	Illustration of the origin of cycle skipping effect. Local optimization converges towards the global minimum for the modeled data with smaller time shift error (left pink circle), whereas it converges towards the local minimum for the modeled wavelet with larger time shift error (red star). The black arrows mark the descent directions.	15
1.8	Same as for Figure 1.7, but the low-frequency content is involved in the observed and modeled data. The cycle skipped case (red star) in Figure 1.7 now converges towards the global minimum with the help of low-frequency.	16
1.9	The principle of Born modeling: (a) calculation of traveltime; (b) Kirchhoff hyperbola on seismogram. (adapted from Pyun and Shin, 2008) . .	17
1.10	Description of imaging condition for a simple reflector.	18
1.11	Description of iterative migration.	19

1.12	Illustration indicating that migration cannot completely reproduce the data: (a) observed data, (b) migrated image with correct macro velocity model, and (c) modeled data with image (b) in the correct macro velocity model. The artificial direct arrivals in panel (c) correspond to migration artifacts in panel (b). The observed and modeled data are in phase, but their amplitudes are different.	19
1.13	Conventional RTM (a) and least-squares RTM (b) images of the synthetic salt model based on a GOM survey (adapted from Zeng et al., 2017).	21
1.14	Prestack common-offset data before (a) and after (b) migration. Seismic data set computed in 2D Marmousi model (a), migrated to surface-offset domain CIGs (b) by prestack depth migration using the true velocity model. A and C are common-offset gather and prestack migrated image, respectively, at zero-offset. B and D are common midpoint gather and CIG, respectively, at fixed position. Accurate velocity model results in flatness on the CIG panel D. (from Chauris, 2000)	27
1.15	Examples illustrating the impact of migration artifacts on image differences, for a single reflector embedded in a homogeneous macro velocity model. (a) corresponds to the exact reflectivity model, (b) to the image migrated in the correct macro velocity model for a single shot marked by green point in (a), and (c) to the difference between image (b) and the image for an adjacent shot. (d–e) are similar to (b–c), but for an incorrect model of which the velocity is 5% lower than the correct value. Black ellipses correspond to the upward-curved events attributed to velocity inaccuracies, and green ellipse to those attributed to migration smiles.	31
1.16	Same as for Figure 1.15, but for the Marmousi model.	32
2.1	Description of the grid points needed to update (v_x, v_x) (a) and u (b) fields (adapted from Thorbecke, 2015). For both cases, four neighbouring points are needed to compute one central point.	41
2.2	Snapshots of wavefields for the case of stable result (a), of instability (b), and of dispersion (c) (adapted from Thorbecke, 2015).	42
2.3	The construction of synthetic data: (a) the source wavelet, (b) the random pulses and (c) the convolution between the wavelet and random pulses.	46

2.4	Investigation of the shape of the ℓ_2 norm OF: (a) observed data (blue) and calculated data with shifted phase (red); (b) misfit curve related to different phase shifts (blue curve) and the misfit corresponding to calculated data in (a) (red dot); (c) observed data (blue) and calculated data with varying amplitude terms (red); (d) misfit curve associated to different amplitude ratios (blue curve) and the misfit corresponding to calculated data in (a) (red dot).	47
2.5	Same as Figure 2.4, but for the normalized crosscorrelation OF.	49
2.6	Same as Figure 2.4, but for the AWI OF.	50
2.7	Same as Figure 2.4, but for the logarithmic envelop OF. (a,c,d,f) correspond to (a-d) in Figure 2.4, respectively. (b,e) illustrate the envelopes with respect to data in (a,d), respectively. Note that the vertical axis of (f) is in logarithmic scale. For the amplitude case, the minimum is not reached for 1 due to the pre-whitening applied to the envelope.	52
2.8	Same as Figure 2.4, but for the envelop integral OF. (a,c,d,f) correspond to (a-d) in Figure 2.4, respectively. (b,e) illustrate the envelop integrals with respect to data in (a,d), respectively.	54
2.9	Same as Figure 2.4, but for the auxiliary bump OF. (a,c,d,f) correspond to (a-d) in Figure 2.4, respectively. (b,e) illustrate the auxiliary bump form of data in (a,d), respectively.	55
2.10	An illustration comparing the behavior of FWI and RWI. Sensitivity kernel of FWI for a single reflector embedded in the homogeneous background velocity model and its different subkernels (adapted from Chi et al., 2015): (a) full kernel of FWI, (b) transmission kernel, (c) migration ellipse, (d) source-side reflection kernel, and (e) receiver-side reflection kernel. For RWI, the kernel related to macro model update includes only (d) and (e), introducing the famous rabbit ear shapes. . . .	59
2.11	CIGs related to incorrect (left) and correct (right) macro velocity models (from Chauris et al., 2002b).	62
2.12	Schematic view of the wavefields (a, from the source s ; b, from the receiver r) contributing to the gradient in the presence of a horizontal scattering source (e.g. a single reflector) at depth z_0 (adapted from Chauris and Cocher, 2017). The parameters S_0 and λ_s as well as R_0 and λ_r correlate between the surface and depth z_0	64
2.13	Description of the horizontal subsurface-offset h	64
2.14	(a) CIGs and (b) associated gradients of the MVA OF computed with classic migration for a too low (left), correct (middle), and too high (right) velocity model. Blue, white and red colors correspond to negative, null and positive values, respectively (from Lameloise, 2015). . .	67
2.15	Same as Figure 2.14 with a quantitative (ray-based) migration instead of a classic (adjoint) migration (from Lameloise, 2015).	68

2.16	The MVA gradients for a homogeneous model with a single horizontal reflector before (a) and after (b) removing gradient artifacts by horizontal contraction (from Fei and Williamson, 2010).	69
2.17	A generic workflow for solving the inverse problem in both data and image domains (a). The differences among various approaches (b) in terms of data and model descriptions, and the criteria for the definition of OF.	71
3.1	Ray parameters \mathbf{p}_s and \mathbf{p}_r are the slowness vectors at the image point, associated to angles θ_s and θ_r . β_s and β_r are oriented angles at the source and receiver points. We define the opening angle as $\theta = \theta_s - \theta_r$	81
3.2	Stacked image sections obtained with migration (a, c) and inversion (b, d) for a single reflector at 0.6 km in incorrect homogeneous models, respectively. Compared to the exact model, the velocity is 0.5 km/s higher for (a, b) and 0.5 km/s lower for (c, d), respectively.	85
3.3	CIGs for position $x = 2.4$ km associated to images for Figure 3.2.	85
3.4	Gradients obtained with classical MVA (a, d), IVA with $\alpha = 0$ (b, e) and $\alpha = 1$ (c, f) for a single reflector at 0.6 km in incorrect invariant models, respectively. Compared to the exact model, the velocity is 0.5 km/s higher for (a–c) and 0.5 km/s lower for (d–f), respectively. Blue, white, red colors mean negative, zero, positive values, respectively.	86
3.5	Same as for Figure 3.4, after the Gaussian smoothing over half a wavelength.	87
3.6	(a) Exact velocity model, (b) stacked migrated section in $c_0 = 2.5$ km/s for all shots and (c) associated MVA gradient after a Gaussian smoothing over half a wavelength. From inside to outside of the Gaussian circle in panel (a), the velocity ranges from 1.6 to 2.5 km/s. Black stars and dashed line refer to shot position in panel (a). Red corresponds to positive and blue to negative value in panel (c).	88
3.7	For a single shot at $s = 1.6$ km, (a) observed data, (b) computed data after inversion and modelling in an incorrect model of constant velocity $c_0 = 2.5$ km/s, and (c) residuals at the same scale. Black arrows indicate the triplicated events.	89
3.8	Traces extracted from Figure 3.7, for surface-offset (a) -0.5, (b) 0, and (c) +0.5 km. The solid red line refers to the observed data and the dashed blue line to the data after inversion and modelling in an incorrect model of constant velocity $c_0 = 2.5$ km/s.	90

3.9	(a) Stacked inversion section in $c_0 = 2.5$ km/s for all shots, (b) first gradient after a Gaussian smoothing over half a wavelength, (c) final inverted model after 10 non-linear iterations, and (d) stacked inversion section in the final model. Red corresponds to positive and blue to negative value in panel (c). From inside to outside of the inverted circle in panel (c), the velocity ranges from 1.86 to 2.54 km/s. Panel (c) and Figure 3.6a are displayed with the same color scale.	91
3.10	CIGs associated to (a) the true Gaussian lens model, (b) the initial homogeneous model and (c) the updated model for position x , ranging from 0.9 to 2.7 km, every 0.3 km. Dashed black ellipses indicate the energy related to the triplication effect. Black arrows mark the same event corresponding to different macro models.	92
3.11	(a) Exact velocity model and (b) observed data for a single shot at $s = 1.6$ km. From inside to outside of the Gaussian circle in panel (a), the velocity ranges from 1.9 to 2.5 km/s. Black stars and dashed line refer to shot position in panel (a).	93
3.12	Same as for Figure 3.9, but for the case displayed in Figure 3.11. From inside to outside of the inverted circle in panel (c), the velocity ranges from 2.05 to 2.53 km/s.	94
3.13	Same as for Figure 3.10, but for the case displayed in Figure 3.11. . . .	95
3.14	Rays and wavefronts calculated from a source at $s_x = 1.6$ km. (a–d) correspond to Figures 3.6a, 3.9c, 3.11a and 3.12c, respectively. Black arrows mark the triplications. The macromodels obtained with IVA contain no triplications (b,d). All images are displayed with the same scale.	96
3.15	Same as for Figure 3.7, but the computed data is obtained in the correct background model.	96
3.16	Same as for Figure 3.7, but the traces are extracted from Figure 3.15. Black arrows mark the major differences between exact and reconstructed data.	97
3.17	Gradients obtained with classical MVA (a, b, e and f) and with IVA for $\alpha = 1$ (c, d, g and h) for the Marmousi model. The initial background velocity is a constant gradient model (from shallow to deep part: 2.0–5.0 km/s) with velocity higher than the true velocity (a, c, e and g). The initial background velocity is a 1.5 km/s homogeneous model for (b, d, f and h). Images on the right are obtained from images on the left after a Gaussian smoothing over half a wavelength distance.	98
3.18	Convergence curve for the IVA optimization performed on the Marmousi model. The vertical axis is in logarithmic scale. Blue line represents the misfit value of objective function in the iteratively updated model and red dashed line represents the cost value for the exact velocity model.	99

- 3.19 True Marmousi macromodel (a) and total updated model ($c_0 + \delta c$) after 100 IVA iterations (b). White stars and dashed line indicate the shot position extension. (c–f) are vertical velocity profiles at positions 2.6, 3.6, 4.6 and 5.6 km, respectively. Solid blue, dashed-dot red and solid red lines correspond to true, initial and updated macromodels, respectively. 100
- 3.20 Stacked inverted images associated to (a) true, (b) initial and (c) updated (Figure 3.19) macromodels, respectively. 101
- 3.21 Profiles of stacked inverted images associated with true, initial and updated (Figure 3.19) macromodels. From left to right, columns are related to positions 3.0, 4.0, 5.0 and 6.0 km, respectively. Blue lines refer to image profiles for true model. The red lines represent profiles for initial model in the top rows and for updated model in the bottom rows. 102
- 3.22 CIGs associated to (a) the true Marmousi model, (b) the initial homogeneous model and (c) the updated (Figure 3.19) macromodel for position x , ranging from 2.5 to 6.5 km, every 0.5 km. We apply a taper on the image above the depth indicated by dashed black line, to exclude the associated contribution to the objective function. Black ellipse marks a local coherent event for image position $(x, z) = (4.5, 1.0)$ km. 103
- 3.23 Single-shot inversion test in the Marmousi model: (a) observed data, (b) inverted image, (c) reconstructed data and (d) single trace data comparison for trace at $x = 5$ km indicated by the black dashed line, respectively. The blue and red lines correspond to the observed and reconstructed data, respectively. 105
- 3.24 Shot gathers of (a) observed data and (b–d) modeled data for the shot at 4.6 km. Modeled data are computed for (b) true, (c) the initial and (d) the updated background models, with the associated stacked inverted images after summation over all shots. Curves in (e–g) represent the comparison between observed and modeled data for a trace at 4.6 km. The blue curves correspond to the observed data and the red dashed curves to the modeled data. The black arrow in panel (e) marks the event corresponding to the one in the ellipse in Figure 3.22. 106
- 3.25 True Marmousi full model (a) and updated model (b) after 70 FWI iterations starting from the final IVA result (Figure 3.19b). (c–f) are vertical velocity profiles at positions 2.6, 3.6, 4.6 and 5.6 km, respectively. Solid blue and dashed-dot red lines represent true and updated models, respectively. 107

- 3.26 Single-shot test in the Marmousi model for a shot located at 4.6 km at the surface. (a) True reflectivity model; (b) Data modeled from (a); (c) Migrated image using the method of [Qin et al. \(2015\)](#); (d) Data modeled from (c); (e) Migrated image using our method; (f) Data modeled from (e). (a), (c) and (e) are plotted with the same scale. (b), (d) and (f) are all modeled in a correct smooth background model and plotted at the same scale. 115
- 3.27 Comparison between exact and reconstructed single trace data. (a–d) are examples using the method of [Qin et al. \(2015\)](#). (e–h) are examples using our method. From left to right, the column refers to the single trace data located at receiver positions 2, 3, 5 or 6 km, respectively. Blue curve represents exact data. Red curve represents reconstructed data. 116
- 3.28 RMS misfit between observed and reconstructed data for each receiver position. Blue curve represents [Qin et al. \(2015\)](#)'s method. Red curve represents the approach developed here. 116
- 4.1 Image-domain and data-domain IVA. Left is observed shot gather. Middle is the constructed CIG after the application of the inverse operator to the an observed data. Red, blue and green dashed lines correspond to the cases of under-estimated, correct and over-estimated background velocity models, respectively. Right are the data modeled from the differential images in the given background velocity. Different background velocity correspond to modeled data of almost the same kinematics and they are distinguished by their amplitudes. 127
- 4.2 Stacked image sections obtained with inversion operator B^\dagger for a single reflector at 0.6 km in different constant velocity models, respectively. Compared to the exact model, the velocity is 0.3 km/s lower for (a), equivalent for (b) and 0.3 km/s higher for (c), respectively. 130
- 4.3 CIGs for position $x = 2.4$ km (a–c) associated to images for Figure 4.2. (d–f) are the residuals between adjacent traces in CIGs (a–c). Green dashed lines illustrate the curvatures of events related to different model inaccuracies. (a–c) and (d–f) are displayed with the same scale, respectively. 131
- 4.4 Observed data for shot position $s_x = 4.0$ km (a), and scaling data for shot position $s_x = 4.0$ km (b–d) associated to image incoherency for (d–f) in Figure 4.3, respectively. Green and black arrows correspond to the direct arrival events and the reflected events, respectively. 132

- 4.5 Gradients obtained with data-domain IVA ($\beta = 0$) for a homogeneous macromodel with velocity too low (a, d), with correct velocity (b, e) and with velocity too high (c, f), respectively. We apply the Gaussian smoothing over half a wavelength to (a–c) to derive (d–f), respectively. All images are displayed with the same scale. 132
- 4.6 Decomposition of macromodel gradients obtained with image-domain IVA or data-domain IVA for a homogeneous macromodel with velocity too low. (a) corresponds to the additional term related to $\alpha = 1$ for image-domain IVA. (b–d) correspond to $\tilde{\lambda}_s \star S_0$, $\tilde{\lambda}_r \star R_0$ and the oscillatory term in equation 4.10 for image-domain IVA ($\alpha = 0$), respectively. (e–h) correspond to $\lambda_d \star S_d$, $\lambda_s \star S_0$, $\lambda_r \star R_0$ and the oscillatory term in equation 4.11 for data-domain IVA ($\beta = 0$), respectively. (a–d) and (e–h) are displayed with same scale, respectively. 134
- 4.7 Gradients obtained with image-domain IVA ($\alpha = 0$) (a), image-domain IVA ($\alpha = 0$) (b) and data-domain IVA ($\beta = 0$) (c) for a homogeneous macromodel with velocity too low, after a Gaussian smoothing over half a wavelength. (a) corresponds to the summation of Figures 4.6b-4.6d, (b) to the summation of Figures 4.6a-4.6d, and (c) to the summation of Figures 4.6e-4.6h. 134
- 4.8 Same as for Figure 4.6, but for a homogeneous macromodel with velocity too high. 135
- 4.9 Same as for Figure 4.7, but for a homogeneous macromodel with velocity too high. 135
- 4.10 Gradients obtained with image-domain IVA ($\alpha = 0$) (a), image-domain IVA ($\alpha = 0$) (b) and data-domain IVA ($\beta = 0$) (c) with respect to the correct macro velocity model. The annihilator in all methods are replaced by the identity factor. We apply the Gaussian smoothing over half a wavelength to (a–c) to derive (d–f), respectively. (a,b,d,e) and (c,f) are displayed with the same scale, respectively. 136
- 4.11 Gradients obtained with data-domain IVA for the Marmousi model. The initial model is a 1.5 km/s homogeneous model for (a,c,e). The initial model is a constant gradient model (from shallow to deep part: 2.0–5.0 km/s) with velocity higher than the true velocity (b,d,f). We apply the Gaussian smoothing over half a wavelength to (a,b) for (c,d) and over one wavelength for (e,f), respectively. Blue, white and red represent negative, zero and positive values, respectively. 137
- 4.12 Convergence curve for the data-domain IVA performed on the Marmousi model. Blue line represents the misfit value of objective function for the iteratively updated model and red dashed line represents the misfit value for the exact velocity model. 138

- 4.13 Example illustrating the optimization converging towards a local minimum with an improper smoothing strategy. True Marmousi background model (a) and updated model after 100 data-domain IVA iterations (b). White stars and dashed line indicate the shot position extension. (c–f) are vertical velocity profiles at positions 2.6, 3.6, 4.6 and 5.6 km, respectively. Solid blue, dashed-dot red and solid red lines correspond to true, initial and updated models, respectively. 139
- 4.14 True Marmousi background model (a) and updated model after 100 data-domain IVA iterations (b). White stars and dashed line indicate the shot position extension. (c–f) are vertical velocity profiles at positions 2.6, 3.6, 4.6 and 5.6 km, respectively. Solid blue, dashed-dot red and solid red lines correspond to true, initial and updated models, respectively. . . 140
- 4.15 Stacked inverted images associated to (a) true (Figure 4.14a), (b) initial and (c) updated (Figure 4.14b) models, respectively. 142
- 4.16 Profiles of stacked inverted images associated with true, initial and updated (Figure 4.14b) models. From left to right, columns are associated to positions 2.6, 3.6, 4.6 and 5.6 km, respectively. Blue lines refer to image profiles for true model. The red lines represent profiles for initial model in the top rows and for updated model in the bottom rows. 143
- 4.17 CIGs associated to (a) the true Marmousi model, (b) the initial homogeneous model and (c) the updated model (Figure 4.14b) for position x , ranging from 2.5 to 6.5 km, every 0.5 km. We apply a taper on the image above the depth indicated by dashed black line, to exclude the associated contribution to our objective. 144
- 4.18 The same as Figure 4.17, but for the image residuals computed by comparing adjacent traces in CIGs. All images are displayed with the same scale. 145
- 4.19 The same as Figure 4.17, but for the time-domain scaling data sets, computed for shot position s_x , ranging from 3.4 to 5.8 km, every 0.4 km. All images are displayed with the same scale. 146
- 4.20 Rays and wavefronts calculated from a source at surface with $s_x = 6.0$ km in (a) true and (b) inverted (Figure 4.14b) Marmousi macromodels, respectively. All images are displayed with the same scale. 147
- 4.21 True Marmousi full model (a) and updated model (b) after 70 FWI iterations starting from the final data-domain IVA result (Figure 4.14b). (c–f) are vertical velocity profiles at positions 2.6, 3.6, 4.6 and 5.6 km, respectively. Solid blue and dashed-dot red lines represent true and updated models, respectively. 148

- 4.22 The first macromodel gradient, after a Gaussian smoothing over half a wavelength, obtained with image-domain IVA for $\alpha = 0$ (a,e) and for $\alpha = 1$ (b,f), and with data-domain IVA (c,g). (d,h) are (c,g) multiplied by the value of depth z , respectively. Compared to the exact model, the velocity is 0.5 km/s lower for (a–d) and 0.5 km/s higher for (e–h), respectively. 156
- 4.23 Profiles at position $x = 6$ km for the first gradients displayed in Figure 4.22. (a,b) are related to image-domain IVA ($\alpha = 0$). In plots (c,d), solid blue, solid red and dashed red lines correspond to image-domain IVA ($\alpha = 1$), and data-domain IVA before and after preconditioning, respectively. Compared to the exact model, the velocity is 0.5 km/s lower for (a,c) and 0.5 km/s higher for (b,d), respectively 157
- 4.24 Profiles at $x = 4.6$ km extracted from background models obtained with different methods after few nonlinear iterations. Dashed black lines correspond to the exact macromodel. Solid blue, solid red, and dashed red lines correspond to the results obtained with image-domain IVA ($\alpha = 1$), and with data-domain IVA before and after preconditioning, respectively. The gradients for data-domain IVA is multiplied by z in (a,b), and by z^2 in (c,d). (a,c) and (b,d) correspond to the results after 1 iteration and 2 iterations, respectively. 159
- 4.25 Macromodels obtained with different methods after 5 nonlinear iterations (a) and associated CIGs (b–d). (a) is the same as Figure 4.24d, but for the case after 5 iterations. (b) is associated to the correct macromodel, and (c–e) to the results of non-preconditioned data-domain IVA, data-domain IVA preconditioned by z^2 and non-preconditioned image-domain IVA, respectively. 160
- 4.26 The background model obtained after performing 100 nonlinear iterations with imaged-domain IVA (a) and with data-domain IVA (b). Red arrows are at same positions for (a,b) and mark the area exhibiting differences. (c–f) are vertical velocity profiles at positions 3.5, 4.5, 5.5 and 6.5 km, respectively. Dashed black, solid blue and solid red lines correspond to true Marmousi macromodel, image-domain result (preconditioner z^2) and data-domain result (preconditioner z^2), respectively. Green arrow is at the same position as in Figure 4.28. 161
- 4.27 CIGs for starting macromodel (a,f), exact macromodel (b,e) and the final results obtained with image-domain IVA (c,g) and data-domain IVA (d,h). (a–d) correspond to shot position 4.6 km and (f–h) to shot position 6.2 km. Dashed green rectangles are at the same location for (e–h) to mark the differences. 162
- 4.28 Same as for Figure 4.26, but for the results after subsequent FWI. 163

-
- 5.1 Description of the impact of edge effects on macromodel gradient obtained with image-domain IVA. Black dashed lines mark the boarder between well illuminated area and edges. 174

List of Tables

1.1	Non exhaustive review of published direct inversion approaches. For subsurface-offset and common-angle domains, the reader is referred to section 1.3.2.	22
3.1	Dimensions of the data, extended and shot domains. s and r are the source and receiver coordinates, respectively. t is the time. (x,y,z) are the spatial coordinates. h is the subsurface-offset.	76
3.2	Non exhaustive references related to different MVA/IVA approaches. . .	78
3.3	Comparison of different approaches in terms of memory and computational requirements for calculating the gradient of the associated objective function. (n_z, n_x) denote the size of the 2D model and n_t the sample size of the time for the data. n_h is the number of the subsurface-offset sampling and n_s the number of the shot. For a given source term, q_{mod} and q_{cross} are the costs required to solve the wave equation once and to crosscorrelate two wavefields once, respectively.	110
4.1	Comparison between different approaches in terms of computational requirements for calculating the gradient of associated objective functions with respect to model. n_s is the number of shots. For a given source term, q_{mod} is the CPU cost required to solve the wave equation once. The CPU cost for crosscorrelation is neglected here.	129
4.2	Strategies proposed for a successful implementation of image-domain and data-domain IVA, at least for the case of Marmousi model.	164

Chapter 1

Introduction

Contents

1.1	Seismic imaging principles	3
1.1.1	Seismic data	4
1.1.2	Scale separation	6
1.1.3	Forward modeling	9
1.1.4	Inverse problem	10
1.2	Data-domain methods	12
1.2.1	Full Waveform Inversion	13
1.2.2	Linearized waveform inversion	16
1.2.3	Alternative methods	21
1.3	Image-domain methods	24
1.3.1	Surface-oriented MVA	26
1.3.2	Depth-oriented MVA	28
1.4	Motivations and thesis organization	29
1.4.1	Motivation I: towards common-shot inversion velocity analysis – a robust approach	30
1.4.2	Motivation II: investigating the link between data-domain and image-domain methods	31
1.4.3	Thesis organization	32
1.4.4	Contributions	34

Résumé du chapitre 1

L'imagerie sismique est très utilisée pour caractériser les structures géologiques du sous-sol au travers de l'analyse des données sismiques, pour ainsi révéler de possibles ressources souterraines. Le modèle de sub-surface est constitué de paramètres physiques comme les vitesses des ondes de pression et de cisaillement ou encore la densité des roches. Ces paramètres contrôlent la propagation des ondes dans la Terre. En raison des variations spatiales de ces paramètres, les ondes sismiques sont réfractées, réfléchies et diffractées au cours de la propagation. Des capteurs à la surface enregistrent les ondes qui ont interagit avec la sub-surface et fournissent ainsi les données pour l'imagerie sismique. Le traitement géophysique a pour objectif de convertir ces données en images du sous-sol. Cette étape cruciale est formulée comme un problème inverse pour déterminer les meilleurs paramètres du modèle Terre. Un bon modèle sert ensuite pour l'interprétation géologique, la détermination des forages et le positionnement des puits. Le problème inverse est difficile à résoudre en pratique. Parmi plusieurs raisons, la forte non-linéarité entre les données et les paramètres du modèle joue un rôle important. C'est le cas lorsque le modèle correspond aux grandes longueurs d'onde de la vitesse : un macro-modèle qui augmente de 10% ne conduit pas à un champ de pression qui change dans les mêmes proportions, mais modifie le champ d'onde complet, et en particulier les temps d'arrivées. Cette thèse se focalise sur cette thématique.

En ce qui concerne la nature des ondes sismiques, la plupart des méthodes d'imagerie considère l'approximation acoustique avec les ondes de pression, afin de simplifier les formulations mathématiques et réduire le coût de calcul. Cette thèse est une contribution à la construction du modèle de vitesse dans le cadre de l'approximation acoustique isotrope.

Dans l'introduction, je résume brièvement les principes de base autour de l'imagerie sismique dans le contexte de l'exploration géophysique, et regarde plus particulièrement les approches classiques de la résolution du problème inverse dans les domaines des données et des images. Finalement, je motive les développements et j'explique les principales limitations actuelles des méthodes dans le domaine image et l'intérêt à faire le lien entre les domaines des données et des images.

Seismic imaging is widely used to characterize the geological structures in the subsurface from the analysis of seismic data, and thus to reveal possible resource bearing formations. The subsurface model consists of a set of physical parameters such as pressure and shear wave velocities or rock density that control the wave propagation in the Earth. Due to the spatial variations of model parameters, emitted seismic waves are refracted, reflected and diffracted during propagation. Sensors deployed at the surface record the waves traveling back to the surface after having interacted with the subsurface, to provide the data used in seismic imaging. Geophysical processing can convert those surface measurements into the images of the subsurface. This crucial step is formulated as the inverse problem aiming at determining the model parameters. An accurate subsurface model is important for subsequently interpreting the geology, determining the drilling location, and correctly positioning the wells. The inverse problem, however, is difficult to solve. Among others, one reason is the highly nonlinear relationship between data and model parameters. This is the case when the model corresponds to the large-scale part of the velocity model: a macro velocity model increased by 10% does not scale the related pressure field by the same amount, but modifies the total wavefields, including the arrival times. This thesis especially focuses on this issue.

Regardless of the elastic nature of waves, most seismic imaging methods are based on acoustic assumptions and involve only pressure waves, to simplify the mathematical formulation and greatly reduce the computational cost. This thesis is a contribution related to the velocity model building under the isotropic acoustic approximation.

In this introduction, I first briefly summarize the basic principles behind seismic imaging in the context of exploration geophysics, and then pay attention to the classic approaches addressing the inverse problem defined in the data-domain or the image-domain. Finally, I motivate the research developments by explaining the main limitations of current image-domain methods and by indicating the interest of seeking the links between image-domain and data-domain methods.

1.1 Seismic imaging principles

In exploration geophysics, seismic experiments are widely used for the subsurface imaging and the reservoir management. The seismic waves can penetrate into the Earth and thus bring information about the geological structures involved in industrial production, located at a depth of a few kilometers in the subsurface. Conventional seismic exploration mainly takes advantage of the active seismic experiments, meaning the sources are artificially triggered, in opposition to the passive earthquake sources used in global seismology (Lay and Wallace, 1995; Dahlen and Tromp, 1998; Aki and Richards, 2002). The passive methods, such as interferometry have also been taken into account for exploration problems nowadays (Schuster et al., 2004; Schuster, 2016). In general, the study of a physical system can involve three essential elements: observed data, for-

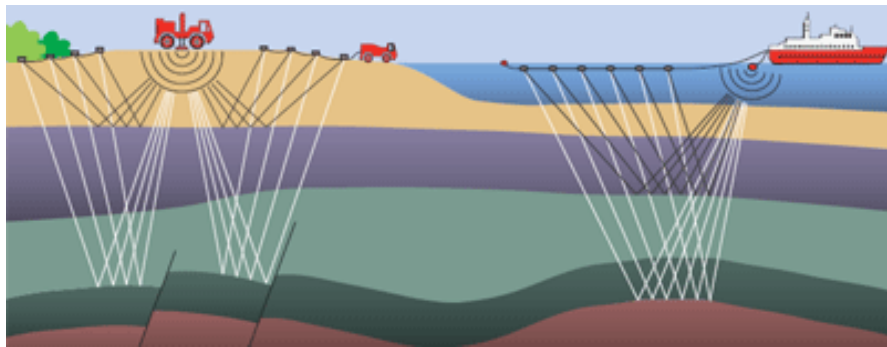
ward modeling and inverse problem (Tarantola, 2005). In the following, I review the principles of seismic imaging for exploration problems. Additionally, I also introduce the concept of scale separation (Claerbout, 1985) distinguishing different wavenumber components of the model, as it is the basis for many seismic imaging techniques. The reader is referred to Sheriff and Geldart (1995); Yilmaz (2001) for a broad introduction to seismic imaging.

1.1.1 Seismic data

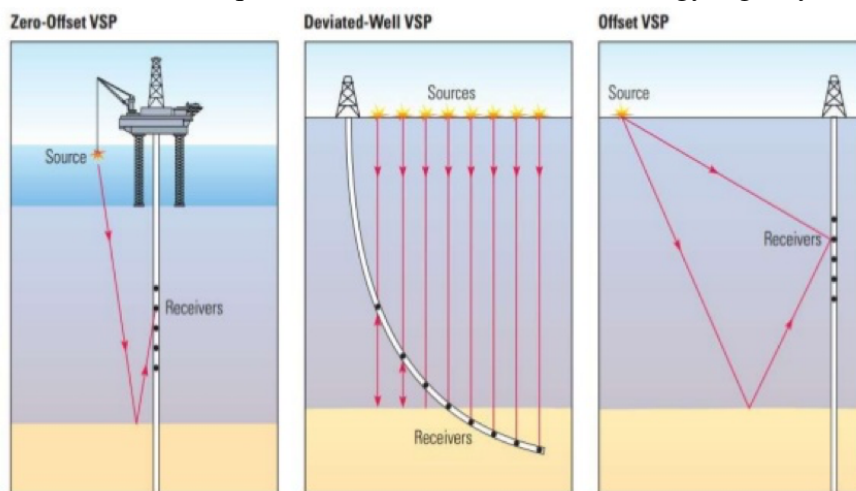
Seismic data can be acquired in land and marine environments. In the land case, the source is usually a truck-mounted seismic vibrator and the receivers called geophones record the motion of particles. The receivers are typically deployed on both sides of the source at the surface (Figure 1.1a). In a marine acquisition, the source is an air-gun and the receivers called hydrophones measure the pressure. The receivers are located on one side of the source, along several streamers towed by a marine seismic vessel (Figure 1.1a). Alternatively, the receivers can also be located at the sea floor in an Ocean Bottom Cable configuration (MacLeod et al., 1999; Plessix and Perkins, 2010). In both land and marine cases, the presence of a drilled borehole can provide a different acquisition system called Vertical Seismic Profiling (VSP) (Balch and Lee, 1984; Hardage, 1985; Soni, 2014). The source is similar to the one in conventional experiments but receivers are located within a well (Figure 1.1b). In the cross-well configuration, the source is a shot-hole dynamite, and sources and receivers are located in two different wells (Rector, 1995; Zhou et al., 1995; Plessix et al., 2000).

A *trace* of seismic data is the discrete-time signal measured at a single receiver. A group of traces recorded for the same source is usually displayed on a panel called *common-shot gather* or simply *shot gather*, with the time on the vertical axis and source-receiver distance (i.e. offset) on the horizontal axis. The shot gather records various types of structural responses to the excited source (Figure 1.2). Particular events, for example surface waves in land data and ghosts in marine data, are commonly considered as noise in seismic exploration and have to be removed by preprocessing (Yilmaz, 2001), whereas global or engineering seismology may benefit from surface waves to characterize the near-surface structure (Xia et al., 1999; Socco and Strobbia, 2004; Shapiro et al., 2005; Pérez Solano, 2013). Among others, body waves are mainly considered in seismic imaging and can be categorized according to the paths that connect the source and the receiver. One group is labeled as *transmitted waves*, involving

- *Direct wave* – This type of wave travels across the superficial part of the model. The associated wave path is a straight line if the very shallow zone of the model is homogeneous;
- *Diving wave* – Wave-paths can be curved in the case of increasing velocities with depth. The effect may bend the waves back to the surface and the associated



(a) Surface acquisitions (taken from Danish Energy Agency)



(b) VSP acquisitions (taken from Schlumberger)

Figure 1.1 – Acquisition geometries for land and marine environments.

recording is called diving waves. Consequently, this type of waves only has a limited penetrating depth, especially for short offsets. They may arrive at the receiver with shorter time than the direct waves at large offsets;

- *Refracted wave* – With sharp property contrasts existing in the medium, the propagated waves can be refracted for the critical angles along these interfaces and then later return to the surface.

The other group is identified as *scattered waves*, consisting of

- *First-order scattered wave* – The surface excited waves can be diffracted or reflected when it propagates through the strong variations of medium properties. Interfaces generate reflections and sharp edges trigger diffractions. The first-order scattered waves are reflected or diffracted only once during propagation. They are

directly linked to the rapid changes of model parameters. I will focus on this type of waves in this thesis;

- *High-order scattered wave* – The multi-scattering is caused by strong medium property contrasts such as the sea surface or the water bottom. The energy of up-going first-scattered waves is partly sent back to the subsurface and can reflect or diffract a few times before being recorded. Their contributions are commonly either neglected or removed by the preprocessing (Verschuur et al., 1992; Verschuur, 2006). Alternatively, recent investigations propose to extend seismic imaging to multiples (Guitton, 2002; Jiang et al., 2005, 2007; Berkhout, 2012; Verschuur and Berkhout, 2015; Cocher, 2017).

The preprocessing step is an essentially preliminary procedure before determining the Earth's model parameters. It consists of selecting the interested events for the subsequent analysis and of attenuating noise inherent to field data (see Yilmaz, 2001 for more details). Some developments attempt to take advantage of all information included in the data, such as transmissions, multiples, or more generally the full wavefield (see section 1.2.1). However, many seismic imaging techniques still rely on primary reflection data only. In this case, the other events like multiples are considered as noise and should be removed before the further analysis. The removal of multiples is an intense research topic (Verschuur, 2006). Here, the thesis only relies on the primary reflected data.

1.1.2 Scale separation

It is difficult in practice to fully reconstruct the velocity model c with limited acquisitions (Virieux and Operto, 2009). Conventional surface acquisition system usually provides insufficient low frequencies, limited offsets and/or restricted azimuth coverage in the observed data. As demonstrated by Claerbout (1985), the model reconstructed from seismic data lacks intermediate spatial frequencies (i.e. wavenumbers). The recovered model mainly consists of two separate ranges in the spectrum (Figure 1.3), leading to the concept of *scale separation* that distinguishes between perturbation (high wavenumbers) and background (low wavenumbers) models (Jannane et al., 1989). The velocity model c is thus split into two parts,

$$c(\mathbf{x}) = c_0(\mathbf{x}) + \delta c(\mathbf{x}), \quad (1.1)$$

where δc is the *reflectivity* model and c_0 the *macro velocity* or *background velocity* model (Figure 1.4), according to the associated wavenumber contents. They all depend on the spatial coordinate \mathbf{x} . The full model c can be inverted without scale separation, as in Full Waveform Inversion (FWI) presented in section 1.2.1. Alternatively, two main categories of seismic imaging methods have been established aiming at recovering different scales (Mora, 1989):

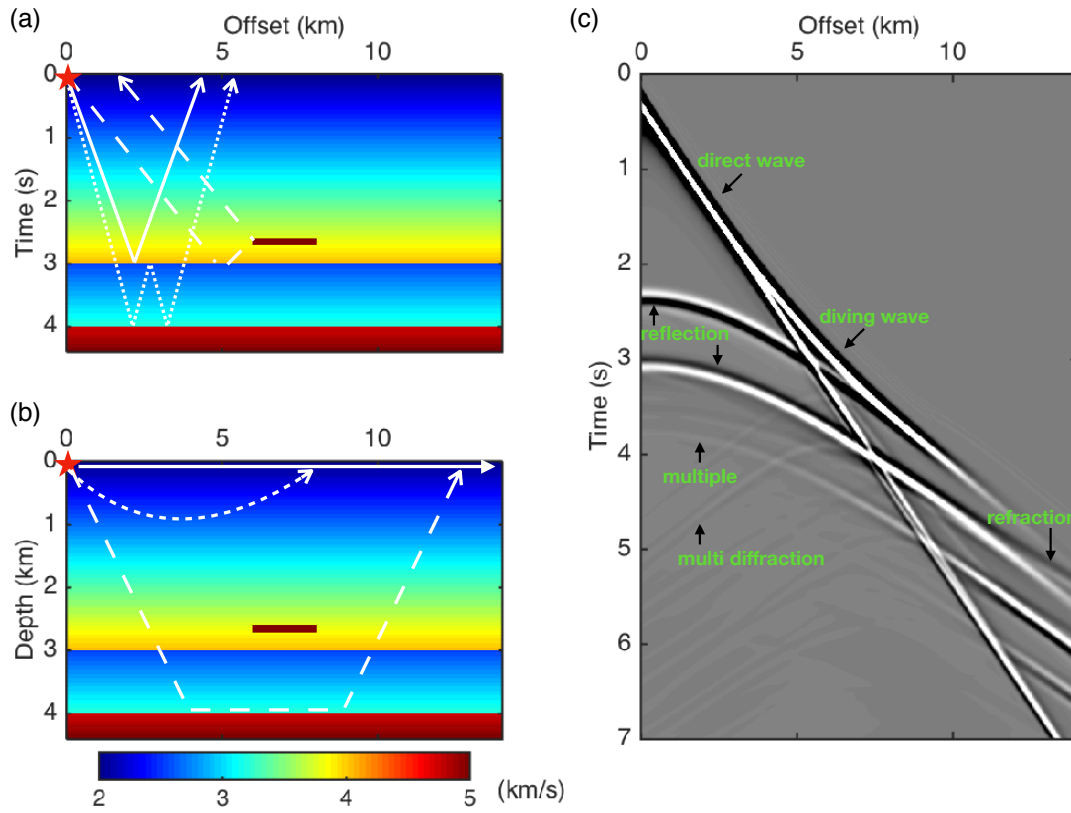


Figure 1.2 – A synthetic example of marine seismic data acquired at the surface: (a) scattered paths in the model; (b) transmitted paths in the model; (c) shot gather recorded at the surface. The red star denotes the position of the source. Solid, dot dashed and dashed white lines in (a) denote first-reflected, multi-reflected and multi-diffracted waves, respectively. Solid, dot dashed and dashed white lines in (b) denote direct, diving and refracted waves, respectively.

- *Migration* aims at recovering the reflectivity model δc (the high-wavenumber part). Under the Born approximation, the determination of δc is a linear inverse problem. Such approaches require an accurate background velocity model c_0 . I will review migration schemes in section 1.2.2;
- *Macro velocity estimation* seeks the reconstruction of the macro velocity model c_0 (the low-wavenumber part) that controls the kinematics of the recorded seismic responses. The macro velocity model is determined with *tomographic* methods (in a broad sense) formulated either in the data-domain or in the image-domain (i.e. before or after migration). I will review tomographic strategies in sections 1.2.3 and 1.3.

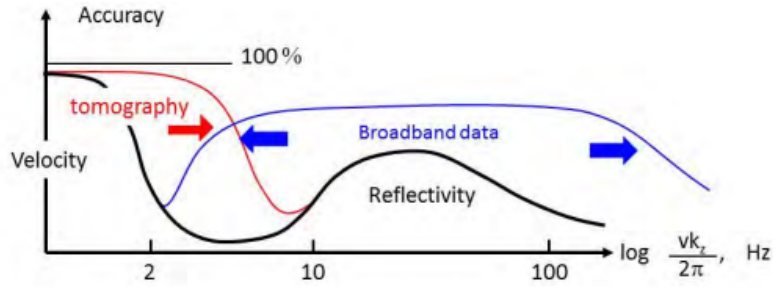


Figure 1.3 – In black, the classic sketch by [Claerbout \(1985\)](#) illustrating the spatial frequencies that can be resolved from seismic data. The gap is now filled by the improved resolution of tomography (red curve) and by the impact of broadband acquisition on imaging (blue curve) (from [Lambaré et al., 2014](#)).

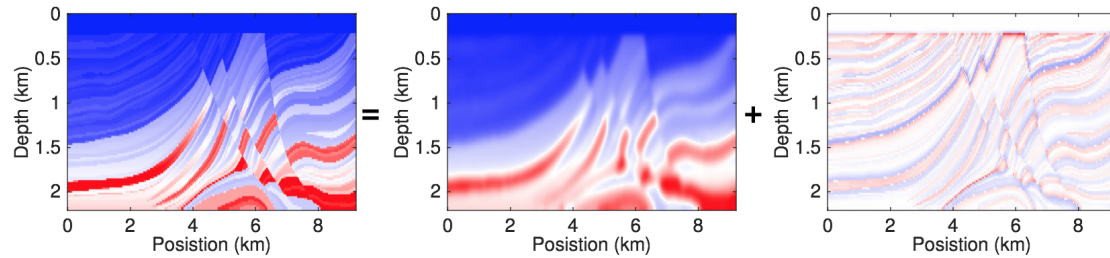


Figure 1.4 – A synthetic example of the scale separation for the Marmousi model. The full velocity model (left) is decomposed into a smooth background velocity model (middle) and a velocity perturbation model (right).

In seismic imaging, it is common to deal with two types of information included in the surface measurements: the *dynamic* and the *kinematics* aspects. The dynamic aspect has a direct impact on the amplitudes of the seismic waves. It is related to the reflection/transmission coefficient, but also depends on the source wavelet, the source and receiver distributions, etc. On the other hand, the kinematics of wave propagation are mainly controlled by the long-wavelength part of the velocity model, and the main consequence on the seismic wave is the arrival time. It also has an impact on the amplitudes due to the geometrical spreading, the attenuation in a dissipative media, etc.

Note that the spectrum gap is significantly filled nowadays ([Nichols, 2012](#); [Lambaré et al., 2014](#)) (Figure 1.3): (1) the improvement of the acquisition system allows recording seismic data with longer offset, wider azimuth and broader frequency band; (2) the advanced imaging tools can recover the background velocity model with more details. Nevertheless, scale separation remains the theoretical basis of many conventional and new imaging methods. This thesis explicitly uses such assumption for the parameterization of the model.

Before discussing data-domain and image-domain methods, I first review the main elements in forward modeling and in the resolution of the inverse problem.

1.1.3 Forward modeling

For subsequent imaging procedures, it is essential to define a physical law that allows to properly describe the link between the chosen model parameters and the associated data. It is mathematically formulated as a hyperbolic partial differential equation called wave equation, which simulates the wave propagation giving a set of model parameters. In the real Earth, the wave propagation is affected by many subsurface properties. Due to the high computational cost of simulating the visco-elastic anisotropic wave propagation, I consider only the isotropic acoustic approximation in this thesis. Anisotropy, elasticity and attenuation are not included. The Earth's model is thus parameterized by the pressure wave velocity V_p and the rock density ρ . The reflections are generated from rapid changes of acoustic *impedance* $I_p = V_p \rho$. In the constant density case, only the pressure wave velocity needs to be specified to solve the wave equation. Even for this simplest assumption, one still has to input a large number of parameters since the model parameters depend on the spatial coordinates.

The wave equation is commonly resolved by numerical modeling schemes including finite-difference method (FDM) (Virieux, 1986; Levander, 1988; Etgen and O'Brien, 2007), finite-element method (FEM) (Smith, 1975; Marfurt, 1984), spectrum-element method (SEM) (Komatitsch and Vilotte, 1998; Komatitsch and Tromp, 1999), etc. In exploration geophysics, the two main forward modeling tools are FDM and FEM. Moczo et al. (2010, 2011) compared the two schemes and concluded that the accuracy of the two methods are comparable with fine samplings. Virieux et al. (2011) have reviewed the efficiency and complexity of the two methods, and indicated that FDM is widely used due to its simplicity to implement and the relatively lower computational cost. The main advantage of FEM is its flexibility in meshing to deal with the boundary conditions and irregular structures. Compared to FDM, it is more numerically expensive and more difficult to implement. SEM is more applied for global seismology scale problems (Komatitsch and Vilotte, 1998; Komatitsch and Tromp, 1999; Fichtner, 2011). Some investigations proposed to use SEM on the exploration scale (Luo et al., 2009, 2013), but SEM is very demanding in terms of the computational cost and its grid meshing requires a priori knowledge of the subsurface structure. This thesis will concentrate on simulating the wavefield with FDM.

The modeling can also rely on the ray theory, which is based on the high-frequency asymptotic approximation of the wave equation (Červený, 2005). The propagation of waves in the subsurface is described by rays sharing the similar propagation laws used in optics (Červený et al., 1977; Chapman, 2004). For this scheme, the Green's function can be decomposed into three parts: one corresponding to the traveltime, one to the amplitude and one to the source signature. In a given velocity model, the traveltime can

be simulated either by ray-tracing strategy (Julian, 1977; Červený, 1987), or by solving the Eikonal equation (Vidale, 1988; Podvin and Lecomte, 1991; Qin et al., 1992).

In this thesis, I only focus on wave-equation-based modeling methods (in particular FDM) and ray-based modeling methods will not be considered, except for analyzing the main effects of wave-equation operators on the data.

1.1.4 Inverse problem

Objective function

The inverse problem consists of defining a scalar *objective function* (also called *cost function*) to evaluate the accuracy of model parameters used for modeling. It is designed such that the best model corresponds to the global maximum or minimum in the objective function. In the data-domain, the objective functions can be defined by measuring the misfit between observed and modeled data, typically in the least-squares sense (Tarantola, 1984a; Pratt et al., 1998). Alternatively, the data misfit can also be assessed using other schemes such as crosscorrelation (Luo and Schuster, 1991; Van Leeuwen and Mulder, 2010), deconvolution (Luo and Sava, 2011; Warner and Guasch, 2016), optimal transport distance (Métivier et al., 2016), Huber norm (Guitton and Symes, 2003), etc. On the other hand, one can first migrate the data and then formulates the objective functions in the image-domain, with semblance principle (see section 1.3 for details), measuring the level of coherency or focusing of reconstructed reflectivity images of the Earth (Al-Yahya, 1989; Symes and Carazzone, 1991; Sava and Biondi, 2004; Symes, 2008).

In seismic imaging, the inverse problem generally seeks a set of model parameters minimizing the objective function. However, it is in practice an ill-posed problem (Tarantola, 1984a; Scales et al., 1990; Virieux and Operto, 2009). The solution is not unique, meaning that there are probably several models that can perfectly explain the data for a given noise level. Some regularizations are conventionally applied to make the inverse problem better posed (Tikhonov et al., 1977; Menke, 1984; Hansen, 2000; Asnaashari et al., 2013). The non-uniqueness problem results from various reasons including the imperfect acquisition and the crosstalk among different physical parameters (Tarantola, 2005). Most of the seismic surveys are located at the surface of the Earth and the offsets are limited, leading to insufficient illumination of the subsurface, especially in the area with complex geological structures (e.g. salt body and gas clouds). Consequently, some of the subsurface model parameters may be poorly or even not sampled, such that they cannot be retrieved by analyzing the measurements. For example, the local change of seismic velocity and attenuation property may produce similar reflection coefficients (Mulder and Hak, 2009; Hak and Mulder, 2010). To reduce the crosstalk, a proper parameterization is needed, for which the model parameters are less coupled (Virieux and Operto, 2009; Zhou et al., 2015; He and Plessix, 2017).

Optimization scheme

Inverse problem aims at determining a set of model parameters minimizing the objective function through an optimization procedure. One possible solution is the global optimization methods (also called the global search methods) (Sen and Stoffa, 2013), which explore the whole model space to determine the optimum choice. Such schemes require the ability to calculate the value of the objective function and to properly sample the model space. For example, Monte Carlo methods (Jin and Madariaga, 1994; Sambridge and Mosegaard, 2002), simulated annealing (Ingber, 1989; Mosegaard and Vestergaard, 1991; Scales et al., 1992; Misra and Sacchi, 2008) and generic algorithms (Gallagher et al., 1991; Sambridge and Drijkoningen, 1992; Jin and Madariaga, 1993) have been applied to the geophysical inverse problems. The main limitation of these schemes is the computational cost because they require evaluating the objective functions numerous times.

On the other hand, a realistic alternative is the local optimization methods (also called gradient-based methods), applicable if the objective function does not contain local minima. Examples of this family are steepest descent methods (Lines and Treitel, 1984; Tarantola, 1984a), nonlinear conjugate gradient methods (Fletcher and Reeves, 1964; Portniaguine and Zhdanov, 1999; Luo and Schuster, 1991), Gaussian-Newton methods (Shin, 1988), quasi-Newton methods (Nocedal, 1980; Nash and Nocedal, 1991) and Newton methods (Santosa et al., 1987; Pratt et al., 1998; Métivier et al., 2013). For these schemes, the model is iteratively updated such that the value of the objective function decreases with iterations (Figure 1.5). The direction of the model update at each iteration is determined by calculating the gradient of the objective function with respect to model parameters, and possibly the inverse of the Hessian to take into account the local curvature of the objective function. The adjoint-state technique (Chavent, 1974; Plessix, 2006) provides an efficient way to derive of the gradient of the objective function with respect to model parameters.

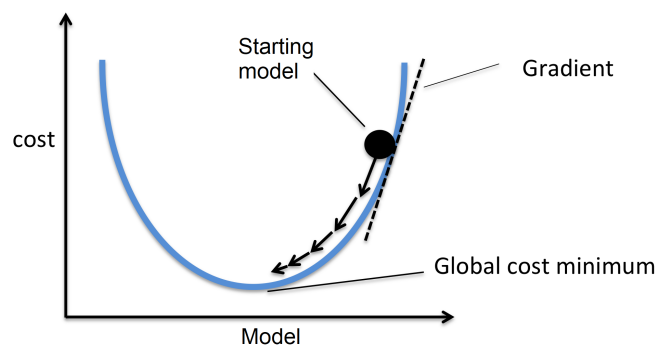


Figure 1.5 – Schematic for local optimization methods.

The nonlinear relationship between data and model parameters is an intrinsic difficulty for the local optimization procedure: the objective function is not necessarily convex and it may converge towards a local minimum instead of the global one (Bunks et al., 1995). Therefore, it is important to start from an initial model close enough to the true model (Beydoun and Tarantola, 1988; Pratt et al., 2008). The alternative is to develop a multi-scale approach and to start with the low-frequency content of the data. The presence of low-frequency content in the observed data can relax the requirement for the initial model (Sirgue, 2006): for example, the classical least-squares objective function has an enlarged basin of attraction. I will explain the nonlinear issues more in section 1.2.1. Strategies to define the objective functions with better convex property will be reviewed in sections 1.2.3 and 1.3.

I have reviewed the seismic imaging principle in this section. In the following section, I review the inverse problems formulated in data-domain and image-domain, as I will investigate the links between two families later in Chapter 4. Although modeling aspects with ray theory were reviewed in section 1.1.3, I will only investigate the wave-equation-based methods in the following Chapters, except for analyzing the wave-equation operators. The rays can emphasize the kinematics of wave propagation in the subsurface. However, ray-based methods poorly behave for sampling the shadow zone (a low velocity anomaly area) (Moser, 1991). Ray-based methods have built-in disadvantages including the limited resolution of recovered model (a ray travels through a infinitesimally narrow path inside the model) and the instability dealing with strong heterogeneities (Díaz et al., 2014; Wang et al., 2014). I now detail the classic inverse problems defined in the data-domain, and the principle of image-domain techniques will be reviewed in section 1.3. Note that a tomographic mode corresponds to the update of macromodel c_0 and a migration mode to the update of model perturbation δc as mentioned in section 1.1.2.

1.2 Data-domain methods

The data-domain methods measure the misfit between observed and modeled data to evaluate the quality of an estimated model (Tarantola, 2005). Under the isotropic acoustic approximation, the model c corresponds to the pressure wave velocity field. In section 1.2.1, I first review the FWI strategy, which uses the full recorded data to reconstruct c without any scale separation. Then, in section 1.2.2, I introduce the migration technique as a linearized waveform inversion problem. The method relies on the scale separation assumption and is dedicated to recovering the small scale structure δc . The reconstruction of the background model c_0 is challenging. Traveltime tomography has become a standard for estimating the macro velocity model c_0 in the oil and gas industry since last 90s (Woodward et al., 2008; Nichols, 2012; Lambaré et al., 2014). However,

it is a ray-based technique and will not be discussed here. I will detail alternative wave-equation-based strategies which behave in a tomographic mode in section 1.2.3. More technical details will be provided in Chapter 2.

1.2.1 Full Waveform Inversion

Formulation and resolution

Full Waveform Inversion (FWI) is a technique for seismic imaging which has gained popularity with the increase of computer powers. It defines a nonlinear inverse problem in the data-domain, which seeks a set of model parameters minimizing the misfit between observed and modeled data in the least-squares sense. Originally, Lailly et al. (1983); Tarantola (1984a) introduced the data misfit as a similarity estimation in the time-domain. Alternatively, the problem is formulated in the frequency-domain Pratt and Worthington (1990). The frequency-domain FWI is equivalent to time-domain FWI when all frequencies are considered simultaneously (Pratt et al., 1998). The reader is referred to Virieux and Operto (2009); Fichtner (2011) for more details.

FWI considers the full wavefields, and the recovered subsurface image should include both the large and small scale structures of the velocity model. The contribution of different types of waves in the estimation of the velocity model can be analyzed via the following relationship (Devaney, 1982; Miller et al., 1987),

$$\mathbf{k} = \frac{2\omega}{c} \cos\left(\frac{\theta}{2}\right)\mathbf{n}, \quad (1.2)$$

linking the wavenumber \mathbf{k} at an spatial point \mathbf{x} to the angular frequency ω and to the opening angle θ related to a source-receiver pair; \mathbf{n} is the normalized vector of \mathbf{k} (Figure 1.6). The low-frequency content in data can thus induce small-wavenumber (i.e. long-wavelength) updates of the model, whereas the high-frequency corresponds to the large-wavenumber (i.e. short-wavelength) updates. On the other hand, the data from long-offset is related to large scattering angles and can result in small-wavenumber updates. Hence, low frequencies and large offsets are preferred for a tomographic update (Alkhalifah and Plessix, 2014). However, the recorded frequencies and surface offset ranges are limited in practice. In the shallow part of the subsurface, the diving waves correspond to large diffraction angles, allowing to retrieve small-wavenumber contents of the subsurface model. Differently, the deep part is mainly sampled by reflected waves with a small diffraction angle due to the limited surface offset range. Thus, only the high-wavenumber part of the model can be recovered, meaning that FWI behaves in a migration mode in this case (Mora, 1989). In section 1.2.3 and 1.3. I will present some approaches that can extract the long-wavelength part of the velocity model from the reflected data.

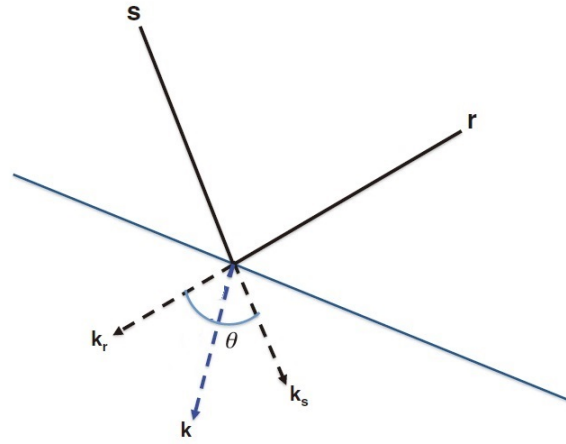


Figure 1.6 – Illustration of the relationship between the wavenumber $\mathbf{k} = \mathbf{k}_s + \mathbf{k}_r$ and the opening angle θ (adapted from Alkhalifah and Plessix, 2014).

Multi-parameter inversion

FWI includes in principle all types of waves (both transmitted and scattered waves) for resolving different model parameters (velocity, density, anisotropy, attenuation). The multi-parameter estimation requires a modeling engine capable of reproducing the physics of wave propagation as accurate as possible (Warner et al., 2013). This is a first difficulty, since the real Earth is a complex poro-visco-elastic medium, and the implementation is more computational demanding than for the acoustic approximation of the wave equation. The simultaneous inversion of several parameters is a more severe challenge, as different parameters may have coupling effects and sensitivities with various orders of magnitude, making the inversion poorly conditioned. Operto et al. (2013); Prieux et al. (2013a); He and Plessix (2017) proposed guidelines to design multi-parameter strategies for FWI, typically by only inverting for v_p in a first step. In practice, many approaches only invert for v_p , with other fixed model parameters, because of the difficulties for multi-parameter estimation (Operto et al., 2013).

Cycle skipping effects

It is well known that the FWI objective function suffers from local minima due to the nonlinear relationship between data and model parameters (Gauthier et al., 1986). This issue is so-called *cycle-skipping effects* (Bunks et al., 1995). In theory, global optimization methods may properly deal with this problem, but this is not yet a realistic choice for FWI due to the heavy computational cost, except recent attempts in 2D (Sambridge and Mosegaard, 2002; Misra and Sacchi, 2008; Sen and Stoffa, 2013). local optimization methods are less expensive but they require a starting model close enough to the

true model (Beydoun and Tarantola, 1988; Pratt et al., 2008). More precisely, the phase mismatch between observed and modeled data should be less than half a period recorded in the data. For example, Figure 1.7 presents three examples with different phase shifts in data: less than half a period, equal to zero and more than half a period. The first example mimics the data generated from the model sufficiently close to the true model and can converge towards the global minimum, whereas the third one is cycle skipped. The presence of low-frequency contents in observed data can relax the requirement for the initial model (Sirgue, 2006), as the associated objective function has an enlarged basin of attraction around the correct velocity (Figure 1.8). In theory, the basin of attraction of the misfit function is inversely proportional to the central frequency of the data (Bunks et al., 1995; Pratt et al., 1996; Sirgue and Pratt, 2004).

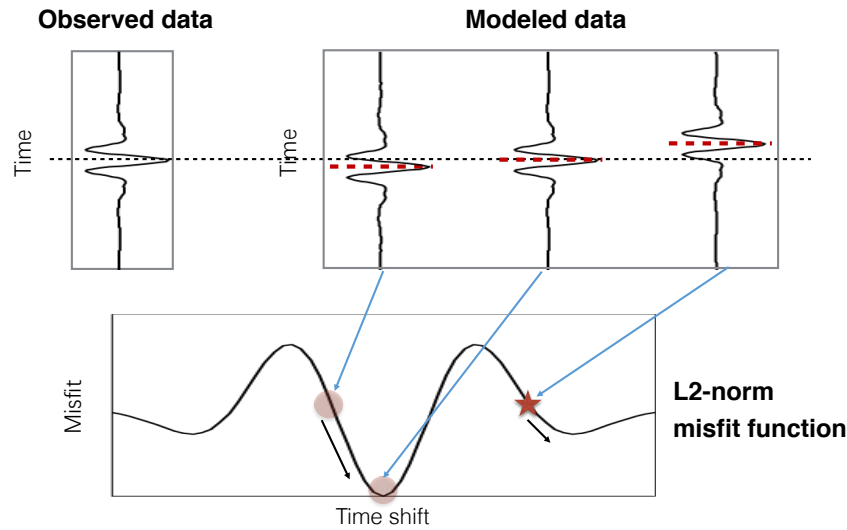


Figure 1.7 – Illustration of the origin of cycle skipping effect. Local optimization converges towards the global minimum for the modeled data with smaller time shift error (left pink circle), whereas it converges towards the local minimum for the modeled wavelet with larger time shift error (red star). The black arrows mark the descent directions.

To deal with this issue, the multi-scale strategy can mitigate the nonlinearity, by progressively performing the inversion on data from low-frequency to high-frequency and from long-offset to short-offset (Bunks et al., 1995; Shipp and Singh, 2002; Sirgue, 2006; Virieux and Operto, 2009; Wang et al., 2014). As already mentioned in the resolution part, data recorded with the low-frequency and large-offset allow recovering the long-wavelength part of the velocity model, such that the attraction basin of the objective function can be enlarged. In addition, the large-offset recording introduces more transmitted waves in data (Shipp and Singh, 2002). This multi-scale strategy has been successfully applied to real data in marine (Sirgue et al., 2009) and land (Plessix et al.,

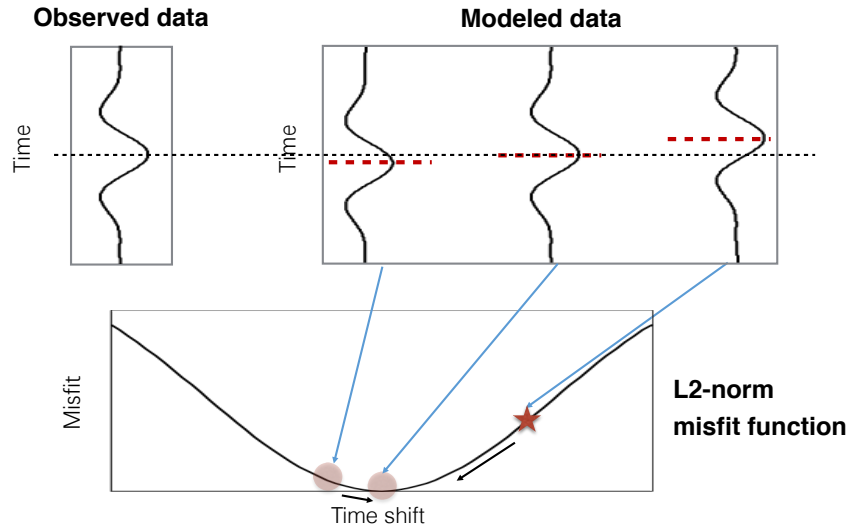


Figure 1.8 – Same as for Figure 1.7, but the low-frequency content is involved in the observed and modeled data. The cycle skipped case (red star) in Figure 1.7 now converges towards the global minimum with the help of low-frequency.

2010) cases. However, the required low frequencies (< 1 Hz) for exploration problems are very difficult to acquire during controlled-source experiments (Virieux and Operto, 2009), and the large-offset acquisition is expensive since the penetration depth of diving waves is only one third to one sixth of the largest offset, depending on the experience (Zhou, 2016).

Alternatively, the definition of the FWI objective function can be modified to enlarge the basin of attraction around the correct velocity model, avoiding falling into the local minimum. Several alternative methods will be detailed in section 1.2.3.

1.2.2 Linearized waveform inversion

I first introduce the migration technique as a linearized waveform inversion problem. The migration approach developed later in Chapter 3 is a wave-equation-based scheme, but I will review some ray-based migration techniques in this section, as ray theory has played an essential role in the historical development of migration, including recently for the derivation of direct approaches.

Classic migration

The purpose of migration techniques is to recover the reflectivity image as velocity perturbations of the Earth's interior. These methods are based on the scale separation concept and assume that the macro velocity model is known. They were mainly designed

for primary reflections, and can be extended to multiple reflections (Guittou, 2002; Jiang et al., 2005; Muijs et al., 2007; Cocher et al., 2015).

The early migration techniques are based on ray theory. The ray-based family mainly refers to ray+Kirchhoff (French, 1975; Schneider, 1978; Gray, 1992) or ray+Born (Beydoun and Mendes, 1989; Thierry et al., 1998; Operto et al., 2000) migration. The former is an integral approximation to the wave equation, explicitly expressing how recorded traces contribute to the reflectivity images (Figure 1.9). The ray+Born approximation differs from ray+Kirchhoff method by the description of the perturbation components of the model (i.e. in terms of velocity and density perturbations or in terms of specular reflectivity, respectively) (Lambaré et al., 2003; Operto et al., 2003). Ray-based methods are flexible and efficient, but rays are only asymptotic solutions of the wave equation, such that they are not always adequately effective in imaging extremely complex structures, for example in the presence of hard layers and salt bodies.

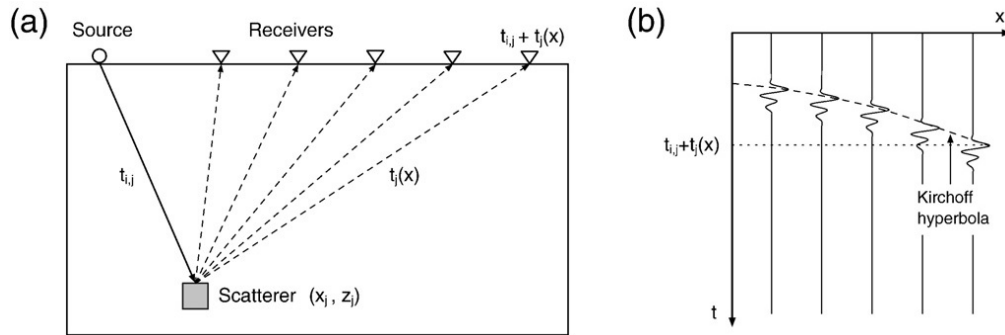


Figure 1.9 – The principle of Born modeling: (a) calculation of traveltime; (b) Kirchhoff hyperbola on seismogram. (adapted from Pyun and Shin, 2008)

Later, reflectivity imaging was reformulated as wave-equation-based migration techniques. These approaches can be further divided into one-way wave-equation migration (namely wavefield extrapolation migration, WEM) (Bamberger et al., 1988; Ehinger et al., 1996) and two-way wave-equation reverse time migration (RTM) (McMechan, 1983; Baysal et al., 1983), depending on how the lateral variations of velocities are introduced. WEM methods use paraxial approximation of the wave equation. WEM is properly defined within certain angles around the main direction but fails to handle wider angles, especially those near or beyond 90° . RTM solves the full wave equation and shows its superiority over other methods in dealing with steep dipping angles and complex velocity models (Mulder and Plessix, 2004). The wave propagation in the subsurface can be characterized more accurately using the two-way propagator compared to the case of one-way. (Mulder, 2008; Gao and Symes, 2009). Biondi and Shan (2002) showed the potential of RTM for imaging overturned reflections.

These wave-equation-based methods are commonly implemented through *imaging*

condition concept introduced by Claerbout (1971). The reflectivity model is updated where a downgoing incident wavefield coincides in time and space with an upgoing receiver wavefield (Figure 1.10). The procedure consists of three steps: (1) propagate the source wavelet from the source position to determine the downgoing source wavefield; (2) backpropagate the observed data from the receiver position to determine the upgoing receiver wavefield; (3) apply the imaging condition to determine the reflectivity image of the subsurface. Many formulations exist for the imaging condition and the most common one is the zero-lag crosscorrelation of the source and receiver wavefields. The procedure is repeated for every shot, with a summation over sources to reduce the signal to noise ratio, or with an image per shot. On the other hand, the consistency among different images can also be used as information. This is the basic principle of MVA techniques presented in section 1.3.

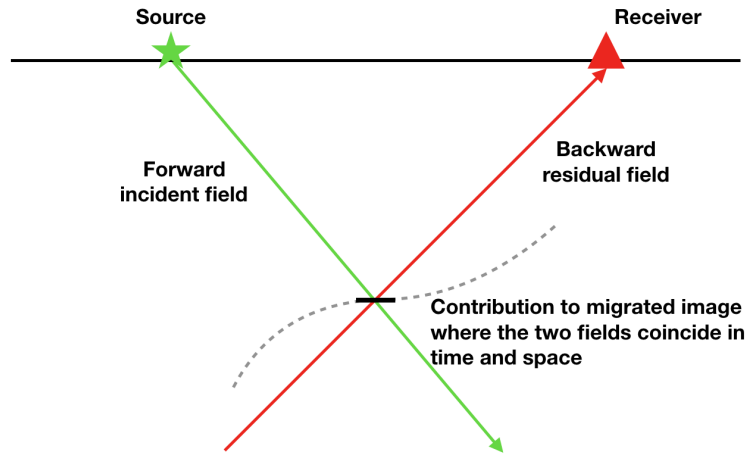


Figure 1.10 – Description of imaging condition for a simple reflector.

This migration algorithm was recognized as the first iteration of FWI, the least-squares data-fitting inverse problem, by Lailly et al. (1983); Tarantola (1984b) (Figure 1.11). Under the first-order Born approximation, it is formulated as a *linearized waveform inversion* (Østmo et al., 2002). Assuming the background velocity model c_0 is known, it aims at determining the model perturbation δc that best reproduces observed data. Compared to FWI introduced in the previous section, the forward modeling is linear in the migration case.

The conventional migration is in fact the adjoint operator of the Born modeling (Claerbout, 1992). Although the adjoint is useful to qualitatively estimate the perturbation model, it is not exactly an inverse and cannot correctly estimate the amplitudes of migrated images. Consequently, one cannot accurately reproduce the amplitudes of seismic data from the migrated images, even if a correct background model is given. In

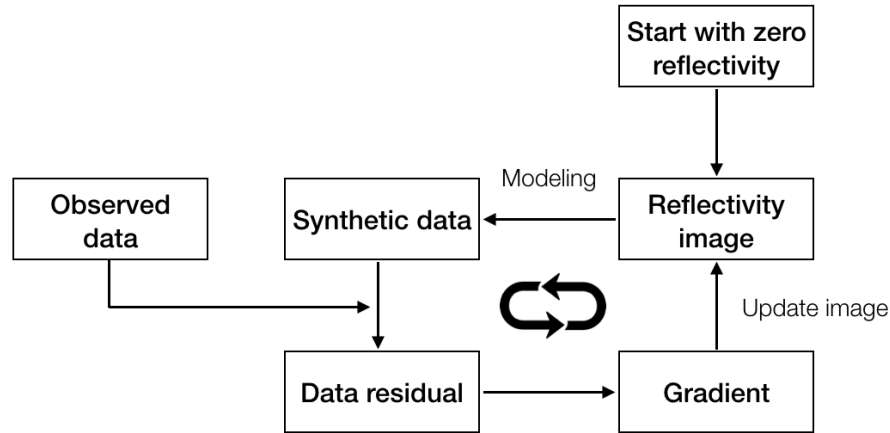


Figure 1.11 – Description of iterative migration.

practice, the reproduced data is only kinematically correct, in the sense that the major event is located at the same position as the original data, but the amplitudes are of different magnitudes (Figure 1.12). Moreover, the reproduced data also suffers from artificial direct arrivals (marked by arrows in Figure 1.12c) related to the migration artifacts (marked by arrows in Figure 1.12b). Classical migration methods lead to artifacts, such as migration smiles, due to the limited aperture in observed data. Briefly, the migrated images are not the solution minimizing the objective function for linearized waveform inversion.

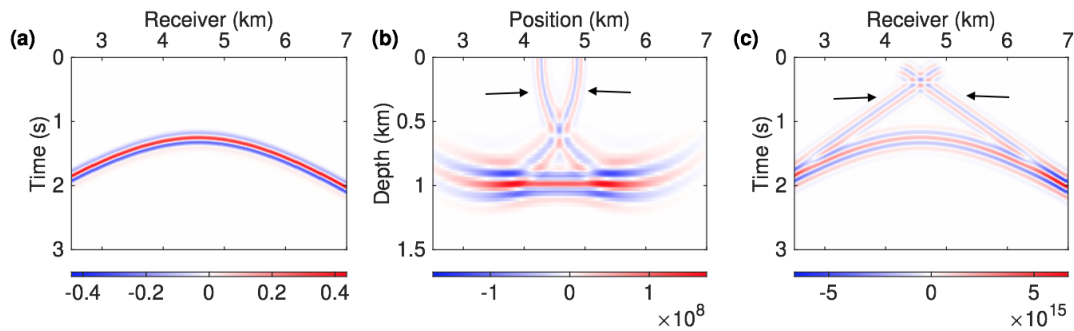


Figure 1.12 – Illustration indicating that migration cannot completely reproduce the data: (a) observed data, (b) migrated image with correct macro velocity model, and (c) modeled data with image (b) in the correct macro velocity model. The artificial direct arrivals in panel (c) correspond to migration artifacts in panel (b). The observed and modeled data are in phase, but their amplitudes are different.

