Multi-parameter seismic linear waveform imaging

Milad Farshad

PSL
UNIVERSITÉ PARIS

THÈSE DE DOCTORAT
DE L'UNIVERSITÉ PSL

Préparée à MINES ParisTech

Multi-parameter seismic linear waveform imaging

Imagerie sismique multi-paramètres : approche linéarisée par équation d'onde

Soutenue par
Milad Farshad
Le 9 Décembre 2021

École doctorale n°398
Géosciences, Ressources Naturelles et Environnement

Spécialité
Géosciences et Géoingénierie

Composition du jury :

M. Stéphane Operto Directeur de recherche, CNRS	*Président du jury*
M. Gilles Lambaré Ingénieur de recherche, CGG	*Rapporteur*
M. René-Édouard Plessix Ingénieur de recherche, Shell	*Rapporteur*
Mme. Hélène Barucq Directrice de recherche, INRIA	*Examinateur*
M. Romain Brossier Maître de Conférences, Univ. Grenoble Alpes	*Examinateur*
M. Hervé Chauris Professeur, MINES ParisTech	*Directeur de thèse*

MINES ParisTech | PSL

Abstract

Seismic imaging is one of the most effective methods for estimating the Earth's physical parameters from seismic data. Based on the assumption of scale separation, several imaging methods 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. The goal of Migration techniques is to determine the reflectivity in a given background model. Among different migration algorithms, reverse time migration (RTM) has become the method of choice for seismic imaging in complex geologic structures. By definition, RTM is the adjoint of the linearized Born modeling operator and suffers from various migration artifacts. Recent developments proposed an asymptotic inversion in the context of RTM. They determine a direct method to invert the Born modeling operator, providing quantitative results within a single iteration. The direct inverse, also known as pseudoinverse Born operator, automatically compensates for uneven illuminations and geometrical spreading, removing in practice migration artifacts. This operator is expressed in the concept of the extended model domain, such that the dimensions for the model and data spaces fit. There are several possibilities for the extension. The additional parameter used in this work is the horizontal subsurface offset, providing panels called Common Image Gathers (CIGs), as a function of depth and of the subsurface-offset. The application of the pseudoinverse Born operator is based on constant-density acoustic media, which is a limiting factor for practical applications.

In this thesis, I first extend the applicability of linearized direct inverse from constant-density acoustic to variable-density acoustic and elastic media. These extensions consist of three main steps: application of the pseudoinverse Born operator, transforming the subsurface-offset domain CIG to the angle-domain CIG using a Radon transform, and physical decomposition. The last part ("Physical decomposition") is based on a weighted least-squares (WLS) objective function and by definition is either a bivariate (variable-density acoustic) or trivariate (variable-density elastic) curve fitting problem. The proposed method not only provides more robust results, but also offers the flexibility to include constraints in the objective function in order to suppress migration artifacts. Only the first step contains wave-equation terms. I analyze how the approach can effectively deal with complex models, in particular in presence of inaccurate background models and Gaussian noisy data. In the concept of the multi-parameter imaging, the main limitation is the non-uniqueness of the inversion results.

To tackle the ill-posedness of the inverse problem, I propose to add independently ℓ_1-norm constraints to each inverted parameter as regularization terms in the WLS method.

I illustrate that regularization is essential especially in the presence of sparse acquisition geometry. Furthermore, I utilize the pseudoinverse Born operator to accelerate the convergence rate of multi-parameter least-squares RTM (LSRTM). Using 2D synthetic models, I illustrate that the proposed preconditioning scheme not only largely reduces the required number of iterations but also significantly increases the quality of the inverted images even in presence of strong parameter cross-talk and inaccurate migration background models.

Finally, I investigate the importance of accounting for density variations in linearized direct waveform inversion using a marine real dataset from the Eastern Nankai Trough, offshore Japan. I demonstrate that the variable-density direct inversion generates subsurface images with higher resolution but also a better reconstruction of the field data than does the constant-density approach.

The main contributions of this work are: (1) extending the linearized direct waveform inversion beyond constant-density acoustics; (2) implementing sparsity-promoting regularization techniques for a more robust scheme; (3) preconditioning multi-parameter LSRTM using linearized direct inverse; (4) validating the results and investigating the effect of density variations using a real dataset. The next step would be to study the impact of these developments on the background model estimation.

Résumé

L'imagerie sismique est l'une des méthodes les plus effectives pour estimer les paramètres physiques de la Terre à partir de données sismiques. Sous l'approximation d'une séparation d'échelle, plusieurs méthodes d'imagerie séparent le modèle de vitesse entre un modèle lisse de référence qui contrôle la cinématique de propagation des ondes, et un modèle de réflectivité qui caractérise les changements rapides des paramètres élastiques. L'objectif des techniques de migration est de déterminer la réflectivité pour un modèle lisse de référence donné. Parmi différents algorithmes de migration, la « Reverse Time Migration » (RTM en anglais) est devenue la méthode de choix pour l'imagerie sismique des structures géologiques complexes. Par définition, la RTM est l'adjoint de l'opérateur de modélisation linéarisée de Born et souffre d'un certain nombre d'artefacts. Des développements récents proposent un inverse de l'opérateur de modélisation ; il donne des résultats quantitatifs en une seule itération. L'inverse direct, aussi connu sous le nom d'opérateur pseudoinverse de Born, compense automatiquement pour l'illumination non homogène et pour l'amplitude géométrique. En pratique, il attenue fortement les artefacts de migration. Cet opérateur est développé dans le concept de « modèles étendus », pour lesquels les dimensions des tailles des modèles et des données sont les mêmes. Plusieurs possibilités existent pour les extensions. Le paramètre additionnel utilisé ici dans ce travail est le « subsurface offset » horizontal, qui permet de créer des panneaux « Common Image Gathers » (CIGs), comme une fonction de la profondeur et du paramètre d'extension. L'application de l'opérateur pseudoinverse suppose un milieu à densité constante, ce qui est un facteur limitant pour les applications pratiques.

Dans cette thèse, j'étends d'abord l'applicabilité de l'inverse direct linéarisé depuis les milieux acoustiques à densité constante vers les milieux acoustiques à densité variable et vers les milieux élastiques. Trois étapes sont proposées pour ces extensions : application de l'opérateur pseudoinverse, transformation des CIGs du domaine « subsurface offset » en CIGs en angle par une transformée de Radon, et décomposition physique des paramètres. La dernière partie (« décomposition physique ») est basée une fonction objective pondérée au sens des moindres carrés (WLS, « Weighted Least-Squares » en anglais) qui revient à estimer deux ou trois classes de paramètres (cas densité variable et élastique). La méthode proposée non seulement apporte des résultats plus robustes, mais aussi offre la flexibilité d'inclure des contraintes dans la fonction objective pour supprimer les artefacts de migration. Seule la première étape contient des termes liés au calcul de la propagation des ondes. J'analyse comment cette approche peut effectivement fonctionner dans les modèles complexes, en particulier pour des milieux de

référence non parfaits et en présence de bruit Gaussien. Dans le contexte de l'imagerie multi-paramètre, la principale limitation est la non-unicité de la solution.

Pour prendre en compte ce problème mal posé, je propose d'ajouter des contraintes ℓ_1, indépendamment sur chacun des paramètres à inverser comme un terme de régularisation dans la méthode WLS. J'illustre que cette régularisation est essentielle, tout particulièrement pour des géométries d'acquisition creuses. De plus, j'utilise le pseudoinverse pour accélérer la convergence de la LSRTM multi-paramètre (« Least-Squares RTM »). Au travers de modèles synthétiques 2D, j'illustre que le schéma de préconditionement non seulement réduit le nombre d'itérations nécessaires mais aussi augmente de manière significative la qualité des images inversées, même en présence de « cross-talks » entre les paramètres ou encore pour des vitesses de références imparfaites.

Enfin, j'étudie l'importance de prendre en compte les variations de densité sur un jeu de données réelles marines du Japon (« Eastern Nankai Trough »). Je montre que l'inverse direct à densité variable apporte des images du sous-sol avec une meilleure résolution mais aussi une plus grande fidélité aux données que peut le faire une approche avec une densité purement constante.

Les principales contributions de ce travail sont : (1) l'extension de l'imagerie linéarisée directe au-delà du cas à densité constante ; (2) l'implémentation de techniques de régularisation avec la promotion de la sparsité, pour des schémas plus robustes ; (3) le préconditionement de la LSRTM multi-paramètre avec l'inverse direct ; (4) la validation sur données réelles et l'étude des effets de variations de densité. L'étape suivante sera de comprendre l'impact de ces développements sur l'estimation du modèle de vitesse de référence.

Acknowledgments

First of all, I would like to thank my supervisor Prof. Hervé Chauris for his continuous support, motivation, and guidance during the running of this project. Hervé continuously provided encouragement and was always willing and enthusiastic to assist in any way he could throughout the research project. I am very grateful to him for helping me contact companies to peruse my passion in the industry. I could not have imagined having a better supervisor for my Ph.D. study.

My gratitude extends to the members of my jury. I thank René-Édouard Plessix and Gilles Lambaré for reviewing my thesis in details with professional comments and fruitful recommendations. I thank Stéphane Operto for chairing the thesis defense, and Romain Brossier, Hélène Barucq and François Audebert for examining my work.

Many thanks to the members of the geophysics team at MINES ParisTech. I especially thank Véronique Lachasse who extremely facilitated the administrative affairs since my first day at the office. I thank the permanent members of our lab, Mark Noble, Alexandrine Gesret and Pierre Dublanchet for their kindness.

I would like to once more thank Mark Noble for providing the real dataset, and more importantly for his guidance during processing the dataset and interpreting the results. As a colleague, he is more like a good friend of mine. I can recall the moment when we changed a flat tire using a seismic sledge hammer.

Many thanks to my dear colleagues and friends that have encouraged and supported me in the past years. I specially thank Rita Abou Jaoude, Tianyou Zhou, Michelle Almakari, Emmanuel Cocher, Mohamadhasan Mohomadian Sarvandani, Hao Jiang, Christine El Khoury, Jinlin Jiang, Léo de Souza, Antoine Collet, Laure Pizzella, Warren Guistel, Carlos André Martins de Assis. I am deeply grateful to Tianyou Zhou, who has been truly a good friend and colleague. I express my thanks to Michelle Almakari, Rita Abou Jaoude, Christine El Khoury and Léo de Souza who served me as french translators whenever I needed them. I assure you all have a place in my heart.

I would like to acknowledge my friends in Institute of Geophysics, University of Tehran for their kindness and support. I thank Reza dolatabadi and my best friend Kamal Aghazade for their pure friendship and many research discussions. I would like to thank Ali Gholami and Hossein Aghamiri for their wise counsel and for introducing and recommending me to Hervé.

My deep and sincere gratitude to my family, my father, mother, my parents-in-law, my sister Mina, my brothers-in-law Saleh, Erfan and Emran, and my sweet nephews Sajedeh and Sameh for their continuous and unparalleled love, help and support. Lastly,

I wish to thank my loving and supportive wife, Elham. Thank you for unconditional love and endless care of me. You will be the most important and attractive research in my whole life!

I sincerely apologize if I forgot some names, but there were many others who helped me during my Ph.D.

Dedicated to the memory of my grandfather

Contents

List of Figures **xiii**

List of Tables **xxvii**

Part I – Background

1 Introduction **1**
- 1.1 Seismic imaging . . . 3
  - 1.1.1 Seismic surveys . . . 3
  - 1.1.2 Forward problem . . . 6
  - 1.1.3 Inverse Problem . . . 7
- 1.2 Full-waveform inversion . . . 12
  - 1.2.1 Principles . . . 12
  - 1.2.2 Applicability and practical aspects . . . 15
- 1.3 Linearized full-waveform inversion . . . 18
  - 1.3.1 Migration: history . . . 18
  - 1.3.2 True-amplitude migration: an iterative scheme . . . 20
  - 1.3.3 An alternative to iterative scheme: direct inversion . . . 22
- 1.4 Motivation, strategy and thesis organization . . . 24
  - 1.4.1 Thesis organization . . . 27
  - 1.4.2 Contributions . . . 28

2 Review of constant-density direct methods **31**
- 2.1 Abstract . . . 34
- 2.2 Introduction . . . 34
- 2.3 MVA principle . . . 36
- 2.4 Focusing principle . . . 41
  - 2.4.1 Relationships between velocity errors and focusing . . . 41
  - 2.4.2 Expected convex shape . . . 48
  - 2.4.3 Gradient artifacts . . . 49
- 2.5 Extensions to a more complex physics . . . 51
  - 2.5.1 Beyond the constant-density acoustic case . . . 51

2.5.2 Beyond the single scattering approximation 54
2.6 Implementation aspects 58
2.6.1 Reducing the number of shots 59
2.6.2 Target-oriented MVA 60
2.7 Discussion 61
2.7.1 Resolution 62
2.7.2 Tomographic Hessian 63
2.7.3 Multi-parameter imaging 63
2.8 Conclusions 64
2.9 Acknowledgements 64
2.10 Appendix I: Gradient derivation 64
2.11 Appendix II: Approximate inverse 66
2.12 Appendix III: Equivalences between data and image domains 69

Part II – Direct methods – Beyond constant-density acoustics

3 Extension to variable-density acoustic media 75
3.1 Abstract 77
3.2 Introduction 77
3.3 Preliminaries 79
3.3.1 Variable density acoustic Born modeling 80
3.3.2 Pseudoinverse Born modeling 82
3.3.3 Angle domain CIGs 87
3.4 Variable density pseudoinverse modeling 89
3.4.1 Two-trace method 89
3.4.2 Weighted least-squares (WLS) method 89
3.4.3 Taylor expansion 91
3.5 Numerical experiments 92
3.5.1 Simple model 92
3.5.2 Marmousi2 model 102
3.6 Discussion 105
3.7 Conclusions 112
3.8 Acknowledgments 112

4 Extension to variable-density elastic media 113
4.1 Abstract 115
4.2 Introduction 115
4.3 Methodology 118
4.3.1 Elastic Born modeling 118
4.3.2 Reverse time migration 119
4.3.3 Elastic pseudoinverse Born operator 120

4.4 Numerical experiments . . . 125
4.4.1 Simple model . . . 125
4.4.2 Dipping layer model . . . 134
4.4.3 Marmousi2 model . . . 142
4.5 Discussion . . . 151
4.6 Conclusion . . . 153
4.7 Acknowledgments . . . 153
4.8 Appendix I: Gradient derivation for elastic WLS method . . . 154
4.9 Appendix II: Acoustic WLS method . . . 154

Part III – Iterative methods

5 Sparsity-promoting regularization 157
5.1 Abstract . . . 159
5.2 Introduction . . . 159
5.3 Theory . . . 162
5.3.1 Forward Born modeling . . . 162
5.3.2 Pseudoinverse Born operator . . . 163
5.3.3 Extension to variable density . . . 164
5.4 Numerical experiments . . . 167
5.4.1 Discontinuous interfaces . . . 168
5.4.2 The Marmousi2 model . . . 172
5.4.3 2004 BP salt model . . . 179
5.5 Discussion . . . 187
5.6 Conclusion . . . 189
5.7 Acknowledgments . . . 190

6 Preconditioning least-squres reverse time migration 191
6.1 Abstract . . . 193
6.2 Introduction . . . 193
6.3 Reverse time migration . . . 197
6.4 Least-squares reverse time migration . . . 200
6.5 Preconditioning with pseudoinverse Born operator . . . 200
6.6 Numerical experiments . . . 202
6.6.1 Simple model . . . 203
6.6.2 Dipping layer model . . . 211
6.6.3 Marmousi2 model . . . 219
6.7 Discussion . . . 226
6.8 Conclusion . . . 230
6.9 Appendix I: Derivation of Hessian . . . 230
6.10 Appendix II: Alternative approach . . . 231

Part IV – Real data applications

7 A case study from the Eastern Nankai Trough **233**

7.1 Abstract . . . 236
7.2 Introduction . . . 236
7.3 Geological background, survey description and seismic data . . . 238
7.3.1 Geological background . . . 238
7.3.2 Survey description and Seismic data . . . 239
7.4 Processing . . . 239
7.4.1 Velocity-stack inversion . . . 241
7.4.2 Automatic velocity picking . . . 243
7.4.3 Normal-moveout-free stacking . . . 244
7.4.4 Interval-velocity building and time to depth conversion . . . 245
7.4.5 Data filtering and wavelet estimation . . . 246
7.5 Review of multi-parameter pseudoinverse Born inversion . . . 247
7.5.1 Application of pseudoinverse Born operator . . . 248
7.5.2 Transforming the subsurface offset to the diffraction angle . . . 249
7.5.3 Solving a weighted least-squares objective function . . . 251
7.6 Results and validation . . . 253
7.7 Discussion . . . 264
7.8 Conclusion . . . 265
7.9 Acknowledgments . . . 265
7.10 Appendix: Gradient derivation for WLS . . . 265

Part V – Conclusions and Perspectives

8 Conclusions and Perspectives **267**

8.1 Conclusion . . . 269
8.1.1 Beyond constant density . . . 269
8.1.2 Iterative approaches . . . 270
8.1.3 A case study from the Eastern Nankai Trough . . . 271
8.2 Perspectives . . . 272
8.2.1 Extension to anisotropic media . . . 272
8.2.2 Extension to multi-component seismic data . . . 274
8.2.3 Coupling multi-parameter direct inverse with migration velocity analysis . . . 275
8.2.4 Coupling multi-parameter direct inverse with reflection waveform inversion . . . 276
8.2.5 Extension to 3D . . . 276

References **279**

List of Figures

1.1 Acquisition geometries for different types of seismic surveys. a) Land (left) and marine (right) acquisitions (from Danish Energy Agency), and b) VSP acquisition (from Blackburn et al. (2007)). 5

1.2 Illustration of the scale separation for the Marmousi model. The full velocity model (left) is decomposed into a background model (middle) and a reflectivity model (right). 8

1.3 In black the famous sketch by Claerbout (1985) summarizing the spatial frequencies that can be resolved from seismic data. Nowadays the mid frequency gap (2-10 Hz) is filled by the improved resolution brought by advanced tomography methods (red curve) and development of high-quality broadband data (blue curve) (from Lambaré et al. (2014)). . . . 9

1.4 Examples of radiation patterns as a function of diffraction angle for different parametrization in an acoustic media. 12

1.5 Wavenumber illumination. One source-receiver pair (s, r), one scattering point $\mathbf{x}$, and one frequency in the data provide one wavenumber in the image space (equation 1.3). 14

1.6 Illustration of cycle skipping effect. Local optimization converges towards the global minimum for the reconstructed data with a smaller time shift error (first and second rows), whereas it converges towards a local minimum for a larger time shift error (third row). The black arrows denote the descent directions. 16

1.7 A standard RTM computation animation. Forward propagation of the source wavelet (left panels), back propagation of the observed data (middle panels) and application of imaging condition (right panels). 20

1.8 Constant density acoustic media. a) Observed data for a single layer Earth model and the corresponding CIG inverted via pseudoinverse Born modeling operator with b) correct and c) incorrect background model (high velocity). 24

1.9 A two layered Earth model a) without density variations and the corresponding b) shot gather, inverted c) ξ and d)$(h/h_{max})\xi$. (e,f,g,h) Same as (a,b,c,d) with density variations in the model (from Farshad and Chauris (2020b)). 26

1.10 Twenty iterations results of IVA for a maximum frequency equal to 10 Hz. a) background velocity model, b) extracted CIGs from $x = 7, 13, 19, 25, 30$ km, c) perturbation velocity model and d) full velocity model equal to the sum of (a) and (c). Dashed ellipses in (b) indicate the defocused energy may be related to density effects (from Zhou (2020)). 26
1.11 Sketch showing the interaction and organization of the thesis chapters. . 29

2.1 Schematic illustration of the migration of three locally coherent events picked in the data domain (left), with a common surface offset (middle) and subsurface offset (right) schemes. The red, blue and green colors correspond to under-, correct and over-estimated velocity values. 38
2.2 Shape of events in CIGs for the subsurface case and for homogeneous velocity values between 1500 to 2500 m/s, every 100 m/s The local horizontal axis for each panel refers to the subsurface offset, from -100 to +100 m/s. 39
2.3 Interpretation of the tomographic gradient: a non-zero contribution is obtained when there is a $\mathbf{y}$ position such that $A_n^2(\mathbf{h})\xi(\mathbf{y}-\mathbf{h},\mathbf{y}) \neq 0$ and when there is another $\mathbf{x}$ position along the ray joining $\mathbf{y}+\mathbf{h}$ and a receiver position $\mathbf{r}$, here possibly $\mathbf{r}_0$. The same holds for the source side. The second condition implies that in a smooth macro-model, the tomographic update is only above the reflector (left) and not below (right). 40
2.4 Gradient principle for the MVA case: the interface is created as a correlation between the source (solid red lines) and the receiver (solid blue lines) wavefields (top). The macro-model gradient is obtained as the sum of two contributions (see the text for details), leading to a large-scale perturbation above the interface (middle and bottom, gray zone). . 42
2.5 The a) reflectivity section, associated b) CIG for $x = 2000$ m and c) tomographic gradient, for an underestimated velocity model and a single interface, from Chauris and Cocher (2017). 46
2.6 Same as for Figure 2.5, but for the inverse case. In addition, the tomographic gradients are displayed for c) $\beta = 0$ and d) $\beta = -3/2$, from Chauris and Cocher (2017). 47
2.7 "Gradient artifacts", with the standard MVA approach in the case of discontinuous reflectors, from Fei and Williamson (2010). The horizontal and vertical axis refer to the z and x coordinates. 50
2.8 Discontinuous multilayer case, with the migration (a, c, and e) and pseudoinverse (b, d, and f) approaches, for reflectivity section (a and b); macro-model update (c and d); and the same update after Gaussian smoothing over the expected wavelength (e and f), from Chauris and Cocher (2017). 50

2.9 A two layered Earth model a) without density variations and the corresponding b) shot gather, inverted c) ξ and d)$(h/h_{max})\xi$. (e,f,g,h) Same as (a,b,c,d) but with density variations in the model, from Farshad and Chauris (2020b). 51

2.10 Comparison between the inverted parameters corresponding to sparse shots (sourse spacing is 432 m) for Marmousi2 model. Inverted a) inverse Bulk modulus and b) density parameters without regularization and inverted c) inverse Bulk modulus and d) density parameters with ℓ_1-norm regularization, from Farshad and Chauris (2021c). 53

2.11 The reconstructed shot via the inverted physical parameters obtained a) without regularization and b) with ℓ_1-norm regularization. The dashed lines in (a) correspond to the extracted traces at the surface offsets c) 0.0 km and d) 2.5 km, from Farshad and Chauris (2021c). 53

2.12 Extension beyond reflected waves: CIG for reflected waves (a, b, and c), body waves traveling along the surface (d, e, and f) and transmitted waves (g, h, and i), for an under (a, d, and g), exact (b, e, and h) and over-estimated (c, f, and i) macro-model, from Lameloise and Chauris (2016). 55

2.13 MVA cross-well configuration with the full wavefield. a) CIG in the initial homogeneous model, b) CIG in the final model, c) exact velocity model and d) inverted model. In the final model, energy is localized around $h = 0$, from Lameloise and Chauris (2016). 56

2.14 Final macro-model in m/s (a), model perturbation in s^2/m^2 (b) and total model in m/s (c), as well as the CIGs (d) for the Marmousi case, from Zhou et al. (2020b) 67

2.15 Kinematic invariants for a common shot migration in possibly two different macro-models, adapted from Chauris et al. (2002a). 73

3.1 The analytical diffraction pattern for an acoustic medium parameterized by (β, ρ). 82

3.2 Constant density acoustic media. a) Observed data for a single layer Earth model and the corresponding CIG inverted via pseudoinverse Born modeling operator (equation 3.9) with b) correct and c) incorrect background model (high velocity). 85

3.3 Left: reconstructed shots with a) correct velocity extended domain, b) incorrect velocity extended domain and c) incorrect velocity physical domain. The RMS error between synthetic and observed data is written on each panel. right: extracted traces for near and far offsets. Shots are plotted in the same scale. 86

3.4 Variable density acoustic media. a) Observed data for a single layer Earth model and the corresponding b) AVA for the interface, and c) CIG inverted via pseudoinverse Born modeling operator (equation 3.9) with correct background model. 87
3.5 a) a CIG with three nonzero points and b) its Radon transform showing the path of integration for different points. 88
3.6 Weighting mask $W(\mathbf{x}, \gamma)$. The red and blue dashed lines defines the W for $\alpha = 1$ and $\alpha = 0.85$, respectively. 90
3.7 Exact a) β , b) ρ and c) reflection coefficient as a function of angle for each interface corresponding to model 1. 93
3.8 The initial correct background a) β_0 and b) ρ_0 models used for pseudoinverse Born modeling. Inverted c) ξ and its d) $\mathcal{R}_\xi$ at position $x = 1500$ m corresponding to Figure 3.7. The green and black dashed lines in (d) corresponds to the limit for acquired angles and the used ones in inversion, respectively. 94
3.9 Left: observed and reconstructed shots obtained via different methods corresponding to correct background models, right: extracted traces. The RMS error between synthetic and observed data is written on each panel. Shots are plotted in the same scale. 95
3.10 Left: observed and reconstructed shot obtained via Taylor expansion method corresponding to short-offset acquisition, right: extracted traces. Shots are plotted in the same scale. 96
3.11 Inversion results via two-trace method for different angles. 97
3.12 RMSE corresponding to a) inverted β and ρ and b) shot gathers for different angles via two-trace method comparing to WLS method. The gray dashed line in (a) corresponds to maximum available angle for the 4th event. 98
3.13 Comparison of inversion results for β, ρ and I_p obtained via a) two-trace, b) WLS and c) Taylor expansion corresponding to Figure 3.8. The solid black and dashed red lines correspond to the true and inverted values, respectively. 99
3.14 Inverted a) ξ and its b) $\mathcal{R}_\xi$ at position $x = 1500$ m corresponding to Figure 3.7. The green and black dashed lines in (b) corresponds to the limit for acquired angles and the used ones in inversion, respectively. . . 100
3.15 Left: observed and reconstructed shots obtained via different methods corresponding to incorrect background models, right: extracted traces. Shots are plotted in the same scale. 101
3.16 Comparison of inversion results for β, ρ and I_p obtained via a) two-trace and b) WLS corresponding to Figure 3.14. The solid black and dashed red lines correspond to the true and inverted values, respectively. 102

3.17 The exact a) velocity and b) density corresponding to the Marmousi2 model. 103
3.18 The initial incorrect background a) velocity and b) density model and the initial correct background c) velocity and d) density model. 104
3.19 Inverted a) ξ and its b) $\mathcal{R}_\xi$ corresponding to incorrect background model and inverted c) ξ and its d) $\mathcal{R}_\xi$ corresponding to correct background model at position $x = 5.2$ km. The green and black dashed lines in (b) and (d) corresponds to the limit for acquired angles and the used ones in inversion, respectively. 105
3.20 Shot gathers of a) observed data and reconstructed data for the shot at position 5.2 km with b) incorrect background models and c) correct background models. The extracted traces at different offsets are shown in left of each shot. Shots are plotted in the same scale. 106
3.21 The exact perturbation model for a) β, b) ρ and c) I_p, and the inverted perturbation model for d) β, e) ρ and f) I_p corresponding to incorrect background models and g) β, h) ρ and i) I_p corresponding to correct background models. The dashed line corresponds to extracted traces. . . 107
3.22 The extracted traces from perturbation models (Figure 3.21) for a) 2.2 km, b) 5.2 km and c) 8.5 km. The solid black and dashed red lines correspond to the true and inverted values, respectively. 108
3.23 The inverted parameters for different parameterization classes as a) (β, ρ), b) (I_p, ρ), c) (V_p, ρ) and d) (V_p, I_p). The third parameter in each panel is inferred by combination of the first two parameters. 109
3.24 Inverted a) ξ and its b) $\mathcal{R}_\xi$ corresponding to incorrect background model and 480 m extension of subsurface offset. c) Reconstructed data for the shot at position 5.2 km with incorrect background models. The extracted traces at different offsets are shown in left on the shot gather. The green and black dashed lines in (b) corresponds to the limit for acquired angles and the used ones in inversion, respectively. 111

4.1 P- to P-wave elastic amplitude diffraction pattern. 123
4.2 Simple model experiment: exact a) P-impedance, b) S-impedance and c) density models. 127
4.3 Simple model experiment: true perturbation models of a) P-impedance, b) S-impedance and c) density. 128
4.4 Elastic and acoustic shot-domain comparisons for the simple model experiment: a) elastic observed shot versus b) acoustic observed shot. The yellow arrows indicate the main differences between the elastic and acoustic simulations (AVO effects or converted waves). 129

4.5 Simple model experiment using the elastic observed data: a) inverted δm using elastic pseudoinverse Born operator and b) its Radon transform, c) inverted δm using acoustic pseudoinverse Born operator and d) its Radon transform. The difference between (a,c) and (b,d) are shown in (e) and (f), respectively. The blue and red lines in (b,d,f) correspond to the theoretical limit for acquired angles and the practical ones used in inversion, respectively. 130

4.6 Elastic and acoustic reconstructed shot comparisons for the simple model experiment: a) elastic reconstructed shot versus b) acoustic reconstructed shot. Traces are extracted from two positions (indicated as black dashed line in (a)) for surface offset c) 0 and 3 km. Shots are plotted at the same scale as the observed shot (Figure 4.4a). 131

4.7 Elastic and acoustic inverted image-domain comparisons for the simple model experiment: inverted a) ζ_{I_p}, b) ζ_ρ and c) ζ_{I_s} using elastic WLS, d) ζ_{I_p} and e) ζ_ρ using acoustic WLS. 132

4.8 Vertical logs extracted at $x = 3.4$ km (Figure 4.7) comparing the true models (black) and estimated ones using acoustic WLS (blue) and elastic WLS (red). The blue line for ζ_{I_s} is not displayed as it is not reconstructed. 133

4.9 Dipping layer model experiment: exact models of a) P-impedance, b) S-impedance and c) density, and background models of d) P-impedance, e) S-impedance and f) density. A and B in (c) indicate gas-charged sand channel and water-wet sand channels, respectively. 134

4.10 Dipping layer model experiment. a) Inverted δm and b) its Radon transform. The blue and red lines in (b) correspond to the limit for acquired angles and the used ones in inversion, respectively. 135

4.11 The a) observed and b) reconstructed shots for the dipping layer model experiment. Shots are plotted in the same scale. Traces are extracted from two positions (indicated as black dashed lines in (a)) for surface offset c) 0 and 3.8 km. The solid black and dashed red lines correspond to the true and reconstructed values, respectively. 136

4.12 Exact and inverted image-domain comparisons for the dipping layer model experiment: exact a) ζ_{I_p}, b) ζ_{I_s} and c) ζ_ρ, and inverted d) ζ_{I_p}, e) ζ_{I_s} and f) ζ_ρ. The dashed lines in (a,b,c) indicate the position of the extracted traces. Inverted models are plotted in the same scale as the exact models. 137

4.13 Vertical logs passing through the two target areas from inverted dipping layer model (Figure 4.12) for a) 3.4 km and b) 5.8 km. A and B arrows in (a) and (b) indicate the gas charged sand channel and water wet sand channels, respectively. 138

4.14 Evaluation of a) data misfit with respect to error in different background models, and NCC in image domain with respect to error in b) P-wave velocity background model, c) S-wave velocity background model and d) density background model corresponding to dipping layer model. . . 139
4.15 Same as Figure 4.10, but for inaccurate P-wave velocity background model (8% error): a) inverted δm and b) its Radon transform. 140
4.16 Same as Figure 4.12, but for inaccurate P-wave velocity background model (8% error): inverted a) ζ_{I_p}, b) ζ_{I_s} and c) ζ_ρ. Inverted model are plotted in the same scale as the exact models. 141
4.17 Marmousi2 model experiment: exact models of a) P-impedance, b) S-impedance and c) density, and background models of d) P-impedance, e) S-impedance and f) density. 142
4.18 The a) observed and b) reconstructed shots for the Marmousi2 model experiment. Shots are plotted in the same scale. Traces are extracted at two positions (indicated as black dashed lines in (a)) for surface offset c) 0.5 and 4.5 km. The solid black and dashed red lines correspond to the true and reconstructed values, respectively. 143
4.19 Exact and inverted image-domain comparisons for the Marmousi2 model experiment: exact a) ζ_{I_p}, b) ζ_{I_s} and c) ζ_ρ, and inverted d) ζ_{I_p}, e) ζ_{I_s} and f) ζ_ρ. The dashed lines in (a,b,c) indicate the position of the extracted traces. Inverted models are plotted in the same scale as the exact models. 144
4.20 Vertical logs from inverted Marmousi2 model (Figure 4.19) for a) 3.4 km and b) 6.1 km. The green arrow in (b) indicates the position of the gas sand. 145
4.21 Radio of a) 2nd/1st and b) 3rd/1st eigenvalues of the matrix $\mathbf{M}$ at each spatial coordinate corresponding to the Marmousi2 model. 146
4.22 The Marmousi2 noisy observed shot with a) SNR -1 dB, b) SNR -5 dB and c) SNR -10 dB, and the reconstructed shots corresponding to noisy data with d) SNR -1 dB, e) SNR -5 dB and f) SNR -10 dB. Shots are plotted in the same scale as the noiseless observed shot (Figure 4.18a). . 147
4.23 The extracted traces at $x = 0.5$ km from Figure 4.22 for a) SNR -1 dB, b) SNR -5 dB and c) SNR -10 dB. 148
4.24 The inverted image-domain comparisons for the Marmousi2 model using noisy data: inverted (a,d,g) ζ_{I_p}, (b,e,h) ζ_{I_s} and (c,f,i) ζ_ρ using noisy data with (a,b,c) SNR -1 dB, (d,e,f) SNR -5 dB and (g,h,i) SNR -10 dB. 149
4.25 The localized NCC values corresponding to Figure 4.24: for (a,d,g) ζ_{I_p}, (b,e,h) ζ_{I_s} and (c,f,i) ζ_ρ using noisy data with (a,b,c) SNR -1 dB, (d,e,f) SNR -5 dB and (g,h,i) SNR -10 dB. 150

5.1 The diffraction pattern for $(\beta - \rho)$ parameterization. 165

5.2 The exact a) V_p (m/s) and b) ρ (kg/m^3) models corresponding to discontinuous interfaces. 169
5.3 The a) offset-domain CIG (δm) and b) its angle-domain response corresponding to Figure 5.2. The green and black dashed lines define the W for $\alpha = 1.0$ and $\alpha = 0.6$, respectively. 170
5.4 Comparison between the inverted parameters for discontinuous interfaces. Exact a) $\delta\beta/\beta_0$ and b) $\delta\rho/\rho_0$. Inverted c) $\delta\beta/\beta_0$ and d) $\delta\rho/\rho_0$ via the WLS method, and inverted e) $\delta\beta/\beta_0$ and f) $\delta\rho/\rho_0$ via the RWLS method. The green dashed lines in (a) indicate the positions of the extracted traces. The truncation artifacts are circled by the yellow ellipses in (c). 171
5.5 Comparison between the extracted traces from Figure 5.4 at a) $\mathtt{x} = 1.4$ km, b) $\mathtt{x} = 2.4$ km and c) $\mathtt{x} = 2.9$ km. 172
5.6 The exact a) V_p (m/s), b) ρ (kg/m^3), c) $\delta\beta/\beta_0$ and d) $\delta\rho/\rho_0$ models corresponding to the Marmousi2 model. 173
5.7 The Marmousi2 dense shots experiment. The a) observed data and reconstructed data for the shot at position 5.2 km via the b) WLS and c) RWLS methods. The dashed lines in (a) correspond to the d) extracted traces. Shots are plotted at the same scale. 173
5.8 Comparison between the inverted parameters corresponding to dense shots (292 shots) for Marmousi2 model. Inverted a) $\delta\beta/\beta_0$ and b) $\delta\rho/\rho_0$ vis the WLS method, c) $\delta\beta/\beta_0$ and d) $\delta\rho/\rho_0$ via the RWLS method. The dashed lines in (a) correspond to the extracted traces. 175
5.9 Comparison between the extracted inverted parameters corresponding to dense shots (292 shots) for Marmousi2 model. Inverted $\delta\beta + \beta_0$ and $\delta\rho + \rho_0$ for a) $\mathtt{x} = 4.5$ km and b) $\mathtt{x} = 7.0$ km. 175
5.10 Same as for Figure 5.7, but for the decimated shots experiment. The reconstructed data for the shot at position 5.2 km via the a) WLS and b) RWLS methods. The dashed lines in (a) correspond to the extracted traces at surface offset c) 0.0 km and d) 2.5 km. Shots are plotted at the same scale as the observed shot (Figure 5.7a). 176
5.11 Same as for Figure 5.8, but for the decimated shots experiment. Inverted a) $\delta\beta/\beta_0$ and b) $\delta\rho/\rho_0$ via the WLS method, c) $\delta\beta/\beta_0$ and d) $\delta\rho/\rho_0$ via the RWLS method. 177
5.12 Same as for Figure 5.9, but for the decimated shots experiment. Inverted $\delta\beta + \beta_0$ and $\delta\rho + \rho_0$ for a) $\mathtt{x} = 4.5$ km and b) $\mathtt{x} = 7.0$ km. 177
5.13 The Marmousi2 decimated shots experiment. Wavenumber spectra of $\delta\beta/\beta_0$ corresponding to a) true, b) WLS and c) RWLS, and $\delta\rho/\rho_0$ corresponding to d) true, e) WLS and f) RWLS. Panels are plotted on the same amplitude scale as (a) and (d). The box and arrows highlight the observed artifacts. 178

5.14 Evaluation of a) NCC in image domain and b) RMSE in data domain with respect to the source spacing corresponding to Marmousi2 model. The gray dashed line in (b) indicates the abrupt change in behavior of the WLS method, and the one in (a) is deduced from the one in (b). . . . 179

5.15 The exact a) V_p (m/s), b) ρ (kg/m^3), c) $\delta\beta/\beta_0$ and d) $\delta\rho/\rho_0$ corresponding to the modified 2004 BP salt model. 180

5.16 The 2004 BP salt model. The a) observed and constructed data via the b) WLS and c) RWLS methods corresponding to the noiseless data. (d), (e), (f) same as (a), (b), (c) corresponding to the noisy data ($\mathrm{SNR} = -8$ dB). The dashed line in (a) corresponds to the extracted trace in Figure 5.17. Shots are plotted at the same scale. 181

5.17 The 2004 BP salt model. The zero-offset traces from shots in Figure 5.16 for the a) noiseless and b) noisy data. The NCC values corresponding to each trace are written on the panels. 182

5.18 The inverted parameters corresponding to the noiseless 2004 BP salt model. Inverted a) $\delta\beta/\beta_0$ and b) $\delta\rho/\rho_0$ via the WLS method, and inverted c) $\delta\beta/\beta_0$ and d) $\delta\rho/\rho_0$ via the RWLS method. The dashed lines in (a) correspond to the extracted traces. 183

5.19 The inverted parameters corresponding to the noisy 2004 BP salt model. Inverted a) $\delta\beta/\beta_0$ and b) $\delta\rho/\rho_0$ via the WLS method, and inverted c) $\delta\beta/\beta_0$ and d) $\delta\rho/\rho_0$ via the RWLS method. The yellow dashed boxes indicate the salt boundary. 184

5.20 The final inverted parameters corresponding to the (a,b) noiseless and (c,d) noisy 2004 BP salt model. Inverted $\delta\beta + \beta_0$ and $\delta\rho + \rho_0$ for (a,c) $\mathtt{x} = 3.8$ km and (b,d) $\mathtt{x} = 7.0$ km. 185

5.21 Comparison between the histogram of the true model components and the inverted ones (Figure 5.19). a) $\delta\beta/\beta_0$ and b) $\delta\rho/\rho_0$ corresponding to the true models, c) $\delta\beta/\beta_0$ and d) $\delta\rho/\rho_0$ inverted via the WLS method, e) $\delta\beta/\beta_0$ and f) $\delta\rho/\rho_0$ inverted via the RWLS method. The red lines indicate the probability density functions fitted to the histograms. 186

5.22 Evaluation of a) NCC in image domain and b) RMSE in data domain with respect to noise corresponding to 2004 BP salt model. 187

5.23 The angle-domain CIGs positioned at a) 2.6 km, b) 5.2 km and c) 7.8 km corresponding to the Marmousi2 model, and at d) 2.6 km, e) 5.2 km and f) 7.8 km corresponding to the 2004 BP salt model. The green and black dashed lines define the W for $\alpha = 1$ and $\alpha = 0.48$, respectively. 188