True-amplitude migration: iterative migration versus direct inversion

Standard migration methods are designed to produce kinematically accurate images of the subsurface, but are not aimed at preserving the amplitudes in migrated images. The quantitative methods are preferred as they compensate for geometrical spreading losses, uneven illuminations and deconvolution. The quantitative migration can be retrieved by seeking a perturbation model in which the modeled data best fit the observed data in the least-squares sense (Ronen and Liner, 2000). To solve the minimization problem, a first solution is to use the inverse of the Hessian matrix, the second derivatives of the objective function with respect to model parameters, which is prohibitively expensive to compute and to store directly. It is more feasible, either to use a gradient-based iterative algorithm as in the FWI approach, or to directly solve the inverse problem by approximating the impact of the Hessian. This leads to two types of true-amplitude migration methods: *iterative migration* versus *direct inversion*.

The initial implementation of iterative migration was carried out with ray-based migration (Lambaré et al., 1992; Nemeth et al., 1999; Duquet et al., 2000). Then, it was developed for WEM (Duquet, 1996; Kühl and Sacchi, 2003; Clapp et al., 2005), and more recently for RTM (Dai et al., 2012; Liu et al., 2013; Zhang et al., 2014a; Xue et al., 2014). Such approaches can be extended from acoustic to elastic (Forgues and Lambaré, 1997; Yan and Sava, 2008) and to visco-elastic (Ribodetti and Virieux, 1998; Dutta and Schuster, 2014). It has been demonstrated that iterative migration significantly improves the resolution of the migrated image and effectively attenuates migration artifacts (Figure 1.13), especially in the case of incomplete observed data (Nemeth et al., 1999; Kühl and Sacchi, 2003) and of complex geology Zeng et al. (2017). However, the strategy is expensive in terms of computational cost as the optimization usually requires several iterations to converge. Thus, it is very computationally demanding to couple iterative migration to macro velocity estimation techniques (Zhou, 2016; Cocher et al., 2017b), and in practice the proper regularization and preconditioner are required for a faster convergence (Fomel, 2007; Dai et al., 2012; Cocher et al., 2017b).

Direct inversion replaces migration, the adjoint of the Born modeling operator, by an explicit pseudo-inverse formula. Assuming an infinite recording aperture, these schemes derive the formula represented by an integral equation similar as in the migration case. This inverse formula makes the Hessian matrix close to a Dirac function leading to a better conditioned problem. The quantitative weight can either be directly used (Beylkin, 1985; Bleistein, 1987; ten Kroode, 2012; Hou and Symes, 2015), or act as a preconditioner in conjunction with iterative migration to boost the convergence (Qin et al., 2015; Hou and Symes, 2016b; Duprat and Baina, 2016; Cocher, 2017), leading to a more efficient application. Direct inversion was initially proposed for ray-based methods (Beylkin, 1985; Bleistein, 1987; Lambaré et al., 1992; Lameloise, 2015). Later, it was developed for WEM (Zhang et al., 2005, 2007; Joncour et al., 2011), and more recently for RTM (Op't Root et al., 2012; ten Kroode, 2012; Zhang et al., 2014a; Hou

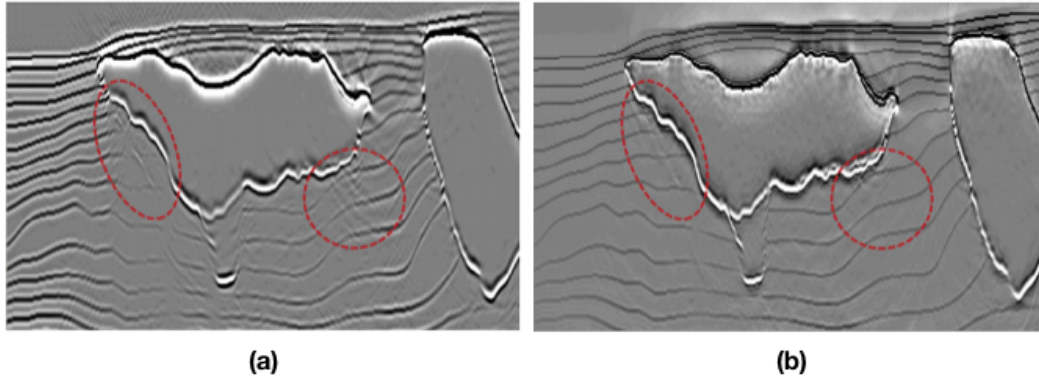


Figure 1.13 – Conventional RTM (a) and least-squares RTM (b) images of the synthetic salt model based on a GOM survey (adapted from [Zeng et al., 2017](#)).

and Symes, 2015; Lameloise, 2015; Qin et al., 2015; Duprat and Baina, 2016; Chauris and Cocher, 2017). Table 1.1 presents a non exhaustive review of published direct inversion approaches. Note that some of the direct inversion schemes require an accurate background velocity model ([Zhang et al., 2007](#); [Zhang and Sun, 2009](#)), and the others do not ([ten Kroode, 2012](#); [Hou and Symes, 2015](#); [Li and Chauris, 2017](#)). For the latter, the model space is extended with an extra parameter (e.g. shot position) such that the dimensions for model and data spaces are equivalent.

For the works of [ten Kroode \(2012\)](#); [Hou and Symes \(2015\)](#); [Qin et al. \(2015\)](#); [Chauris and Cocher \(2017\)](#), the asymptotic of ray theory justifies the derivation of the inverse formula, but no ray quantities are present in the final result. [Bleistein et al. \(2005\)](#) explained the remarkable fact: the calculation of Beylkin determinants (Jacobian of transformation between model parameters and acquisition surface coordinates) produces the reciprocal geometrical amplitudes and other ray-dependent terms. It is a way to remove the effect of geometrical spreading. In this thesis, I develop a common-shot direct inversion in Chapter 3 and then couple it to various velocity estimation techniques for a better robustness in Chapters 3 and 4.

1.2.3 Alternative methods

Alternatives to FWI objective function

To get rid of the cycle-skipping effects, alternative techniques have been developed, modifying the definition of the classic FWI least-squares objective function to enlarge the basin of attraction around the correct velocity model. In this subsection, I consider methods that do not rely on the scale separation. Many researchers have proposed alternative functions to converge towards the global minimum in the absence of low-frequency and long-offset information ([Van Leeuwen and Mulder, 2010](#); [Bozdağ et al.,](#)

Table 1.1 – Non exhaustive review of published direct inversion approaches. For subsurface-offset and common-angle domains, the reader is referred to section 1.3.2.

	Subsurface-offset or common-angle domain	Surface-offset or common-shot domain
Ray-based	Xu and Lambaré (2004) ; Lameloise (2015) , etc.	Beylkin (1985) ; Bleistein (1987) ; Lambaré et al. (1992) , etc.
One-way wave-equation-based	Zhang et al. (2007) ; Joncour et al. (2011) , etc.	Zhang et al. (2005) , etc.
Two-way wave-equation-based	ten Kroode (2012) ; Zhang et al. (2014a) ; Hou and Symes (2015) ; Chauris and Cocher (2017) , etc.	Op’t Root et al. (2012) ; Qin et al. (2015) ; Duprat and Baina (2016) ; Li and Chauris (2017) , etc.

2011; [Liu et al., 2011b](#); [Wu et al., 2014](#); [Warner and Guasch, 2016](#); [Métivier et al., 2016](#)). There are two distinguished principles to design such an objective function ([Bharadwaj et al., 2015](#)): (1) give more weight to the kinematic error, or (2) transfer the data to a reduced form (e.g. envelope). Both aim at inferring a tomographic update for velocity model. They allow to mitigate the cycle-skipping issues but the retrieved model generally has a lower resolution than conventional FWI.

[Van Leeuwen and Mulder \(2010\)](#); [Luo and Sava \(2011\)](#); [Chi et al. \(2015\)](#) proposed the crosscorrelation-based objective function by computing the crosscorrelation factor between observed and modeled data. The correct velocity model is associated to focusing energy at zero temporal delay. This function does not properly handle multiple arrivals. [Bozdağ et al. \(2011\)](#); [Wu et al. \(2014\)](#); [Chi et al. \(2014\)](#) proposed to design the objective function evaluating the misfit between observed and modeled data after the application of the envelope transform. The major advantage is that the reduced seismic signals are less oscillating, reducing cycle-skipping effects. But the technique is sensitive to small amplitude noise. [Warner and Guasch \(2014, 2016\)](#) developed Adaptive Waveform Inversion (AWI), defining an objective function based on the non-causal Wiener filter. Wiener coefficients are retrieved by deconvolving modeled data with observed data. The method penalizes the Wiener coefficients with a weighting factor defined as the temporal lag. Since the adjoint source of AWI is mathematically in phase with modelled data, an immunity to cycle skipping effect is then expected. [Warner and Guasch \(2016\)](#) showed that AWI can properly deal with synthetic data dominated by the reflected arrivals. They further indicate that Wiener filter based AWI is less sensitive to errors in the source wavelet than conventional FWI. [Luo and Sava \(2011\)](#) compared the behaviors between crosscorrelation-based and deconvolution-based objective functions. [Bharadwaj et al. \(2015\)](#) came up with a strategy using an auxiliary bump functional

for inversion. The approach squares and subsequently smooths the data to enhance the long-period content of the data. The bump function is not sensitive to the sign of data, such that it is only used as a guide for conventional FWI to converge at a global minimum. More recently, [Métivier et al. \(2016\)](#) introduced the optimal transport distance to evaluate the data residual. More technical aspects about these methods will be provided in Chapter 2.

Wave-equation traveltimes tomography

Wave-equation traveltimes tomography method (WETT) extends the conventional traveltimes tomography by taking into account the finite-frequency nature of seismic data. The readers are referred to [Woodward et al. \(2008\)](#); [Lambaré \(2008\)](#); [Lambaré et al. \(2014\)](#) for recent reviews of conventional traveltimes tomography techniques, as the focus of this thesis is related to waveform inversion. [Luo and Schuster \(1991\)](#) proposed to crosscorrelate the first arrival events in the observed and modeled data trace by trace, to extract the temporal lags maximizing the crosscorrelation. Then they seek a velocity model minimizing these temporal lags in the least-squares sense. This procedure is similar to minimizing the picked traveltimes residuals, but the traveltimes error might not be equivalent to the extracted time-lag because of possible errors in the source wavelet used for modeled data ([Van Leeuwen and Mulder, 2010](#)). WETT uses the same forward modeling kernel and optimization scheme as FWI, whereas the objective functions are different (time-lag versus data misfit). Although WETT behaves in a tomographic mode similar to the conventional traveltimes tomography, it is more realistic due to the wave-equation-based modeling and the finite-frequency attribute ([Zhao et al., 2000](#)). WETT proves to be stable in the presence of low velocity anomalies in the model ([Priolo and Chiaruttini, 2003](#)).

Reflection Waveform Inversion

Reflection Waveform Inversion (RWI) is a technique dedicated to extract the macro velocity model from reflection data ([Xu et al., 2012](#); [Brossier et al., 2015](#); [Wu and Alkhalifah, 2015](#); [Zhou et al., 2015](#)). It is derived from migration based traveltimes tomography (MBTT) ([Chavent et al., 1994](#); [Plessix et al., 1995](#)). RWI either assumes the scale separation between short-wavelength and long-wavelength parts of the velocity model ([Xu et al., 2012](#)) or relies on the parameterization of the impedance and the velocity models to naturally separate different scales ([Zhou, 2016](#)). The short-wavelength and long-wavelength parts are updated with an iterative relaxation method. The reflectivity is first inverted in a given macro velocity model with only the short-offset data. Then, the modeled data with longer offset are produced from the inverted reflectivity. Finally, the macro velocity model is updated by minimizing the misfit between observed and predicted reflection data. One repeats these steps until convergence ([Zhou, 2016](#)). RWI

aims at reproducing the data from the reflectivity model such that a true-amplitude migration scheme must be considered. The conventional migration (an adjoint operator) is not enough as it can only reproduce the kinematics of the data without properly dealing with the amplitudes, even if the given macro velocity model is correct.

In terms of resolution, [Alkhalifah and Wu \(2017\)](#) indicate that RWI is similar to the velocity analysis scheme proposed by [Symes and Carazzone \(1991\)](#), behaving in a tomographic mode. RWI retrieves the information about the macro model between the image points and the surface, allowing to recover the long-wavelength part of the velocity model in deep areas only reached by reflected data. RWI can be extended to elastic ([Guo and Alkhalifah, 2016](#)) and multiple ([Staal, 2015; Zhou, 2016](#)) cases.

Differential waveform inversion

Differential Waveform Inversion (DWI) is a technique to assess the macro velocity model quality using the reflection data ([Chauris and Plessix, 2012, 2013](#)). The approach is formulated in the data domain and relies on the scale separation assumption. As for RWI, it consists of two main steps: an image section is migrated from a given shot gather and used to calculate synthetic data for the next shot; then the macro model is updated by minimizing the difference between the predicted shot and the observed shot at the next shot position. These two steps formulate a nested optimization loop. This procedure is inspired by the image quality assessing criteria established for image-domain methods presented in section 1.3. In Chapter 4, I investigate the relationship between objective functions defined in the data-domain and the image-domain along the same line as [Chauris and Plessix \(2012\)](#).

DWI is based on a combination of successive migration and demigration steps. In that perspective, it is similar to the MBTT method ([Chavent et al., 1994; Plessix et al., 1998](#)), to the crosscorrelation method ([Van Leeuwen and Mulder, 2010](#)) and to the approach proposed by [Staal and Verschuur \(2012\)](#). More importantly, it is similar to RWI due to several reasons: (1) both methods rely on the scale separation assumption; (2) true-amplitude migration is essential for both methods; (3) they all explicitly or implicitly introduce the image domain coherency criteria into data-domain; (4) DWI is a nested optimization procedure, whereas RWI ([Xu et al., 2012](#)) assumes the reflectivity and macromodel are independent. The alternative of DWI will be investigated in Chapter 4 to discuss the possible links between image-domain and data-domain methods. The differences between RWI and DWI will also be discussed in Chapter 4.

1.3 Image-domain methods

I present in this section the inverse problem formulated in image-domain. This is the class of methods which will be mainly investigated in this thesis. Image-domain meth-

ods rely on the fact that seismic data are redundant. Given an macro velocity model, a set of migrated images can be constructed with different subsets of the data, for example one image corresponding to a shot gather (Al-Yahya, 1989). With a correct macro model, the recovered reflectivity images corresponding to different shot experiments should be kinematically coherent, in the sense that the reflectors/diffractors in different images should be located at the same position of the subsurface if they correspond to the same structures (Al-Yahya, 1989). Otherwise, the incoherencies among different images are attributed to an incorrect macro velocity model. Migration Velocity Analysis (MVA) is the family of techniques based on this principle (Symes, 2008).

Historically, MVA emerged as an extension of the Normal Move Out (NMO) correction procedure. The data set is first sorted into panels called common midpoint gathers, depending on surface-offset and time. Assuming the Earth's structure is horizontally homogeneous, in a common midpoint gather, the reflection traveltimes are a function of surface-offset following hyperbolic shapes. The process of NMO correction seeks a 1D velocity model that best explains the hyperbolas. The reader is referred to Yilmaz (2001) for more details. This method is limited to the case of simple models and Al-Yahya (1989) extended it to the image-domain to handle more complex structures.

Many wave-equation based MVA approaches have been proposed (Biondi and Sava, 1999; Sava and Biondi, 2004; Biondi and Symes, 2004; Mulder, 2008, among others). MVA aims to retrieve the macro velocity model, relying on the scale separation and the Born approximation. According to the formalism introduced by Symes (2008), different MVA approaches can be distinguished along these lines:

- *the choice of redundant parameter*: the data is migrated to the image-domain parameterized with an additional parameter representing the redundancy of seismic data. For example in 2D, the migrated sections $\delta c(x, z, s)$ are parameterized by lateral position x , depth z , and the redundant parameter – shot position s . I discuss in section 1.3.2 alternatives for the redundant parameters;
- *the choice of a coherency or focusing criterion* assessing the image quality: it depends on the choice of the redundant parameter. For example, the coherency principle will be considered if the redundant parameter corresponds to the shot position, and the analysis is performed on the panel called Common Image Gather (CIG) representing a section of the reflectivity volume a fixed lateral position as a function of depth and the extra parameter. The observed data provide a natural reference for assessing the quality of velocity model in the data-domain, whereas there is no such obvious reference in the image-domain. Instead, the objective function is defined based on the coherency or focusing criterion, and the partial derivative of this function with respect to the image is commonly called *image residual* (versus data misfit).

The technical aspects will be detailed in Chapter 2.

In the following, I present the principle of two MVA families: the redundant parameter is related to the acquisition parameters in the *surface-oriented* approach, whereas the extra parameter is an artificial parameter introduced during the construction of the migrated image in the *depth-oriented* approach.

1.3.1 Surface-oriented MVA

A choice for the redundant parameter related to the acquisition is the shot position (Al-Yahya, 1989; Symes and Kern, 1994; Huang and Symes, 2015) or the source-receiver offset (distance between source and receiver) (Chauris and Noble, 2001; Rickett and Sava, 2002; Mulder and ten Kroode, 2002). The image volume can be constructed from different subsets of seismic data, such as common-shot or common-offset gather. The surface-oriented CIGs are then built by collecting images upon the redundant parameter. The idea of surface-oriented MVA is that images obtained from several experiments are supposed to be independent of the extra parameter. Consequently, events on these CIGs should be horizontal if they are constructed with the correct macro velocity model. An example of this procedure for the source-receiver offset case is presented in Figure 1.14.

Events in CIGs are not horizontal any more if the macro velocity model is incorrect. For a single horizontal reflector embedded in homogeneous background velocity model, they curve upward for a too low velocity model and downward for a too high velocity model. Then, CIGs can be analyzed by *semblance criterion* (Chavent and Jacewitz, 1995), which first stacks images over the extension parameter and then defines an objective function to measure the energy of the stack. For the correct macro velocity model, the images from different subsets of data are coherent and thus the stack energy is expected to be maximal, meaning that the best model indeed corresponds the maximum of the objective function. This objective function has a large basin of attraction around the correct macro velocity model. However, it may exhibit oscillations away from the correct model (Chauris and Noble, 2001). One can also correct the Residual Move Out (RMO) in prestack depth migrated common image point gathers for velocity analysis (Woodward et al., 1998; Xie and Yang, 2008).

Alternatively, Symes and Carazzone (1991) proposed Differential Semblance Optimization (DSO) strategy, defining an objective function that computes the derivative of images with respect to the redundant parameter, to measure if the events on CIGs are horizontal (Chauris and Noble, 2001). Compared to semblance criteria, the DSO objective function demonstrates a more convex behavior that allows the local optimization scheme converging towards the global minimum, at least for simple 1D models (Stolk and Symes, 2002; Van Leeuwen and Mulder, 2009). The DSO principle is justified by various studies in 2D models (Chauris and Noble, 2001; Mulder and ten Kroode, 2002; Mulder, 2008; Shen and Symes, 2008).

In practice, the coherency principle of MVA may break down due to imaging artifacts. For example, the migration of multiples produces artificial interfaces under the

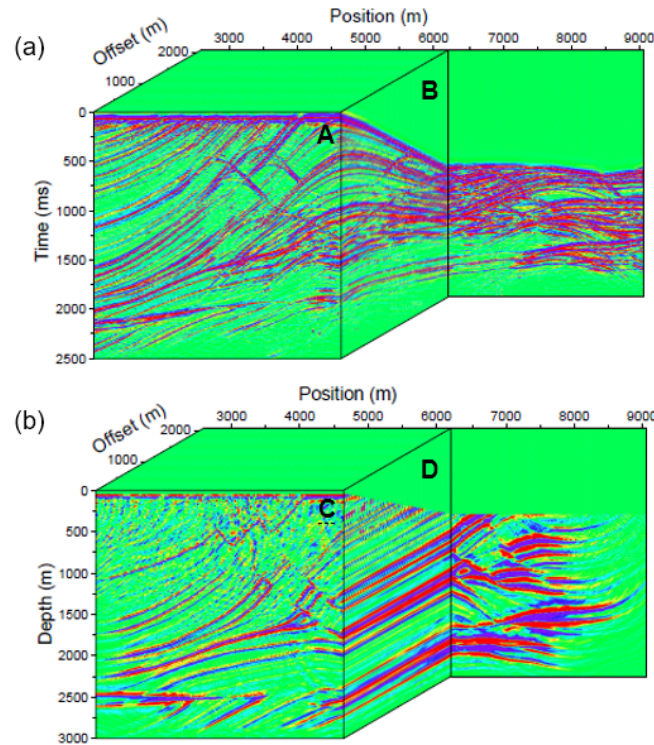


Figure 1.14 – Prestack common-offset data before (a) and after (b) migration. Seismic data set computed in 2D Marmousi model (a), migrated to surface-offset domain CIGs (b) by prestack depth migration using the true velocity model. A and C are common-offset gather and prestack migrated image, respectively, at zero-offset. B and D are common midpoint gather and CIG, respectively, at fixed position. Accurate velocity model results in flatness on the CIG panel D. (from [Chauris, 2000](#))

first-order Born approximation ([Verschuur and Berkhout, 2015](#); [Cocher, 2017](#)). The limited acquisition can result in uneven illuminations in migrated images, leading to incoherent stretching shapes of events in CIGs ([Xie et al., 2005](#)). The evaluation of the DSO objective function is sensitive to coherent noise and thus proper filtering is required ([Chauris and Noble, 2001](#)). The characterization of coherent events over the offset range can be tedious such that [Chauris et al. \(2002a\)](#) proposed to pick the slope of locally coherent events in CIGs to evaluate the quality of velocity model, and proved this to be equivalent to the slope tomography approaches ([Billette and Lambaré, 1998](#); [Billette et al., 2003](#); [Lambaré et al., 2004](#); [Lambaré, 2008](#); [Prioux et al., 2013b](#)). In surface-oriented methods, each image is constructed from only a subset of the data: kinematic artifacts may appear when complex wavepaths such as triplicated wavefields are involved in the problem ([Nolan and Symes, 1996](#); [Prucha et al., 1999](#); [Xu et al.,](#)

2001; [Stolk and Symes, 2004](#)). This issue will be discussed and further investigated for common-shot MVA in Chapter 3.

1.3.2 Depth-oriented MVA

In the depth-oriented formulation of MVA, the redundant parameter is not related to the acquisition but is introduced during the construction of the migrated images. The image space is *extended* by introducing, during the construction of the images, a space-lag ([Rickett and Sava, 2002](#); [Shen et al., 2005](#); [Sava and Vasconcelos, 2011](#); [Lameloise et al., 2015](#); [Chauris and Cocher, 2017](#)) or a time-delay ([Sava and Fomel, 2006](#); [Yang and Sava, 2011](#); [Sava and Vasconcelos, 2011](#)) as the extra parameter. The space-lag is commonly referred to as *subsurface-offset*. The main difference between the surface-oriented approach is that the input data cannot be split according to acquisition parameters. MVA formulated in the extended domain measures the image quality via the focusing principle, in the sense that the model is optimal when the energy of CIGs is focused at zero spatial and temporal delay. The reason is that the image section at zero delay corresponds to the image of the physical reflectivity obtained with classic migration after summation over all sources and receivers, meaning that the energy focused at non-zero value of extension parameters is not physical and can carry information about the inaccuracies of the macro velocity model.

In practice, the extension is usually made along a single extension parameter, to keep the migrated image size within a reasonable limit. The horizontal subsurface-offset is usually chosen ([Shen et al., 2005](#); [Shen and Symes, 2008](#); [Lameloise et al., 2015](#); [Hou and Symes, 2016a](#); [Alkhalifah and Wu, 2017](#)). In this case, the energy of CIGs spreads over non-zero subsurface offset for an incorrect macro model. Events curve downward (upward) for a too low (high) velocity model ([Mulder, 2014](#)).

Alternatively, a scattering angle can be considered as the extension parameter to build the angle-domain CIGs (ADCIGs) ([Xu et al., 2001](#); [Sava and Fomel, 2003](#); [Biondi and Symes, 2004](#)). Hence, the velocity model is updated by increasing the coherency of images across angles. Constructing ADCIGs with ray theory is straightforward, as the scattering angle can be easily deduced from the ray parameter ([Xu et al., 2001](#)). Alternatively, it is feasible to extract ADCIGs from the prestack wavefield downward-continued using the wave equation ([Prucha et al., 1999](#)). Subsurface-offset CIGs can be linked to ADCIGs through the Radon transform ([Sava and Fomel, 2003](#); [Montel and Lambaré, 2013](#); [Silvestrov et al., 2016](#)).

Despite an attractive formulation, there are not so many published applications on real data ([Chauris and Noble, 2001](#); [Mulder and ten Kroode, 2002](#); [Alkhalifah, 2005](#); [Shen and Symes, 2008](#); [Mulder, 2014](#)). MVA suffers from gradient artifacts caused by the limited acquisition and the complex structure (e.g. discontinuous reflectors and salt bodies) ([Fei and Williamson, 2010](#); [Chauris and Lameloise, 2014](#); [Lameloise, 2015](#)). More details about the gradient artifacts will be reported in Chapter 2. One solution is

too compensate for uneven illumination in the migration procedure. To attenuate the migration smiles related to the gradient artifacts, the classic migration can be replaced by the true-amplitude migration – the iterative migration or the direct inversion (Lameloise, 2015; Hou and Symes, 2016a; Chauris and Cocher, 2017; Cocher et al., 2017a). Both alternatives significantly mitigate the spurious oscillations presented in the gradient. This new approach is called Inversion Velocity Analysis (IVA). In terms of applications, the iterative migration is more numerically demanding. In this thesis, I will extensively discuss the IVA where the inversion is obtained by a direct (wave-equation-based) approach. Recently, MVA has been extended to transmission (Shen and Symes, 2013; Chauris et al., 2013; Lameloise and Chauris, 2016) and multiple (Cocher et al., 2017a) cases.

Compared to FWI formulated in data-domain, inverse problems formulated in the image-domain provides a more linear approach for estimating the macro velocity model. This results from the fact that the objective functions for MVA are more convex than for FWI. On the other hand, FWI is always capable of providing high-resolution results. It is of great interest to investigate the links between data-domain and image-domain methods for formulating a high-resolution strategy in a robust way. Among others, Biondi and Almomin (2012, 2014) proposed to combine FWI and MVA objective functions to formulate tomographic FWI. Symes (2017) introduced a similar approach to simultaneously invert perturbation and macro velocity model with a joint objective function. Note that DSO was initially defined as a regularization term for FWI. RWI and DWI, as mentioned in section 1.2.3, aim at determining the large-scale structures of the velocity model and therefore should be linked to MVA in some ways (Chapter 4). This study will discuss the possibility of extending MVA to data-domain in Chapter 4.

1.4 Motivations and thesis organization

In the previous sections, I have first introduced different aspects of seismic imaging: seismic data, scale separation, modeling and inversion. Then, I have detailed the seismic imaging methods, among others, formulated in both data-domain and image-domain. The thesis mainly focuses on the wave-equation-based methods under the constant-density isotropic acoustic assumption. There are two major objectives: (1) *seeking a robust and efficient velocity estimation approach*; (2) *investigating the possible links between image-domain and data-domain methods*. For the former, in particular, I consider common-shot MVA, with the extension to Inversion Velocity Analysis (IVA) by coupling direct inversion to velocity analysis. For the latter, I further extend the IVA approach from the image-domain to the data-domain and, along this line, investigate the possible links between two domains. I now introduce the motivations, thesis organization and contributions.

1.4.1 Motivation I: towards common-shot inversion velocity analysis – a robust approach

Macro velocity estimation remains challenging in exploration geophysics. MVA is expected to be a powerful automatic imaging tool for determining the background velocity model. However, only few applications on real data have been published (Chauris and Noble, 2001; Mulder and ten Kroode, 2002; Alkhalifah, 2005; Shen and Symes, 2008; Mulder, 2014). Among others, the major reasons are: (1) subsurface-offset MVA, which is now popular in current literature (Symes, 2008), is very numerically demanding even in 2D, preventing from current extension to 3D; (2) migration artifacts can bias the velocity update: for example, in the subsurface-offset case, the optimization converges towards a velocity lower than the correct velocity and local minima are present in the objective function (Lameloise, 2015). For the first issue, it motivates the need to consider conventional common-shot MVA to reduce the numerical requirements, as the common-shot scheme benefits from a more natural parallel implementation. For the second issue, I propose to couple direct inversion to common-shot MVA, formulating a common-shot IVA approach. Direct inversion is a true-amplitude scheme, which automatically compensates for geometrical spreading losses and uneven illuminations. Thus, it suppresses migration artifacts visible in constructed images. Common-shot IVA is expected to be a robust macro velocity estimation method. I will study common-shot IVA through synthetic tests in Chapter 3.

Before the detailed investigations in Chapter 3, I indicate here the reason why conventional migration is not sufficient for velocity analysis in a simple case. I consider a homogeneous model with a single reflector embedded in (Figure 1.15a). In a correct background model, I first migrate the data to a reflectivity image (Figure 1.15b) for a single shot. Then, I migrate the data at an adjacent shot position and derive the difference between two images. Such images related to adjacent shots are expected to be consistent in a correct background model, however, upward curvatures are visible in the residual (Figure 1.15c). I repeat the same steps in an incorrect velocity (too low) model, and there are two main contributions in the image difference (Figure 1.15e): black ellipses correspond to upward-curved events attributed to model inaccuracies and green ellipse to upward-curved events attributed to migration smiles. One can not converge towards the correct velocity in a MVA procedure, as the image difference is not optimal even if the background velocity is correct. In practice, the macromodel will be over estimated, as the image differences attributed to migration smiles behave as if the estimated velocity is too low. In the case of the Marmousi model (1.16), the impact of migration smiles is even more severer: more events related to migration artifacts are visible due to the increased complexity of structures. The robustness of common-shot MVA will be significantly affected if these migration artifacts are not removed from the images.

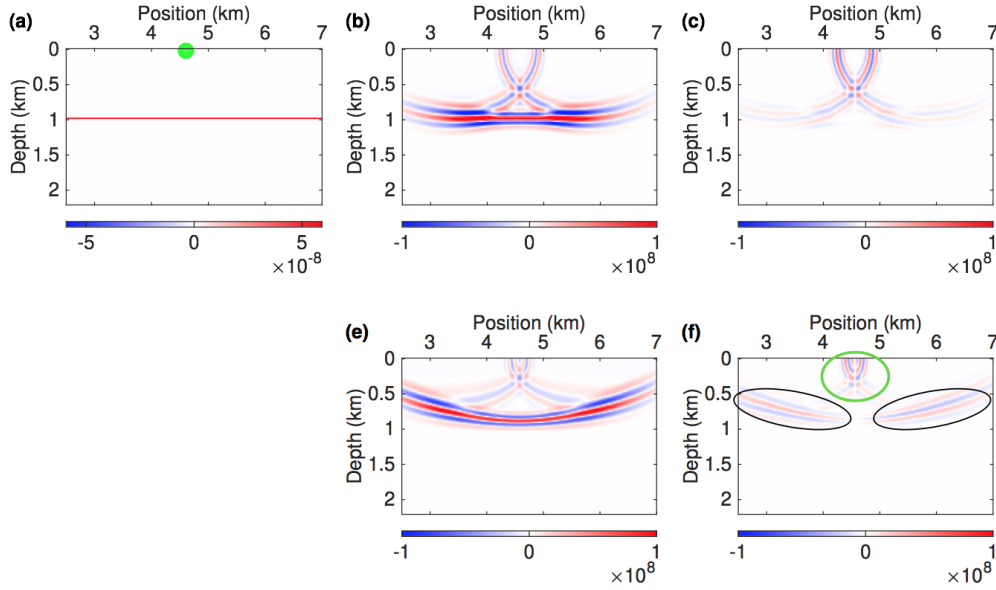


Figure 1.15 – Examples illustrating the impact of migration artifacts on image differences, for a single reflector embedded in a homogeneous macro velocity model. (a) corresponds to the exact reflectivity model, (b) to the image migrated in the correct macro velocity model for a single shot marked by green point in (a), and (c) to the difference between image (b) and the image for an adjacent shot. (d–e) are similar to (b–c), but for an incorrect model of which the velocity is 5% lower than the correct value. Black ellipses correspond to the upward-curved events attributed to velocity inaccuracies, and green ellipse to those attributed to migration smiles.

1.4.2 Motivation II: investigating the link between data-domain and image-domain methods

FWI can generate high-resolution models (Lailly et al., 1983; Tarantola, 1984a) but suffers from cycle-skipping effects due to the highly nonlinear relationship between data and model parameters (Bunks et al., 1995; Virieux and Operto, 2009). Thus, an starting model close enough to the true one is essential for FWI. It motivates the need for a powerful macromodel estimation technique. Relying on the Born approximation and scale separation, a tomographic approach can either be formulated in the data-domain (Biondi and Almomin, 2012; Xu et al., 2013; Chauris and Plessix, 2013; Liu et al., 2014a) or in the image-domain (Symes and Carazzone, 1991; Sava and Biondi, 2004; Biondi and Symes, 2004; Lameloise et al., 2015; Chauris and Cocher, 2017). Although many authors have tried to introduce the MVA concept into waveform inversion (Chauris and Plessix, 2013; Biondi and Almomin, 2012; Alkhalifah and Wu, 2016; Symes, 2017), the relationship between image-domain and data-domain methods has not been

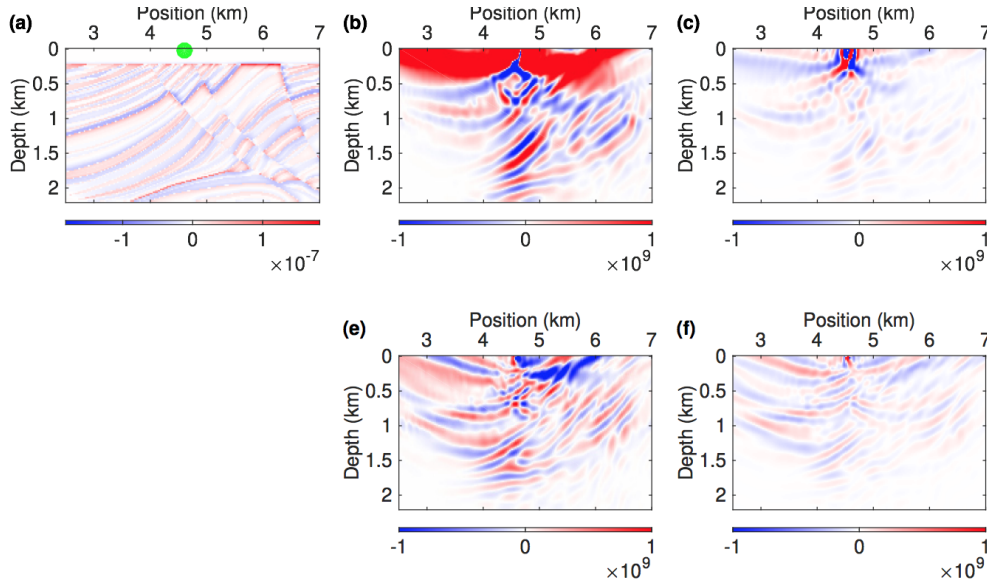


Figure 1.16 – Same as for Figure 1.15, but for the Marmousi model.

fully understood. [Chauris et al. \(2002a\)](#) established the link between MVA ([Symes and Carazzone, 1991](#); [Mulder and ten Kroode, 2002](#)) and slope tomography ([Billette and Lambaré, 1998](#); [Lambaré, 2008](#)), but it was for ray-based methods. It motivates the need to further investigate the possible links between two domains for wave-equation-based approaches. Therefore, in Chapter 4, I propose to extend the IVA approach from image-domain to data-domain and, along this line, investigate the links between data-domain and image-domain IVA methods. Data-domain IVA is expected to be close to RWI/DWI, in the sense that both methods rely on the scale separation and attempt to assess the estimated macro velocity model using the equivalence of image coherency criteria in the data-domain. It is of interest to better understand the relationship between data-domain and image-domain methods. I will explain more mathematical details of different approaches in Chapter 2. The formal comparison between data-domain and image-domain methods will be investigated in Chapter 4. In particular, I compare data-domain and image-domain IVA through numerical experiments. In addition, with theoretical analysis, I also compare DSO, a MVA strategy defined in image-domain, to DWI, a data-fitting scheme defined in data-domain.

1.4.3 Thesis organization

The thesis is organized as followed:

- In Chapter 2, I introduce the methodologies of forward modeling and popular wave-equation-based methods formulated in both data-domain or image-domain.

The data-domain part consists of FWI, RWI, and DWI, while the image-domain part includes common-shot MVA, subsurface-offset MVA, and the limitations of current MVA approaches. I mainly introduce the mathematical definition of the objective functions and the related gradients. In the FWI part, various alternatives to FWI will be compared on a simple numerical example to analyze the advantages and disadvantages;

- In Chapter 3, I first develop common-shot true-amplitude migration, which is an explicit direct inverse formula. It is an alternative to the work of [Qin et al. \(2015\)](#). The two approaches only differ at the source position and our method preserves early arrivals more accurately. Then, the direct inversion is coupled to common-shot velocity analysis. The scheme is formulated as a nested optimization, in which the inner loop for reflectivity imaging is directly solved through the direct inversion and outer loop updates the background velocity model with a nonlinear optimization. I analyze the first gradient for a horizontal reflector embedded in a homogeneous model. A simple modification of the DSO objective function is proposed by simply applying an additional operator to the final CIGs before evaluating the coherency. The operator is a function of the background velocity model and yields gradients free of spurious oscillations. Then, the impact of the triplicated wavefields is studied through a model consisting of a low velocity anomaly zone. Even the common-shot strategy is known to suffer from multipathing caused by triplications, the proposed approach proves to be robust to retrieve the shape of anomaly zone after several iterations. Finally, on the Marmousi model, the result justifies the superior robustness of IVA over MVA. By starting from the IVA result, subsequent FWI produces non cycle-skipped high resolution result, at least for the area where IVA supplies a correct background velocity update;
- In Chapter 4, I extend the common-shot IVA from image-domain to data-domain, leading to a more linearized waveform inversion approach. I calculate the differentiate images by comparing adjacent traces in CIGs. The data are produced by simply applying the Born modeling operator to images after the application of annihilator. The values of residual images are expected to be around zero for a correct background model, such that the remodeled data are expected to be minimal. The method is tested on the same models as in the common-shot IVA case. The major advantage of data-domain is that the additional operator for images is not necessary any more, and the spurious oscillation can be removed after a simple smoothing. This new approach is expected to be equivalent to DWI and, along this line, I compare the image-domain and data-domain approaches to establish the formal links between data-fitting principle and image-coherency criteria. The comparison is carried out through both theoretical analysis and numerical experiments;

- In Chapter 5, the main conclusions of the thesis are summarized and I propose several directions for future studies, in particular: edge effects, introduction of more physics (e.g. density, multiples, transmissions, elastic, attenuation, etc.), extension to 3D, and real data application.

1.4.4 Contributions

I first modified an existing finite-difference acoustic forward propagation code, by implementing the 4th order staggered grid (Virieux, 1986) and by adding the convolutional perfect matching layer (CPML) (Komatitsch and Martin, 2007) to absorb artificial reflections. Then, I have developed two methods, common-shot IVA in both image and data domains, for the estimation of the background velocity update presented in Chapter 3 and 4, as well as the direct inversion formula used here as a reflectivity imaging tool. The gradient derived with the adjoint state method (Plessix, 2006) is checked with the finite-difference test (Claerbout, 1985). My main contributions are

- the development of common-shot direct inversion scheme for reflectivity imaging, leading to CIGs with less migration artifacts, as detailed in Chapter 3;
- the coupling of direct inversion with velocity analysis, formulating a more robust approach – the common-shot IVA scheme, as detailed in Chapter 3;
- the extension of common-shot IVA to the data-domain, leading to a more linearized waveform inversion equivalent to DWI, as detailed in Chapter 4;
- the qualitative and quantitative comparisons between approaches formulated in the image-domain and the data-domain, establishing the formal link between data-fitting principle and image coherency criteria, as detailed in Chapter 4.

These developments allow a better understanding of the behavior of common-shot IVA and of the possible links between image-domain and data-domain methods. I have first proposed strategies for a robust and efficient approach – common-shot IVA. Then, this scheme has been extended to data-domain and, along this line, the relationship between data fitting and image coherency has been investigated. Part of the results has been presented in

- Chauris, H., Li, Y. and Cocher, E. [2017] Image-domain versus data-domain velocity analysis based on true-amplitude subsurface extended migration. 79th EAGE Conference & Exhibition Workshop, Paris, France;
- Li, Y. and Chauris, H. [2017] An alternative true-amplitude common-shot reverse time migration. 79th EAGE Conference & Exhibition, Paris, France;

- Li, Y. and Chauris, H. [2017] Coupling true-amplitude common-shot reverse time migration to velocity analysis. 79th EAGE Conference & Exhibition, Paris, France;
- Chauris, H. and Li, Y. [2015] Stereo-wave tomography: a new strategy for seismic imaging. SIAM Conference on Imaging Science, Stanford, United States,

and/or submitted to publication,

- Li, Y. and Chauris, H. [2017] Coupling direct inversion to common-shot velocity analysis. Geophysics. Submitted;
- Li, Y. and Chauris, H. [2017] Comparison between image-domain and data-domain inversion velocity analysis via a synthetic case study. Geophysics. In preparation.

Chapter 2

Methodology

Contents

2.1	Introduction	39
2.2	Forward problem	39
2.2.1	Wave equation	39
2.2.2	Numerical solution	40
2.2.3	Stability and dispersion	41
2.2.4	Boundary condition	42
2.3	Data fitting principle	43
2.3.1	Full Waveform Inversion and alternatives	43
2.3.2	Reflection Waveform Inversion	56
2.3.3	Differential Waveform Inversion	58
2.4	Image coherency criteria	60
2.4.1	Surface-oriented MVA	62
2.4.2	Depth-oriented MVA	63
2.4.3	Limitations of MVA	65
2.5	Summary	70

Résumé du chapitre 2

Ce chapitre est dédié à l'introduction des méthodologies d'imagerie sismique dans les domaines des données et des images. Plus tard au chapitre 4, le lien entre ces deux familles sera investigué avec plus de détails mathématiques. Je considère ici seulement les méthodes basées sur l'équation des ondes complètes, même si je me réfère aux méthodes de rais pour l'analyse de certains operateurs d'équation d'onde. Avant d'introduire ces différentes approches, je revois les aspects de modélisation par différences finies en tant que méthodes numériques. Les méthodes liées au domaine des données sont l'inversion des formes d'onde (FWI), la « Reflection Waveform Inversion » (RWI) et la « Differential Waveform Inversion » (DWI), tandis que celles définies dans le domaine image incluent l'analyse de vitesse par migration (MVA) et sa version données. Les références bibliographiques sont données au cours du chapitre. J'introduis principalement les fonctions objectives et le gradient associé. Pour la partie FWI, des alternatives sont analysées, avec une discussion sur les effets de saut de phase. Je discute aussi les limitations actuelles des méthodes MVA.

2.1 Introduction

This chapter is dedicated to introducing the methodologies of seismic imaging methods defined in both data and image domains. Later in chapter 4 the link between two families will be investigated with more mathematical details. I only deal here with wave-equation-based methods even if I may refer to ray-based methods for the analysis of some wave-equation operators. The data-domain part consists of Full Waveform Inversion (FWI), Reflection Waveform Inversion (RWI) and Differential Waveform Inversion (DWI), while those defined in the image-domain include surface-oriented Migration Velocity Analysis (MVA) and depth-oriented MVA. The appropriate references are given below in this Chapter. I mainly introduce here the definition of the objective function (OF) and the associated gradient. In the FWI part, the alternative definitions of the classic OF will be analyzed to discuss the cycle skipping effects. The image-domain part involves the discussion about the current limitations of MVA. Before introducing the different approaches, I first review the finite difference modeling aspects used for the numerical solutions.

2.2 Forward problem

2.2.1 Wave equation

The forward problem consists of solving the wave equation to model the seismic data propagation given a set of model parameters. For the scope of this thesis, I consider the constant-density isotropic acoustic approximation of the wave propagation, for which the wave equation reads:

$$\mathcal{L}u(\mathbf{s}, \mathbf{x}, t) = \delta(\mathbf{x} - \mathbf{s})\Omega(t), \quad (2.1)$$

where u is the pressure wavefield depending on the source position \mathbf{s} , the image position \mathbf{x} and the time t . u provides the observed data d_{cal} at receiver positions. Ω is the source wavelet, a function of time. The Helmholtz operator \mathcal{L} denotes $\frac{\partial^2}{\partial t^2} \frac{1}{c^2} - \Delta$, where $c(\mathbf{x})$ is the velocity and Δ the Laplacian operator. In the 2D case, this second-order partial differential equation is equivalent to the equations defined through the first-order linearized system of Newton's and Hooke's law, reading:

$$\frac{\partial v_x}{\partial t} = -\frac{\partial u}{\partial x}, \quad (2.2)$$

$$\frac{\partial v_z}{\partial t} = -\frac{\partial u}{\partial z}, \quad (2.3)$$

$$\frac{1}{c^2} \frac{\partial u}{\partial t} = -\left(\frac{\partial v_x}{\partial x} + \frac{\partial v_z}{\partial z}\right) + \delta(\mathbf{x} - \mathbf{s}) \int_t dt \Omega, \quad (2.4)$$

where v_x, v_z are the particle velocity components in x and z directions, respectively. To simulate the wave propagation, I solve the system of first-order equations instead of equation 2.1. The main reason is the introduction of absorbing boundary (section 2.2.4).

2.2.2 Numerical solution

The acoustic wave equation can be solved in the time domain (Tarantola, 1984a; Mora, 1989) or in the frequency domain (Pratt and Worthington, 1990; Sirgue and Pratt, 2004). As indicated in chapter 1, the numerical resolution of the partial differential equations can be achieved with many schemes including finite-difference method (FDM) (Virieux, 1986; Levander, 1988; Operto et al., 2007) and finite-element method (FEM) (Smith, 1975; Marfurt, 1984). Virieux et al. (2011) have reviewed the efficiency and complexity of the different numerical modeling methods.

FDM is a natural way to solve partial differential equations. It provides a relatively efficient implementation compared to FEM for exploration problems. The method directly estimates each differential term using the Taylor expansion on a regular grid. Suppose that the 2D spatial samplings are δx and δz and the time step is δt , then the first-order derivatives in the spatial coordinates (lateral position x and depth position z) are approximated by the so-called centralized fourth-order Crank-Nicolson approximation (Kreiss et al., 1973; Robertsson et al., 1994), reading

$$\frac{\partial u_{i,j}^n}{\partial x} \approx \frac{-u_{i+\frac{3}{2},j}^n + 27u_{i+\frac{1}{2},j}^n - 27u_{i-\frac{1}{2},j}^n + u_{i-\frac{3}{2},j}^n}{24\delta x}, \quad (2.5)$$

$$\frac{\partial u_{i,j}^n}{\partial z} \approx \frac{-u_{i,j+\frac{3}{2}}^n + 27u_{i,j+\frac{1}{2}}^n - 27u_{i,j-\frac{1}{2}}^n + u_{i,j-\frac{3}{2}}^n}{24\delta z}, \quad (2.6)$$

where i, j are the 2D spatial coordinates and n the temporal coordinate. Note that $\pm\frac{1}{2}$ means that the position of the grid is shifted by half a grid point, corresponding to staggered grids (Virieux, 1986). Then, the first-order derivative in time is approximated by

$$\frac{\partial u_{i,j}^n}{\partial t} \approx \frac{u_{i,j}^{n+1} - u_{i,j}^n}{\delta t}. \quad (2.7)$$

The partial derivatives of u are given as examples, but note that the same approximations can be applied to the partial derivatives of v_x and v_z . The reader is referred to Fornberg (1988) for a detailed derivation. In this thesis, I use exactly the scheme with fourth-order in space and second-order in time for simulating the wave propagation.

Virieux (1986) first introduced the FDM on a staggered grid to increase the order of the numerical scheme. With such scheme, velocity c and u are sampled at the same position, whereas v_x and v_z grids are shifted by half a grid point (Figure 2.1). There remain two problems for FDM: high velocity contrasts in models lead to instabilities,

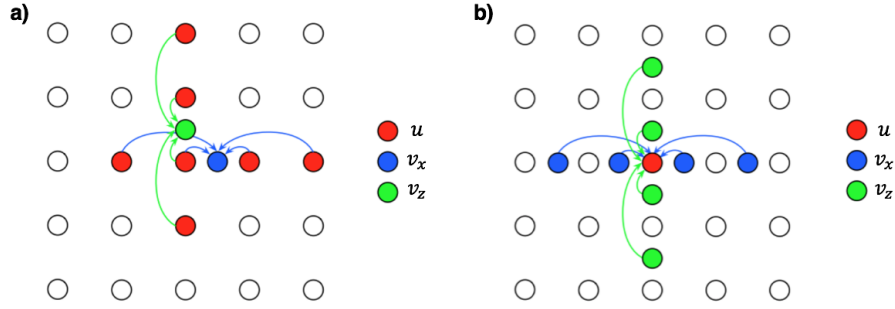


Figure 2.1 – Description of the grid points needed to update (v_x, v_x) (a) and u (b) fields (adapted from Thorbecke, 2015). For both cases, four neighbouring points are needed to compute one central point.

and high Poisson ratio to dispersion. Stability and dispersion will be detailed in the next section.

The spatial discretization for FDM is achieved on the grid with a regular rectangle shape, making it simple and straightforward to apply. Nevertheless, the implementation of regular grids also gives rise to certain limitations, in particular when dealing with topography. The spatial sampling is fixed and not adapted to local physical properties, thus the method is not as flexible as FEM which uses irregular grids. Marfurt (1984) observes that FEM better simulates curvilinear interfaces and thin beds than FDM. However, for the scope of this work, I only consider FDM.

2.2.3 Stability and dispersion

To ensure the stability of modeling scheme, the Courant–Friedrichs–Lewy (CFL) condition is a necessary condition for convergence while solving partial differential equations by FDM. The CFL condition is used to restrict the time-step in numerical simulations. For example, if a wave is crossing a discrete grid distance (δx), then the time-step must be less than the time needed for the wave to travel to an adjacent grid point, otherwise the simulation will produce incorrect results. The discretization must satisfy in 2D:

$$\delta t \leq \frac{1}{c_{max} \sqrt{\frac{1}{(\delta x)^2} + \frac{1}{(\delta z)^2}}}, \quad (2.8)$$

where c_{max} is the maximum velocity. With $\delta h = \delta x = \delta z$, equation 2.8 reads

$$\delta t \leq \frac{1}{\sqrt{2}} \frac{\delta h}{c_{max}}. \quad (2.9)$$

If this condition is not satisfied, the simulation will output unstable results as shown in Figure 2.2b.

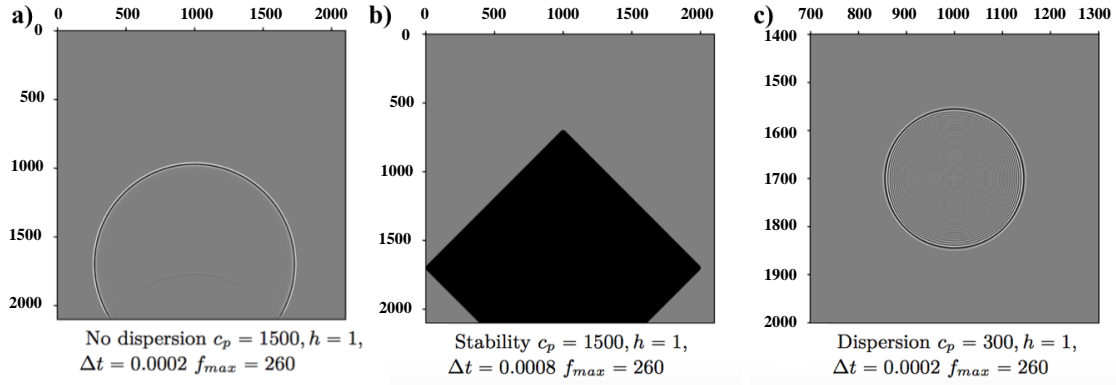


Figure 2.2 – Snapshots of wavefields for the case of stable result (a), of instability (b), and of dispersion (c) (adapted from [Thorbecke, 2015](#)).

On the other hand, dispersion can also occur during the simulation of wave propagation. The spatial sampling δh should be defined such that:

$$\delta h \leq \frac{\lambda_{min}}{n_\lambda} = \frac{c_{min}}{f_{max}} \frac{1}{n_\lambda}, \quad (2.10)$$

where λ_{min} corresponds to the smallest wavelength, which depends on the maximum frequency f_{max} and the minimum velocity c_{min} . n_λ denotes the number of grid points per wavelength. The number of points per wavelength depends on the spatial and temporal orders in the Taylor expansion: $n_\lambda = 5$ is usually chosen for fourth-order in space and second-order in time ([Alford et al., 1974](#)). In the case of numerical dispersion, the simulation will produce dispersive waves, as shown in Figure 2.2c. Note that numerical dispersion is not the physical dispersion of visco-elastic waves.

2.2.4 Boundary condition

Due to limited computational capacities, the models used in simulation are usually limited to a finite domain. If no boundary condition is defined at the borders of the model, waves will be artificially reflected from the numerical borders. To attenuate these spurious reflections, absorbing boundary conditions must be applied. The idea is that absorbing boundary conditions allow the waves propagating out of the model but prevent waves from propagating back into the model. Historically, absorbing boundary conditions introduce additional layers around the model, and these layers work as sponges that can attenuate the waves ([Cerjan et al., 1985](#)). However, these sponge layers are not always effective. Waves entering with a normal incident angle can be appropriately absorbed by the sponges, however spurious reflections still appear for other incident angles ([Brossier et al., 2009](#)).

Bérenger (1994) introduced a more effective technique, the Perfectly Matched Layer (PML), for electromagnetic. Later, among others Komatitsch and Tromp (2003) introduced this scheme for the simulation of seismic wave propagation. It is proven that PML can properly absorb the waves reaching the numerical artificial borders, regardless of their incidence angle. Furthermore, the behavior of PML at grazing incidence can be improved using an unsplit convolutional approach proposed by Komatitsch and Martin (2007).

To be realistic, it also requires the introduction of the Earth-air interface by a free surface boundary condition, for the simulation of reflections from the Earth-air interface. However, in this thesis, I use convolutional PML all around the model and the free surface is not considered (i.e. surface multiples should be removed from the data in a pre-processing step).

2.3 Data fitting principle

I detail in this section data-domain methods addressing the inverse problem under the isotropic constant-density acoustic approximation of the wave equation, meaning that the unknown is the pressure wave velocity field $c(\mathbf{x})$. The FWI strategy is first presented, considering the complete data set to reconstruct the model $c(\mathbf{x})$. However, such a scheme may be trapped into a local minimum due to the cycle-skipping effects, for which the solution is either to provide an initial model accurate enough or to include the low-frequency content in data. Alternatively, the least-squares OF can be replaced by other distance estimation metrics to enlarge the basin of attraction around the correct velocity model, such that the optimization converges towards a global minimum. Then, alternative approaches (Reflection Waveform Inversion, RWI and Differential Waveform Inversion, DWI) are detailed: they rely on the scale separation assumption, which divides $c(\mathbf{x})$ into the background velocity model $c_0(\mathbf{x})$ and velocity perturbation model $\delta c(\mathbf{x})$. Reflected data are used to update the large-scale structures of the subsurface. Note that the recovery of the macro velocity model $c_0(\mathbf{x})$ (i.e. the tomographic mode) is the main focus in this study, which remains challenging in the literature of seismic imaging.

2.3.1 Full Waveform Inversion and alternatives

I first introduce the derivation of the gradient of a generic OF with respect to the model. It is useful for the remaining sections of this thesis. Then, I review the classic FWI approach with a least-squares OF (Lailly et al., 1983; Tarantola, 1984a; Virieux and Operto, 2009; Fichtner, 2011). Current alternatives of FWI are then introduced through other OF definitions and the associated gradients with respect to velocity model. I test the different approaches on very simple synthetic data sets to discuss if they are prone

to cycle skipping effects. The least-squares OF is evaluated on a very simple synthetic data where the necessary low-frequency content is not present (Pratt et al., 1996; Sirgue and Pratt, 2004).

2.3.1.1 Deriving the model gradient of a generic OF

This introduction is useful for deriving the gradient of all kinds of data-fitting based objective functions. The gradient of a generic OF \mathcal{J}_0 can be computed via the widely used adjoint-state technique (Chavent, 1974; Plessix, 2006). We extend \mathcal{J}_0 with the Lagrangian formulation, denoted by \mathcal{J}_{ext} , reading

$$\mathcal{J}_{ext} = \mathcal{J}_0[d_{cal}(\mathbf{s}, \mathbf{r}, t)] - \langle \lambda_{gen}(\mathbf{s}, \mathbf{x}, t), \mathcal{L}u(\mathbf{s}, \mathbf{x}, t) - \delta(\mathbf{x} - \mathbf{s})\Omega(t) \rangle_{\mathbf{s}, \mathbf{x}, t}, \quad (2.11)$$

where λ_{gen} is the generic adjoint source and u the source wavefield. The modeled data is the response of source wavefield recorded at receiver positions: $d_{cal}(\mathbf{s}, \mathbf{r}, t) = \int d\mathbf{x} u(\mathbf{s}, \mathbf{x}, t)\delta(\mathbf{x} - \mathbf{r})$. The total derivative of \mathcal{J}_{ext} with respect to the velocity model is given by,

$$\frac{d\mathcal{J}_{ext}}{dc} = \frac{\partial \mathcal{J}_{ext}}{\partial c} + \frac{\partial u}{\partial c} \frac{\partial \mathcal{J}_{ext}}{\partial u} + \frac{\partial \lambda_{gen}}{\partial c} \frac{\partial \mathcal{J}_{ext}}{\partial \lambda_{gen}}. \quad (2.12)$$

The adjoint-state wavefield λ_{gen} is obtained by taking $(\frac{\partial \mathcal{J}_{ext}}{\partial u}, \frac{\partial \mathcal{J}_{ext}}{\partial \lambda_{gen}}) = 0$, such that the computation of Fréchet derivatives $\frac{\partial u}{\partial c}$ and $\frac{\partial \lambda_{gen}}{\partial c}$ are avoided. The gradient of \mathcal{J}_{ext} with respect to c reads,

$$\frac{d\mathcal{J}_{ext}}{dc(\mathbf{x})} = \frac{\partial \mathcal{J}_{ext}}{\partial c(\mathbf{x})} = \frac{2}{c^3(\mathbf{x})} \iint d\mathbf{s} dt \frac{\partial^2}{\partial t^2} \lambda_{gen}(\mathbf{s}, \mathbf{x}, t) u(\mathbf{s}, \mathbf{x}, t), \quad (2.13)$$

To derive λ_{gen} , I differentiate \mathcal{J}_{ext} with respect to u , yielding

$$\mathcal{L}^* \lambda_{gen}(\mathbf{s}, \mathbf{x}, t) = \frac{\partial \mathcal{J}_0}{\partial u(\mathbf{s}, \mathbf{x}, t)}. \quad (2.14)$$

where $*$ represents the adjoint operator such that \mathcal{L}^* is the back propagating operator. The velocity model can be updated using a gradient-based optimization scheme, for example the conjugate gradient method. In the following sections, various methods rely on the strategy to derive the gradient of the objective function with respect to velocity model.

2.3.1.2 Classic FWI

Lailly et al. (1983); Tarantola (1984a) use a ℓ_2 norm to measure the data misfit, reading

$$\mathcal{J}_{fwi}[c(\mathbf{x})] = \frac{1}{2} \|d_{cal}[c(\mathbf{x})](\mathbf{s}, \mathbf{r}, t) - d_{obs}(\mathbf{s}, \mathbf{r}, t)\|^2, \quad (2.15)$$

where d_{cal} and d_{obs} are the modeled and observed data, respectively. In practice, FWI is an ill-posed inverse problem; one must include additional information to find a feasible solution through a process called *regularization* (Tikhonov et al., 1977; Menke, 1984; Engl and Zou, 2000; Asnaashari et al., 2013). For example, one can enforce the smoothness of the recovered model, or ensure the consistency between model and a priori information (Asnaashari et al., 2013). Chauris et al. (2015); Cocher (2017) indicated the importance of regularization for the linearized waveform inversion. The readers are referred to Castellanos-Lopez (2014) for a detailed review of regularization technique for waveform inversion.

As before, the gradient of J_{fwi} with respect to the velocity model $c(\mathbf{x})$ is derived through the adjoint-state technique (Plessix, 2006). It is obtained by crosscorrelating two wavefields, reading

$$\frac{\partial J_{fwi}}{\partial c(\mathbf{x})} = \frac{2}{c^3(\mathbf{x})} \iint d\mathbf{s} dt \frac{\partial^2}{\partial t^2} \lambda(\mathbf{s}, \mathbf{x}, t) u(\mathbf{s}, \mathbf{x}, t), \quad (2.16)$$

where λ represents the back propagation wavefield using data residual as the new source term,

$$\mathcal{L}^* \lambda(\mathbf{s}, \mathbf{x}, t) = \int d\mathbf{r} \delta(\mathbf{x} - \mathbf{r}) (d_{cal}(\mathbf{s}, \mathbf{r}, t) - d_{obs}(\mathbf{s}, \mathbf{r}, t)). \quad (2.17)$$

To illustrate the behavior of the FWI OF, a synthetic data set is generated by convolving a Ricker source wavelet, for which the maximum frequency is 27 Hz and frequencies below 8 Hz are filtered, with several instantaneous pulses (Figure 2.3). We consider it as the observed data. This scheme mimics the wave propagation effect, as the reflected/transmitted waves can be roughly explained as a convolution between the source wavelet and reflection/transmission coefficients (Cooke and Schneider, 1983; Mendel, 2013).

When the starting model is not close enough to the exact model, the classic FWI can be cycle skipped due to the lack of low-frequency contents in the data. The method is very sensitive to kinematic errors (i.e. phase shift). In the seismic data, the phases and the amplitudes mainly correspond to the kinematics and the dynamics controlled by the velocity model, respectively. I thus mimic the kinematic and the dynamic errors by shifting the phases and by modifying the amplitudes in data, respectively.

Two cases are investigated to check the behavior of the ℓ_2 norm OF with respect to phase shifts and amplitude changes. In the first test, the modeled data d_{cal} is generated by shifting the observed data with a time lag (Figure 2.4a), simply

$$d_{cal}[t, \tau] = d_{obs}[t - \tau], \quad (2.18)$$

where τ is the time lag, and the misfit curve associated to ℓ_2 norm (equation 2.15) reads

$$f_{fwi}[\tau] = \frac{1}{2} \|d_{cal}[t, \tau] - d_{obs}[t]\|^2. \quad (2.19)$$

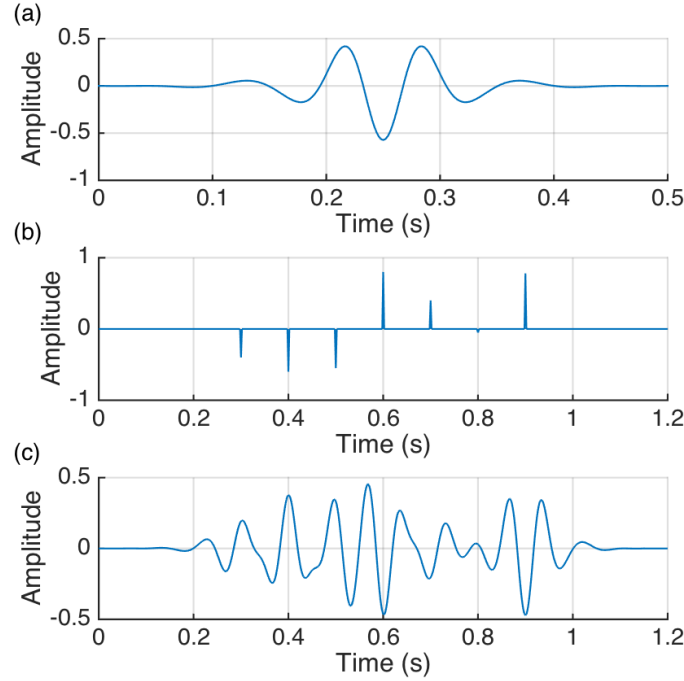


Figure 2.3 – The construction of synthetic data: (a) the source wavelet, (b) the random pulses and (c) the convolution between the wavelet and random pulses.

The misfit curve exhibits many local minima (Figure 2.4b). The time shift mimics the error of the kinematics controlled by the velocity model, such that a large time lag corresponds to a model far away from the exact model. If the starting model generates data as marked by the red dot in Figure 2.4b, the FWI will be cycle skipped since the modeled and observed data differ more than half a period in phase (Beydoun and Tarantola, 1988; Pratt et al., 2008). The ℓ_2 norm suffers from local minima due to the nonlinear relationship between model and data (Gauthier et al., 1986; Virieux and Operto, 2009). In the second test, I modify the amplitudes of modeled data from zero to twice the original amplitude (Figure 2.4c) according to

$$d_{cal}[R_{amp}] = R_{amp}d_{obs}, \quad (2.20)$$

to build the misfit curve,

$$f_{fwi}[R_{amp}] = \frac{1}{2} \|d_{cal}[R_{amp}] - d_{obs}\|^2, \quad (2.21)$$

where R_{amp} is the amplitude ratio. The amplitude change mimic the errors of dynamic information included in data. For example, in the case of reflected data, large amplitudes of data correspond to strong interface contrasts. The associated misfit curve is convex

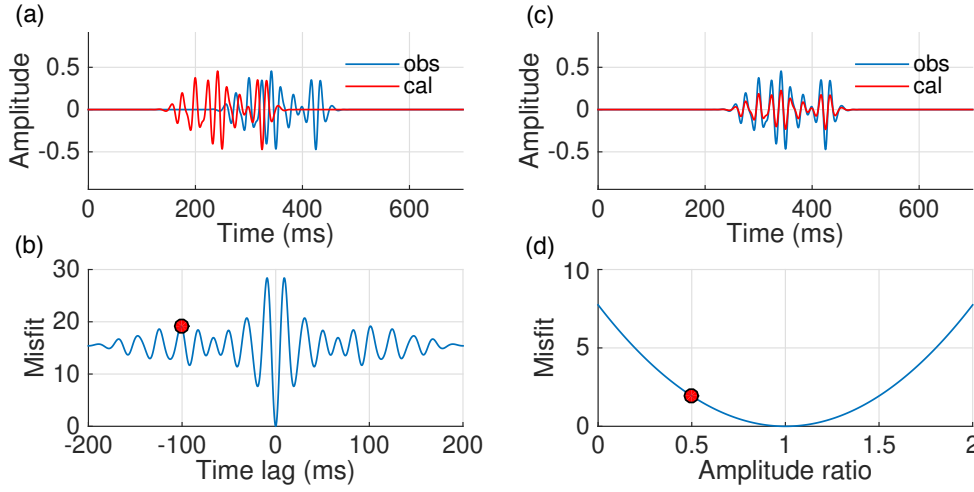


Figure 2.4 – Investigation of the shape of the ℓ_2 norm OF: (a) observed data (blue) and calculated data with shifted phase (red); (b) misfit curve related to different phase shifts (blue curve) and the misfit corresponding to calculated data in (a) (red dot); (c) observed data (blue) and calculated data with varying amplitude terms (red); (d) misfit curve associated to different amplitude ratios (blue curve) and the misfit corresponding to calculated data in (a) (red dot).

(Figure 2.4d) such that the optimization converges towards the correct solution. Assuming the physics of wave propagation is properly described, the ℓ_2 norm is able to retrieve the correct amplitudes but is very sensitive to phase shifts.

2.3.1.3 Current alternative OFs

I introduced several alternative OFs that are less prone to cycle skipping effects than FWI. The related mathematics is detailed now. For each method, I mainly review two steps, the definition of OF and the corresponding gradient, two essential elements to formulate an optimization procedure for solving the inverse problem.

Crosscorrelation OF

Crosscorrelation is a common tool to estimate the kinematic differences between two time series. The correlation-based OF has been introduced in seismic waveform inversion to get rid of cycle skipping effects (Luo and Schuster, 1991; Woodward, 1992). Initially, Luo and Schuster (1991) proposed to measure the crosscorrelation lag between real and modeled data and then minimized this time lag to update the velocity model. They applied it to a cross-well data. Later, Van Leeuwen and Mulder (2010) further proposed to use the shifted crosscorrelation between real and modeled data as the OF

and then penalized the focusing of the energy. They applied it to the transmitted waves on synthetic cases. [Chi et al. \(2015\)](#); [Wang et al. \(2015\)](#); [Luo et al. \(2016\)](#), among others, extended the correlation-based OF to reflected data. [Wang et al. \(2015\)](#) applied the method to a 2D elastic ultra-long streamer data set recorded in marine environment and [Luo et al. \(2016\)](#) implemented the method on a 2D field data set.

Referred to [Van Leeuwen and Mulder \(2010\)](#), an alternative OF measures the data difference using the crosscorrelation-based scheme,

$$\mathcal{J}_{\text{crss}}[c(\mathbf{x})] = \frac{1}{2} \|A(\tau)C(\mathbf{s}, \mathbf{r}, \tau)\|^2, \quad (2.22)$$

where C is the crosscorrelation between observed and modeled data, reading

$$C(\mathbf{s}, \mathbf{r}, \tau) = \int d\omega d_{\text{cal}}(\mathbf{s}, \mathbf{r}, \omega) d_{\text{obs}}^*(\mathbf{s}, \mathbf{r}, \omega) e^{2i\omega\tau}, \quad (2.23)$$

and $A(\tau) = \tau$, acting as a penalizing operator to adjust the weight of correlation energy corresponding to different time-lags τ . ω denotes the angular frequency. Similar to FWI, the gradient with respect to velocity can be derived using adjoint-state technique, reading

$$\frac{\partial \mathcal{J}_{\text{crss}}}{\partial c(\mathbf{x})} = \frac{2}{c^3(\mathbf{x})} \iint d\mathbf{x} d\omega (i\omega)^2 u(\mathbf{s}, \mathbf{x}, \omega) \lambda_{\text{crss}}^*(\mathbf{s}, \mathbf{x}, \omega), \quad (2.24)$$

where λ_{crss} represents the back propagated wavefield, reading

$$\mathcal{L}^* \lambda_{\text{crss}}(\mathbf{s}, \mathbf{x}, \omega) = \iint d\mathbf{r} d\tau \delta(\mathbf{x} - \mathbf{r}) A(\tau)^2 C(\mathbf{s}, \mathbf{r}, \tau) d_{\text{obs}}(\mathbf{s}, \mathbf{r}, \omega) e^{-2i\omega\tau}. \quad (2.25)$$

In practice, the observed and modeled data should be normalized before crosscorrelation ([Routh et al., 2011](#); [Zhang et al., 2014a](#)). Alternatively, one can formulate a normalized version by modifying equation 2.22:

$$\mathcal{J}_{\text{crss}}[c(\mathbf{x})] = \frac{1}{2} \frac{\|A(\tau)C(\mathbf{s}, \mathbf{r}, \tau)\|^2}{\|C(\mathbf{s}, \mathbf{r}, \tau)\|^2}. \quad (2.26)$$

To remove the effect of amplitudes, the energy of the crosscorrelation in the denominator acts as a scaling factor similar to the one used in image-domain methods ([Chauris and Noble, 2001](#); [Mulder and ten Kroode, 2002](#)), and thus the correlation OF mainly evaluates the kinematic errors included in data ([Van Leeuwen and Mulder, 2010](#); [Luo and Sava, 2011](#)).

The same tests are applied to evaluate the behavior of this method with respect to phase shifts and amplitude changes (Figure 2.5). In the first test, the calculated data are generated by giving different phase shifts to the observed data (equation 2.18 and

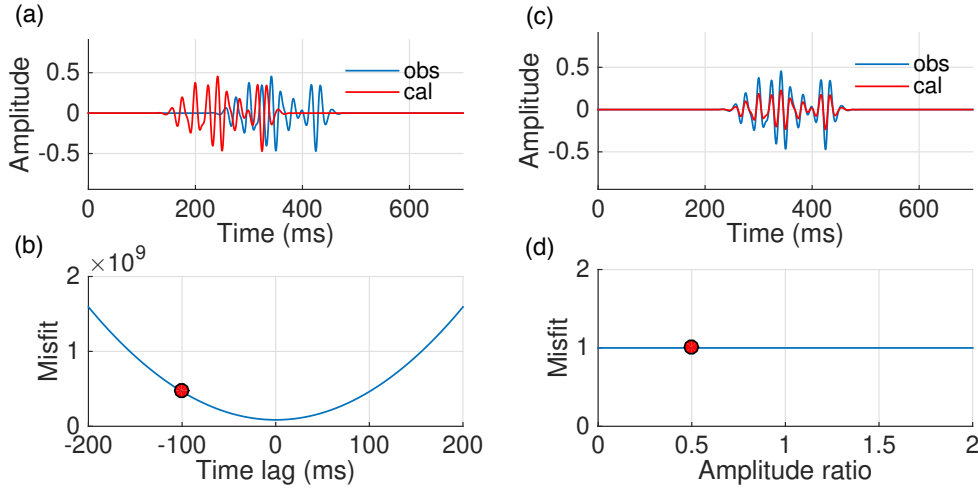


Figure 2.5 – Same as Figure 2.4, but for the normalized crosscorrelation OF.