5.24 The evaluation of the cost function (equation 5.15) with respect to λ_1 and λ_2 for the Marmousi2 model. 190

6.1 a) Typical surface 2D acquisition setting. The *star* and the *triangles* indicate the seismic source and the receiver positions, respectively. The lines emerged from the source represent the ray path of the primary reflections to illustrate the wave propagation. b) The reflected wavefield recorded by the receivers. 194
6.2 Sketch of the subsurface offset extension. The red and blue arrows indicate the interaction between the source and receiver wavefields for the non-extended and extended modeling operators, respectively (Claerbout, 1985). 198
6.3 Simple model experiment: the exact a) P-impedance and b) density models, and background c) P-impedance and d) density models. 204
6.4 Relative model perturbation ζ (equation 6.6) for the exact a) impedance ζ_{I_p} and b) density ζ_ρ, and inverted c) ζ_{I_p} and d) ζ_ρ after 60 iterations of CG-LSRTM, inverted e) ζ_{I_p} and f) ζ_ρ after 15 iterations of PCG-LSRTM. Each inverted model is plotted with the same scale as for the exact model. 206
6.5 Comparison of inverted impedance perturbation ζ_{I_p} images after a) 1, b) 15 and c) 30 iterations of CG-LSRTM, d) 1, e) 5 and f) 10 iterations of PCG-LSRTM. 207
6.6 Same as Figure 6.5, but for the density model. Inverted density perturbation ζ_ρ images after a) 1, b) 15 and c) 30 iterations of CG-LSRTM, d) 1, e) 5 and f) 10 iterations of PCG-LSRTM. 208
6.7 Comparison of AVA response for the first two layers of Figure 6.4 after 60 iterations of CG-LSRTM and 15 iterations of PCG-LSRTM. 209
6.8 Evaluation of a) relative data misfit, and b) normalized correlation coefficients of P-impedance and density models with respect to the iteration number. 210
6.9 Dipping layer model experiment: exact models of a) P-impedance and b) density, and background models of c) P-impedance and d) density. The black arrows indicate the position of the gas-charged sand channel. 211
6.10 The exact model perturbations a) ζ_{I_p} and b) ζ_ρ, and inverted c) ζ_{I_p} and d) ζ_ρ via 40 iterations of CG-LSRTM, inverted e) ζ_{I_p} and f) ζ_ρ via 15 iterations of PCG-LSRTM. Each inverted model is plotted with the same scale as for the exact model. The dashed line in (a) indicates the position of the extracted trace. 212
6.11 Dipping layer model experiment: extracted vertical profiles at $\mathrm{x} = 3.75$ km from inverted P-impedance (left) and density (right) models. The black arrow indicates the gas-charged sand channel. 213
6.12 Dipping layer model experiment: evaluation of a) relative data misfit, and b) normalized correlation coefficients of P-impedance and density models with respect to the iteration number. 214

6.13 Dipping layer model experiment: inaccurate migration background models for a) P-impedance and b) density. 215

6.14 Dipping model experiment: the estimated (a,c) subsurface offset-domain CIG and (b,d) its angle-domain CIG after 15 iterations of PCG-LSRTM assuming (a,b) accurate and (c,d) inaccurate background models. The blue and red lines in the angle-domain CIGs correspond to the theoretical limit for acquired angles ($\alpha = 1$, equation 6.9) and the practical ones used in inversion ($\alpha = 0.85$), respectively. 216

6.15 Same as Figure 6.10, but for inaccurate background models: inverted a) ζ_{I_p} and b) ζ_ρ via 40 iterations of CG-LSRTM, inverted c) ζ_{I_p} and d) ζ_ρ via 15 iterations of PCG-LSRTM. Each inverted model is plotted with the same scale as for the exact model. 217

6.16 Same as Figure 6.12, but for inaccurate background models: evaluation of a) relative data misfit, and b) normalized correlation coefficients of P-impedance and density models with respect to the iteration number. . 218

6.17 The Marmousi2 experiment: exact models of a) P-impedance and b) density, and background models of c) P-impedance and d) density. . . . 219

6.18 The Marmousi2 experiment: a) true angle-domain CIGs and the ones built by application of the Radon transform using inverted model via b) 40 CG-LSRTM iterations and c) 20 PCG-LSRTM iterations. The two successive vertical lines (from left to right) indicate angles of range 0 to 60 degrees. The arrows highlight the main differences. 221

6.19 The exact model perturbations a) ζ_{I_p} and b) ζ_ρ, and inverted c) ζ_{I_p} and d) ζ_ρ via 40 CG-LSRTM iterations, inverted e) ζ_{I_p} and f) ζ_ρ via 20 PCG-LSRTM iterations. The arrows highlight the main differences. 222

6.20 The localized NCC values corresponding to Figure 6.19: for (a,c) ζ_{I_p} and (b,d) ζ_ρ via (a,b) CG-LSRTM and (c,d) PCG-LSRTM. 223

6.21 The Marmousi2 experiment: wavenumber spectra of impedance corresponding to a) true, b) CG-LSRTM and c) PCG-LSRTM, and density corresponding to d) true, e) CG-LSRTM and f) PCG-LSRTM. All panels are plotted with the same amplitude scale. 224

6.22 The Marmousi2 experiment: evaluation of a) relative data misfit, and b) normalized correlation coefficients of P-impedance and density models with respect to the iteration number. 225

6.23 The multi-parameter approximated Hessian: a) without preconditioner and b) with preconditioner. The yellow boxes delineate a close-up structure of the Hessian. 228

6.24 One element of the (a,c) impedance-impedance and (b,d) density-density block matrix of the Hessian in distance-depth domain corresponding to (a,b) without preconditioner and (c,d) with preconditioner. The first, second, and third rows correspond to the shallow, middle and deep part of the model, respectively. The yellow dashed circles denote the position of the perturbation. 229

7.1 Location map of the studied area. The red line indicates the investigated seismic line in this study (Nouzé et al., 2004). 239
7.2 An example of observed common-shot gather. 240
7.3 The CMP fold coverage. 241
7.4 Velocity-stack inversion: (a,c,e) CMP gathers and (b,d,f) their corresponding absolute values of sparse Radon transform plane. The red dashed curves in (b,d,f) show the slowness model calculated by equation (7.4). 243
7.5 a) The area shown in white are those points whose velocity value was calculated. b) RMS velocity model. 245
7.6 The normal-moveout-free stacked section. 246
7.7 The interval velocity section obtained with Dix formula. 246
7.8 The a) frequency spectrum of the original shot illustrated in Figure 7.2, and b) one extracted trace at the location illustrated by dashed line in (a). 247
7.9 The a) shot gather and b) its spectrum after low-pass filtering with cut frequency of 45 Hz. 248
7.10 The extracted traces around the water bottom reflection from the filtered shot (Figure 7.9) for offsets a) 225 m and b) 1175 m. The solid, dashed and dot-dashed lines correspond to the shots positioned at 3.4 km, 5.6 km and 8.1 km, respectively. 249
7.11 The analytical diffraction pattern for a variable-density acoustic medium parameterized by impedance and density. 251
7.12 The a) impedance and b) density background models. 254
7.13 The extracted a) subsurface-offset domain and b) angle-domain CIG. The position of each CIG is written on each panel. 255
7.14 The (a,d,g) observed and (b,c,e,f,h,i) reconstructed data assuming (b,e,h) constant-density and (c,f,i) variable-density acoustics for the shots positioned at (a,b,c) 3.1 km, (d,e,f) 5.6 km and (g,h,i) 8.1 km. The NCC and RMS values are written on each panel. The dashed yellow boxes denote the areas with main differences in AVO responses at each shot position. 256
7.15 The extracted traces from the observed and reconstructed shots (Figure 7.14) at (a,b,c) 225 m and (d,e,f) 1180 m for shots positioned at (a,d) 3.1 km, (b,e) 5.6 km and (c,f) 8.1 km. The RMS values between the observed and reconstructed traces are written on each panel. 257

7.16 The common-offset sections comparison: (a,b) the observed common-offset sections versus the reconstructed ones assuming (c,d) constant-density and (e,f) variable-density acoustics for offset (a,c,e) 287.5 m and (b,d,f) 756.25 m. The NCC and RMS values are written on each panel. 258
7.17 The residual common-offset sections comparison: the difference between the observed and reconstructed common-offset sections (Figure 7.16) assuming (a,b) constant-density and (c,d) variable-density acoustics for offset (a,c) 287.5 m and (c,d) 756.25 m. 259
7.18 The localized NCC values corresponding to Figure 7.16: NCC panels for reconstructed common-offset sections assuming (a,b) constant-density and (c,d) variable-density acoustics for offset (a,c) 287.5 m and (c,d) 756.25 m. 260
7.19 The inverted a) δI_p assuming constant-density acoustics, and b) δI_p and c) $\delta\rho$ assuming variable-density acoustics. The dashed yellow boxes point to the areas for the magnified views. 261
7.20 Magnified views of the inverted impedance models corresponding to a) constant-density and b) variable-density acoustics, and the main structures detected overlaid on the impedance section for c) constant-density and d) variable-density acoustics. 262
7.21 The eigenvalue/eigenvector decomposition of the matrix $\mathbf{M}$ versus depth: a) 1st eigenvalue, b) 2nd eigenvalue, c) 1st eigenvector and d) 2nd eigenvector. 263

8.1 The a) observed shot assuming isotropic acoustic medium and its b) inverted CIG, and c) observed shot assuming VTI anisotropic acoustic medium and its d) inverted CIG. The inverted CIGs are obtained with isotropic pseudoinverse Born operator. 273
8.2 The Hess VTI model experiment: exact a) δV_p, b) δV_h and c) δV_n, and inverted δV_p image using d) RTM, e) isotropic pseudoinverse Born operator and f) anisotropic pseudoinverse Born operator. The yellow arrow in (e) and (f) denote the parameter cross-talk. 275

List of Tables

1.1 Non exhaustive examples of methods that mitigate cycle-skipping by modifying/reformulating the FWI objective function. 17

2.1 Review of the approximate inverse operator expressions for different approaches (Chauris and Cocher, 2018). 67
2.2 Kinematic invariants. The fitted quantity is zero by construction of the migration scheme. The mid-point slope p_m and subsurface-offset slope p_H are defined as $p_m = p_s + p_r$ and $p_H = p_s - p_r$, respectively. The weights α_s and α_r are specified in the text. For the subsurface offset case, the fit to the two slopes are always satisfied, but possibly from different subsurface locations. 70

3.1 Dimension of the data and model domains. s and r are the source and receiver coordinates; t is the time; x , y and z are the spatial coordinates; and h is the subsurface offset. 80
3.2 The different diffraction patterns for different parameterization classes. . 107

4.1 Main parameter configuration for different numerical experiments. . . . 126
4.2 Modeling engine for each Figure corresponding to simple model experiment. 127

6.1 Main parameter configuration for the different numerical experiments. . 205
6.2 Comparison of the number of iterations and total computational time (in hours) for CG- and PCG-LSRTM. 205
6.3 Maximum value of each block of the Hessian matrix. 228

Chapter 1

Introduction

Contents

1.1 Seismic imaging **3**
1.1.1 Seismic surveys 3
1.1.2 Forward problem 6
1.1.3 Inverse Problem 7
1.2 Full-waveform inversion **12**
1.2.1 Principles 12
1.2.2 Applicability and practical aspects 15
1.3 Linearized full-waveform inversion **18**
1.3.1 Migration: history 18
1.3.2 True-amplitude migration: an iterative scheme 20
1.3.3 An alternative to iterative scheme: direct inversion 22
1.4 Motivation, strategy and thesis organization **24**
1.4.1 Thesis organization 27
1.4.2 Contributions 28

Résumé

L'imagerie sismique est très utilisée pour caractériser les structures géologiques du sous-sol à partir de l'analyse des sismogrammes, et potentiellement pour mieux connaitre les ressources de la Terre. Les images sont des représentations du sous-sol, par exemple au travers de la vitesse, densité ou des paramètres visco-élastiques. Dans le contexte de l'exploration pétrolière, une caractérisation précise de la subsurface est importante pour interpréter la géologie, déterminer la position des puits de forage. Dans ce chapitre, je résume les principaux éléments de l'imagerie sismique et de la résolution du problème inverse. Les deuxièmes et troisièmes parties de ce chapitre donnent un aperçu général des méthodes non-linéaires et linéaires d'imagerie sismique. Finalement, je conclus avec les principales limites actuelles et les questions ouvertes en imagerie sismique, comme sources de motivation pour ce travail. Je présente l'organisation globale du document, avec les principales contributions de chaque chapitre. Ici, j'introduis quelques notations et quelques équations. Les détails techniques sont présentés au Chapitre 2.

Seismic imaging is widely utilized to characterize the Earth's geological structures in the subsurface from the analysis of observed seismograms, and thus to expose possible resource-bearing formations. The images are well-defined representations of the Earth's interior, for instance, velocity, density, or viscoelastic parameters. In oil and gas exploration, a precise subsurface model is essential for subsequently interpreting the geology, determining the drilling location, and accurately positioning the wells.

In this chapter, I first summarize the key steps in seismic imaging and the resolution of the inverse problem. The second and third parts of this chapter is an overview of the seismic non-linear and linear imaging methods. Finally, I conclude with the current limitations and open questions in seismic imaging, motivation for my research, and the organization of the thesis. In this chapter, I only introduce a few notations and equations. More technical details are presented in Chapter 2.

1.1 Seismic imaging

Seismic imaging is a procedure for estimating the Earth's physical parameters from seismic data. These physical parameters can be represented by the spatial distribution of, e.g., P-wave velocity, S-wave velocity, density, porosity, or anisotropic parameters (Schuster, 2011). In general, seismic imaging consists of three main elements: (1) data acquisition through seismic surveys, (2) forward problem and (3) inverse problem (Tarantola, 2005).

In the following, I briefly review the main theoretical aspects of the seismic surveys, forward and inverse problems. The reader is referred to Sheriff and Geldart (1995); Yilmaz (2001); Schuster (2017) for a detailed introduction to seismic imaging.

1.1.1 Seismic surveys

"Seismic survey is a program for mapping geologic structure by observation of seismic waves, especially by creating seismic waves with artificial sources and observing the arrival time of the waves reflected from acoustic-impedance contrasts or refracted through high velocity members" (Sheriff, 2002). Investigated subsurface geologic structures which contain oil and gas reservoirs are found beneath land and sea environments. Thus the seismic data can be acquired either in land or sea. Although they share the same goal, being imaging the subsurface, the two contexts require unique technology and terminology.

In the land acquisition (Figure 1.1a, right), the sources are either dynamites or seismic vibrators. The later is a truck-mounted or buggy-mounted device that is capable of injecting low-frequency vibrations into the Earth. In the marine acquisition (Figure 1.1a, left), the sources are either air guns or water guns, with the air gun being by far the most common. In both environments, sources and receivers are arranged along

coinciding lines. The main difference, as far as geometry is concerned, is that in marine acquisition an end-on geometry (source is deployed at one end of the receiver line) is used; whereas in land data acquisition a split-spread geometry (source is deployed in the center of the receiver lines) is also possible (Sheriff and Geldart, 1995). In both environments, borehole seismic acquisition, called Vertical Seismic Profiling (VSP), is also possible, with sources at the surface and receivers within a well (Figure 1.1b) (Hardage, 1985; Gilpatrick and Fouquet, 1989). Compared to surface seismic acquisitions, VSP provides locally higher resolution images of the subsurface which is used for signal calibration with surface seismic data. In general, different types of VSP exist, depending on the well configuration, the numbers, positions and types of the sources and receivers (Figure 1.1b) (Blackburn et al., 2007). Seabed acquisition (also referred as ocean bottom acquisition) is also a rapidly growing area of the seismic acquisition. In ocean bottom acquisition, receivers can be either ocean bottom cables (OBC) or ocean bottom nodes (OBN). Compared to conventional streamer acquisition on the water surface, ocean bottom has lots of benefits, for instance, recording 3C or 4C elastic wavefield and providing wider azimuth (Beaudoin and Ross, 2007).

Whatever the acquisition design, the recorded Earth's response to the triggered source in the acquisition is called *seismogram* (also known as *common-shot gather* or simply *shot gather*).

In the context of seismic imaging, waves can be classified into two main types: (1) *body waves*, which travel through the Earth's interior, (2) *surface waves*, which travel only along the surface of the Earth. In seismic exploration, these two types of waves are automatically recorded, but surface waves are commonly considered as noise for subsequent imaging steps. The body waves can also be split into two main groups: (1) *P-wave* (also known as *primary/compressional wave*) which oscillates along the direction of wave propagation, (2) *S-wave* (also known as *secondary wave/shear wave*) which oscillates perpendicularly to the direction of wave propagation.

When body waves propagate through the Earth they are categorized according to their propagation path, as:

- *Transmitted waves* – Waves that travel between the source and receivers without being reflected;
  - *Direct waves*: For a homogeneous velocity in the superficial part of the model, direct waves travel across the shallow part of the Earth from the source to the receivers following a straight wave path;
  - *Diving waves* (also known as *turning waves*): Assuming a positive vertical velocity gradient in the shallow part of the model, diving waves travel through the Earth and naturally bend back to the surface fallowing a curved wave path. These waves only have a limited penetrating depth, especially for short offsets, and may arrive at the receivers earlier than the direct waves;

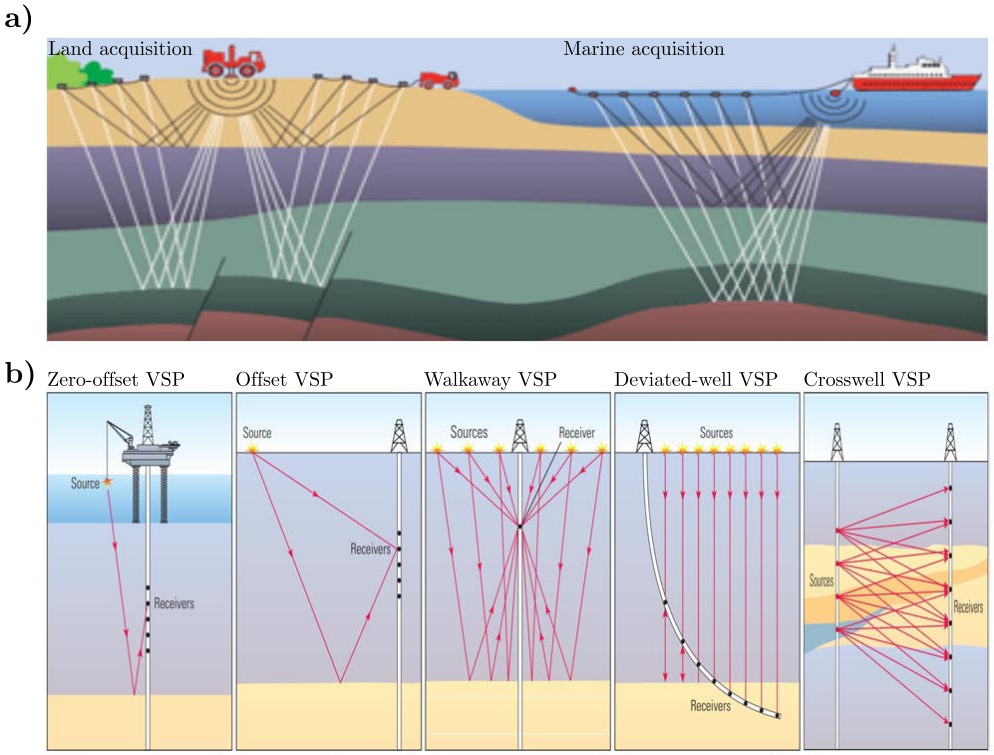

Figure 1.1: Acquisition geometries for different types of seismic surveys. a) Land (left) and marine (right) acquisitions (from Danish Energy Agency), and b) VSP acquisition (from Blackburn et al. (2007)).

- *Refracted waves* – Similar to diving waves, but in a model with interfaces, when the waves strike an interface with higher velocity, part of wavefronts propagating tangent to the interface serve as secondary sources and trigger plane waves that propagate to the surface. They may arrive at the receivers with shorter time than the direct waves at far offsets;
- *Scattered waves* – Waves that are generated due to discontinuities of the Earth;
  - *Reflected waves*: When the waves strike an interface with impedance contrast, part of their energy is reflected backwards to the surface. For a single reflection, these waves are categorized as first-order scattered waves;
  - *Diffracted waves*: When the waves strike a sharp edge (particle, object or obstacle), the edge acts as a secondary source (*Huygens' principle*) and generates diffracted waves. Multi-scattering exists, depending on the number of times the wavefield is diffracted.
  - *Multiples*: These waves are high-order reflected/diffracted waves that are

bounced back and forth for several times between the surface and interfaces before being recorded. They are caused by strong reflectors (i.e. the sea surface or the water bottom) and commonly recognized as unwanted events and hence removed during processing even if they provide potentially higher resolution.

1.1.2 Forward problem

Once the data are acquired, one of the key elements in almost every aspect of exploration seismology is the modeling of seismic data. Fundamentally, it provides a means of understanding the characteristics of observed seismic data. In addition, modeling is the core of many seismic data-processing and imaging/inversion algorithms (section 1.1.3).

Mathematically, the propagation of the waves through the Earth is formulated as a partial differential equation (PDE) called wave equation, which simulates the wave propagation giving a set of model parameters. The general viscoelastic anisotropic wave equation involves density, attenuation of P and S waves and the 21 elastic coefficients of the stiffness tensor relating the stress tensor to the strain tensor. Note that all these quantities are spatially varying. Due to the high computational cost of simulating the viscoelastic anisotropic wave propagation, it is usually simplified by reducing the parameters to be considered in the propagation. But the more fundamental reason is that one cannot independently estimate all these parameters.

I briefly review the numerical methods for solving the wave equation (Virieux et al., 2011). The most popular ones are finite-difference method (FDM) (Kelly et al., 1976; Virieux, 1986; Levander, 1988; Operto et al., 2007), and finite-element method (FEM) (Smith, 1975). Moczo et al. (2010) compared the two methods and concluded that the accuracy of the two methods are comparable with dense samplings. The main advantage of FEM over FDM is its flexibility in meshing to deal with the boundary conditions and irregular structures. In a 3D complex model, the construction of a suitable mesh can be complicated. Virieux et al. (2011) have reviewed FDM and FEM from the point of view of efficiency and complexity, and indicated that FDM is widely used due to its simplicity to implement and the relatively lower computational cost.

The methods such as FDM and FEM are based on direct numerical solutions to the wave equation. The modeling can also rely on ray theory, i.e., a high-frequency approximation of the wave equation (Červenỳ et al., 1977). The ray theory decomposes the Green's function, a specific solution of wave equation, into three elements: (1) traveltimes, (2) amplitudes and (3) source signature. Therefore, new equations are required to be solved: the Eikonal equation for traveltimes (Vidale, 1988; Podvin and Lecomte, 1991) and the transport equation for amplitudes along the calculated rays (Babich and Buldyrev, 1989). An ideal modeling approach should consist of generating complete synthetic seismograms, including all types of waves (section 1.1.1). "Wave-equation-based modeling" (i.e. without high-frequency approximation) provides complete seis-

mograms, but in a "black-box" manner. The ray-based modeling not only rapidly and accurately provides synthetic seismograms for user-selected waves, but also additional information such as traveltime, amplitude and polarization (Lecomte et al., 2015). However, finite-frequency aspects are not taken into account.

This thesis only focuses on wave-equation-based modeling methods (in particular FDM) and does not consider ray-based modeling methods, except for analyzing the wave-equation operators.

1.1.3 Inverse Problem

The inverse problem in exploration geophysics can be defined as fitting the observed data (d^{obs}) to synthetic data (d) by adjusting the model parameters (m). The synthetic data is obtained by

$$d = G(m), \tag{1.1}$$

where G denotes the forward modeling operator relating d to m (section 1.1.2). There are several types of observed seismic data, i.e. traveltimes, or more generally phase information, and waveform information (Schuster, 2017).

Objective function

The inverse problem consists of defining a scalar *objective function*, also called *cost function*. The goal of the objective function is to evaluate the accuracy of model parameters used for modeling. It is designed such that the optimal model is a global minimizer/maximizer of the function. The most common choice for objective function is measuring the differences between observed and reconstructed data in the least-squares sense assuming an additive white noise (Tarantola, 1984).

In practice, seismic inverse problem is an ill-posed problem, meaning that an infinite number of models approximatively matches the data (Tarantola, 2005). This is also referred as nonuniqueness of the inverse problem. Ill-posedness may result from various reasons. In a typical seismic survey, the sources and receivers are located only over a small portion of the zone of interest, leading to insufficient illumination of the subsurface. In addition, in the framework of multi-parameter reconstruction, different combinations of parameters may lead to the same seismic response, such that they cannot be retrieved by only analyzing the observed data (Operto et al., 2013). Preconditioning or regularization techniques may alleviate the nonuniqueness of the ill-posed inverse problem and steer the solution to one with preferred characteristics, such as sparsity promoting norms, smooth and/or blocky models (Menke, 1984; Tarantola, 2005; Scales et al., 1990; Asnaashari et al., 2013). A way to guide the inversion is to separate the model parameters into different spatial components as discussed in the next section.

Scale separation

The separation of scales between the velocity wavenumbers, which are responsible for the *kinematics* and *dynamics* aspects, was first noticed by Claerbout (1985) and Jannane et al. (1989) as

- The smooth slowly varying (small-wavenumber) components of the model, also called *background model* or *macro model*, governing the kinematics (traveltimes) of wave propagation (Figure 1.2, middle);
- The singularities (large wavenumbers) of the model, also called *reflectivity*, governing the dynamics (amplitudes) of wave propagation (Figure 1.2, left).

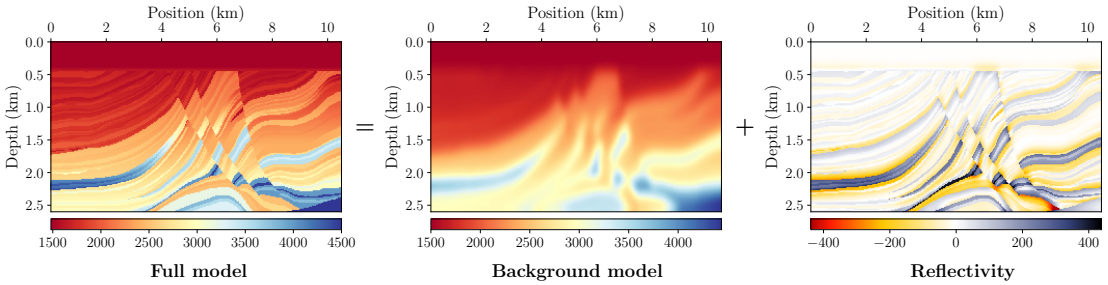

Figure 1.2: Illustration of the scale separation for the Marmousi model. The full velocity model (left) is decomposed into a background model (middle) and a reflectivity model (right).

In general, the seismic imaging workflow, based on the scale separation, can be subdivided into two distinguished tasks (Wapenaar, 1996): (1) building the background model, a highly non-linear problem, followed by (2) the reconstruction of reflectivity, usually a linear problem. The first task is carried out by methods so called *tomographic* methods, and the second one by *migration* methods (Mora, 1989). The quality of the migration depends on the quality of the background model provided by the tomographic methods. Note that both tomography (in a broad sense) and migration can be implemented with a wave-equation-based approach or with an asymptotic approach. The alternative is to recover the full model (Figure 1.2, left) without scale separation. This method is known as Full Waveform Inversion (section 1.2).

Under the concept of scale separation, Claerbout (1985) demonstrated that the models retrieved from seismic data lack intermediate wavenumbers (black curve in Figure 1.3). However, nowadays this gap is significantly filled due to improvements from the acquisition and processing sides: designing wide-azimuth long-offset acquisition geometries, using high-quality broadband data, and applying more advanced high-resolution imaging methods (red and blue curves in Figure 1.3) (Nichols, 2012; Lambaré et al., 2014).

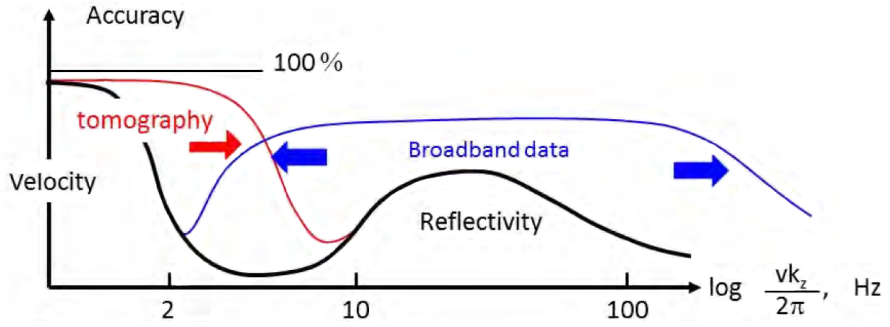

Figure 1.3: In black the famous sketch by Claerbout (1985) summarizing the spatial frequencies that can be resolved from seismic data. Nowadays the mid frequency gap (2-10 Hz) is filled by the improved resolution brought by advanced tomography methods (red curve) and development of high-quality broadband data (blue curve) (from Lambaré et al. (2014)).

Optimization

Mathematically, optimization is an iterative process seeking the minimization/maximization of a objective function. There is no universal optimization algorithm but rather a collection of algorithms, each of which is tailored to a specific type of optimization problem. The responsibility of choosing appropriate algorithm for a particular application often falls on the user. This choice is an important one, as it may determine whether the problem is solved rapidly or slowly and, indeed, whether the solution is found at all, with related uncertainties or not (Nocedal and Wright, 2006). In seismic imaging, the objective function is typically a nonlinear function of the model parameters. The main difficulty with such scheme is that the objective function is not necessarily convex and often plagued by many local minima. In general, there are two classes of optimization methods: global optimization and local optimization (Gill et al., 1981). The global optimization methods evaluates the all candidate solutions within the whole model space to obtain the optimal solution (Sen and Stoffa, 2013). The methods from this class such as Monte Carlo methods (Jin and Madariaga, 1994; Cordua et al., 2012; Stuart et al., 2019), simulated annealing (Mosegaard and Vestergaard, 1991; Dariu et al., 2003; Tran and Hiltunen, 2012; Datta et al., 2019) and genetic algorithms (Gallagher et al., 1991; Boschetti et al., 1996; Mazzotti et al., 2016; Luu et al., 2017) have been investigated in geophysical inverse problems. Although robust in practice, the main bottleneck of this class is its expensive computational cost, as it requires numerous assessment of the objective function. Applications are mostly related to ray-based approaches, although some tentatives have been tested in the context of full waveform inversion (Mazzotti et al., 2016; Datta et al., 2019).

The second class of methods is local optimization methods, also referred as gradient-based methods (Nocedal and Wright, 2006). Contrary to the first class, this class requires

an initial starting model. The local optimization methods iteratively searches for models along the downward direction of the objective function, and usually stops at the first minimum encountered, hopefully the global minimum. The gradient of the objective function with respect to model parameters, which can be efficiently calculated by using the adjoint-state technique (Plessix, 2006b), determines an approximation of the ideal direction of the model update (Newton update) at each iteration. The advantage of these schemes compared to the global optimization methods is that they have much less computational cost, which makes them more applicable in practice. However unlike the global optimization methods, they are susceptible to be trapped in a local minimum. Examples from this class investigated in geophysical inverse problems are steepest-descent method (Lines and Treitel, 1984; Tarantola, 1984), conjugate-gradient method (Mora, 1987; Crase et al., 1990; Rodi and Mackie, 2001; Hu et al., 2011), quasi-Newton methods (Brossier et al., 2009; Ma and Hale, 2012; Wu et al., 2015) and Newton methods (Pratt et al., 1998; Akcelik, 2003; Métivier et al., 2013). The convergence rate for the Newtonian methods is generally faster than the gradient-based methods, but at the expense of regularization and evaluating the impact of the Hessian matrix (second-order derivative of the objective function) at each iteration which is computationally expensive (both computing time and memory consumption). An alternative to Newton methods are quasi-Newton methods, which build up an approximation of the Hessian or its inverse from previous gradients. Implementing an efficient and correct version of this method, such as the BFGS algorithm (named after Broyden, Fletcher, Goldfarb and Shanno), exceeds a few lines of code (Byrd et al., 1995). The reader is referred to Nocedal and Wright (2006) for an extensive review of optimization methods.

Before concluding this section, I would like to discuss the multi-parameter inversion, an important aspect in this thesis.

Multi-parameter

In the concept of seismic modeling and inversion, the most basic assumption about Earth's subsurface is to consider an isotropic acoustic medium, parametrized by the P-wave velocity. The main advantage of this simplification is the affordable computational burden, as well as a relatively less ill-posed inverse problem.

From the point of view of seismic interpretation, P-wave velocity models can provide useful information for geological interpretation. But for detecting the hydrocarbon potential of a reservoir, more realistic physical properties of the subsurface, and in particular elastic properties, are required:

> *Density*, which is closely related to the porosity, is one of these physical properties. The importance of introducing a density or a pseudodensity term (an additional term for a better data fit) in seismic imaging has been frequently highlighted (Plessix et al., 2013; Chen and Sacchi, 2018).

Impedance, the product of velocity with density, mostly affects through its contrast the amplitude of reflected seismic waves (especially for short offsets). Traditionally this parameter is a preferred choice for the reflectivity reconstruction (Mora, 1987; Forgues and Lambaré, 1997; Brossier et al., 2015; Zhou et al., 2015). But contrary to widely accepted ideas, Prieux et al. (2013) showed that using P-wave velocity and density allows us to take advantage of the isotropic radiation pattern for P-wave velocity and hence its broadband reconstruction. Then to absorb the small coupling effect between P-wave velocity and density, we can form impedance a posteriori.

Seismic attenuation, the term to characterize the seismic energy loss through the subsurface media, is another physical property of the Earth. This parameter can provide useful information for reservoir characterization, such as the physical state, the degree of saturation and lithology (Aki and Richards, 1980; Wang, 2008).

S-wave velocity contains different rock information than the P-wave velocity. The interpretation of P-wave velocity alone has been unable to detect the gas reservoirs because the P-wave velocities of both the gas and the surrounding shale are very similar (Ensley, 1984). On the other hand, S-wave velocity in combination with P-wave velocity provides additional information about lithology and pore fluid of the subsurface assisting in detection of possible gas reservoirs (Shi et al., 2007; Brossier et al., 2009).

Seismic anisotropy, the term to describe the directional dependency of the seismic velocity, affected by fracturing and strength behavior of the subsurface rocks. Accounting for anisotropy in seismic imaging can improve reservoir detection in oil and gas exploration (Pratt and Shipp, 1999; Operto et al., 2009; Prieux et al., 2011).

We cannot retrieve all the Earth's elastic parameters from seismic data. The inverse problem is an ill-posed problem. Including more parameters into inversion increases the ill-posedness of the inverse problem, since more degree of freedom are considered (Virieux and Operto, 2009). The effect of different parameter classes are coupled, meaning that they have a similar impact on seismic data. This issue, known as *cross-talk* or *trade-off* between parameters, can be analyzed by plotting radiation patterns (Figure 1.4) (Forgues and Lambaré, 1997; Operto et al., 2013). As pointed out before, this is indeed due to the insufficient illumination of the subsurface caused by limited acquisition at the surface. The possible solution is to consider a hierarchical approach where we first invert the most dominant parameters and estimate the remaining in a sequential manner (Tarantola, 1986; Brossier et al., 2009; He et al., 2018).

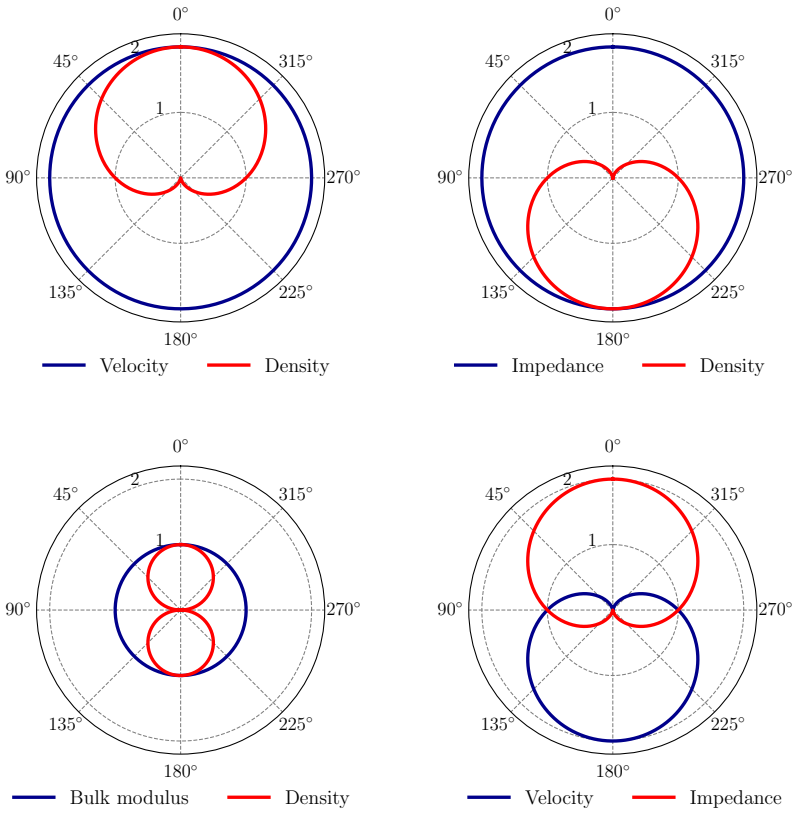

Figure 1.4: Examples of radiation patterns as a function of diffraction angle for different parametrization in an acoustic media.

I have provided an overview of the main elements for seismic imaging. In the next two sections, I briefly explain why full waveform inversion is the reference imaging tool, but also that it may be complicated to obtain meaningful results. I also review the linearized version, which is indeed the core of this thesis. The objective is to have all the elements to understand the validity of the questions raised in this thesis (section 1.4).

1.2 Full-waveform inversion

I review the main principles as well as the practical aspects for successful applications.

1.2.1 Principles

During the eighties, Full-waveform inversion (FWI) was introduced by Tarantola (1984) and Lailly (1983) as a non-linear inverse problem. FWI is an iterative approach which

exploits all types of waves by minimizing the weighted least-squares differences between the observed (d^{obs}) and reconstructed data $d(m)$ as (Chavent, 1974; Lailly, 1983; Tarantola, 1984)

$$J_{FWI}(m) = \frac{1}{2} \left\| d(m) - d^{obs} \right\|^2, \tag{1.2}$$

where J is the objective function and m denotes the model parameters. Exploiting all types of waves such as direct waves, diving waves, refractions, reflections and multiples, provides information of the full wavefield: kinematics (phase and traveltimes) and dynamics (amplitudes). In other words, there is no need to decompose the model in term of kinematics and dynamic effects (scale separation, section 1.1.3).

Resolution

The resolution of FWI can be analyzed by defining the scattering wavenumber $\mathbf{k}$ as (Miller et al., 1987; Devaney, 1982)

$$\mathbf{k} = \frac{2f}{v_0} \cos\left(\frac{\theta}{2}\right)\mathbf{n}, \tag{1.3}$$

where f is the frequency, θ denotes the diffraction angle and $\mathbf{n}$ is the unit vector normalizing vector $\mathbf{k}$ (Figure 1.5). Several key conclusions can be derived from equation 1.3 (Virieux and Operto, 2009):

(1) Small diffraction angle induces large wavenumber (i.e. short wavelength), whereas large diffraction angle obtains small wavenumber (i.e. long wavelength);

(2) The low-frequency content in the data resolves small wavenumbers, whereas high-frequency content resolves large wavenumbers.

In this context, large diffraction angles (which correspond to far-offset data and thus transmitted waves) and low-frequencies recover small wavenumbers (i.e. the large-scale structure of the velocity model, tomographic mode). On the other hand, small diffraction angles (which correspond to short-offset data) and high-frequencies recover the large wavenumbers (i.e. the singularities of the velocity model, migration mode).

(3) The maximum resolution of half the wavelength can be obtained by using the maximum frequency and recording normal-incidence reflections ($\theta = 0°$).

In practice, due to the limited acquisitions, the deeper part of the subsurface is investigated mainly by reflected waves with limited diffraction angles.

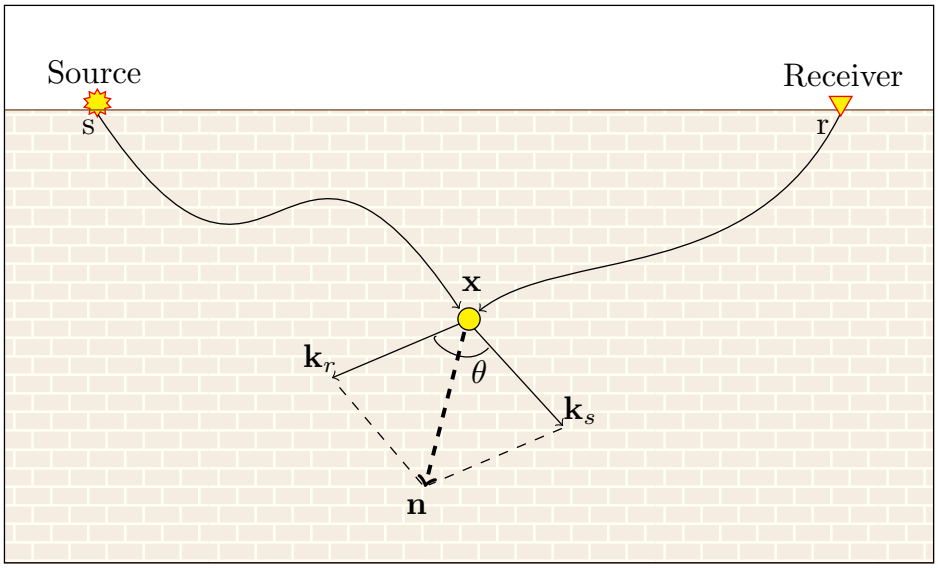

Figure 1.5: Wavenumber illumination. One source-receiver pair (s, r), one scattering point **x**, and one frequency in the data provide one wavenumber in the image space (equation 1.3).