Figure 2.5a). The misfit curve presents a convex behavior in this case (Figure 2.5b). If one starts from the red dot in Figure 2.5b, the method can perfectly converge towards the global minimum. In the second test, the amplitude is changed from zero to twice the original amplitude (equation 2.20 and Figure 2.5c). The crosscorrelation-based method is not sensitive to amplitude changes due to the introduction of normalization (Figure 2.5d). These two tests indicate that crosscorrelation OF suffers less from kinematic errors and its evaluation is not affected by amplitudes after normalization, at least in this simple case.

Adaptive Waveform Inversion OF

Deconvolution is an alternative to crosscorrelation for measuring the kinematic differences of two signals. Recently, the deconvolution-based OF has been introduced in seismic waveform inversion (Luo and Sava, 2011; Warner and Guasch, 2014, 2016). In early stages, Luo and Sava (2011) tested the approach on transmission data in a synthetic case. Then, Warner and Guasch (2014) proposed Adaptive Waveform Inversion (AWI) using the Wiener filter to perform the deconvolution between observed and modeled data. AWI works well with reflection-dominated data; then Guasch and Warner (2014) applied the approach formulated in 2D to an OBC anisotropic data set. Later, Warner and Guasch (2015) further developed the method into 3D and applied it to a 3D field data set, and also applied the approach on the Chevron model (designed for the 2014 FWI SEG workshop) with a modeling engine using the visco-elastic two-way wave equation.

The OF of AWI is defined via the Wiener transform filter (Warner and Guasch,

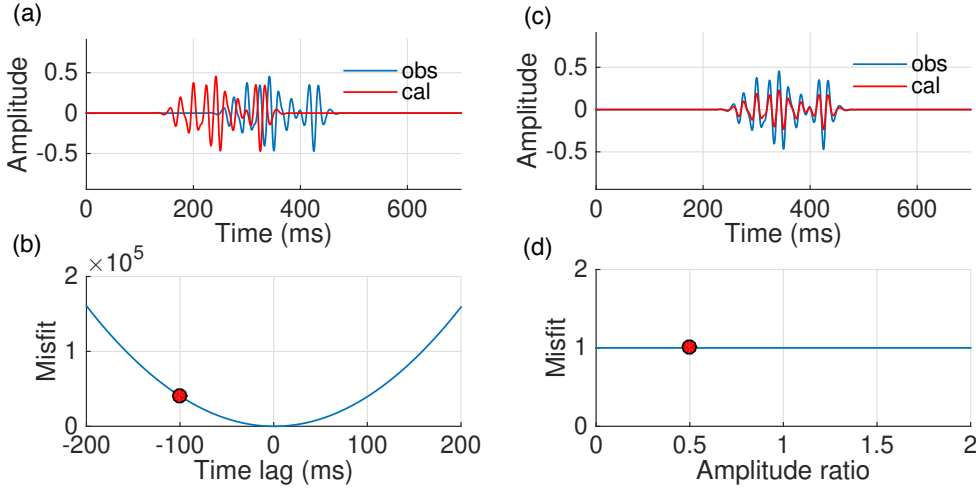


Figure 2.6 – Same as Figure 2.4, but for the AWI OF.

2016),

$$\mathcal{J}_{\text{awi}}[c(\mathbf{x})] = \frac{1}{2} \frac{||A(t)w[t]||^2}{||w[t]||^2}, \quad (2.27)$$

where $A(t) = t$ represents the penalty factor. $w[\omega] = \frac{d_{\text{cal}}[\omega]}{d_{\text{obs}}[\omega]}$ denotes the acausal Wiener filter that transforms a single observed trace into an associated modeled trace, that is, the deconvolution between calculated and observed data. The gradient of OF and the associated adjoint source read,

$$\frac{\partial \mathcal{J}_{\text{awi}}}{\partial c(\mathbf{x})} = \frac{2}{c^3(\mathbf{x})} \int dt ds \frac{\partial^2}{\partial t^2} u(\mathbf{s}, \mathbf{x}, t) \lambda_{\text{awi}}(\mathbf{s}, \mathbf{x}, t), \quad (2.28)$$

$$\mathcal{L}^* \lambda_{\text{awi}} = \int d\mathbf{r} \delta(\mathbf{x} - \mathbf{r}) \left[\frac{\partial w[t]}{\partial d_{\text{cal}}[t]} \left(\frac{A(t)^2 - 2\mathcal{J}_{\text{awi}}}{||w[t]||^2} \right) w[t] \right], \quad (2.29)$$

I apply the synthetic tests to the AWI OF with respect to phase shifts (Figure 2.6a) and amplitude changes (Figure 2.6c) as well. The misfit curve for the first test has a convex behavior (Figure 2.6b). The second test leads to a misfit curve with a constant value (Figure 2.6d). The AWI OF is normalized by the inner product of the Wiener filter such that it is not sensitive to any amplitude changes (equation 2.27). The AWI OF is based on the deconvolution of data. Similar to crosscorrelation based method, AWI uses penalizing factor (i.e. the temporal lag A) to enhance the weight of kinematic errors. The method is effective to get rid of cycle skipping effect but is not aimed at matching the amplitudes of data.

Envelope OF

Envelope is a reduced form of the signal: the modified data contain low frequencies even if the original signal only contains high frequencies. It has been introduced in seismic imaging to design an OF with an enlarged basin of attraction (Bozdağ et al., 2011; Liu et al., 2011b; Chi et al., 2014; Wu et al., 2014). In the early stages, Liu et al. (2011b); Chauris et al. (2012) applied envelop-based approach to transmission synthetic data. Then, Chi et al. (2014); Wu et al. (2014) extended it to reflected data in 2D synthetic case. Luo and Wu (2015) showed that envelope inversion is insensitive to Gaussian noise. In all cases, the low-frequency contents are not present in data but envelope-based waveform inversion converges well.

The envelope can be extracted from the seismic signal $d[t]$ via its analytic form $\tilde{d}[t]$ that is constructed by introducing the Hilbert transform \mathcal{H} of the data (Farnbach, 1975; Luo et al., 2003; Bozdağ et al., 2011),

$$\tilde{d}[t] = d[t] + i\mathcal{H}\{d[t]\}. \quad (2.30)$$

Then, the envelope $E[t]$ is the modulus of $\tilde{d}[t]$, reading

$$E[t] = \sqrt{\Re\{\tilde{d}[t]\}^2 + \Im\{\tilde{d}[t]\}^2}, \quad (2.31)$$

where \Re denotes the real part and \Im the imaginary part. Bozdağ et al. (2011) provide a possible definition of the envelope OF, defined as the difference between logarithms of envelopes of the modeled and observed data, reading

$$\mathcal{J}_{\text{env}}[c(\mathbf{x})] = \frac{1}{2} \left\| \ln \frac{E_{\text{obs}}(\mathbf{s}, \mathbf{r}, t)}{E_{\text{cal}}(\mathbf{s}, \mathbf{r}, t)} \right\|^2, \quad (2.32)$$

where E_{cal} and E_{obs} are envelopes of observed and modeled data, respectively. In practice, a pre-whitening factor is needed to avoid zero values at the denominator. The gradient of the envelope misfit with respect to the model and the adjoint source are

$$\frac{\partial \mathcal{J}_{\text{env}}}{\partial c(\mathbf{x})} = \frac{2}{c^3(\mathbf{x})} \iint d\mathbf{s} dt \frac{\partial^2}{\partial t^2} u(\mathbf{s}, \mathbf{x}, t) \lambda_{\text{env}}(\mathbf{s}, \mathbf{x}, t), \quad (2.33)$$

$$\mathcal{L}^* \lambda_{\text{env}} = \int d\mathbf{r} \delta(\mathbf{x} - \mathbf{r}) \left[\ln \left(\frac{E_{\text{obs}}}{E_{\text{cal}}} \right) \frac{d_{\text{cal}}}{E_{\text{cal}}^2} - \mathcal{H} \left\{ \ln \left(\frac{E_{\text{obs}}}{E_{\text{cal}}} \right) \frac{\mathcal{H} d_{\text{cal}}}{E_{\text{cal}}^2} \right\} \right]. \quad (2.34)$$

The same tests are also conducted to evaluate the behavior of envelope OF with respect to phase shifts (Figures 2.7a and 2.7b) and amplitude changes (Figures 2.7c and 2.7d). The envelop contains much lower frequencies than the original data such that the misfit curve presents no local minimum with respect to the phase shifts (Figure 2.7c). A pre-whitening factor ϵ is added to E_{cal} to avoid zero denominator in equation 2.32.

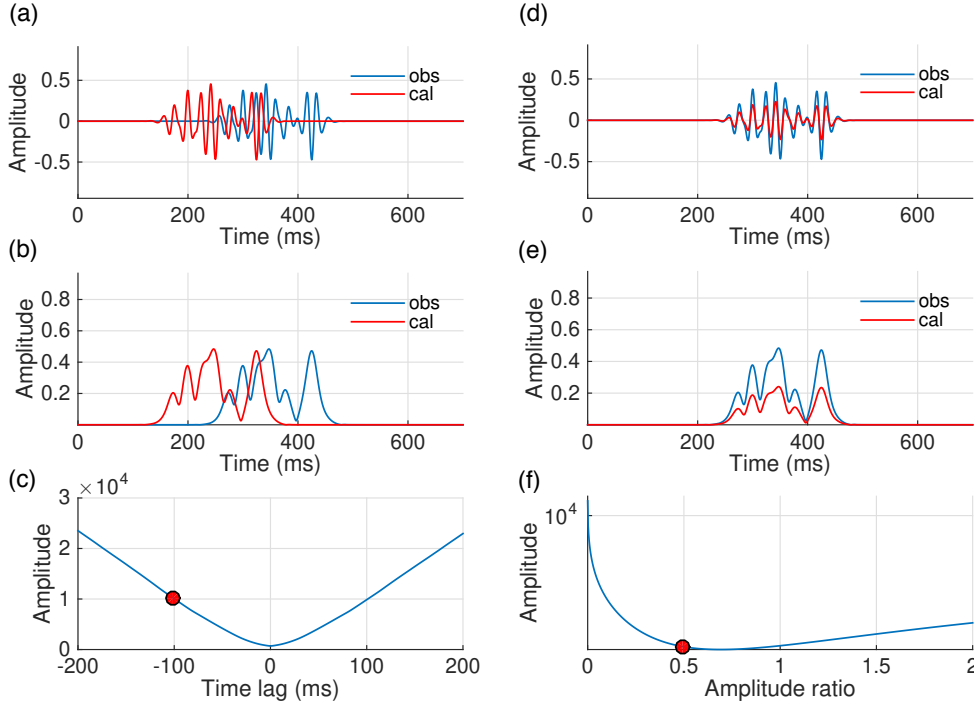


Figure 2.7 – Same as Figure 2.4, but for the logarithmic envelope OF. (a,c,d,f) correspond to (a-d) in Figure 2.4, respectively. (b,e) illustrate the envelopes with respect to data in (a,d), respectively. Note that the vertical axis of (f) is in logarithmic scale. For the amplitude case, the minimum is not reached for 1 due to the pre-whitening applied to the envelope.

However, the log function is sensitive to small changes of amplitudes. In the amplitude case, the global minimum shifts to a wrong position in the misfit curve (Figure 2.7f) due to the pre-whitening, that is, $\ln(E_{obs}/E_{cal})$ is not 0 even if the amplitude of E_{cal} and E_{obs} are identical. Without a cautious prewhitening, the envelope OF can not converge towards a model producing data with correct amplitudes. Besides, the envelope OF is limited to distinguish the difference between two data with reversed amplitudes (i.e. not sensitive to the sign of the data).

Alternatively, [Liu et al. \(2011b\)](#) proposed to calculate the time integral of envelopes for observed and modeled data and then defined a normalized envelope-based OF. I present here the version without normalization:

$$\mathcal{J}_{ienv} = \frac{1}{2} ||Q_{cal}(\mathbf{s}, \mathbf{r}, t) - Q_{obs}(\mathbf{s}, \mathbf{r}, t)||^2. \quad (2.35)$$

where

$$Q_{cal}(\mathbf{s}, \mathbf{r}, t) = \int_0^t d\tau E_{cal}(\mathbf{s}, \mathbf{r}, \tau). \quad (2.36)$$

The model gradient of equation 2.35 and the related adjoint source read

$$\frac{\partial \mathcal{J}_{ienv}}{\partial c(\mathbf{x})} = \frac{2}{c^3(\mathbf{x})} \iint d\mathbf{s} dt \frac{\partial^2}{\partial t^2} u(\mathbf{s}, \mathbf{x}, t) \lambda_{ienv}(\mathbf{s}, \mathbf{x}, t), \quad (2.37)$$

$$\mathcal{L} \lambda_{ienv}^*[t] = \iint d\mathbf{r} d\sigma \delta(\mathbf{x} - \mathbf{r}) \frac{\partial Q_{cal}[\sigma]}{\partial d_{cal}[t]} [Q_{cal}[\sigma] - Q_{obs}[\sigma]]. \quad (2.38)$$

where

$$\frac{\partial Q_{cal}[\sigma]}{\partial d_{cal}[t]} = \begin{cases} \int_0^\sigma d\tau d_{cal}[t]/E_{cal}[t], & \text{if } t = \tau \\ \int_0^\sigma d\tau \mathcal{H}\{d_{cal}[\tau]\}/E_{cal}[\tau]/\pi(t - \tau), & \text{if } t \neq \tau \end{cases} \quad (2.39)$$

The preliminary tests show that this OF present a convex misfit curve for both the phase shift and amplitude change cases (Figure 2.8). Note that the envelope of data is not sensitive to the sign of the data such that this OF may still be limited to distinguish the difference between two data sets with reversed amplitude.

Bump OF

In a similar way as for the envelope case, an alternative possibility to artificially create low-frequency contents is to use the absolute value or squared form of the data (Liu et al., 2011b; Bharadwaj et al., 2015). First, Liu et al. (2011b) proposed to calculate the normalized integral of the absolute value of data and then measured the difference between real and modeled data. This approach was applied to a 2D synthetic transmitted data set. Then, Donno et al. (2013) used the same strategy, except the normalized integral is performed on squared data, for a 2D synthetic cross-hole data set. Bharadwaj et al. (2015) extended the approach to reflected data in 2D synthetic cases. Bharadwaj et al. (2016) indicated that the squared form of data after smoothing can be regarded as a generalized envelope.

Bharadwaj et al. (2016) define a bump OF

$$\mathcal{J}_{bump}[c(\mathbf{x})] = \frac{1}{2} \left\| b(t) *_t (d_{cal}(\mathbf{s}, \mathbf{r}, t)^2 - d_{obs}(\mathbf{s}, \mathbf{r}, t)^2) \right\|^2, \quad (2.40)$$

where $*_t$ denotes the convolution in time and b a blurring function. b is in practice a low-pass filter which does not necessarily affect the original data set due to the lack of low frequencies, such that the data need to be squared. The gradient and related adjoint

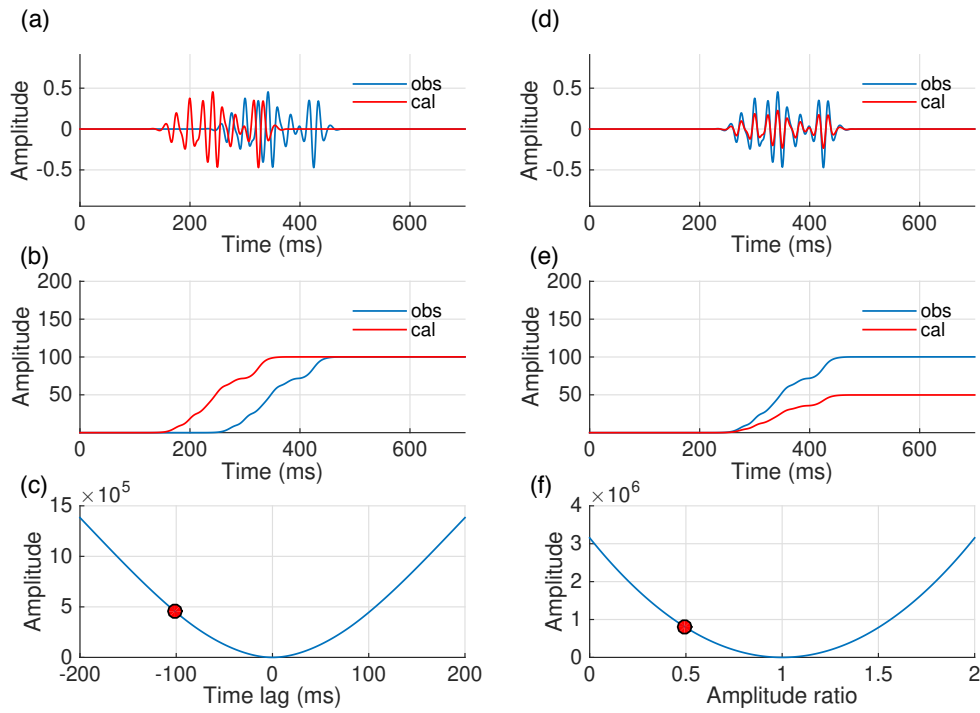


Figure 2.8 – Same as Figure 2.4, but for the envelop integral OF. (a,c,d,f) correspond to (a–d) in Figure 2.4, respectively. (b,e) illustrate the envelop integrals with respect to data in (a,d), respectively.

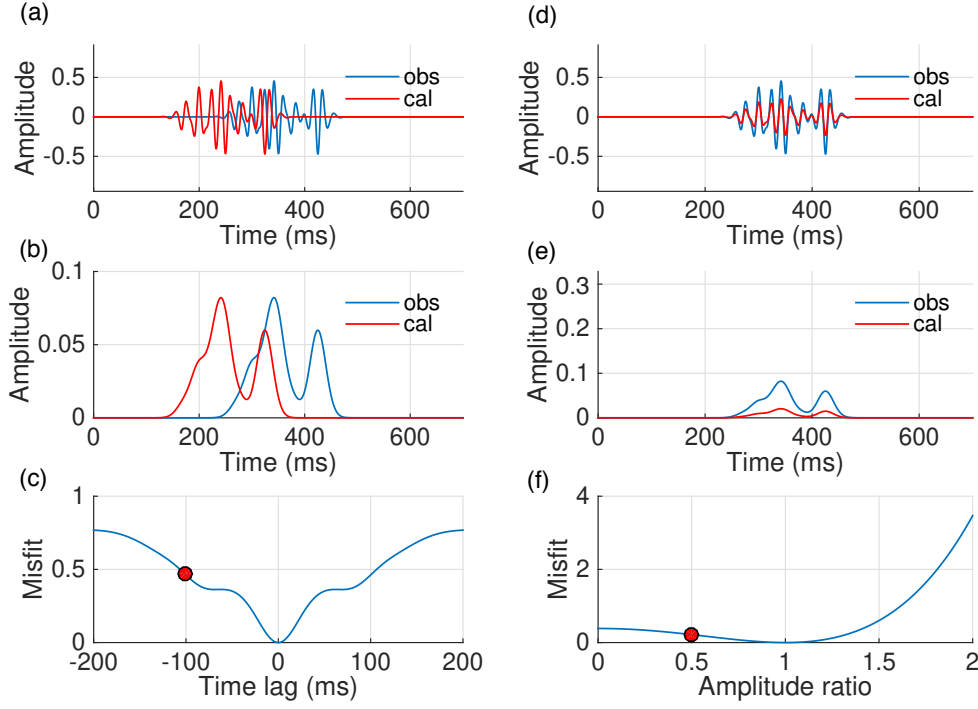


Figure 2.9 – Same as Figure 2.4, but for the auxiliary bump OF. (a,c,d,f) correspond to (a–d) in Figure 2.4, respectively. (b,e) illustrate the auxiliary bump form of data in (a,d), respectively.

source read,

$$\frac{\partial \mathcal{J}_{\text{bmp}}}{\partial c(\mathbf{x})} = \frac{2}{c^3(\mathbf{x})} \iint d\mathbf{s} dt \frac{\partial^2}{\partial t^2} \lambda_{\text{bmp}}(\mathbf{s}, \mathbf{x}, t) u(\mathbf{s}, \mathbf{x}, t), \quad (2.41)$$

$$\begin{aligned} \mathcal{L}^* \lambda_{\text{bmp}}(\mathbf{s}, \mathbf{x}, t) = & \int d\mathbf{r} \delta(\mathbf{x} - \mathbf{r}) b(t) \star_t d_{\text{cal}}(\mathbf{s}, \mathbf{r}, t) \\ & \times b(t) \star_t (d_{\text{cal}}(\mathbf{s}, \mathbf{r}, t)^2 - d_{\text{obs}}(\mathbf{s}, \mathbf{r}, t)^2), \end{aligned} \quad (2.42)$$

where \star_t denotes the crosscorrelation in time.

The sensitivity tests for the auxiliary bump OF are applied with respect to phase shifts (Figure 2.9a) and amplitude changes (Figure 2.9c). The misfit curve does not contain any local minima with respect to the phase shift but is not convex (Figure 2.9b). The second test shows that the auxiliary bump OF does not have the severe converging problem for fitting the dynamic information (Figure 2.9d). However, similar to the envelope OF, the method is not sensitive to the sign of data.

2.3.1.4 Conclusion

The behaviors of different OFs have been tested. The numerical tests displayed have several limitations: (1) the data used in the tests are not obtained by solving the wave equation; (2) I mimic the velocity model changes by varying either the phases or the amplitudes of the data, but the amplitudes and phases change simultaneously for different velocity models in practice; (3) the tests are performed in 1D. Despite the very rude approximations, they can still provide an insight into the behavior of different OFs. In particular, it roughly shows that these alternative OFs are less cycle skipped compared to standard FWI.

The conventional FWI seeks a velocity model that best explains the observed data, but may converge towards a local minimum due to the cycle skipping effects, such that the large-scale part of structure is difficult to be retrieved by the inversion procedure. One can modify the OF by giving more weight to evaluate the kinematic errors (e.g. the crosscorrelation or the Wiener filter) and/or using an altered data form (e.g. the envelop or the squared data), attempting to find the global optimal velocity model. These alternatives present misfit curves with a better convex property, at least for simple cases. In the following, I detail the reflection-based waveform inversion approaches relying on the scale separation and Born approximation.

2.3.2 Reflection Waveform Inversion

Reflection Waveform Inversion (RWI) (Xu et al., 2012; Alkhalifah, 2014; Brossier et al., 2015; Zhou et al., 2015) is an approach similar to FWI developed to retrieve the large-scale structure from reflected events. It is directly inspired by the Migration-Based Traveltime Tomography (MBTT) (Chavent et al., 1994; Plessix et al., 1995). FWI considers the velocity model c , whereas RWI either assumes an explicit scale separation between low and high wavenumber contents of the velocity model ($c = c_0 + \delta c$) (Xu et al., 2012) or relies on a parametrization with both the velocity and the impedance to naturally facilitate an implicit scale separation (Zhou et al., 2015; Zhou, 2016). I consider the former to formulate RWI seeking the optimal models by fitting the reflected data, reading

$$\mathcal{J}_{\text{rwi}}[c_0, \delta c] = \frac{1}{2} \|\delta d_{\text{cal}}[c_0, \delta c] - \delta d_{\text{obs}}\|^2, \quad (2.43)$$

where δd_{cal} and δd_{obs} are modeled and observed reflected data, respectively. c_0 and δc are the background velocity and the perturbation model, respectively. The modeling of reflection data can be linearized using the first-order Born approximation, reading

$$\delta d_{\text{cal}}(\mathbf{s}, \mathbf{r}, \omega) = \int d\mathbf{x} (i\omega)^2 \Omega(\omega) G_0(\mathbf{s}, \mathbf{x}, \omega) \frac{2\delta c(\mathbf{x})}{c_0(\mathbf{x})^3} G_0(\mathbf{r}, \mathbf{x}, \omega), \quad (2.44)$$

where Ω is the function of source wavelet and G_0 the Green's function associated to the background velocity model c_0 . Then, the augmented function read

$$\begin{aligned} \mathcal{J}_{\text{rwi}}^{\text{ext}}[c_0, \delta c] = & \frac{1}{2} \|\delta d_{\text{cal}}[c_0, \delta c] - \delta d_{\text{obs}}\|^2 - \iiint \text{d}\mathbf{s} \text{d}\mathbf{x} \text{d}\omega \delta R^* [\mathcal{L}_0 S_0 - \Omega \delta(\mathbf{x} - \mathbf{s})] \\ & - \iiint \text{d}\mathbf{s} \text{d}\mathbf{x} \text{d}\omega R_0^* [\mathcal{L}_0 \delta S - (i\omega)^2 S_0 \frac{2\delta c}{c_0^3}], \end{aligned} \quad (2.45)$$

where \mathcal{L}_0 denotes the Helmholtz operator associated to c_0 . R_0 and δR are adjoint state variables. S_0 and δS represent forward propagated and scattered wavefields in a given macromodel c_0 , respectively, reading

$$\mathcal{L}_0 S_0(\mathbf{s}, \mathbf{x}, \omega) = \Omega(\omega) \delta(\mathbf{x} - \mathbf{s}) \quad (2.46)$$

$$\mathcal{L}_0 \delta S(\mathbf{s}, \mathbf{x}, \omega) = (i\omega)^2 S_0(\mathbf{s}, \mathbf{x}, \omega) \frac{2\delta c(\mathbf{x})}{c_0(\mathbf{x})^3}, \quad (2.47)$$

Note that δd_{cal} record the response of δS at receiver positions \mathbf{r} : $\delta d_{\text{cal}} = \int \text{d}\mathbf{x} \delta S \delta(\mathbf{x} - \mathbf{r})$. Using the adjoint-state technique (Plessix, 2006), the gradients of OF with respect to c_0 and δc respectively read,

$$\frac{\partial \mathcal{J}_{\text{rwi}}}{\partial \delta c} = \frac{2}{c_0^3} \iint \text{d}\mathbf{s} \text{d}\omega (i\omega)^2 S_0^* R_0 \quad (2.48)$$

$$\frac{\partial \mathcal{J}_{\text{rwi}}}{\partial c_0} = \frac{2}{c_0^3} \iint \text{d}\mathbf{s} \text{d}\omega (i\omega)^2 (S_0^* \cdot \delta R + R_0^* \cdot \delta S - \frac{3\delta c}{c_0} S_0 R_0^*). \quad (2.49)$$

Note that I assume here that δc and c_0 are independent for such a derivation, following the strategy of Xu et al. (2012). The third term in equation 2.49 corresponds to the imprints of velocity perturbations in the gradient, and is an oscillatory term. Wu and Alkhalifah (2015) proposed to multiply this term by a factor to reduce its weight in the gradient. Adjoint state variables can be derived by taking $(\frac{\partial J_{\text{rwi}}^{\text{ext}}}{\partial R_0}, \frac{\partial J_{\text{rwi}}^{\text{ext}}}{\partial \delta R}) = 0$, reading

$$\mathcal{L}_0^* R_0(\mathbf{s}, \mathbf{x}, \omega) = \int \text{d}\mathbf{r} (\delta d_{\text{cal}}(\mathbf{s}, \mathbf{r}, \omega) - \delta d_{\text{obs}}(\mathbf{s}, \mathbf{r}, \omega)) \delta(\mathbf{x} - \mathbf{r}), \quad (2.50)$$

$$\mathcal{L}_0^* \delta R(\mathbf{s}, \mathbf{x}, \omega) = (i\omega)^2 R_0(\mathbf{s}, \mathbf{x}, \omega) \frac{2\delta c(\mathbf{x})}{c_0(\mathbf{x})^3}. \quad (2.51)$$

Adjoint variables R_0 and δR can be explained as back propagated and back scattered wavefields in an estimated macromodel c_0 , respectively. The conventional FWI generates a gradient (Figure 2.10a) containing different types of subkernels, such as the transmission kernel (Figure 2.10b) and migration ellipse (Figure 2.10c). The transmission wave that penetrates into the deep subsurface requires a ultra-long offset acquisition to record. The migration ellipse dominates the contribution of reflected data in the FWI

gradient. In areas not reached by diving waves, RWI seeks a way to extract the large-scale information of the velocity model by constructing transmission wavepaths between the reflectors and the source as well as between the reflectors and the receivers (the famous rabbit ear shapes, see Figures 2.10d and 2.10e) (Xu et al., 2012; Alkhalifah, 2014; Brossier et al., 2015; Wu and Alkhalifah, 2015).

In a given macromodel c_0 , the model perturbation δc is estimated via equation 2.48, which is equivalent to the formulation of reverse time migration (RTM). Then, reflected data are modeled from $(\delta c, c_0)$ and the residuals between modeled and observed data are measured to iteratively update the background model. These two procedures are repeated until converging (Wu and Alkhalifah, 2015; Zhou et al., 2015). In practice, the estimation of perturbation only uses short-offset (even zero-offset) data and the recovery of macromodel involves reflected data of long-offset. The idea behind is to build a perturbation model using data with limited offset and then to predict long-offset data. Thus, the macromodel can be updated by fitting these predicted data with the observed data. Zhou et al. (2015); Alkhalifah and Wu (2016) proposed strategies to combine the information extracted from both transmission data and reflection data to update the background velocity model.

2.3.3 Differential Waveform Inversion

In the context of macro velocity estimation, Chauris and Plessix (2012) proposed the Differential Waveform Inversion (DWI) approach defined in the data domain, which is inspired by Differential Semblance Optimization (DSO) technique (see Symes, 2008 for a review) formulated in the image domain. In a given background velocity model c_0 , DWI first constructs the velocity perturbation δc from reflected data of a single shot s by minimizing

$$\mathcal{J}_{\text{dwi}}^{(0)}[\delta c] = \frac{1}{2} \|\delta d_{\text{cal}}^{(s)}[c_0, \delta c] - \delta d_{\text{obs}}^{(s)}\|^2, \quad (2.52)$$

where $\delta d_{\text{cal}}^{(s)}$ and $\delta d_{\text{obs}}^{(s)}$ are the computed and observed reflection data at shot position s . Similar to least-squares migration, a migrated section $\delta c[c_0, s]$ is extracted after minimization. The second step consists of minimizing the misfit between the shot couples $(s + \delta s)$ for the optimal reflectivity δc derived from equation 2.52

$$\mathcal{J}_{\text{dwi}}^{(1)}[c_0] = \frac{1}{2} \|\delta d_{\text{cal}}^{(s+\delta s)}[c_0, \delta c[c_0, s]] - \delta d_{\text{obs}}^{(s+\delta s)}\|^2, \quad (2.53)$$

Due to the finite frequency band of the data, the calculated and observed gathers $s + \delta s$ should match within less than half a wavelength if the interval δs is not too large. This can ensure a larger basin of attraction than in the case of classic FWI. Depending on the authors, the dependency of δc on c_0 is not always considered. This approach accounts for

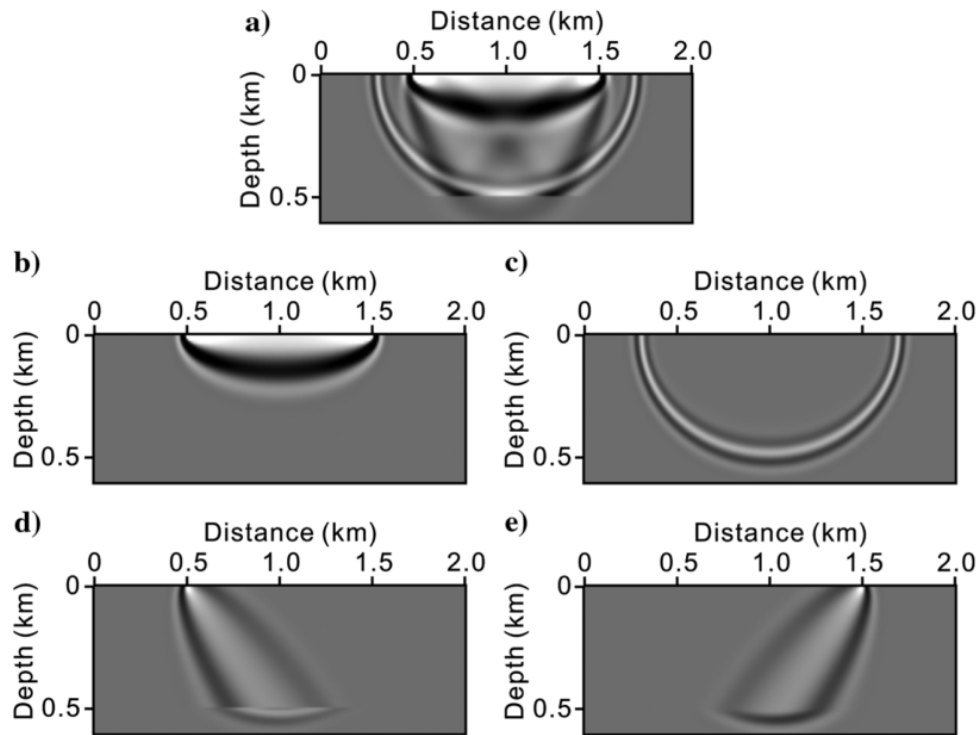


Figure 2.10 – An illustration comparing the behavior of FWI and RWI. Sensitivity kernel of FWI for a single reflector embedded in the homogeneous background velocity model and its different subkernels (adapted from [Chi et al., 2015](#)): (a) full kernel of FWI, (b) transmission kernel, (c) migration ellipse, (d) source-side reflection kernel, and (e) receiver-side reflection kernel. For RWI, the kernel related to macro model update includes only (d) and (e), introducing the famous rabbit ear shapes.

such dependency, formulating a nested optimization. I develop an equivalent approach in chapter 4 and the derivation of gradient will be detailed in that part.

A main interest of this strategy resides in the possibility to take into account surface-related multiples (Chauris and Plessix, 2013). Similar to the iterative MVA that can cope with the multiples proposed by Cocher et al. (2017b), the key point is the iterative migration. In practice, the iterative migration can be replaced by direct inverse (Chapter 3) to significantly reduce the computational cost.

Discussion

FWI, RWI and DWI are different data-fitting techniques. The idea behind FWI is that a model is optimal if it can perfectly reproduce the complete data set. RWI and DWI use only reflected data and rely on the scale separation explicitly or implicitly. Similar to FWI, RWI seeks the consistency between modeled and observed data. In addition, it introduces the idea that, only in a correct macro model, one can perfectly reproduce the long-offset data with the perturbation model constructed from short-offset data with limited offset range. In this sense, DWI is similar to RWI: it predicts the data at a single shot position with the perturbation model constructed from the data at the previous single shot position. Two elements are important to distinguish between RWI and DWI: (1) offset versus shot; (2) predicting all data sets versus predicting next shot gather. In chapter 4, I will extend the image-domain method to data-domain formulating an approach similar to DWI, and then investigate the link between image-domain and data-domain methods. I conduct a more detailed analysis of DWI in Chapter 4.

2.4 Image coherency criteria

In this section, I introduce more mathematical details on the image-domain methods which will be mainly investigated in this thesis. The principle of image-domain methods is to exploit the redundancy of seismic data after migration. As mentioned in section 1.3, a collection of migrated images can be constructed with different subsets of the data, for example one image for each shot gather, in a given background velocity model. Due to the fact that the Earth's model is unique, the correct velocity leads to coherent images, meaning that the recovered reflectors resulting from different shot experiments should be consistent. Otherwise, the images present discrepancies due to inaccuracies in the estimated macro velocity model. Such a class of macro model estimation methods based on the migration technique is called Migration Velocity Analysis (MVA). According to the formalism introduced by Symes and Carazzone (1991); Symes and Kern (1994); Chauris et al. (2002a); Biondi and Symes (2004); Symes (2008); Sava and Vasconcelos (2011) among others, the general framework of a MVA procedure consists of three essential steps:

- the definition of the OF: in contrary to the natural data-fitting strategy, the image-domain method seeks an OF assessing the quality of velocity model by measuring the image quality through the coherency or focusing criterion. A generic definition reads

$$\mathcal{J}_{\text{mva}}[c_0] = \frac{1}{2} \|A(\sigma)\xi[c_0](\mathbf{x}, \sigma)\|^2 \quad (2.54)$$

where operator A is the annihilator as a function of the redundant parameter σ . I define reflectivity ξ as $2\delta c/c_0^3$ in this study. \mathbf{x} denotes the spatial coordinate (x, z) in 2D. A and σ are chosen based on the selected criterion and associated parameterization detailed later. The OF is known to have a convex behavior over a large range of velocities (Stolk and Symes, 2002; Van Leeuwen and Mulder, 2009; Symes, 2008), such that one can minimize the value of \mathcal{J}_{mva} with the gradient based optimization method. The analysis is performed on panels called Common Image Gathers (CIGs) consisting of a section of reflectivity volume as a function of depth z and of the redundant parameter σ for a fixed lateral position x in 2D.

- the construction of images ξ parameterized with an extra parameter σ representing the redundancy of the data: for example in 2D, the migrated volume is $\xi(x, z, \sigma)$, a function of lateral position x , depth z and the redundant parameter σ . Then the image and the data have the same dimension – (x, z, σ) versus (s_x, r_x, t) . This is a key point of MVA strategies: even in an incorrect background velocity model, all information may be preserved during migration of the data, in the sense that data are supposed to be able to be re-modeled in the same background velocity model using the migrated volume. However, in practice, migration is the adjoint operator of modeling such that the reproduced data are only kinematically correct (see Figure 1.12 from Chapter 1);
- the iterative update of the background velocity model c_0 : for a single iteration, this includes the computation of the gradient of \mathcal{J}_{mva} with respect to c_0 which can be efficiently derived with the adjoint-state method (Plessix, 2006), and the subsequent velocity update using the local optimization scheme such as a conjugate gradient scheme.

As mentioned in Chapter 1, MVA is formulated either as a *surface-oriented* approach in which the redundant parameter is related to the acquisition (e.g. shot position, shot-receiver offset, etc.) (Al-Yahya, 1989; Symes and Carazzone, 1991; Chauris and Noble, 2001; Mulder and ten Kroode, 2002; Shen and Symes, 2008 among others) or a *depth-oriented* approach where the migration is extended during the construction of the images (e.g. space lag, time lag, etc.) (Sava and Biondi, 2004; Biondi and Symes, 2004; Symes, 2008; Fei and Williamson, 2010; Shen and Symes, 2013; Mulder, 2014; Shen and Symes, 2015 among others). I first introduce the details of the essential steps for the two families, and then discuss the current limitations of MVA.

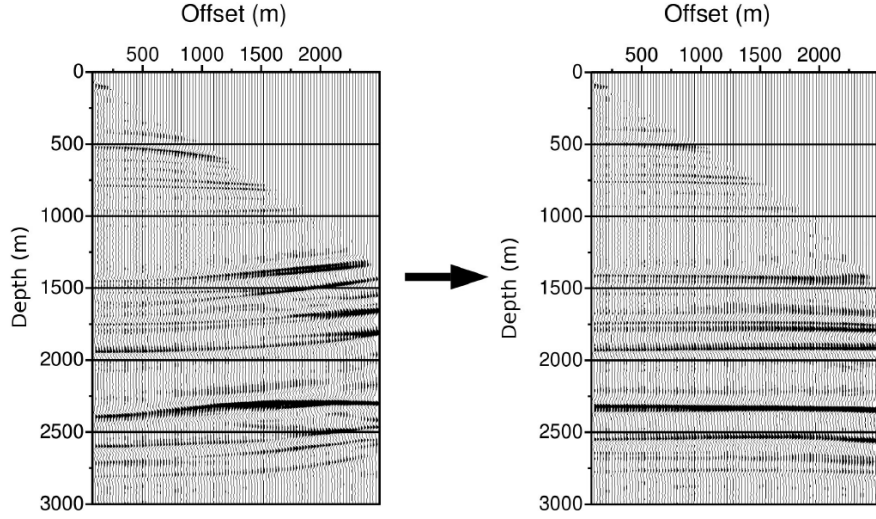


Figure 2.11 – CIGs related to incorrect (left) and correct (right) macro velocity models (from [Chauris et al., 2002b](#)).

2.4.1 Surface-oriented MVA

A natural choice of the extra parameter is the acquisition-related parameter such as the surface-offset (source-receiver distance) ([Chauris and Noble, 2001](#); [Rickett and Sava, 2002](#); [Mulder and ten Kroode, 2002](#) among others) or the shot position ([Symes and Kern, 1994](#); [Huang and Symes, 2015](#) among others). The images should be identical when they are constructed from several experiments related to different values of the selected extra parameter. Consequently, events on the associated CIGs should be horizontal for the correct velocity model (Figure 2.11). The model inaccuracy can thus be described by analyzing the CIGs: the events curve upward for velocity model too slow and downward for a velocity model too fast. The DSO technique ([Symes and Carazzone, 1991](#)) measures the coherency of events on CIGs,

$$\mathcal{J}_{\text{smva}}[c_0] = \frac{1}{2} \|D_{s_x} \xi[c_0](\mathbf{x}, \mathbf{s})\|^2, \quad (2.55)$$

where the annihilator A and the extra parameter σ in equation 2.54 are replaced by the horizontal partial derivative D_{s_x} and the shot position s_x , respectively. This type of OF computes the derivative of the images with respect to redundant parameter to measure the consistency of events in CIGs.

In the surface-oriented case, the reflectivity ξ can be derived as a prestack RTM approach,

$$\xi[c_0](\mathbf{x}, \mathbf{s}) = \int dt \frac{\partial^2}{\partial t^2} S_0(\mathbf{s}, \mathbf{x}, t) R_0(\mathbf{s}, \mathbf{x}, t), \quad (2.56)$$

which represents the crosscorrelation between forward and backward propagated wavefields similar to equation 2.48. The main difference is that there is no integral over the source position \mathbf{s} in equation 2.56. The gradient of $\mathcal{J}_{\text{smva}}$ with respect to c_0 is derived using the adjoint-state technique, reading

$$\frac{\partial \mathcal{J}_{\text{smva}}}{\partial c_0(\mathbf{x})} = \frac{2}{c_0(\mathbf{x})^3} \iint d\mathbf{s} d\mathbf{x} \frac{\partial^2}{\partial t^2} (\lambda_s(\mathbf{s}, \mathbf{x}, t) S_0(\mathbf{s}, \mathbf{x}, t) + \lambda_r(\mathbf{s}, \mathbf{x}, t) R_0(\mathbf{s}, \mathbf{x}, t)), \quad (2.57)$$

where λ_s and λ_r are the adjoint state variables denoted by,

$$\mathcal{L}_0 \lambda_s(\mathbf{s}, \mathbf{x}, t) = \lambda_\xi(\mathbf{s}, \mathbf{x}) \frac{\partial^2}{\partial t^2} R_0(\mathbf{s}, \mathbf{x}, t), \quad (2.58)$$

$$\mathcal{L}_0^* \lambda_r(\mathbf{s}, \mathbf{x}, t) = \lambda_\xi(\mathbf{s}, \mathbf{x}) \frac{\partial^2}{\partial t^2} S_0(\mathbf{s}, \mathbf{x}, t), \quad (2.59)$$

$$\lambda_\xi(\mathbf{s}, \mathbf{x}) = \frac{\partial J_{\text{smva}}}{\partial \xi(\mathbf{s}, \mathbf{x})} = -\frac{\partial^2}{\partial s_x^2} \xi(\mathbf{s}, \mathbf{x}). \quad (2.60)$$

The *image residual* λ_ξ acts comparably as the scattering structures in RWI and DWI such that λ_r and λ_s have the same behavior as the forward and back scattered wavefields, respectively. The formulation of gradient (e.g. equations 2.48, 2.49 and 2.57) always contains crosscorrelations between different wavefields. The crosscorrelation contributes to the gradient in a migration mode if two wavefields only coincide with each other at interfaces, whereas the contribution to gradient is in a tomographic mode if two wavefields share the same wavepath. The wavepaths of λ_r and R_0 (λ_s and S_0) are shown in Figure 2.12a (Figure 2.12b) corresponding to the latter, and thus lead to a tomographic update of the model.

However, the conventional surface-oriented method is known to suffer from kinematic artifacts when complex wavepaths are considered (Xu et al., 2001; Stolk and Symes, 2004). The depth-oriented extended model presented in the next paragraph introduce a better way to construct images, less prone to those artifacts (Stolk and de Hoop, 2005). In this thesis, I propose a new approach by coupling common-shot true-amplitude migration to MVA in Chapter 3.

2.4.2 Depth-oriented MVA

In the depth-oriented formulation of MVA, one introduces the redundant parameter in the wavefield crosscorrelation during the construction of migrated images. The image-domain can be extended with the spatial lag (Rickett and Sava, 2002; Shen et al., 2005; Lameloise et al., 2015; Chauris and Cocher, 2017), the temporal shift (Sava and Fomel, 2006; Yang and Sava, 2011) or the scattering angle (Sava and Fomel, 2003; Biondi and Symes, 2004). The general extended migration formula with both the spatial lag \mathbf{h} and

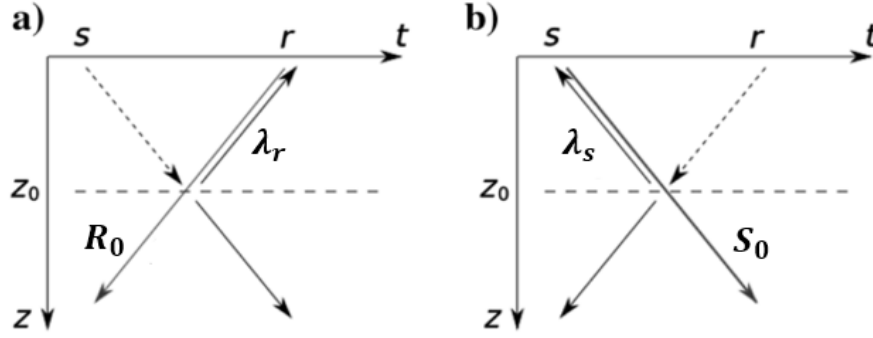


Figure 2.12 – Schematic view of the wavefields (a, from the source s ; b, from the receiver r) contributing to the gradient in the presence of a horizontal scattering source (e.g. a single reflector) at depth z_0 (adapted from [Chauris and Cocher, 2017](#)). The parameters S_0 and λ_s as well as R_0 and λ_r correlate between the surface and depth z_0 .

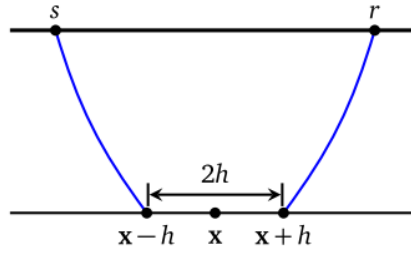


Figure 2.13 – Description of the horizontal subsurface-offset h .

the temporal shift τ ([Sava and Vasconcelos, 2011](#)) read

$$\xi[c_0](\mathbf{x}, \mathbf{h}, \tau) = \iint dt ds \frac{\partial^2}{\partial t^2} S_0(\mathbf{s}, \mathbf{x} - \mathbf{h}, t - \tau) R_0(\mathbf{s}, \mathbf{x} + \mathbf{h}, t + \tau), \quad (2.61)$$

The image section at $\mathbf{h} = 0$ and $\tau = 0$ corresponds to the migrated section obtained with RTM formula 2.48 with summation over all sources and receivers. Energy focusing at non-zero values of the extension parameters corresponds to inaccuracies of the background velocity model. This is the so-called image focusing principle ([Symes, 2008](#)).

In practice, the extension is usually made along a single extension parameter only, to limit the image size and the computational time ([Shen and Symes, 2008](#)). The spatial lag \mathbf{h} is commonly referred to as *subsurface-offset* and $\mathbf{h} = (h_x, h_z)$ in 2D. In the case of surface acquisition and mainly horizontal structures, \mathbf{h} is constrained to the horizontal direction $(h_x, 0)$ (Figure 2.13) and this is the common choice ([Shen and Symes, 2008](#)).

With $\mathbf{h}=(h_x, 0)$, CIGs represent the migrated section at fixed lateral position x and depend on the depth z and the subsurface-offset h . The defocused energy spread over non-zero \mathbf{h} carries the information about inaccuracies of the estimated macro velocity

model. Events are supposed to have a downward (upward) curvature for a velocity model too low (too high) (Mulder, 2014; Lameloise and Chauris, 2014). The quality of velocity model can be assessed by an OF that penalizes the defocused energy, reading

$$\mathcal{J}_{\text{dmva}}[c_0] = \frac{1}{2} \|A(\mathbf{h})\xi[c_0](\mathbf{x}, \mathbf{h})\|^2, \quad (2.62)$$

where $A(\mathbf{h}) = |\mathbf{h}|$ (i.e. h_x for $\mathbf{h} = (h_x, 0)$). Similar to surface-oriented MVA, one minimizes the OF using the local optimization scheme and derive the gradient with adjoint-state method, reading

$$\frac{\partial \mathcal{J}_{\text{dmva}}}{\partial c_0(\mathbf{x})} = \frac{2}{c_0(\mathbf{x})^3} \iint \text{dsd}\mathbf{x} \frac{\partial^2}{\partial t^2} (\lambda_s(\mathbf{s}, \mathbf{x}, t) S_0(\mathbf{s}, \mathbf{x}, t) + \lambda_r(\mathbf{s}, \mathbf{x}, t) R_0(\mathbf{s}, \mathbf{x}, t)), \quad (2.63)$$

where the adjoint state variables λ_s , λ_r and λ_ξ are redefined by

$$\mathcal{L}_0 \lambda_s(\mathbf{s}, \mathbf{x}, t) = \int d\mathbf{h} \lambda_\xi(\mathbf{x} + \mathbf{h}, \mathbf{h}) \frac{\partial^2}{\partial t^2} R_0(\mathbf{s}, \mathbf{x} + 2\mathbf{h}, t), \quad (2.64)$$

$$\mathcal{L}_0^* \lambda_r(\mathbf{s}, \mathbf{x}, t) = \int d\mathbf{h} \lambda_\xi(\mathbf{x} - \mathbf{h}, \mathbf{h}) \frac{\partial^2}{\partial t^2} S_0(\mathbf{s}, \mathbf{x} - 2\mathbf{h}, t), \quad (2.65)$$

$$\lambda_\xi(\mathbf{x}, \mathbf{h}) = \frac{\partial J_{\text{dmva}}}{\partial \xi(\mathbf{x}, \mathbf{h})} = a(\mathbf{h})^2 \xi(\mathbf{x}, \mathbf{h}). \quad (2.66)$$

In the inverse theory, one should account for both the gradient and the Hessian of the OF to update the model, but the computation of the Hessian matrix is prohibitively expensive for industrial application. Shen and Symes (2015) proposed a method to approximate the Hessian in an efficient way. An alternative is the quasi-Newton strategy (Nash and Nocedal, 1991) and the truncated Newton approach (Métivier et al., 2013).

The components in equations 2.63–2.66 indicate that the surface-oriented and the depth-oriented MVA methods behave in the similar tomographic mode but with two different descriptions of the macromodel inaccuracies, in the sense that the formulations of image residual λ_ξ are different. Despite a tomographic update, the gradient of those MVA methods are not always smooth, meaning that artifacts are present in the gradient. I will explain the reasons and possible solutions in the following section.

2.4.3 Limitations of MVA

MVA was initially proposed for the reflected data (Symes and Carazzone, 1991). Despite the Born approximation, recent studies have extended MVA to transmitted waves (Chauris et al., 2013; Shen and Symes, 2013; Biondi and Almomin, 2014; Lameloise and Chauris, 2016) and to multiples (Staal and Verschuur, 2012; Cocher et al., 2015, 2017b). However, not so many applications on real data have been published (Chauris and Noble, 2001; Mulder and ten Kroode, 2002; Alkhalifah, 2005; Shen and Symes, 2008; Mulder, 2014; Lameloise and Chauris, 2016 among others). This section reviews several challenges of MVA which prevent from broader adoption.

Computational cost

In general, the MVA technique is more numerically demanding than FWI in terms of computation and memory requirements (van Leeuwen et al., 2015). In the case of subsurface-offset MVA, the model is extended only along the horizontal component of the subsurface-offset in order to limit the memory requirements. But, one needs to store the full migrated volume for all spatial and subsurface-offset coordinates. Thus, it is prohibitively expensive to extend the subsurface-offset MVA to 3D applications, where two extra dimensions are in principle needed for the model space to match the data size. Yang and Sava (2015) propose to compute CIGs only at a limited number of image points rather than the whole image. van Leeuwen et al. (2015) construct CIGs with only a random choice of data traces. Fu and Symes (2015) investigate a multiscale strategy in which less and less subsurface-offset is required to be considered during the iterative update. Note that even if these techniques may decrease the computational cost for subsurface-offset MVA, the numerical requirements are still heavy. We investigate in Chapter 3 if and how common-shot MVA can be less numerically expensive and provide a more efficient implementation.

Migration smiles

Conventional MVA uses the reflectivity resulting from the classic migration which is the adjoint of the Born modeling operator as already mentioned. This kind of image is only the first gradient of the ℓ_2 norm OF minimizing the data misfit. As an inaccurate solution to the inverse problem, it may suffer from uneven illumination in a velocity model containing complex structures. In the presence of salt body, Yang et al. (2013) indicate that the uneven illumination results in obviously defocused energy visible at non-zero spatial lag in a correct velocity model.

In the subsurface-offset domain, migration smiles are visible in CIGs, even for a very simple model such as a single reflector embedded in the homogeneous velocity (Figure 2.14a). The smiles always present upward curvatures and do not focus for the correct velocity. As a result, these smiles introduces a bias in the estimation of the optimal velocity model, meaning that the OF is minimal for a velocity model usually slightly lower than the correct velocity model (Mulder, 2014; Lameloise, 2015): the curvature of smiles is similar to the one related to too high velocity model. The associated gradient thus can not provide a satisfactory homogeneous update (Figure 2.14b).

To solve this problem, the iterative migration is preferred (Liu et al., 2014b; Chauris et al., 2015) since it provides indeed an inverse of the modeling operator, rather than the classic migration, which is only the adjoint operator. However, the iterative migration usually requires several loops before convergence and thus MVA becomes a nested optimization problem with a more complex gradient computation (Cocher et al., 2017b). This increases the overall numerical cost for a successful MVA. It may also lead to

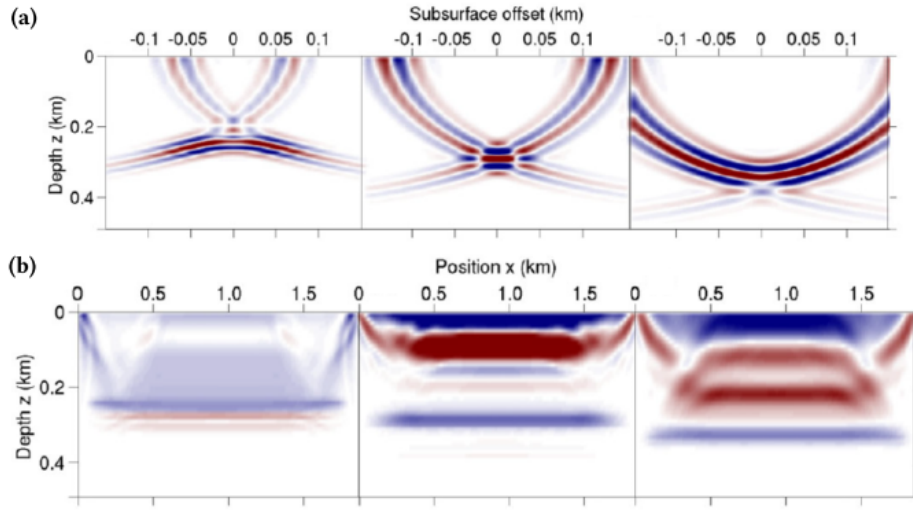


Figure 2.14 – (a) CIGs and (b) associated gradients of the MVA OF computed with classic migration for a too low (left), correct (middle), and too high (right) velocity model. Blue, white and red colors correspond to negative, null and positive values, respectively (from [Lameloise, 2015](#)).

unstable results ([Cocher, 2017](#)) On the other hand, a sophisticated solution consists of replacing the classic migration by the direct inverse formula ([Lameloise and Chauris, 2014](#); [Hou and Symes, 2015](#)). [Lameloise and Chauris \(2014\)](#) develop the ray+Born inversion approach proposed by [Lambaré et al. \(1992\)](#) to the subsurface-offset domain and [ten Kroode \(2012\)](#); [Hou and Symes \(2015, 2016b\)](#); [Chauris and Cocher \(2017\)](#) further extend it to wave-equation-based operators as a direct inverse. In practice, the quantitative weights compensate for uneven illumination and geometrical spreading losses in the migrated images such that smiles are significantly attenuated (Figure 2.15a) and the associated gradient is much cleaner (Figure 2.15b). Moreover, it can be either regarded as a direct true-amplitude migration approach ([Hou and Symes, 2016b](#); [Chauris and Cocher, 2017](#)) or used as the preconditioner boosting the iterative migration ([Hou and Symes, 2016b](#); [Cocher et al., 2017b](#)) or FWI ([Qin et al., 2015](#)), leading to a more efficient implementation.

The impact of migration smiles can be observed in the surface-oriented domain as well as in the subsurface-offset domain (Chapter 3). In this thesis, I will follow the same strategy discussed above, to extend a ray-based quantitative migration ([Beylkin, 1985](#); [Bleistein, 1987](#)) into a wave-equation-based inverse formula for the common-shot gathers, which is a minor modification on the work of [Qin et al. \(2015\)](#), and more importantly to couple it to the velocity analysis in Chapter 3.

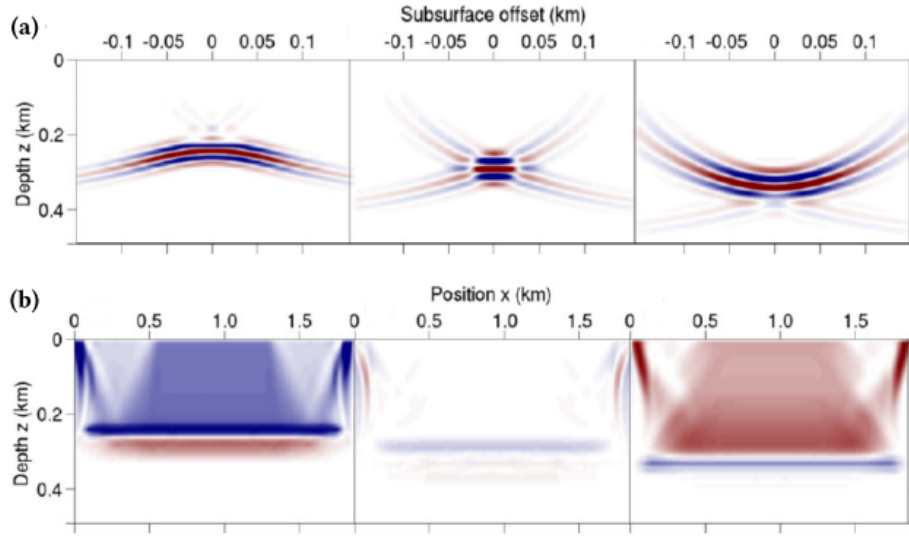


Figure 2.15 – Same as Figure 2.14 with a quantitative (ray-based) migration instead of a classic (adjoint) migration (from Lameloise, 2015).

Gradient "artifacts"

In the subsurface-offset MVA gradient, strong oscillations are present (Figure 2.16a), especially around diffractors and/or along the reflectors (Vyas et al., 2010; Fei and Williamson, 2010). They are known as *gradient artifacts* which prevent the MVA gradient from being directly used for updating the velocity model. In the case of a complex velocity model, these oscillations will be present everywhere, even altering the gradient direction (Lameloise et al., 2015). A simple prior smoothing on the gradient cannot completely destroy these oscillations (Chauris and Cocher, 2017). I will discuss the equivalent issues in the case of common-shot MVA in Chapter 3.

Note that these oscillations are not really artifacts of the gradient: their presence is due to the fact that when the macromodel velocities increase, the positions of the reflectors are deeper.

In the context of subsurface-offset MVA, Fei and Williamson (2010) propose the introduction of a h derivative in the image residual λ_ξ before the computation of the gradient of the MVA OF. Shen and Symes (2015) recognize it as a warping technique, namely *horizontal contraction*. The h derivative makes CIGs contracted in the h direction, leading to the energy more focused at zero subsurface-offset. The trick significantly improves the quality of the velocity update (Figure 2.16b), but the modification is no more the gradient of an OF. More importantly, such a technique does not properly handle low velocity anomalies (Shen and Symes, 2015). Alternatively, Chauris and Cocher (2017) propose to multiply the recovered images with the velocity at a specific

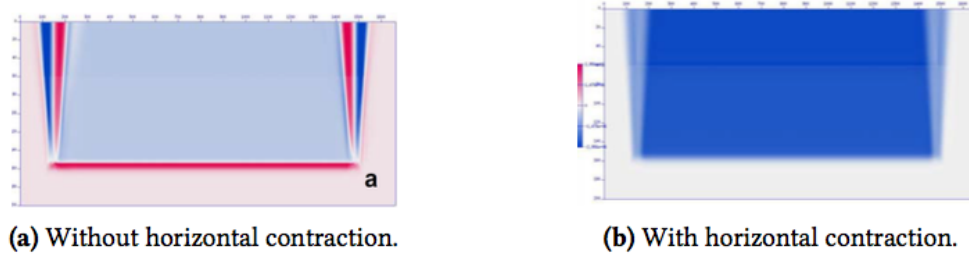


Figure 2.16 – The MVA gradients for a homogeneous model with a single horizontal reflector before (a) and after (b) removing gradient artifacts by horizontal contraction (from [Fei and Williamson, 2010](#)).

power before evaluating the MVA OF. Then, the oscillatory terms around the interface positions can be removed, and the artifacts related to discontinuities can be attenuated using the Gaussian smoothing. I will introduce this modification for common-shot MVA in Chapter 3.

Discussion

In the subsurface-oriented domain, one can avoid some limitations of MVA by replacing the migration, an adjoint operator of modeling, by the inverse, to formulate a new approach called Inversion Velocity Analysis (IVA). Such an inverse can be derived by iterative migration ([Liu et al., 2014b](#); [Chauris et al., 2015](#); [Cocher et al., 2017b](#)) or direct inversion ([Hou and Symes, 2015](#); [Chauris and Cocher, 2017](#); [Hou and Symes, 2016a](#)). For the former, it is important to introduce a regularization in the migration step ([Huang and Symes, 2015](#); [Hou and Symes, 2016b](#); [Chauris et al., 2015](#); [Cocher, 2017](#)) to stabilize the derivation of the optimal macromodel. For the latter, the key is to derive a formula which is the inverse of Born modeling operator ([ten Kroode, 2012](#); [Hou and Symes, 2015](#); [Chauris and Cocher, 2017](#)). For the inner loop of a nested optimization problem, direct inversion is favored over iterative migration due to its relatively lower numerical cost.

However, the subsurface-offset method is very demanding in terms of the computational cost and the memory requirement, due to the extended dimension, preventing from 3D implementation. Briefly, migration smiles and gradient artifacts motivate the need for considering an inverse operator, and the numerical requirements motivate the need to consider direct inverse scheme and common-shot gathers. Consequently, in Chapter 3, I develop a direct inverse formula for common-shot gathers and then couple it to MVA for better robustness and efficiency. The numerical requirements between common-shot and subsurface-offset methods are compared as well in Chapter 3. It is known that common-shot approach is biased in the presence of triplicated wavefields ([Stolk and Symes, 2004](#)). There is no reason why the direct inverse would solve this

issue. We investigate in Chapter 3 how triplications impact MVA when direct common-shot is coupled to velocity analysis.

2.5 Summary

We reviewed various velocity estimation approaches formulated in either data or image domain in this Chapter. In section 1.2, I explained the ideas behind FWI/RWI/DWI and indicated the importance of investigating the link between data-domain and image-domain methods. Note that such a DSO functional was initially defined as the regularization term for FWI (Symes and Kern, 1994). The investigations are detailed in Chapters 4. In section 1.3, I presented the current limitations of MVA

All introduced methods can be summarized as a generic workflow (Figure 2.17a). Defining the inverse problem consists of several essential steps: parameterization, modeling, OF and the computation of the gradient of the OF. Such a framework can easily distinguish different macro velocity estimation approaches (Figure 2.17a).

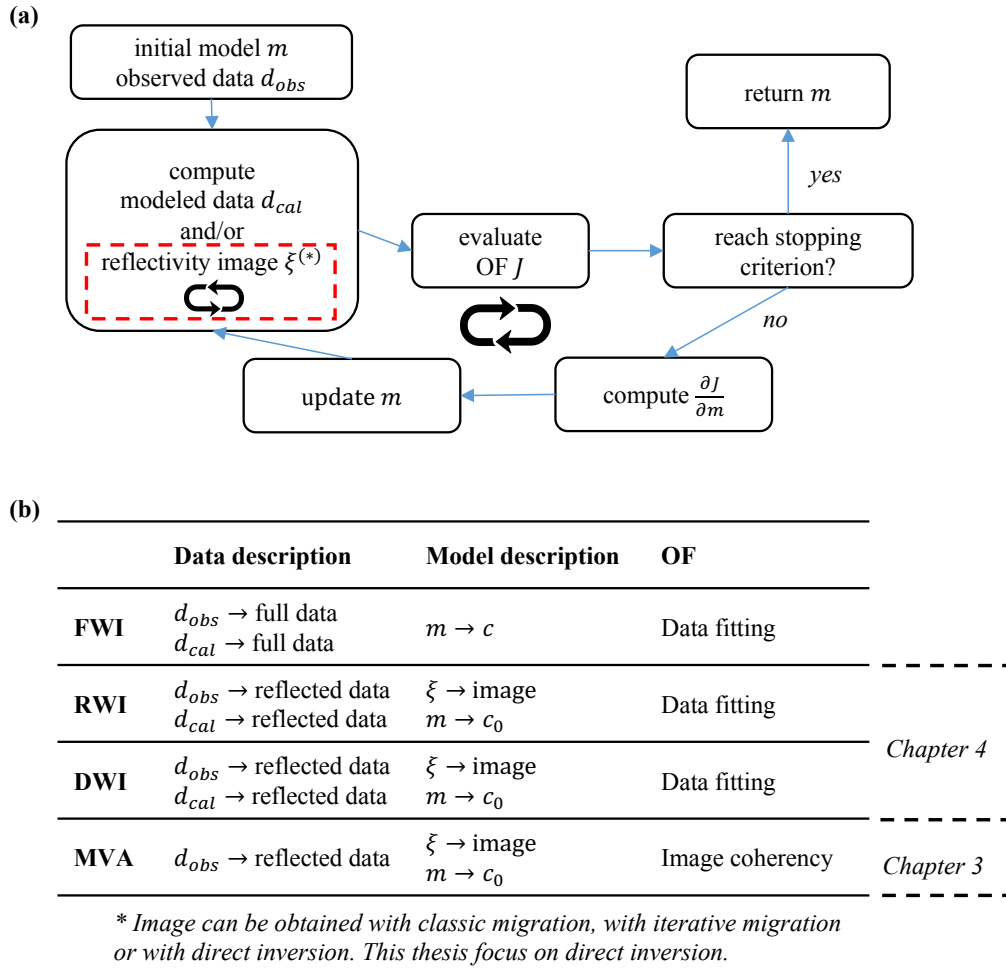


Figure 2.17 – A generic workflow for solving the inverse problem in both data and image domains (a). The differences among various approaches (b) in terms of data and model descriptions, and the criteria for the definition of OF.

Chapter 3

Common-shot Inversion Velocity Analysis in the image-domain

Contents

3.1	Introduction	75
3.2	From migration to inversion	79
3.3	From MVA to IVA	82
3.4	Numerical examples	84
3.4.1	Homogeneous model	84
3.4.2	Low-velocity anomaly	86
3.4.3	Marmousi model	97
3.5	Discussions	108
3.6	Conclusions	111
3.7	Appendix I: Common-shot inversion scheme	112
3.8	Appendix II: Gradient derivation for IVA	117

(Adapted from the article submitted to *Geophysics*.)

Résumé du chapitre 3

L'analyse de vitesse par migration est une technique pour estimer les grandes structures du modèle de vitesse de propagation des ondes. C'est ce modèle qui contrôle la cinématique de la propagation. Pour des résultats plus stables, des travaux récents ont proposé de remplacer la migration, l'adjoint de l'opérateur de modélisation par Born, par un inverse direct et ceci dans le contexte du domaine étendu avec l'offset en profondeur. En suivant la même stratégie, nous avons développé un schéma d'analyse de vitesse par inversion pour la migration classique par point de tir. Nous utilisons la semblance différentielle pour évaluer la qualité des images qui dépendent de la position de la source et pour obtenir le gradient de la fonction objective, un élément essentiel pour remettre à jour le macro-modèle de vitesse. Nous discutons des avantages et des limites au travers d'applications sur des données synthétiques en 2D, d'abord sur des modèles très simples avec un seul réflecteur au sein d'un modèle homogène, puis sur le modèle Marmousi. L'inverse direct atténue les sourires de migration car il compense pour le facteur d'atténuation géométrique et pour l'illumination variable de la sub-surface. Des oscillations parasites autour des interfaces dans les gradients de vitesse peuvent être supprimées en multipliant la réflectivité inversée par la vitesse à une certaine puissance, avant de mesurer la cohérence des images. Nous regardons de près la présence de triplications du champ d'onde qui viennent d'anomalies lentes de vitesse. Il apparaît que l'analyse de vitesse par inversion est robuste, même si des artefacts sont observés dans les images migrées. Le gradient de vitesse conduit à une mise-à-jour stable, tout particulièrement après un filtrage Gaussien sur une longueur d'onde. Le couplage de l'inverse par point de tir avec l'analyse de vitesse offre de nouvelles possibilités pour l'extension à la 3D dans le futur.

Abstract

Migration Velocity Analysis is a technique to estimate the large-scale structure of the subsurface velocity model controlling the kinematics of wave propagation. For more stable results, recent studies proposed to replace migration, adjoint of Born modeling, by the direct inverse of the modeling operator in the context of extended subsurface-offset domain. Following the same strategy, we develop a two-way wave-equation-based Inversion Velocity Analysis approach for the original and more efficient surface-oriented shot gathers. We use the Differential Semblance Optimization objective function to evaluate the quality of inverted images depending on shot positions and to derive the associated gradient, an essential element to update the macro-model. We discuss the advantages and limitations through applications of 2D synthetic data sets, first on simple models with a single-reflector embedded in various background velocities and then on the Marmousi model. The direct inverse attenuates migration smiles by compensating for geometrical spreading and uneven illuminations. Spurious oscillations around interface positions in the velocity gradient can be removed by multiplying the inverted reflectivity with the velocity at a specific power before measuring the consistency of the migrated images. We pay attention to the presence of triplicated wavefields caused by low velocity anomalies. It appears that Inversion Velocity Analysis is robust even if artifacts are observed in the seismic migrated sections. The velocity gradient leads to a stable update, especially after a Gaussian smoothing over a wavelength distance. Coupling common-shot direct inversion to velocity analysis offers new possibilities for the extension to 3D in the future.

3.1 Introduction

Migration Velocity Analysis (MVA) is a technique to determine the large-scale structure of the subsurface velocity from seismic surface acquisitions. It assumes that the model parameters can be split into a large-scale model or macromodel controlling the kinematics of wave propagation and a model perturbation generating scattering waves. Under the Born approximation, the reflected data linearly depend on the reflectivity model. Classical images are obtained by crosscorrelating the forward wavefield with the backward wavefield ([Claerbout, 1971](#)). As the real Earth model is unique, images corresponding to different subsets of the data are supposed to be consistent in a correct background model, at least in the well illuminated regions ([Symes, 2008](#)). Migration Velocity Analysis can be associated to different migration schemes, such as common-shot, common-offset or common-angle gathers ([Symes and Carazzone, 1991](#); [Chauris and Noble, 2001](#); [Mulder and ten Kroode, 2002](#); [Biondi and Symes, 2004](#); [Sava and Biondi, 2004](#); [Stolk et al., 2009](#)). More recent alternatives introduced before crosscorrelation a spatial or a temporal shift in the imaging condition, for example the extended

Table 3.1 – Dimensions of the data, extended and shot domains. s and r are the source and receiver coordinates, respectively. t is the time. (x, y, z) are the spatial coordinates. h is the subsurface-offset.