Implementation

The first implementation of FWI was in the time domain (Lailly, 1983; Tarantola, 1984). Later in nineties, it was formulated in the frequency domain (Pratt and Worthington, 1990; Pratt, 1999). In theory, if a dense series of discrete frequencies is chosen from frequency spectrum, the two implementations lead to the same results. The two approaches are indeed equivalent if all frequencies are inverted simultaneously (Pratt et al., 1998). But the transposition in the frequency domain led to the recognition of the key role played by frequencies to avoid local minima (see section 1.2.2). Strategies have been proposed to select the frequencies (Sirgue and Pratt, 2004; Sirgue, 2006). In case of sparse selection of the frequencies, the frequency-domain FWI is preferred as it circumvents the computational burden of the time-domain FWI; whereas for dense selection of frequencies, the time-domain implementation is preferred as it brings the flexibility to select one type of arrival by time-windowing (Shipp and Singh, 2002; Virieux and Operto, 2009; Chauris, 2019). In 3D, the time-domain remains the preferred choice (Chauris, 2019).

1.2.2 Applicability and practical aspects

Model parameterization

In the case of multi-parameter inversion, FWI becomes highly non-linear as different classes of parameters have different influences on the data (Virieux and Operto, 2009). Although the combination of different parameter classes is theoretically equivalent, the choice of the parameters is not neutral. In other words, if they are not adequately chosen, the numerical algorithms in the inversion can be inefficient (Tarantola, 1986). For example, the kinematics of the model (corresponding to long wavelengths/small wavenumbers) are governed by P- and S-wave velocities, whereas the dynamics of the model (corresponding to short wavelengths/large wavenumbers) are governed by P- and S-wave impedances and density. These effects are coupled as a function of the diffraction angle (θ) (Forgues and Lambaré, 1997). As indicated in section 1.1.3, the influence (also referred as trade-off) of different parameter classes can be studied by either plotting radiation patterns and sensitivity kernels, or analyzing the eigenvalues and eigenvectors of the Hessian (Forgues and Lambaré, 1997; Ribodetti and Virieux, 2000; Prieux et al., 2013; Operto et al., 2013; Bharadwaj et al., 2018).

Avoiding cycle skipping

The nonconvex nature of FWI objective function is one of the main obstacles for the successful implementation of FWI. This is due to the comparison between oscillating signals, and complexified by the non-linear relationship between the model and data in the FWI objective function (Gauthier et al., 1986). An important manifestation of this nonconvexity is cycle skipping, which occurs as incorrect phases match, for example when the model is not accurate enough to estimate the observed traveltimes with an error less than half of the dominant period (Bunks et al., 1995; Virieux and Operto, 2009). The effect of this issue is illustrated in Figure 1.6.

Mitigating cycle-skipping has been remaining one of the primary challenges for a successful FWI implementation and has been tackled in several ways. These can be categorized into three main classes. The first class is based on multiscale techniques, in which the velocity model is built over multiple inversion runs by staging over frequencies in a frequency continuation strategy (Bunks et al., 1995; Sirgue and Pratt, 2004; Brossier et al., 2009). Since lower frequencies provides wider half cycle, the inversion starts from the lowest oscillating frequency in the observed data and then it is increased sequentially. This will indeed provide wider basis of attraction in FWI objective function. A rule of thumb is that the longest wavelength associated to the lowest frequency covers the deepest depth of interest.

The second class of algorithm consists of reformulation of FWI objective function creating a much broader convex region around the global minimum than that of standard FWI. Table 1.1 presents a non exhaustive examples of this class.

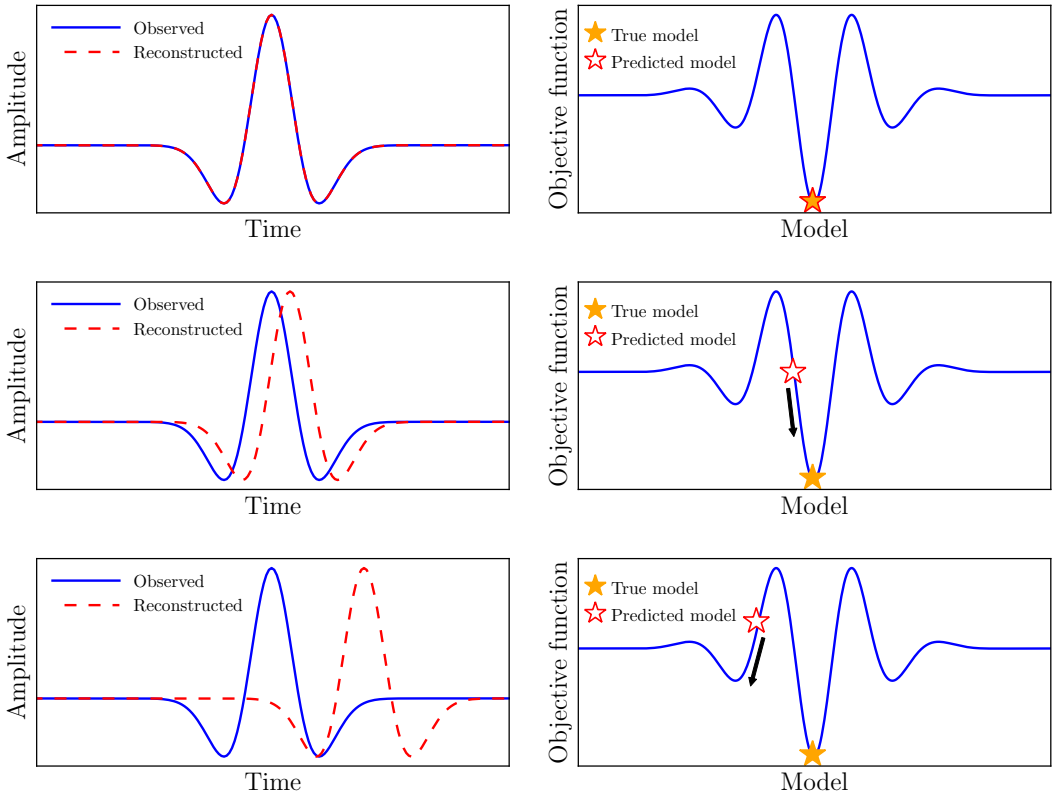

Figure 1.6: Illustration of cycle skipping effect. Local optimization converges towards the global minimum for the reconstructed data with a smaller time shift error (first and second rows), whereas it converges towards a local minimum for a larger time shift error (third row). The black arrows denote the descent directions.

The third category of methods are based on explicit scale separation, enforced in the formulization through an alternating optimization scheme. Two famous methods from this category are Reflection Waveform Inversion (RWI) originally referred to as Migration Based Traveltime Tomography (MBTT) (Chavent et al., 1994; Xu et al., 2012b; Brossier et al., 2015) and Migration Velocity Analysis (MVA) (Al-Yahya, 1989; Sava et al., 2005; Symes, 2008b). Both methods use reflection data and take the form of two nested optimization problems:

- In the inner loop, a reflectivity model is determined by minimizing the data misfit given a fixed background velocity model. This problem is linear or quasi-linear (section 1.3);

- In the outer loop, given the reflectivity model obtained in the inner loop, the background velocity model is determined by minimizing a new (data- or image-

Table 1.1: Non exhaustive examples of methods that mitigate cycle-skipping by modifying/reformulating the FWI objective function.

Method	**Description**
Optimal-transport-distance-based FWI	Minimizes a new objective function based on optimal transport distance using proximal splitting techniques (Métivier et al., 2016a,b; Messud et al., 2021)
Laplace- and Laplace Fourier-domain FWI	The damped wavefields are included to constrain the objective function (Shin and Min, 2006; Kim et al., 2013)
Crosscorrelation- and deconvolution-based FWI	Phase shifts are imparted into the reconstructed traces (Van Leeuwen and Mulder, 2010; Luo and Sava, 2011; Zhang et al., 2018)
Adaptive waveform inversion	Analogous to deconvolution-based FWI, but uses Wiener filters to match the observed and reconstructed data (Van Leeuwen and Mulder, 2010; Luo and Sava, 2011)
Extended FWI	Enlarges parameter search space by penalizing FWI objective function with a least-squares norm of the wave-equation error (van Leeuwen and Herrmann, 2013, 2015; Biondi and Almomin, 2014; Wang et al., 2016a; Huang et al., 2018a,b; Aghamiry et al., 2020a)

domain) objective function which evaluates the quality of the data fit or focusing. This problem is non-linear. But there are reasonable proofs that the objective function is convex (Symes, 2008b).

Although RWI and MVA are similar in spirit, they differ partly in the inner and outer loop:

- In the inner loop, RWI minimizes the data misfit at zero- or short-offsets, whereas MVA minimizes the data misfit at all offsets;

- In the outer loop, RWI minimizes the same objective function as the inner loop (data domain) but at far offsets, whereas MVA minimizes a new objective function defined in the model domain, using a focusing or coherency criterion.

The main principle in MVA (also known as image-domain method) is that the seismic data are redundant: for example, a migrated image can be obtained per each shot position (Al-Yahya, 1989). The image-domain objective function is evaluated on panels called Common Image Gathers (CIGs) representing a section of the reflectivity volume as a

function of depth and of an extra parameter representing the redundancy of seismic data (Symes, 2008b). This scheme will be more detailed in the following sections.

In this thesis, I will concentrate on the linearized full-waveform inversion (migration part), now presented.

1.3 Linearized full-waveform inversion

Seismic migration by definition, being based on scale separation, is a technique for imaging the subsurface structure from observed seismic data. As explained later, the straightforward implementation only provides qualitative (structural) results. This is insufficient for oil and gas exploration which also needs amplitude information (quantitative information) to reconstruct the model parameters, here mainly the contrasts. In the eighties, Lailly (1983) and Tarantola (1984) recognized that migration could be formulated as an inverse problem. In this section, I first briefly review the qualitative migration methods, and then continue with quantitative aspects as a linearized waveform inversion problem. In this thesis, I mainly develop the wave-equation-based migration scheme. However, I will also briefly review ray-based migration methods since they played an essential role not only in the historical development of migration algorithms, but also more recently in the development of wave-equation-based operators (ten Kroode, 2012; Hou and Symes, 2015, 2017; Chauris and Cocher, 2017).

1.3.1 Migration: history

Under the *Born approximation*, the primary refection data (δd) linearly depend on the reflectivity model (δm) (Born, 1926; Beydoun and Tarantola, 1988; Symes, 1995):

$$\delta d = F\delta m, \tag{1.4}$$

where F denotes the linearized forward modeling operator, i.e., Born modeling operator. Standard migration operator can be regarded as the adjoint of the linearized forward modeling operator as:

$$\delta m = F^T \delta d^{obs}, \tag{1.5}$$

where ^{T} denotes the transpose (adjoint) operator, and δd^{obs} is the observed primary reflections.

The earliest form of migration techniques was based on ray theory (Gardner et al., 1974a; French, 1975; Schneider, 1978; Keho and Beydoun, 1988). In ray theory, an asymptotic approximation of the exact solution of the wave equation is obtained by searching for a solution in the form of a ray series (Červeny̌ et al., 1977). The major drawback of such theory is the high frequency assumption, which leads to incorrect estimation of the amplitudes in the vicinity of caustic points and complex structures

of the subsurface. Later, wave-equation-based migration techniques have been proposed by using finite-difference solution of the wave equation, instead of approximating it by ray tracing. Depending on the definition of lateral variations of velocity, wave-equation-based migration can be further classified into one-way wave-equation migration (OWEM) (Claerbout, 1971, 1985), also known as wavefield extrapolation migration (WEM), and two-way wave-equation migration, also known as reverse time migration (RTM) (Baysal et al., 1983; McMechan, 1983; Biondi and Shan, 2002).

Among different migration algorithms, RTM has become the method of choice for seismic imaging in complex geologic structures (Baysal et al., 1983; Mulder and Plessix, 2004). The imaging principle at the core of the RTM is that the reflectors exist where the source and receiver wavefields coincide in time and space (Claerbout, 1971). This concept is also known as *Imaging Condition*. The most common imaging condition is the zero-lag crosscorrelation between the forward- and backward-propagated wavefields. Therefore, suppose that we have estimated the source wavelet of the observed data, a standard RTM algorithm can be summarized into three steps:

- Forward propagation of the source wavelet from the source position to obtain the source wavefield (Figure 1.7, left);
- Back propagation of the data residuals from the receiver positions to compute the receiver wavefield (Figure 1.7, middle);
- Application of the imaging condition to the source and receiver wavefields to obtain an image of the subsurface (Figure 1.7, right);

This algorithm is performed over all seismic sources in the acquisition, by stacking all contributions to increase the signal to noise ratio. Instead of summing over images, one may asses the consistency between the estimated images. This is indeed the basic principle of MVA mentioned in section 1.2.2.

Conventionally, RTM can be formulated as the adjoint of the linearized forward modeling operator, i.e., adjoint Born modeling operator. The adjoint operator (equation 1.5) can only correctly calculate kinematics (phase), and does not preserve amplitudes: if one performs a new modeling from the RTM results, the shots are not correctly reconstructed in terms of dynamics. In practice, the seismic data are also subject to significant aliasing, noise, irregular source and receiver sampling and finite recording aperture: the final migrated images suffer from artifacts and low-resolution (Mulder and Plessix, 2004). In other words, the migrated images obtained by application of the adjoint operator, are not the solutions minimizing the objective function for linearized waveform inversion.

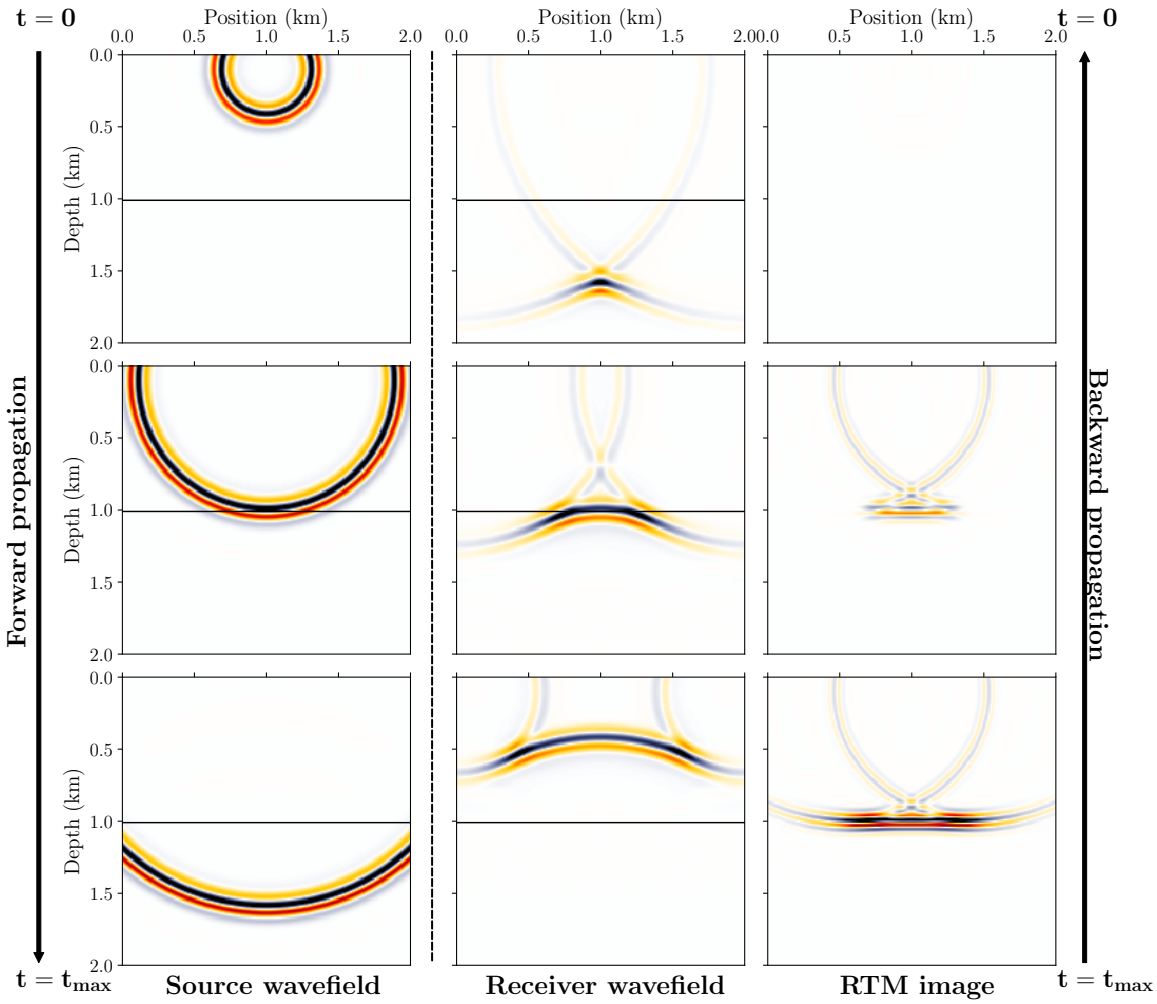

Figure 1.7: A standard RTM computation animation. Forward propagation of the source wavelet (left panels), back propagation of the observed data (middle panels) and application of imaging condition (right panels).

1.3.2 True-amplitude migration: an iterative scheme

To remove the acquisition footprint and to improve the quality of seismic imaging, least-squares migration (LSM) has been proposed to seek a quantitative inverted image, which generates the simulated data best matching the amplitudes of the observed data (LeBras and Clayton, 1988; Nemeth et al., 1999). Since the beginning of the nineties, LSM has been implemented using ray-based migration (Lambaré et al., 1992; Nemeth et al., 1999; Duquet et al., 2000; Fomel et al., 2008), OWEM (Kühl and Sacchi, 2003; Clapp et al., 2005; Zhu et al., 2018) and RTM (Dai et al., 2012; Zhang et al., 2015b; Xue et al., 2016). Assuming a fixed estimate of the background model (m_0), LSM objective

function consists of determining the model perturbation (δm) which best reconstructs the observed primary reflection (δd^{obs}):

$$J_{Mig}(\delta m) = \frac{1}{2} \left\| \delta d - \delta d^{obs} \right\|^2, \tag{1.6}$$

where $\delta d = F\delta m$ represents the synthetic reflection data estimated by Born modeling from δm. The optimal solution is obtained when the gradient $\partial J_{Mig}(\delta m)/\partial \delta m$ is equal to zero, such that (Lailly, 1983; Tarantola, 1984):

$$\delta m^{true} = (F^T F)^{-1} \delta m. \tag{1.7}$$

Here, $F^T F$ is the Hessian matrix. Since the Born operator is not unitary, the Hessian matrix is not equal to the Identity matrix. Therefore, the migrated δm is considered as a Hessian-blurred version of the true-amplitude image δm^{true} (Schuster, 1993; Nemeth et al., 1999; Chavent and Plessix, 1999).

Sensitivity to background velocity model

LSM is formulated based on the assumption that an accurate background model velocity model is available to correctly predict the traveltime (kinematics) of the observed data. The amplitude differences between the observed and modeled data are then matched through an iterative inversion process. The quality and accuracy of migration images depends greatly on the accuracy of the background velocity model, such that errors in this background model can lead to an incoherent, defocused migrated images.

To partially reduce the sensitivity of LSM to background velocity model, Luo and Hale (2014) proposed a simple modification to the objective function, to minimize it after correcting for nonzero traveltime shifts between the reconstructed and observed data using dynamic warping. An alternative option is to implement LSM in the extended domain (Symes, 2008b; Li and Chauris, 2018). The most recent conventional choice for the extension is the subsurface offset (Rickett and Sava, 2002; Hou and Symes, 2018), introduced as an offset between the sunken sources and receivers by Claerbout (1985). The other alternatives for extension can be time shifts (Sava and Fomel, 2006; Yang and Sava, 2011) and scattering angle (Rickett and Sava, 2002; Sava and Fomel, 2003). This extension potentially decouples the data fitting in LSM from the choice of the background velocity model, and provides the key principle of MVA methods: even in an incorrect background velocity model, no information is lost during migration of the data. This implies that the data can be reconstructed in the same background velocity model using the migrated volume and that the quality of the background velocity model can be investigated equivalently in the data and image domains (Symes, 2008b; Chauris et al., 2017).