Dimension	Data domain	Extended domain	Shot domain
1D	t	z	z
2D	(s, r, t)	(x, z, h)	(x, z, s)
3D	(s_x, s_y, r_x, r_y, t)	(x, y, z, h_x, h_y)	(x, y, z, s_x, s_y)

subsurface-offset, instead of splitting the input data set before migration (De Bruin et al., 1990; Sava and Fomel, 2006; Symes, 2008) (Table 3.1).

MVA is expected to be a powerful automatic imaging tool. Historically, MVA has been developed for ray-based methods and then extended to wave-equation-based strategies. Wave-equation MVA is in principle capable of reconstructing complex velocity structures (Sava and Biondi, 2004). Among others, Differential Semblance Optimization (DSO) defines an objective function to evaluate the quality of the background velocity model used for migration and such an estimation is known to have a convex behavior (Symes and Carazzone, 1991; Symes, 2008). The inverted long-wavelength model can provide the initial model for subsequent Full Waveform Inversion (FWI), a technique for generating high-resolution models (Lailly et al., 1983; Tarantola, 1984a). The first MVA step is important to avoid FWI to suffer from cycle-skipping effects (Bunks et al., 1995; Virieux and Operto, 2009). Despite the scale separation and the associated linearization, MVA also has the potential to be extended to transmitted waves (Chauris et al., 2013; Shen and Symes, 2013; Biondi and Almomin, 2014; Lameloise, 2015; Lameloise and Chauris, 2016) or to multiples (Staal and Verschuur, 2012; Cocher et al., 2015, 2017a,b). Nevertheless, Symes (2008); Lameloise et al. (2015) reviewed that only few applications on real data have been published yet. MVA still faces a number of challenges: (1) it is very demanding in terms of computation and memory requirements (van Leeuwen et al., 2015). In the subsurface-offset configuration, the CPU cost related to the crosscorrelations for different space lags is not negligible. Moreover, one needs to store the full migrated volume for all spatial and subsurface offset coordinates, currently preventing from 3D applications except on a very limited size (Chauris and Cocher, 2017). (2) Migration smiles visible in CIGs introduce a bias in the estimation of the optimal velocity model (Mulder, 2014; Lameloise et al., 2015). In the subsurface-offset case, velocities are underestimated: lower velocities introduce a moveout curvature with an opposite sign as the one associated to migration smiles. (3) The gradient of DSO objective function with respect to the macro velocity contains spurious oscillations around the reflector positions, especially in the case of discontinuous reflectivities (Vyas et al., 2010). They are not artifacts, and exist simply for the reason that higher (lower) veloc-

ities introduce deeper (shallower) depth for the reflectivity model. [Fei and Williamson \(2010\)](#) proposed the horizontal contraction method to get rid of these oscillations by artificially modifying the gradient: after modification, it is no more the gradient of any objective function ([Shen and Symes, 2015](#)). [Chauris and Cocher \(2017\)](#) also explains why a simple smoothing on the gradient would not destroy these oscillations without changing the kinematics of wave propagation.

To overcome the difficulties related to the impact of migration smiles on the macro-model gradient, the first attempt is to consider iterative least-squares migration ([Nemeth et al., 1999](#); [Liu et al., 2014b](#); [Chauris et al., 2015](#)) as the quality of migrated images has a significant influence on the velocity analysis. Conventional migration is only designed to produce accurate kinematic images of the subsurface, but is not aimed at preserving the amplitudes in migrated images, whereas inversion promises that the modeled data from the inverted reflectivity perfectly match the observed data. This kind of approaches is a priori rather expensive since the derivation of the optimal reflectivity normally requires several iterations. As an alternative to the iterative schemes, true-amplitude theory or direct approaches were initially proposed for ray-based methods ([Beylkin, 1985](#); [Bleistein, 1987](#); [Lambaré et al., 1992](#); [Xu and Lambaré, 2004](#)) and is a pseudo-inverse of the Born modeling operator. Then, true-amplitude was extended to elastic case ([Forgues and Lambaré, 1997](#)) and to attenuation case ([Ribodetti and Virieux, 1998](#)). Compared to conventional migration, the true-amplitude methods are preferred because the retrieved amplitudes give an approximate estimation of the reflection coefficient and automatically compensate for geometrical-spreading and uneven illumination.

Ray-based methods are flexible and efficient but rays are asymptotic solutions of the wave equation. Such approaches are not always effective in imaging complex structures. One-way wave-equation-based methods use paraxial approximation of the wave equation. One-way wave-equation-based true-amplitude methods have been proposed for both common-shot and common-angle migrations ([Zhang et al., 2005, 2007](#)). These approaches are properly defined within certain angles around the main direction but fail to handle wider angles, especially those near or beyond 90 degrees. Conversely, two-way wave-equation-based reverse time migration (RTM) solves the full wave equation and handle more properly steep dipping angles and complex velocity models. [Op't Root et al. \(2012\)](#); [Qin et al. \(2015\)](#) have proposed a similar explicit asymptotic inverse for shot-profile. [Zhang et al. \(2014b\)](#) have studied how to implement true-amplitude RTM in angle-domain common-image gathers. [ten Kroode \(2012\)](#); [Hou and Symes \(2015\)](#); [Chauris and Cocher \(2017\)](#) have investigated the extension to subsurface-offset domain. These approaches provide asymptotic inversion even if the final expression only contains wave-equation-based operators. For example, the cosine of the take-off angle at the source position is obtained through the derivative of the Green's function with respect to the vertical position of the source. Note that these direct approaches provide proper preconditioner to reduce the number of iterations ([Hou and Symes, 2016b](#); [Duprat and Baina, 2016](#); [Cocher et al., 2017b](#); [Hou and Symes, 2017](#)).

Table 3.2 – Non exhaustive references related to different MVA/IVA approaches.

	Surface-oriented: common-shot, surface-offset, etc.	Depth-oriented: common-angle, subsurface-offset and/or time-lag, etc.
MVA	Al-Yahya (1989); Symes and Kern (1994); Chauris and Noble (2001); Mulder and ten Kroode (2002), etc.	Sava and Biondi (2004); Biondi and Symes (2004); Symes (2008); Fei and Williamson (2010); Shen and Symes (2013); Mulder (2014); Shen and Symes (2015), etc.
IVA	?	Liu et al. (2014b); Chauris et al. (2015); Lameloise et al. (2015); Hou and Symes (2016a); Chauris and Cocher (2017), etc.

Inversion Velocity Analysis (IVA) can be established by coupling such asymptotic inversion to velocity analysis. Artifacts in CIGs are largely attenuated by inversion (Hou and Symes, 2016a; Chauris and Cocher, 2017). IVA has been developed so far in the extended subsurface-offset domain (Liu et al., 2014b; Chauris et al., 2015; Hou and Symes, 2016a,b; Chauris and Cocher, 2017), whereas only a few studies consider the more traditional surface-oriented domains (Table 3.2). The objective of this study is to fill this gap. As discussed later, a main advantage is that shot-profile approach, beyond the natural parallel computation, requires less memory than extended subsurface-offset if considering 3D extension. However, common-shot scheme is known to suffer from migration smiles and cannot properly handle triplicated wavefields (Stolk and Symes, 2004). The first aspect will be compensated by direct inversion. The problem related to caustics will not be solved by the asymptotic inverse, but the coupling with velocity analysis will be discussed here though numerical investigations.

In this article, we first propose a slightly different common-shot wave-equation inversion scheme compared to Qin et al. (2015). This is still a pseudo-inverse as it is an inverse in the asymptotic sense, but better preserves early arrivals in the reconstructed data. We then couple true-amplitude common-shot reverse time migration to velocity analysis, namely the shot-profile IVA and explain how to derive the gradient with respect to the macromodel. The approach is applied to three synthetic cases to discuss the advantages and limitations. First, we consider a single reflector embedded in homogeneous background velocity models, to investigate the impact of inverse operator over classical migration. We pay attention to reflectivity images, CIGs and the first macromodel gradients. We then consider positive and negative Gaussian circular anomaly zones to study the sensitivity of IVA with respect to a high or low velocity anomaly area. In particular,

we discuss the influence of triplicated wavefields. Finally, we perform IVA on the Marmousi model, starting from a homogeneous model. The large-scale structure is retrieved after several iterations. We subsequently perform FWI starting from the final IVA result, to extract the fine details within the well illuminated region.

3.2 From migration to inversion

We derive here a direct inverse scheme for common-shot wave equation migration. This is an alternative to the expression proposed by [Qin et al. \(2015\)](#). We first review the classical Born modeling and its corresponding adjoint operator. We then express the true-amplitude version based on high-frequency approximation, and finally explain how to remove the ray quantities for a pure wave equation based expression. We consider a velocity model c as the sum of the background model c_0 which controls the kinematics of the wave propagation and the model perturbation δc which triggers reflected and diffracted waves. We define the reflectivity as $\xi = 2\delta c/c_0^3$. The model c_0 is iteratively determined by minimizing the DSO objective function in the outer loop procedure, whereas for each c_0 model, an optimal ξ is obtained by solving the inner loop problem, for which we discuss two possibilities, migration and inversion.

Under the Born approximation, modeled reflected data d linearly depend on the perturbation ξ . For common-shot gathers, the classic definition of Born modeling operator \mathcal{B}_0 reads ([Symes, 2008](#))

$$(\mathcal{B}_0\xi)(\mathbf{s}, \mathbf{r}, \omega) = \int d\mathbf{x} (i\omega)^2 \Omega(\omega) G_0(\mathbf{s}, \mathbf{x}, \omega) \xi(\mathbf{x}, \mathbf{s}) G_0(\mathbf{x}, \mathbf{r}, \omega), \quad (3.1)$$

where Ω is the source wavelet depending on the angular frequency ω . The two G_0 are the Green's functions computed in model c_0 . The composition $\mathcal{B}_0\xi$ denotes computed data d in (ξ, c_0) . We consider here the 2D constant density acoustic wave-equation. The model perturbation ξ depends on the spatial coordinates $\mathbf{x} = (x, z)$. Source and receiver positions are denoted by \mathbf{s} and \mathbf{r} , respectively.

For a specific c_0 , migration is introduced by minimizing for each shot the misfit between modeled reflection d and observed reflection d^{obs}

$$\mathcal{J}_0[\xi] = \frac{1}{2} \|d[\xi] - d^{obs}\|^2. \quad (3.2)$$

The first possibility for determining ξ defines the migration ξ_{mig} as $-\frac{\partial \mathcal{J}_0}{\partial \xi}|_{\xi=0}$, the gradient with respect to the reflectivity ξ for $\xi = 0$ ([Lailly et al., 1983](#); [Tarantola, 1984a](#)), yielding

$$(\mathcal{B}_0^T d^{obs})(\mathbf{x}, \mathbf{s}) = \int d\omega \tilde{S}_0^*(\mathbf{s}, \mathbf{x}, \omega) \tilde{R}_0(\mathbf{s}, \mathbf{x}, \omega), \quad (3.3)$$

where \mathcal{B}_0^T denotes the adjoint of the modeling operator \mathcal{B}_0 in model c_0 . There is no summation over the source coordinates to later address the velocity analysis part. The forward and backward wavefields read

$$\tilde{S}_0(\mathbf{s}, \mathbf{x}, \omega) = (i\omega)^2 G_0(\mathbf{s}, \mathbf{x}, \omega) \Omega(\omega), \quad (3.4)$$

$$\tilde{R}_0(\mathbf{s}, \mathbf{x}, \omega) = \int d\mathbf{r} G_0^*(\mathbf{r}, \mathbf{x}, \omega) d^{obs}(\mathbf{s}, \mathbf{r}, \omega). \quad (3.5)$$

For the true-amplitude migration, inversion operator \mathcal{B}_0^\dagger is an inverse of \mathcal{B}_0 in the asymptotic sense. It requires the composition of two operators $\mathcal{B}_0 \circ \mathcal{B}_0^\dagger$ applied to any data outputs equivalent reconstructed data, even in an incorrect background model. It is possible to derive such an inverse formula under the high frequency approximation. We follow the same strategy as the one proposed by [Qin et al. \(2015\)](#); the differences with their approach are introduced later. We first replace the Green's functions in equation 3.1 by their asymptotic version, then derive an inverse formula, and finally replace all terms associated to ray quantities by wave-equation terms. The Green's function is the solution of wave-equation associated with a Dirac distribution in space and time as a source term. With ray-based operators, its analytic formula reads

$$G_0(\mathbf{s}, \mathbf{x}, \omega) = K(\omega) A(\mathbf{s}, \mathbf{x}) e^{i\omega\tau(\mathbf{s}, \mathbf{x})}, \quad (3.6)$$

where $K(\omega) = 1/\sqrt{i\omega}$ in 2D. A is geometrical spreading term, solution of the transport equation, and τ is travel time, solution of the Eikonal equation ([Červený, 1987](#)). Consequently, the Born modeling operator (equation 3.1) reads

$$\begin{aligned} (\mathcal{B}_0 \xi)(\mathbf{s}, \mathbf{r}, \omega) &= \int d\mathbf{x} (i\omega)^2 \Omega(\omega) K^2(\omega) \xi(\mathbf{x}) \\ &\times A(\mathbf{s}, \mathbf{x}) A(\mathbf{x}, \mathbf{r}) e^{i\omega(\tau(\mathbf{s}, \mathbf{x}) + \tau(\mathbf{x}, \mathbf{r}))}. \end{aligned} \quad (3.7)$$

Let us now introduce the associated true-amplitude common-shot inversion formula derived by [Beylkin \(1985\)](#); [Bleistein \(1987\)](#). It contains ray-based terms:

$$\begin{aligned} (\mathcal{B}_0^\dagger d^{obs})(\mathbf{x}, \mathbf{s}) &= - \iint d\mathbf{r} d\omega \frac{4}{c_0^2(\mathbf{x})} \frac{\cos \beta_r}{c_r} \cos^2\left(\frac{\theta}{2}\right) \frac{d^{obs}(\mathbf{s}, \mathbf{r}, \omega)}{\Omega(\omega)} \\ &\times \frac{K^*(\omega) A(\mathbf{x}, \mathbf{r}) e^{-i\omega\tau(\mathbf{x}, \mathbf{r})}}{K(\omega) A(\mathbf{s}, \mathbf{x}) e^{i\omega\tau(\mathbf{s}, \mathbf{x})}}, \end{aligned} \quad (3.8)$$

where c_r is background velocity at the receiver position. β_r is the take-off angle at the receiver position and θ the opening angle (Figure 3.1). $\mathcal{B}_0^\dagger d^{obs}$ is the inverted reflectivity ξ_{inv} in model c_0 for shot gathers d^{obs} . The ratio between the Green's function at the receiver and at the source is a deconvolution imaging condition to balance the amplitudes ([Jones, 2014](#)). Under the high frequency, \mathcal{B}_0^\dagger proves to be the inverse of \mathcal{B}_0 (see

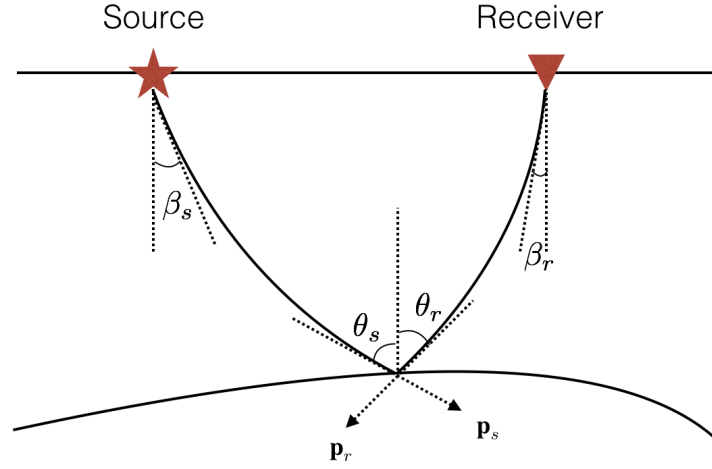


Figure 3.1 – Ray parameters \mathbf{p}_s and \mathbf{p}_r are the slowness vectors at the image point, associated to angles θ_s and θ_r . β_s and β_r are oriented angles at the source and receiver points. We define the opening angle as $\theta = \theta_s - \theta_r$.

Appendix 3.7). The additional weights are referred to as compensation for illuminations (Schleicher et al., 2008; Liu et al., 2011a).

The next step for developing the direct inversion scheme is to remove the ray-based quantities in equation 3.8. The main strategy consists of neglecting the derivatives of the amplitude terms when we apply a spatial gradient to the Green's functions (ten Kroode, 2012; Hou and Symes, 2015; Qin et al., 2015). By substituting those expressions (see details in Appendix 3.7), we finally rearrange equation 3.8 with only wave-equation-based terms, reading

$$(\mathcal{B}_0^\dagger d^{obs})(\mathbf{x}, \mathbf{s}) = 4 \int d\omega (|S_0(\mathbf{s}, \mathbf{x}, \omega)|^2 + \epsilon)^{-1} \\ \times \left(\nabla_{\mathbf{x}} S_0^*(\mathbf{s}, \mathbf{x}, \omega) \cdot \nabla_{\mathbf{x}} R_0(\mathbf{s}, \mathbf{x}, \omega) + \left(\frac{i\omega}{c_0(\mathbf{x})} \right)^2 S_0^*(\mathbf{s}, \mathbf{x}, \omega) R_0(\mathbf{s}, \mathbf{x}, \omega) \right), \quad (3.9)$$

where ϵ denotes a stabilization parameter to avoid divisions by zero, here applied to the Green's functions. Forward and backward propagated wavefields are respectively defined as

$$S_0(\mathbf{s}, \mathbf{x}, \omega) = (i\omega)^3 G_0(\mathbf{s}, \mathbf{x}, \omega) \Omega(\omega), \quad (3.10)$$

$$R_0(\mathbf{s}, \mathbf{x}, \omega) = \int d\mathbf{r} D_{r_z} G_0^*(\mathbf{r}, \mathbf{x}, \omega) d^{obs}(\mathbf{s}, \mathbf{r}, \omega), \quad (3.11)$$

where D_{r_z} denotes the partial vertical derivative at receiver point, here applied to the Green's function. The final formula only contains wave-equation-based operators. It

is similar to the one provided by [Qin et al. \(2015\)](#). They only differ around the source position and in the way to handle early arrivals (see Appendix 3.7 for more details). [Duprat and Baina \(2016\)](#) propose to consider the formula in the time domain as a powerful preconditioner to accelerate least-squares migration. We implement it here in the frequency domain as a simple approach to handle the deconvolution part.

The main differences between migration and inversion are as followed: (1) instead of an imaging condition based on crosscorrelation, the deconvolution imaging condition is applied to balance the amplitudes (equation 3.9). It implies the implementation of an inverted source term Ω^{-1} in inversion; (2) vertical derivative at receiver positions must be applied on the Green's function. It is equivalent to a multiplication by the cosine of the take-off angle at receiver positions. Such a weight emphasizes the vertical rays at receiver positions; (3) the combinations of partial derivatives with respect to the image points, such as gradient and Laplacian, are applied on the Green's functions at the numerator. Such derivatives act as if the deconvolution kernel was multiplied by $\cos^2(\theta/2)$, the squared cosine of the half opening angle at the image points. It reinforces the contribution of short offsets. We now couple the direct inversion to velocity analysis.

3.3 From MVA to IVA

In the nested optimization, the reflectivity is determined by either ξ_{mig} or ξ_{inv} for the inner loop. The outer loop for updating the background model is based on the consistency in the image-domain ([Symes, 2008](#)). For the DSO approach, we minimize the residuals between adjacent traces in CIGs constructed from neighboring shot gathers as

$$\mathcal{J}_{mig}[c_0] = \frac{1}{2} \|D_{s_x} \xi_{mig}\|^2, \quad (3.12)$$

$$\mathcal{J}_{inv}^\alpha[c_0] = \frac{1}{2} \|c_0^\alpha D_{s_x} \xi_{inv}\|^2, \quad (3.13)$$

where D_{s_x} denotes the derivative with respect to the horizontal source position. Similar to the strategy proposed by [Chauris and Cocher \(2017\)](#) in the extended subsurface-offset domain, an extra term $c_0^\alpha(\mathbf{x})$ has been introduced in the inversion case. It allows more flexibility in the shape of the gradient $\frac{\partial \mathcal{J}_{inv}^\alpha}{\partial c_0}$: the additional contribution to the gradient for a non-zero α is localized around the reflectivities and may attenuate the spurious oscillations. The determination of an optimal power α will be later indicated in a case study. We use the adjoint-state technique ([Plessix, 2006](#)) to efficiently derive the gradient (see details in Appendix 3.8). It is a general method to compute the gradient of an objective function that depends on model parameters through a set of state variables – $\tilde{\lambda}_s$ and $\tilde{\lambda}_r$ for migration, or λ_s , λ_r and λ_ξ for inversion. The adjoint-state variables are the solutions of an adjoint linear system. They can be introduced as Lagrangian constraints.

The associated gradients read

$$\begin{aligned} \frac{\partial \mathcal{J}_{mig}}{\partial c_0}(\mathbf{x}) = 2c_0^{-3}(\mathbf{x}) & \left[\iint d\mathbf{s} d\omega (i\omega)^2 \tilde{\lambda}_s^*(\mathbf{s}, \mathbf{x}, \omega) \tilde{S}_0(\mathbf{s}, \mathbf{x}, \omega) \right. \\ & \left. + \iint d\mathbf{s} d\omega (i\omega)^2 \tilde{\lambda}_r^*(\mathbf{s}, \mathbf{x}, \omega) \tilde{R}_0(\mathbf{s}, \mathbf{x}, \omega) \right], \end{aligned} \quad (3.14)$$

$$\begin{aligned} \frac{\partial \mathcal{J}_{inv}^0}{\partial c_0}(\mathbf{x}) = 2c_0^{-3}(\mathbf{x}) & \left[\iint d\mathbf{s} d\omega (i\omega)^2 \lambda_s^*(\mathbf{s}, \mathbf{x}, \omega) S_0(\mathbf{s}, \mathbf{x}, \omega) \right. \\ & + \iint d\mathbf{s} d\omega (i\omega)^2 \lambda_r^*(\mathbf{s}, \mathbf{x}, \omega) R_0(\mathbf{s}, \mathbf{x}, \omega) \\ & \left. - \iint d\mathbf{s} d\omega (i\omega)^2 \lambda_\xi(\mathbf{x}, \mathbf{s}) \frac{4S_0^*(\mathbf{s}, \mathbf{x}, \omega) R_0(\mathbf{s}, \mathbf{x}, \omega)}{|S_0(\mathbf{s}, \mathbf{x}, \omega)|^2 + \epsilon} \right]. \end{aligned} \quad (3.15)$$

The parameters λ_s and λ_r ($\tilde{\lambda}_s$ and $\tilde{\lambda}_r$ for migration) are the scattering wavefields triggered by the product of λ_ξ , the adjoint source, and the back or forward propagated wavefields. Cross-correlations $\lambda_s \star S_0$ and $\lambda_r \star R_0$ contribute to a tomographic update because they have a similar kernel as Migration Based Traveltime Tomography (MBTT) (Chavent et al., 1994) and Reflection Waveform Inversion (RWI) (Xu et al., 2012; Wu and Alkhalifah, 2015). MVA holds a similar tomographic mode update, although migration suffers from smile artifacts and spurious oscillations present in the associated gradient. The third integration term in IVA gradient formula is similar to the deconvolution migration and localized around interfaces. Those components are oscillatory and can be removed by spatially smoothing the gradient.

The gradient for a non-zero α simply reads

$$\frac{\partial \mathcal{J}_{inv}^\alpha}{\partial c_0}(\mathbf{x}) = c_0^{2\alpha}(\mathbf{x}) \left[\frac{\partial \mathcal{J}_{inv}^0}{\partial c_0}(\mathbf{x}) + \alpha \int d\mathbf{s} \frac{(D_{s_x} \xi_{inv})^2}{c_0(\mathbf{x})} \right]. \quad (3.16)$$

Here, the term $\int_{\mathbf{s}} (D_{s_x} \xi_{inv})^2$ always holds positive value and is only present around interface positions. In practice, we first compute the gradient with respect to $\alpha = 0$ and then add the additional term up to different α values to compare their influences on the gradient, without additional resolution of the wave-equation.

MVA contains a tomographic mode but its gradient still suffers from two main issues: (1) migration smiles contribute to the gradient. The associated upwards curvatures lead to over velocity estimations; (2) Spurious oscillations do exist around interfaces (Fei and Williamson, 2010). A simple smoothing does not remove them (Chauris and Cocher, 2017). IVA provides the solution for both problems: (1) we replace migration by inversion that reduces the amplitude of migration smiles; (2) the extra term c_0^α gives

the flexibility to modify the gradient around reflector positions such that we are able to attenuate the oscillations. From our experience, we do not introduce c_0^α in the MVA case because artifacts associated with migration smiles have a larger impact than spurious oscillations. In the next section, we compare the results of MVA and IVA through numerical examples.

3.4 Numerical examples

We investigate three 2D synthetic cases. The first example, a single horizontal reflector embedded in a homogeneous model, gives insights into (1) how the direct inversion influences the shape of the gradient compared migration and (2) what value of α is an optimal choice. The selected α will be used in all subsequent tests. Then, the second case contains a low-velocity anomaly, in which triplicated events exist in the macro-model. Common-shot approaches are a priori not suited in that case (Nolan and Symes, 1996; Prucha et al., 1999; Xu et al., 2001; Stolk and Symes, 2004). Finally, the last test is performed on the Marmousi model containing complex structures, such as discontinuities and rough interfaces. We display only the first gradient for the first example and perform nonlinear iterations to update the velocity model for the last two. In all cases, the data are generated with a 4th-order staggered-grid finite-difference scheme including the Convolution Perfectly Matched Layer (CPML) absorbing boundary (Virieux, 1986; Levander, 1988; Komatitsch and Martin, 2007).

3.4.1 Homogeneous model

We start with a simple model containing a single horizontal reflector at depth $z = 0.6$ km embedded in a homogeneous velocity. We trigger 81 shots every 0.04 km from 2.4 to 5.6 km at the surface. The source function is a Ricker wavelet with a maximum frequency of 10 Hz. Receivers are symmetrically deployed every 0.02 km on both sides of shots with offsets ranging from -1.6 to 1.6 km.

We first migrate and then invert the data in incorrect constant velocity models. In an ideal stacked image section, we expect only horizontal reflectivities below (for velocity too high) or above (for velocity too low) the correct depth. The migration case in reference velocity $c_e = 2.5$ km/s and incorrect velocity $c_0 = 3.0$ km/s exhibits a flat interface at around $z = 0.7$ km (Figure 3.2a). The same holds for exact velocity $c_e = 3.0$ km/s and incorrect velocity $c_0 = 2.5$ km/s except that the interface is located at around $z = 0.5$ km (Figure 3.2c). The main visible difference on the stacked section with the inversion scheme is the deconvolution aspect (Figure 3.2b and 3.2d). More differences are visible on CIGs (Figure 3.3). For the macromodel with a too high velocity, the CIG in the migration case displays events with an upward-curved migration smiles not related to the incorrect model (Figure 3.3a), whereas inversion largely removes those

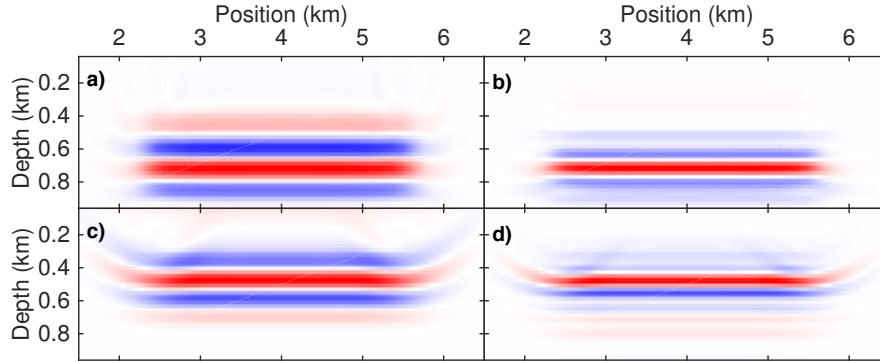


Figure 3.2 – Stacked image sections obtained with migration (a, c) and inversion (b, d) for a single reflector at 0.6 km in incorrect homogeneous models, respectively. Compared to the exact model, the velocity is 0.5 km/s higher for (a, b) and 0.5 km/s lower for (c, d), respectively.

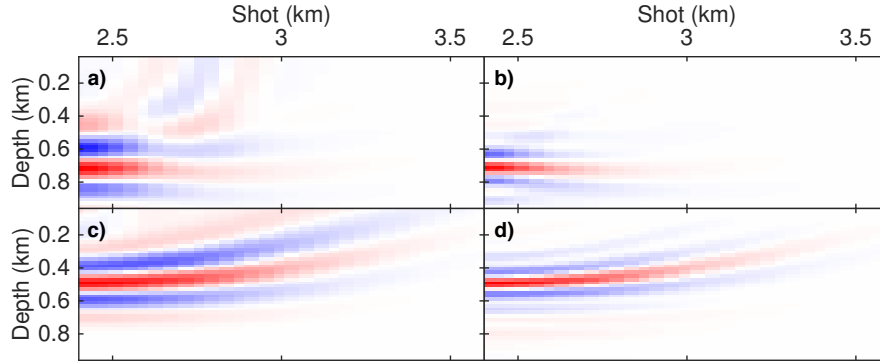


Figure 3.3 – CIGs for position $x = 2.4$ km associated to images for Figure 3.2.

artifacts in CIGs (Figure 3.3b) with the introduction of additional weights such as the take-off and the opening angles (equation 3.8). If the macromodel velocity is too low, the migrated CIG seems to be clean, but the inconsistent energy is partly superimposed to migration smiles (Figure 3.3c). The inverse formula removes this superimposition and thus produces cleaner CIGs (Figure 3.3d).

In an ideal tomographic approach, we expect only homogeneous positive (negative) values for a background velocity higher (lower) than the correct velocity. In practice, MVA gradients are not homogeneous above the reflector, where large-scale positive and negative values are visible (Figures 3.4a and 3.4d). We replace the migration by inversion and build the IVA associated gradients with $\alpha = 0$ (Figures 3.4b and 3.4e). IVA gradients are much more homogeneous than in the MVA case above the interface. The reason is that the inverse operator compensates for amplitude decay and illumination.

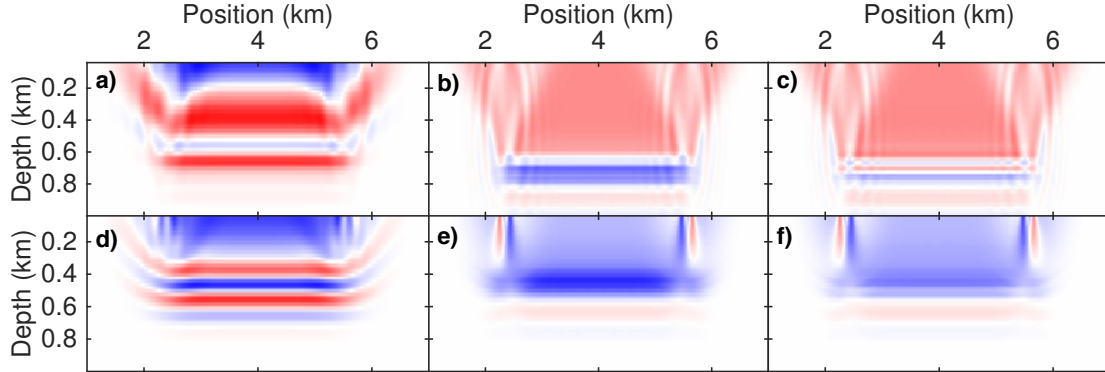


Figure 3.4 – Gradients obtained with classical MVA (a, d), IVA with $\alpha = 0$ (b, e) and $\alpha = 1$ (c, f) for a single reflector at 0.6 km in incorrect invariant models, respectively. Compared to the exact model, the velocity is 0.5 km/s higher for (a–c) and 0.5 km/s lower for (d–f), respectively. Blue, white, red colors mean negative, zero, positive values, respectively.

However, the gradient still has a negative (blue) contribution around the reflector (Figure 3.4b). The reason is that when the velocity is modified, the reflector depth is also modified. Chauris and Cocher (2017) propose an optimal $\alpha = -3/2$ to remove the imprint of the interface but it was derived for the extended domain formulated with the squared slowness model. As we deal here with velocity, an equivalent optimal $\alpha = 1$ is selected to further attenuate oscillations around the reflector position (Figures 3.4c and 3.4f). By optimal, we mean a value such that the gradient has a pure tomographic update after a Gaussian smoothing over half a wavelength of the data (Figure 3.5). The Gaussian smoothing filter is defined as

$$g_{\sigma}(r_0) = \frac{1}{\sqrt{\pi\sigma^2}} e^{-\frac{r_0^2}{2\sigma^2}}, \quad (3.17)$$

where r_0 is the distance. We choose the smoothing parameter $\sigma = \lambda_{data}/2$, where λ_{data} is the mean wavelength of the data.

We here conclude that the IVA approach with $\alpha = 1$ can produce the gradient in a tomographic mode after smoothing, at least for simple models with a continuous interface. The same α value is used for following tests.

3.4.2 Low-velocity anomaly

We now investigate the case of three horizontal reflectors at depth $z = 0.6, 0.8$ and 1.0 km embedded in a velocity model containing a low-velocity zone of up to -0.9 km/s and located close to the surface (Figure 3.6a), leading to multipathing in observed data.

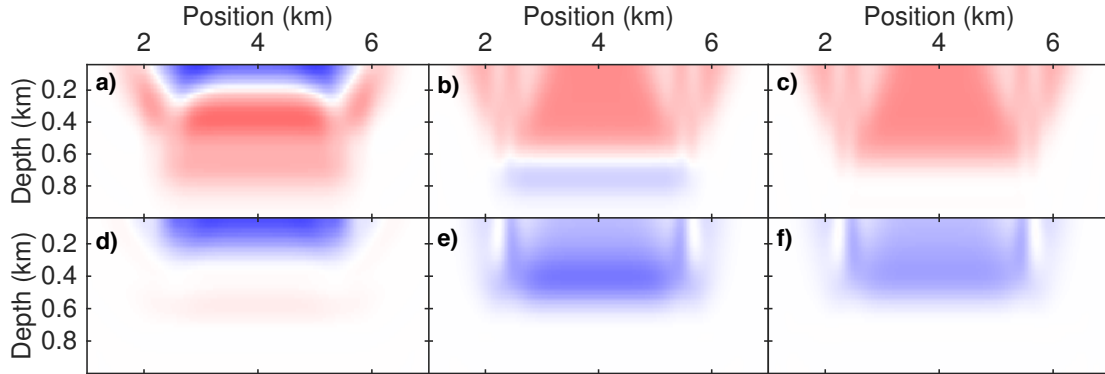


Figure 3.5 – Same as for Figure 3.4, after the Gaussian smoothing over half a wavelength.

The anomaly zone has the shape of a Gaussian lens of characteristic size $r = 0.3$ km. For the subsurface-offset domain, similar examples are discussed in [Shen and Symes \(2015\)](#); [Chauris and Cocher \(2017\)](#): horizontal contraction does not provide a correct answer, whereas IVA in the extended domain properly deals with the issue.

In this example, we trigger 61 shots every 0.04 km from 0.6 to 3.0 km at the surface. Receivers are deployed at fixed positions from 0.0 to 3.6 km every 0.02 km at the surface. After classical migration in a homogeneous model $c_0 = 2.5$ km/s, the stacked migrated section displays deflected energy, especially below the anomaly zone (Figure 3.6b). The associated MVA gradient is expected to display a red (positive) contribution, only at the central part around the anomaly position. This is not the case in Figure 3.6c: the gradient is oscillatory, similarly as a migrated section provides. The positive gradient value has a wider distribution beyond just the center, meanwhile the shallow part of central area shows a blue (positive) update.

Before coupling the inverse scheme to the non-linear optimization, we first check the quality of inversion by comparing observed data and computed data. Note that the reconstruction is computed after inversion and modeling in an incorrect model of constant velocity $c_0 = 2.5$ km/s. The effect of the triplication is visible on the observed common-shot gather (Figure 3.7a). By correct inverse scheme, we mean that inversion followed by modeling in the same model for a single shot can reconstruct data, perfectly matching the observed data. This is here the case (Figure 3.7b and c). Extracted traces for different offsets show that both the phases and amplitudes are correctly recovered (Figure 3.8). With the inversion approach, the stacked image section is similar to the one produced by the migrated case (Figure 3.9a), but the gradient displays a expected red circular zone around the position of the low velocity anomaly (Figure 3.9b). This is a satisfactory first step in model building. After 10 iterations on the background model with a classic nonlinear conjugate gradient scheme, the inverted background model contains a main

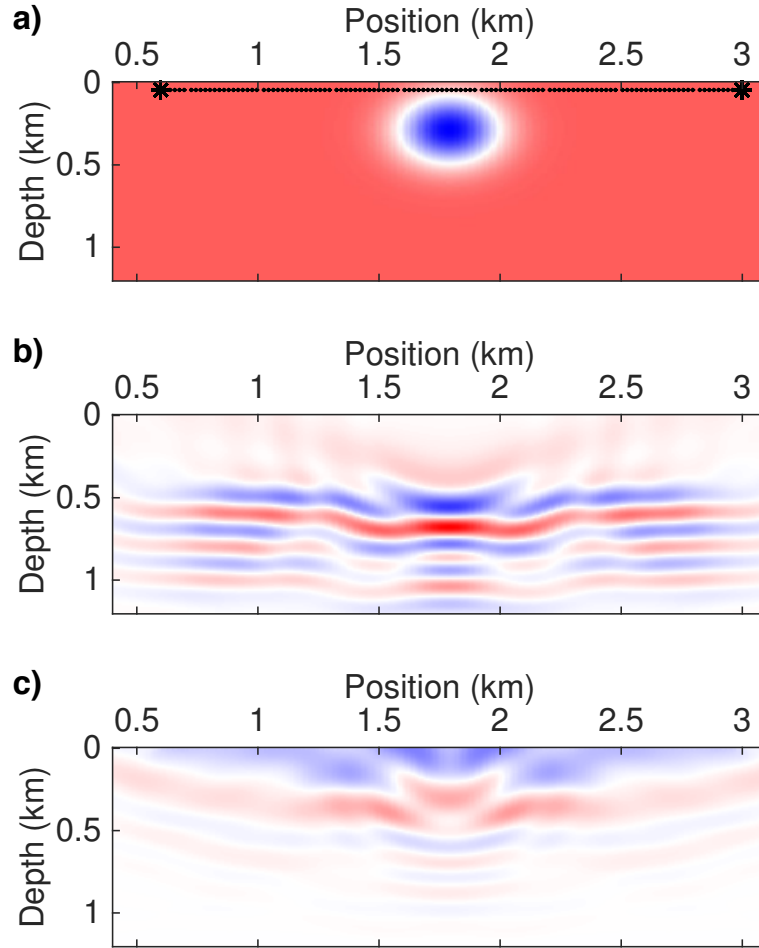


Figure 3.6 – (a) Exact velocity model, (b) stacked migrated section in $c_0 = 2.5$ km/s for all shots and (c) associated MVA gradient after a Gaussian smoothing over half a wavelength. From inside to outside of the Gaussian circle in panel (a), the velocity ranges from 1.6 to 2.5 km/s. Black stars and dashed line refer to shot position in panel (a). Red corresponds to positive and blue to negative value in panel (c).

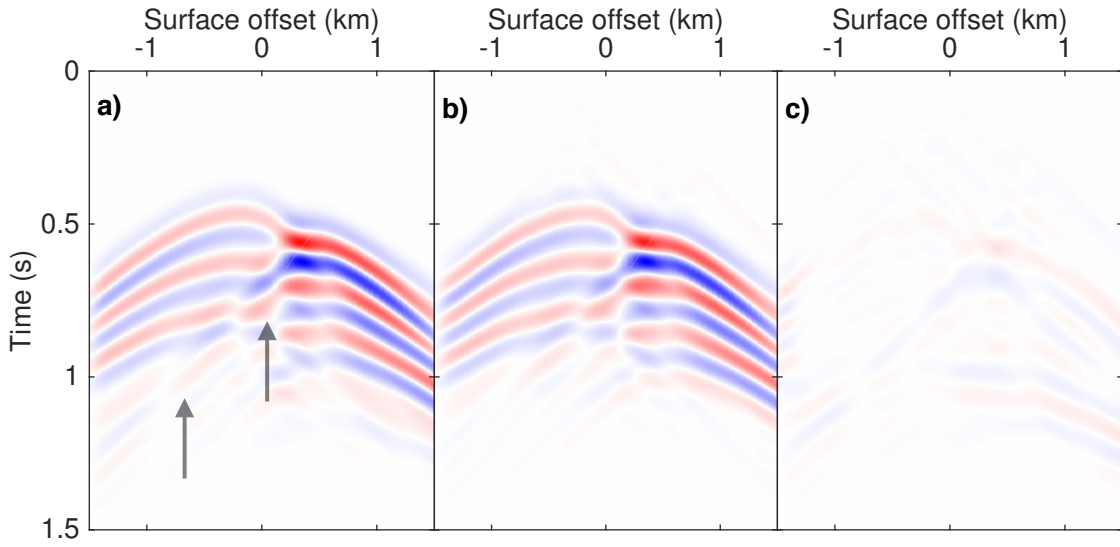


Figure 3.7 – For a single shot at $s = 1.6$ km, (a) observed data, (b) computed data after inversion and modelling in an incorrect model of constant velocity $c_0 = 2.5$ km/s, and (c) residuals at the same scale. Black arrows indicate the triplicated events.

velocity anomaly (Figure 3.9c). Because the model only contains three reflectors, we do not have enough reflections to perfectly constrain the shape of the anomaly. Here, the inverted circle has a narrower width than the true one. Nevertheless, it is overall localized around the correct position. The inverted anomaly of the lens is not as low as the exact one -0.7 versus -0.9 km/s). It is known that tomography retrieves more easily the shape of the anomaly than the value of the velocity in the case of low velocity anomaly as rays tend to avoid the zone. The stacked image section in the final model displays rather horizontal shapes (Figure 3.9d). The consistency is largely improved after updating the background model as indicated by the CIGs (see arrows in Figure 3.10). Events with steep dips, present in the CIGs for the true model, are due to the triplications (see ellipses in Figure 3.10a) as discussed now.

To better understand the impact of multipathing, we consider a less complicated case by reducing the velocity anomaly inside the Gaussian lens (Figure 3.11a). The shot gather (Figure 3.11b) and the stacked image section in the incorrect initial model (Figure 3.12a) show much weaker multipathing effects than in the previous case. The first gradient (3.12b) is still provides a correct direction. Similar to the previous test, the value of the lens anomaly in the final updated model (Figure 3.12c), is not as low as for the exact anomaly (-0.5 versus -0.6 km/s), even the inverted anomaly is approximately located at the correct position. Compared to the severe triplicated case, the stacked inverted image (Figure 3.12d) is cleaner below the anomaly zone, meaning the effect of the triplication here is not that significant. The main difference between the CIGs

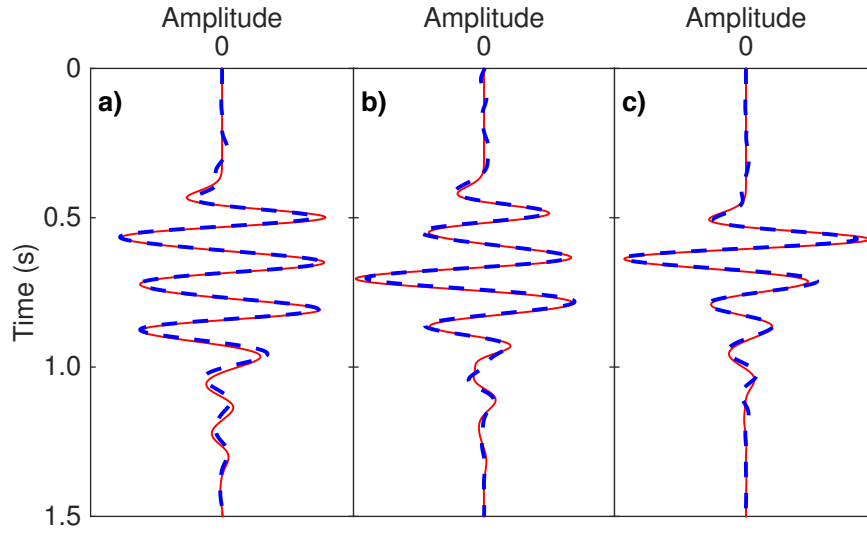


Figure 3.8 – Traces extracted from Figure 3.7, for surface-offset (a) -0.5, (b) 0, and (c) +0.5 km. The solid red line refers to the observed data and the dashed blue line to the data after inversion and modelling in an incorrect model of constant velocity $c_0 = 2.5$ km/s.

associated to the two low-anomaly tests, is the amplitude of the artifacts due to the triplications (see the eclipses on Figure 3.10 and 3.13), consistent with Xu et al. (2001). This is confirmed by shooting rays in the different models (Figures 3.14). The correct macromodel contains clear triplicated wavefields (Figures 3.14a), whereas it is not the case for the macromodel obtained with IVA (Figures 3.14b). Inversion followed by modeling in the correct background model for a single shot cannot represent the data, especially for the triplicated events (Figures 3.15 and 3.16), where inversion breaks down.

Many studies have reported that common-shot methods suffer from triplicated wavefields (Nolan and Symes, 1996; Prucha et al., 1999; Xu et al., 2001; Stolk and Symes, 2004). This remains true for inversion approach (Figure 3.10), but the coupling between inverse and velocity analysis is robust. Note that the updated velocity anomaly, derived by the outer loop optimization, is not as low as the one in the exact model. Although we were not capable of deriving the exact value of low-anomaly velocity, the results here indicate that common-shot IVA can effectively find the position of low-anomaly zone and leads to a correct velocity update direction. We conclude here that direct inversion breaks down due to the impact of triplications but the common-shot IVA approach stills provides seismic plausible results.

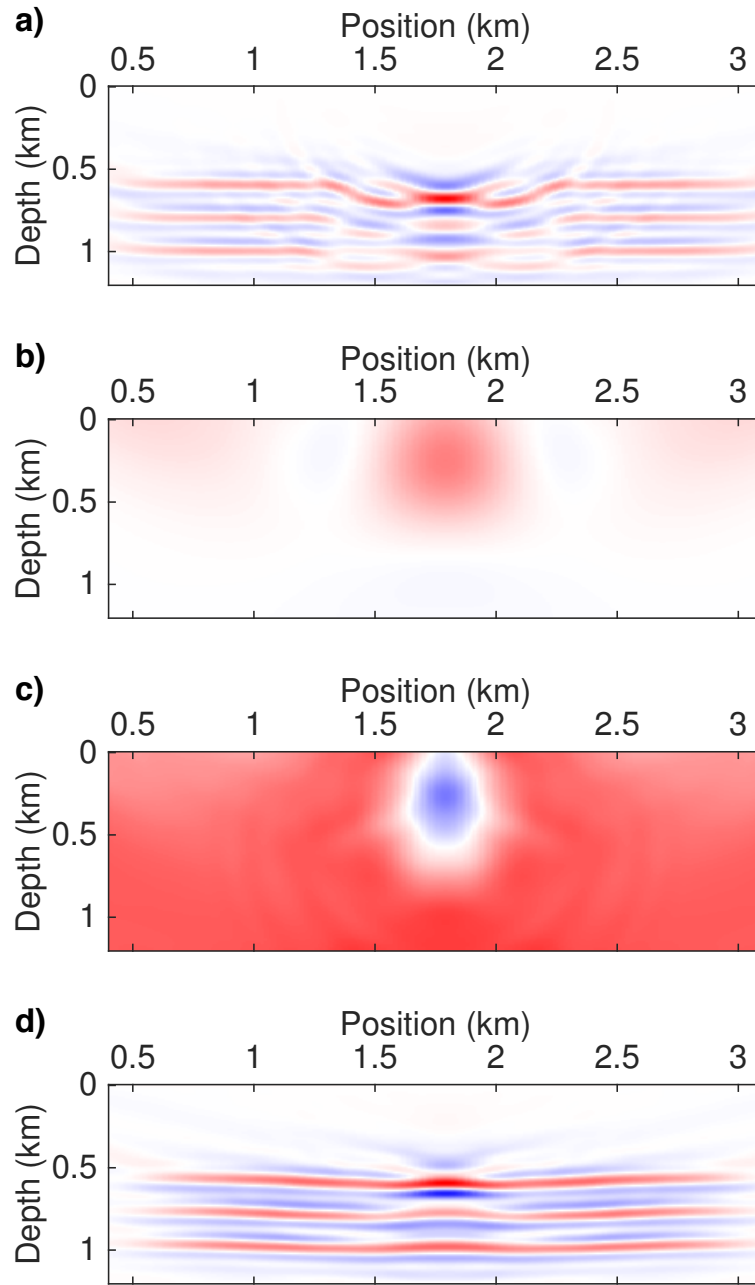


Figure 3.9 – (a) Stacked inversion section in $c_0 = 2.5$ km/s for all shots, (b) first gradient after a Gaussian smoothing over half a wavelength, (c) final inverted model after 10 non-linear iterations, and (d) stacked inversion section in the final model. Red corresponds to positive and blue to negative value in panel (c). From inside to outside of the inverted circle in panel (c), the velocity ranges from 1.86 to 2.54 km/s. Panel (c) and Figure 3.6a are displayed with the same color scale.

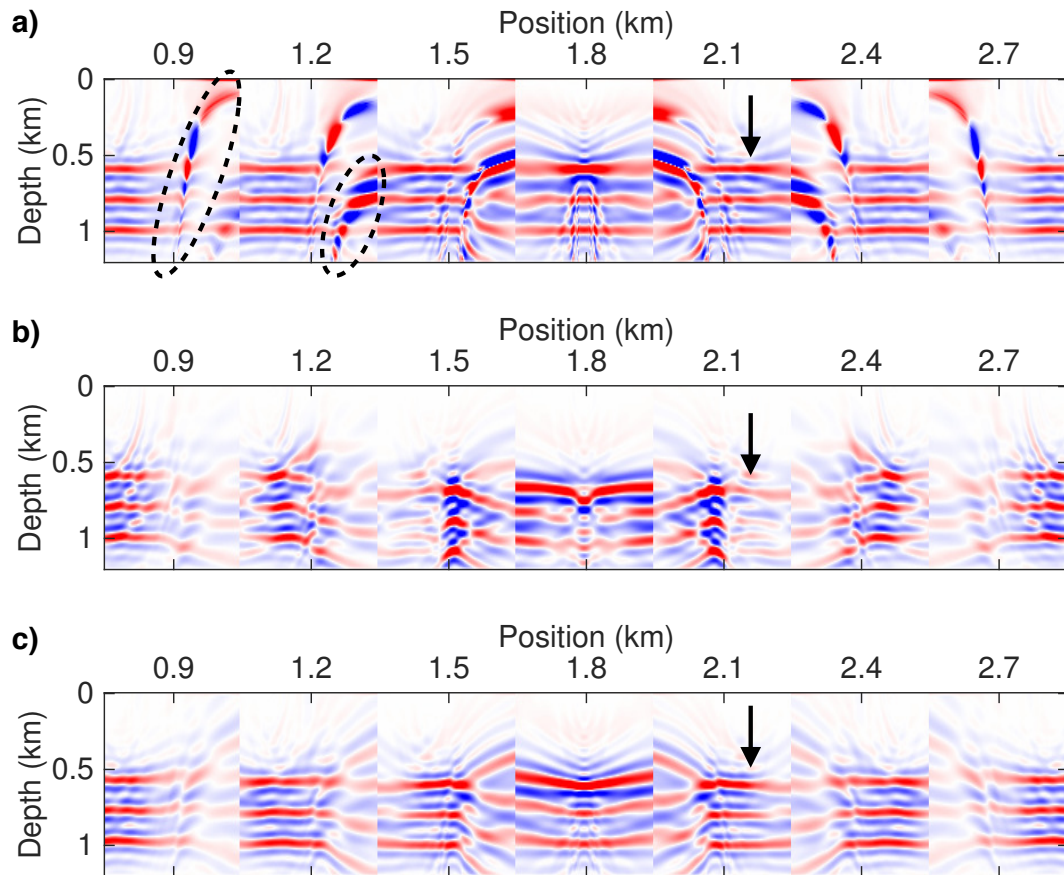


Figure 3.10 – CIGs associated to (a) the true Gaussian lens model, (b) the initial homogeneous model and (c) the updated model for position x , ranging from 0.9 to 2.7 km, every 0.3 km. Dashed black ellipses indicate the energy related to the triplication effect. Black arrows mark the same event corresponding to different macro models.

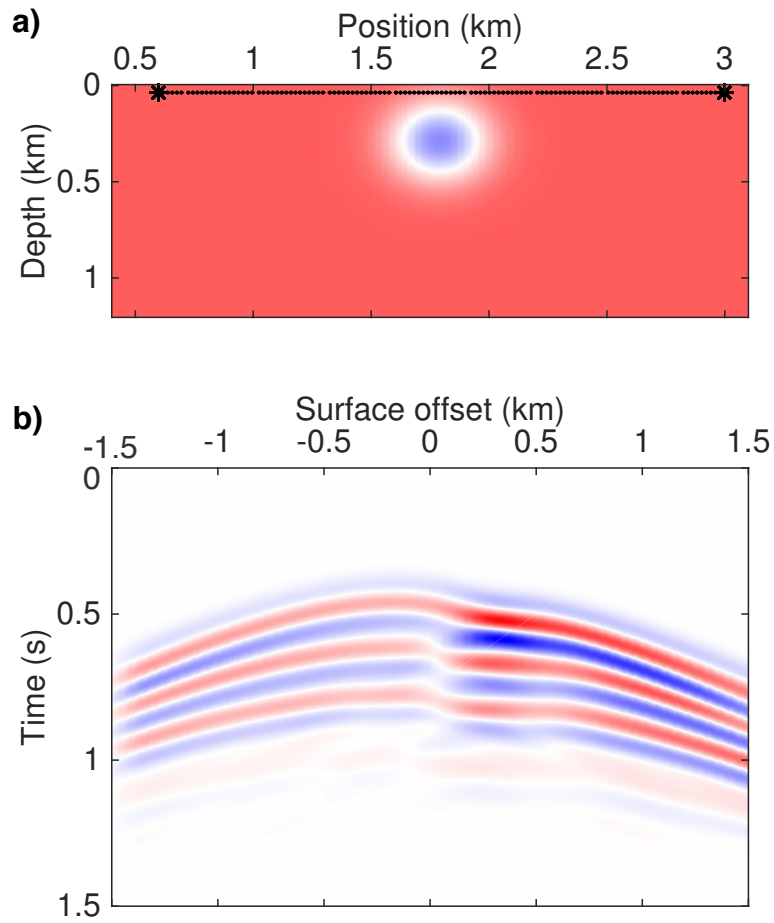


Figure 3.11 – (a) Exact velocity model and (b) observed data for a single shot at $s = 1.6$ km. From inside to outside of the Gaussian circle in panel (a), the velocity ranges from 1.9 to 2.5 km/s. Black stars and dashed line refer to shot position in panel (a).

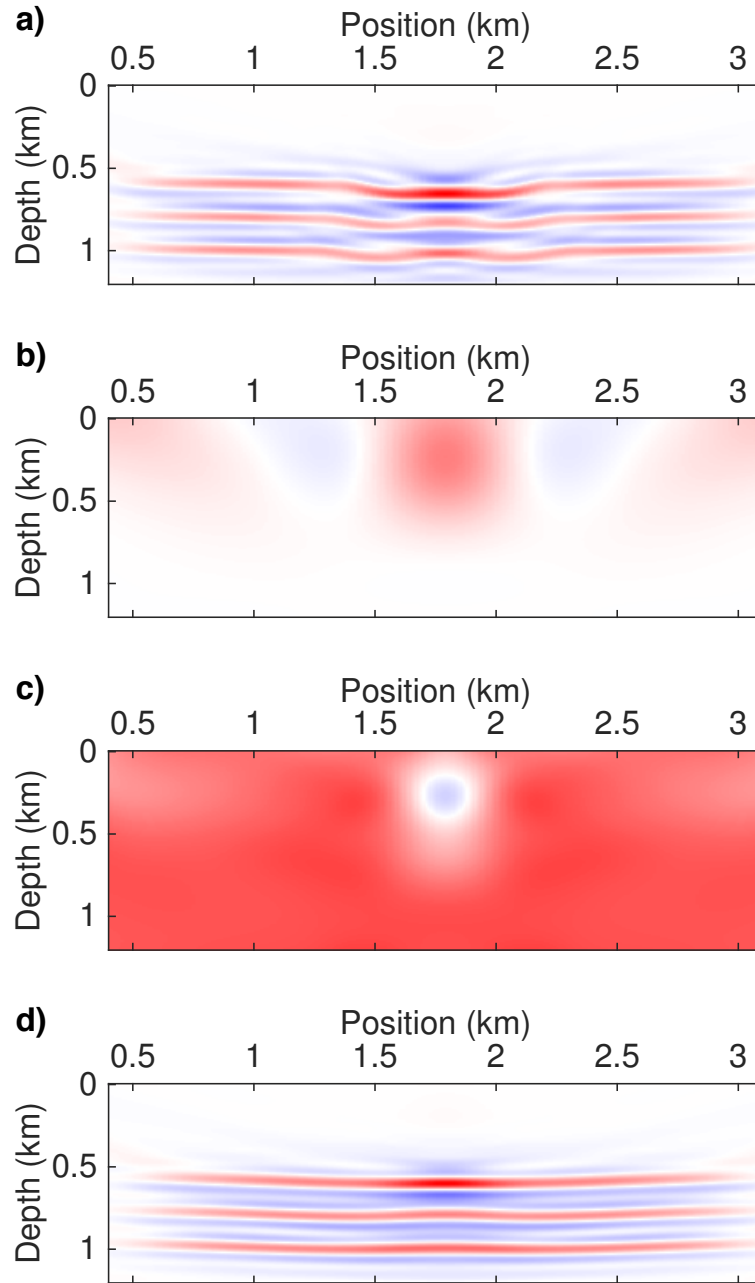


Figure 3.12 – Same as for Figure 3.9, but for the case displayed in Figure 3.11. From inside to outside of the inverted circle in panel (c), the velocity ranges from 2.05 to 2.53 km/s.

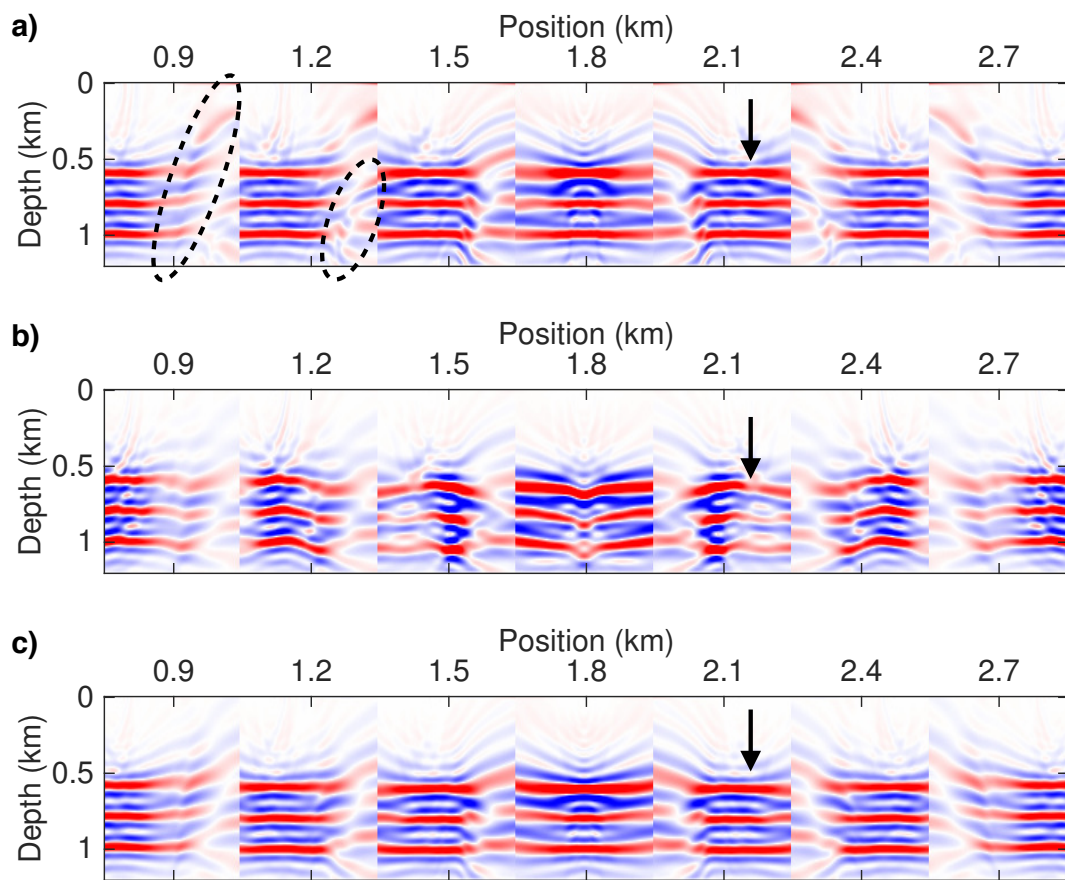


Figure 3.13 – Same as for Figure 3.10, but for the case displayed in Figure 3.11.

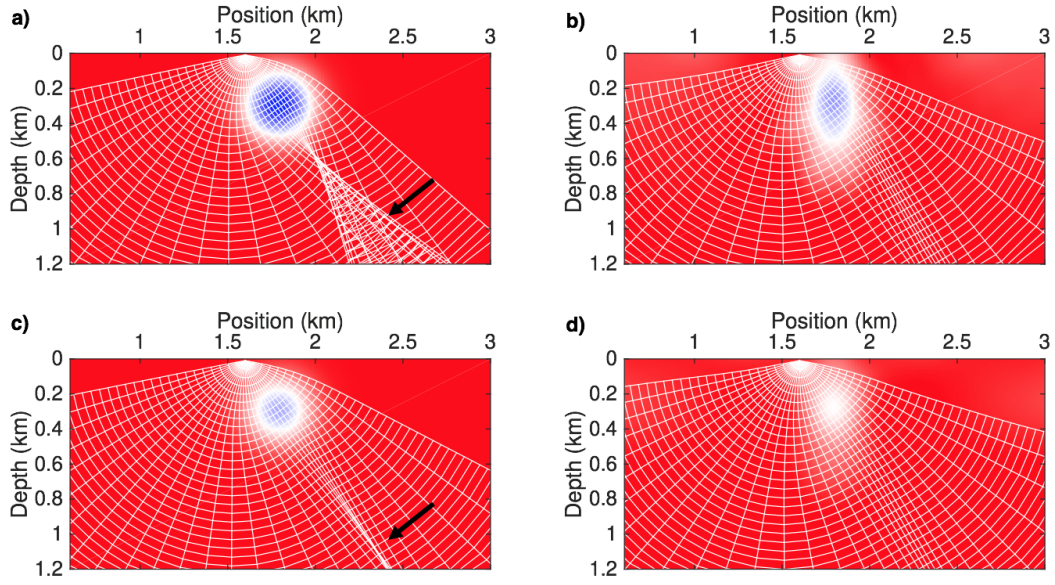


Figure 3.14 – Rays and wavefronts calculated from a source at $s_x = 1.6$ km. (a–d) correspond to Figures 3.6a, 3.9c, 3.11a and 3.12c, respectively. Black arrows mark the triplications. The macromodels obtained with IVA contain no triplications (b,d). All images are displayed with the same scale.

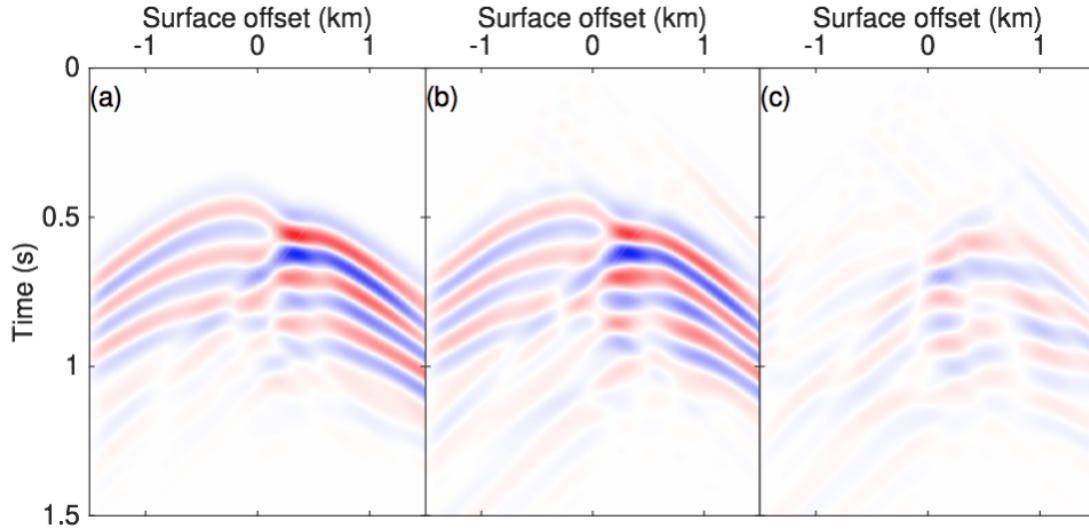


Figure 3.15 – Same as for Figure 3.7, but the computed data is obtained in the correct background model.

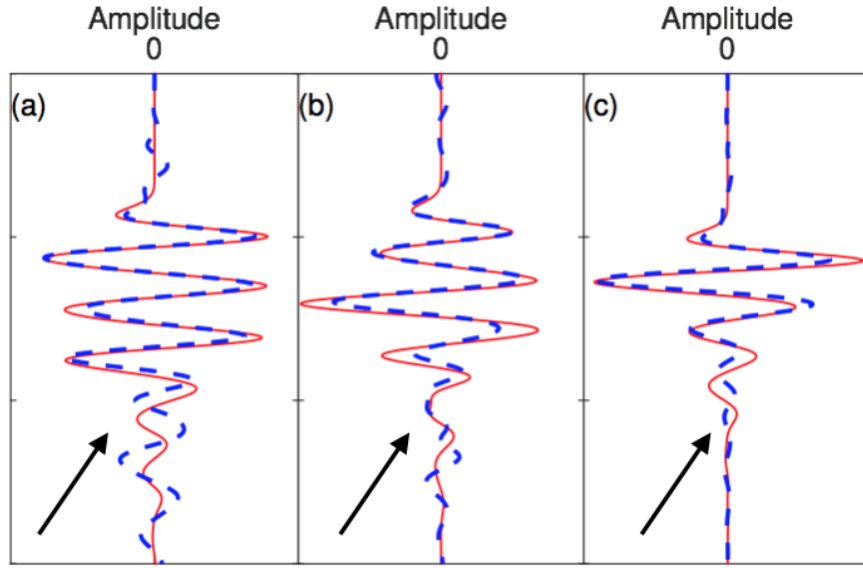


Figure 3.16 – Same as for Figure 3.7, but the traces are extracted from Figure 3.15. Black arrows mark the major differences between exact and reconstructed data.

3.4.3 Marmousi model

We apply the common-shot IVA scheme on the central part of the Marmousi model (Versteeg, 1994). For the background model, we apply a Gaussian smoothing, with a characteristic length of 60 m, to the original model. Shots are located at depth $z = 0.04$ km. We use the same acquisition geometry (split spread) as for the single interface case, except with a larger maximum shot-receiver offset of 3.0 km for the Marmousi model. Before the iterative optimization, we check the first gradient of data-domain IVA. The initial incorrect velocity models are a homogeneous model at 1.5 km/s (velocity too low) and a constant gradient model (velocity too high) starting from 2.0 km/s at the surface and 4.5 km/s at 2.0 km depth. The gradients are oscillating (Figures 3.17a and 3.17b). After the Gaussian smoothing over half a wavelength, inversion results (Figures 3.17g and 3.17h) only exhibit positive (negative) values for too high (too low) velocity model, indicating that the long wavelengths of the gradient are meaningful even for such model that consists of many discontinuous interfaces.

We start the nested IVA optimization from a homogeneous initial model at 1.5 km/s. Optimization is carried out with a hybrid conjugate gradient algorithm (Touati-Ahmed and Storey, 1990). As a preconditioner, the gradient is multiplied by the squared depth value z^2 . 100 iterations are performed (Figure 3.18) to retrieve more details even if most of the structure is constructed after 25 iterations (not shown here). The value characterizing the Gaussian smoothing on the gradient, is reduced by half every 10 iterations

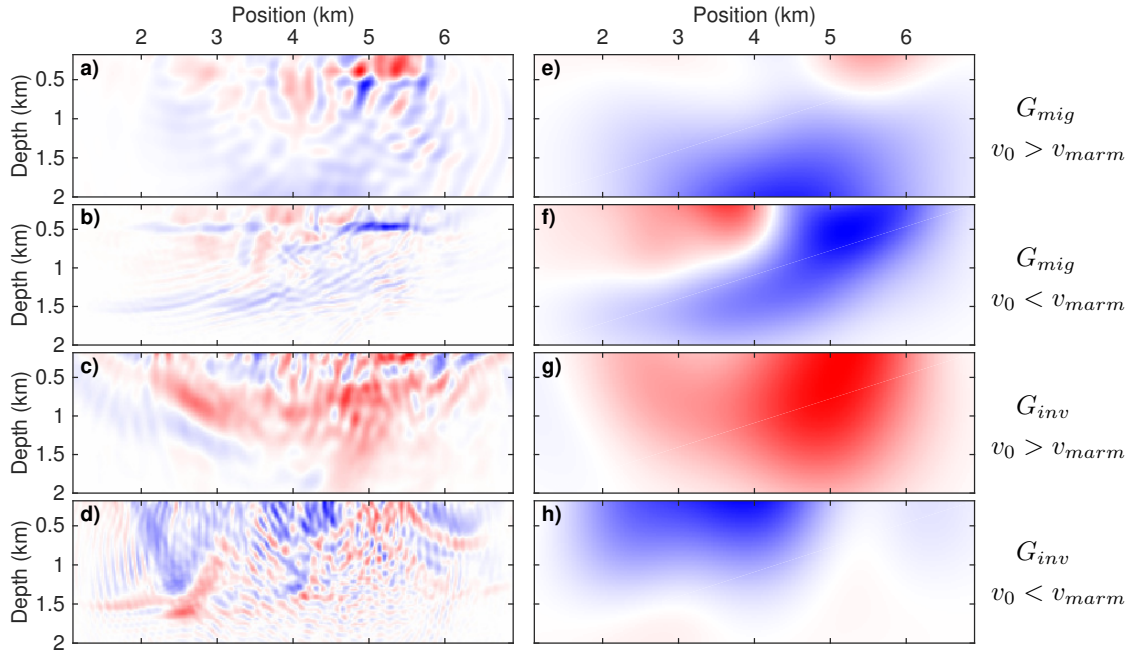


Figure 3.17 – Gradients obtained with classical MVA (a, b, e and f) and with IVA for $\alpha = 1$ (c, d, g and h) for the Marmousi model. The initial background velocity is a constant gradient model (from shallow to deep part: 2.0-5.0 km/s) with velocity higher than the true velocity (a, c, e and g). The initial background velocity is a 1.5 km/s homogeneous model for (b, d, f and h). Images on the right are obtained from images on the left after a Gaussian smoothing over half a wavelength distance.

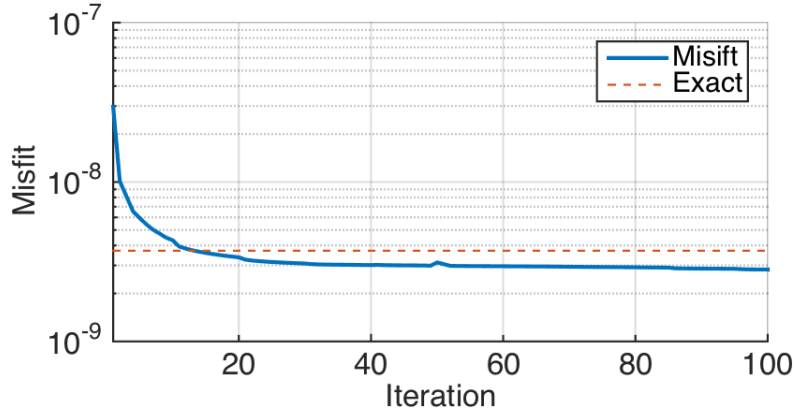


Figure 3.18 – Convergence curve for the IVA optimization performed on the Marmousi model. The vertical axis is in logarithmic scale. Blue line represents the misfit value of objective function in the iteratively updated model and red dashed line represents the cost value for the exact velocity model.

and eventually we do not smooth the gradient at all after 50 iterations, explaining why we can have small oscillations in the misfit curve (e.g. for the 51st iteration). The objective function value goes slightly beyond the value associated to the correct velocity, meaning that the objective function is not minimum for the true velocity model. This effects may result from kinematic ambiguity in the images (Nolan and Symes, 1996; Prucha et al., 1999; Stolk and Symes, 2004): in the presence of multipathing, the CIGs are not perfectly consistent even for a correct velocity model. Residual artifacts can also be responsible for that effect, as observed in Hou and Symes (2015). We compare the true background model to the final updated model after 100 iterations containing of the sum of the macromodel c_0 and the model perturbation δc (Figure 3.19a and 3.19b). The macro structure of the Marmousi model is well reconstructed, especially in the well illuminated regions (Figures 3.19d and 3.19e).