Multi-parameter LSM

Recent advances in high-performance computing had promoted LSM to extend from pure-acoustic approximation to more realistic approximation of the Earth, such as variable-density acoustic (Yang et al., 2016a), anisotropic acoustic (Duveneck et al., 2008; Zhang et al., 2011; Duveneck and Bakker, 2011), viscoacoustic (Dutta and Schuster, 2014; Dai et al., 2015), Elastic (Stanton and Sacchi, 2017; Feng and Schuster, 2017; Chen and Sacchi, 2018), anisotropic elastic (Rocha and Sava, 2018; Yang et al., 2019b) and viscoelastic (Guo and McMechan, 2018; Hu et al., 2019). Although the acoustic approximation can be questioned in the framework of LSM due to the unreliable amplitudes in the modeling engine, one advantage of acoustic LSM is dealing with less computationally expensive forward modeling than in other cases. Moreover, pure-acoustic LSM is better posed than multi-parameter LSM, thus it does not suffer from cross-talk artifacts as only the dominant parameter P-wave velocity is involved in the inversion. This feature is the same as FWI (Virieux and Operto, 2009).

Cost-effective LSM

The major drawback of LSM is its computational cost as it should be solved iteratively. Each iteration of LSM requires an application of forward modeling (also known as *demigration*) and adjoint (migration) modeling operators. During time, different techniques have been proposed to possibly accelerate the convergence rate of LSM, such as multi-source approach with either random-phase encoding or plane-wave encoding (Dai et al., 2012; Dai and Schuster, 2013; Xue et al., 2016), and scaling the linear system with proper pre-conditioners such as the approximate inverse normal operator, also known as Hessian operator (Shin et al., 2001; Rickett, 2003; Herrmann et al., 2009; Huang et al., 2016). In this thesis, I further investigate this latter aspect.

1.3.3 An alternative to iterative scheme: direct inversion

An interesting alternative to the iterative scheme is the direct inversion, which can be achieved by replacing migration, the adjoint of the Born modeling operator, by its pseudoinverse formula (inverse under some conditions detailed below). This formula makes the Hessian matrix close to a Dirac distribution leading to a better conditioned problem. It was initially proposed for ray-based migration (Beylkin, 1985; Bleistein, 1987; Lambaré et al., 1992), and later for WEM (Zhang et al., 2003, 2007) and more recently for RTM (ten Kroode, 2012; Hou and Symes, 2015, 2017; Duprat and Baina, 2016; Chauris and Cocher, 2017; Li and Chauris, 2018).

It historically started with the asymptotic approximation (ray-based formulation) via the generalized Radon transform, i.e., not summing over linear or parabolic curves, but along isochrones (Beylkin, 1985). Migration/inversion formulas, also known as

true-amplitude prestack depth migration (PSDM), have been proposed either based on ray+Born or ray+Kirchhoff linearized approximations to consider quantitative properties of reflection coefficients (Bleistein, 1987; de Hoop and Bleistein, 1997). In both approaches, the model was split into a unknown reflector/diffractor component and the supposed known background component where the ray tracing was performed. But, they differed by the description of the perturbation components of the model, in which in terms of velocity and density perturbations for ray+Born approach, or in terms of specular reflectivity, for the ray+Kirchhoff approach (Lambaré et al., 2003; Operto et al., 2003).

The methods based on RTM are the recommended approaches as they solve the full wave equation and are able to handle more properly steep dipping angles and complex velocity models. As for the ray-based approaches, the pseudoinverse formula for RTM is expressed in the concept of extended model domain, such that the dimensions for model and data spaces are equivalent. This extension is either surface-oriented (i.e. shot position) (Duprat and Baina, 2016; Li and Chauris, 2018) or subsurface-oriented (i.e. subsurface offset) (ten Kroode, 2012; Hou and Symes, 2015, 2017; Chauris and Cocher, 2017). Quantitative results can be obtained only within a single iteration. Note that the one proposed by ten Kroode (2012) is based on Kirchhoff modeling. This thesis mainly focuses on the pseudoinverse formula proposed by Chauris and Cocher (2017), which is based on subsurface extension.

It appears that the direct inversion is very similar to the standard migration scheme, with only additional weights in the imaging operator (Hou and Symes, 2015, 2017; Chauris and Cocher, 2017). Although the derivations are performed under the high-frequency approximation, the final formulas do not contain any ray quantities but only time and spatial derivatives, which keeps the implementation simple and explicit. In Chapter 2, more details will be provided. It is remarkable that such inversion formulas exist. These direct approaches can also be used as a pre-conditioner to accelerate LSM (Hou and Symes, 2016a; Chauris and Cocher, 2018). Another interesting application is to replace LSM with direct inverse in the inner loop of MVA, to achieve so called inversion velocity analysis (IVA) (Hou and Symes, 2016c; Chauris and Cocher, 2017; Li and Chauris, 2018). The effectiveness of direct inverse as a preconditioner has also been investigated for FWI (Métivier et al., 2015; Hou and Symes, 2016b). The possible technology transfer from IVA to RWI has not been investigated yet.

Although the pseudoinverse formula minimizes the objective function even in the presence of incorrect background velocity model, the estimated CIGs are sensitive to the background velocity: energy is focused at zero subsurface-offset for the correct velocity model and spreads over non-zero offsets otherwise. We illustrate this with a numerical result (Figure 1.8) extracted from Chapter 3. This is indeed the concept of subsurface-oriented MVA/IVA objective function, which penalizes the defocused energy in the CIG (Figure 1.8c) to update the background velocity model.

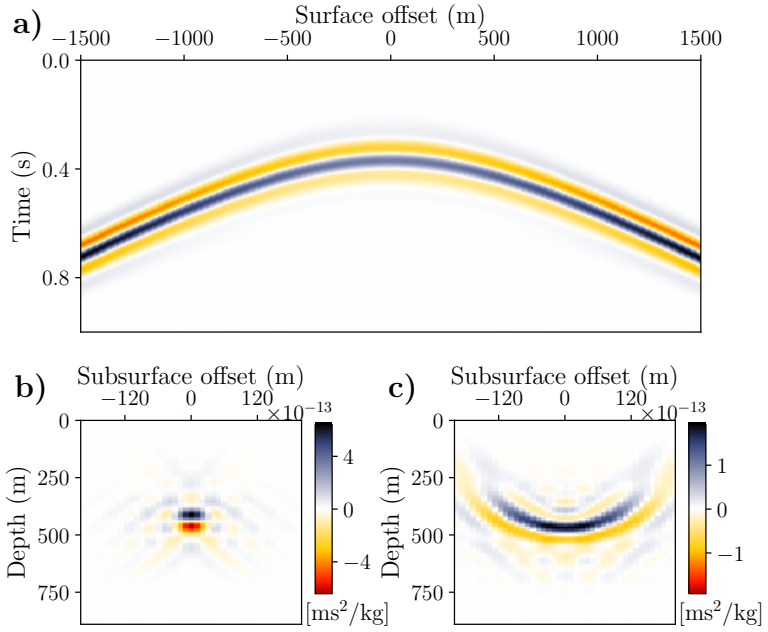

Figure 1.8: Constant density acoustic media. a) Observed data for a single layer Earth model and the corresponding CIG inverted via pseudoinverse Born modeling operator with b) correct and c) incorrect background model (high velocity).

Limitations

The derivative of pseudoinverse Born operator is based on some important hypotheses in which some of them impose limitations in practice. These limitations consists of (ten Kroode, 2012; Hou and Symes, 2015, 2017; Chauris and Cocher, 2017):

- Assumption of infinite dense source/receiver acquisition geometry;
- Absence of turning waves;
- Absence of grazing rays (horizontal tangent direction for the rays connecting a source/receiver location to a subsurface location);
- Assumption of constant-density acoustic media.

In this work, I will mainly address the last issue related to constant-density assumption.

1.4 Motivation, strategy and thesis organization

In the previous sections, I have first reviewed different aspects of seismic imaging: seismic data, scale separation, modeling and inversion. Then, I have detailed the seismic

imaging methods. This thesis mainly focuses on the wave-equation-based methods in the context of linearized inversion. There are two major closely related objectives currently under development: (1) the estimation of multi-parameters beyond the constant density acoustic case; (2) derivation of efficient pre-conditioners for a fast LSRTM. In section 1.3.3, I reviewed the direct inversion as an alternative to iterative scheme. The limitations of pseudoinverse Born operator are also reviewed. The consequence of violation of these limitations has not been fully understood yet. For example:

Q1 : *How does the pseudoinverse operator behave when the input data are not dense at the surface (typically only a few sparse shot gathers)?*

Q2 : *Is the approach always valid in complex media as the theory seems to indicate?*

In reality, density of the Earth is not homogeneous. Although a smooth density model does not affect the kinematics, it modifies the amplitude of the seismic waves, which may lead to complex amplitude variations with offset/angle (AVO/AVA) yielding the question:

Q3 : *How does the density and/or elasticity variations affect the final migrated result obtained by direct inverse? How to develop a direct inverse for the multi-parameter case?*

We illustrate this issue with a numerical result (Figure 1.9) extracted from Farshad and Chauris (2020b). We compare the CIGs (ξ) obtained by application of the pseudoinverse operator using the correct background models (both velocity and density). As expected, in the case of constant density, the energy is focused around zero subsurface offset (Figure 1.9c); whereas it is defocused in case of variable density even in presence of correct background models (Figure 1.9f). The quality of the focusing is measured by energy of the $(h/h_{max})\xi$ for each CIG (Figure 1.9d-h), where h denotes the subsurface offset. The message illustrated in Figure 1.9 is the following: energy is expected to be focused around $h = 0$ for the correct background velocity model. This is not the case in the presence of density variations that would be interpreted as an incorrect background velocity model.

We notice that the same observation was made by Zhou (2020), who proposed the guidelines for the parameters to be used in MVA/IVA. We extracted an image from Zhou (2020) (chapter 5), about the application of the IVA on real data set (Figure 1.10). As Zhou (2020) mentioned in the discussion part, some of the inconsistent events in the CIGs (indicated by dashed ellipses in Figure 1.10b) probably correspond to the impact of the density perturbations.

As discussed in section 1.1.3, adding more parameters into the inversion increases the ill-posedness of the inverse problem (Virieux and Operto, 2009). Ill-posedness is another significant pathology of multi-parameter pseudoinverse Born operator. Overcoming this issue requires robust and versatile regularized optimization approaches, which rises another question:

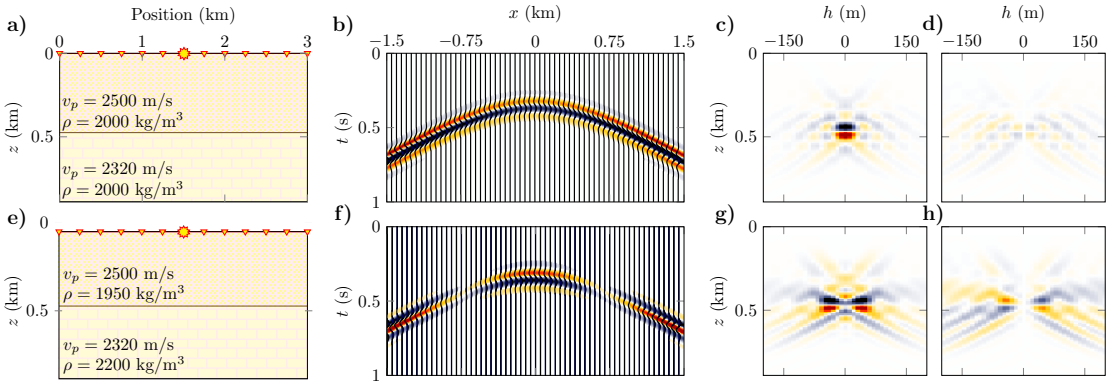

Figure 1.9: A two layered Earth model a) without density variations and the corresponding b) shot gather, inverted c) ξ and d)$(h/h_{max})\xi$. (e,f,g,h) Same as (a,b,c,d) with density variations in the model (from Farshad and Chauris (2020b)).

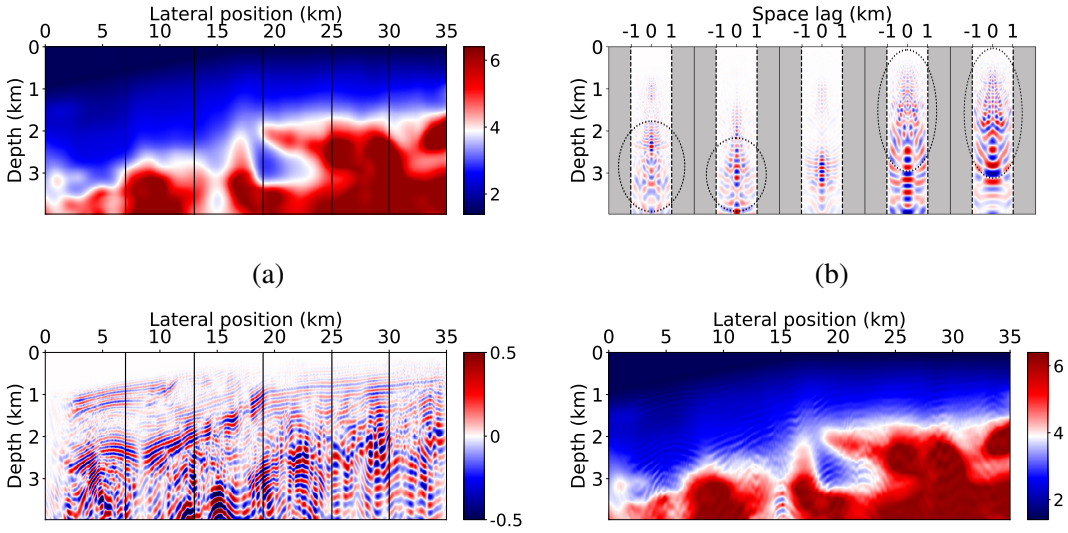

Figure 1.10: Twenty iterations results of IVA for a maximum frequency equal to 10 Hz. a) background velocity model, b) extracted CIGs from $x = 7, 13, 19, 25, 30$ km, c) perturbation velocity model and d) full velocity model equal to the sum of (a) and (c). Dashed ellipses in (b) indicate the defocused energy may be related to density effects (from Zhou (2020)).

Q4 : How to incorporate regularization terms in the multi-parameter direct inverse case?

Our objectives of the thesis are mainly answering the mentioned question. They can be summarized as: (1) investigating the limitations of pseudoinverse operator, (2) ex-

tending the approach beyond pure-acoustic approximation to achieve an efficient multi-parameter imaging tool, (3) incorporating regularization terms to suppress various migration artifacts (i.e. including the ones correspond to parameters cross-talk, sparse shot acquisition geometry, noisy data and complex structures), (4) utilizing multi-parameter pseudoinverse operator as a preconditioner for LSRTM and (5) validating our approach on a real dataset.

1.4.1 Thesis organization

The thesis consists of 5 main parts as follows (Figure 1.11):

- Part I – Background (Chapters 1 and 2),
- Part II – Direct methods – Beyond constant-density acoustics (Chapters 3 and 4),
- Part III – Iterative methods (Chapters 5 and 6),
- Part IV – Real data application (Chapter 7),
- Part V – Conclusions and Perspectives (Chapter 8).

These parts are broken down into chapters, summarized as follows:

- In Chapter 2, I review the essential ingredients in MVA and different quantitative imaging schemes. Further, I review the recent advances, in particular the introduction of direct inverses for a more robust approach. Then I analyze the implementation and practical issues. Finally, I give a brief overview of real data applications, before discussing different perspectives and challenges. This study has been submitted to *Geophysics*.
- In Chapter 3, I extend the applicability of direct inverse from constant-density to variable-density acoustic media. I propose a weighted least-squares method as a generalization of the previously proposed "two-trace method". Using synthetic examples, I compare and discuss these approaches. This study has been published in *Geophysics*.
- In Chapter 4, I extend the applicability of direct inverse from acoustic to elastic media. I restrict the extension to marine environment, with the recording of pressure waves at the receiver positions. Through synthetic numerical experiments, I first demonstrate the consequences of acoustic approximation on elastic data. Then I illustrate that our method can simultaneously invert for elastic parameters in the presence of complex uncorrelated structures, inaccurate background models, and Gaussian noisy data. This study has been published in *Geophysics*.

- In Chapter 5, I show how to implement sparsity-promoting regularization terms in variable-density direct inverse. The fast iterative shrinkage-thresholding algorithm is used to carry out the optimization problem. Through synthetic numerical experiments, I investigate the robustness of my method against different migration artifacts including the parameter crosstalk, interfaces with abrupt truncations, sparse shot acquisition geometry, noisy data, and high-contrast complex structures. This study has been published in *Geophysics*.

- In Chapter 6, I show how to leverage the variable-density direct inverse as a preconditioner to accelerate the convergence rate of LSRTM. I numerically assess the capability of the proposed method in mitigating parameter cross-talk while accelerating the convergence rate. Furthermore, I analyze the shape of the multi-parameter Hessian matrix without and with preconditioner. This study has been published in *Computational Geosciences*.

  In Chapter 7, I apply the variable-density direct inversion scheme to the 2D marine Nankai trough real dataset. I present the preprocessing steps and the advantages of including density parameter in the imaging process. This study has been published in *Geophysical Journal International*.

  In Chapter 8, I summarize the main conclusions and propose issues inviting future research.

The interaction between the thesis chapters is illustrated in Figure 1.11. Chapters 2, 3, 4, 5, 6 and 7 of this thesis have been published/submitted in different journals, meaning that each chapter is self-contained and can be read in isolation.

1.4.2 Contributions

Starting from an existing finite-difference acoustic forward/adjoint propagation code in Fortran, I have built an interface connecting Python and Fortran languages. This provided a balance between computational efficiency (Fortran) and implementation time and effort (Python). I have extended the constant-density pseudoinverse Born operator to account for density variations from the theoretical and numerical points of view. Then I have developed P-SV elastic forward, adjoint and pseudoinverse modeling operators. In both acoustic and elastic cases, the modeling engine for the solution of the 2-D wave equation is staggered-grid finite-difference with flexible accuracy in time and space implemented with perfectly matched layers (PML) absorbing boundary conditions (Bérenger, 1994). The implementation of adjoint operators has been verified with the dot-product test (Claerbout, 2014) and gradient derived with the adjoint-state method (Plessix, 2006b) have been compared with finite-difference computations. My main contributions are

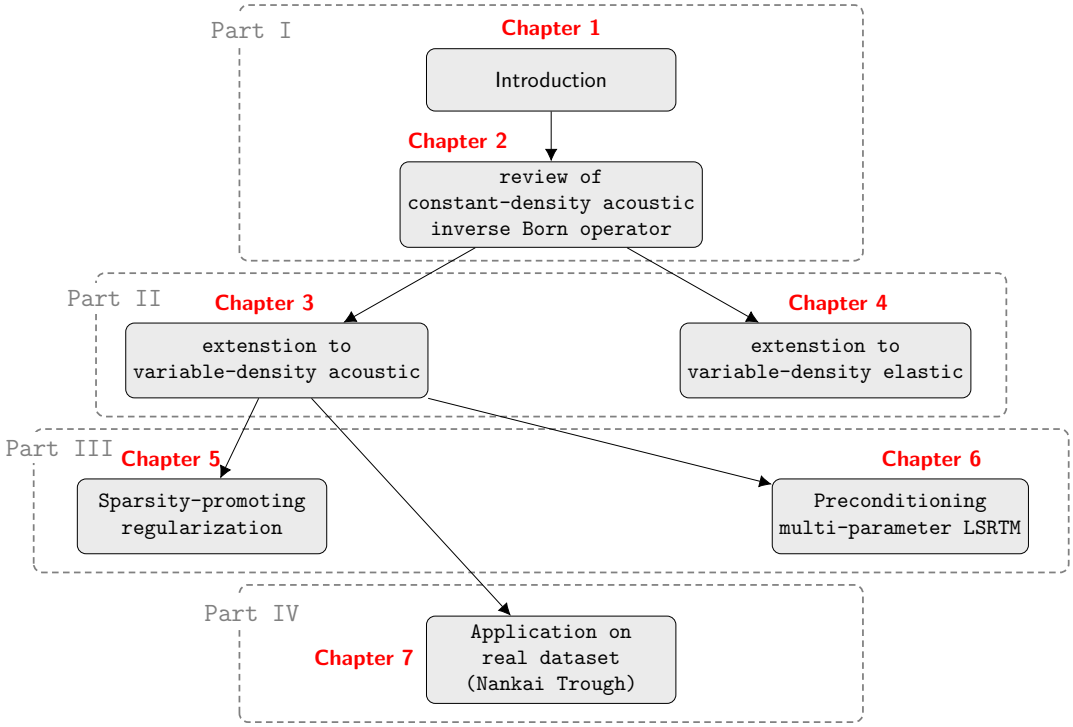

Figure 1.11: Sketch showing the interaction and organization of the thesis chapters.