To evaluate the reliability, we compute the reflectivity models according to equation 3.9 in these different macro-models (Figure 3.20). The final retrieved background model significantly improves the quality of the stacked image (Figure 3.20c), approximating the result obtained for the true model (Figure 3.20a). Careful examination on the vertical profiles indicates that the result of IVA is reasonable: the phases and amplitudes of reflectors are similar for the images obtained in the final and in the true background velocity models, especially in the central areas (Figure 3.21). The profiles close to edges show small distortions since the model is incompletely illuminated for these parts. CIGs are extracted from the image volume to evaluate the quality of the final result, producing coherent CIGs and demonstrating the similar performance as the true model (Figure 3.22). Note that a mask is applied in the first 0.2 km to prevent from updating the shallow part. The limit of the mask (dashed line) is illustrated on Figure 3.22.

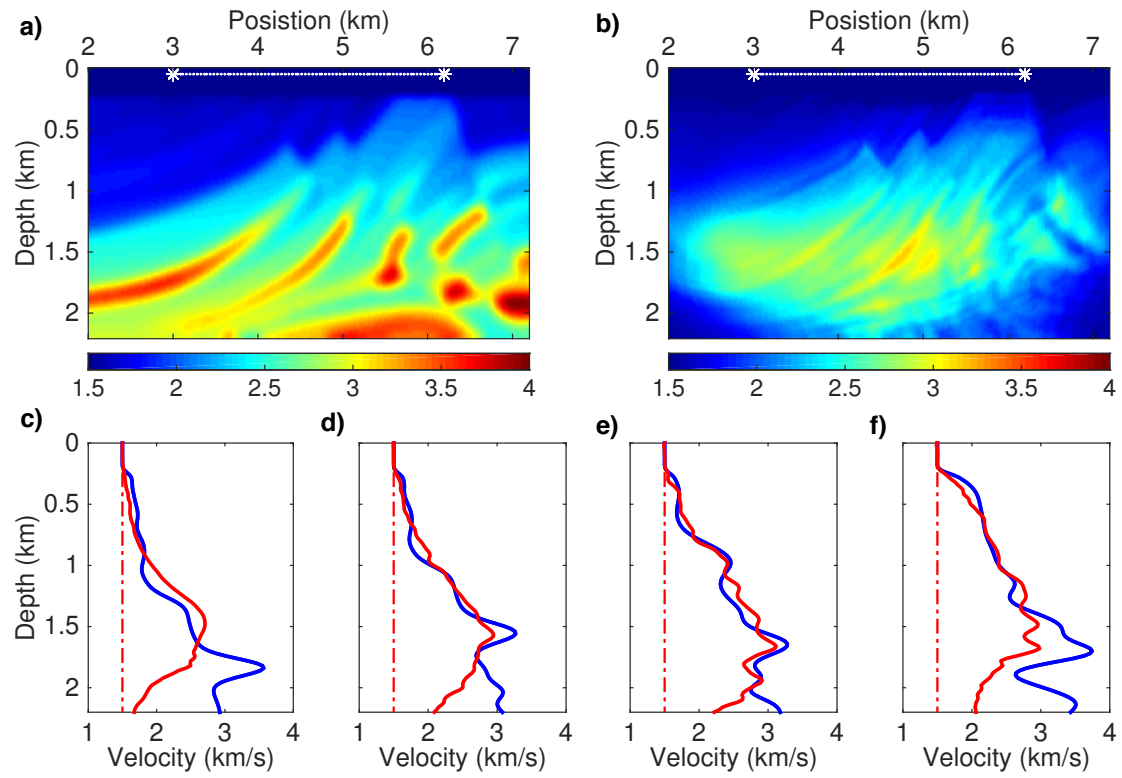


Figure 3.19 – True Marmousi macromodel (a) and total updated model ($c_0 + \delta c$) after 100 IVA iterations (b). White stars and dashed line indicate the shot position extension. (c–f) are vertical velocity profiles at positions 2.6, 3.6, 4.6 and 5.6 km, respectively. Solid blue, dashed-dot red and solid red lines correspond to true, initial and updated macromodels, respectively.

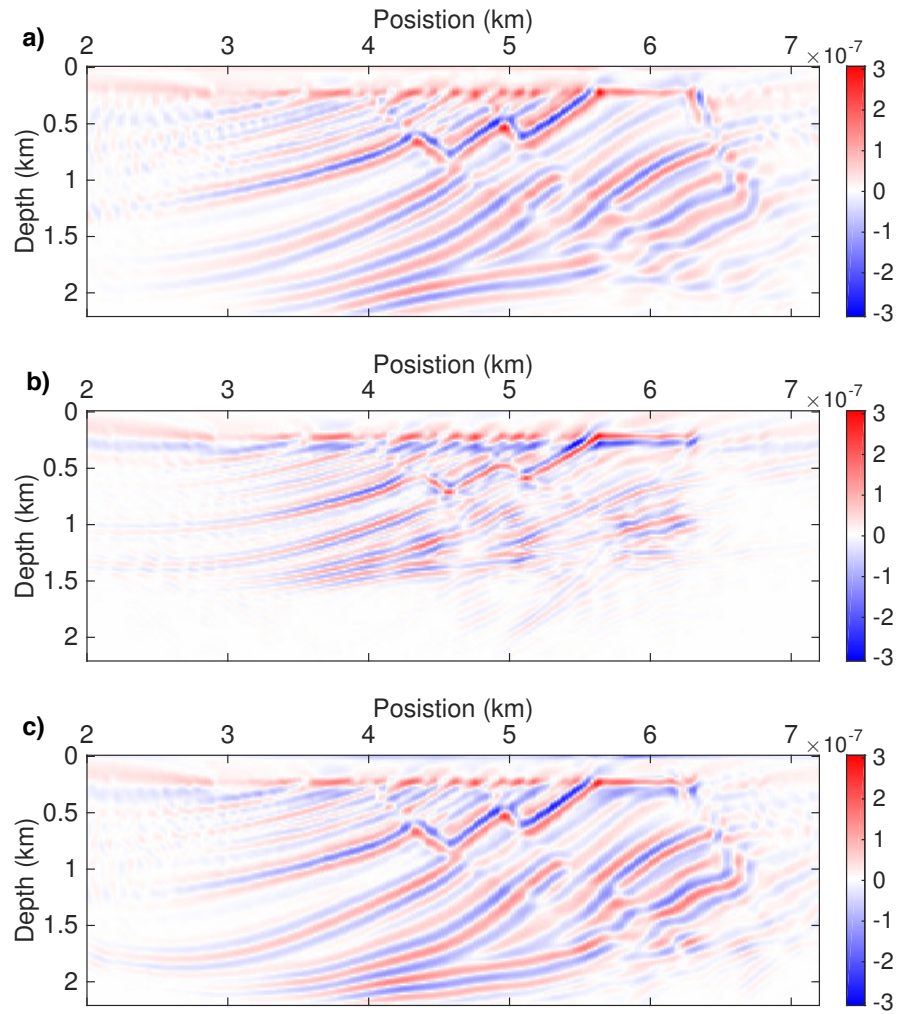


Figure 3.20 – Stacked inverted images associated to (a) true, (b) initial and (c) updated (Figure 3.19) macromodels, respectively.

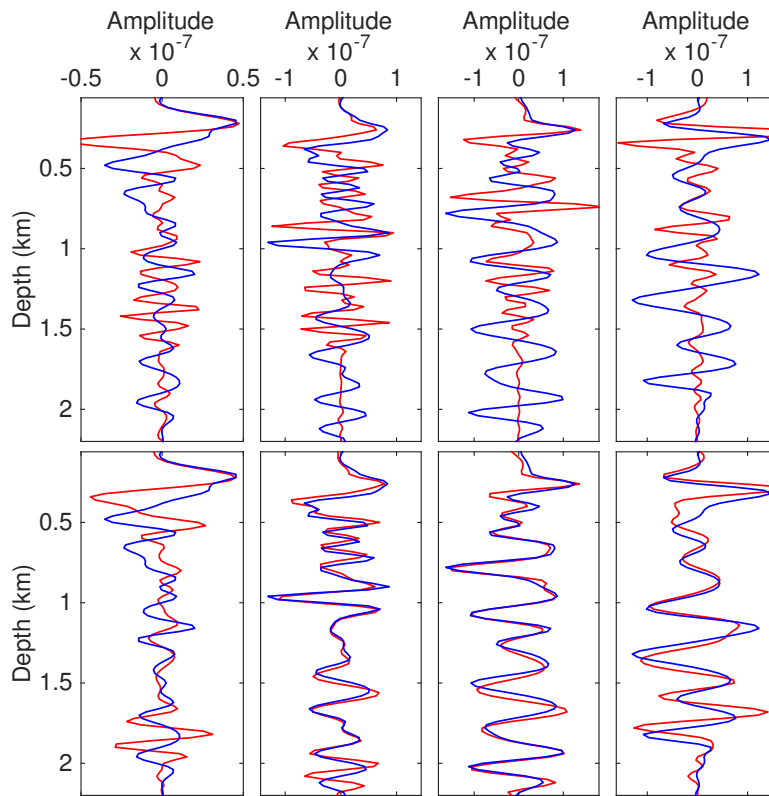


Figure 3.21 – Profiles of stacked inverted images associated with true, initial and updated (Figure 3.19) macromodels. From left to right, columns are related to positions 3.0, 4.0, 5.0 and 6.0 km, respectively. Blue lines refer to image profiles for true model. The red lines represent profiles for initial model in the top rows and for updated model in the bottom rows.

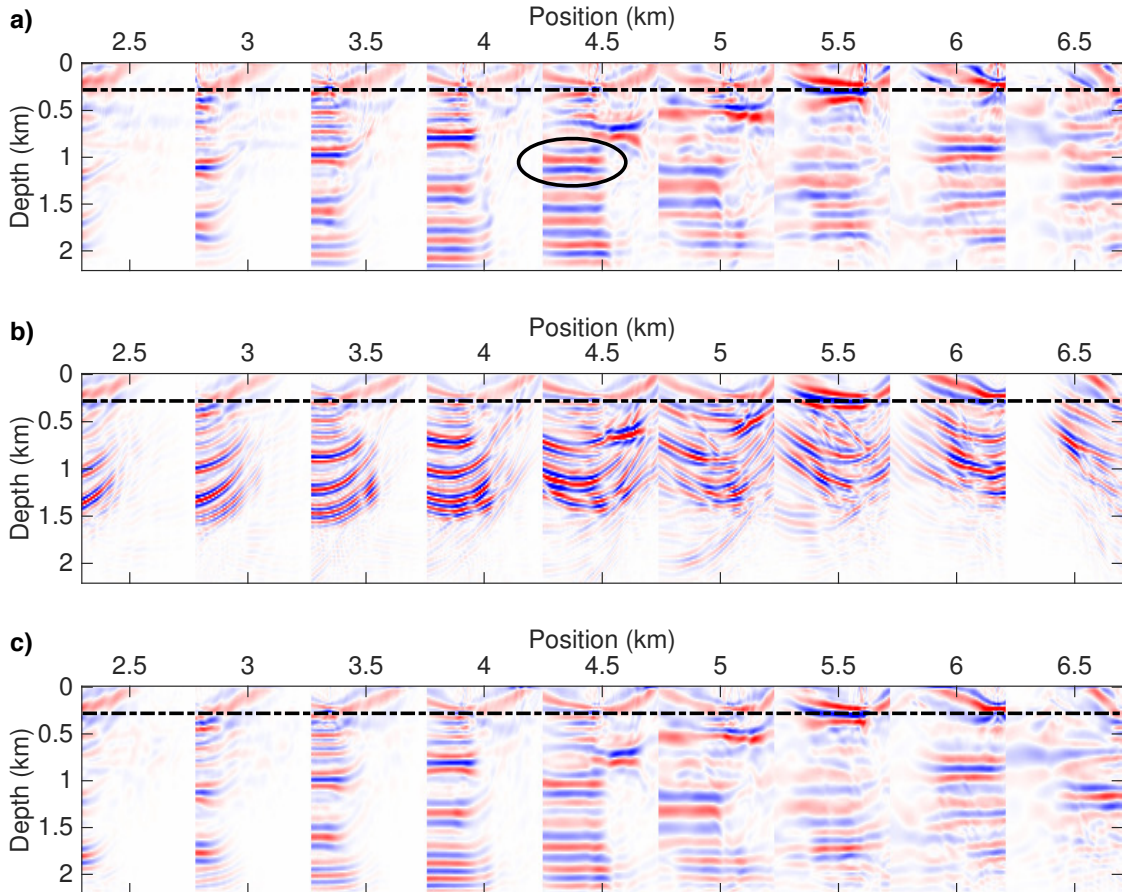


Figure 3.22 – CIGs associated to (a) the true Marmousi model, (b) the initial homogeneous model and (c) the updated (Figure 3.19) macromodel for position x , ranging from 2.5 to 6.5 km, every 0.5 km. We apply a taper on the image above the depth indicated by dashed black line, to exclude the associated contribution to the objective function. Black ellipse marks a local coherent event for image position $(x, z) = (4.5, 1.0)$ km.

Similar to Figure 3.7, we are able to produce modeled data that perfectly match the observed data, through the inversion followed by the Born modeling in an incorrect homogeneous background model (Figure 3.23). After a successful IVA optimization, the events in CIGs are flattened such that the images can be stacked to build a unique reflectivity model (i.e. the stacked image section) with the improved signal-to-noise ratio. Since the born modeling is a linear system, we expect such stacked image section can produce the reconstructed data that fully match the observation in terms of amplitudes and phases, resembling the performance of the prestack inverted images (Figure 3.23d). However, the amplitudes of the data modeled from the stacked image are not completely recovered, whereas the phase information is accurately preserved (Figure 3.24). The reason is that the inverted reflectivity models corresponding to different shot positions are not exactly the same even in the correct background model, in particular regarding their lateral extensions. In this example, we divide the stacked image by the number of shots (81) as a rough normalization, but the proper factor must be determined by the number of shots that effectively illuminate the image position. Due to the complexity of geological structure, an image position cannot be fully illuminated by all shots. For example, the image position $(x, z) = (4.5, 1.0)$ km is only illuminated by half of the shots in the true background model case (see the ellipse in Figure 3.22), such that the normalization factor is supposed to be around 40 instead of 81. Consequently, the amplitude of reconstructed data, particularly corresponding to the discussed event (see the arrow in Figure 3.24e), is roughly half of the observed data. The amplitudes of modeled data is always lower than the observed data (Figure 3.24a), not only for the incorrect background model (Figure 3.24c and 3.24f), but also for the true background model (Figure 3.24b and 3.24e) and the final updated model (Figure 3.24d and 3.24g). A subsequent FWI can thus refine the dynamic information.

Finally, the model from Figure 3.19b is used as the initial model for FWI to further improve the dynamic details. FWI suffers from cycle skipping effects, unless low-frequency data are recorded or the initial model is close enough to the correct one (Virieux and Operto, 2009). In this test, shots are located every 0.08 km ranging from 1 to 6.2 km, and receivers everywhere at the surface ($z = 0.04$ km). Frequencies below 2 Hz are not present in the observed data. Consequently, the accuracy of the initial model is the essential element to avoid cycle skipping effects. The structure of Marmousi model is well reconstructed after 70 iterations (Figure 3.25). Even the quality of the ill-illuminated edges are improved. We hardly observe cycle skipping effects in the area of which IVA supplies a reasonable macro model. Note that we compare the FWI results with the true complete model c rather than background model c_0 in this case.

We conclude from these different numerical tests that the inversion formulas developed for the 2D acoustic constant density case (equation 3.9) is indeed inverse instead of adjoint: the data modeled from inverted images nicely fit the observed data, even in an incorrect velocity model. More interestingly, inversion highly attenuates the migration smiles visible in CIGs. The coupling between inversion and velocity analysis produces a

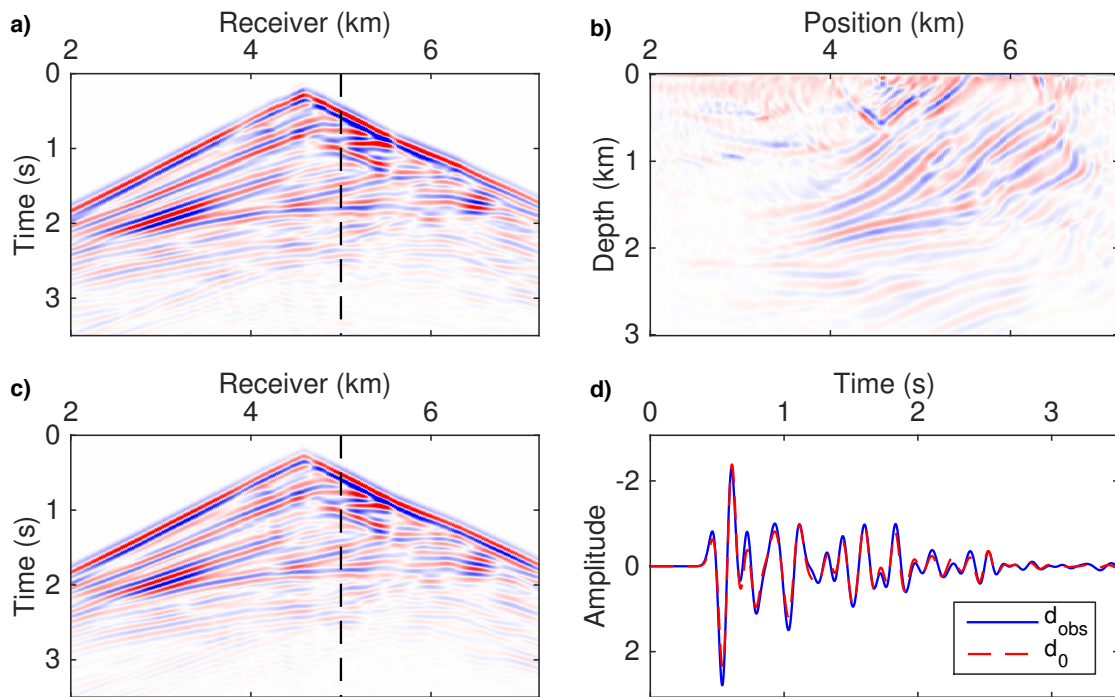


Figure 3.23 – Single-shot inversion test in the Marmousi model: (a) observed data, (b) inverted image, (c) reconstructed data and (d) single trace data comparison for trace at $x = 5$ km indicated by the black dashed line, respectively. The blue and red lines correspond to the observed and reconstructed data, respectively.

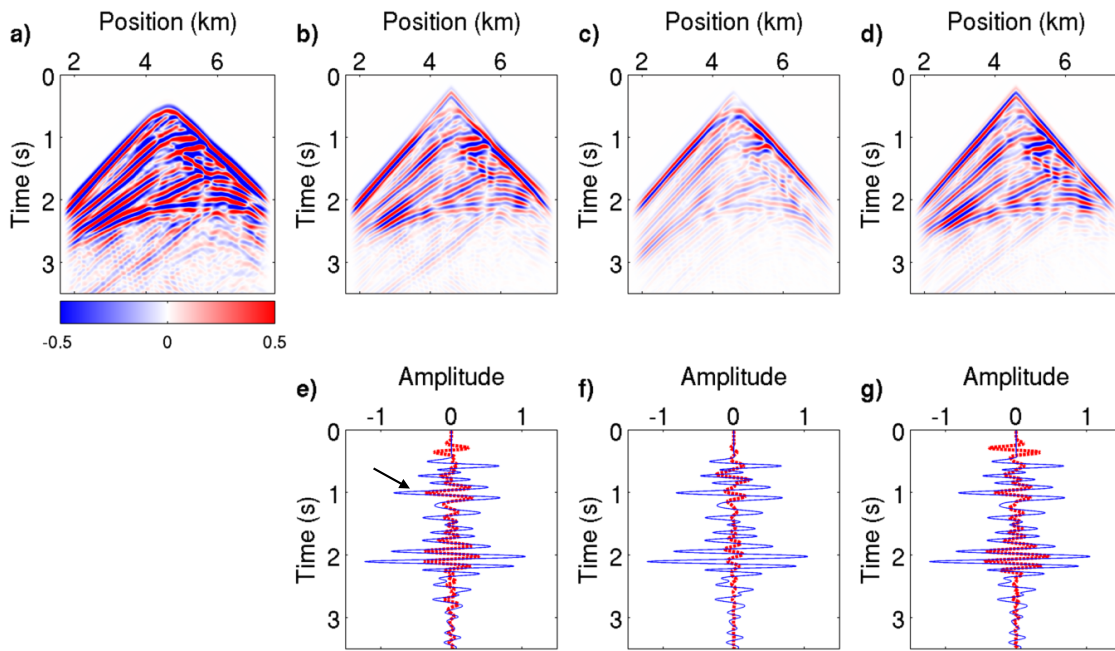


Figure 3.24 – Shot gathers of (a) observed data and (b–d) modeled data for the shot at 4.6 km. Modeled data are computed for (b) true, (c) the initial and (d) the updated background models, with the associated stacked inverted images after summation over all shots. Curves in (e–g) represent the comparison between observed and modeled data for a trace at 4.6 km. The blue curves correspond to the observed data and the red dashed curves to the modeled data. The black arrow in panel (e) marks the event corresponding to the one in the ellipse in Figure 3.22.

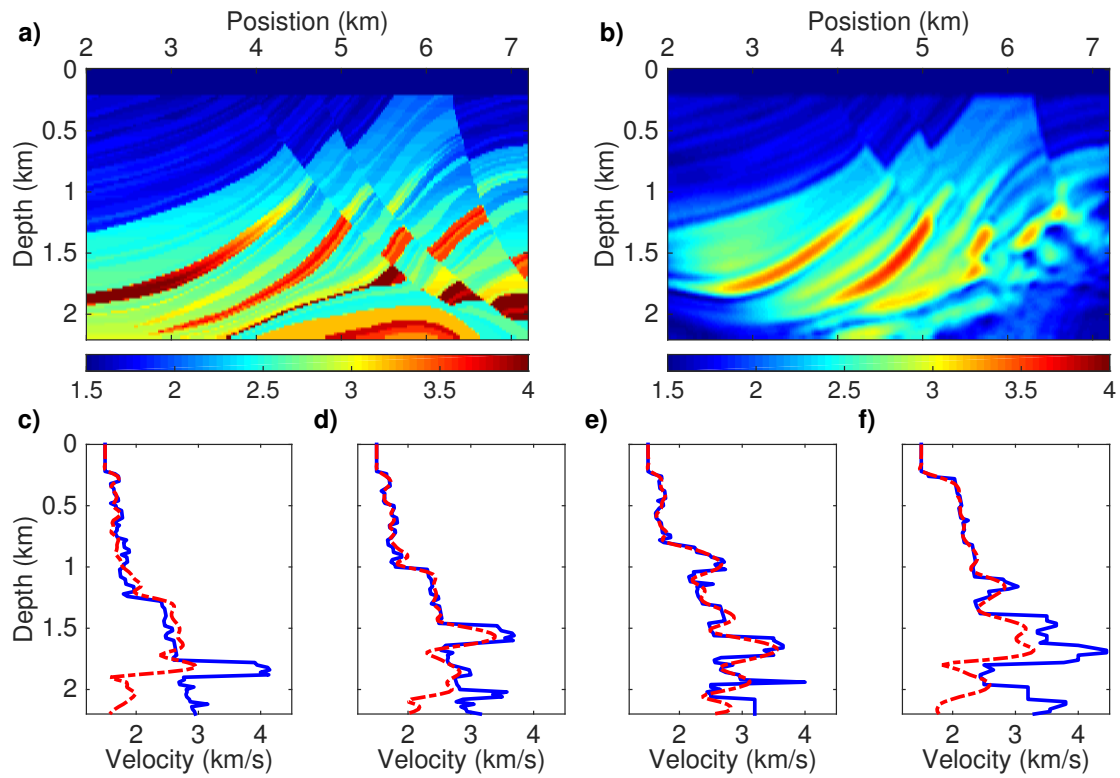


Figure 3.25 – True Marmousi full model (a) and updated model (b) after 70 FWI iterations starting from the final IVA result (Figure 3.19b). (c–f) are vertical velocity profiles at positions 2.6, 3.6, 4.6 and 5.6 km, respectively. Solid blue and dashed-dot red lines represent true and updated models, respectively.

gradient that does not suffer from the imprints of reflectivity. We propose to incorporate c_0^α , with $\alpha = 1$, in the definition of the DSO objective function for outer optimization to evaluate the quality of the consistency in CIGs. Afterwards, the imprint artifacts are reduced in the gradient smoothed over a characteristic length proportional to the half wavelength of the data. We discuss the impact of multipathing in the next section. IVA is capable of dealing with complex structure, at least for the Marmousi model.

3.5 Discussions

We derived the inversion formula by assuming that the main contribution in equation 3.19 is provided by terms around the diagonal $x = x_0$. Thus, a linearized version (equation 3.20) can be obtained by the Taylor expansion. The final formula 3.9 is slightly different from the one provided by [Qin et al. \(2015\)](#) (equation 3.28). The difference only concerns the energy at source locations. We apply vertical receiver derivative on the Green function, whereas [Duprat and Baina \(2016\)](#) accomplish that by applying an modified boundary condition. Note that the integration over receivers is only performed along the horizontal component of the receiver position and not the vertical one, preventing from integrating by part. [Rocha et al. \(2016\)](#) propose to compensate for illumination by introducing the energy norm imaging condition. The energy norm weight is the Laplacian filter. We give a justification for the introduction of this weight that is explicitly defined in the numerator of our inversion formula. The weight is a high-pass filter, attenuating the most dipping events in CIGs, especially vertical events associated to low wavenumbers. [Alkhalifah \(2014\)](#) uses a modified low-pass filter to provide only long wavelength update for RWI. Nevertheless, such weight only removes kinematic artifacts but does not recover correct amplitudes. We have additionally introduced wavefield deconvolution to preserve amplitude information.

In the case of the low-velocity anomaly tests, the Gaussian lens anomaly model leads to clear triplicated wavefields for which the conventional common-shot scheme breaks down, whereas our method is robust. The multipathing effect mainly leads to imaging artifacts corresponding to the reflector location ambiguity: a single event recorded in the observation at one surface receiver can be attributed to reflectors at two or more image points ([Stolk and Symes, 2004](#)). The imaging ambiguity contributes to non-flat events in CIGs, thus altering conventional velocity analysis ([Nolan and Symes, 1996](#)). This is still the case for the asymptotic inversion formula. However, numerical tests (the second synthetic example) show that IVA is relatively robust for different reasons. The non-triplicated events contribute more to the velocity update more than those artifacts do, and CIGs for the final inverted macromodel does not contain very strong artifacts. Even if the derivation of the inverse formula assumes the absence of multipathing, the coupling with velocity analysis is relatively robust as for the extended subsurface-offset approaches ([Stolk et al., 2009](#); [ten Kroode, 2012](#); [Hou and Symes,](#)

2015).

For the nested optimization system, we solve the inner loop by the direct inversion, and the outer loop by the nonlinear iterative scheme. Beyond the scope of this article, there is another strategy: instead of direct inversion, one can retrieve the inverted image by iteratively minimizing the inner loop (equation 3.2). Such an approach is computational expensive since it generally requires several iterations to converge (Cocher et al., 2017a). However, the scheme can incorporate multiples if they are modeled by the forward engine, whereas the asymptotic inversion is limited to primary reflection only. The direct inversion could be used as a preconditioner for the iterative scheme (Lambaré et al., 1992; Qin et al., 2015; Hou and Symes, 2016b). Considering the limited acquisition, iterative migration has advantages over direct inversion that assumes an infinite acquisition and a complete illumination of the subsurface. To fully solve the outer loop, it is essential to estimate the Hessian associated to the objective function (equation 3.13). For industry-scale problems, iterative schemes for inverting the Hessian are always prohibitively expensive though. In the context of horizontal contraction, Shen and Symes (2015) propose an economical estimation of the sum of the Hessian matrix. More studies are needed to derive an equivalent for common-shot IVA approach.

We now discuss the computational cost and the memory requirement by comparing classic FWI (Lailly et al., 1983; Tarantola, 1984a), subsurface-offset IVA proposed by Chauris and Cocher (2017) and common-shot IVA detailed here. For all methods, the main computational step consists of evaluating the objective function and of calculating its gradient with respect to the velocity model. In practice, the former is included in the derivation of the latter, as the modeled data or inverted images are also required for computing the macromodel gradient. We thus analyze only the procedure of macromodel gradient computation. We evaluate the computational cost as a function of q_{mod} and q_{cross} . For a given source term, q_{mod} and q_{cross} indicate the computational costs required to solve the wave equation once and to crosscorrelate two wavefields once, respectively. q_{mod} is proportional to the number of grid points n_t , n_z and n_x . Compared to q_{mod} , q_{cross} is negligible in 2D, but its cost can significantly increase in 3D.

First, FWI needs to solve the forward and backward propagation problems and then crosscorrelate the two wavefields to derive the gradient, such that only two wavefields are required. The first column in Table 3.3 gives a summary of required memory and CPU cost for FWI. Second, subsurface IVA requires computing four vectors of wavefields: the forward and backward propagated wavefields and the associate two adjoint state variables. Besides, the method also needs storing all the inverted images, which depend on the subsurface-offset, to compute the contribution of a single source to the macromodel gradient. This scheme needs the computation of wavefields already calculated in the direct inversion (see details in Cocher et al., 2017b), because it is necessary to integrate over all shots for deriving the inverted image corresponding to a given subsurface-offset value h . Thus, the wavefields are 4D arrays depended on (n_t, n_z, n_x, n_s) , too large to be stored in the memory: they need to be recomputed

Table 3.3 – Comparison of different approaches in terms of memory and computational requirements for calculating the gradient of the associated objective function. (n_z, n_x) denote the size of the 2D model and n_t the sample size of the time for the data. n_h is the number of the subsurface-offset sampling and n_s the number of the shot. For a given source term, q_{mod} and q_{cross} are the costs required to solve the wave equation once and to crosscorrelate two wavefields once, respectively.

	FWI	IVA (subsurface-offset)	IVA (common-shot)
Memory	$2(n_z, n_x, n_t)$ -vectors	$4(n_z, n_x, n_t)$ -vectors, (n_z, n_x, n_h) -vector	$4(n_z, n_x, n_t)$ -vectors, $(n_z, n_x, 3)$ -vector
CPU	$2n_s q_{mod} +$	$6n_s q_{mod} +$	$4n_s q_{mod} +$
cost	$n_s q_{cross}$	$(n_h + 2)n_s q_{cross}$	$5n_s q_{cross}$

twice for each shot during the gradient derivation. Briefly, the subsurface-offset IVA requires $2n_s q_{mod}$ calculating wavefields for the direct inversion, $2n_s q_{mod}$ for the recomputed wavefields and $2n_s q_{mod}$ for the adjoint state variables. In terms of crosscorrelation, $n_h n_s q_{cross}$ is required for inverting the images and $2n_s q_{cross}$ for deriving the gradient (see the inversion and gradient formulations in [Chauris and Cocher, 2017](#)). The second column in Table 3.3 provides a summary for subsurface-offset IVA. Finally, common-shot IVA also requires four wavefields similar to subsurface-offset IVA. However, we need only three images ($\xi(\mathbf{x}, s - \delta s)$, $\xi(\mathbf{x}, \delta s)$ and $\xi(\mathbf{x}, s + 1)$), for which the memory requirement is less than in the subsurface-offset case, to compute λ_ξ (equation 3.34) for a given source s and the associated contribution to gradient. In terms of CPU cost, common-shot IVA avoids the recomputation of wavefields in subsurface-offset, as a result of the more natural shot-by-shot implementation. Consequently, $2n_s q_{mod}$ modeling are saved. For crosscorrelation, $3n_s q_{cross}$ is required for inverting the images (equation 3.9) and $2n_s q_{cross}$ for deriving the gradient (equation 3.15). n_h is much larger than 3 such that $(n_h + 2)n_s q_{cross}$ is greater than $5n_s q_{cross}$. However, the numerical requirements for storing images and for crosscorrelating wavefields are negligible in the 2D case. Consequently, common-shot and subsurface-offset IVA approaches in practice have similar memory requirements in 2D. The third column in Table 3.3 shows a summary of required memory and CPU cost for common-shot IVA.

If we consider a 3D implementation, subsurface-offset IVA will become prohibitively expensive due to the significantly increased memory requirements (from (n_z, n_x, n_h) to $(n_z, n_y, n_x, n_{h_x}, n_{h_z})$), whereas the common-shot IVA requires less (from $(n_z, n_x, 3)$ to $(n_z, n_y, n_x, 5)$). Additionally, the cost of crosscorrelation is no more negligible for the subsurface-offset case, as two extra dimensions are introduced such that $n_h n_s q_{cross}$ becomes $n_{h_x} n_{h_z} n_{s_x} n_{s_z} q_{cross}$. The common-shot IVA requires less computational cost and memory than in the subsurface-offset case: it may be extended to 3D.

3.6 Conclusions

We have proposed to couple direct inversion to common-shot image-domain velocity analysis. The asymptotic inverse operator automatically compensates for uneven illumination and removes migration smiles. As a result, the IVA gradient has a smoother behavior and suffers less from oscillations around interface positions than in the MVA case. To get rid of the imprint for reflectivity in the macromodel update, we have multiplied the inverted reflectivity by the velocity at a specific power, determined in a two dimension case study, before evaluating the objective function. Compared to conventional common-shot methods, triplicated waves still alter the images, but the coupling with velocity analysis is robust. The non-linear optimization, performed on the Marmousi model, proves its capability to deal with complex media. By starting from the final IVA results as the initial model, the subsequent FWI successfully derives a non-cycle skipped high resolution model, at least for the area of which IVA supplies a correct background velocity update. Three aspects are essential for a successful IVA: inverse instead of adjoint, reflectivity multiplied by the velocity at a specific power in the evaluation of the coherency, and Gaussian smoothing over half a wavelength distance. The future work consists of comparing IVA approaches in the surface-offset and subsurface-offset domains.

3.7 Appendix I: Common-shot inversion scheme

The objectives of this appendix are the following: (1) demonstrate that \mathcal{B}_0^\dagger is an inverse of \mathcal{B}_0 in the asymptotic sense; (2) provide the detailed derivation from ray-based inversion formula to an expression containing only wave-equation-based operators.

We first introduce $H\xi$ as the integrand of the product of the two operators, reading

$$(\mathcal{B}_0^\dagger \circ \mathcal{B}_0)\xi(\mathbf{x}, \mathbf{s}) = \int d\mathbf{x}_0 H(\mathbf{x}, \mathbf{x}_0, \mathbf{s})\xi(\mathbf{x}_0, \mathbf{s}). \quad (3.18)$$

Such expression indicates that \mathcal{B}_0^\dagger is indeed the inverse of \mathcal{B}_0 if H is a spatial Dirac distribution $\delta(\mathbf{x} - \mathbf{x}_0)$ (Lambaré et al., 1992; Lameloise et al., 2015; Hou and Symes, 2015; Chauris and Cocher, 2017). Substituting equations 3.7 and 3.8 into equation 3.18, the operator H reads:

$$H(\mathbf{s}, \mathbf{x}, \mathbf{x}_0) = \iint d\omega d\mathbf{r} \frac{4\omega}{c_0^2(\mathbf{x})} \frac{\cos \beta_r}{c_r} \cos^2\left(\frac{\theta}{2}\right) \frac{A_r}{A_s} A_{s_0} A_{r_0} e^{-i\omega\Delta\tau}, \quad (3.19)$$

where we simplify amplitude and traveltime terms: $A_s = A(\mathbf{s}, \mathbf{x})$, $A_{s_0} = A(\mathbf{s}, \mathbf{x}_0)$, $A_r = A(\mathbf{r}, \mathbf{x})$, $A_{r_0} = A(\mathbf{r}, \mathbf{x}_0)$ and $\Delta\tau = \tau(\mathbf{s}, \mathbf{x}) + \tau(\mathbf{r}, \mathbf{x}) - \tau(\mathbf{s}, \mathbf{x}_0) - \tau(\mathbf{r}, \mathbf{x}_0)$. The half integration is denoted by $K(\omega) = 1/\sqrt{i\omega}$ in the 2D case. We are interested in the terms around the diagonal of operator H . Using the Taylor expansion, H becomes:

$$H(\mathbf{s}, \mathbf{x}, \mathbf{x}_0) \approx \iint d\omega d\mathbf{r} \frac{4\omega}{c_0^2(\mathbf{x})} \frac{\cos \beta_r}{c_r} \cos^2\left(\frac{\theta}{2}\right) A_{r_0}^2 e^{-i(\mathbf{k} \cdot (\mathbf{x} - \mathbf{x}_0))}, \quad (3.20)$$

where $\mathbf{k} \cdot (\mathbf{x} - \mathbf{x}_0) = \omega\Delta\tau$. Wavenumber \mathbf{k} is a 2D vector, whose components are k_x and k_z . The traveltime terms τ are linked to wavenumber terms \mathbf{k} through slowness terms. The slowness vector is the spatial derivative of the traveltime. The cosine and sine functions of the oriented angle terms θ_s and θ_r can also be transferred to partial derivatives of traveltimes τ_s and τ_r , and the link is given in Hou and Symes (2015). Changing variables from (\mathbf{r}, ω) to (k_x, k_z) , equation 3.20 reads:

$$H(\mathbf{s}, \mathbf{x}, \mathbf{x}_0) \approx \int d\mathbf{k} \frac{4\omega}{c_0^2(\mathbf{x})} \frac{\cos \beta_r}{c_r} \cos^2\left(\frac{\theta}{2}\right) A_{r_0}^2 \left| \frac{\partial(\mathbf{r}, \omega)}{\partial \mathbf{k}} \right| e^{-i(\mathbf{k} \cdot (\mathbf{x} - \mathbf{x}_0))}. \quad (3.21)$$

Here, the 2x2 Jacobian determinant in 2D can easily be obtained since all wavenumber terms have explicit expressions, yielding

$$\begin{aligned} \left| \frac{\partial \mathbf{k}}{\partial(\mathbf{r}, \omega)} \right| &= \left| \begin{array}{cc} \frac{\partial k_x}{\partial \mathbf{r}} & \frac{\partial k_x}{\partial \omega} \\ \frac{\partial k_z}{\partial \mathbf{r}} & \frac{\partial k_z}{\partial \omega} \end{array} \right| = \left| \begin{array}{cc} \omega \frac{1}{c_0(\mathbf{x})} \cos \theta_r \frac{\partial \theta_r}{\partial \mathbf{r}} & \frac{1}{c_0(\mathbf{x})} (\sin \theta_s + \sin \theta_r) \\ -\omega \frac{1}{c_0(\mathbf{x})} \sin \theta_r \frac{\partial \theta_r}{\partial \mathbf{r}} & \frac{1}{c_0(\mathbf{x})} (\cos \theta_s + \cos \theta_r) \end{array} \right| \\ &= \omega \frac{2}{c_0^2(\mathbf{x})} \left| \frac{\partial \theta_r}{\partial \mathbf{r}} \right| \cos^2 \frac{\theta}{2}. \end{aligned} \quad (3.22)$$

The expression for the amplitude term are given in [Zhang et al. \(2005\)](#); [ten Kroode \(2012\)](#); [Hou and Symes \(2015\)](#):

$$A_r^2 = \frac{1}{8\pi^2} \frac{c_r}{\cos \beta_r} \left| \frac{\partial \theta_r}{\partial \mathbf{r}} \right|. \quad (3.23)$$

We substitute equations 3.22 and 3.23 into equation 3.21 such that most of ray terms are canceled out, yielding

$$\begin{aligned} H(\mathbf{s}, \mathbf{x}, \mathbf{x}_0) &= \int d\mathbf{k} \frac{1}{4\pi^2} e^{-i(\mathbf{k} \cdot (\mathbf{x} - \mathbf{x}_0))} \\ &= \delta(\mathbf{x} - \mathbf{x}_0). \end{aligned} \quad (3.24)$$

To obtain equation 3.9 from equation 3.8, we introduce the following approximations

$$\nabla_{\mathbf{x}} G_0(\mathbf{s}, \mathbf{x}, \omega) \simeq \frac{i\omega}{c_0(\mathbf{x})} (\sin \theta_s, \cos \theta_s) G_0(\mathbf{s}, \mathbf{x}, \omega), \quad (3.25)$$

$$\nabla_{\mathbf{x}} G_0(\mathbf{r}, \mathbf{x}, \omega) \simeq \frac{i\omega}{c_0(\mathbf{x})} (\sin \theta_r, \cos \theta_r) G_0(\mathbf{r}, \mathbf{x}, \omega), \quad (3.26)$$

$$D_{r_z} G_0(\mathbf{r}, \mathbf{x}, \omega) \simeq i\omega \frac{\cos \beta_r}{c_r} G_0(\mathbf{r}, \mathbf{x}, \omega), \quad (3.27)$$

where c_r is the velocity at the receiver position. θ_s and θ_r are the angles at the image points \mathbf{x} associated with the source and receiver positions, respectively. We replace high-frequency terms in equation 3.8 with Green's function and then substitute the cosine and sine functions with equations 3.25–3.27. Given the fact that $\theta = \theta_s + \theta_r$, ray-based inversion formula is finally rearranged to equation 3.9, a formula containing only wave-equation-based operators.

For the purpose of accelerating the convergence of FWI, [Qin et al. \(2015\)](#); [Qin and Lambaré \(2016\)](#) have developed an equivalent inversion formula, reading

$$\begin{aligned} \mathcal{B}_{inv}^{qin}[d^{obs}](\mathbf{x}, \mathbf{s}) &= 4 \int d\omega (|S_0(\mathbf{s}, \mathbf{x}, \omega)|^2 + \epsilon)^{-1} \\ &\times (\nabla_{\mathbf{x}} S_0^*(\mathbf{s}, \mathbf{x}, \omega) \cdot \nabla_{\mathbf{x}} R_0(\mathbf{s}, \mathbf{x}, \omega) + R_0(\mathbf{s}, \mathbf{x}, \omega) \Delta_{\mathbf{x}} S_0^*(\mathbf{s}, \mathbf{x}, \omega)). \end{aligned} \quad (3.28)$$

The difference is only the second term in the numerator, that is, $\Delta_{\mathbf{x}} S_0^*(\mathbf{s}, \mathbf{x}, \omega)$ versus $\frac{i\omega}{c_0} S_0^*(\mathbf{s}, \mathbf{x}, \omega)$. These two terms are the same, except at the source position. Consequently, we expect better preservation of early arrivals for our approach.

To validate the reliability of our method, we apply our approach to the constant density isotropic acoustic Marmousi model. The data are generated by finite difference Born modeling in the correct background velocity and reflectivity models (Figure 3.26a). We use a single shot located at 4.6 km at the surface. Receivers are symmetrically deployed on both sides of shot with offsets ranging from -4 to 4 km (Figure 3.26b).

We then migrate observed data in the correct background velocity using methods of [Qin et al. \(2015\)](#) and ours (Figures 3.26c and 3.26e), respectively. We use migrated images as reflectivity models to generate reconstructed data (Figures 3.26d and 3.26f). Both methods provide the same result, except around the first arrivals, due to non-zero energy at the source position in the inverted images. Our method is closer to the exact data than [Qin et al. \(2015\)](#) for the early arrivals. [Qin et al. \(2015\)](#)'s approach adds an artificial value in the inverted image at the source position such that the direct arrival in reconstructed data is less accurate than ours.

The exact and reconstructed data should perfectly match if an inverse of the Born modeling is used to retrieve the reflectivity image. We select 4 traces at positions 2, 3, 5 and 6 km (Figure 3.27), to compare the quality of reconstructions in details for both approaches. Our method preserves amplitudes better than [Qin et al. \(2015\)](#). The observed and modeled data are consistent even at large shot-receiver offset (Figures 3.27e and 3.27h). We also compute the root-mean-square error between exact and reconstructed data trace by trace (Figure 3.28). Our method presents an overall lower RMS error, which means our reconstruction is closer to exact data. This is due to missing energy around the first arrivals.

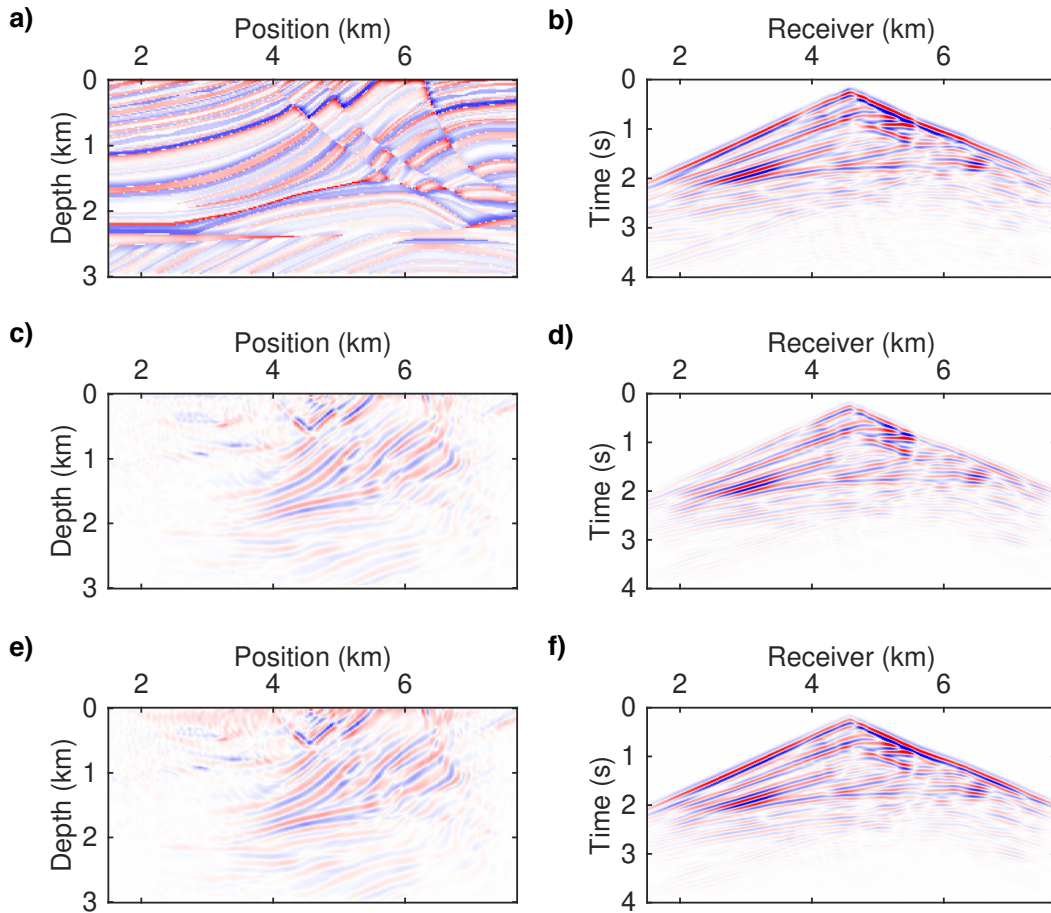


Figure 3.26 – Single-shot test in the Marmousi model for a shot located at 4.6 km at the surface. (a) True reflectivity model; (b) Data modeled from (a); (c) Migrated image using the method of [Qin et al. \(2015\)](#); (d) Data modeled from (c); (e) Migrated image using our method; (f) Data modeled from (e). (a), (c) and (e) are plotted with the same scale. (b), (d) and (f) are all modeled in a correct smooth background model and plotted at the same scale.

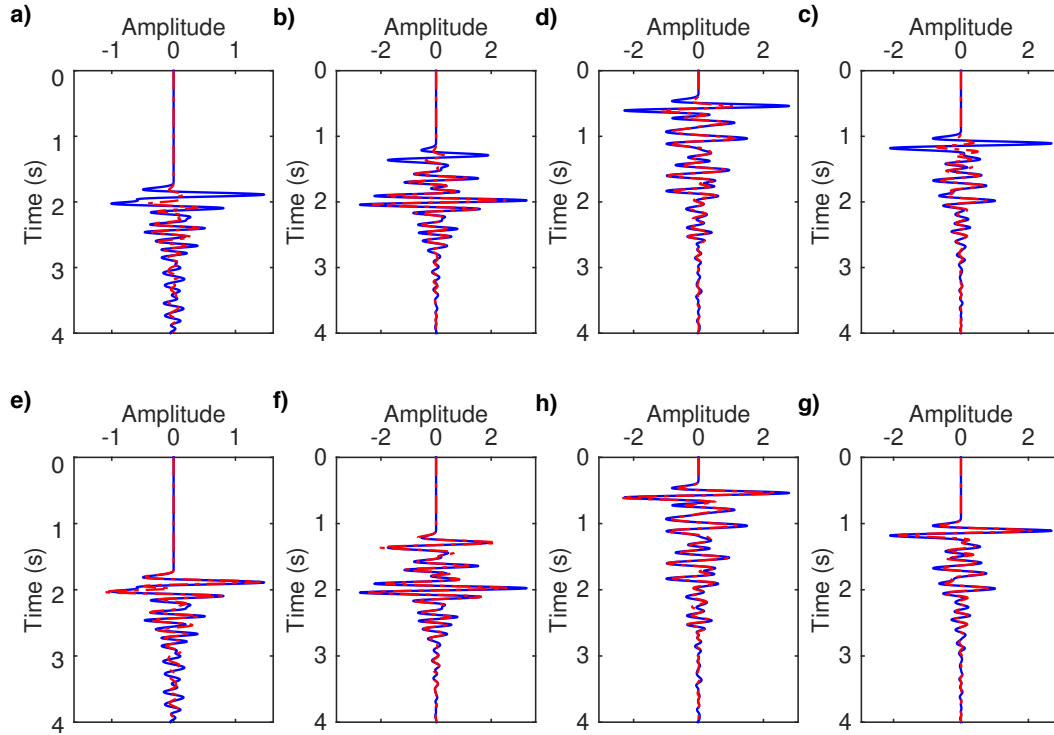


Figure 3.27 – Comparison between exact and reconstructed single trace data. (a–d) are examples using the method of [Qin et al. \(2015\)](#). (e–h) are examples using our method. From left to right, the column refers to the single trace data located at receiver positions 2, 3, 5 or 6 km, respectively. Blue curve represents exact data. Red curve represents reconstructed data.

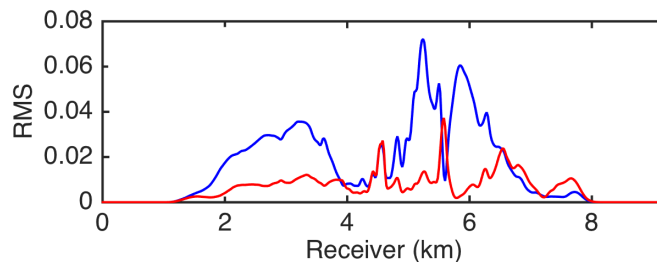


Figure 3.28 – RMS misfit between observed and reconstructed data for each receiver position. Blue curve represents [Qin et al. \(2015\)](#)'s method. Red curve represents the approach developed here.

3.8 Appendix II: Gradient derivation for IVA

We explain here how to compute the gradient of the IVA objective function (equation 3.13) using the adjoint-state technique (Plessix, 2006). We consider the source and receiver wavefields defined in equations 3.10 and 3.11. In the following, we note the 2D acoustic Helmholtz operator:

$$\mathcal{L}_0 = -\frac{\omega^2}{c_0^2(\mathbf{x})} - \Delta. \quad (3.29)$$

We extend equation 3.13 with Lagrangian formulation for $\alpha = 0$, denoted by \mathcal{J}_{ext}^0 , depending on $(c_0, \xi_{inv}, S_0, R_0, \lambda_\xi, \lambda_s, \lambda_r)$ and reading

$$\begin{aligned} \mathcal{J}_{ext}^0 = & \frac{1}{2} \|D_{s_x} \xi_{inv}(\mathbf{x}, \mathbf{s})\|^2 \\ & - \iint d\mathbf{x} d\mathbf{s} \lambda_\xi(\mathbf{x}, \mathbf{s}) [\xi_{inv}(\mathbf{x}, \mathbf{s}) - 4 \int d\omega (|S_0(\mathbf{s}, \mathbf{x}, \omega)|^2 + \epsilon)^{-1} \\ & \quad \times (\nabla_{\mathbf{x}} S_0^*(\mathbf{s}, \mathbf{x}, \omega) \cdot \nabla_{\mathbf{x}} R_0(\mathbf{s}, \mathbf{x}, \omega) + (\frac{i\omega}{c_0(\mathbf{x})})^2 S_0^*(\mathbf{s}, \mathbf{x}, \omega) R_0(\mathbf{s}, \mathbf{x}, \omega))] \\ & - \iiint d\mathbf{x} d\mathbf{s} d\omega \lambda_s^*(\mathbf{s}, \mathbf{x}, \omega) [\mathcal{L}_0 S_0(\mathbf{s}, \mathbf{x}, \omega) - (i\omega)^3 \Omega(\omega) \delta(\mathbf{x} - \mathbf{s})] \\ & - \iiint d\mathbf{x} d\mathbf{s} d\omega \lambda_r^*(\mathbf{s}, \mathbf{x}, \omega) [\mathcal{L}_0^* R_0(\mathbf{s}, \mathbf{x}, \omega) - \int d\mathbf{r} d^{obs}(\mathbf{s}, \mathbf{r}, \omega) \frac{\partial \delta(\mathbf{x} - \mathbf{r})}{\partial r_z}], \end{aligned} \quad (3.30)$$

where λ_ξ , λ_s and λ_r are adjoint state variables associated with ξ , S_0 and R_0 , respectively. The different terms can be interpreted as constraints. The total derivative of \mathcal{J}_{ext} with respect to the velocity is given by

$$\begin{aligned} \frac{d\mathcal{J}_{ext}}{dc_0} = & \frac{\partial \mathcal{J}_{ext}}{\partial c_0} + \frac{\partial \xi_{inv}}{\partial c_0} \frac{\partial \mathcal{J}_{ext}}{\partial \xi_{inv}} + \frac{\partial S_0}{\partial c_0} \frac{\partial \mathcal{J}_{ext}}{\partial S_0} + \frac{\partial R_0}{\partial c_0} \frac{\partial \mathcal{J}_{ext}}{\partial R_0} \\ & + \frac{\partial \lambda_\xi}{\partial c_0} \frac{\partial \mathcal{J}_{ext}}{\partial \lambda_\xi} + \frac{\partial \lambda_s}{\partial c_0} \frac{\partial \mathcal{J}_{ext}}{\partial \lambda_s} + \frac{\partial \lambda_r}{\partial c_0} \frac{\partial \mathcal{J}_{ext}}{\partial \lambda_r}. \end{aligned} \quad (3.31)$$

We develop the IVA gradient by taking $(\frac{\partial \mathcal{J}_{ext}}{\partial \xi_{inv}}, \frac{\partial \mathcal{J}_{ext}}{\partial S_0}, \frac{\partial \mathcal{J}_{ext}}{\partial R_0}, \frac{\partial \mathcal{J}_{ext}}{\partial \lambda_\xi}, \frac{\partial \mathcal{J}_{ext}}{\partial \lambda_s}, \frac{\partial \mathcal{J}_{ext}}{\partial \lambda_r}) = 0$. It avoids the computation of Fréchet derivatives $(\frac{\partial \xi_{inv}}{\partial c_0}, \frac{\partial S_0}{\partial c_0}, \frac{\partial R_0}{\partial c_0}, \frac{\partial \lambda_\xi}{\partial c_0}, \frac{\partial \lambda_s}{\partial c_0}$ and $\frac{\partial \lambda_r}{\partial c_0})$. The partial derivatives $\frac{\partial \mathcal{J}_{ext}}{\partial \lambda_\xi}$, $\frac{\partial \mathcal{J}_{ext}}{\partial \lambda_s}$ and $\frac{\partial \mathcal{J}_{ext}}{\partial \lambda_r}$ equal to zero leads to the definitions of inverted image, forward wavefield and backward wavefield, respectively. The derivatives of \mathcal{J}_{ext} with respect to ξ , S_0 and R_0 , imposed to 0, give the expressions to calculate the adjoint

state variables λ_ξ , λ_s and λ_r , respectively:

$$\begin{aligned} \mathcal{L}_0^* \lambda_s(\mathbf{s}, \mathbf{x}, \omega) = & -4 \left[\nabla_{\mathbf{x}} \cdot (U(\mathbf{s}, \mathbf{x}, \omega) \nabla_{\mathbf{x}} R_0(\mathbf{s}, \mathbf{x}, \omega)) \right. \\ & - \left(\frac{i\omega}{c_0(\mathbf{x})} \right)^2 U(\mathbf{s}, \mathbf{x}, \omega) R_0(\mathbf{s}, \mathbf{x}, \omega) \\ & \left. + 2S_0(\mathbf{s}, \mathbf{x}, \omega) U(\mathbf{s}, \mathbf{x}, \omega) \Re \left\{ \tilde{\xi}_{inv}(\mathbf{s}, \mathbf{x}, \omega) \right\} \right], \end{aligned} \quad (3.32)$$

$$\begin{aligned} \mathcal{L}_0 \lambda_r(\mathbf{s}, \mathbf{x}, \omega) = & -4 \left[\nabla_{\mathbf{x}} \cdot (U(\mathbf{s}, \mathbf{x}, \omega) \nabla_{\mathbf{x}} S_0(\mathbf{s}, \mathbf{x}, \omega)) \right. \\ & \left. - \left(\frac{i\omega}{c_0(\mathbf{x})} \right)^2 U(\mathbf{s}, \mathbf{x}, \omega) S_0(\mathbf{s}, \mathbf{x}, \omega) \right], \end{aligned} \quad (3.33)$$

$$\lambda_\xi(\mathbf{x}, \mathbf{s}) = -D_{sx}^2 \xi_{inv}(\mathbf{x}, \mathbf{s}), \quad (3.34)$$

where $U = \lambda_\xi / (|S_0|^2 + \epsilon)$ for simplification and $\tilde{\xi}_{inv}(\mathbf{s}, \mathbf{x}, \omega)$ is the integrand of $\xi_{inv}(\mathbf{s}, \mathbf{x})$. Equations 3.32–3.34 are solved in reverse order. Subsequently, the final gradient is obtained by inserting the values of the forward and backward propagated wavefields, inverted images, and associated adjoint state variables into equation 3.15.

Chapter 4

Investigating the links between image and data domains

Contents

4.1	Introduction	121
4.2	Common-shot IVA in the data-domain	125
4.2.1	Objective function	125
4.2.2	Gradient of the objective function	127
4.2.3	Numerical examples	129
4.3	Comparisons between image and data domains	149
4.3.1	Equivalence between data-domain IVA and DWI	149
4.3.2	Data fitting versus image coherency	150
4.3.3	Numerical comparisons	154
4.4	Conclusions	164
4.5	Appendix I: Gradient derivation for IVA in the data-domain	166

Résumé du chapitre 4

Dans ce chapitre, nous comparons trois approches différentes :

- L'analyse de vitesse par inversion dans le domaine image (introduite au Chapitre 3) ;
- Une extension proposée ici, définie dans le domaine des données ;
- L'inversion d'onde différentielle (introduite dans le Chapitre 2), une méthode définie dans le domaine des données et basée sur les arrivées réfléchies.

L'analyse de vitesse par migration est une technique pour la détermination des grandes structures de la sub-surface. Elle est définie dans le domaine image et évalue la qualité du macro-modèle au travers de la cohérence des images migrées. L'analyse de vitesse par inversion est une alternative : elle remplace la migration, l'adjoint de l'opérateur de Born, par un inverse direct. Ceci conduit à des résultats plus stables. Pour établir le lien entre les méthodes dans les domaines des données et des images, nous proposons d'étendre l'analyse de vitesse par inversion au domaine des données, pour une version d'inversion des formes d'onde linearisée. Pour cela, nous appliquons l'opérateur de modélisation sur les images différenciées. Par images différenciées, nous entendons des images qui ont été multipliées par l'annihilateur. La qualité du macro-modèle est estimée par la norme de ces données. Nous discutons le comportement de cette nouvelle approche au travers d'applications en 2D, d'abord dans un milieu homogène avec un seul réflecteur, puis sur le modèle Marmousi. Ensuite, nous établissons des liens quantitatifs entre la RWI, une technique de minimisation des résidus des données, et la semblance différentielle, basée sur la cohérence des images. Enfin, nous comparons numériquement les approches dans les domaines des données et des images. Les deux approches sont robustes et donnent des résultats comparables, au moins dans le cas de Marmousi. Ainsi, un lien est établi entre les deux approches. Il est cohérent avec l'analyse théorique. Ces travaux donnent une meilleure compréhension des relations entre les méthodes définies dans les domaines données et images.

Abstract

In this Chapter, we compare three different approaches:

- common-shot Inversion Velocity Analysis in the image-domain (introduced in Chapter 3);
- an extension (proposed here) defined in the data-domain (data-domain Inversion Velocity Analysis);
- Differential Waveform Inversion (introduced in Chapter 2), a reflection-based approach defined in the data-domain.

Migration Velocity Analysis is a technique for the determination of the large-scale structures of subsurface. It is defined in the image-domain and assesses the quality of the estimated macromodel by measuring the image coherency. Inversion Velocity Analysis is an alternative: it replaces the migration, adjoint of Born modeling, by the direct inverse of the modeling operator, leading to more stable results. To investigate the links between image-domain and data-domain methods, we propose to extend Inversion Velocity Analysis to the data-domain, formulating a linearized waveform inversion approach. For that, we map the differentiated image back to the data space by simply applying the modeling operator onto it. By differentiated image, we mean the image multiplied by the annihilator. We assess the quality of an estimated macromodel by measuring the norm of these residual data. We discuss the behavior of this new approach through applications to 2D synthetic data sets, first on a homogeneous model with a single-reflector embedded in, and then on the Marmousi model. Then, we establish the quantitative links between reflection-based waveform inversion, a technique based on data fitting principle, and differential semblance optimization, a scheme measuring the image coherency. Finally, we numerically compare the results of image-domain and data-domain IVA. The two approaches are robust and provide comparable results, at least for the Marmousi model. Therefore, a quantitative equivalence between image-domain and data-domain methods is established. This is consistent with the theoretical analysis. These investigations provide a better understanding for the relationships between image-domain and data-domain methods.

4.1 Introduction

An accurate velocity model is an essential requirement for a successful seismic imaging procedure. Full Waveform Inversion (FWI) seeks a velocity model that can best explain the complete data sets (Tarantola, 1984a; Pratt et al., 1998). The major problem of FWI is that the relationship between data and velocity model is highly nonlinear. Therefore, FWI suffers from a severe cycle-skipping problem, meaning that its objective function

has many local minima (Gauthier et al., 1986; Bunks et al., 1995). This leads to the requirements for including low-frequency contents in the data and/or for starting from an estimated velocity model close enough to the true one (Beydoun and Tarantola, 1988; Sirgue, 2006; Pratt et al., 2008). However, the required low frequencies, typically below 1 Hz in the exploration context, are very difficult to acquire during controlled-source experiments (Virieux and Operto, 2009). It is thus needed to obtain a good starting model with other approaches for subsequent FWI.

Mora (1989) recognized that seismic inverse problem has a migration mode and a tomographic mode. Assuming a scale separation of the subsurface model (Claerbout, 1985), the migration and tomographic modes correspond to the determination of the small-scale (i.e. the model perturbation) and of the large-scale (i.e. the background model or macromodel) parts of the model, respectively. Since the migration mode only characterizes the local rapid variations of the model, the approaches used to build a starting model for FWI should behave in a tomographic mode to retrieve the large-scale structure of the subsurface. In the regions illuminated by transmitted waves, FWI itself can retrieve the large-scale part of the model (Alkhalifah and Plessix, 2014; Brossier et al., 2015; Zhou et al., 2015), but the penetration depth of transmitted events is limited due to the finite range of the acquisition offset in practice (Zhou, 2016). Thus, reflection-based tomographic approaches are particularly important to provide a starting model for FWI in the region not reached by transmitted waves. These tomographic approaches can be formulated either in the image-domain or in the data-domain (Díaz et al., 2014). Many tomographic approaches do not split the model into the perturbation model and the macromodel (Van Leeuwen and Mulder, 2010; Liu et al., 2011b; Wu et al., 2014; Warner and Guasch, 2014; Bharadwaj et al., 2015 among others). However, the following section focuses on those techniques relying on the scale separation assumption and on the Born approximation.

Image-domain methods mainly refer to Migration Velocity Analysis (MVA), a technique to determine the background model controlling the kinematics of wave propagation. Under the Born approximation, the reflected data linearly depend on the model perturbation. In an estimated macromodel, the perturbation images are obtained through migration defined as the crosscorrelation between source and receiver wavefields (Claerbout, 1971). Once the data are migrated, one can evaluate the quality of the macromodel by measuring the coherency of these images. Conventional MVA splits the data into subsets, such as common-shot (Al-Yahya, 1989; Symes and Carazzone, 1991) or common-offset (Chauris and Noble, 2001; Mulder and ten Kroode, 2002), and then migrate them separately in the same background model. If all images are consistent, the current macromodel is considered to be optimal. Moveout residuals in Common-Image Gathers (CIGs) carry the information about the inaccuracies of the macromodel. Differential Semblance Optimization (DSO) (Symes and Carazzone, 1991; Symes, 2008) is a way to evaluate the quality of the model by comparing adjacent images.

Later, other extensions were proposed for MVA. Instead of splitting data into sub-

sets, the full data set is migrated and then a spatial or temporal shift is introduced in the imaging condition (Faye and Jeannot, 1986; Sava and Fomel, 2006; Symes, 2008; Sava and Vasconcelos, 2011). Thus, the estimated model is optimal if the energy in associated CIGs focuses around the zero spatial or temporal shift. The spatial shift is also called subsurface-offset. Alternatively, a scattering angle can be considered as the extension parameter to build the angle-domain CIGs (ADCIGs) (Xu et al., 2001; Sava and Fomel, 2003; Biondi and Symes, 2004). One important issue is that migration is by definition the adjoint operator of the Born modeling; in practice, artifacts are present in migrated images. In the extended domain, it was recently proposed to replace migration by an inversion scheme: first iterative migration (Liu et al., 2014b; Chauris et al., 2015; Cocher et al., 2015), then direct inversion (ten Kroode, 2012; Hou and Symes, 2015; Lameloise et al., 2015; Chauris and Cocher, 2017). In that case, the migration artifacts are highly attenuated, leading to a more robust macromodel estimation approach, namely Inversion Velocity Analysis (IVA). However, the extended subsurface-offset approach is very demanding in terms of the memory requirement, preventing from 3D extension (Yang and Sava, 2011; van Leeuwen et al., 2015; Chauris and Cocher, 2017). We proposed to apply the same strategy to common-shot MVA in Chapter 3 for the feasibility of future 3D extension.

Data-domain methods here mainly refer to the techniques attempting to introduce the principle of MVA into waveform inversion. One possibility is to introduce the concept of *extended modeling* proposed by Symes (2008) to formulate a forward modeling engine: the wave equation is linearized based on an extension of the perturbation model along the spatial (Almomin and Biondi, 2012; Liu et al., 2014a) or temporal (Biondi and Almomin, 2012, 2014) shift axis. This extra parameter is same as the one used to extend the image space in depth-oriented MVA. Based on this extension, they define an objective function that has an image focusing term in addition to the conventional FWI data fitting term. The approaches based on spatial or temporal shift are recognized as Extended Waveform Inversion (EWI) (Liu et al., 2014a; Fu and Symes, 2015; Symes, 2017) and Tomographic Full Waveform Inversion (TFWI) (Almomin and Biondi, 2012; Biondi and Almomin, 2012, 2014). Alternatively, Chauris et al. (2017) proposed to obtain the reflectivity images with a true-amplitude migration (direct inversion) scheme in the subsurface-offset domain. After applying an annihilator to inverted images, one can model the associated data in an estimated macromodel. Then, the optimal macromodel is determined by minimizing the norm of this new data. Thus, the image focusing criteria is extended to the data-domain. Despite a different objective function, this approach also introduces the linearized modeling extended by the subsurface-offset parameter. One can view this family as the extension of depth-oriented IVA to data-domain.

On the other hand, one can directly minimize the misfit between modeled and observed reflected data in the least-squares sense, to formulate Reflection Waveform Inversion (RWI) (Xu et al., 2012; Brossier et al., 2015; Zhou et al., 2015; Wu and Alkhalifah, 2015). It is inspired by migration based traveltimes tomography (MBTT) (Chavent

et al., 1994; Plessix et al., 1995). The implementation of Xu et al. (2012) assumes the model perturbation and macro velocity model are independent parameters. The short-wavelength and long-wavelength parts of the model are updated by minimizing the same objective function. One first determines the model perturbation in a given background velocity model, and then updates the macromodel with only short-offset data (even zero-offset) (Alkhalifah and Wu, 2016; Zhou, 2016). The modeled data with larger offset range are predicted from the inverted reflectivities in the estimated macromodel. Then, the data misfit is used to update the macromodel. However, the model perturbation depends on the background model, such that the two parameters are supposed to be updated in a nested optimization procedure, as for MVA (Cocher et al., 2017b). One can alternatively formulate a reflection-based approach defined in the data-domain, called Differential Waveform Inversion (DWI) (Chauris and Plessix, 2012, 2013). It consists of two main steps: an image section is retrieved from a given shot gather and used to calculate synthetic data for the next shot; then the macro model is updated by minimizing the difference between the predicted shot gather and the observed shot gather at the next shot position. This procedure is inspired by DSO (Symes and Kern, 1994; Chauris and Plessix, 2012). These approaches are all reflection-based waveform inversion, relying on the scale separation and on Born approximation. We shall see in this Chapter that one can consider them as the extension of surface-oriented IVA to the data-domain.

Despite many approaches introducing the concept of MVA to waveform inversion, the relationship between image-domain and data-domain methods has not been fully understood. Chauris et al. (2002a) investigated the link between MVA Symes and Carazzone (1991); Sava and Biondi (2004) and slope tomography (Billette and Lambaré, 1998; Lambaré, 2008) in the context of ray-based methods. In this Chapter, we first extend the common-shot IVA approach from image-domain to data-domain following the same strategy proposed by (Chauris and Cocher, 2017). We construct the reflectivity images with a direct inverse applied to common-shot gathers. Then, the data are computed from the differential images, which is derived by comparing images for neighbouring shots. We define the objective function as a ℓ_2 -norm of this new data and explain how to derive the gradient with respect to the background velocity model. This procedure formulates a nested optimization, where the inner loop is dedicated to retrieving the reflectivity and the outer loop to updating the macromodel. Numerical examples in synthetic cases are presented to discuss the advantages and the limitations. We then investigate the relation between the new approach and reflection-based waveform inversion. The purpose is to investigate the explicit link between image-domain and data-domain for wave-equation-based methods.

4.2 Common-shot IVA in the data-domain

We consider a velocity model c as the sum of the background model c_0 which controls the kinematics of the wave propagation and the model perturbation δc which triggers scattered waves. We define the reflectivity as $\xi = 2\delta c^3/c_0$. We update δc and c_0 in a nested optimization procedure for the new approach.

4.2.1 Objective function

We introduce here the new objective function defined in the data-domain. The model c_0 is iteratively determined by minimizing the new objective function in the outer loop procedure, whereas in an estimated c_0 model, an optimal ξ is obtained by solving the inner loop problem, for which we consider the direct inversion developed in Chapter 3.