- extending the pseudoinverse Born operator from constant-density acoustic to variable-density acoustic and variable-density elastic media, as detailed in Part II through Chapters 3 and 4,
- implementing sparsity-promoting regularization terms to suppress various migration artifacts, as detailed in Part III through Chapter 5,
- utilizing pseudoinverse Born operator as a preconditioner in LSRTM, as detailed in Part III through Chapter 6,
- validating the results and investigating the effect of density variations using a real dataset, as detailed in Part IV through Chapter 7.

As explained in this thesis, an important contribution is to introduce a scheme to handle density and regularization, with almost no additional CPU/memory costs: that part indeed does not imply the resolution of direct/adjoint wave equation resolution (more details in Part II and III).

Part of the results has been presented/accepted in

- Farshad, M. and Chauris, H. [2020] From constant- to variable-density inverse extended Born modeling. Geophysics, **85**(4), S217-S232, doi: 10.1190/geo2019-0489.1;
- Farshad, M. and Chauris, H. [2020] Reconstruction of quantitative variable density acoustic reflectivity in the context of velocity model building. 82th EAGE Conference & Exhibition, Workshop Programme, WS09–08;
- Farshad, M. and Chauris, H. [2020] Sparsity-promoting pseudoinverse Born operator in the presence of density variations: an efficient multi-parameter imaging tool. 82th EAGE Conference & Exhibition, Extended Abstracts, Th–Dome5–17;
- Farshad, M. and Chauris, H. [2021], Sparsity-promoting multi-parameter pseudoinverse Born inversion in acoustic media, Geophysics, **86**(3), S205–S220, doi: 10.1190/geo2020-0527.1;
- Farshad, M. and Chauris, H. [2021], From acoustic to elastic inverse extended Born modeling: a first insight in the marine environment, Geophysics, 2021, doi: 10.1190/geo2020-0916.1;
- Farshad, M. and Chauris, H. [2021], Accelerating the multi-parameter least-squares reverse time migration using an appropriate preconditioner, Computational Geosciences, 2021, doi: 10.1007/s10596-021-10089-4;
- Farshad, M., Chauris, H. and Noble, M. [2021], The importance of including density in multi-parameter asymptotic linearized direct waveform inversion: a case study from the Eastern Nankai Trough, Geophysical Journal International, doi: 10.1093/gji/ggab405;
- Farshad, M. and Chauris, H. [2021], Asymptotic Born inversion in elastic media. 82nd EAGE Conference & Exhibition, Extended Abstracts;
- Farshad, M. and Chauris, H. [2021], Preconditioning the variable-density least-squares reverse time migration via asymptotic Born inversion, 82nd EAGE Conference & Exhibition, Extended Abstracts;
- Farshad, M. and Chauris, H. [2021], Elastic versus acoustic asymptotic extended Born inversion for marine surveys, 82nd EAGE Conference & Exhibition, Workshop Programme, WS02;

and/or submitted to publication,

- Chauris, H. and Farshad, M. [2021], Seismic image-domain wave-equation-based methods – an overview, Geophysics, submitted.

Chapter 2

Review of constant-density direct methods

Contents

2.1 Abstract **34**
2.2 Introduction **34**
2.3 MVA principle **36**
2.4 Focusing principle **41**
2.4.1 Relationships between velocity errors and focusing 41
2.4.2 Expected convex shape 48
2.4.3 Gradient artifacts 49
2.5 Extensions to a more complex physics **51**
2.5.1 Beyond the constant-density acoustic case 51
2.5.2 Beyond the single scattering approximation 54
2.6 Implementation aspects **58**
2.6.1 Reducing the number of shots 59
2.6.2 Target-oriented MVA 60
2.7 Discussion **61**
2.7.1 Resolution 62
2.7.2 Tomographic Hessian 63
2.7.3 Multi-parameter imaging 63
2.8 Conclusions **64**
2.9 Acknowledgements **64**
2.10 Appendix I: Gradient derivation **64**

2.11 Appendix II: Approximate inverse **66**
2.12 Appendix III: Equivalences between data and image domains . . . **69**

Résumé

Ce chapitre comprend l'article intitulé « Seismic automatic wave-equation Migration Velocity Analysis – an overview », soumis pour publication à la revue *Geophyics*. Il contient quelques nouvelles notions extraites des résultats des travaux présentés dans les chapitres 3,4, 5 et 6.
Les techniques d'Analyse de Vitesse par Migration (MVA en anglais) ont pour but de déterminer les structures du sous-sol en deux étapes. Dans un premier temps, l'objectif est d'obtenir une image de la subsurface (partie migration). Dans un second temps, il s'agit d'analyser la qualité de cette image pour mettre à jour le macro-modèle (partie tomographie). Un avantage majeur de cette approche est que la fonction objective associée est en principe convexe et que l'optimisation entièrement automatique. Cependant, il existe une grande différence entre les propriétés théoriques et les aspects pratiques : seules peu d'applications automatiques sur données réelles ont été publiées et l'Analyse de Vitesse par Migration n'est pas une routine standard dans l'industrie. Nous donnons un aperçu de trois grands aspects pour expliquer le fossé entre théorie et pratique. (1) Nous revisitons le principe de focalisation qui dit que la focalisation est obtenue pour le vrai macro-modèle. Ce n'est pas vrai en général, si par exemple la compensation pour l'illumination irrégulière du sous-sol n'est pas corrigée lors de la phase d'imagerie. Les propriétés de convexité sont associées à des méthodes qui sont très sensibles aux bruits cohérents. (2) Pour les applications standards, la physique la plus simple est employée (cas acoustique à densité constante), avec l'hypothèse de diffraction simple : c'est un second facteur limitant. (3) Nous discutons aussi des aspects d'implémentation, avec actuellement des besoins prohibitifs pour les applications 3D réalistes. Dans chaque cas, nous proposons des solutions. Nous donnons des arguments comme quoi les applications 3D ne sont possibles que si les stratégies les plus efficaces sont combinées dans une seule approche. Les travaux futurs incluent des pré-traitements spécifiques sur les panneaux de focalisation pour traiter le bruit cohérence et la décimation du nombre de points de tir. Plus de recherche est aussi nécessaire pour considérer le champ d'onde complet.

Seismic image-domain wave-equation-based methods – an overview
Hervé Chauris and Milad Farshad
Geophysics, 2021, submitted

2.1 Abstract

Migration Velocity Analysis techniques aim at determining the structure of the subsurface in a two-step process. It consists of first obtaining an image of the subsurface (migration part) and then analyzing its quality to update the macro-model (tomographic part). As a major advantage, the associated objective function is expected to be convex and the optimization process be fully automatic. A large discrepancy exists between the attractive theoretical properties and the practical applications: only few automated real data applications have been yet published and Migration Velocity Analysis is not a routine processing within the industry.
We review here three main aspects to explain such a gap between the theory and the practice. (1) We revisit the focusing principle claiming that the focusing is obtained for the correct macro-model. This is not true in general if for example the compensation for uneven illumination is not properly taken into account during the derivation of the image. The expected convex properties are associated with methods that are most sensitive to coherent noise. (2) For the standard applications, the simplest physics is considered (constant-density acoustic wave equation), with the assumption of a single scattering approximation: this is a second major limiting factor. (3) We also discuss the implementation aspects, with currently an excessive computational requirement for realistic 3D applications. In each case, we propose some remedies. We argue that 3D applications will only be possible if the most advanced and efficient strategies are combined within a single approach. Future work includes some specific preprocessing steps on the focusing panels to handle coherent noise and shot decimation. More research is also needed for considering the complete wavefield.

2.2 Introduction

Migration Velocity Analysis (MVA) family is a two-step technique to image the subsurface: in a given macro-model, seismic images are first derived from different subsets of the input data. The optimal result is then obtained when these images are consistent (Symes, 2008b). Despite the fact that such approaches are limited to the use of reflected waves, the MVA principle is simple and attractive as an automatic subsurface imaging tool. For some specific objective functions, the formulation is expected to have a convex shape, meaning that the minimum obtained with a gradient-based approach is

the global minimum (Gockenbach and Symes, 1995; Symes, 1998a, 2008b; Stolk and Symes, 2003; Sun and Schuster, 2001); the method fully exploits the data redundancy for dense acquisition systems; individual reflectors are more easily interpretable in the image domain than in the original data domain; the migration part is an efficient stacking tool to enhance the signal to noise ratio; finally, MVA approaches do not necessarily need a picking step. But *currently, no fully automated wave-equation-based MVA approaches really exist at an industrial scale*, at it is the case for Reverse Time Migration (RTM) or Full Waveform Inversion (FWI) (Etgen et al., 2009; Virieux and Operto, 2009; Fichtner, 2010). In this introduction, we first give a general MVA history in brief. We then discuss a few 3D real data MVA applications published in the literature, as a motivation for this work. The next section ("MVA principle") provides more technical details on the automatic wave-equation-based MVA approaches mainly considered here.

The first general MVA approaches were defined in the '70s (Gardner et al., 1974a; Sattlegger, 1975). We may underline four main alternatives: ray-based versus wave-equation-based imaging schemes, picking versus semblance, different kinds of focusing panels and from simple to more complex velocity models. Initially, ray-based methods were developed for the velocity model building from the analysis of reflected data (Woodward et al., 2008). Later, the wave-equation approaches were introduced for which the sensitivity kernels are frequency dependent (Woodward, 1992). The historical approaches include a picking phase of the reflected events to characterize the quality of the residual move-out, (among others Al-Yahya (1989); Liu and Bleistein (1995); Liu (1997); Lafond and Levander (1993); Eckhardt (1994); Wang et al. (1995); Audebert et al. (1997); Meng et al. (1999b,a)). Then fully automatic schemes were developed, through semblance analysis or more generally focusing principles. The third aspect is the precise definition of the focusing panels, also known as Common Image Gathers (CIGs). They are a function of an extension parameter, with several choices, from shot or surface offset coordinate (surface binning), to angle, time-lag, or subsurface distance (survey sinking) (Clapp and Biondi, 2000; Biondi and Symes, 2004; Sava and Biondi, 2004; Symes, 2008b, 2009). Finally, the applications ran from 1D acoustic (Symes, 1993; Gockenbach and Symes, 1995; Symes, 1998b; Li and Symes, 2007) to 3D anisotropic visco-elastic models, from one-way (Shen and Calandra, 2005; Albertin et al., 2006) to two-way wave equation approaches (Kern and Symes, 1994). All these notions will be detailed in the next section, with a particular attention to the picking-free wave-equation-based approaches.

We now provide an overview of some 3D real data applications published in the literature. We indicate the main aspects that let the approach feasible, without entering to much in the technical details but to underline some of the compromises for a successful application. One can first note that only a few wave-equation-based MVA approaches have been published with real data applications, here mainly 2D applications (Chauris and Noble, 2001; Mulder and ten Kroode, 2002; Alkhalifah, 2005; Shen and Symes, 2008; Weibull and Arntsen, 2013, 2014; Mulder, 2014; Li et al., 2014; Lameloise et al.,

2015; Zhang et al., 2015a; Diaz and Sava, 2017; Li et al., 2019; Zhou, 2020), including passive data (Shabelansky et al., 2015; Witten and Shragge, 2017). The 3D real data applications are even less frequent. Guerra and Biondi (2011) discussed a North Sea example in the presence of salt. As a trade-off, they assumed a sparse reflector distribution with the use of a generalization of the exploding reflector technique. Yang and Sava (2013) presented some results on another North Sea OBC data set, with a total of 250 receiver gathers, used for reciprocity reasons as shot gathers during the processing in practice. For efficiency reasons, Yang and Sava (2013) rely on a 1D extension (time-lag extension, defined in the next section), instead of a 2D extension that would theoretically be needed here. The second important ingredient for the 3D application was to select sparse zones where the focusing is evaluated, reducing the computational cost and memory requirements. On a Gulf of Mexico marine data set in a salt context, Tang and Biondi (2013) used a target-oriented MVA approach, assuming the macro-model to be known outside of the zone of interest. Li et al. (2016) presented a 3D application in the same geological context. The interesting aspect is the VTI anisotropy. Once the migrated images have been obtained, only one third of the total number of shots are used to update the macro-model. Moreover, they used a specific focusing criteria (the stack power maximization approach detailed in the next section) that requires a relatively correct initial macro-model. Finally, Shen et al. (2018) performed MVA on a 3D North Sea data set, with the evaluation of velocity and attenuation fields in the context of shallow subsurface gas chimneys. The one-way wave-equation was used to limit the CPU cost, reducing the maximum possible dips to be imaged.

This short review, without being exhaustive nor without yet entering into the details, has indicated that only rare 3D automatic MVA real data applications at an large scale are presented. In each case, a compromise was required between quality and computational costs (Fei et al., 2009). The main objective in this paper is to better understand if the fundamental reasons are related to theoretical or practical issues. The article is organized as follows: we first review the MVA principles (wave propagation, imaging principle, focusing panels and quality criteria), concentrating on the wave-equation-based picking-free schemes. In the next three sections, we then discuss the current limitations and possible remedies along three axis: (1) assumed focusing principle for the correct macro model, (2) limited physics and (3) implementation aspects. In all cases, we start with known issues and propose some existing or to be further developed elements for addressing them. We conclude by discussing a plan for the future developments, from simple modifications to more complex investigations.

2.3 MVA principle

This section provides the main MVA elements. In particular, it introduces the notion of “extended domain” to analyze the quality of the focusing panels. The limitations are

analyzed in the three next sections. MVA is a two-step process (Symes, 2008b, 2009; Sava and Biondi, 2004). In the first stage, the input data, consisting of only reflected waves, is migrated to obtain partial images. For a given spatial position at the surface, CIGs are created. The quality of the focusing panels is then analyzed to evaluate the consistency between the partial images (Algorithm 1).

Algorithm 1: MVA algorithm

```
1 Outer loop: for each tomographic iteration do
2     Inner loop: for each reflectivity update do
3         Construct (iteratively) the partial images
4     Evaluate the quality of the focusing
5     Update the macro-model
```

For the inner loop, the reference or macro-velocity model is fixed. Migrated images are obtained by combining source and receiver wavefields through an imaging condition (Gardner et al., 1974a; Yilmaz, 2001). We refer to Etgen et al. (2009) for a review on the migration part. In this article, we mainly concentrate on full-wave propagation with two-way operators, including Reverse-Time Migration and its variants (Baysal et al., 1983; McMechan, 1983; Whitmore, 1983). For each shot, the incident wavefield and the backward receiver wavefield should be computed. Claerbout (1971) proposes two imaging condition: the cross-correlation and the deconvolution one. The impact of the imaging condition is analyzed in the "Focusing principle" section.

For a possible evaluation of the quality of the macro-model used for migration, it is essential to design a reversible mapping between the data and image domains: the application of these operators between the two domains should preserve the information contained in the signal (Symes, 2008b). In practice, it means that the dimensions of the data and image spaces should be the same. For example in 3D, the data domain is described by five variables (s_x, s_y, r_x, r_y, t) (the two source and receiver positions, the time axis), where the standard model space by only three $\mathbf{x} = (x, y, z)$. The extended model space depends on $(x, y, z, \mathbf{e})$, where $\mathbf{e}$ is an extended parameter. Several possibilities have been proposed (Sava and Vasconcelos, 2011). The binning-based approaches include common shot and offset schemes. [1] Common shot migration ($\mathbf{e} = (s_x, s_y)$ in 3D) is the original strategy (Kern and Symes, 1994; Huang, 2016). It is directly compatible with the acquisition design and the traditional shot-oriented wave-equation-based implementation. [2] Common offset migration ($\mathbf{e} = (s_x - r_x, s_y - r_y)$) is another surface-oriented approach (Deregowski, 1990). It does not suffer from edge effects but is not efficient for wave-equation-based approaches, at least for a limited number of surface offset H values. The other strategies (survey-sinking) are subsurface-oriented. In particular, [3] common angle migration ($\mathbf{e} = (\theta, \varphi)$) refers to the two imaging angles in the depth domain in 3D (Xu et al., 2001; Rickett and Sava, 2002; Sava and Fomel,

2003; Biondi and Symes, 2004; Stolk and Symes, 2004). A more recent alternative is the subsurface offset (or space lag) $\mathbf{e} = \mathbf{h}$, different from the surface offset H. The imaging condition considers two different positions $\mathbf{x} - \mathbf{h}$ and $\mathbf{x} + \mathbf{h}$, where $\mathbf{h} = (h_x, h_y)$ usually in 3D. Finally, Sava and Fomel (2006); Yang and Sava (2011) also proposed a time shift extension τ, with a 1D extension. The recommended choices are discussed in the "Focusing principle" section.

Let $\xi(\mathbf{x}, \mathbf{e})$ be the migrated image at a position $\mathbf{x}$ and for a given extension $\mathbf{e}$. It depends on the observed reflected data in a linear way but in a non-linear way on the macro-model. A CIG panel is defined for a fixed (x, y) position and varying $(z, \mathbf{e})$ values. As pointed out by Montel and Lambaré (2019); Montel and Lambaré (2019), the shape of the events in the CIGs depend on the way the panels are built. For an intuitive understanding, we compare from the high frequency asymptotic point of view the shape of events for different homogeneous macro-models, corresponding to under-estimated, correct and over-estimated velocities (Figure 2.1). We investigate the surface (H) and subsurface (h) 2D cases. Suppose one would pick three locally coherent events with a typical hyperbolic shape in the data domain (Figure 2.1, left). They are characterized by a (s, r, t) position as well as two slopes along the source and receiver positions (Billette and Lambaré, 1998; Lambaré, 2008). We only represent here the receiver and time axis. In the extended domain, migration is a mapping from the data to the image domains (Chauris et al., 2002a). In both H and h cases, the migrated depths increase with the mean velocity values. The surface offset corresponds to the case where the upward curve becomes horizontal for the correct velocity model and downward for an over-estimated value (Figure 2.1, middle). The behavior is different for the subsurface offset case, with a focusing of energy around $h = 0$ in the optimal case (Figure 2.1, right).

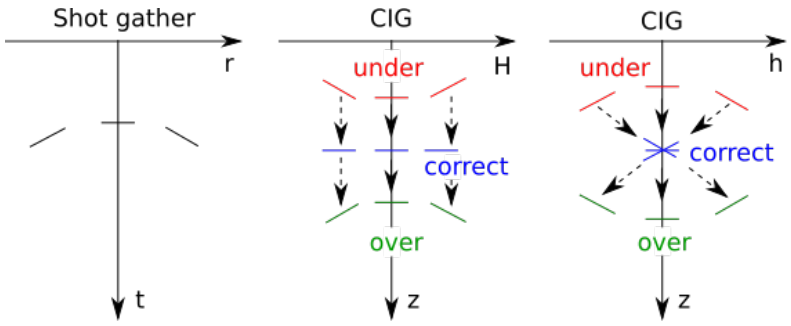

Figure 2.1: Schematic illustration of the migration of three locally coherent events picked in the data domain (left), with a common surface offset (middle) and subsurface offset (right) schemes. The red, blue and green colors correspond to under-, correct and over-estimated velocity values.

The same behavior is observed when the modeling and adjoint operators are computed with a wave-equation-based approach. We consider a model consisting of three

interfaces embedded in a homogeneous model. For the subsurface offset case, we test 11 values between 1500 to 2500 m/s (Figure 2.2). The optimal focusing is obtained for the correct value (2000 m/s). A more detailed analysis of the shape, including the amplitude aspect, can be found in Mulder (2014); Montel and Lambaré (2019).

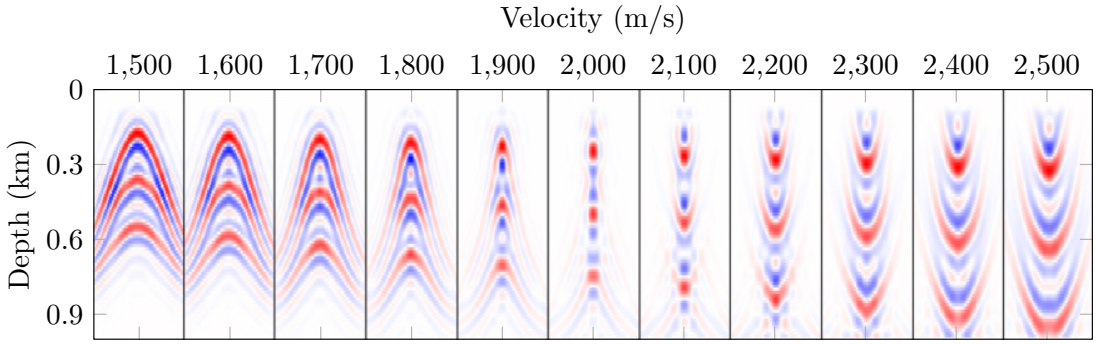

Figure 2.2: Shape of events in CIGs for the subsurface case and for homogeneous velocity values between 1500 to 2500 m/s, every 100 m/s The local horizontal axis for each panel refers to the subsurface offset, from -100 to +100 m/s.