In a given macro velocity model c_0 , we define the Born modeling as a linear relationship between reflectivity ξ and reflected data d for every single shot. Here, the reflectivity ξ , depending on the spatial coordinates \mathbf{x} and on the shot coordinates \mathbf{s} , is obtained with the convolution of two Green's functions

$$\begin{aligned} d(\mathbf{s}, \mathbf{r}, \omega) &= (\mathcal{B}_0 \xi)(\mathbf{s}, \mathbf{r}, \omega) \\ &= \int d\mathbf{x} (i\omega)^2 \Omega(\omega) G_0(\mathbf{s}, \mathbf{x}, \omega) \xi(\mathbf{x}, \mathbf{s}) G_0(\mathbf{x}, \mathbf{r}, \omega), \end{aligned} \quad (4.1)$$

where \mathbf{r} is the receiver coordinates and Ω the source wavelet, function of the angular frequency ω . The linear operation between the data d and the model perturbation ξ is denoted by \mathcal{B}_0 , equation 3.1 in Chapter 3. The migrated section ξ_{mig} is obtained by applying the adjoint \mathcal{B}_0^T of \mathcal{B}_0 to the observed data. For common-shot gathers, we came up with a pseudo-inverse \mathcal{B}_0^\dagger of the Born modeling operator in Chapter 3, which is an alternative to the formulations proposed by [Qin et al. \(2015\)](#); [Duprat and Baina \(2016\)](#). It is a true-amplitude migration scheme in the asymptotic sense (see section 3.7). By true-amplitude migration, we mean the inverted images can perfectly reproduce the observed data. Note that the true-amplitude scheme used here is a direct inversion rather than an iterative migration. By definition, we have

$$\xi_{mig} = \mathcal{B}_0^T d^{obs}, \quad (4.2)$$

$$\xi_{inv} = \mathcal{B}_0^\dagger d^{obs}. \quad (4.3)$$

The adjoint operator \mathcal{B}_0^T represents the conventional reverse time migration, defined as the crosscorrelation between source and receiver wavefields ([Claerbout, 1971](#); [Baysal](#)

et al., 1983). Referred to equation 3.9 in Chapter 3, the inverse \mathcal{B}_0^\dagger is given by

$$\begin{aligned} \xi_{inv}(\mathbf{x}, \mathbf{s}) &= (\mathcal{B}_0^\dagger d^{obs})(\mathbf{x}, \mathbf{s}) = 4 \int d\omega (|S_0(\mathbf{s}, \mathbf{x}, \omega)|^2 + \epsilon)^{-1} \\ &\times \left(\nabla_{\mathbf{x}} S_0^*(\mathbf{s}, \mathbf{x}, \omega) \cdot \nabla_{\mathbf{x}} R_0(\mathbf{s}, \mathbf{x}, \omega) + \left(\frac{i\omega}{c_0(\mathbf{x})} \right)^2 S_0^*(\mathbf{s}, \mathbf{x}, \omega) R_0(\mathbf{s}, \mathbf{x}, \omega) \right), \end{aligned} \quad (4.4)$$

with modified source and receiver wavefields

$$S_0(\mathbf{s}, \mathbf{x}, \omega) = (i\omega)^3 G_0(\mathbf{s}, \mathbf{x}, \omega) \Omega(\omega), \quad (4.5)$$

$$R_0(\mathbf{s}, \mathbf{x}, \omega) = \int d\mathbf{r} D_{r_z} G_0^*(\mathbf{r}, \mathbf{x}, \omega) d^{obs}(\mathbf{s}, \mathbf{r}, \omega), \quad (4.6)$$

where D_{r_z} denotes the partial vertical derivative at the receiver position. Although the derivation relies on ray-based expressions, this common-shot inversion formula contains only wave-equation-based terms. It is defined such that $\mathcal{B}_0 \mathcal{B}_0^\dagger d = d$ even in an incorrect background velocity model.

In the 2D case, we develop the objective function of common-shot IVA from the image-domain to the data-domain

$$\mathcal{J}_{image}^\alpha[c_0] = \frac{1}{2} \|c_0^\alpha D_{s_x} \xi_{inv}\|_\xi^2 = \frac{1}{2} \|c_0^\alpha D_{s_x} \mathcal{B}_0^\dagger d^{obs}\|_\xi^2, \quad (4.7)$$

$$\mathcal{J}_{data}^\beta[c_0] = \frac{1}{2} \|\mathcal{B}_0 c_0^\beta D_{s_x} \xi_{inv}\|_{\mathcal{D}}^2 = \frac{1}{2} \|\mathcal{B}_0 c_0^\beta D_{s_x} \mathcal{B}_0^\dagger d^{obs}\|_{\mathcal{D}}^2. \quad (4.8)$$

The subscripts ξ and \mathcal{D} indicate that the ℓ_2 -norms are evaluated in the image-domain and data-domain, respectively. D_{s_x} is the derivative with respect to the horizontal source position, same as for the case of DSO (Symes and Kern, 1994). For simplification, we define $d^\beta = \mathcal{B}_0 c_0^\beta D_{s_x} \xi_{inv}$ in the following. The new data d^β is a function of $(\mathbf{s}, \mathbf{r}, \omega)$. We call it *scaling data*. These two objective functions both evaluate the consistency of the reflectivity sections. Equation 4.7 has already been discussed in Chapter 3. Equation 4.8 corresponds to the new approach defined in the data-domain. In practice, after the application of derivative D_{s_x} to the inverted reflectivity model ξ_{inv} , scaling data are modeled from the differentiate images under the Born approximation (equation 4.1). It is expected that the ℓ_2 -norm of these data is minimal for the correct velocity model. The inaccuracies of macromodel correspond to the non-horizontal events in CIGs for image-domain IVA, whereas the non-zero values of scaling data d^β indicate that the macromodel should be updated for data-domain IVA (Figure 4.1). As already discussed in Chapter 3, the introduction of c_0^α in equation 4.7 allows more flexibility in the shape of the gradient $\frac{\partial}{\partial c_0} \mathcal{J}_{image}^\alpha$: the additional contribution to the gradient for a non-zero α is localized around the interfaces and may attenuate the spurious oscillations related to interface imprints. For the same purpose, c_0^β is introduced in data-domain IVA. However, we explain later that $\beta = 0$ is sufficient in practice.

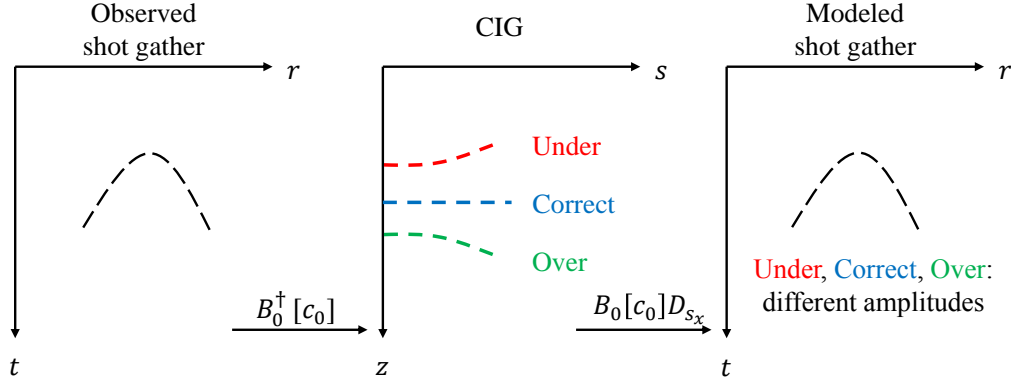


Figure 4.1 – Image-domain and data-domain IVA. Left is observed shot gather. Middle is the constructed CIG after the application of the inverse operator to the an observed data. Red, blue and greed dashed lines correspond to the cases of under-estimated, correct and over-estimated background velocity models, respectively. Right are the data modeled from the differential images in the given background velocity. Different background velocity correspond to modeled data of almost the same kinematics and they are distinguished by their amplitudes.

4.2.2 Gradient of the objective function

We use the adjoint-state technique (Chavent, 1974; Plessix, 2006) to efficiently derive the gradient (see details in Appendix 4.5). It is a general method to compute the gradient of an objective function with respect to model parameters through a set of state variables: $\tilde{\lambda}_s$, $\tilde{\lambda}_r$ and $\tilde{\lambda}_\xi$ for image-domain IVA, or λ_s , λ_r , λ_ξ and λ_d for data-domain IVA. These adjoint-state variables are the solutions of an adjoint linear system. They can be introduced as Lagrangian constraints. We first define a wavefield term S_d , reading

$$S_d(\mathbf{s}, \mathbf{x}, \omega) = \int d\mathbf{y} (i\omega)^2 \Omega(\omega) G_0(\mathbf{s}, \mathbf{y}, \omega) G_0(\mathbf{y}, \mathbf{x}, \omega) D_{s_x} \xi(\mathbf{y}, \mathbf{s}). \quad (4.9)$$

where y is the auxiliary spatial coordinates. S_d is the forward scattering wavefield triggered by $D_{s_x} \xi$ in the background velocity c_0 . The scaling data d^0 is the response of S_d at receiver positions.

The macromodel gradient for image-domain IVA is already introduced in Chapter 3. For the reader's convenience, we write here the expressions of associated gradients for

both image-domain and data-domain IVA

$$\begin{aligned} \frac{\partial \mathcal{J}_{image}^0}{\partial c_0(\mathbf{x})} = 2c_0^{-3}(\mathbf{x}) & \left[\iint d\mathbf{s} d\omega (i\omega)^2 \tilde{\lambda}_s^*(\mathbf{s}, \mathbf{x}, \omega) S_0(\mathbf{s}, \mathbf{x}, \omega) \right. \\ & + \iint d\mathbf{s} d\omega (i\omega)^2 \tilde{\lambda}_r^*(\mathbf{s}, \mathbf{x}, \omega) R_0(\mathbf{s}, \mathbf{x}, \omega) \\ & \left. - \iint d\mathbf{s} d\omega (i\omega)^2 \tilde{\lambda}_\xi(\mathbf{x}, \mathbf{s}) \frac{4S_0^*(\mathbf{s}, \mathbf{x}, \omega) R_0(\mathbf{s}, \mathbf{x}, \omega)}{|S_0(\mathbf{s}, \mathbf{x}, \omega)|^2 + \epsilon} \right], \end{aligned} \quad (4.10)$$

$$\begin{aligned} \frac{\partial \mathcal{J}_{data}^0}{\partial c_0(\mathbf{x})} = 2c_0^{-3}(\mathbf{x}) & \left[\iint d\mathbf{s} d\omega (i\omega)^2 \lambda_d^*(\mathbf{s}, \mathbf{x}, \omega) S_d(\mathbf{s}, \mathbf{x}, \omega) \right. \\ & + \iint d\mathbf{s} d\omega (i\omega)^2 \lambda_s^*(\mathbf{s}, \mathbf{x}, \omega) S_0(\mathbf{s}, \mathbf{x}, \omega) \\ & + \iint d\mathbf{s} d\omega (i\omega)^2 \lambda_r^*(\mathbf{s}, \mathbf{x}, \omega) R_0(\mathbf{s}, \mathbf{x}, \omega) \\ & \left. - \iint d\mathbf{s} d\omega (i\omega)^2 \lambda_\xi(\mathbf{x}, \mathbf{s}) \frac{4S_0^*(\mathbf{s}, \mathbf{x}, \omega) R_0(\mathbf{s}, \mathbf{x}, \omega)}{|S_0(\mathbf{s}, \mathbf{x}, \omega)|^2 + \epsilon} \right]. \end{aligned} \quad (4.11)$$

Similar to $(\tilde{\lambda}_s, \tilde{\lambda}_r, \tilde{\lambda}_\xi)$ for image-domain IVA, parameters λ_s and λ_r are the scattering wavefields triggered by the product of λ_ξ , the adjoint source, and the back or forward propagated wavefields. Crosscorrelations $\lambda_s \star S_0$ and $\lambda_r \star R_0$ contribute to a tomographic update for data-domain IVA same as for image-domain IVA. λ_d is the receiver wavefield related to scaling data d^0 . Crosscorrelation $\lambda_d \star S_d$ has a similar behavior as $\lambda_r \star R_0$. The fourth integration term on the right side of equation 4.11 is similar to deconvolution migration and localized around interfaces. Those components are oscillatory and can be removed by spatially smoothing the gradient. For the macromodel gradient, the major differences between image-domain and data-domain IVA are that: (1) λ_ξ is related to the scaling data d^0 for data-domain IVA, whereas $\tilde{\lambda}_\xi$ is the partial derivative of equation 4.7 with respect to image ξ ; (2) the expressions of macromodel gradient for data-domain IVA has one more tomographic component ($\lambda_d \star S_d$) than for image-domain IVA. These additional terms require solving wave equation twice more, leading to a more CPU time-consuming implementation. The readers are referred to table 4.1 for a comparison of FWI, image-domain IVA and data-domain IVA in terms of the CPU cost for calculating the gradient of the objective function. The mathematical definitions of adjoint variables are given in appendix 4.5.

In the next section, we investigate how IVA is extended from image-domain to data-domain in practice. Then, we pay attention to the shape of the gradients, especially around the reflector location, through numerical examples. As the example will show that $\beta = 0$ is sufficient for data-domain IVA, we did not give the generic formulation for gradient of \mathcal{J}_{data}^β with respect to the macromodel.

Table 4.1 – Comparison between different approaches in terms of computational requirements for calculating the gradient of associated objective functions with respect to model. n_s is the number of shots. For a given source term, q_{mod} is the CPU cost required to solve the wave equation once. The CPU cost for crosscorrelation is neglected here.

	FWI/RTM	Image-domain IVA	Data-domain IVA
CPU cost	$2n_s q_{mod}$	$4n_s q_{mod}$	$6n_s q_{mod}$

4.2.3 Numerical examples

We investigate two 2D synthetic cases. The first example, a single horizontal reflector embedded in a homogeneous model, gives insights into (1) what is measured by the new data-domain objective function and (2) why $\beta = 0$ is sufficient for the data-domain IVA. In this case, we only display the first gradient. Then, the second test is performed on the Marmousi model containing complex structures, such as discontinuities and rough interfaces. We perform nonlinear iterations to update the macro velocity model for this case. In all cases, the data are generated with a 4th-order staggered-grid finite-difference scheme including the Convolution Perfectly Matched Layer (CPML) absorbing boundary (Virieux, 1986; Levander, 1988; Komatitsch and Martin, 2007).

Homogeneous model

We start with a simple model containing a single horizontal reflector at depth $z = 1.0$ km embedded in a homogeneous model, for which the velocity is 2.5 km/s. We trigger 81 shots every 0.04 km from 2.4 to 5.6 km at the surface. The source function is a Ricker wavelet with a maximum frequency of 10 Hz. Receivers are symmetrically deployed every 0.02 km on both sides of shots with offsets ranging from -1.6 to 1.6 km.

After the application of direct inverse (equation 4.4), we invert the data in three constant velocity models, for which the velocities are 2.2, 2.5 and 2.8 km/s, respectively. In an ideal stacked image section, we expect only horizontal reflectivities below (for velocity too high), around (for accurate velocity) or above (for velocity too low) the correct depth.

The three examples exhibit flat interfaces at around $z = 0.8$ km (Figure 4.2a), at around $z = 0.6$ km (Figure 4.2b) and at around $z = 1.2$ km (Figure 4.2c). More differences to distinguish three cases are visible on CIGs. For the macromodel of too low velocity, the CIG displays an upward-curved event related to the incorrect model (Figure 4.3a). For the correct background velocity, the event is almost horizontal (Figure 4.3b), even if the edges (related to the spatial positions far away from the shot position) are slightly upward-curved due to the limited acquisition. If the macromodel velocity is too high, the corresponding event exhibits a downward curvature (Figure 4.3c). The images are constructed through the direct inversion. Therefore, the migration artifacts in CIGs are

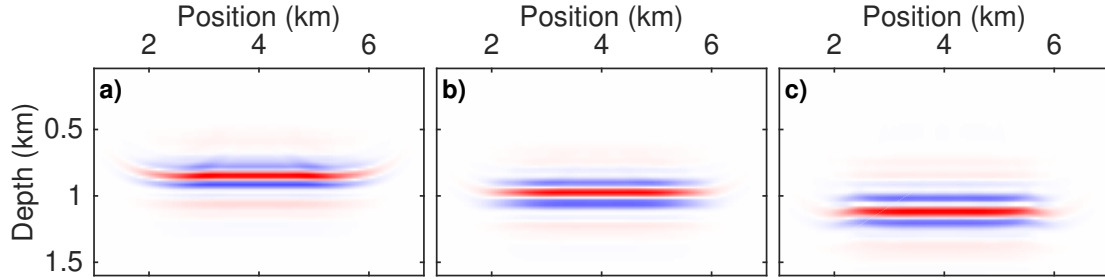


Figure 4.2 – Stacked image sections obtained with inversion operator B^\dagger for a single reflector at 0.6 km in different constant velocity models, respectively. Compared to the exact model, the velocity is 0.3 km/s lower for (a), equivalent for (b) and 0.3 km/s higher for (c), respectively.

highly attenuated such that we can distinguish the events associated to model inaccuracies from migration smiles. For the image-domain IVA, in practice, we compute the differential images by comparing adjacent traces in CIGs. For the cases of velocity too low and too high, the moveout residuals result from non-horizontal events in CIGs (Figure 4.3d and 4.3f). These residuals can be used to update the background velocity model. However, residuals are also visible in the CIG associated to the correct macro-model even if they are weaker than the cases of incorrect macromodels (Figure 4.3e). They are mainly caused by the stretching shape of the event in the CIG and the upward-curved event at the edges related to limited acquisition. We use this numerical experiment to explain the idea of image-domain common-shot IVA (Figure 4.1). The next step consists of transferring these image incoherency to scaling data for the introduction of the new approach defined in the data-domain.

The shot gather contains only reflection which exhibits a hyperbolic shape (indicated by the curved green line in Figure 4.4a). If we simply apply the modeling operator \mathcal{B}_0 to the differential images (Figures 4.3d, 4.3e and 4.3f), it is expected to output reflected events with different amplitudes at the same location as the observed data (Figure 4.1), because the same macromodel is used for the modeling and inversion parts. In practice, the scaling data related to different background velocity models indeed consist of reflected events of hyperbolic shapes with modified amplitudes: the reflected data for the two incorrect velocity cases are of reversed amplitudes and of almost the same magnitudes, and the correct velocity case corresponds to the smallest amplitudes (marked by green arrows in Figures 4.4b, 4.4c and 4.4d). In addition to the reflection, events similar to the direct arrivals are also present in the scaling data (marked by black arrows in Figures 4.4b, 4.4c and 4.4d). In the case of correct macromodel, they result from the residual artifacts exhibited in Figure 4.3e, and the reasons for those artifacts are already discussed in the previous paragraph. For the case of background model with too high velocity, the events marked by dashed green straight lines are caused by migration ar-

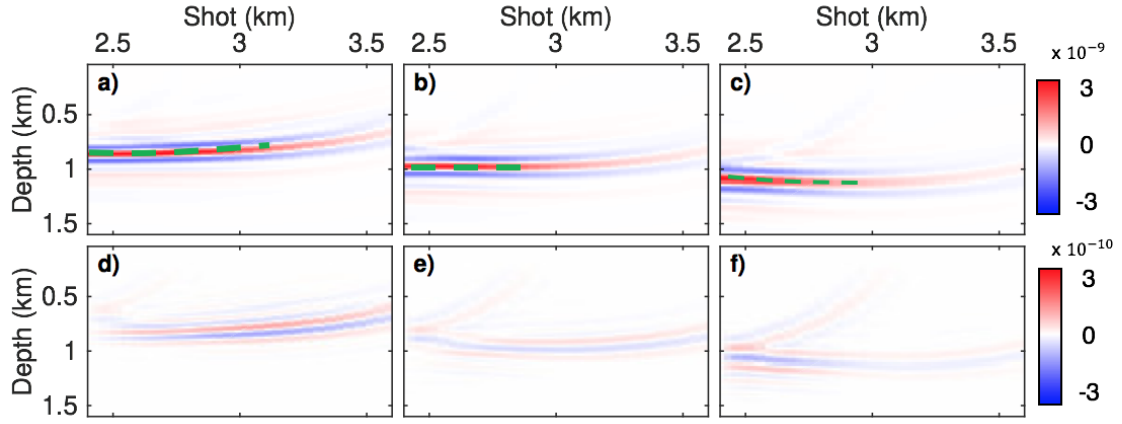


Figure 4.3 – CIGs for position $x = 2.4$ km (a–c) associated to images for Figure 4.2. (d–f) are the residuals between adjacent traces in CIGs (a–c). Green dashed lines illustrate the curvatures of events related to different model inaccuracies. (a–c) and (d–f) are displayed with the same scale, respectively.

tifacts which are not fully removed from the images. These artifacts are very weak but still visible in CIGs (e.g. around the left vertical axis of Figure 4.3f): they are upward-curved events but not related to inaccuracies of the model. In the next paragraph, we investigate the shape of macromodel gradient to check if the velocity update would be significantly biased.

In an ideal tomographic approach, we expect zero value gradient for the correct background velocity model, and only homogeneous positive (negative) values for a background velocity higher (lower) than the correct velocity. For image-domain IVA, we need an optimal $\alpha = 1$ in equation 4.7 to remove the spurious oscillations visible in the macromodel gradient around the interface positions (section 3.4.1). On the other hand, $\beta = 0$ is sufficient for data-domain IVA, in the sense that the associated gradients have the expected shapes after the application of Gaussian smoothing over half a wavelength. The definition of Gaussian smoothing filter is referred to equation 3.17. We build the data-domain IVA associated gradients with $\beta = 0$ (Figures 4.5a, 4.5b and 4.5c). These gradients are almost homogeneous above the interface. Around the reflector, spurious oscillations are still present even it is much weaker than for the case of image-domain IVA with $\alpha = 0$ (see Chapter 3). After the application of the Gaussian smoothing over half a wavelength, the value in gradient for correct macromodel is almost zero (Figure 4.5e), and the gradient for incorrect macromodel shows a pure tomographic update (Figures 4.5d and 4.5f).

A gradient obtained with data-domain IVA can be split into different subkernels. For a simple comparison, we decompose the gradients for both image-domain and data-

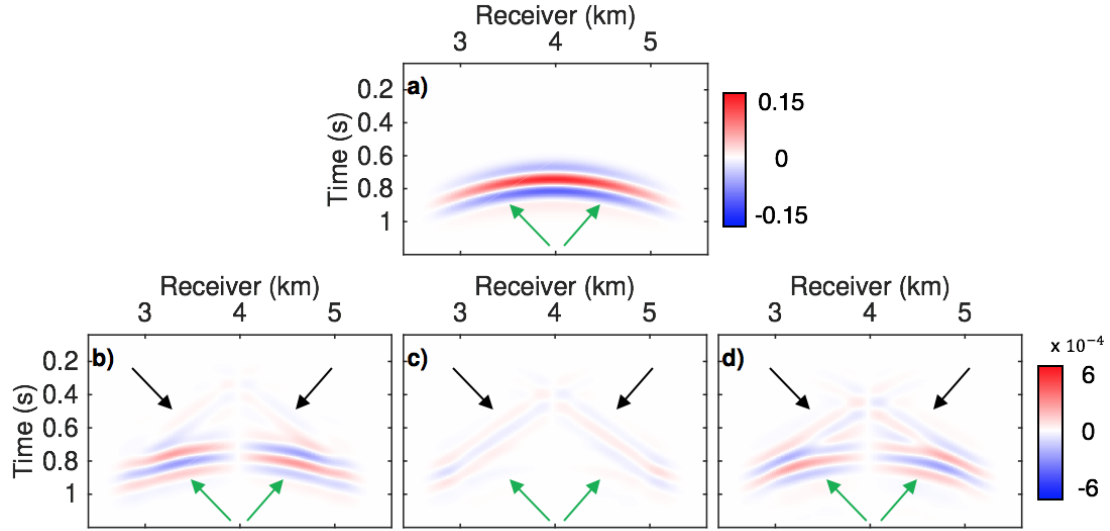


Figure 4.4 – Observed data for shot position $s_x = 4.0$ km (a), and scaling data for shot position $s_x = 4.0$ km (b–d) associated to image incoherency for (d–f) in Figure 4.3, respectively. Green and black arrows correspond to the direct arrival events and the reflected events, respectively.

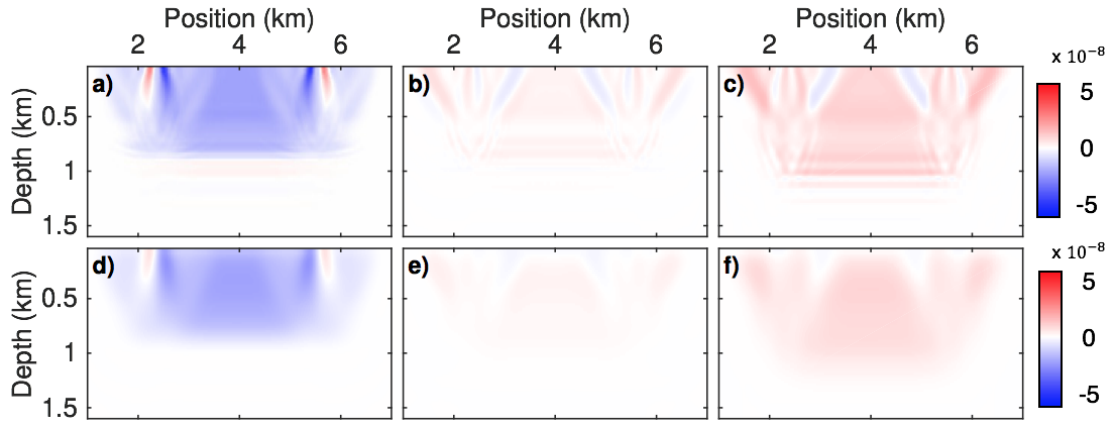


Figure 4.5 – Gradients obtained with data-domain IVA ($\beta = 0$) for a homogeneous macromodel with velocity too low (a, d), with correct velocity (b, e) and with velocity too high (c, f), respectively. We apply the Gaussian smoothing over half a wavelength to (a–c) to derive (d–f), respectively. All images are displayed with the same scale.

domain IVA according to equations 4.10 and 4.11. Note that the comparison between two approaches will be further detailed in section 4.3.3. We first consider a too low background velocity for example. For image-domain IVA, the contributions of $\tilde{\lambda}_s \star S_0$ and $\tilde{\lambda}_r \star R_0$ have tomographic behaviors (Figures 4.6b and 4.6c). The subkernel related to the oscillatory term is localized around the interface (Figure 4.6d), and can be reduced by the introduction of $\alpha = 1$, for which the contribution is localized at the same position but with opposite sign (Figure 4.6a). For data-domain IVA, the contributions of $\lambda_s \star S_0$ and $\lambda_r \star R_0$ (Figures 4.6f and 4.6g) have tomographic behaviors similar to the case of image-domain IVA above the interface. However, Figure 4.6f contains less oscillations than Figure 4.6b around the interface, whereas Figure 4.6g has stronger positive values around the interface compared to Figure 4.6c. In addition, data-domain IVA has one more tomographic subkernel related to $\lambda_d \star S_d$ (Figure 4.6e), which behaves similar to the contribution of $\lambda_r \star R_0$. For data-domain IVA, we do not introduce any additional component related β (i.e. $\beta = 0$) in the gradient to cancel the contribution of the oscillatory term (Figure 4.6h), as it would be attenuated after summing different contributions and the subsequent Gaussian smoothing can simply break the spurious oscillations in gradient. Back to the complete gradients, the oscillations are visible around the interface position for image-domain IVA ($\alpha = 0$) even after the application of the Gaussian smoothing (Figure 4.7a), whereas image-domain ($\alpha = 1$) and data-domain ($\beta = 0$) both provide more homogeneous update (Figures 4.7b and 4.7c). Similar results are observed for a background model with too high velocity (Figures 4.8 and 4.9).

We now explain why $\beta = 0$ is optimal for data-domain IVA. If the annihilator D_{s_x} is replaced by the identity operator I in equation 4.8, we expect $\mathcal{B}_0 \mathcal{B}_0^\dagger d^{obs} = d^{obs}$ (definition of the inverse) with $\beta = 0$, such that the gradient of modified \mathcal{J}_{data}^0 with respect to c_0 should be zero. In fact, this is only valid for the long wavelengths as the inversion is derived in an asymptotic sense. We change the annihilator D_{s_x} in equations 4.7 and 4.8 and compute the gradients of modified \mathcal{J}_{image}^0 , \mathcal{J}_{image}^1 and \mathcal{J}_{data}^0 with respect to correct c_0 . They all have small oscillations localized around the interface (Figures 4.10a, 4.10b and 4.10c). After the application of a Gaussian smoothing over half a wavelength, the values of macromodel gradients for image-domain IVA ($\alpha = 1$) and data-domain IVA ($\beta = 0$) are almost zero (Figures 4.10e and 4.10f), whereas it is not the case for image-domain IVA ($\alpha = 0$) (Figure 4.10d). It indicates that $\beta = 0$ is optimal for data-domain IVA: data-domain IVA ($\beta = 0$) is superior over image-domain IVA ($\alpha = 0$) and is equivalent to image-domain IVA ($\alpha = 1$).

We provide another argument why data-domain IVA ($\beta = 0$) can remove the spurious oscillations in the macromodel gradient. With $\alpha = 0$, the gradient obtained with image-domain IVA contains spurious oscillations. These oscillations are in fact the imprints of interfaces: the depth of interface changes when the background velocity model is updated. On the other hand, for data-domain IVA, the scaling data (i.e. d^β) can be seen as a reflectivity defined in the time-domain as for the MBTT case (Chavent et al., 1994; Plessix et al., 1995). A modification of the macromodel only leads to amplitude

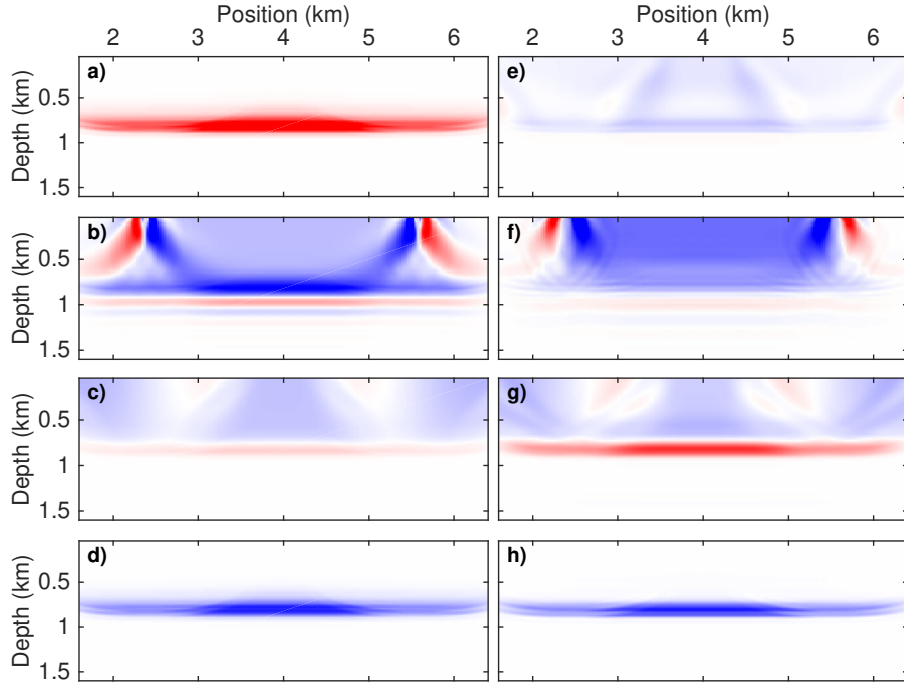


Figure 4.6 – Decomposition of macromodel gradients obtained with image-domain IVA or data-domain IVA for a homogeneous macromodel with velocity too low. (a) corresponds to the additional term related to $\alpha = 1$ for image-domain IVA. (b–d) correspond to $\tilde{\lambda}_s \star S_0$, $\tilde{\lambda}_r \star R_0$ and the oscillatory term in equation 4.10 for image-domain IVA ($\alpha = 0$), respectively. (e–h) correspond to $\lambda_d \star S_d$, $\lambda_s \star S_0$, $\lambda_r \star R_0$ and the oscillatory term in equation 4.11 for data-domain IVA ($\beta = 0$), respectively. (a–d) and (e–h) are displayed with same scale, respectively.

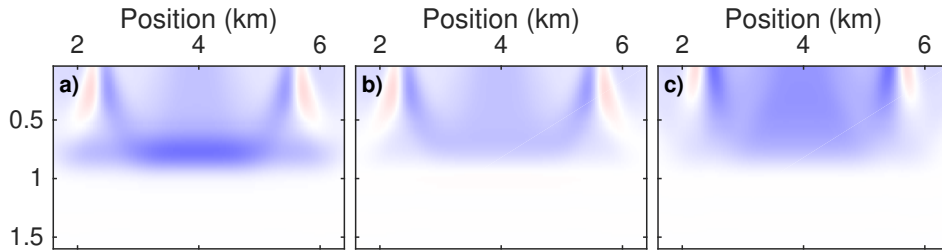


Figure 4.7 – Gradients obtained with image-domain IVA ($\alpha = 0$) (a), image-domain IVA ($\alpha = 0$) (b) and data-domain IVA ($\beta = 0$) (c) for a homogeneous macromodel with velocity too low, after a Gaussian smoothing over half a wavelength. (a) corresponds to the summation of Figures 4.6b–4.6d, (b) to the summation of Figures 4.6a–4.6d, and (c) to the summation of Figures 4.6e–4.6h.

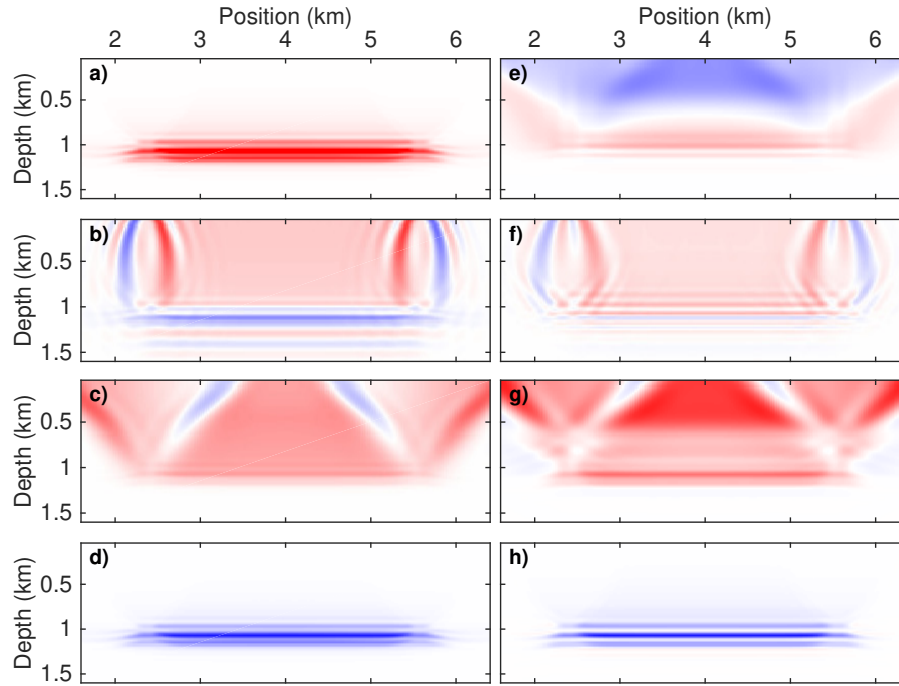


Figure 4.8 – Same as for Figure 4.6, but for a homogeneous macromodel with velocity too high.

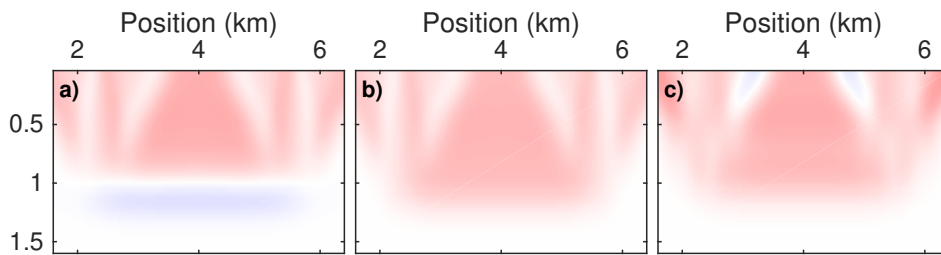


Figure 4.9 – Same as for Figure 4.7, but for a homogeneous macromodel with velocity too high.

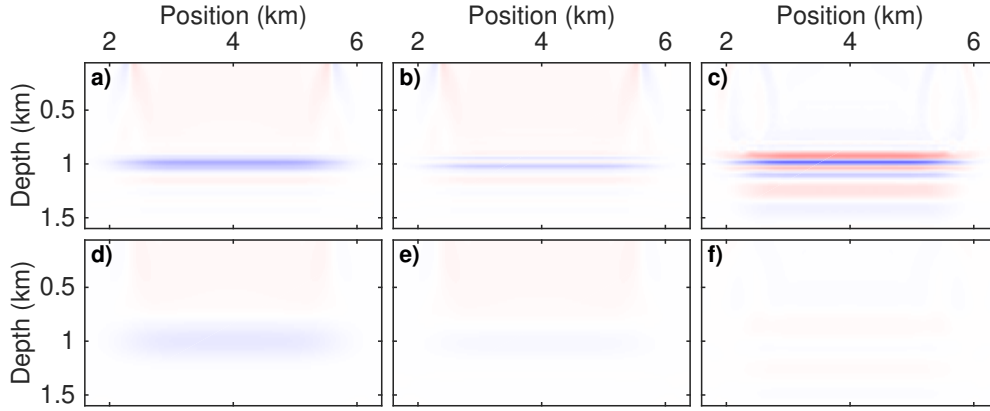


Figure 4.10 – Gradients obtained with image-domain IVA ($\alpha = 0$) (a), image-domain IVA ($\alpha = 0$) (b) and data-domain IVA ($\beta = 0$) (c) with respect to the correct macro velocity model. The annihilator in all methods are replaced by the identity factor. We apply the Gaussian smoothing over half a wavelength to (a–c) to derive (d–f), respectively. (a,b,d,e) and (c,f) are displayed with the same scale, respectively.

variations in this time-domain reflectivity.

the applications of \mathcal{B}_0 and \mathcal{B}_0^\dagger in equation 4.8 are at the same position, such that we could expect a macromodel gradient free of interface imprints. However, $\mathcal{B}_0 \mathcal{B}_0^\dagger d^{obs} = d^{obs}$ is in the asymptotic sense. Thus, only the long-wavelength part of the gradient reaches our expectation.

We conclude here that the data-domain IVA approach with $\beta = 0$ can produce the gradient in a tomographic mode after smoothing, at least for simple models with a continuous interface. The same β value is used for following tests.

Marmousi model

We apply the data-domain IVA scheme on the central part of the Marmousi model (Versteeg, 1994). The true background model is obtained after applying a Gaussian smoothing with a characteristic length of 60 m to the original model. Shots are located at depth $z = 0.04$ km. We use the same acquisition geometry (split spread) as for the single interface case, except with a larger maximum shot-receiver offset of 2.4 km for the Marmousi model.

Before the iterative optimization, we check the first gradient of data-domain IVA. The initial incorrect velocity models are a homogeneous model at 1.5 km/s (velocity too low) and a constant gradient model (velocity too high) starting from 2.0 km/s at the surface and 4.5 km/s at 2.0 km depth. The gradients of objective function with respect to macromodel exhibit many vertical strips behaving in a tomographic mode, but many oscillations are still visible (Figures 4.11a and 4.11b). After the application of a Gaus-

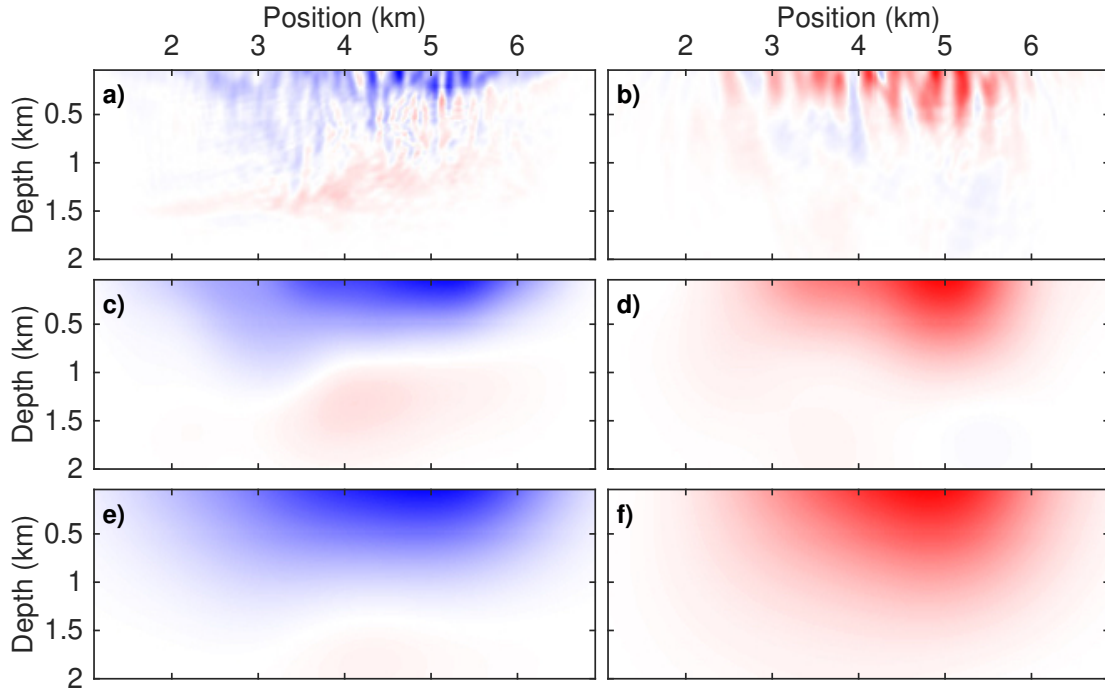


Figure 4.11 – Gradients obtained with data-domain IVA for the Marmousi model. The initial model is a 1.5 km/s homogeneous model for (a,c,e). The initial model is a constant gradient model (from shallow to deep part: 2.0-5.0 km/s) with velocity higher than the true velocity (b,d,f). We apply the Gaussian smoothing over half a wavelength to (a,b) for (c,d) and over one wavelength for (e,f), respectively. Blue, white and red represent negative, zero and positive values, respectively.

sian smoothing over half a wavelength, the gradients are much reliable (Figures 4.11c and 4.11d), but the values are not only negative for the background model of too low velocities (Figures 4.11c) and these positive values in the deeper part can probably bias the update after the application of a depth preconditioner to the gradient. Increasing the characteristic wavelength of the Gaussian smoothing to one wavelength, the gradients are finally more consistent, meaning that the results only exhibit positive (negative) values for too high (too low) velocity model (Figures 4.11e and 4.11f). It indicates that the gradient has a smooth behavior even for such model that consists of many discontinuous interfaces. In the case of complex structure, the Gaussian smoothing over more than half a wavelength may be needed.

We start the nested IVA optimization from a homogeneous initial model of 1.5 km/s. Optimization is carried out with a hybrid conjugate gradient algorithm (Touati-Ahmed and Storey, 1990). As a preconditioner, the gradient is multiplied by the squared depth value z^2 . 100 iterations are performed and the misfit curve presents the convergence after

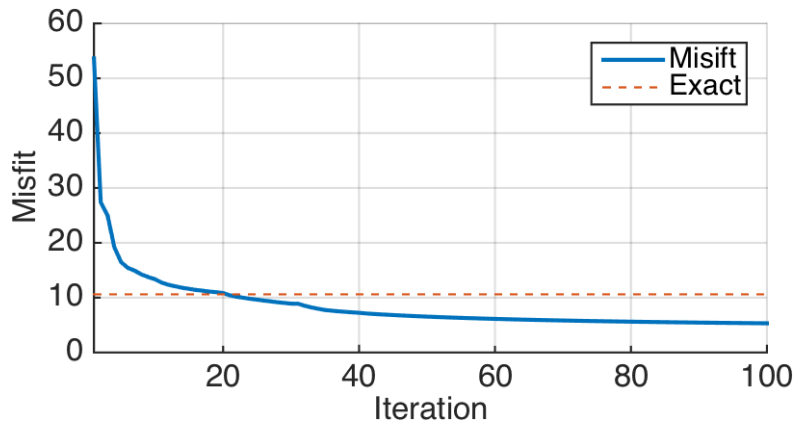


Figure 4.12 – Convergence curve for the data-domain IVA performed on the Marmousi model. Blue line represents the misfit value of objective function for the iteratively updated model and red dashed line represents the misfit value for the exact velocity model.

around 80 iterations (Figure 4.12). IVA converges much slower in the data-domain than in the image-domain. As mentioned in Chapter 3, the same test requires only around 25 iterations to converge. The reason of different convergence speeds will be explained in section 4.3.3.

The final value of the objective function goes beyond the value associated to the correct velocity, meaning that the objective function is not minimal even for the true velocity model. This was also observed in Hou and Symes (2016a). The effects directly link to the fact that the values of scaling data are not exactly zero even for the correct macromodel. The non-zero value of scaling data can result from kinematic ambiguity in the image (Nolan and Symes, 1996; Prucha et al., 1999; Stolk and Symes, 2004): in the presence of multipathing, the CIGs are not perfectly consistent even for a correct velocity model.

During the iterative update, the characteristic length for the Gaussian smoothing applied to the gradient is one wavelength at the very beginning and is reduced every 10 iterations. Eventually, we keep using a Gaussian smoothing over a quarter wavelength after 30 iterations. Note that this strategy of smoothing is different from the one applied to image-domain IVA. If the gradient is not smoothed any longer after several iterations, the optimization could be trapped into a local minimum and the estimated macro velocity model exhibits spurious oscillations. For example, Figure 4.13 shows the result of an updated model if the gradient is no more smoothed after several iterations: the information about the deeper part of the macro velocity model is not retrieved, and spurious oscillations are present everywhere. This approach is not necessary recommended as the short-wavelength information is introduced in the macromodel.

Different from the previous case, we keep applying in practice a Gaussian smoothing

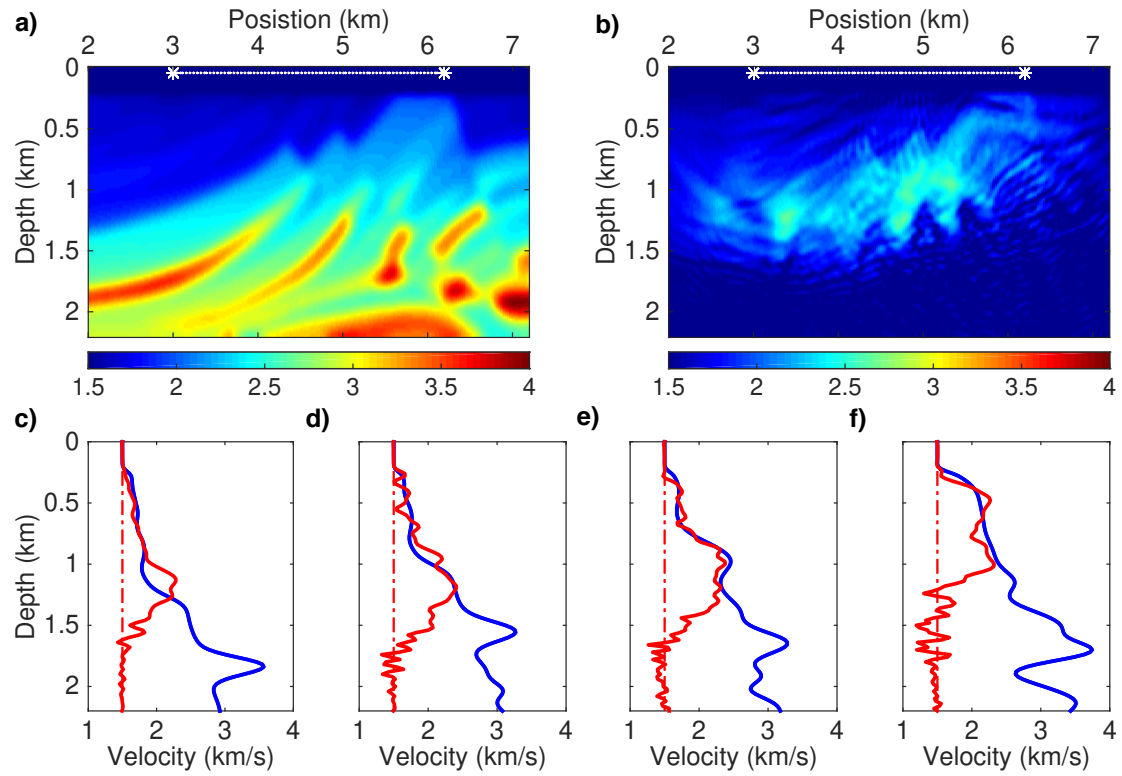


Figure 4.13 – Example illustrating the optimization converging towards a local minimum with an improper smoothing strategy. True Marmousi background model (a) and updated model after 100 data-domain IVA iterations (b). White stars and dashed line indicate the shot position extension. (c–f) are vertical velocity profiles at positions 2.6, 3.6, 4.6 and 5.6 km, respectively. Solid blue, dashed-dot red and solid red lines correspond to true, initial and updated models, respectively.

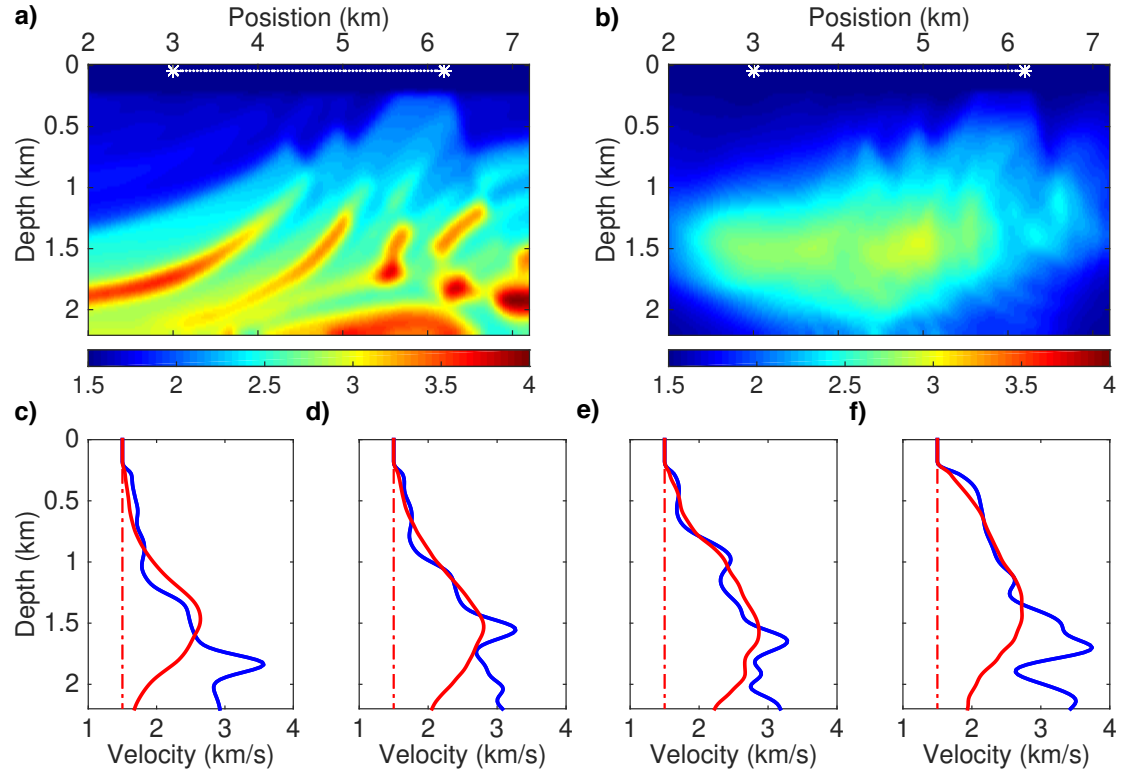


Figure 4.14 – True Marmousi background model (a) and updated model after 100 data-domain IVA iterations (b). White stars and dashed line indicate the shot position extension. (c–f) are vertical velocity profiles at positions 2.6, 3.6, 4.6 and 5.6 km, respectively. Solid blue, dashed-dot red and solid red lines correspond to true, initial and updated models, respectively.

over a quarter wavelength after 30 iterations. We compare the true background model to the final updated model after 100 nonlinear iterations (Figure 4.14a and 4.14b). The macro structure of the Marmousi model is well reconstructed, especially the central part that is well illuminated (Figures 4.14d and 4.14e). The structures at edges are less appropriately recovered due to the limited acquisition geometry (Figures 4.14c and 4.14f).

To evaluate the quality of the final result, we compute the reflectivity models according to equation 4.4 for the exact, initial and estimated macromodels, respectively (Figure 4.15). The final retrieved background model significantly improves the quality of the stacked image (Figure 4.15c), close to the result obtained for the true model (Figure 4.15a). Cautious examination on the vertical profiles indicates that the result of data-domain IVA is reasonable: the phases and amplitudes of reflectors are similar for the images obtained in the final and in the true background velocity models, especially in the central areas (Figure 4.16). The profiles close to edges show small distortions since

the model is not fully illuminated for these deep regions. CIGs are extracted from the image volume to evaluate the consistency of the included events over different shot positions. The final result produces coherent CIGs and has similar behavior to the true model (Figure 4.17). Note that a mask on the inverted images and the macromodel gradient is applied in the first 0.2 km to prevent from updating the shallow part, as an additional preconditioner. The limit of the mask (dashed line) is illustrated on Figure 4.17. These indicate the consistency between the data-domain and image-domain IVA approaches: the final results improve the quality of images despite objective functions defined in different domains. A more detailed comparison will be provided in section 4.3.3.

After a successful data-domain IVA optimization procedure, the events in CIGs are more horizontal. Therefore, the image incoherency should be significantly reduced, leading to scaling data with weaker amplitudes. The differential images are computed by comparing the neighboring traces in CIGs (Figure 4.18). The scaling data are modeled in the estimated macromodels using the image-domain residuals as the reflectivity models. The image incoherency associated to data-domain IVA results (Figure 4.18c) are significantly reduced compared to the case of the starting model (Figure 4.18b), and is even smaller than for the correct velocity model in some areas (Figure 4.18a). The associated scaling data transfer these residuals from image-domain to data-domain. Theoretically, we expect that the scaling data related to the correct velocity model is minimal. In practice, some spurious direct arrivals (Figure 4.19a) are still visible even for the case of true macromodel. These direct arrivals also exist in the scaling data for the starting model (Figure 4.19b), indicated by black arrows. The reflected events in scaling data carry the main information about the inaccuracies of the macromodel, whereas the direct arrivals correspond to residual artifacts in CIGs close to the same position (Figure 4.18a). After the nonlinear optimization of data-domain IVA, both the energy of the direct arrivals and reflections in scaling data are minimized (Figure 4.19c), as it is explicitly defined in the objective function. This explains the reason why the misfit for the data-domain IVA result is even lower than the value related to the exact macromodel in Figure 3.18. Despite the artifacts, numerical results have shown that the large-scale structures of the Marmousi model are properly retrieved by minimizing the new objective function defined in the data-domain.

We shoot rays in the correct and updated Marmousi macromodels (Figures 4.20). The correct macromodel contains clear triplicated wavefields (Figures 4.20a), whereas it is not the case for the macromodel obtained with IVA (Figures 4.20b). We refer to Chapter 3 for more discussions about triplications.

Finally, the model from Figure 4.14b is used as the starting model for FWI to further improve the dynamic details. To avoid the cycle skipping effects for FWI, low-frequency data should be recorded or the starting model should be close enough to the correct one (Virieux and Operto, 2009). In this test, shots are located every 0.08 km ranging from 1 to 6.2 km, and receivers everywhere at the surface ($z = 0.04$ km). Frequencies below 2 Hz are not recorded in the observed data. Therefore, the accuracy of the starting model

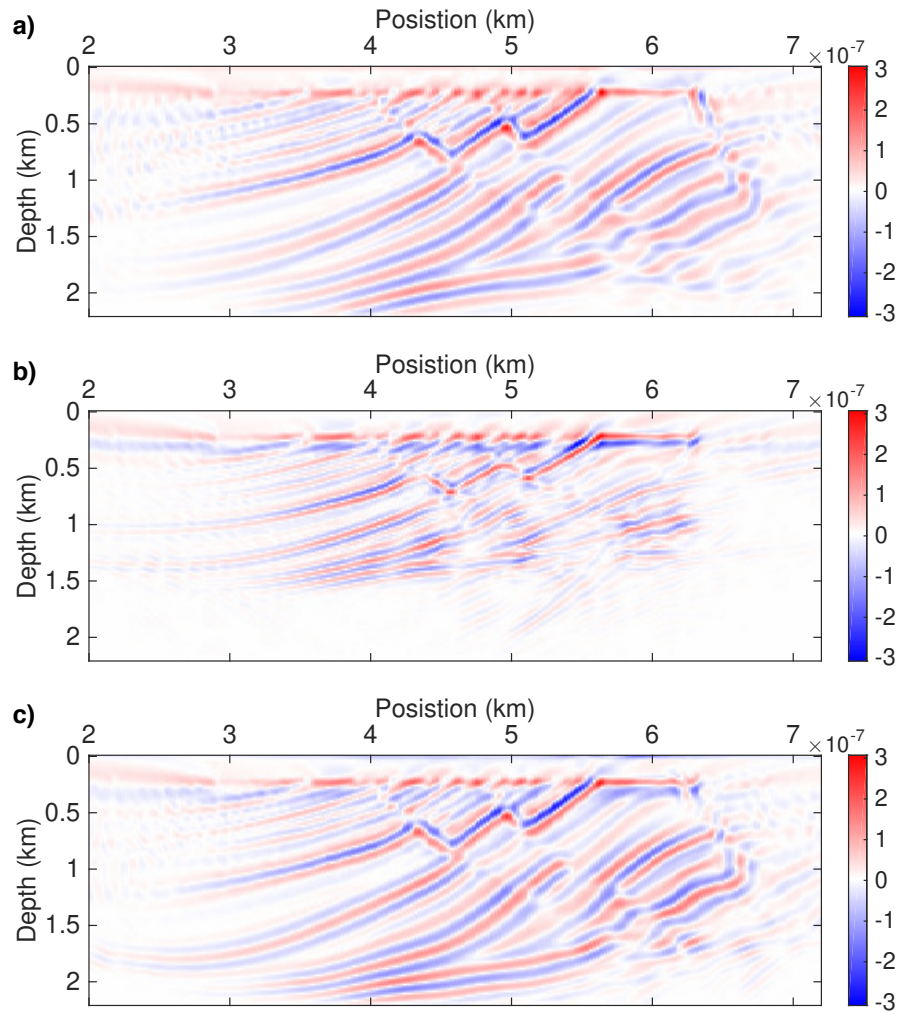


Figure 4.15 – Stacked inverted images associated to (a) true (Figure 4.14a), (b) initial and (c) updated (Figure 4.14b) models, respectively.

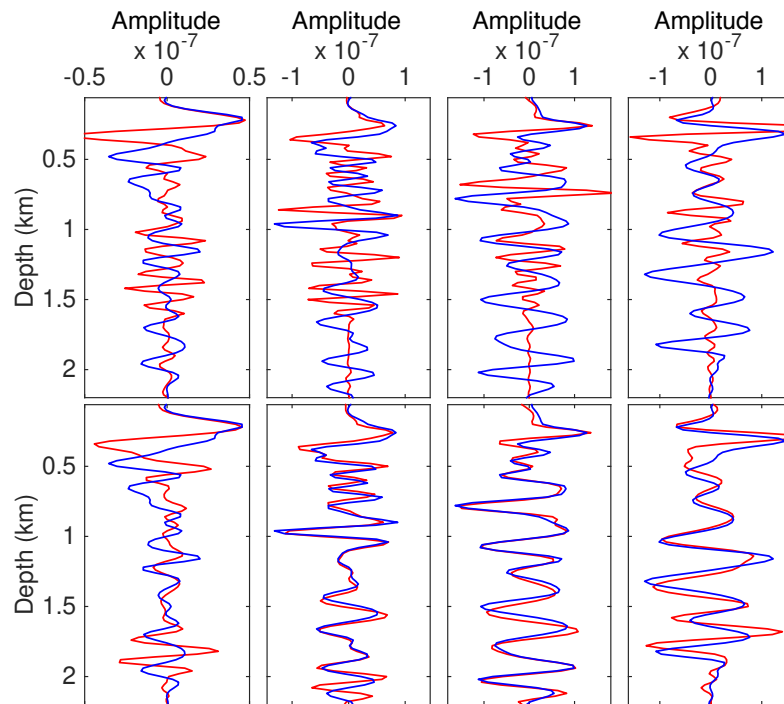


Figure 4.16 – Profiles of stacked inverted images associated with true, initial and updated (Figure 4.14b) models. From left to right, columns are associated to positions 2.6, 3.6, 4.6 and 5.6 km, respectively. Blue lines refer to image profiles for true model. The red lines represent profiles for initial model in the top rows and for updated model in the bottom rows.

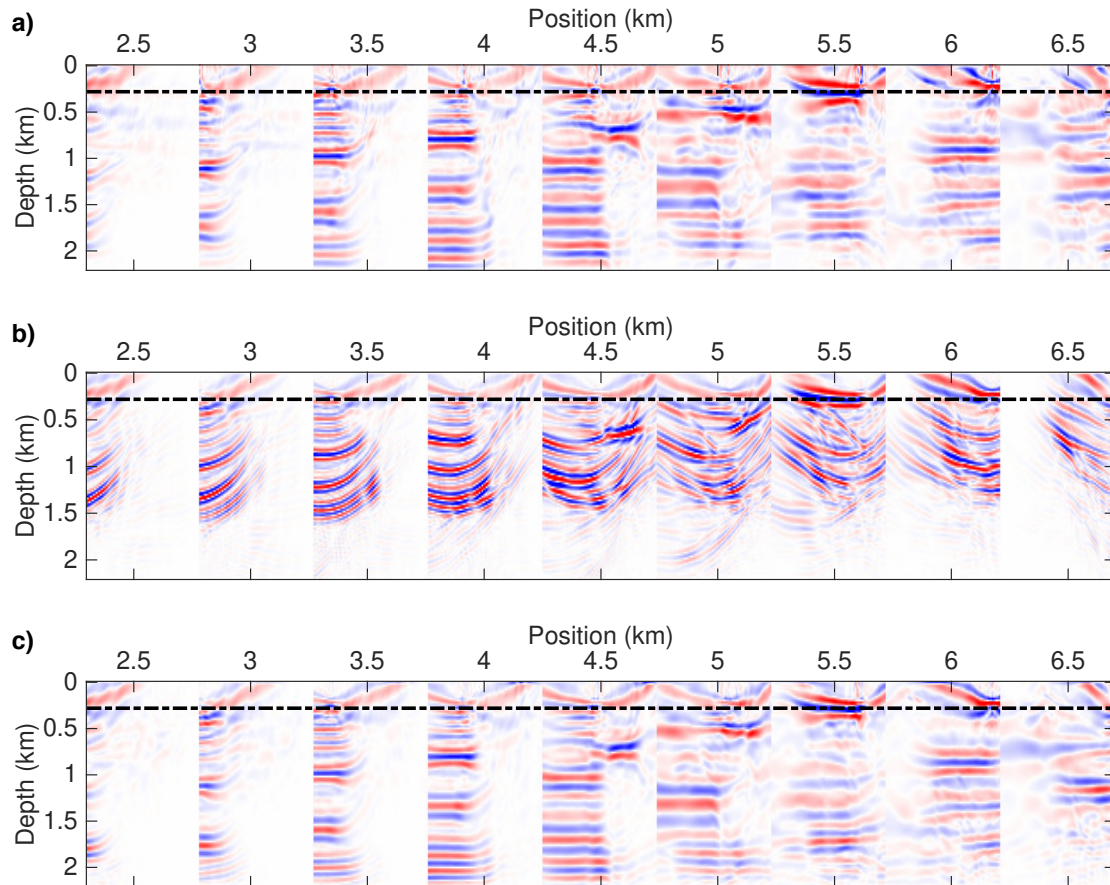


Figure 4.17 – CIGs associated to (a) the true Marmousi model, (b) the initial homogeneous model and (c) the updated model (Figure 4.14b) for position x , ranging from 2.5 to 6.5 km, every 0.5 km. We apply a taper on the image above the depth indicated by dashed black line, to exclude the associated contribution to our objective.

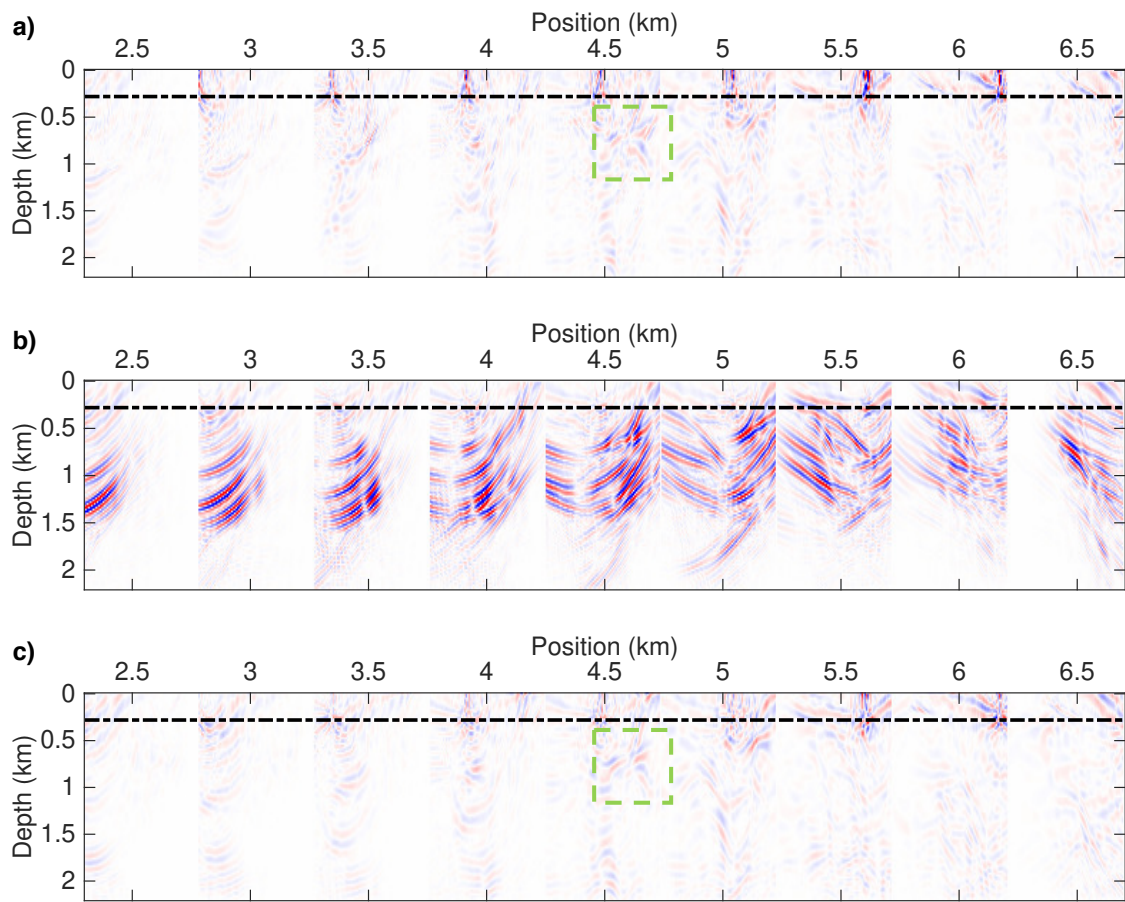


Figure 4.18 – The same as Figure 4.17, but for the image residuals computed by comparing adjacent traces in CIGs. All images are displayed with the same scale.

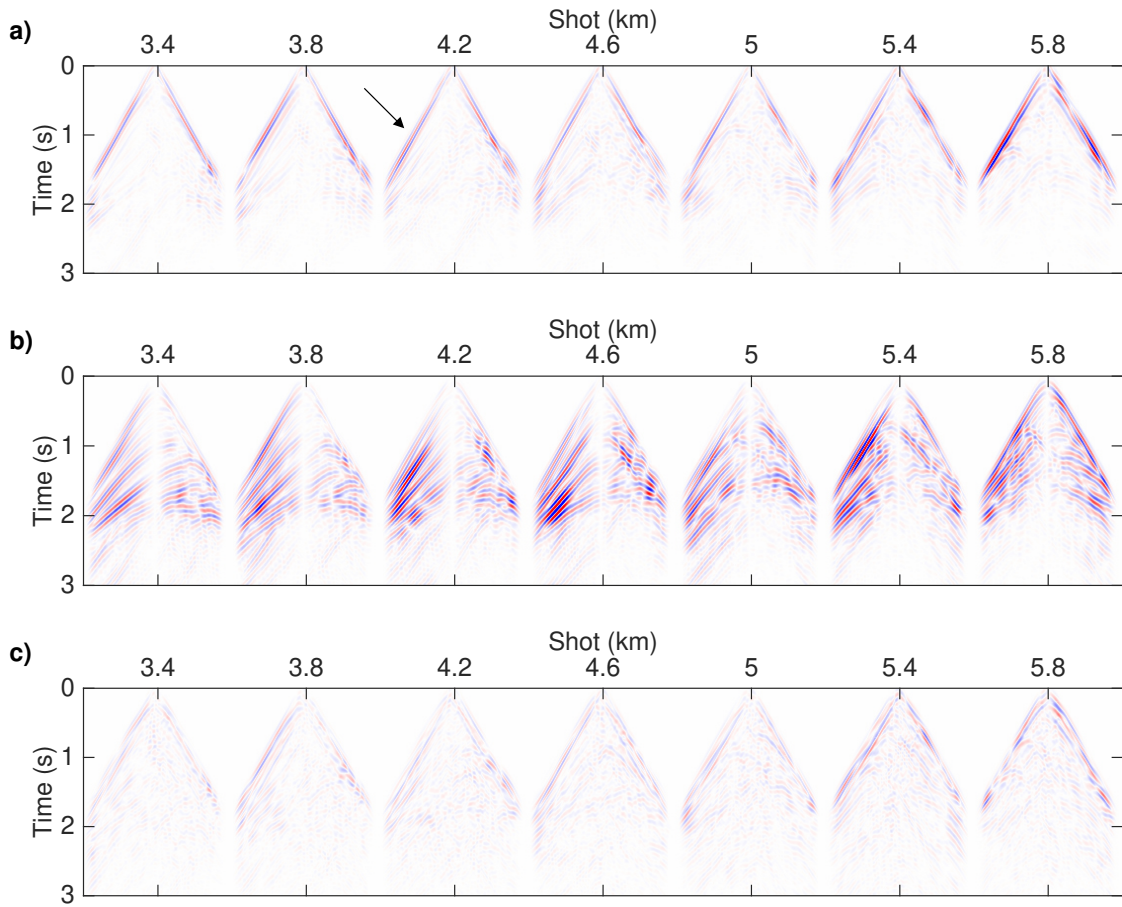


Figure 4.19 – The same as Figure 4.17, but for the time-domain scaling data sets, computed for shot position s_x , ranging from 3.4 to 5.8 km, every 0.4 km. All images are displayed with the same scale.

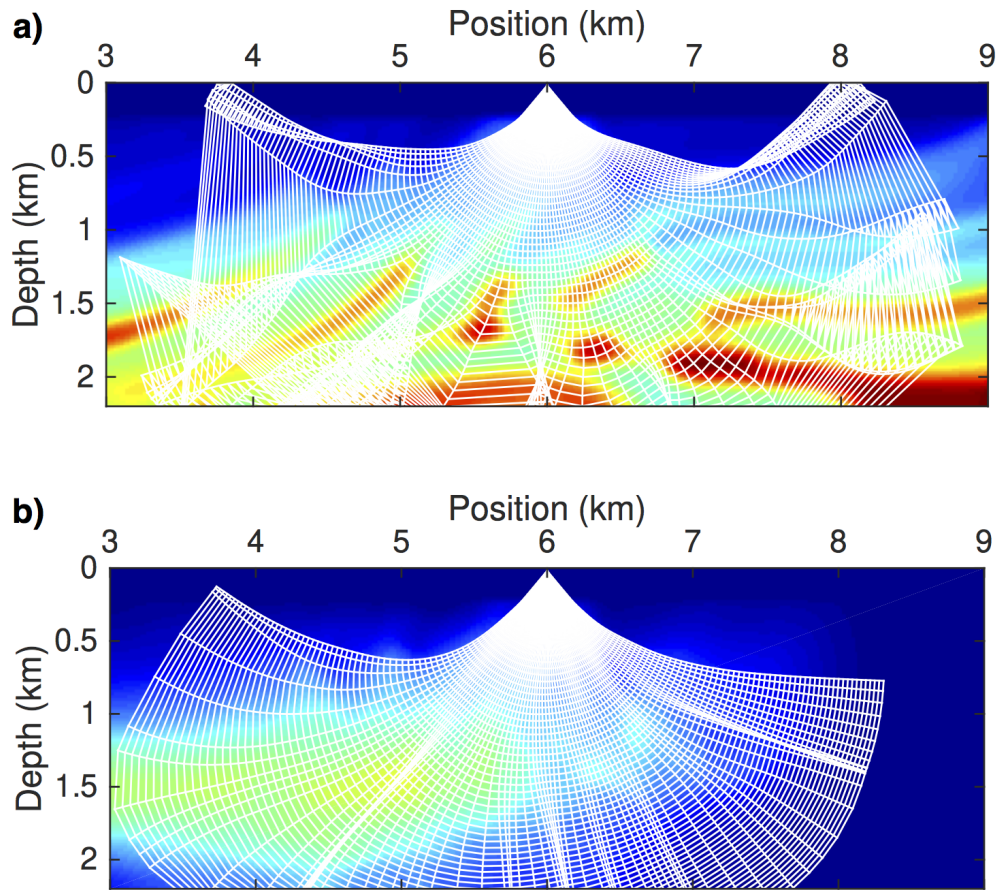


Figure 4.20 – Rays and wavefronts calculated from a source at surface with $s_x = 6.0$ km in (a) true and (b) inverted (Figure 4.14b) Marmousi macromodels, respectively. All images are displayed with the same scale.

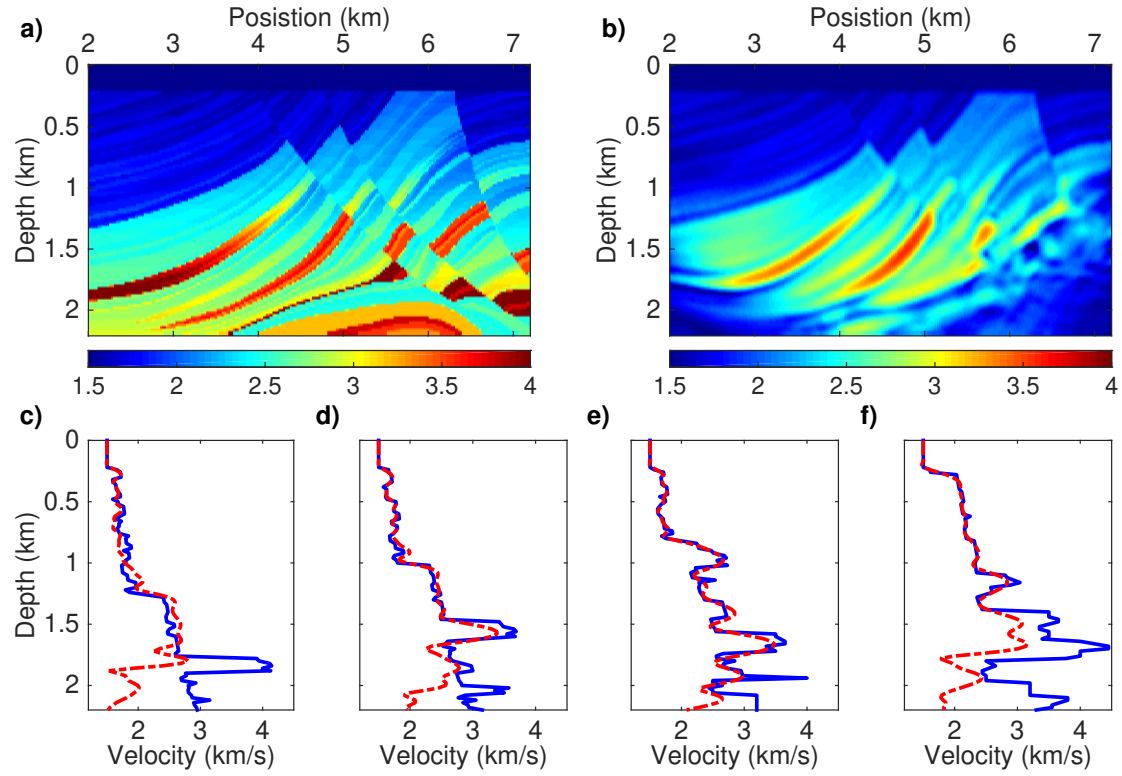


Figure 4.21 – True Marmousi full model (a) and updated model (b) after 70 FWI iterations starting from the final data-domain IVA result (Figure 4.14b). (c–f) are vertical velocity profiles at positions 2.6, 3.6, 4.6 and 5.6 km, respectively. Solid blue and dashed-dot red lines represent true and updated models, respectively.

is the key element to get rid of cycle skipping effects. The structures of Marmousi model are well retrieved after 70 nonlinear iterations of FWI (Figure 4.21). Cycle skipping effects are hardly observed in the area of which IVA supplies a reasonable macromodel. Note that we compare the FWI results with the true complete model c rather than background model c_0 in this case.

Summary

We conclude from these different numerical tests that the new data-domain approach (equation 4.8) is robust. In contrary to incorporating c_0^α , with $\alpha = 1$, to the objective function 4.7 for the image-domain approach, we only use $\beta = 0$ for data-domain IVA. The new approach produces a macromodel gradient that does not suffer from the imprints of reflectivity, at least after the Gaussian smoothing over half a wavelength. For a more complex model, a larger smoothing may be needed. IVA is capable of dealing

with complex structures, at least for the Marmousi model. The macromodel gradient should always be smoothed for a robust iterative updating. Minimizing the norm of scaling data increases the coherency of images at the same time. The final result of data-domain IVA provides a well-estimated starting model for subsequent FWI, leading to non cycle-skipped results, at least for well-illuminated regions.

4.3 Comparisons between image and data domains

For the methods relying on the Born approximation, both data-domain and image-domain aim at determining an optimal background model. The objective of this section is to better understand the possible links between different approaches. We first establish the link between data-domain IVA and DWI, under the condition that the reflectivity model is properly estimated. More generally, we investigate, in a generic way, the relationships between data fitting principle and image coherency criteria. Finally, we compare the numerical results between image-domain and data-domain IVA methods.

4.3.1 Equivalence between data-domain IVA and DWI

We first compare data-domain IVA to Differential Waveform Inversion (DWI) (Chauris and Plessix, 2012). Two approaches are both developed from the DSO approach. Here, we consider only two shot gathers at positions s_k and s_{k+1} with $k \in [1, N_s]$, where N_s is the total shot numbers. In the DWI approach, one first determines the reflectivity image for a single shot position in a given macromodel, and then predict data at the next shot position. We build a reflectivity image for s_k in a given background velocity model c_0

$$\xi(\mathbf{x}, s_k) = \xi[s_k] = \mathcal{B}_0^\dagger[s_k] d^{obs}[s_k]. \quad (4.12)$$

Then, we calculate the modeled (demigrated) data at the next shot position using the reflectivity retrieved from the previous shot gather

$$d_{dwi}^{cal}[s_{k+1}] = \mathcal{B}_0[s_{k+1}] \xi[s_k]. \quad (4.13)$$

Note that \mathcal{B}_0 and \mathcal{B}_0^\dagger are referred to equations 4.1 and 4.4, respectively. To assess the quality of the estimated background model, we evaluate the misfit between modeled and observed data

$$J_{dwi}[c_0] = \frac{1}{2} \left\| d^{obs}[s_{k+1}] - d_{dwi}^{cal}[s_{k+1}] \right\|^2. \quad (4.14)$$

If the background model is correct, the reflectivity images retrieved from two shot gathers should be consistent such that the data at s_{k+1} can be predicted from $\xi[s_k]$. This

is equivalent to the image coherency principle. In the asymptotic sense, the inverse operator is defined such that

$$\mathcal{B}_0[s_{k+1}]\xi[s_k] = \mathcal{B}_0[s_{k+1}]\mathcal{B}_0^\dagger[s_{k+1}]d^{obs}[s_{k+1}] \simeq d^{obs}[s_{k+1}] \quad (4.15)$$

and then we substitute equations 4.12, 4.13 and 4.15 into equation 4.14, reading

$$\begin{aligned} J_{dwi}[c_0] &\simeq \frac{1}{2} \left\| \mathcal{B}_0[s_{k+1}](\xi[s_{k+1}] - \xi[s_k]) \right\|^2 \\ &= \frac{1}{2} \left\| \mathcal{B}_0[s_{k+1}]D_{s_k}\xi[s_k]\delta s \right\|^2, \end{aligned} \quad (4.16)$$

where δs represents the shot interval. This objective function is equivalent to equation 4.8. Therefore, the two approaches are proved to be consistent if a true amplitude migration scheme is applied, in the sense that \mathcal{B}_0^\dagger is indeed an inverse of \mathcal{B}_0 , and it is indeed the case for our direct inverse formula (equation 4.4). Note that the inverse formula is derived under the first-order Born approximation, meaning that the multi-scattered wavefields are not considered. To cope with multiples, one can replace direct inverse by iterative migration (Chauris and Plessix, 2013; Cocher et al., 2017b), even if iterative scheme is more computational intensive. In the next section, we generalize the results for different migration schemes.

4.3.2 Data fitting versus image coherency

DWI is inspired by common-shot DSO (Symes and Carazzone, 1991; Symes and Kern, 1994; Chauris and Plessix, 2012). Its idea can be extended to other migration schemes, to be explained later, formulating different reflection-based waveform inversion approaches. We establish the link between the DWI method, a data fitting technique defined in the data-domain, and the DSO approach, a MVA strategy defined in the image-domain. The comparison is performed for different migration schemes (common-shot, common-offset, time-delay, common-angle and subsurface-offset cases). True-amplitude migration is the key element to link two families.

Common-shot migration scheme

We express the calculated shot gather $d_{dwi}^{cal}[s_{k+1}]$ for the reflectivity $\xi[s_k] = \xi(\mathbf{x}, s_k)$ with a first-order approximation

$$d_{\text{dwi}}^{\text{cal}}(s_{k+1}, \mathbf{r}, \omega) = (i\omega)^2 \Omega(\omega) \int d\mathbf{x} G_0(s_{k+1}, \mathbf{x}, \omega) \xi[s_k] G_0(\mathbf{r}, \mathbf{x}, \omega) \quad (4.17)$$

$$\begin{aligned} &\simeq (i\omega)^2 \Omega(\omega) \int d\mathbf{x} G_0(s_k, \mathbf{x}, \omega) \xi[s_k] G_0(\mathbf{r}, \mathbf{x}, \omega) \\ &\quad + (i\omega)^2 \Omega(\omega) \int d\mathbf{x} \frac{\partial G_0(s_k, \mathbf{x}, \omega)}{\partial s_k} \xi[s_k] G_0(\mathbf{r}, \mathbf{x}, \omega) \delta s. \end{aligned} \quad (4.18)$$

The first term on the right side of the equation 4.18 corresponds to $d^{\text{obs}}[s_k]$. We replace the second term using the sum rule in differentiation

$$\begin{aligned} d_{\text{dwi}}^{\text{cal}}[s_{k+1}] &\simeq d^{\text{obs}}[s_k] + (i\omega)^2 \Omega(\omega) \frac{\partial}{\partial s_k} \int d\mathbf{x} G_0(s_k, \mathbf{x}, \omega) G_0(\mathbf{r}, \mathbf{x}, \omega) \xi[s_k] \delta s \\ &\quad - (i\omega)^2 \Omega(\omega) \int d\mathbf{x} G_0(s_k, \mathbf{x}, \omega) G_0(\mathbf{r}, \mathbf{x}, \omega) \frac{\partial \xi[s_k]}{\partial s_k} \delta s. \end{aligned} \quad (4.19)$$

The second term on the right side of the equation is recognized as the derivative of the observed data

$$\begin{aligned} d_{\text{dwi}}^{\text{cal}}[s_{k+1}] &\simeq d^{\text{obs}}[s_k] + \frac{\partial d^{\text{obs}}[s_k]}{\partial s_k} \delta s \\ &\quad - (i\omega)^2 \Omega(\omega) \int d\mathbf{x} G_0(s_k, \mathbf{x}, \omega) G_0(\mathbf{r}, \mathbf{x}, \omega) \frac{\partial \xi[s_k]}{\partial s_k} \delta s. \end{aligned} \quad (4.20)$$

The first two terms on the right side read the first-order approximation of observed data at shot position s_{k+1} . We thus finally have

$$d^{\text{obs}}[s_{k+1}] - d_{\text{dwi}}^{\text{cal}}[s_{k+1}] = (i\omega)^2 \Omega(\omega) \int d\mathbf{x} G_0(s_k, \mathbf{x}, \omega) \frac{\partial \xi[s_k]}{\partial s_k} G_0(\mathbf{r}, \mathbf{x}, \omega) \delta s. \quad (4.21)$$

Minimizing the difference between the calculated data at shot position s_{k+1} and for $\xi[s_{k+1}]$ and the observed data at s_{k+1} is the same as minimizing the derivative of the reflectivity $\xi[s_k]$ with respect to s_k (for common-shot DSO, we refer to [Symes and Kern, 1994](#)). Thus the DWI scheme is equivalent to the common-shot DSO approach if they both obtain $\xi(\mathbf{x}, s_k)$ after true-amplitude migration (iterative migration or direct inverse). Note that the reflectivity section results from a first iteration of iterative migration in the conventional DSO approach. The first method is defined in the data-domain, whereas the second is defined in the model space. For equation 4.21, the time-domain data misfit is minimal when amplitudes are correctly predicted – this is a reason to explain why true-amplitude migration is needed in the image-domain even if the absolute values of are not required.

Common-offset migration scheme

We now consider the common-offset migration scheme. For the reflectivity section $\xi[\mathbf{x}_h] = \xi(\mathbf{x}_h, \mathbf{x})$, a function of \mathbf{x} and a shot-receiver offset \mathbf{x}_h . Following the same strategy for common-shot migration, the minimization of the data misfit $d^{obs}[x_h^{k+1}] - d_{dwi}^{cal}[x_h^{k+1}]$ is equivalent to the minimization of the derivative of $\xi[x_h^k]$ with respect to x_h^k (for common-offset DSO, we refer to [Chauris and Noble, 2001](#); [Mulder and ten Kroode, 2002](#)). This common-offset DWI scheme behaves in a tomographic similarly to the conventional RWI ([Xu et al., 2012](#); [Wu and Alkhalifah, 2015](#); [Brossier et al., 2015](#)). However, the two approaches have some major differences: (1) DWI is formulated as a nested optimization to update δc and c_0 , whereas RWI assumes two parameters are independent, at least for the approach of [Xu et al. \(2012\)](#); (2) DWI predicts the a single common-offset gather using the perturbation model constructed from the previous common-offset gather, whereas RWI builds δc model from data with a limited range of shot-receiver offsets (e.g. $x_h \in (0, h_0)$, h_0 is the value of the maximum offset) and then predicts data with larger range of offsets (e.g. $x_h \in (0, h_1)$, h_1 is the value of the maximum offset and $h_1 \ll h_0$).

Time-delay migration scheme

We extend the previous results to time-delay migration scheme. For the reflectivity section $\xi[\tau] = \xi(\mathbf{x}, \tau)$, a function of \mathbf{x} and the time-delay τ , the data d is given by

$$d(\mathbf{s}, \mathbf{r}, \omega) = (i\omega)^2 \Omega(\omega) \iint d\mathbf{x} d\tau G_0(\mathbf{s}, \mathbf{x}, \omega) G_0(\mathbf{r}, \mathbf{x}, \omega) \xi[\tau] e^{-i\omega\tau} \quad (4.22)$$

$$= (i\omega)^2 \Omega(\omega) \int d\mathbf{x} G_0(\mathbf{s}, \mathbf{x}, \omega) G_0(\mathbf{r}, \mathbf{x}, \omega) \hat{\xi}_\tau(\mathbf{x}, \omega) \quad (4.23)$$

The variable $\hat{\xi}_\tau[\omega]$ is the Fourier transform of $\xi[\tau]$ over τ . Note that the usual equation for time-delay expansion expresses $\xi[\tau]$ as a function of the input data ([Sava and Fomel, 2006](#)). Equation 4.22 is the expression of extended modeling proposed by [Symes \(2008\)](#); [Biondi and Almomin \(2014\)](#), being the adjoint version of the other.

Following the same strategy as for common-shot migration, we retrieve reflectivity section $\xi[\omega]$ and then define the modeled data for frequency $\omega + \delta\omega$ here

$$d_{dwi}^{cal}(\mathbf{s}, \mathbf{r}, \omega + \delta\omega) = (i\omega)^2 \Omega(\omega) \int d\mathbf{x} G_0(\mathbf{s}, \mathbf{x}, \omega) \hat{\xi}[\omega] G_0(\mathbf{r}, \mathbf{x}, \omega), \quad (4.24)$$

and the differences between the modeled and observed data for frequency $\omega + \delta\omega$ reads

$$\begin{aligned} d^{obs}[\omega + \delta\omega] - d_{dwi}^{cal}[\omega + \delta\omega] &\simeq (i\omega)^2 \Omega(\omega) \\ &\times \int d\mathbf{x} G_0(\mathbf{s}, \mathbf{x}, \omega) \frac{\partial \hat{\xi}[\omega]}{\partial \omega} G_0(\mathbf{r}, \mathbf{x}, \omega) \end{aligned} \quad (4.25)$$

An objective function measuring data misfit is equivalent to the one assessing the image coherency through the derivative of $\hat{\xi}_\tau(\mathbf{x}, \omega)$ with respect to ω , or alternatively through the norm of $i\tau\xi(\mathbf{x}, \tau)$, as a derivative in the Fourier domain can be expressed as a multiplication in the original space (for time-delay MVA, we refer to [Sava and Fomel, 2006](#)). This provides an key explanation for the introduction of a multiplication by τ in the objective function as an annihilator.

Subsurface-offset migration scheme

The extension for the depth-offset migration case is similar to the previous cases. Once more, we express the data as a function of the reflectivity ξ depending on the image point \mathbf{x} and on the depth-offset \mathbf{h}

$$d(\mathbf{s}, \mathbf{r}, \omega) = (i\omega)^2 \Omega(\omega) \iint d\mathbf{x} d\mathbf{h} G_0(\mathbf{s}, \mathbf{x}-\mathbf{h}, \omega) G_0(\mathbf{r}, \mathbf{x}+\mathbf{h}, \omega) \xi[\mathbf{h}] \quad (4.26)$$

$$= (i\omega)^2 \Omega(\omega) \int d\mathbf{x} G_0(\mathbf{s}, \mathbf{x}, \omega) G_0(\mathbf{r}, \mathbf{x}, \omega) \hat{\xi}_\mathbf{h}(\mathbf{x}, \mathbf{s}, \mathbf{r}, \omega), \quad (4.27)$$

where

$$\hat{\xi}_\mathbf{h}(\mathbf{x}, \mathbf{s}, \mathbf{r}, \omega) = \int d\mathbf{h} \frac{G_0(\mathbf{s}, \mathbf{x}-\mathbf{h}, \omega) G_0^*(\mathbf{s}, \mathbf{x}, \omega) G_0(\mathbf{r}, \mathbf{x}-\mathbf{h}, \omega) G_0^*(\mathbf{r}, \mathbf{x}, \omega)}{|G_0(\mathbf{s}, \mathbf{x}, \omega)|^2 |G_0(\mathbf{r}, \mathbf{x}, \omega)|^2} \xi[\mathbf{h}] \quad (4.28)$$

$$\simeq \int d\mathbf{h} e^{i\mathbf{k} \cdot \mathbf{h}} \xi[\mathbf{h}] = \hat{\xi}[\mathbf{k}] \quad (4.29)$$

with $\mathbf{k} = \omega \left(\frac{\partial \tau_0(\mathbf{s}, \mathbf{x})}{\partial \mathbf{x}} + \frac{\partial \tau_0(\mathbf{r}, \mathbf{x})}{\partial \mathbf{x}} \right)$. Under the high frequency approximation and up to smooth term, the integrand in equation 4.28 can be simplified to a Fourier transform in the spatial-domain ([ten Kroode et al., 1998](#)). Once again, the modified demigration reads

$$d_{\text{dwi}}^{\text{cal}}[\mathbf{k} + \delta k] = (i\omega)^2 \Omega(\omega) \int d\mathbf{x} G_0(\mathbf{s}, \mathbf{x}, \omega) \hat{\xi}[\mathbf{k}] G_0(\mathbf{r}, \mathbf{x}, \omega), \quad (4.30)$$

and the differences between the modeled and observed data for wavenumber $\mathbf{k} + \delta k$ are

$$\begin{aligned} d^{\text{obs}}[\mathbf{k} + \delta k] - d_{\text{dwi}}^{\text{cal}}[\mathbf{k} + \delta k] &\simeq (i\omega)^2 \Omega(\omega) \\ &\times \int d\mathbf{x} G_0(\mathbf{s}, \mathbf{x}, \omega) \frac{\partial \hat{\xi}[\mathbf{k}]}{\partial \mathbf{k}} G_0(\mathbf{r}, \mathbf{x}, \omega) \end{aligned} \quad (4.31)$$

As before, a multiplication by \mathbf{h} is equivalent to a derivative with respect to \mathbf{k} (for subsurface-offset MVA, we refer to [Symes, 2008](#)). With the ray theory, we have $|\mathbf{k}| = 2\omega \cos(\theta)/c_0(\mathbf{x})$, where θ is the half-opening angle at the image point \mathbf{x} . This is thus

equivalent to a constant angle migration (for angle-domain MVA, we refer to [Biondi and Symes, 2004](#)).

We have established the link between a misfit in the data-domain and a residual in the image-domain. The two formulations can be equivalently used under the Born approximation. The essential element to link the two domains are true-amplitude migration. DSO seeks a model which produces kinematically coherent images. Despite migration artifacts, the conventional migration, an adjoint operator of the modeling, can already construct reflectivity images at correct positions in an accurate background model. Therefore, one may argue that true-amplitude migration is not essential for MVA. However, in the DWI approach, observed and modeled data are not comparable if only the adjoint operator is used to reconstruct reflectivity model. Thus, its image-domain alternative, the DSO approach, might produce biased result without introducing the true amplitude migration scheme. This is consistent with the conclusion in [Chapter 3](#): true-amplitude migration removes the artifacts in velocity updates related to migration smiles, leading to a more robust approach. For a deeper analysis, we conduct numerical experiments in the next section.

4.3.3 Numerical comparisons

We now compare the numerical results of image-domain and data-domain IVA. With the same preconditioner, the new approach converges slower compared to image-domain IVA, as presented in [section 4.2.3](#). To analyze the reasons, we first focus on the shape of macromodel gradient in a simple homogeneous model with three reflectors embedded in. We investigate which approach is better conditioned and what preconditioner should be applied. Then, we validate the choice of preconditioner by comparing in details the results for the Marmousi model, including the final macro velocity model for IVA, associated CIGs, and the estimated complete velocity model after subsequent FWI. The purpose is to investigate the quantitative links as well as the differences between the image-domain and data-domain methods.

Multi-reflector model

We start from a homogeneous model, for which the velocity is 2.5 km/s, with three horizontal reflectors (at depth $z = 0.6, 1.3$ and 2.0 km) embedded in. We trigger 181 shots every 0.04 km from 2.4 to 9.6 km at the surface. Receivers are symmetrically deployed every 0.02 km on both sides of shots with offsets ranging from -2.4 to 2.4 km.

We compute the first macromodel gradient for two homogeneous models, for which the velocities are 2.0 and 3.0 km/s, respectively. In an ideal macromodel gradient, we expect only positive (negative) values for a background velocity higher (lower) than the correct velocity. Beyond the sign aspects, the gradients are expected to exhibit a three-layered shape separated by three reflectors. The value of gradient should be stronger for

the shallower part as it contains three contributions from the three layers, and weaker for the deeper part. The gradients obtained with image-domain IVA ($\alpha = 0$) exhibit strong variations around the interface locations (Figure 4.22a and 4.22e). The sign of the gradient values are even altered for some positions around the interfaces ($z = 1.5$ and 2.3 km) in Figure 4.22e. On the other hand, the gradients obtained with image-domain IVA ($\alpha = 1$) are closer to the expectations, consisting of only positive or negative values and of a layered gradient shape (Figure 4.22b and 4.22f). The results obtained with data-domain IVA are comparable to the ones obtained with image-domain IVA ($\alpha = 1$), except for the low sensitivity to the deeper parts of the model (Figure 4.22c and 4.22g). After multiplying the gradients by depth z , the sensitivity to deeper parts of the model is significantly improved. We extract the central profiles of different gradients to compare the associated behaviors in details. Despite some differences around the interface positions, the preconditioned results exhibit almost the same trend as the ones obtained with image-domain IVA ($\alpha = 1$) (Figure 4.23). Note that the profiles are normalized by the respective maximum absolute value.

After preconditioning, the gradient obtained with data-domain IVA approximates the result with image-domain IVA ($\alpha = 1$). We explain the reason through a rough analysis. The gradient of equations 4.7 and 4.8 can be expressed as

$$\frac{\partial \mathcal{J}_{\text{image}}^1}{\partial c_0} = \langle c_0 \frac{\partial D_{s_x} \xi}{\partial c_0}, c_0 D_{s_x} \xi \rangle_s + \langle \frac{\partial c_0}{\partial c_0} D_{s_x} \xi, c_0 D_{s_x} \xi \rangle_s \quad (4.32)$$

$$\frac{\partial \mathcal{J}_{\text{data}}^0}{\partial c_0} = \langle \mathcal{B}_0 \frac{\partial D_{s_x} \xi}{\partial c_0}, \mathcal{B}_0 D_{s_x} \xi \rangle_{s,r,\omega} + \langle \frac{\partial \mathcal{B}_0}{\partial c_0} D_{s_x} \xi, \mathcal{B}_0 D_{s_x} \xi \rangle_{s,r,\omega} \quad (4.33)$$

The right side of the formulations have a consistent form. Compared to the case of image-domain IVA, the macromodel gradient for data-domain IVA has $\mathcal{B}_0^T \mathcal{B}_0$ and $\mathcal{B}_0^T \frac{\partial \mathcal{B}_0}{\partial c_0}$ as additional factors, which are related to four Green's functions. Under the high frequency approximation, these functions correspond to the multiplication by $1/z^2$ with vertical rays. Neglecting other terms of constant or minor values, the gradient for equation 4.33 is decayed by $1/z^2$ compared to the one for equation 4.32. Therefore, the gradient obtained with data-domain IVA should be additionally multiplied by a depth preconditioner, to approximate the one obtained with image-domain IVA. Note that the theoretical explanation is limited to the homogeneous model and vertical rays. Multiplication by z is sufficient for the homogeneous model for the numerical example demonstrated in this section. In the following, we investigate this issue for more complex structures.

Marmousi model

With a homogeneous model, the application of a z preconditioner to the gradient of data-domain IVA is sufficient, in the sense that the gradient shapes for non-preconditioned image-domain IVA and preconditioned data-domain are comparable. However, the first

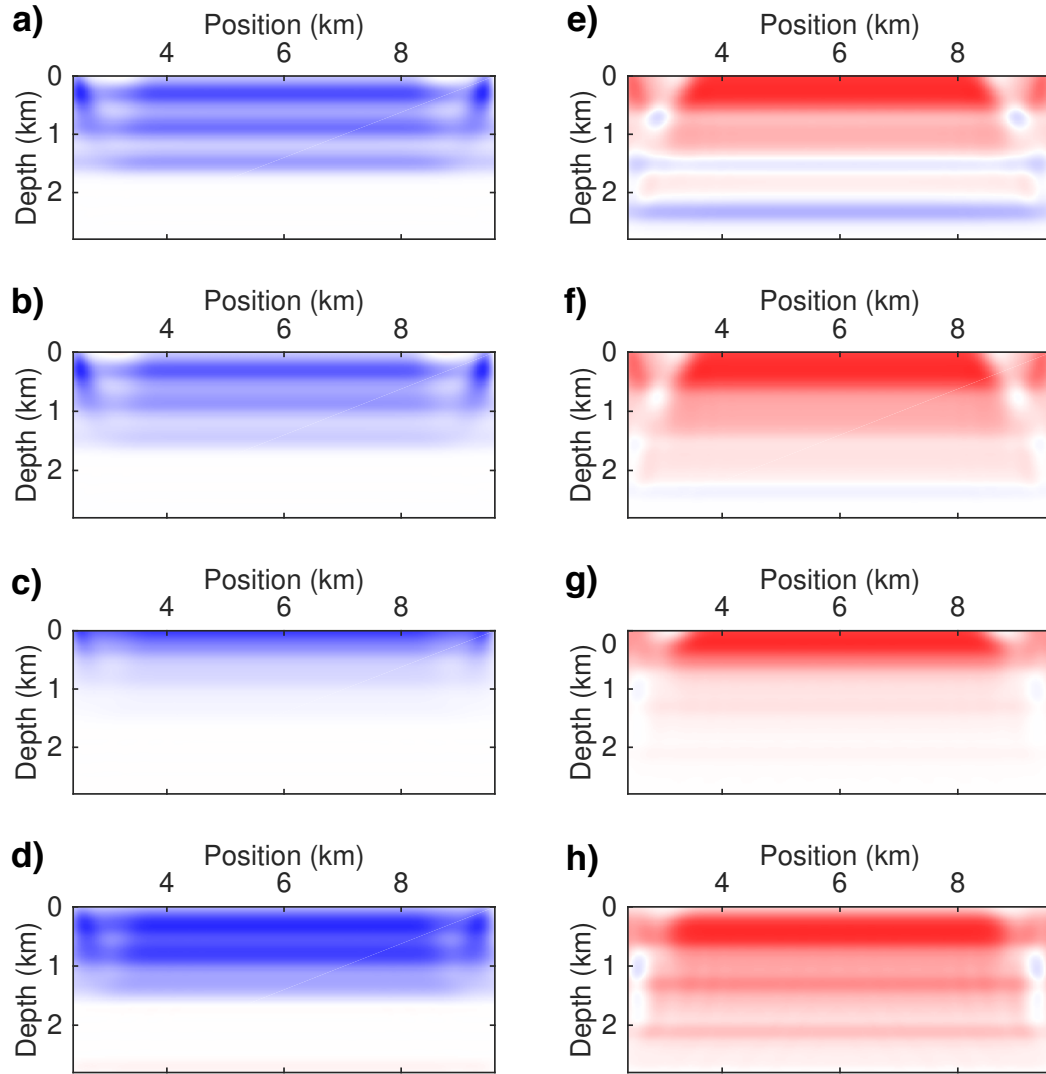


Figure 4.22 – The first macromodel gradient, after a Gaussian smoothing over half a wavelength, obtained with image-domain IVA for $\alpha = 0$ (a,e) and for $\alpha = 1$ (b,f), and with data-domain IVA (c,g). (d,h) are (c,g) multiplied by the value of depth z , respectively. Compared to the exact model, the velocity is 0.5 km/s lower for (a–d) and 0.5 km/s higher for (e–h), respectively.

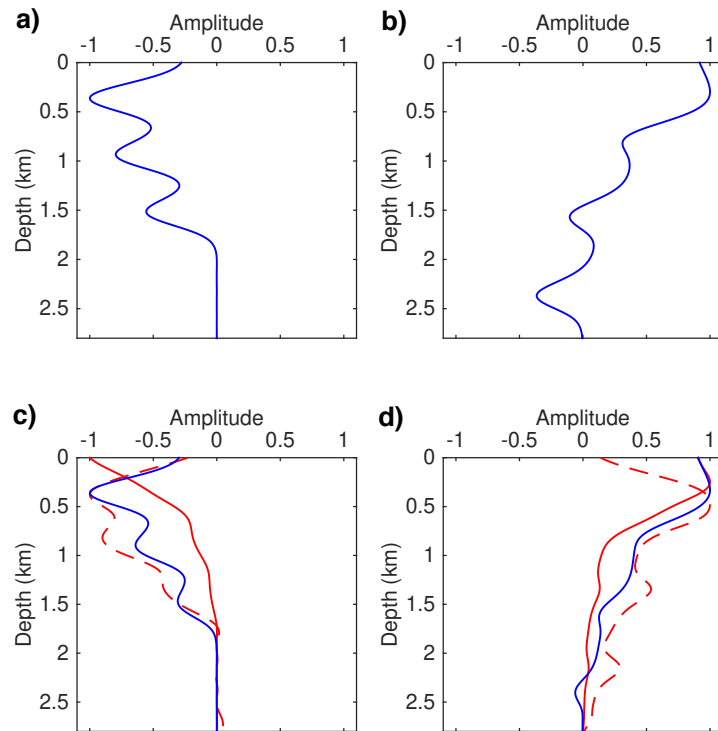


Figure 4.23 – Profiles at position $x = 6$ km for the first gradients displayed in Figure 4.22. (a,b) are related to image-domain IVA ($\alpha = 0$). In plots (c,d), solid blue, solid red and dashed red lines correspond to image-domain IVA ($\alpha = 1$), and data-domain IVA before and after preconditioning, respectively. Compared to the exact model, the velocity is 0.5 km/s lower for (a,c) and 0.5 km/s higher for (b,d), respectively

gradient cannot fully represent the velocity update after nonlinear iterations, and the choice of preconditioner may change in a more complex model. To be more realistic, we further investigate the impact of preconditioner for data-domain IVA in the Marmousi model.

We start from the same homogeneous model ($c_0 = 1.5$ km/s) and perform few nonlinear iterations to update the background velocity with different approaches: non-preconditioned image-domain IVA ($\alpha = 1$), and data-domain IVA before and after preconditioning. By preconditioning, we mean the gradient of the objective function with respect to macromodel is multiplied by z or z^2 . The geometry is same as in section 4.2.3. In all tests, the updated macromodel related to data-domain IVA after preconditioning is closer to the case of non-preconditioned image-domain IVA (Figure 4.24). In the case of the z preconditioner, the velocity for the updated macromodel, obtained with preconditioned data-domain IVA, is still lower than the result of image-domain IVA, after 1 iteration (Figure 4.24a) or 2 iterations (Figure 4.24b). On the other hand, the updated macromodel, obtained with data-domain IVA preconditioned by z^2 , approximates the result of image-domain IVA after 1 iteration (Figure 4.24c). These two macromodels are almost identical after 2 nonlinear iterations (Figure 4.24d). Therefore, z^2 behaves better than z as a preconditioner for data-domain IVA in the Marmousi model, at least after 2 or more iterations. We now focus on the impact of z^2 preconditioner. After 5 nonlinear iterations (Figure 4.25a), the macromodels, retrieved by image-domain IVA and preconditioned data-domain IVA, produce very similar CIGs (Figure 4.25d and 4.25e). These CIGs are different from the case of non-preconditioned data-domain IVA (Figure 4.25c). As the optimization has not converged yet, only the shallower parts of these CIGs are consistent with the case of correct background velocity model (Figure 4.25b). The application of depth preconditioner enhances the response of the macromodel gradient to deeper structures, and thus significantly boosts the convergence. Without the depth preconditioner, the optimization should still converge towards a similar result but at a much slower converging speed. It is the reason why, in section 4.2.3, data-domain IVA uses more iterations (Figure 4.12), than image-domain IVA (Figure 3.18) for convergence, as we applied the same preconditioner to the gradient for both experiments.

We now gather the final results already presented in sections 3.4.3 (for image-domain IVA) and 4.2.3 (for data-domain IVA), for a more detailed comparison between the two approaches. With the application of the same preconditioner z^2 , the inverted macromodels for two approaches are very alike except for the deep part and the edges, as marked by red arrows in Figures 4.26a and 4.26b. In practice, the strategies of two experiments are not completely the same: we do not apply any smoothing to the gradient after several iterations for image-domain IVA, whereas a Gaussian smoothing over at least a quarter wavelength is applied to the gradient for data-domain IVA. Consequently, more details are included in the final result for image-domain IVA than for data-domain IVA. This is clearly visible in profiles of the estimated macromodel (Figure 4.26). Compared to the starting macromodel, the final updated macromodel for both image-domain and data-

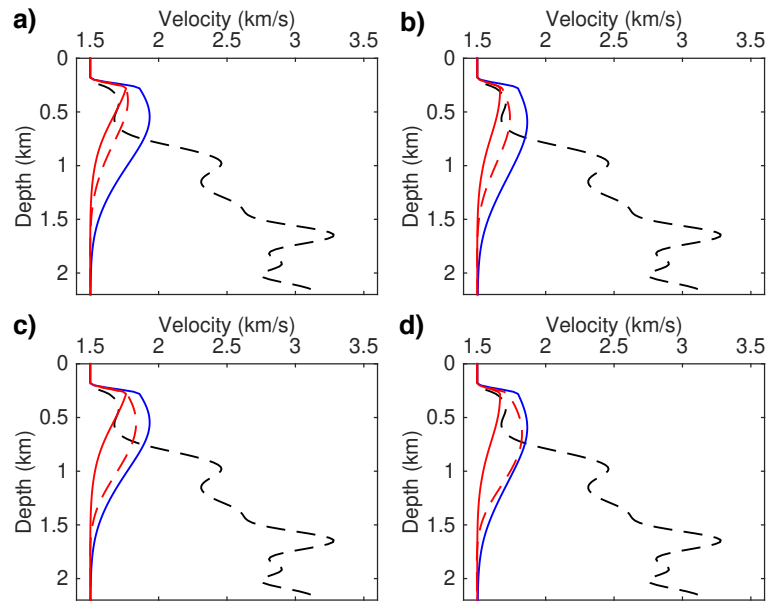


Figure 4.24 – Profiles at $x = 4.6$ km extracted from background models obtained with different methods after few nonlinear iterations. Dashed black lines correspond to the exact macromodel. Solid blue, solid red, and dashed red lines correspond to the results obtained with image-domain IVA ($\alpha = 1$), and with data-domain IVA before and after preconditioning, respectively. The gradients for data-domain IVA is multiplied by z in (a,b), and by z^2 in (c,d). (a,c) and (b,d) correspond to the results after 1 iteration and 2 iterations, respectively.

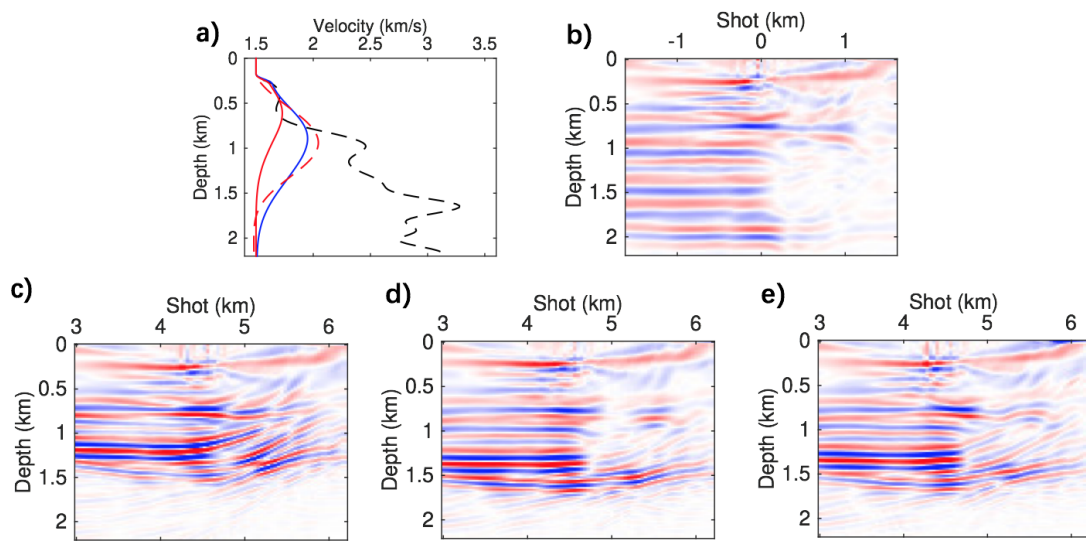


Figure 4.25 – Macromodels obtained with different methods after 5 nonlinear iterations (a) and associated CIGs (b–d). (a) is the same as Figure 4.24d, but for the case after 5 iterations. (b) is associated to the correct macromodel, and (c–e) to the results of non-preconditioned data-domain IVA, data-domain IVA preconditioned by z^2 and non-preconditioned image-domain IVA, respectively.

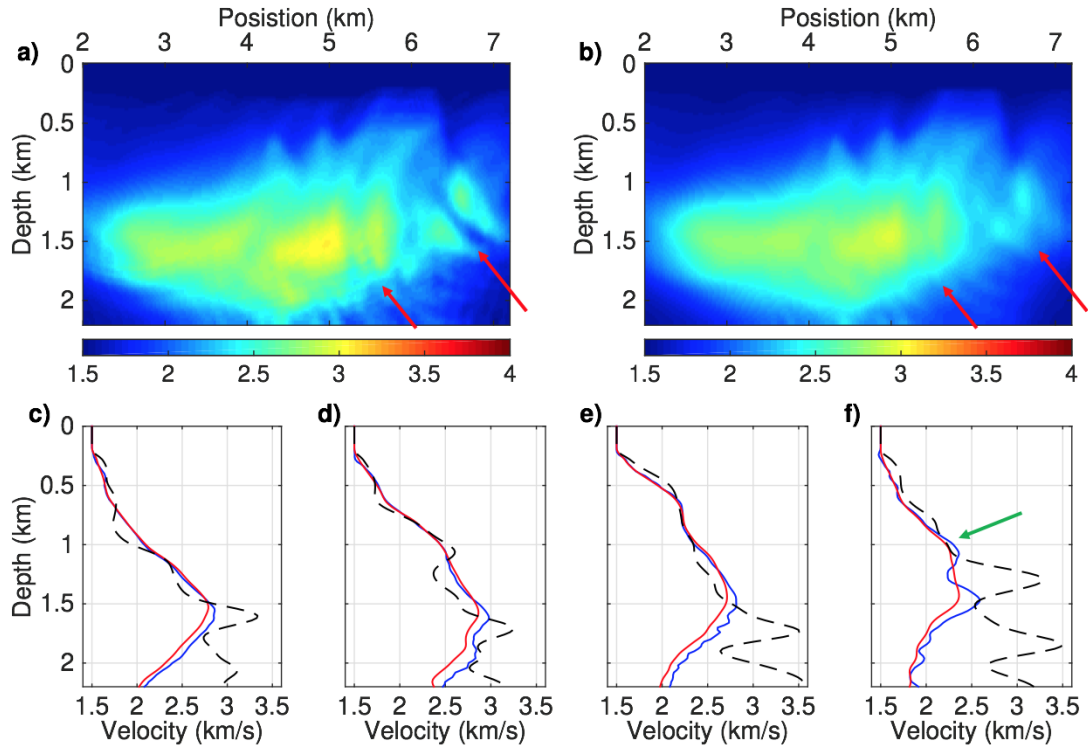


Figure 4.26 – The background model obtained after performing 100 nonlinear iterations with imaged-domain IVA (a) and with data-domain IVA (b). Red arrows are at same positions for (a,b) and mark the area exhibiting differences. (c–f) are vertical velocity profiles at positions 3.5, 4.5, 5.5 and 6.5 km, respectively. Dashed black, solid blue and solid red lines correspond to true Marmousi macromodel, image-domain result (preconditioner z^2) and data-domain result (preconditioner z^2), respectively. Green arrow is at the same position as in Figure 4.28.

domain approaches produce CIGs where events are more coherent, much closer to the case of correct model (Figure 4.27). The differences between CIGs for two methods are mainly visible in the deepest part at a large shot-receiver offset, related to a poor illumination, as shown inside the green rectangles in Figures 4.27f-4.27h. After 70 iterations of FWI, although the areas marked by red arrows in Figures 4.28a and 4.28b still exhibit minor differences, the consistency between two estimated macromodels is significantly further improved. An example is the event marked by the green arrow in Figures 4.26f and 4.28f: the results of image-domain and data-domain approaches are more identical after subsequent FWI.

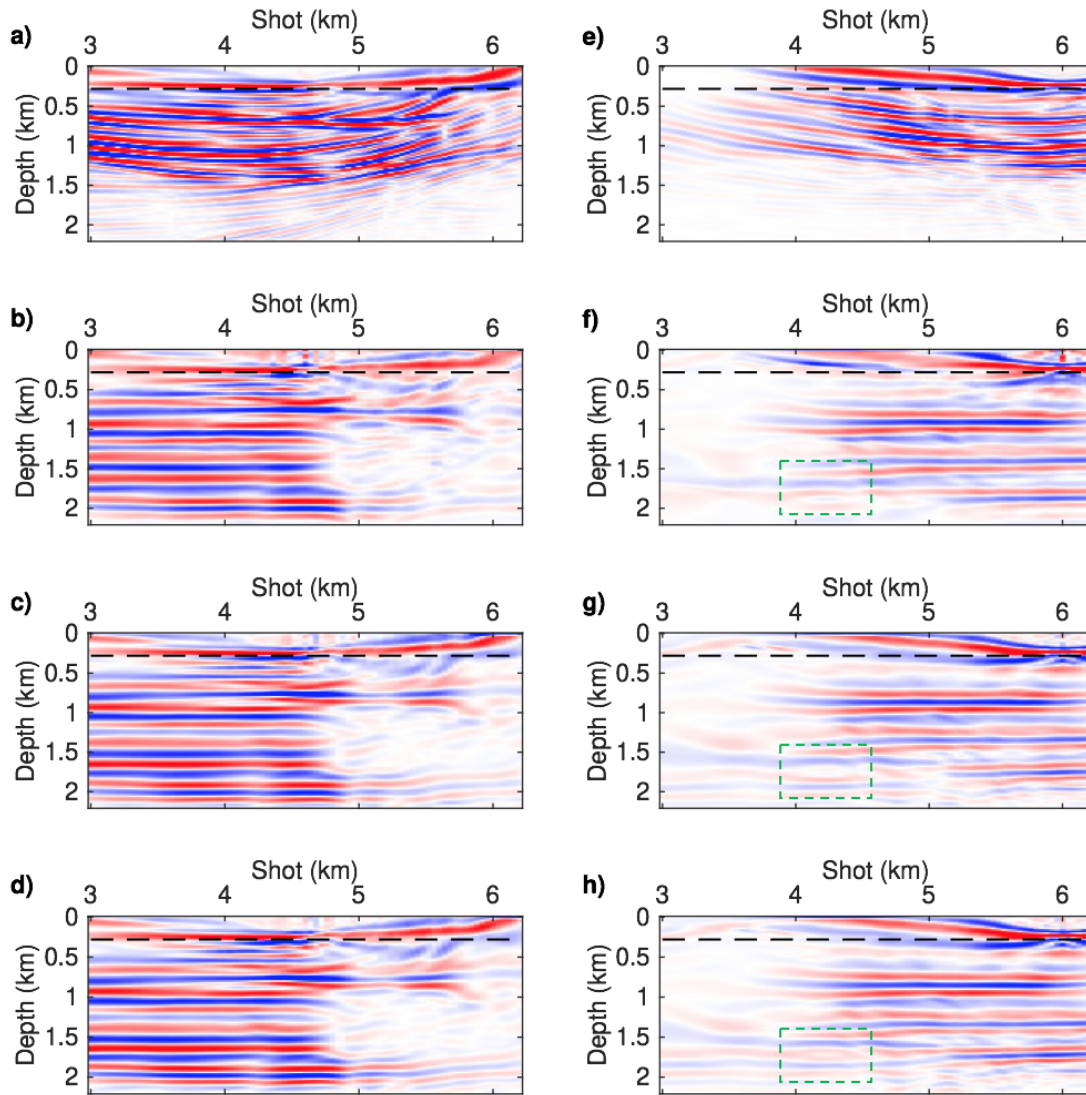


Figure 4.27 – CIGs for starting macromodel (a,f), exact macromodel (b,e) and the final results obtained with image-domain IVA (c,g) and data-domain IVA (d,h). (a–d) correspond to shot position 4.6 km and (f–h) to shot position 6.2 km. Dashed green rectangles are at the same location for (e–h) to mark the differences.

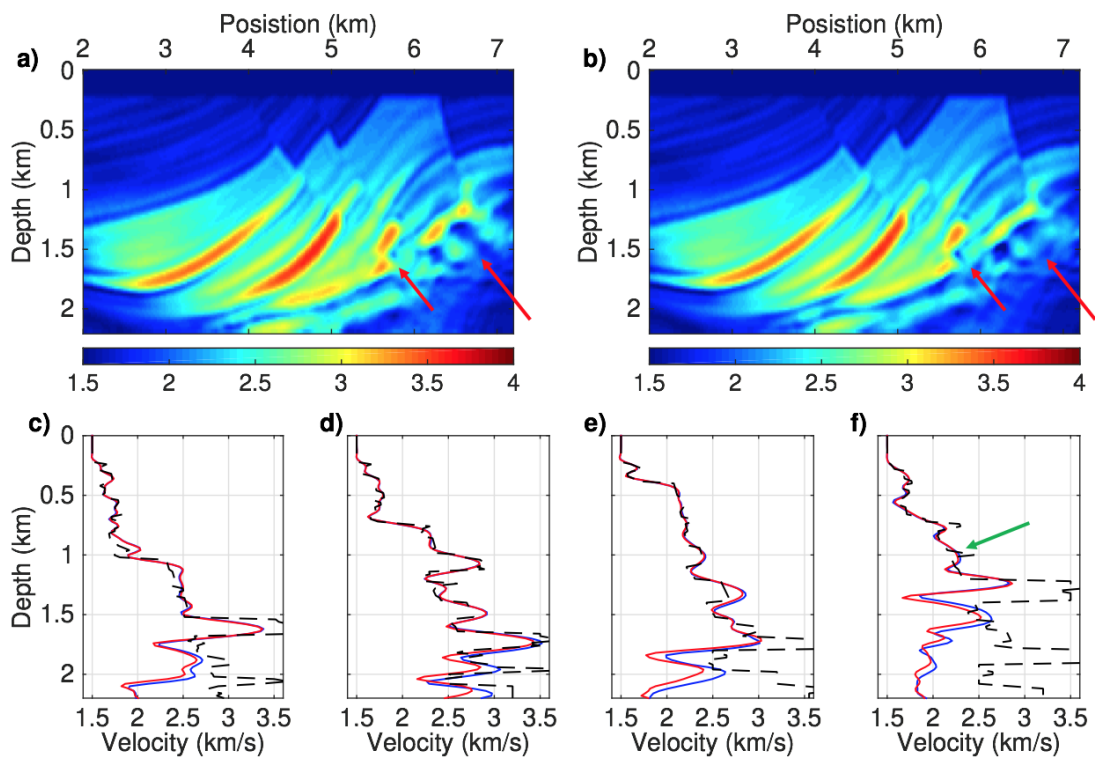


Figure 4.28 – Same as for Figure 4.26, but for the results after subsequent FWI.

Table 4.2 – Strategies proposed for a successful implementation of image-domain and data-domain IVA, at least for the case of Marmousi model.

	Image-domain IVA	Data-domain IVA
c_0^α or $c_0^{\beta(1)}$	$\alpha = 1$	$\beta = 0$
Smoothing ⁽²⁾	no	yes
Preconditioner ⁽³⁾	z^0	z^2

⁽¹⁾ See equations 4.7 and 4.8.

⁽²⁾ For both methods, a Gaussian smoothing over one wavelength is applied to the gradient at the first iteration, and the characteristic length is reduced every 10 iterations. "No" means the gradient is not smoothed after several iterations and "yes" means the gradient is always smoothed;

⁽³⁾ Data-domain IVA preconditioned by z^2 is equivalent to non-preconditioned image-domain IVA; For the numerical experiments, we used z^2 for both approaches.

Summary

We compared the image-domain and data-domain IVA through numerical examples. Several elements are essential to link the image-domain and data-domain IVA: (1) data-domain IVA ($\alpha = 1$) is equivalent to image-domain IVA with $\beta = 0$; (2) data-domain IVA requires a more aggressive preconditioner, in the sense that the macromodel gradient should be multiplied by z or z^2 , to have same behavior as the image-domain IVA. If not preconditioned, data-domain IVA may still requires more iterations; (3) we have observed the macro gradient of data-domain IVA should always be smoothed during iterative optimization. Otherwise, spurious oscillations can present in the updated velocity model, preventing from convergence (discussed in section 4.2.3). One possible reason is that data-domain IVA requires more iterations and we stopped smoothing the gradient before convergence in the case converging towards a local minimum (Figure 4.13). The reader is referred to Table 4.2 for the strategy we followed in practice to deal with the Marmousi model. Note that we preconditioned the both approaches with z^2 in practice, leading to a relative slower convergence speed for data-domain IVA.

4.4 Conclusions

In the first part, we have proposed to extend common-shot IVA from image-domain to data-domain, formulating a new approach by minimizing the norm of scaling data. In an estimated macromodel, these data are modeled from the differentiate images of the inverted reflectivities. For image-domain IVA, we multiplied the inverted reflectivity by the velocity at a specific power to provide more flexibility in the shape of the gradient. This is not necessary needed for the new approach: the associated gradient has a smoother behavior and does not suffer from interface imprint after the Gaussian smoothing. In general, the value of characteristic length of Gaussian smoothing is half a

wavelength, but it may need to be increased for complex structures up to one wavelength during the first iteration. Then, the nonlinear optimization, performed on the Marmousi model, proves its capability to deal with complex media. More efforts are needed to investigate the smoothing strategy applied to the gradient in a nonlinear optimization procedure. By starting from the final data-domain IVA results, the subsequent FWI successfully derives a non-cycle skipped high resolution model, at least for the area where IVA supplies a correct background velocity update.

In the second part, we have compared the image-domain and data-domain methods. Data-domain IVA proves to be equivalent to DWI. DWI was originally proposed for common-shot migration scheme and its idea is now extended to other migration schemes, including common-offset, time-delay, subsurface-offset and common-angle migration schemes. We established a quantitative link between the data fitting principle and the image coherency criteria. It explains why true-amplitude migration is an important aspect in image-domain velocity analysis. Then, we compared the results of image-domain and data-domain IVA approaches. The data-domain approach requires a stronger preconditioner, for example, multiplying the gradient by the value of depth at a higher power. After preconditioning, the velocity update for data-domain IVA is equivalent to the case of image-domain IVA. Otherwise, it converges slower than image-domain IVA. Despite these, both methods prove to be robust and provide a good starting model for FWI. The quantitative equivalence between image-domain and data-domain methods is established through these numerical tests.

The objective function for the new approach defined in data-domain introduces an additional Born modeling operator compared to the case of image-domain IVA. This additional modeling leads to two more wavefield terms in the formulation of the model gradient of the objective function. More precisely, for a single source term, the derivation of model gradient requires computing two wavefields for FWI/RTM, four wavefields for image-domain IVA and six wavefields for data-domain IVA.

The main advantage of data-domain IVA is the definition of the data that can be seen as a reflectivity in time as in the MBTT approach ([Chavent et al., 1994](#); [Plessix et al., 1995](#)). More research is needed to fully exploit this approach.

4.5 Appendix I: Gradient derivation for IVA in the data-domain

We explain here how to compute the gradient of the data-domain IVA objective function (equation 4.8) using the adjoint-state technique (Plessix, 2006). We consider the source and receiver wavefields defined in equations 4.5 and 4.6. In the following, we note the 2D acoustic Helmholtz operator:

$$\mathcal{L}_0 = -\frac{\omega^2}{c_0^2(\mathbf{x})} - \Delta. \quad (4.34)$$

We define a mask operator K such that $d^0(\mathbf{s}, \mathbf{r}, \omega) = K(\mathbf{x}, \mathbf{r})S_d(\mathbf{s}, \mathbf{x}, \omega)$. Then, we extend equation 4.8 with Lagrangian formulation for $\beta = 0$, denoted by \mathcal{J}_{ext}^0 , depending on $(c_0, S_d, \xi_{inv}, S_0, R_0, \lambda_d, \lambda_\xi, \lambda_s, \lambda_r)$ and reading

$$\begin{aligned} \mathcal{J}_{ext}^0 = & \frac{1}{2} \|K(\mathbf{x}, \mathbf{r})S_d(\mathbf{s}, \mathbf{x}, \omega)\|^2 \\ & - \iiint d\mathbf{x}d\mathbf{s}d\omega \lambda_d^*(\mathbf{s}, \mathbf{x}, \omega) \left[\mathcal{L}_0 S_d(\mathbf{s}, \mathbf{x}, \omega) - D_{sx} \xi_{inv}(\mathbf{x}, \mathbf{s}) \frac{S_0(\mathbf{s}, \mathbf{x}, \omega)}{i\omega} \right] \\ & - \iiint d\mathbf{x}d\mathbf{s} \lambda_\xi(\mathbf{x}, \mathbf{s}) \left[\xi_{inv}(\mathbf{x}, \mathbf{s}) - 4 \int d\omega (|S_0(\mathbf{s}, \mathbf{x}, \omega)|^2 + \epsilon)^{-1} \right. \\ & \quad \times \left(\nabla_{\mathbf{x}} S_0^*(\mathbf{s}, \mathbf{x}, \omega) \cdot \nabla_{\mathbf{x}} R_0(\mathbf{s}, \mathbf{x}, \omega) + \left(\frac{i\omega}{c_0(\mathbf{x})} \right)^2 S_0^*(\mathbf{s}, \mathbf{x}, \omega) R_0(\mathbf{s}, \mathbf{x}, \omega) \right) \left. \right] \\ & - \iiint d\mathbf{x}d\mathbf{s}d\omega \lambda_s^*(\mathbf{s}, \mathbf{x}, \omega) \left[\mathcal{L}_0 S_0(\mathbf{s}, \mathbf{x}, \omega) - (i\omega)^3 \Omega(\omega) \delta(\mathbf{x} - \mathbf{s}) \right] \\ & - \iiint d\mathbf{x}d\mathbf{s}d\omega \lambda_r^*(\mathbf{s}, \mathbf{x}, \omega) \left[\mathcal{L}_0^* R_0(\mathbf{s}, \mathbf{x}, \omega) - \int d\mathbf{r} d^{obs}(\mathbf{s}, \mathbf{r}, \omega) \frac{\partial \delta(\mathbf{x} - \mathbf{r})}{\partial r_z} \right], \end{aligned} \quad (4.35)$$

where λ_d , λ_ξ , λ_s and λ_r are adjoint state variables associated with S_d , ξ , S_0 and R_0 , respectively. The different terms can be interpreted as constraints. The total derivative of \mathcal{J}_{ext}^0 with respect to the velocity is given by

$$\begin{aligned} \frac{d\mathcal{J}_{ext}^0}{dc_0} = & \frac{\partial \mathcal{J}_{ext}^0}{\partial c_0} + \frac{\partial S_d}{\partial c_0} \frac{\partial \mathcal{J}_{ext}^0}{\partial S_d} + \frac{\partial \xi_{inv}}{\partial c_0} \frac{\partial \mathcal{J}_{ext}^0}{\partial \xi_{inv}} + \frac{\partial S_0}{\partial c_0} \frac{\partial \mathcal{J}_{ext}^0}{\partial S_0} + \frac{\partial R_0}{\partial c_0} \frac{\partial \mathcal{J}_{ext}^0}{\partial R_0} \\ & + \frac{\partial \lambda_d}{\partial c_0} \frac{\partial \mathcal{J}_{ext}^0}{\partial \lambda_d} + \frac{\partial \lambda_\xi}{\partial c_0} \frac{\partial \mathcal{J}_{ext}^0}{\partial \lambda_\xi} + \frac{\partial \lambda_s}{\partial c_0} \frac{\partial \mathcal{J}_{ext}^0}{\partial \lambda_s} + \frac{\partial \lambda_r}{\partial c_0} \frac{\partial \mathcal{J}_{ext}^0}{\partial \lambda_r}. \end{aligned} \quad (4.36)$$

To avoid computing Fréchet derivatives $(\frac{\partial S_d}{\partial c_0}, \frac{\partial \xi_{inv}}{\partial c_0}, \frac{\partial S_0}{\partial c_0}, \frac{\partial R_0}{\partial c_0}, \frac{\partial \lambda_d}{\partial c_0}, \frac{\partial \lambda_\xi}{\partial c_0}, \frac{\partial \lambda_s}{\partial c_0}, \frac{\partial \lambda_r}{\partial c_0})$, we develop the IVA gradient by taking $(\frac{\partial \mathcal{J}_{ext}^0}{\partial S_d}, \frac{\partial \mathcal{J}_{ext}^0}{\partial \xi_{inv}}, \frac{\partial \mathcal{J}_{ext}^0}{\partial S_0}, \frac{\partial \mathcal{J}_{ext}^0}{\partial R_0}, \frac{\partial \mathcal{J}_{ext}^0}{\partial \lambda_d}, \frac{\partial \mathcal{J}_{ext}^0}{\partial \lambda_\xi}, \frac{\partial \mathcal{J}_{ext}^0}{\partial \lambda_s}, \frac{\partial \mathcal{J}_{ext}^0}{\partial \lambda_r}) =$

0. The partial derivatives $\frac{\partial \mathcal{J}_{ext}^0}{\partial \lambda_d}$, $\frac{\partial \mathcal{J}_{ext}^0}{\partial \lambda_\xi}$, $\frac{\partial \mathcal{J}_{ext}^0}{\partial \lambda_s}$ and $\frac{\partial \mathcal{J}_{ext}^0}{\partial \lambda_r}$ correspond to the definitions of inverted image, forward wavefield and backward wavefield, respectively. The derivatives of \mathcal{J}_{ext}^0 with respect to ξ , S_0 and R_0 , imposed to 0, give the expressions to calculate the adjoint state variables λ_s , λ_r , λ_ξ and λ_d , respectively:

$$\begin{aligned} \mathcal{L}_0^* \lambda_s(\mathbf{s}, \mathbf{x}, \omega) = & -4 \left[\nabla_{\mathbf{x}} \cdot (U(\mathbf{s}, \mathbf{x}, \omega) \nabla_{\mathbf{x}} R_0(\mathbf{s}, \mathbf{x}, \omega)) \right. \\ & - \left(\frac{i\omega}{c_0(\mathbf{x})} \right)^2 U(\mathbf{s}, \mathbf{x}, \omega) R_0(\mathbf{s}, \mathbf{x}, \omega) \\ & \left. + 2S_0(\mathbf{s}, \mathbf{x}, \omega) U(\mathbf{s}, \mathbf{x}, \omega) \Re \left\{ \tilde{\xi}_{inv}(\mathbf{s}, \mathbf{x}, \omega) \right\} \right] \end{aligned} \quad (4.37)$$

$$- D_{sx} \xi_{inv}(\mathbf{x}, \mathbf{s}) \frac{\lambda_d(\mathbf{s}, \mathbf{x}, \omega)}{i\omega}, \quad (4.38)$$

$$\begin{aligned} \mathcal{L}_0 \lambda_r(\mathbf{s}, \mathbf{x}, \omega) = & -4 \left[\nabla_{\mathbf{x}} \cdot (U(\mathbf{s}, \mathbf{x}, \omega) \nabla_{\mathbf{x}} S_0(\mathbf{s}, \mathbf{x}, \omega)) \right. \\ & \left. - \left(\frac{i\omega}{c_0(\mathbf{x})} \right)^2 U(\mathbf{s}, \mathbf{x}, \omega) S_0(\mathbf{s}, \mathbf{x}, \omega) \right], \end{aligned} \quad (4.39)$$

$$\lambda_\xi(\mathbf{x}, \mathbf{s}) = - \int d\omega D_{sx} \frac{S_0(\mathbf{s}, \mathbf{x}, \omega) \lambda_d^*(\mathbf{s}, \mathbf{x}, \omega)}{i\omega}, \quad (4.40)$$

$$\mathcal{L}_0^* \lambda_d(\mathbf{s}, \mathbf{x}, \omega) = \int d\mathbf{r} \delta(\mathbf{x} - \mathbf{r}) K(\mathbf{y}, \mathbf{r}) S_d(\mathbf{s}, \mathbf{y}, \omega), \quad (4.41)$$

where $U = \lambda_\xi / (|S_0|^2 + \epsilon)$ for simplification and $\tilde{\xi}_{inv}(\mathbf{s}, \mathbf{x}, \omega)$ is the integrand of $\xi_{inv}(\mathbf{s}, \mathbf{x})$. Equations 4.38–4.41 are solved in reverse order. Subsequently, the final gradient is obtained by inserting the values of the forward and backward propagated wavefields, inverted images, and associated adjoint state variables into equation 4.11.

Chapter 5

Conclusions and Perspectives

Contents

5.1	Conclusions	171
5.1.1	Common-shot Inversion Velocity Analysis	171
5.1.2	Links between image-domain and data-domain methods . . .	172
5.2	Perspectives	173
5.2.1	Edge effects	173
5.2.2	Introduction of more physics	173
5.2.3	Extension to 3D	174
5.2.4	Application to real data	175

Résumé du chapitre 5

Dans cette thèse, j'ai étudié l'analyse de vitesse par migration sous l'approximation de la propagation des ondes acoustiques en milieu à densité constante. Ce type de technique estime la qualité du macro-modèle en regardant la cohérence des images migrées. Elle est conduite sur des panneaux de focalisation (CIGs). Ils sont construits de deux manières, soit la migration est réalisée sur des sous-ensembles de données comme les points de tir, soit un décalage spatial est introduit lors de la construction des images de réflectivité. Des investigations ont proposé de coupler l'inverse direct à l'analyse de vitesse pour une plus grande robustesse, et dans le cas de l'offset en profondeur. Cependant, l'approche offset en profondeur demande beaucoup de mémoire, même en 2D, et ne peut pas être étendue en 3D actuellement.

Historiquement, [Chauris et al. \(2002a\)](#), ont regardé le lien entre l'analyse de vitesse ([Symes and Carazzone, 1991](#); [Sava and Biondi, 2004](#)), un schéma défini dans le domaine image, et la stéréotomographie ([Billette and Lambaré, 1998](#); [Lambaré, 2008](#)), une approche dans le domaine des données. C'était pour des approches avec les rais. L'optimisation par semblance différentielle (DSO) a été initialement définie comme un terme de régularisation de l'inversion des formes d'onde ([Symes and Kern, 1994](#)). Beaucoup d'approches ont introduit le concept d'analyse de vitesse dans l'inversion ([Biondi and Almomin, 2012](#); [Liu et al., 2014a](#); [Alkhalifah and Wu, 2016](#); [Chauris et al., 2017](#)). Cependant, la relation entre les domaines images et des données demande plus de travaux.

Dans cette thèse, j'ai regardé principalement deux aspects : (1) pour une approche applicable et robuste, j'ai développé l'inversion directe par point de tir, et surtout le couplage avec l'analyse de vitesse ; (2) j'ai étendu l'analyse de vitesse par inversion du domaine image au domaine des données, et par ce fait établi des liens entre ceux deux domaines.

5.1 Conclusions

In this thesis, I have investigated Migration Velocity Analysis (MVA) techniques under the constant-density acoustic wave-equation approximation. This type of technique assesses the quality of the estimated macromodel by measuring the coherency of migrated images. The analysis is carried out on the panels called Common Image Gathers (CIGs). CIGs can be built in two manners: the surface-oriented methods first perform migration on different subsets of input data, such as the common-shot gathers, and then collect images along the redundant parameter; the depth-oriented methods extend the image volume with an additional parameter, for example the subsurface-offset as a spatial shift, inserted during the construction of reflectivity images. Recent investigations propose to couple the direct inversion to MVA in the subsurface-offset domain, introducing better robustness. However, this approach requires large memory, even in 2D, and cannot be currently extended to 3D.

Historically, [Chauris et al. \(2002a\)](#) investigated the link between MVA ([Symes and Carazzone, 1991](#); [Sava and Biondi, 2004](#)), a scheme defined in the image-domain, and stereotomography ([Billette and Lambaré, 1998](#); [Lambaré, 2008](#)), an approach defined in the time-domain. But it was only for ray-based methods. Differential Semblance Optimization (DSO) was initially defined as a regularization term of Full Waveform Inversion (FWI) ([Symes and Kern, 1994](#)). Many approaches introduced the concept of MVA to waveform inversion ([Biondi and Almomin, 2012](#); [Liu et al., 2014a](#); [Alkhalifah and Wu, 2016](#); [Chauris et al., 2017](#)). However, the relationship between image-domain and data-domain methods has not been fully understood.

In this thesis, I have mainly investigated these two issues: (1) for an affordable and robust approach, I have developed the common-shot direct inversion and, more importantly, coupled it to velocity analysis; (2) I have extended the common-shot IVA from image-domain to data-domain and, along this line, established some links between two domains.

5.1.1 Common-shot Inversion Velocity Analysis

In Chapter 3, I have developed an alternative approach to [Qin et al. \(2015\)](#), related to the common-shot direct inversion. It is a pseudo-inverse of the Born modeling operator in the asymptotic sense. The method automatically compensate for geometrical spreading losses and uneven illuminations, such that migration smiles visible in reflectivity images are highly attenuated. Then, I have proposed to couple this direct inverse operator to velocity analysis, leading to Inversion Velocity Analysis (IVA), a nested optimization procedure. The interface imprints are visible around reflector positions in the macromodel gradient for common-shot IVA, and I removed these spurious oscillations by multiplying the inverted reflectivity with the velocity at a specific power ($\alpha = 1$) before measuring the consistency of the migrated images. The velocity gradient behaves in a

pure tomographic mode after the application of a Gaussian smoothing over half a wavelength distance. After, I have paid attention to the impact of triplicated wavefields. It appears that triplicated waves still alter the images, but the coupling with velocity analysis is robust, at least better than for conventional MVA. This approach can deal with complex models, at least in the case of a smooth version of the Marmousi model. By starting from the final IVA results, the subsequent FWI successfully derives a non-cycle skipped high resolution model, at least for the areas of which IVA supplies a correct background velocity update. Three aspects are essential for a successful IVA implementation: inverse instead of adjoint, reflectivity multiplied by the velocity at a specific power in the evaluation of the coherency, and Gaussian smoothing over half a wavelength distance. Common-shot velocity analysis is less numerically expensive than its counterpart in the subsurface-offset domain, offering new possibilities for the extension to 3D in the future.

5.1.2 Links between image-domain and data-domain methods

In Chapter 4, I have proposed to extend common-shot IVA from image-domain to data-domain, formulating a new approach by minimizing the data misfit. In a given macro-model, these data are modeled from the differentiate images of the inverted reflectivities. For image-domain IVA, I multiplied the inverted reflectivity by the background velocity at a specific power to provide more flexibility in the shape of the gradient, and thus removed the interface imprints (spurious oscillations) visible in the gradient. This is not necessary for the data-domain IVA approach: the associated gradient has a smoother behavior and does not suffer from interface imprint after the Gaussian smoothing. In general, the value of characteristic length of Gaussian smoothing is half a wavelength, but it may need to be increased for complex structures. Through numerical experiments, I have shown that data-domain IVA is capable of dealing with complex media, at least for the case of a smooth version of the Marmousi model. By starting from the final data-domain IVA macromodel, the subsequent FWI successfully derives a non-cycle skipped result, at least for the well-illuminated area.

In Chapter 4, I have also compared the image-domain and data-domain methods. Data-domain IVA proves to be equivalent to Differential Waveform Inversion (DWI). DWI was originally proposed for common-shot migration scheme and, in this thesis, its idea is now extended to other migration schemes, including common-offset, time-delay, subsurface-offset and common-angle migration schemes. I established a quantitative link between the data fitting principle and the image coherency criteria. It explains why true-amplitude migration is an important aspect in image-domain velocity analysis. Then, the image-domain and data-domain IVA approaches are compared through numerical examples. The data-domain approach requires a stronger preconditioner, for example, multiplying the gradient by the value of depth at a higher power. After preconditioning, the velocity update for data-domain IVA is equivalent to the case of

image-domain IVA. Otherwise, it requires more iterations. Despite that, both methods prove to be robust and provide a good starting model for FWI. The equivalence between image-domain and data-domain methods is further justified through these numerical tests. More research is needed to fully exploit the structure of data-domain IVA.

In Chapter 4, I have also discussed the numerical cost of data-domain IVA. The objective function defined in data-domain introduces one more Born modeling operator compared to the case of image-domain IVA, leading to a more CPU-time consuming implementation. More precisely, for a single source term, the derivation of model gradient requires computing two wavefields for FWI/RTM, four wavefields for image-domain IVA and six wavefields for data-domain IVA. In terms of computational cost, data-domain IVA is not necessary to be recommended. However, it has a natural tomographic velocity update and has provided a way to investigate the links between image-domain and data-domain and has a more natural tomographic velocity update.

5.2 Perspectives

5.2.1 Edge effects

An important drawback for common-shot schemes is the edge effects of migration. As for each shot, the illumination of subsurface is limited: even if the direct inverse compensates for uneven illuminations, the edging parts still suffer from a poor illumination, leading to migration smiles. After the summation over all shots, the edge effects can be suppressed due to improved signal to noise ratio. However, common-shot IVA requires prestack images, such that the related macromodel gradient also suffers from severe edge effects (Figure 5.1). Only the central part of the gradient has an expected shape and can lead to reliable velocity updates. To cope with this issue, one possibility is to consider iterative migration, which deals with incomplete observations better than asymptotic inversion (Nemeth et al., 1999). The same kind of limitation is also observed for the subsurface-offset case (Cocher, 2017).

5.2.2 Introduction of more physics

I have proposed a direct inverse formula for common-shot gathers under the constant-density isotropic acoustic approximation. In this study, I have used an explicit scale separation between a smooth background velocity model and a model perturbation. Following a strategy proposed by Zhou (2016) in the framework of FWI, one could also investigate a more natural scale separation, using velocity and impedance perturbation to parameterize the model, the velocity controlling the kinematics of wave propagation and impedance accounting for the reflective property of the subsurface instead of

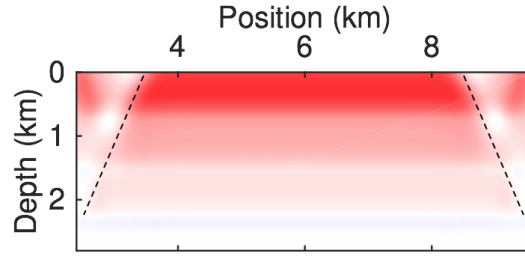


Figure 5.1 – Description of the impact of edge effects on macromodel gradient obtained with image-domain IVA. Black dashed lines mark the boarder between well illuminated area and edges.

the velocity model and model perturbation. [Zhou et al. \(2015\)](#) indicated that these two parameters exhibit less coupling compared to the case of velocity and density.

Under the constant-density acoustic approximation, [Lameloise and Chauris \(2016\)](#) show how transmitted waves can be included in MVA techniques by constructing extended images in a very similar manner to the usual procedure for reflections. They propose a strategy in which these two kinds of events are used successively. This is in the case of subsurface-offset and the extension to common-shot gathers needs further investigations. Transmitted waves are usually used in a first step to update the shallow part of the velocity model; then reflections are inverted to reach the deeper part of the model. Inverting reflection and transmission data simultaneously could potentially better constrain the inversion. As pointed out by [Lameloise et al. \(2015\)](#), an issue is that transmitted events are more energetic than reflections, requiring to introduce weights to balance the amplitudes of both kinds of events. The combination of reflection and transmitted data in a unified framework needs further investigations.

5.2.3 Extension to 3D

Considering the extension to 3D, subsurface-offset MVA will become prohibitively expensive due to the significantly increased memory requirements (from (n_z, n_x, n_h) to $(n_z, n_y, n_x, n_{h_z}, n_{h_z})$), whereas the common-shot IVA requires less (from $(n_z, n_x, 3)$ to $(n_z, n_y, n_x, 5)$). In addition to the storage of 3D images, a crosscorrelation should be performed for each couple (h_x, h_y) and the cost of it is not negligible for the subsurface-offset case any more. The common-shot IVA requires less computational cost and memory than in the subsurface-offset case, and may be extended to 3D. A new inversion formula taking into account the new dimensions in 3D should be defined with a strategy similar to the 2D case (Chapter 3). A main difficulty is that 3D acquisition does not provide dense source and receiver coverage in all directions. With the common-shot migration scheme, it may cause even strong undesirable edge effects on reflectivity image

and associated gradients.

5.2.4 Application to real data

In this thesis, I have only investigated the synthetic cases under the constant-density acoustic approximation. I rely on Born approximations, such that only first reflected waves are present in data. The real case is much more complicated, and I should at least pay attention to four essential elements: (1) in correct estimation of the source wavelet; (2) incomplete physics in the forward modeling engine; (3) limited acquisition; (4) 3D effects.

In the synthetic case, I assumed that the source wavelet is properly estimated, but it may not be the case for real experiments. As a prelude to real data applications, one can apply the proposed IVA approaches to observed data computed with a source wavelet different from the one used during inversion, to check the sensitivity to source estimation errors. For the direct inverse, the approach is robust, whereas this is not the case for the iterative scheme (Cocher, 2017).

More generally, the physics that is not accounted for in the forward modeling engine may lead to inconsistent CIGs and velocity updates. For example, multi-scattered and converted waves are not predicted under the first Born approximation. The impact of attenuation and elastic on the amplitude may also be an issue for IVA strategies. To cope with multiples, Cocher et al. (2017a) has proposed iterative migration under the high-order Born approximation for subsurface-offset MVA, and this strategy should be extended to common-shot schemes. As the next step, one may try to develop a direct inverse formula for the elastic case, and investigate the impacts of attenuation on direct inversion.

Finally the effect of limited acquisition should be further analyzed. Direct inverse formulas are indeed derived with the assumption of an infinite acquisition, while velocity analysis requires a dense shot coverage. In practice, these requirements may not be realistic, especially in 3D the spatial sampling inherent in prestack data can be poor. Once again, iterative migration may be an interesting alternative in this case as it deals more effectively with incomplete observations (Nemeth et al., 1999). For a faster convergence, one can use direct inversion as a preconditioner following the strategy proposed by Qin et al. (2015); Hou and Symes (2016b); Duprat and Baina (2016), and then couple it to velocity analysis.

Bibliography

- Aki, K. and Richards, P.G. [2002] *Quantitative seismology*. University Science Books. (cit. on p. 3.)
- Al-Yahya, K. [1989] Velocity Analysis by Iterative Profile Migration. *Geophysics*, **54**(6), 718–729. (cit. on pp. 10, 25, 26, 61, 78, and 122.)
- Alford, R., Kelly, K. and Boore, D.M. [1974] Accuracy of finite-difference modeling of the acoustic wave equation. *Geophysics*, **39**(6), 834–842. (cit. on p. 42.)
- Alkhalifah, T. [2005] τ -migration and velocity analysis: application to data from the Red Sea. *Geophysical Prospecting*, **53**(5), 643–653. (cit. on pp. 28, 30, and 65.)
- Alkhalifah, T. [2014] Scattering-angle based filtering of the waveform inversion gradients. *Geophysical Journal International*, **200**(1), 363–373. (cit. on pp. 56, 58, and 108.)
- Alkhalifah, T. and Plessix, R.É. [2014] A recipe for practical full-waveform inversion in anisotropic media: An analytical parameter resolution study. *Geophysics*, **79**(3), R91–R101. (cit. on pp. xi, 13, 14, and 122.)
- Alkhalifah, T. and Wu, Z. [2016] The Natural Combination of Full and Image-Based Waveform Inversion. *Geophysical Prospecting*, **64**(1), 19–30. (cit. on pp. 31, 58, 124, 170, and 171.)
- Alkhalifah, T. and Wu, Z. [2017] Migration velocity analysis using pre-stack wave fields. *Geophysical Prospecting*, **65**(3), 639–649. (cit. on pp. 24 and 28.)
- Almomin, A. and Biondi, B. [2012] Tomographic Full Waveform Inversion: Practical and Computationally Feasible Approach. In: *SEG Technical Program Expanded Abstracts 2012*. 1–5. (cit. on p. 123.)
- Asnaashari, A., Brossier, R., Garambois, S., Audebert, F., Thore, P. and Virieux, J. [2013] Regularized seismic full waveform inversion with prior model information. *Geophysics*, **78**(2), R25–R36. (cit. on pp. 10 and 45.)
- Balch, A.H. and Lee, M.W. [1984] *Vertical seismic profiling: technique, applications, and case histories*. IHRDC Press, Boston, MA. (cit. on p. 4.)
- Bamberger, A., Engquist, B., Halpern, L. and Joly, P. [1988] Higher order paraxial wave equation approximations in heterogeneous media. *SIAM Journal on Applied Mathematics*, **48**(1), 129–154. (cit. on p. 17.)
- Baysal, E., Kosloff, D.D. and Sherwood, J.W. [1983] Reverse time migration. *Geo-*

- physics*, **48**(11), 1514–1524. (cit. on pp. [17](#) and [125](#).)
- Béranger, J.P. [1994] A perfectly matched layer for the absorption of electromagnetic waves. *Journal of computational physics*, **114**(2), 185–200. (cit. on p. [42](#).)
- Berkhout, A.J. [2012] Combining Full Wavefield Migration and Full Waveform Inversion, a Glance into the Future of Seismic Imaging. *Geophysics*, **77**(2), S43–S50. (cit. on p. [6](#).)
- Beydoun, W.B. and Mendes, M. [1989] Elastic ray-Born 1 2-migration/inversion. *Geophysical Journal International*, **97**(1), 151–160. (cit. on p. [17](#).)
- Beydoun, W.B. and Tarantola, A. [1988] First Born and Rytov approximations: Modeling and inversion conditions in a canonical example. *The Journal of the Acoustical Society of America*, **83**(3), 1045–1055. (cit. on pp. [12](#), [15](#), [46](#), and [122](#).)
- Beylkin, G. [1985] Imaging of discontinuities in the inverse scattering problem by inversion of a causal generalized Radon transform. *Journal of Mathematical Physics*, **26**(1), 99. (cit. on pp. [20](#), [22](#), [67](#), [77](#), and [80](#).)
- Bharadwaj, P., Mulder, W. and Drijkoningen, G. [2015] An Auxiliary Bump Functional to Overcome Cycle Skipping in Waveform Inversion. In: *77th EAGE Conference & Exhibition*. WS05 D03. (cit. on pp. [22](#), [53](#), and [122](#).)
- Bharadwaj, P., Mulder, W. and Drijkoningen, G. [2016] Full waveform inversion with an auxiliary bump functional. *Geophysical Journal International*, **206**(2), 1076–1092. (cit. on p. [53](#).)
- Billette, F., Bégat, S.L., Podvin, P. and Lambaré, G. [2003] Practical aspects and applications of 2D stereotomography. *Geophysics*, **68**(3), 1008–1021. (cit. on p. [27](#).)
- Billette, F. and Lambaré, G. [1998] Velocity macro-model estimation from seismic reflection data by stereotomography. *Geophysical Journal International*, **135**(2), 671–690. (cit. on pp. [27](#), [32](#), [124](#), [170](#), and [171](#).)
- Biondi, B. and Almomin, A. [2012] Tomographic full waveform inversion (TFWI) by combining full waveform inversion with wave-equation migration velocity analysis. In: *SEG Technical Program Expanded Abstracts 2012*. 1–5. (cit. on pp. [29](#), [31](#), [123](#), [170](#), and [171](#).)
- Biondi, B. and Almomin, A. [2014] Simultaneous inversion of full data bandwidth by tomographic full-waveform inversion. *Geophysics*, **79**(3), WA129–WA140. (cit. on pp. [29](#), [65](#), [76](#), [123](#), and [152](#).)
- Biondi, B. and Sava, P. [1999] Wave-equation migration velocity analysis. In: *SEG Technical Program Expanded Abstracts 1999*. 1723–1726. (cit. on p. [25](#).)
- Biondi, B. and Shan, G. [2002] Prestack imaging of overturned reflections by reverse time migration. In: *SEG Technical Program Expanded Abstracts 2002*. 1284–1287. (cit. on p. [17](#).)
- Biondi, B. and Symes, W.W. [2004] Angle-domain common-image gathers for migration velocity analysis by wavefield-continuation imaging. *Geophysics*, **69**(5), 1283–1298. (cit. on pp. [25](#), [28](#), [31](#), [60](#), [61](#), [63](#), [75](#), [78](#), [123](#), and [154](#).)
- Bleistein, N. [1987] On the imaging of reflectors in the earth. *Geophysics*, **52**(7), 931–

942. (cit. on pp. 20, 22, 67, 77, and 80.)
- Bleistein, N., Zhang, Y., Xu, S., Zhang, G. and Gray, S.H. [2005] Migration/inversion: think image point coordinates, process in acquisition surface coordinates. *Inverse Problems*, **21**(5), 1715. (cit. on p. 21.)
- Bozdağ, E., Trampert, J. and Tromp, J. [2011] Misfit functions for full waveform inversion based on instantaneous phase and envelope measurements. *Geophysical Journal International*, **185**(2), 845–870. (cit. on pp. 21, 22, and 51.)
- Brossier, R., Operto, S. and Virieux, J. [2009] Seismic imaging of complex onshore structures by 2D elastic frequency-domain full-waveform inversion. *Geophysics*, **74**(6), WCC105–WCC118. (cit. on p. 42.)
- Brossier, R., Operto, S. and Virieux, J. [2015] Velocity model building from seismic reflection data by full-waveform inversion. *Geophysical Prospecting*, **63**(2), 354–367. (cit. on pp. 23, 56, 58, 122, 123, and 152.)
- Bunks, C., Saleck, F.M., Zaleski, S. and Chavent, G. [1995] Multiscale seismic waveform inversion. *Geophysics*, **60**(5), 1457–1473. (cit. on pp. 12, 14, 15, 31, 76, and 122.)
- Castellanos-Lopez, C. [2014] *Speed-up and regularization techniques for seismic full waveform inversion*. Ph.D. thesis, Université de Nice Sophia Antipolis. (cit. on p. 45.)
- Cerjan, C., Kosloff, D., Kosloff, R. and Reshef, M. [1985] A nonreflecting boundary condition for discrete acoustic and elastic wave equations. *Geophysics*, **50**(4), 705–708. (cit. on p. 42.)
- Červený, V. [1987] Ray tracing algorithms in three-dimensional laterally varying layered structures. In: *Seismic Tomography*, 5. Springer, 99–133. (cit. on pp. 10 and 80.)
- Chapman, C. [2004] *Fundamentals of seismic wave propagation*. Cambridge university press. (cit. on p. 9.)
- Chauris, H. [2000] *Analyse de vitesse par migration pour l'imagerie des structures complexes en sismique réflexion*. Ph.D. thesis, Ecole des Mines de Paris. (cit. on pp. xii and 27.)
- Chauris, H. and Cocher, E. [2017] From migration to inversion velocity analysis. *Geophysics*, **82**(3), S207–S223. (cit. on pp. xiii, 21, 22, 28, 29, 31, 63, 64, 67, 68, 69, 76, 77, 78, 82, 83, 86, 87, 109, 110, 112, 123, and 124.)
- Chauris, H., Donno, D. and Calandra, H. [2012] Velocity Estimation with the Normalized Integration Method. In: *74th EAGE Conference & Exhibition*. W020. (cit. on p. 51.)
- Chauris, H., Lameloise, C. and Cocher, E. [2015] Inversion Velocity Analysis: the importance of regularisation. In: *Expanded Abstracts. 77th Annual EAGE Meeting and Exposition*, WS05–A02. (cit. on pp. 45, 66, 69, 77, 78, and 123.)
- Chauris, H. and Lameloise, C.A. [2014] Removing Spurious Oscillations in the Gradient of the Differential Semblance Optimization Functional. In: *76th EAGE Conference & Exhibition*. Amsterdam, Th P02 08. (cit. on p. 28.)
- Chauris, H., Lameloise, C.A. and Donno, D. [2013] Migration Velocity Analysis with

- Reflected and Transmitted Waves. In: *75th EAGE Conference & Exhibition*. We P01 01. (cit. on pp. 29, 65, and 76.)
- Chauris, H., Li, Y. and Cocher, E. [2017] Image-Domain versus Data-Domain Velocity Analysis Based on True-Amplitude Subsurface Extended Migration. In: *79th EAGE Conference & Exhibition*. WS09 D02. (cit. on pp. 123, 170, and 171.)
- Chauris, H. and Noble, M. [2001] Two-dimensional velocity macro model estimation from seismic reflection data by local differential semblance optimization: applications to synthetic and real data sets. *Geophysical Journal International*, **144**(1), 14–26. (cit. on pp. 26, 27, 28, 30, 48, 61, 62, 65, 75, 78, 122, and 152.)
- Chauris, H., Noble, M.S., Lambaré, G. and Podvin, P. [2002a] Migration velocity analysis from locally coherent events in 2-D laterally heterogeneous media, Part I: Theoretical aspects. *Geophysics*, **67**(4), 1202–1212. (cit. on pp. 27, 32, 60, 124, 170, and 171.)
- Chauris, H., Noble, M.S., Lambaré, G. and Podvin, P. [2002b] Migration velocity analysis from locally coherent events in 2-D laterally heterogeneous media, Part II: Applications on synthetic and real data. *Geophysics*, **67**(4), 1213–1224. (cit. on pp. xiii and 62.)
- Chauris, H. and Plessix, R.É. [2012] Investigating the Differential Waveform Inversion. In: *74th EAGE Conference & Exhibition*. E07. (cit. on pp. 24, 58, 124, 149, and 150.)
- Chauris, H. and Plessix, R.É. [2013] Differential Waveform Inversion - A Way to Cope with Multiples? In: *75th EAGE Conference & Exhibition*. London, F06. (cit. on pp. 24, 31, 60, 124, and 150.)
- Chavent, G. [1974] Identification of functional parameters in partial differential equations. In: *Joint Automatic Control Conference*, 12. 155–156. (cit. on pp. 11, 44, and 127.)
- Chavent, G., Clément, F. and Gómez, S. [1994] Automatic determination of velocities via migration-based traveltimes waveform inversion: A synthetic data example. In: *SEG Technical Program Expanded Abstracts 1994*. 1179–1182. (cit. on pp. 23, 24, 56, 83, 123, 133, and 165.)
- Chavent, G. and Jacewitz, C. [1995] Determination of Background Velocities by Multiple Migration Fitting. *Geophysics*, **60**(2), 476–490. (cit. on p. 26.)
- Chi, B., Dong, L. and Liu, Y. [2014] Full waveform inversion method using envelope objective function without low frequency data. *Journal of Applied Geophysics*, **109**, 36–46. (cit. on pp. 22 and 51.)
- Chi, B., Dong, L. and Liu, Y. [2015] Correlation-based reflection full-waveform inversion. *Geophysics*, **80**(4), R189–R202. (cit. on pp. xiii, 22, 48, and 59.)
- Claerbout, J.F. [1971] Toward a unified theory of reflector mapping. *Geophysics*, **36**(3), 467–481. (cit. on pp. 18, 75, 122, and 125.)
- Claerbout, J.F. [1985] *Imaging the earth's interior*. Blackwell scientific publications Oxford. (cit. on pp. xi, 4, 6, 8, 34, and 122.)
- Claerbout, J.F. [1992] *Earth soundings analysis: Processing versus inversion*. Blackwell

- Scientific Publications London. (cit. on p. 18.)
- Clapp, M.L., Clapp, R.G. and Biondi, B.L. [2005] Regularized least-squares inversion for 3-D subsalt imaging. In: *SEG Technical Program Expanded Abstracts 2005*. 1814–1817. (cit. on p. 20.)
- Cocher, E. [2017] *Iterative Migration Velocity Analysis: extension to surface-related multiple reflection*. Ph.D. thesis, MINES ParisTech. (cit. on pp. 6, 20, 27, 45, 67, 69, 173, and 175.)
- Cocher, E., Chauris, H. and Lameloise, C.A. [2015] Imaging with Surface-Related Multiples in the Subsurface-Offset Domain. In: *77th EAGE Conference & Exhibition*. We N101 11. (cit. on pp. 17, 65, 76, and 123.)
- Cocher, E., Chauris, H. and Plessix, R.É. [2017a] Robust Iterative Migration Velocity Analysis: Benefits for Imaging with Primaries and with First-Order Surface Multiples. In: *79th EAGE Conference & Exhibition*. Th B4 12. (cit. on pp. 29, 76, 109, and 175.)
- Cocher, E., Chauris, H. and Plessix, R.É. [2017b] Seismic iterative migration velocity analysis: two strategies to update the velocity model. *Computational Geosciences*, **21**, 1–22. (cit. on pp. 20, 60, 65, 66, 67, 69, 76, 77, 109, 124, and 150.)
- Cooke, D.A. and Schneider, W.A. [1983] Generalized linear inversion of reflection seismic data. *Geophysics*, **48**(6), 665–676. (cit. on p. 45.)
- Dahlen, F. and Tromp, J. [1998] *Theoretical global seismology*. Princeton university press. (cit. on p. 3.)
- Dai, W., Fowler, P. and Schuster, G.T. [2012] Multi-source least-squares reverse time migration. *Geophysical Prospecting*, **60**(4), 681–695. (cit. on p. 20.)
- De Bruin, C., Wapenaar, C. and Berkhout, A. [1990] Angle-dependent reflectivity by means of prestack migration. *Geophysics*, **55**(9), 1223–1234. (cit. on p. 76.)
- Devaney, A.J. [1982] A filtered backpropagation algorithm for diffraction tomography. *Ultrasonic imaging*, **4**(4), 336–350. (cit. on p. 13.)
- Díaz, E., Duan, Y., Sava, P. and Pratt, G. [2014] Image-domain and data-domain waveform tomography: A case study. In: *SEG Technical Program Expanded Abstracts 2014*. 1243–1248. (cit. on pp. 12 and 122.)
- Donno, D., Chauris, H. and Calandra, H. [2013] Estimating the background velocity model with the normalized integration method. In: *75th EAGE Conference & Exhibition*. Tu 07 04. (cit. on p. 53.)
- Duprat, V. and Baina, R. [2016] An Efficient Least-squares Reverse-time Migration Using True-amplitude Imaging Condition as an Optimal Preconditioner. In: *78th EAGE Conference & Exhibition*. We SRS2 12. (cit. on pp. 20, 21, 22, 77, 82, 108, 125, and 175.)
- Duquet, B. [1996] *Improving seismic imaging in complex geologic structures*. Ph.D. thesis, University of Paris XIII. (cit. on p. 20.)
- Duquet, B., Marfurt, K.J. and Dellinger, J.A. [2000] Kirchhoff modeling, inversion for reflectivity, and subsurface illumination. *Geophysics*, **65**(4), 1195–1209. (cit. on

- p. 20.)
- Dutta, G. and Schuster, G.T. [2014] Attenuation compensation for least-squares reverse time migration using the viscoacoustic-wave equation. *Geophysics*, **79**(6), S251–S262. (cit. on p. 20.)
- Ehinger, A., Lailly, P. and Marfurt, K.J. [1996] Green’s function implementation of common-offset, wave-equation migration. *Geophysics*, **61**(6), 1813–1821. (cit. on p. 17.)
- Engl, H.W. and Zou, J. [2000] A new approach to convergence rate analysis of Tikhonov regularization for parameter identification in heat conduction. *Inverse Problems*, **16**(6), 1907. (cit. on p. 45.)
- Etgen, J.T. and O’Brien, M.J. [2007] Computational methods for large-scale 3D acoustic finite-difference modeling: A tutorial. *Geophysics*, **72**(5), SM223–SM230. (cit. on p. 9.)
- Farnbach, J.S. [1975] The complex envelope in seismic signal analysis. *Bulletin of the Seismological Society of America*, **65**(4), 951–962. (cit. on p. 51.)
- Faye, J.P. and Jeannot, J.P. [1986] Prestack migration velocities from focusing depth analysis. In: *SEG Technical Program Expanded Abstracts 1986*. 438–440. (cit. on p. 123.)
- Fei, W. and Williamson, P. [2010] On the gradient artifacts in migration velocity analysis based on differential semblance optimization. In: *SEG Technical Program Expanded Abstracts 2010*. 4071–4076. (cit. on pp. xiv, 28, 61, 68, 69, 77, 78, and 83.)
- Fichtner, A. [2011] *Full Seismic Waveform Modelling and Inversion*. Springer. (cit. on pp. 9, 13, and 43.)
- Fletcher, R. and Reeves, C.M. [1964] Function Minimization by Conjugate Gradients. *The Computer Journal*, **7**(2), 149–154. (cit. on p. 11.)
- Fomel, S. [2007] Shaping regularization in geophysical-estimation problems. *Geophysics*, **72**(2), R29–R36. (cit. on p. 20.)
- Forgues, E. and Lambaré, G. [1997] Parameterization study for acoustic and elastic ray plus born inversion. *Journal of Seismic Exploration*, **6**(2–3), 253–277. (cit. on pp. 20 and 77.)
- Fornberg, B. [1988] Generation of finite difference formulas on arbitrarily spaced grids. *Mathematics of computation*, **51**(184), 699–706. (cit. on p. 40.)
- French, W.S. [1975] Computer migration of oblique seismic reflection profiles. *Geophysics*, **40**(6), 961–980. (cit. on p. 17.)
- Fu, L. and Symes, W.W. [2015] Reducing the cost of extended waveform inversion by multiscale adaptive methods. In: *SEG Technical Program Expanded Abstracts 2015*. 1127–1131. (cit. on pp. 66 and 123.)
- Gallagher, K., Sambridge, M. and Drijkoningen, G. [1991] Genetic algorithms: An evolution from Monte Carlo Methods for strongly non-linear geophysical optimization problems. *Geophysical Research Letters*, **18**(12), 2177–2180. (cit. on p. 11.)
- Gao, F. and Symes, W. [2009] Differential semblance velocity analysis by reverse time

- migration: Image gathers and theory. In: *SEG Technical Program Expanded Abstracts 2009*. 2317–2321. (cit. on p. 17.)
- Gauthier, O., Virieux, J. and Tarantola, A. [1986] Two-dimensional nonlinear inversion of seismic waveforms: Numerical results. *Geophysics*, **51**(7), 1387–1403. (cit. on pp. 14, 46, and 122.)
- Gray, S.H. [1992] FREQUENCY-SELECTIVE DESIGN OF THE KIRCHHOFF MIGRATION OPERATOR. *Geophysical prospecting*, **40**(5), 565–571. (cit. on p. 17.)
- Guasch, L. and Warner, M. [2014] Adaptive Waveform Inversion - FWI Without Cycle Skipping - Applications. In: *74th EAGE Conference & Exhibition*. We E106 14. (cit. on p. 49.)
- Guitton, A. [2002] Shot-Profile Migration of Multiple Reflections. In: *SEG Technical Program Expanded Abstracts 2002*. 1296–1299. (cit. on pp. 6 and 17.)
- Guitton, A. and Symes, W.W. [2003] Robust Inversion of Seismic Data Using the Huber Norm. *Geophysics*, **68**(4), 1310–1319. (cit. on p. 10.)
- Guo, Q. and Alkhalifah, T. [2016] Elastic Reflection Based Waveform Inversion in Isotropic Media. In: *78th EAGE Conference & Exhibition*. We SRS2 07. (cit. on p. 24.)
- Hak, B. and Mulder, W.A. [2010] Migration for velocity and attenuation perturbations. *Geophysical Prospecting*, **58**(6), 939–952. (cit. on p. 10.)
- Hansen, P.C. [2000] The L-Curve and its Use in the Numerical Treatment of Inverse Problems. In: *in Computational Inverse Problems in Electrocardiology*. WIT Press, 119–142. (cit. on p. 10.)
- Hardage, B.A. [1985] Vertical seismic profiling. *The Leading Edge*, **4**(11), 59–59. (cit. on p. 4.)
- He, W. and Plessix, R.É. [2017] Analysis of different parameterisations of waveform inversion of compressional body waves in an elastic transverse isotropic Earth with a vertical axis of symmetry. *Geophysical Prospecting*, **65**(4), 1004–1024. (cit. on pp. 10 and 14.)
- Hou, J. and Symes, W. [2016a] Inversion velocity analysis via approximate Born inversion. In: *SEG Technical Program Expanded Abstracts 2016*. 5274–5279. (cit. on pp. 28, 29, 69, 78, and 138.)
- Hou, J. and Symes, W. [2017] True amplitude imaging for incomplete seismic data. In: *SEG Technical Program Expanded Abstracts 2017*. 4534–4539. (cit. on p. 77.)
- Hou, J. and Symes, W.W. [2015] An approximate inverse to the extended Born modeling operator. *Geophysics*, **80**(6), R331–R349. (cit. on pp. 20, 21, 22, 67, 69, 77, 81, 99, 108, 112, 113, and 123.)
- Hou, J. and Symes, W.W. [2016b] Accelerating extended least-squares migration with weighted conjugate gradient iteration. *Geophysics*, **81**, S165–S179. (cit. on pp. 20, 67, 69, 77, 78, 109, and 175.)
- Huang, Y. and Symes, W.W. [2015] Born waveform inversion via variable projection and shot record model extension. In: *SEG Technical Program Expanded Abstracts*

2015. 1326–1331. (cit. on pp. 26, 62, and 69.)
- Ingber, L. [1989] Very Fast Simulated Re-Annealing. *Mathematical and Computer Modelling*, **12**(8), 967–973. (cit. on p. 11.)
- 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., Trezeguet, D. and Xie, M. [1989] Wavelengths of Earth Structures That Can Be Resolved from Seismic Reflection Data. *Geophysics*, **54**(7), 906–910. (cit. on p. 6.)
- Jiang, Z., Sheng, J., Yu, J., Schuster, G.T. and Hornby, B.E. [2007] Migration Methods for Imaging Different-Order Multiples. *Geophysical Prospecting*, **55**(1), 1–19. (cit. on p. 6.)
- Jiang, Z., Yu, J., Schuster, G.T. and Hornby, B.E. [2005] Migration of Multiples. *The Leading Edge*, **24**(3), 315–318. (cit. on pp. 6 and 17.)
- Jin, S. and Madariaga, R. [1993] Background velocity inversion with a genetic algorithm. *Geophysical research letters*, **20**(2), 93–96. (cit. on p. 11.)
- Jin, S. and Madariaga, R. [1994] Nonlinear velocity inversion by a two-step Monte Carlo method. *Geophysics*, **59**(4), 577–590. (cit. on p. 11.)
- Joncour, F., Lambaré, G. and Svay-Lucas, J. [2011] Preserved-amplitude angle domain migration by shot-receiver wavefield continuation. *Geophysical Prospecting*, **59**(2), 256–268. (cit. on pp. 20 and 22.)
- Jones, I. [2014] Tutorial: migration imaging conditions. *First Break*, **32**(12), 45–55. (cit. on p. 80.)
- Julian, B. [1977] Three-dimensional seismic ray tracing. *Journal of Geophysics*, **43**, 95–113. (cit. on p. 10.)
- Komatitsch, D. and Martin, R. [2007] An unsplit convolutional perfectly matched layer improved at grazing incidence for the seismic wave equation. *Geophysics*, **72**(5), SM155–SM167. (cit. on pp. 34, 43, 84, and 129.)
- Komatitsch, D. and Tromp, J. [1999] Introduction to the spectral element method for three-dimensional seismic wave propagation. *Geophysical journal international*, **139**(3), 806–822. (cit. on p. 9.)
- Komatitsch, D. and Tromp, J. [2003] A perfectly matched layer absorbing boundary condition for the second-order seismic wave equation. *Geophysical Journal International*, **154**(1), 146–153. (cit. on p. 43.)
- Komatitsch, D. and Vilotte, J.P. [1998] The spectral element method: An efficient tool to simulate the seismic response of 2D and 3D geological structures. *Bulletin of the seismological society of America*, **88**(2), 368–392. (cit. on p. 9.)
- Kreiss, H., Kreiss, H.O. and Olinger, J. [1973] *Methods for the approximate solution of time dependent problems*. International Council of Scientific Unions, World Meteorological Organization. (cit. on p. 40.)
- ten Kroode, A., Smit, D.J. and Verdel, A. [1998] A microlocal analysis of migration. *Wave Motion*, **28**(2), 149–172. (cit. on p. 153.)
- ten Kroode, F. [2012] A wave-equation-based Kirchhoff operator. *Inverse Problems*,

- 28(11), 115013. (cit. on pp. 20, 21, 22, 67, 69, 77, 81, 108, 113, and 123.)
- Kühl, H. and Sacchi, M.D. [2003] Least-squares wave-equation migration for AVP/AVA inversion. *Geophysics*, **68**(1), 262–273. (cit. on p. 20.)
- Lailly, P. et al. [1983] The seismic inverse problem as a sequence of before stack migrations. In: *Inverse Scattering: Theory and Application*. SIAM, 206–220. (cit. on pp. 13, 18, 31, 43, 44, 76, 79, and 109.)
- Lambaré, G. [2008] Stereotomography. *Geophysics*, **73**(5), VE25–VE34. (cit. on pp. 23, 27, 32, 124, 170, and 171.)
- Lambaré, G., Alerini, M., Baina, R. and Podvin, P. [2004] Stereotomography: a semi-automatic approach for velocity macromodel estimation. *Geophysical Prospecting*, **52**(6), 671–681. (cit. on p. 27.)
- Lambaré, G., Guillaume, P. and Montel, J.P. [2014] Recent Advances in Ray-Based Tomography. In: *76th EAGE Conference & Exhibition*. We G103 01. (cit. on pp. xi, 8, 12, and 23.)
- Lambaré, G., Operto, S., Podvin, P. and Thierry, P. [2003] 3D ray+ Born migration/inversion—Part 1: Theory. *Geophysics*, **68**(4), 1348–1356. (cit. on p. 17.)
- Lambaré, G., Virieux, J., Madariaga, R. and Jin, S. [1992] Iterative asymptotic inversion in the acoustic approximation. *Geophysics*, **57**(9), 1138–1154. (cit. on pp. 20, 22, 67, 77, 109, and 112.)
- Lameloise, C. [2015] *Quantitative migration for velocity analysis and introduction of transmitted waves*. Ph.D. thesis, MINES ParisTech. (cit. on pp. xiii, 20, 21, 22, 28, 29, 30, 66, 67, 68, and 76.)
- Lameloise, C.A. and Chauris, H. [2014] Quantitative Migration for a More Robust Migration Velocity Analysis. In: *76th EAGE Conference & Exhibition*. Amsterdam, Th G103 13. (cit. on pp. 65 and 67.)
- Lameloise, C.A. and Chauris, H. [2016] Extension of migration velocity analysis to transmitted wavefields. *Geophysical Journal International*, **207**(1), 343–356. (cit. on pp. 29, 65, 76, and 174.)
- Lameloise, C.A., Chauris, H. and 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**(2), 391–404. (cit. on pp. 28, 31, 63, 68, 76, 78, 112, 123, and 174.)
- Lay, T. and Wallace, T.C. [1995] *Modern global seismology*. Academic Press. (cit. on p. 3.)
- van Leeuwen, T., Kumar, R. and Herrmann, F. [2015] Affordable Full Subsurface Image Volume - An Application to WEMVA. In: *77th EAGE Conference & Exhibition*. WS05–C01. (cit. on pp. 66, 76, and 123.)
- Levander, A.R. [1988] Fourth-order finite-difference P-SV seismograms. *Geophysics*, **53**(11), 1425–1436. (cit. on pp. 9, 40, 84, and 129.)
- Li, Y. and Chauris, H. [2017] An Alternative True-amplitude Common-shot Reverse Time Migration. In: *79th EAGE Conference & Exhibition*. We P3 05. (cit. on p. 21)

- and 22.)
- Lines, L. and Treitel, S. [1984] Tutorial: A review of least-squares inversion and its application to geophysical problems. *Geophysical prospecting*, **32**(2), 159–186. (cit. on p. 11.)
- Liu, F., Zhang, G., Morton, S.A. and Leveille, J.P. [2011a] An effective imaging condition for reverse-time migration using wavefield decomposition. *Geophysics*, **76**(1), S29–S39. (cit. on p. 81.)
- Liu, J., Chauris, H. and Calandra, H. [2011b] The Normalized Integration Method - An Alternative to Full Waveform Inversion? In: *Near Surface 2011 - 17th EAGE European Meeting of Environmental and Engineering Geophysics*. B07. (cit. on pp. 22, 51, 52, 53, and 122.)
- Liu, Y., Li, Z. and Symes, W.W. [2014a] Extended reflection waveform inversion via differential semblance optimization. In: *SEG Technical Program Expanded Abstracts 2014*. 1232–1237. (cit. on pp. 31, 123, 170, and 171.)
- Liu, Y., Symes, W. and Li, Z. [2014b] Inversion velocity analysis via Differential Semblance Optimization. In: *76th EAGE Conference & Exhibition*. Th P02–07. (cit. on pp. 66, 69, 77, 78, and 123.)
- Liu, Y., Symes, W.W. and Li, Z. [2013] Multisource least-squares extended reverse time migration with preconditioning guided gradient method. In: *SEG Technical Program Expanded Abstracts 2013*. 3709–3715. (cit. on p. 20.)
- Luo, J. and Wu, R.S. [2015] Seismic envelope inversion: reduction of local minima and noise resistance. *Geophysical Prospecting*, **63**(3), 597–614. (cit. on p. 51.)
- Luo, S. and Sava, P. [2011] A deconvolution-based objective function for wave-equation inversion. In: *SEG Technical Program Expanded Abstracts 2011*. 2788–2792. (cit. on pp. 10, 22, 48, and 49.)
- Luo, Y., Al-Dossary, S., Marhoon, M. and Alfaraj, M. [2003] Generalized Hilbert transform and its applications in geophysics. *The Leading Edge*, **22**(3), 198–202. (cit. on p. 51.)
- Luo, Y., Ma, Y., Wu, Y., Liu, H. and Cao, L. [2016] Full-traveltime inversion. *Geophysics*, **81**(5), R261–R274. (cit. on p. 48.)
- Luo, Y. and Schuster, G.T. [1991] Wave-equation traveltime inversion. *Geophysics*, **56**(5), 645–653. (cit. on pp. 10, 11, 23, and 47.)
- Luo, Y., Tromp, J., Denel, B. and Calandra, H. [2013] 3D coupled acoustic-elastic migration with topography and bathymetry based on spectral-element and adjoint methods. *Geophysics*, **78**(4), S193–S202. (cit. on p. 9.)
- Luo, Y., Zhu, H., Nissen-Meyer, T., Morency, C. and Tromp, J. [2009] Seismic modeling and imaging based upon spectral-element and adjoint methods. *The Leading Edge*, **28**(5), 568–574. (cit. on p. 9.)
- MacLeod, M., Hanson, R., Bell, C. and McHugo, S. [1999] The Alba Field ocean bottom cable seismic survey: Impact on development. *The Leading Edge*, **18**(11), 1306–1312. (cit. on p. 4.)

- Marfurt, K.J. [1984] Accuracy of finite-difference and finite-element modeling of the scalar and elastic wave equations. *Geophysics*, **49**(5), 533–549. (cit. on pp. 9, 40, and 41.)
- McMechan, G. [1983] Migration by extrapolation of time-dependent boundary values. *Geophysical Prospecting*, **31**(3), 413–420. (cit. on p. 17.)
- Mendel, J.M. [2013] *Optimal seismic deconvolution: an estimation-based approach*. Elsevier. (cit. on p. 45.)
- Menke, W. [1984] The resolving power of cross-borehole tomography. *Geophysical Research Letters*, **11**(2), 105–108. (cit. on pp. 10 and 45.)
- Métivier, L., Brossier, R., Méridot, Q., Oudet, E. and Virieux, J. [2016] Measuring the misfit between seismograms using an optimal transport distance: application to full waveform inversion. *Geophysical Supplements to the Monthly Notices of the Royal Astronomical Society*, **205**(1), 345–377. (cit. on pp. 10, 22, and 23.)
- Métivier, L., Brossier, R., Virieux, J. and Operto, S. [2013] Full waveform inversion and the truncated Newton method. *SIAM Journal on Scientific Computing*, **35**(2), B401–B437. (cit. on pp. 11 and 65.)
- Miller, D., Oristaglio, M. and Beylkin, G. [1987] A New Slant on Seismic Imaging: Migration and Integral Geometry. *Geophysics*, **52**(7), 943–964. (cit. on p. 13.)
- Misra, S. and Sacchi, M.D. [2008] Global optimization with model-space preconditioning: Application to AVO inversion. *Geophysics*, **73**(5), R71–R82. (cit. on pp. 11 and 14.)
- Moczo, P., Kristek, J., Galis, M., Chaljub, E. and Etienne, V. [2011] 3-D finite-difference, finite-element, discontinuous-Galerkin and spectral-element schemes analysed for their accuracy with respect to P-wave to S-wave speed ratio. *Geophysical Journal International*, **187**(3), 1645–1667. (cit. on p. 9.)
- Moczo, P., Kristek, J., Galis, M. and Pazak, P. [2010] On accuracy of the finite-difference and finite-element schemes with respect to P-wave to S-wave speed ratio. *Geophysical Journal International*, **182**(1), 493–510. (cit. on p. 9.)
- Montel, J. and Lambaré, G. [2013] Asymptotic Analysis of ADCIG from Slant Stacked Subsurface Offset Gather. In: *75th EAGE Conference & Exhibition*. Tu P06 12. (cit. on p. 28.)
- Mora, P. [1989] Inversion= migration+ tomography. *Geophysics*, **54**(12), 1575–1586. (cit. on pp. 6, 13, 40, and 122.)
- Mosegaard, K. and Vestergaard, P.D. [1991] A simulated annealing approach to seismic model optimization with sparse prior information. *Geophysical Prospecting*, **39**(5), 599–611. (cit. on p. 11.)
- Moser, T. [1991] Shortest path calculation of seismic rays. *Geophysics*, **56**(1), 59–67. (cit. on p. 12.)
- Muijs, R., Robertsson, J.O. and Holliger, K. [2007] Prestack depth migration of primary and surface-related multiple reflections: Part I—Imaging. *Geophysics*, **72**(2), S59–S69. (cit. on p. 17.)

- Mulder, W. [2008] Automatic velocity analysis with the two-way wave equation. In: *70th EAGE Conference & Exhibition*. P165. (cit. on pp. 17, 25, and 26.)
- Mulder, W. [2014] Subsurface offset behaviour in velocity analysis with extended reflectivity images. *Geophysical Prospecting*, **62**(1), 17–33. (cit. on pp. 28, 30, 61, 65, 66, 76, and 78.)
- Mulder, W. and Hak, B. [2009] An ambiguity in attenuation scattering imaging. *Geophysical Journal International*, **178**(3), 1614–1624. (cit. on p. 10.)
- Mulder, W. and ten Kroode, A. [2002] Automatic velocity analysis by differential semblance optimization. *Geophysics*, **67**(4), 1184–1191. (cit. on pp. 26, 28, 30, 32, 48, 61, 62, 65, 75, 78, 122, and 152.)
- Mulder, W.A. and Plessix, R.E. [2004] A comparison between one-way and two-way wave-equation migration. *Geophysics*, **69**(6), 1491–1504. (cit. on p. 17.)
- Nash, S.G. and Nocedal, J. [1991] A numerical study of the limited memory BFGS method and the truncated-Newton method for large scale optimization. *SIAM Journal on Optimization*, **1**(3), 358–372. (cit. on pp. 11 and 65.)
- Nemeth, T., Wu, C. and Schuster, G.T. [1999] Least-squares migration of incomplete reflection data. *Geophysics*, **64**(1), 208–221. (cit. on pp. 20, 77, 173, and 175.)
- Nichols, D. [2012] Resolution in Seismic Inversion-Spectral Gap or Spectral Overlap, Which is Harder to Handle? In: *74th EAGE Conference & Exhibition*. E12. (cit. on pp. 8 and 12.)
- Nocedal, J. [1980] Updating quasi-Newton matrices with limited storage. *Mathematics of computation*, **35**(151), 773–782. (cit. on p. 11.)
- Nolan, C.J. and Symes, W.W. [1996] Imaging and coherency in complex structures. In: *SEG Technical Program Expanded Abstracts 1996*. 359–362. (cit. on pp. 27, 84, 90, 99, 108, and 138.)
- Operto, M.S., Xu, S. and Lambaré, G. [2000] Can we quantitatively image complex structures with rays? *Geophysics*, **65**(4), 1223–1238. (cit. on p. 17.)
- Operto, S., Gholami, Y., Prioux, V., Ribodetti, A., Brossier, R., Metivier, L. and Virieux, J. [2013] A guided tour of multiparameter full-waveform inversion with multicomponent data: From theory to practice. *The Leading Edge*, **32**(9), 1040–1054. (cit. on p. 14.)
- Operto, S., Lambaré, G., Podvin, P. and Thierry, P. [2003] 3D ray+ Born migration/inversion—Part 2: Application to the SEG/EAGE overthrust experiment. *Geophysics*, **68**(4), 1357–1370. (cit. on p. 17.)
- Operto, S., Virieux, J., Amestoy, P., L'Excellent, J., Giraud, L. and Ali, H. [2007] 3D Finite-Difference Frequency-Domain Modeling of Visco-Acoustic Wave Propagation Using a Massively Parallel Direct Solver: A Feasibility Study. *Geophysics*, **72**, SM195–SM211. (cit. on p. 40.)
- Op't Root, T.J.P.M., Stolk, C.C. and Maarten, V. [2012] Linearized inverse scattering based on seismic reverse time migration. *Journal de mathématiques pures et appliquées*, **98**(2), 211–238. (cit. on pp. 20, 22, and 77.)

- Østmo, S., Mulder, W. and Plessix, R.E. [2002] Finite-difference iterative migration by linearized waveform inversion in the frequency domain. In: *SEG Technical Program Expanded Abstracts 2002*. 1384–1387. (cit. on p. 18.)
- Pérez Solano, C.A. [2013] *Two-dimensional near-surface seismic imaging with surface waves: alternative methodology for waveform inversion*. Ph.D. thesis, Ecole Nationale Supérieure des Mines de Paris. (cit. on p. 4.)
- Plessix, R.E. [2006] A review of the adjoint-state method for computing the gradient of a functional with geophysical applications. *Geophysical Journal International*, **167**(2), 495–503. (cit. on pp. 11, 34, 44, 45, 57, 61, 82, 117, 127, and 166.)
- Plessix, R.E., Baeten, G., de Maag, J.W., Klaassen, M., Rujie, Z. and Zhifei, T. [2010] Application of acoustic full waveform inversion to a low-frequency large-offset land data set. In: *SEG Technical Program Expanded Abstracts 2010*. 930–934. (cit. on p. 15.)
- Plessix, R.É., De Roeck, Y.H. and Chavent, G. [1995] Automatic and Simultaneous Migration Velocity Analysis and Waveform Inversion of Real Data Using a MBTT/WKB J Formulation. In: *SEG Technical Program Expanded Abstracts 1995*. 1099–1102. (cit. on pp. 23, 56, 124, 133, and 165.)
- Plessix, R.E., De Roeck, Y.H. and Chavent, G. [1998] Waveform inversion of reflection seismic data for kinematic parameters by local optimization. *SIAM Journal on Scientific Computing*, **20**(3), 1033–1052. (cit. on p. 24.)
- Plessix, R.E., Mulder, W. and ten Kroode, A. [2000] Automatic cross-well tomography by semblance and differential semblance optimization: theory and gradient computation. *Geophysical Prospecting*, **48**(5), 913–935. (cit. on p. 4.)
- Plessix, R.E. and Perkins, C. [2010] Thematic Set: Full waveform inversion of a deep water ocean bottom seismometer dataset. *First Break*, **28**(4), 71–78. (cit. on p. 4.)
- Podvin, P. and Lecomte, I. [1991] Finite difference computation of traveltimes in very contrasted velocity models: a massively parallel approach and its associated tools. *Geophysical Journal International*, **105**(1), 271–284. (cit. on p. 10.)
- Portniaguine, O. and Zhdanov, M.S. [1999] Focusing geophysical inversion images. *Geophysics*, **64**(3), 874–887. (cit. on p. 11.)
- Pratt, R., Sirgue, L., Hornby, B. and Wolfe, J. [2008] Crosswell waveform tomography in fine-layered sediments-Meeting the challenges of anisotropy. In: *70th EAGE Conference & Exhibition*. F020. (cit. on pp. 12, 15, 46, and 122.)
- Pratt, R., Song, Z.M., Williamson, P. and Warner, M. [1996] Two-dimensional velocity models from wide-angle seismic data by wavefield inversion. *Geophysical Journal International*, **124**(2), 323–340. (cit. on pp. 15 and 44.)
- Pratt, R.G., Shin, C. and Hick, G. [1998] Gauss–Newton and full Newton methods in frequency–space seismic waveform inversion. *Geophysical Journal International*, **133**(2), 341–362. (cit. on pp. 10, 11, 13, and 121.)
- Pratt, R.G. and Worthington, M. [1990] Inverse theory applied to multi-source cross-hole tomography. Part 1: Acoustic wave-equation method. *Geophysical prospecting*,

- 38(3), 287–310. (cit. on pp. 13 and 40.)
- Prieux, V., Brossier, R., Operto, S. and Virieux, J. [2013a] Multiparameter full waveform inversion of multicomponent ocean-bottom-cable data from the Valhall field. Part 1: Imaging compressional wave speed, density and attenuation. *Geophysical Journal International*, **194**(3), 1640–1664. (cit. on p. 14.)
- Prieux, V., Lambaré, G., Operto, S. and Virieux, J. [2013b] Building starting models for full waveform inversion from wide-aperture data by stereotomography. *Geophysical Prospecting*, **61**(s1), 109–137. (cit. on p. 27.)
- Priolo, E. and Chiaruttini, C. [2003] Analytical and numerical analysis of tomographic resolution with band-limited signals. *Geophysics*, **68**(2), 600–613. (cit. on p. 23.)
- Prucha, M.L., Biondi, B.L. and Symes, W.W. [1999] Angle-domain common image gathers by wave-equation migration. In: *SEG Technical Program Expanded Abstracts 1999*. 824–827. (cit. on pp. 27, 28, 84, 90, 99, and 138.)
- Pyun, S. and Shin, C. [2008] Evaluation of Kirchhoff hyperbola in terms of partial derivative wavefield and virtual source. *Journal of Applied Geophysics*, **65**(1), 50–55. (cit. on pp. xi and 17.)
- Qin, B., Allemand, T. and Lambaré, G. [2015] Full waveform inversion using preserved amplitude reverse time migration. In: *SEG Technical Program Expanded Abstracts 2015*. 1252–1257. (cit. on pp. xvii, 20, 21, 22, 33, 67, 77, 78, 79, 80, 81, 82, 108, 109, 113, 114, 115, 116, 125, 171, and 175.)
- Qin, B. and Lambaré, G. [2016] Joint inversion of velocity and density in preserved-amplitude full-waveform inversion. In: *SEG Technical Program Expanded Abstracts 2016*. 1325–1330. (cit. on p. 113.)
- Qin, F., Luo, Y., Olsen, K.B., Cai, W. and Schuster, G.T. [1992] Finite-difference solution of the eikonal equation along expanding wavefronts. *Geophysics*, **57**(3), 478–487. (cit. on p. 10.)
- Rector, J.W. [1995] Crosswell methods: Where are we, where are we going? *Geophysics*, **60**(3), 629–630. (cit. on p. 4.)
- Ribodetti, A. and Virieux, J. [1998] Asymptotic theory for imaging the attenuation factor Q . *Geophysics*, **63**(5), 1767–1778. (cit. on pp. 20 and 77.)
- Rickett, J.E. and Sava, P.C. [2002] Offset and angle-domain common image-point gathers for shot-profile migration. *Geophysics*, **67**(3), 883–889. (cit. on pp. 26, 28, 62, and 63.)
- Robertsson, J.O., Blanch, J.O. and Symes, W.W. [1994] Viscoelastic finite-difference modeling. *Geophysics*, **59**(9), 1444–1456. (cit. on p. 40.)
- Rocha, D., Tanushev, N. and Sava, P. [2016] Acoustic wavefield imaging using the energy norm. *Geophysics*, **81**(4), S151–S163. (cit. on p. 108.)
- Ronen, S. and Liner, C.L. [2000] Least-squares DMO and migration. *Geophysics*, **65**(5), 1364–1371. (cit. on p. 20.)
- Routh, P., Krebs, J., Lazaratos, S., Baumstein, A., Chikichev, I., Lee, S., Downey, N., Hinkley, D. and Anderson, J. [2011] Full-wavefield inversion of marine streamer data

- with the encoded simultaneous source method. In: *73rd EAGE Conference & Exhibition*. F032. (cit. on p. 48.)
- Sambridge, M. and Drijkoningen, G. [1992] Genetic Algorithms in Seismic Waveform Inversion. *Geophysical Journal International*, **109**(2), 323–342. (cit. on p. 11.)
- Sambridge, M. and Mosegaard, K. [2002] Monte Carlo methods in geophysical inverse problems. *Reviews of Geophysics*, **40**(3). (cit. on pp. 11 and 14.)
- Santosa, F., Symes, W. and Raggio, G. [1987] Inversion of band-limited reflection seismograms using stacking velocities as constraints. *Inverse Problems*, **3**(3), 477. (cit. on p. 11.)
- Sava, P. and Biondi, B. [2004] Wave-equation migration velocity analysis. I. Theory. *Geophysical Prospecting*, **52**(6), 593–606. (cit. on pp. 10, 25, 31, 61, 75, 76, 78, 124, 170, and 171.)
- Sava, P. and Fomel, S. [2006] Time-shift imaging condition in seismic migration. *Geophysics*, **71**(6), S209–S217. (cit. on pp. 28, 63, 76, 123, 152, and 153.)
- Sava, P. and Vasconcelos, I. [2011] Extended imaging conditions for wave-equation migration. *Geophysical Prospecting*, **59**(1), 35–55. (cit. on pp. 28, 60, 64, and 123.)
- Sava, P.C. and Fomel, S. [2003] Angle-domain common-image gathers by wavefield continuation methods. *Geophysics*, **68**(3), 1065–1074. (cit. on pp. 28, 63, and 123.)
- Scales, J.A., Docherty, P. and Gersztenkorn, A. [1990] Regularisation of Nonlinear Inverse Problems: Imaging the near-Surface Weathering Layer. *Inverse Problems*, **6**(1), 115. (cit. on p. 10.)
- Scales, J.A., Smith, M.L. and Fischer, T.L. [1992] Global optimization methods for multimodal inverse problems. *Journal of Computational Physics*, **103**(2), 258–268. (cit. on p. 11.)
- Schleicher, J., Costa, J.C. and Novais, A. [2008] A comparison of imaging conditions for wave-equation shot-profile migration. *Geophysics*, **73**(6), S219–S227. (cit. on p. 81.)
- Schneider, W.A. [1978] Integral formulation for migration in two and three dimensions. *Geophysics*, **43**(1), 49–76. (cit. on p. 17.)
- Schuster, G.T. [2016] Seismic interferometry. In: *Encyclopedia of Exploration Geophysics*. Q1–1. (cit. on p. 3.)
- Schuster, G.T., Yu, J., Sheng, J. and Rickett, J. [2004] Interferometric/Daylight Seismic Imaging. *Geophysical Journal International*, **157**(2), 838–852. (cit. on p. 3.)
- Sen, M.K. and Stoffa, P.L. [2013] *Global optimization methods in geophysical inversion*. Cambridge University Press. (cit. on pp. 11 and 14.)
- Shapiro, N.M., Campillo, M., Stehly, L. and Ritzwoller, M.H. [2005] High-resolution surface-wave tomography from ambient seismic noise. *Science*, **307**(5715), 1615–1618. (cit. on p. 4.)
- Shen, P. and Symes, W.W. [2008] Automatic velocity analysis via shot profile migration. *Geophysics*, **73**(5), VE49–VE59. (cit. on pp. 26, 28, 30, 61, 64, and 65.)
- Shen, P. and Symes, W.W. [2013] Subsurface Domain Image Warping by Horizontal

- Contraction and Its Application to Wave-Equation Migration Velocity Analysis. In: *SEG Technical Program Expanded Abstracts 2013*. 4715–4719. (cit. on pp. 29, 61, 65, 76, and 78.)
- Shen, P. and Symes, W.W. [2015] Horizontal contraction in image domain for velocity inversion. *Geophysics*, **80**(3), R95–R110. (cit. on pp. 61, 65, 68, 77, 78, 87, and 109.)
- Shen, P., Symes, W.W., Morton, S., Hess, A. and Calandra, H. [2005] Differential semblance velocity analysis via shot profile migration. In: *SEG Technical Program Expanded Abstracts 2005*. 2249–2252. (cit. on pp. 28 and 63.)
- Sheriff, R.E. and Geldart, L.P. [1995] *Exploration seismology*. Cambridge university press. (cit. on p. 4.)
- Shin, C. [1988] *Nonlinear elastic wave inversion by blocky representations*. Ph.D. thesis, University of Oklahoma. (cit. on p. 11.)
- Shipp, R.M. and Singh, S.C. [2002] Two-dimensional full wavefield inversion of wide-aperture marine seismic streamer data. *Geophysical Journal International*, **151**(2), 325–344. (cit. on p. 15.)
- Silvestrov, I., Baina, R. and Landa, E. [2016] Poststack diffraction imaging using reverse-time migration. *Geophysical Prospecting*, **64**(1), 129–142. (cit. on p. 28.)
- Sirgue, L. [2006] The importance of low frequency and large offset in waveform inversion. In: *68th EAGE Conference & Exhibition*. A037. (cit. on pp. 12, 15, and 122.)
- Sirgue, L., Barkved, O., Van Gestel, J., Askim, O. and Kommedal, J. [2009] 3D waveform inversion on Valhall wide-azimuth OBC. In: *71st EAGE Conference & Exhibition*. U038. (cit. on p. 15.)
- Sirgue, L. and Pratt, R.G. [2004] Efficient waveform inversion and imaging: A strategy for selecting temporal frequencies. *Geophysics*, **69**(1), 231–248. (cit. on pp. 15, 40, and 44.)
- Smith, W.D. [1975] The application of finite element analysis to body wave propagation problems. *Geophysical Journal International*, **42**(2), 747–768. (cit. on pp. 9 and 40.)
- Socco, L. and Strobbia, C. [2004] Surface-wave method for near-surface characterization: A tutorial. *Near Surface Geophysics*, **2**(4), 165–185. (cit. on p. 4.)
- Soni, A. [2014] *Full Wavefield Migration of Vertical Seismic Profiling Data*. Ph.D. thesis, Delft University of Technology. (cit. on p. 4.)
- Staal, X. and Verschuur, D. [2012] Velocity estimation using internal multiples. In: *SEG Technical Program Expanded Abstracts 2012*. 1–5. (cit. on pp. 24, 65, and 76.)
- Staal, X.R. [2015] *Combined imaging and velocity estimation by Joint Migration Inversion*. Ph.D. thesis, TU Delft University. (cit. on p. 24.)
- Stolk, C.C. and de Hoop, M.V. [2005] Modeling of seismic data in the downward continuation approach. *SIAM Journal on Applied Mathematics*, **65**(4), 1388–1406. (cit. on p. 63.)
- Stolk, C.C., de Hoop, M.V. and Symes, W.W. [2009] Kinematics of shot-geophone migration. *Geophysics*, **74**(6), WCA19–WCA34. (cit. on pp. 75 and 108.)
- Stolk, C.C. and Symes, W.W. [2002] Smooth objective functionals for seismic velocity

- inversion. *Inverse Problems*, **19**(1), 73. (cit. on pp. 26 and 61.)
- Stolk, C.C. and Symes, W.W. [2004] Kinematic artifacts in prestack depth migration. *Geophysics*, **69**(2), 562–575. (cit. on pp. 28, 63, 69, 78, 84, 90, 99, 108, and 138.)
- Symes, B. [2017] Extended Waveform Inversion. In: *79th EAGE Conference & Exhibition*. WS09 A02. (cit. on pp. 29, 31, and 123.)
- Symes, W.W. [2008] Migration velocity analysis and waveform inversion. *Geophysical Prospecting*, **56**(6), 765–790. (cit. on pp. 10, 25, 30, 58, 60, 61, 64, 75, 76, 78, 79, 82, 122, 123, 152, and 153.)
- Symes, W.W. and Carazzone, J.J. [1991] Velocity inversion by differential semblance optimization. *Geophysics*, **56**(5), 654–663. (cit. on pp. 10, 24, 26, 31, 32, 60, 61, 62, 65, 75, 76, 122, 124, 150, 170, and 171.)
- Symes, W.W. and Kern, M. [1994] Inversion of reflection seismograms by differential semblance analysis: Algorithm structure and synthetic examples. *Geophysical Prospecting*, **42**(6), 565–614. (cit. on pp. 26, 60, 62, 70, 78, 124, 126, 150, 151, 170, and 171.)
- Tarantola, A. [1984a] Inversion of seismic reflection data in the acoustic approximation. *Geophysics*, **49**(8), 1259–1266. (cit. on pp. 10, 11, 13, 31, 40, 43, 44, 76, 79, 109, and 121.)
- Tarantola, A. [1984b] Linearized inversion of seismic reflection data. *Geophysical prospecting*, **32**(6), 998–1015. (cit. on p. 18.)
- Tarantola, A. [2005] *Inverse problem theory and methods for model parameter estimation*. SIAM. (cit. on pp. 4, 10, and 12.)
- Thierry, P., Operto, S. and Lambaré, G. [1998] Fast 2-D ray+ Born migration/inversion in complex media. *Geophysics*, **64**(1), 162–181. (cit. on p. 17.)
- Thorbecke, J. [2015] *2D Finite-Difference Wavefield Modelling*. (cit. on pp. xii, 41, and 42.)
- Tikhonov, A., Arsenin, V. and John, F. [1977] *Solutions of ill-posed problems*. Winston Washington, DC. (cit. on pp. 10 and 45.)
- Touati-Ahmed, D. and Storey, C. [1990] Efficient hybrid conjugate gradient techniques. *Journal of Optimization Theory and Applications*, **64**(2), 379–397. (cit. on pp. 97 and 137.)
- Van Leeuwen, T. and Mulder, W. [2009] A comparison of seismic velocity inversion methods for layered acoustics. *Inverse problems*, **26**(1), 015008. (cit. on pp. 26 and 61.)
- Van Leeuwen, T. and Mulder, W. [2010] A correlation-based misfit criterion for wave-equation travelttime tomography. *Geophysical Journal International*, **182**(3), 1383–1394. (cit. on pp. 10, 21, 22, 23, 24, 47, 48, and 122.)
- Červený, V. [2005] *Seismic ray theory*. Cambridge university press. (cit. on p. 9.)
- Červený, V., Molotkov, I.A., Molotkov, I.A. and Pšenčík, I. [1977] *Ray method in seismology*. Univerzita Karlova. (cit. on p. 9.)
- Verschuur, D.J. [2006] *Seismic multiple removal techniques: Past, present and future*.

- EAGE Publications bv. (cit. on p. 6.)
- Verschuur, D.J., Berkhout, A. and Wapenaar, C. [1992] Adaptive surface-related multiple elimination. *Geophysics*, **57**(9), 1166–1177. (cit. on p. 6.)
- Verschuur, D.J. and Berkhout, A.J. [2015] From Removing to Using Multiples in Closed-Loop Imaging. *The Leading Edge*, **34**(7), 744–759. (cit. on pp. 6 and 27.)
- Versteeg, R. [1994] The Marmousi experience: Velocity model determination on a synthetic complex data set. *The Leading Edge*, **13**(9), 927–936. (cit. on pp. 97 and 136.)
- Vidale, J. [1988] Finite-difference calculation of travel times. *Bulletin of the Seismological Society of America*, **78**(6), 2062–2076. (cit. on p. 10.)
- Virieux, J. [1986] P-SV wave propagation in heterogeneous media: Velocity-stress finite-difference method. *Geophysics*, **51**(4), 889–901. (cit. on pp. 9, 34, 40, 84, and 129.)
- Virieux, J., Calandra, H. and Plessix, R.É. [2011] A review of the spectral, pseudo-spectral, finite-difference and finite-element modelling techniques for geophysical imaging. *Geophysical Prospecting*, **59**(5), 794–813. (cit. on pp. 9 and 40.)
- Virieux, J. and Operto, S. [2009] An overview of full-waveform inversion in exploration geophysics. *Geophysics*, **74**(6), WCC1–WCC26. (cit. on pp. 6, 10, 13, 15, 16, 31, 43, 46, 76, 104, 122, and 141.)
- Vyas, M., Geco, W. and Tang, Y. [2010] Gradients for wave-equation migration velocity analysis. In: *SEG Technical Program Expanded Abstracts 2010*. 4077–4081. (cit. on pp. 68 and 76.)
- Wang, H., Singh, S.C., Audebert, F. and Calandra, H. [2015] Inversion of seismic refraction and reflection data for building long-wavelength velocity models. *Geophysics*, **80**(2), R81–R93. (cit. on p. 48.)
- Wang, H., Singh, S.C. and Calandra, H. [2014] Integrated inversion using combined wave-equation tomography and full waveform inversion. *Geophysical Journal International*, **198**(1), 430–446. (cit. on pp. 12 and 15.)
- Warner, M. and Guasch, L. [2014] Adaptive Waveform Inversion - FWI Without Cycle Skipping - Theory. In: *76th EAGE Conference & Exhibition*. We E106 13. (cit. on pp. 22, 49, and 122.)
- Warner, M. and Guasch, L. [2015] Robust adaptive waveform inversion. In: *SEG Technical Program Expanded Abstracts 2015*. 1059–1063. (cit. on p. 49.)
- Warner, M. and Guasch, L. [2016] Adaptive waveform inversion: Theory. *Geophysics*, **81**(6), R429–R445. (cit. on pp. 10, 22, and 49.)
- Warner, M., Nangoo, T., Shah, N., Umpleby, A. and Morgan, J. [2013] Full-waveform inversion of cycle-skipped seismic data by frequency down-shifting. In: *SEG Technical Program Expanded Abstracts 2013*. 903–907. (cit. on p. 14.)
- Woodward, M., Farmer, P., Nichols, D. and Charles, S. [1998] Automated 3D tomographic velocity analysis of residual moveout in prestack depth migrated common image point gathers. In: *SEG Technical Program Expanded Abstracts 1998*. 1218–1221. (cit. on p. 26.)

- Woodward, M.J. [1992] Wave-equation tomography. *Geophysics*, **57**(1), 15–26. (cit. on p. 47.)
- Woodward, M.J., Nichols, D., Zdraveva, O., Whitfield, P. and Johns, T. [2008] A decade of tomography. *Geophysics*, **73**(5), VE5–VE11. (cit. on pp. 12 and 23.)
- Wu, R.S., Luo, J. and Wu, B. [2014] Seismic envelope inversion and modulation signal model. *Geophysics*, **79**(3), WA13–WA24. (cit. on pp. 22, 51, and 122.)
- Wu, Z. and Alkhalifah, T. [2015] Simultaneous inversion of the background velocity and the perturbation in full-waveform inversion. *Geophysics*, **80**(6), R317–R329. (cit. on pp. 23, 57, 58, 83, 123, and 152.)
- Xia, J., Miller, R.D. and Park, C.B. [1999] Estimation of near-surface shear-wave velocity by inversion of Rayleigh waves. *Geophysics*, **64**(3), 691–700. (cit. on p. 4.)
- Xie, X.B., Wu, R.S., Fehler, M. and Huang, L. [2005] Seismic resolution and illumination: A wave-equation-based analysis. In: *SEG Technical Program Expanded Abstracts 2005*. 1862–1865. (cit. on p. 27.)
- Xie, X.B. and Yang, H. [2008] The finite-frequency sensitivity kernel for migration residual moveout and its applications in migration velocity analysis. *Geophysics*, **73**(6), S241–S249. (cit. on p. 26.)
- Xu, S., Chauris, H., Lambaré, G. and Noble, M. [2001] Common-angle migration: A strategy for imaging complex media. *Geophysics*, **66**(6), 1877–1894. (cit. on pp. 27, 28, 63, 84, 90, and 123.)
- Xu, S., Chen, F., Lambaré, G. and Zhang, Y. [2013] Full waveform inversion of reflected seismic data. *Journal of Seismic Exploration*, **22**(5), 449–462. (cit. on p. 31.)
- Xu, S. and Lambaré, G. [2004] Fast migration/inversion with multivalued ray fields: Part 1—Method, validation test, and application in 2D to Marmousi. *Geophysics*, **69**(5), 1311–1319. (cit. on pp. 22 and 77.)
- Xu, S., Wang, D., Chen, F., Lambaré, G. and Zhang, Y. [2012] Inversion on Reflected Seismic Wave. In: *SEG Technical Program Expanded Abstracts 2012*. 1–7. (cit. on pp. 23, 24, 56, 57, 58, 83, 123, 124, and 152.)
- Xue, Z., Chen, Y., Fomel, S. and Sun, J. [2014] Imaging incomplete data and simultaneous-source data using least-squares reverse-time migration with shaping regularization. In: *SEG Technical Program Expanded Abstracts 2014*. 3991–3996. (cit. on p. 20.)
- Yan, J. and Sava, P. [2008] Isotropic angle-domain elastic reverse-time migration. *Geophysics*, **73**(6), S229–S239. (cit. on p. 20.)
- Yang, T. and Sava, P. [2011] Wave-equation migration velocity analysis with time-shift imaging. *Geophysical prospecting*, **59**(4), 635–650. (cit. on pp. 28, 63, and 123.)
- Yang, T. and Sava, P. [2015] Image-domain wavefield tomography with extended common-image-point gathers. *Geophysical Prospecting*, **63**(5), 1086–1096. (cit. on p. 66.)
- Yang, T., Shragge, J. and Sava, P. [2013] Illumination compensation for image-domain wavefield tomography. *Geophysics*, **78**(5), U65–U76. (cit. on p. 66.)

- Yilmaz, Ö. [2001] *Seismic data analysis: Processing, inversion, and interpretation of seismic data*. Society of Exploration Geophysicists. (cit. on pp. 4, 6, and 25.)
- Zeng, C., Dong, S. and Wang, B. [2017] A guide to least-squares reverse time migration for subsalt imaging: Challenges and solutions. *Interpretation*, **5**(3), SN1–SN11. (cit. on pp. xii, 20, and 21.)
- Zhang, Y., Duan, L. and Xie, Y. [2014a] A stable and practical implementation of least-squares reverse time migration. *Geophysics*, **80**(1), V23–V31. (cit. on pp. 20, 22, and 48.)
- Zhang, Y., Ratcliffe, A., Roberts, G. and Duan, L. [2014b] Amplitude-preserving reverse time migration: From reflectivity to velocity and impedance inversion. *Geophysics*, **79**(6), S271–S283. (cit. on p. 77.)
- Zhang, Y. and Sun, J. [2009] Practical issues in reverse time migration: true amplitude gathers, noise removal and harmonic source encoding. *First Break*, **27**(1295), 204–204. (cit. on p. 21.)
- Zhang, Y., Xu, S., Bleistein, N. and Zhang, G. [2007] True-amplitude, angle-domain, common-image gathers from one-way wave-equation migrations. *Geophysics*, **72**(1), S49–S58. (cit. on pp. 20, 21, 22, and 77.)
- Zhang, Y., Zhang, G. and Bleistein, N. [2005] Theory of true-amplitude one-way wave equations and true-amplitude common-shot migration. *Geophysics*, **70**(4), E1–E10. (cit. on pp. 20, 22, 77, and 113.)
- Zhao, L., Jordan, T.H. and Chapman, C.H. [2000] Three-dimensional Fréchet differential kernels for seismic delay times. *Geophysical Journal International*, **141**(3), 558–576. (cit. on p. 23.)
- Zhou, C., Cai, W., Luo, Y., Schuster, G.T. and Hassanzadeh, S. [1995] Acoustic wave-equation traveltime and waveform inversion of crosshole seismic data. *Geophysics*, **60**(3), 765–773. (cit. on p. 4.)
- Zhou, W. [2016] *Velocity model building by full waveform inversion of early arrivals & reflections and case study with gas cloud effect*. Ph.D. thesis, Université Grenoble Alpes. (cit. on pp. 16, 20, 23, 24, 56, 122, 124, and 173.)
- Zhou, W., Brossier, R., Operto, S. and 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. (cit. on pp. 10, 23, 56, 58, 122, 123, and 174.)

Résumé

Les expériences sismiques actives sont largement utilisées pour caractériser la structure de la subsurface. Les méthodes dites d'analyse de vitesse par migration ont pour but la détermination d'un macro-modèle de vitesse, lisse, et contrôlant la cinématique de propagation des ondes. Le modèle est estimé par des critères de cohérence d'image ou de focalisation d'image. Les images de réflectivité obtenues par les techniques de migration classiques sont cependant contaminées par des artefacts, altérant la qualité de la remise à jour du macro-modèle. Des résultats récents proposent de coupler l'inversion asymptotique, qui donne des images beaucoup plus propres en pratique, avec l'analyse de vitesse pour la version offset en profondeur. Cette approche cependant demande des capacités de calcul et de mémoire importantes et ne peut actuellement être étendue en 3D.

Dans ce travail, je propose de développer le couplage entre l'analyse de vitesse et la migration plus conventionnelle par point de tir. La nouvelle approche permet de prendre en compte des modèles de vitesse complexes, comme par exemple en présence d'anomalies de vitesses plus lentes ou de réflectivités discontinues. C'est une alternative avantageuse en termes d'implémentation et de coût numérique par rapport à la version profondeur. Je propose aussi d'étendre l'analyse de vitesse par inversion au domaine des données pour les cas par point de tir. J'établis un lien entre les méthodes formulées dans les domaines données et images. Les méthodologies sont développées et analysées sur des données synthétiques 2D.

Mots Clés

imagerie sismique • problème inverse • analyse de vitesse par migration quantitative • point de tir • domaine donnée

Abstract

Active seismic experiments are widely used to characterize the structure of the subsurface. Migration Velocity Analysis techniques aim at recovering the background velocity model controlling the kinematics of wave propagation. The first step consists of obtaining the reflectivity images by migrating observed data in a given macro velocity model. The estimated model is then updated, assessing the quality of the background velocity model through the image coherency or focusing criteria. Classical migration techniques, however, do not provide a sufficiently accurate reflectivity image, leading to incorrect velocity updates. Recent investigations propose to couple the asymptotic inversion, which can remove migration artifacts in practice, to velocity analysis in the subsurface-offset domain for better robustness. This approach requires large memory and cannot be currently extended to 3D.

In this thesis, I propose to transpose the strategy to the more conventional common-shot migration based velocity analysis. I analyze how the approach can deal with complex models, in particular with the presence of low velocity anomaly zones or discontinuous reflectivities. Additionally, it requires less memory than its counterpart in the subsurface-offset domain. I also propose to extend Inversion Velocity Analysis to the data-domain, leading to a more linearized inverse problem than classic waveform inversion. I establish formal links between data-fitting principle and image coherency criteria by comparing the new approach to other reflection-based waveform inversion techniques. The methodologies are developed and analyzed on 2D synthetic data sets.

Keywords

seismic imaging • inverse problem • quantitative migration velocity analysis • common-shot gathers • data domain