The annihilator is an operator A_{n} acting on to the extended migrated result $\xi(\mathbf{x}, \mathbf{e})$ such that the norm of $A_{\text{n}}\xi$ is minimum for the correct macro-model. For the surface parameters in 2D and the shot s or surface offset H cases, A_{n} simply reads $\partial/\partial s$ and $\partial/\partial H$: the optimal model is obtained when events are horizontal within the focusing panel (Figure 2.1, middle). The same expression also applies to the image angle $\partial/\partial\theta$. For the other subsurface parameters (subsurface offset h and time delay τ), the objective is to focus energy around $\mathbf{e} = 0$. This leads to a different formulation, namely a simple multiplication by $A_{\text{n}} = h$ or $A_{\text{n}} = \tau$. Despite different expressions, there is a clear connection: by considering $A_{\text{n}} = ih$ instead of $A_{\text{n}} = h$, leading to the same norm $\|A_{\text{n}}\xi\|$, a multiplication by ih can also be seen as a derivative with respect to k_h in the Fourier domain, where k_h would be the dual variable of h (Biondi et al., 2019).

Based on the annihilator, the MVA outer loop objective function reads (Symes, 2008b)

$$J_{\text{MVA}} = \frac{1}{2}\frac{\left\|A_{\text{n}}\xi\right\|^2}{\left\|\xi\right\|^2}. \tag{2.1}$$

Some formulations do not necessarily include the denominator term. The reason for the normalization is to decouple the focusing effects from the energy aspects (Chauris et al., 2002a): the modification of the macro-model has also an influence on the energy of the migrated section due to the geometrical spreading terms in the Green's function and to the fact that some events for large velocity values may be localized beyond the

maximum target depth. The normalization provides a more symmetric objective function with respect to the optimal model.

The standard adjoint state technique applies for the derivation of the tomographic gradient, that is the gradient of the J_{MVA} objective function with respect to the macro model m_0 (Plessix, 2006a; Plessix et al., 2000; Weibull and Arntsen, 2013; Chauris and Cocher, 2017). This model is also known as background model. We refer to Appendix 2.10 for a more detailed derivation. As the image (or reflectivity section) ξ is obtained as the cross-correlation between the incident wavefield and the back-propagated wavefield, it contains two Green's functions that each depends on the macro-model, leading to two main terms in the tomographic gradient. In the expressions presented in Appendix 2.10, we recognize the need for solving two additional wavefields that are later correlated with the source and receiver wavefields. The final quantities are proportional to $A_{\text{n}}^T A_{\text{n}} \xi$. In practice, this is a second derivative in 2D for the surface offset $-\partial^2/\partial^2 H$ or a multiplication by h^2 for the subsurface offset case.

Ray quantities are useful to interpret the gradient (Figure 2.3). Let consider a position $\mathbf{y}$ in the subsurface and a given subsurface offset $\mathbf{h}$. The first contribution of the gradient in $\mathbf{x}$ connects $\mathbf{y}+\mathbf{h}$ to $\mathbf{x}$ and $\mathbf{x}$ to the receiver position. The contribution is specular (i.e. adds in a high frequency asymptotic sense) only if $\mathbf{x}$ belongs to the ray joining $\mathbf{y}+\mathbf{h}$ to the receiver position, meaning that the update in a smooth macro-model is located above the reflector $\mathbf{y}$. The same holds for the second contribution to the gradient on the source side (Symes, 1998b).

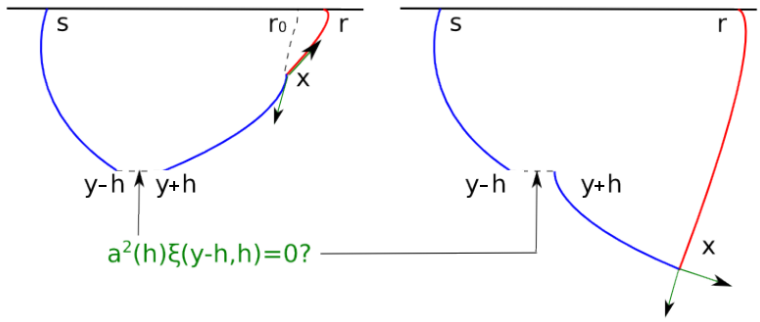

Figure 2.3: Interpretation of the tomographic gradient: a non-zero contribution is obtained when there is a $\mathbf{y}$ position such that $A_{\text{n}}^2(\mathbf{h})\xi(\mathbf{y}-\mathbf{h},\mathbf{y}) \neq 0$ and when there is another $\mathbf{x}$ position along the ray joining $\mathbf{y}+\mathbf{h}$ and a receiver position $\mathbf{r}$, here possibly $\mathbf{r}_0$. The same holds for the source side. The second condition implies that in a smooth macro-model, the tomographic update is only above the reflector (left) and not below (right).

Here is a complementary view to explain how MVA is able to create long wavelengths from the analysis of short wavelengths contained in ξ. We consider a single

source and two receivers (Figure 2.4). The interface in ξ (dashed-dotted line) is obtained as the correlation between the source (red solid lines) and the receiver wavefields (blue solid lines) (Figure 2.4, top). For the first part of the macro-model gradient, the source incident wavefield excites the interface that acts as a secondary source, generating the red dashed lines (here one up-going and another down-going wave). This new wavefield is correlated with the original receiver wavefield (blue solid lines) (Figure 2.4, middle). Compared to the construction of the interface where the wavefields intercept along the interface, here the correlation is non-zero above the interface and zero below, leading to long-wavelength updates (gray zone). The same effect is obtained for the second part of the gradient, where now the receiver wavefields excite a secondary wave (dashed blue line), correlated with the incident source field (solid red line) for the part above the reflector.

We now have the main structural elements to analyze in the next sections three aspects currently limiting the practical MVA applications at a large scale and related to focusing principle, simple physics and computational requirements.

2.4 Focusing principle

We first discuss the assumption that the best focusing is obtained for the correct macro-model. We then comment on the choice of the objective function concerning convexity and robustness arguments. Finally, we discuss the shape of the macro-model velocity update in a gradient-based optimization approach. In each case, we propose some remedies.

2.4.1 Relationships between velocity errors and focusing

Even in the correct macro-model, the optimal focusing is not necessarily obtained in practice. The first reason is the presence of triplications, possibly existing even in a smooth model (Červený, 2001). For surface-oriented CIGs, significant part of the energy may travel along multiple paths between source, receiver and the scattered image point (Nolan and Symes, 1997; Brandsberg-Dahl et al., 2003; Stolk and Symes, 2004). The proposed solution is to consider the survey-sinking approaches, such as the subsurface offset extension for which such artifacts are not present, under the condition that there is no diving waves (i.e. turning energy from the source that reaches the receiver without being reflected, Stolk and Symes (2004) and de Hoop and Stolk (2006)).

The second issue is linked to the dip limitation: even with two-way wave-equation operators, the annihilator of the standard subsurface Differential Semblance Optimization (DSO) approach is a multiplication by $\sqrt{h_x^2 + h_y^2}$ in 3D, where (h_x, h_y) are the space lags in the horizontal plane. In that case, vertical dips are not properly handled. The first possibility would be to introduce an extra dimension h_z, leading to

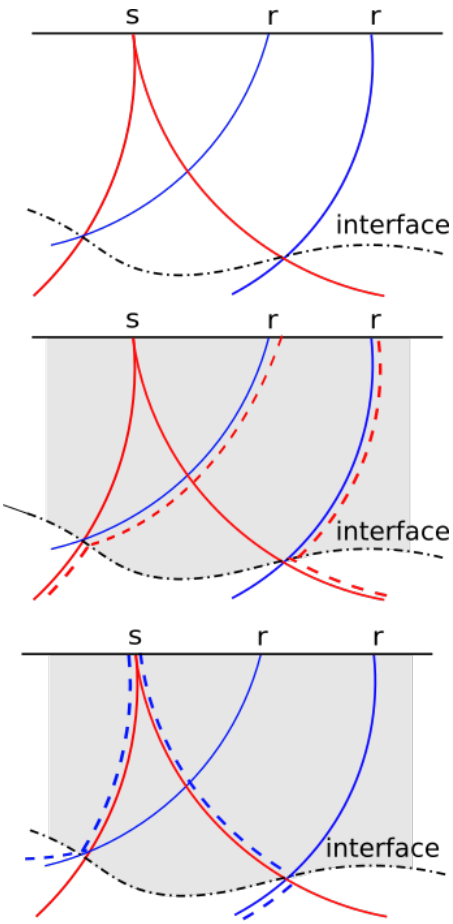

Figure 2.4: Gradient principle for the MVA case: the interface is created as a correlation between the source (solid red lines) and the receiver (solid blue lines) wavefields (top). The macro-model gradient is obtained as the sum of two contributions (see the text for details), leading to a large-scale perturbation above the interface (middle and bottom, gray zone).

$\sqrt{h_x^2 + h_y^2 + h_z^2}$ or to combine with the time-lag τ as $\sqrt{h_x^2 + h_y^2 + (\tau v)^2}$, where v is the typical macro-velocity value at the considered spatial location (Díaz and Sava, 2015; Biondi and Symes, 2004; Shan and Biondi, 2005; Sava and Fomel, 2006). For both solutions, the implementation becomes too prohibitive in 3D (see the section on "Implementation aspects").

The third issue is related to an incomplete acquisition geometry (Mulder, 2014). For a dense but truncated acquisition, the migrated image suffers from edge effects, whose size is controlled by half the maximum surface offset H. For the macro-model gradient, the imprint of the acquisition is amplified, and the 1D invariance expected in a purely 1D model would only be observed within a zone reduced by a distance H on both sides. Along the same line, the user parameters play an important role (Mulder, 2014). Zhou et al. (2020b) demonstrated that the parameters such as the maximum surface offset available or the maximum selected subsurface offset can lead up to a change in the sign of the macro-model update. In the first case (limited surface offsets), there is not enough data to obtain focused energy within the CIGs: residual migration smiles are visible, pushing towards artificial lower velocity models. For the second case (limited subsurface values), energy is truncated in the CIGs, also leading to modified macro-model updates. Following the analysis by Mulder (2014), the authors provide guidelines for an optimal choice, based on the analysis in a homogeneous models with embedded horizontal reflectors. The same rules apply to more complex models such as the Marmousi case (Versteeg and Grau, 1991; Zhou et al., 2020b).

More generally, even for regular acquisition geometries, wave propagation in a non-homogeneous model leads to some uneven illumination of the subsurface. Before addressing this aspect, we point out the last reason leading to apparent defocusing events related to an incorrect physics: for a purely constant-density acoustic imaging tool, it has been observed that CIGs associated with variable-density observed data, leading to amplitude variations along offset, are not focused for the correct velocity macro-model (Dafni and Symes, 2018; Farshad and Chauris, 2020b). This issue will be addressed in the section "Extension to a more complex physics".

We now detail the part on the illumination part. Uneven illumination of the subsurface leads to defocusing effects indistinguishable from those due to velocity errors (Yang et al., 2013; Yang and Sava, 2015). The general solution is to consider the Hessian associated with the inner loop (construction of the image or migration part) to deblur the standard adjoint image. The first possibility consists of successfully applying a modeling operator F followed by its adjoint F^t to the image m_0, leading to $m_1 = F^t F m_0$. This was for example applied to one-way operators for a true-amplitude migration scheme by Kiyashchenko et al. (2007). One recognizes the migration Hessian operator $F^t F$. The same approach is used to study the effect of the Point Spread Functions (PSFs) (Lecomte and Gelius, 1998; Wapenaar and van der Neut, 2010). Then an optimal B operator is determined, typically by minimizing $||Bm_1 - m_0||^2$ and applied to the m_0 image (Guitton,

2017). The B operator is an approximation of the inverse of the Hessian. Valenciano et al. (2006); Tang and Biondi (2011) proposed a target-oriented approach combined with a phase-encoding strategy for a more affordable computational scheme.

A more general approach is to further apply the forward F and adjoint F^t operators, leading to the least-squares iterative migration (LSM) result (Nemeth et al., 1999; Clément et al., 2001; Hicks and Pratt, 2001; Rickett, 2003; Guitton, 2004; Clapp et al., 2005; Yang et al., 2013; Fletcher et al., 2015). Once the image part is defined, the outer loop updates the macro-model. The gradient with respect to the macro-model can be obtained with the adjoint state method (Plessix, 2006b), either after stopping the iterative process after a given number of iterations or by assuming the convergence of the inner loop (Cocher et al., 2017). The first approach requires storing all intermediate steps and can be achieved with an automatic differentiation code, whereas the latter is easier to implement. More fundamentally, instabilities have been observed on the MVA gradient (Liu et al., 2014; Huang, 2016; Cocher et al., 2017). The reason is that the optimal image in the inner loop never completely achieved. Then the tomographic gradient contains an additional term involving a second-order derivative of the modeling operator, leading to unstable tomographic scheme (Huang, 2016). A more intuitive explanation is related to existence of the null space: little energy at far subsurface offset h could appear to slightly decrease the data fit. However, it has a large impact on the adjoint source in the MVA scheme because of the multiplication by h^2, leading to strong distortion in the macro-model gradient. Different regularizations have been proposed (Liu et al., 2014; Huang, 2016). The most promising one is to modify the definition of the tomographic objective function, while preserving the preconditioned iterative migration part (Cocher et al., 2017).

As an alternative to the iterative migration and the intrinsic unstable tomographic gradient, the direct inverse strategy (Virieux et al., 2016) was recently proposed (ten Kroode, 2012; Hou and Symes, 2015, 2017; Chauris and Cocher, 2017). By direct inverse or approximate or asymptotic inverse or also pseudo-inverse, we mean an operator that provides an approximate inverse of the inner-loop modeling operator within a single iteration.

A long history on quantitative imaging started with approximate inverse in the high frequency asymptotic context (Beylkin, 1985; Bleistein, 1987; Schleicher et al., 1993; Lambaré et al., 1992; Jin et al., 1992; Chavent and Plessix, 1999; Thierry et al., 1999; Operto et al., 2000; Lambaré et al., 2003). The standard adjoint operator (migration scheme) was modified such that the Hessian of the data fit objective function becomes close to the identity operator. It consists mainly of introducing additional weights in the formulation. Then the same approach was developed for one-way operators (Plessix and Mulder, 2004; Zhang et al., 2005, 2007; Joncour et al., 2011) and more recently for two-way operators (ten Kroode, 2012; Hou and Symes, 2015; Qin et al., 2015; Qin and Lambaré, 2016; Hou and Symes, 2017; Chauris and Cocher, 2017). In both cases, the inverse should be understood as an approximate inverse as it is only an inverse in

the high frequency asymptotic sense, even if the final formulation does not contain ray quantities.

Let define $F^\dagger$ be the approximate inverse. It means that for any data d, we expect $d \simeq F\,F^\dagger d$. In the first step, the application of the (right) approximate inverse creates a reflectivity model $F^\dagger d$. The data modeled in it should be the same as the original data. We now detail the different formulations relying on the two-way wave-equation operators (ten Kroode, 2012; Hou and Symes, 2015, 2017; Chauris and Cocher, 2017). They all propose an inverse formula for the Born or Kirchhoff modeling approaches in the extended domain. The main aspects shared by the different strategies are the following: (1) the inverse is defined in a high frequency asymptotic sense; (2) the final formulation only contain wave-equation terms (typically vertical spatial derivatives and time derivatives) and no ray-based terms (such as travel times or slopes); (3) their expression is very similar to the one for the adjoint.

The main ingredients for the derivation are the following. (1) Replace the Green's functions G_0 by their asymptotic expression and apply the stationary phase approximation to consider specular contributions; (2) compute the determinant related to the change of variables, here in 2D, from the data domain (s, r, ω) to the dual wave number domain (k_x, k_z, k_h) associated with (x, z, h); (3) acknowledge the fact that the geometrical spreading amplitudes terms A in the Green's function cancel out by some terms in the determinant, more precisely that

$$A^2(\mathbf{s}, \mathbf{x}) = \frac{1}{8\pi} \frac{v_0(\mathbf{s})}{\cos\beta_s} \left| \frac{\partial \theta_s}{\partial s_x} \right|, \tag{2.2}$$

where β_s is the emerging angle at the source position, θ_s the angle at the image point and $v_0(\mathbf{s})$ the velocity at the source position (Zhang et al., 2005; ten Kroode, 2012; Hou and Symes, 2015); (4) apply additional time and spatial derivatives to remove remaining ray quantities. The final result is multiplied by a term that only depends on the macro-model v_0. Typically, the $\cos\theta_s$ term is removed by using

$$\partial_{s_z} G_0(\mathbf{s}, \mathbf{x}, \omega) \simeq \frac{i\omega}{v_0(\mathbf{s})} \cos\beta_s G_0(\mathbf{s}, \mathbf{x}, \omega). \tag{2.3}$$

where ∂_{s_z} refers to the partial derivative with respect to the vertical component of the source position. More details on the approximate inverse, including a summary of four different schemes, are provided in the Appendix 2.11 section. The main limitation of the high frequency asymptotic inverse is the assumption of an infinite dense acquisition system. However, Hou and Symes (2016b); Chauris and Cocher (2018); Farshad and Chauris (2021a) indicated how to use the approximate inverse as a preconditioner. The simplest strategy is to rely on a weighted linear conjugate gradient, with modified norms in the data and model spaces (Hou and Symes, 2016a).

The substitution of F^T by $F^\dagger$ has a clear impact, not only on the migration results (Hou and Symes, 2016b) but also on the MVA result (Lameloise et al., 2015; Chauris and

Cocher, 2017; Hou and Symes, 2018). This leads to the definition of Inverse Velocity Analysis (IVA, Hou and Symes (2018)), where the adjoint or the standard image is replaced by the inverse. We provide a simple example in the case of the subsurface offset. The migrated image ξ is first obtained through the application of the adjoint, here in a model with under-estimated velocities (Figure 2.5). The focusing panel displays a main downward event, consistent with the expected shape (Figure 2.1), but also events with opposite curvatures. These are related to the maximum offset and truncations in the input data. In practice, they can be attenuated by tapering the input data, but here we want to stress that the limitation of the MVA approach. The main reason is that the adjoint source is obtained by multiplication by h^2, enhancing noisy data at large subsurface offset values. This leads to an objective function with a minimum shifted towards lower values here, as a compromise between the focusing of the real event and the focusing of the artifacts, always with positive curvature (Figure 2.5b). As a consequence, the macro-model gradient exhibits strong oscillations above the interface (Figure 2.5c). A pure tomographic update would be a homogeneous zone, similar to the gray zone (Figure 2.4).

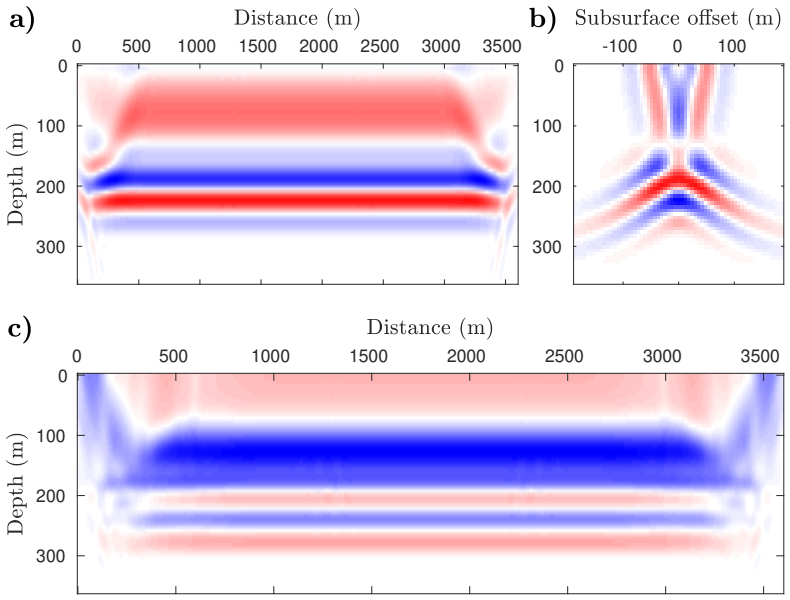

Figure 2.5: The a) reflectivity section, associated b) CIG for $x = 2000$ m and c) tomographic gradient, for an underestimated velocity model and a single interface, from Chauris and Cocher (2017).

With the introduction of the approximate inverse in the inner loop, the reflectivity selection is more focused around the expected depth (Figure 2.6a). This is due to the deconvolution (introduction of the inverse of the seismic wavelet) but also to the compensation for amplitudes. The migration smiles are largely attenuated, especially in the

CIGs (Figure 2.6b). As illustrated in Chauris and Cocher (2017), the tomographic gradient is much more homogeneous above the interface (Figure 2.6c) and the shape of the objective function is more symmetric with respect to the optimal value. The important aspect is that the minimum is now at its correct position.

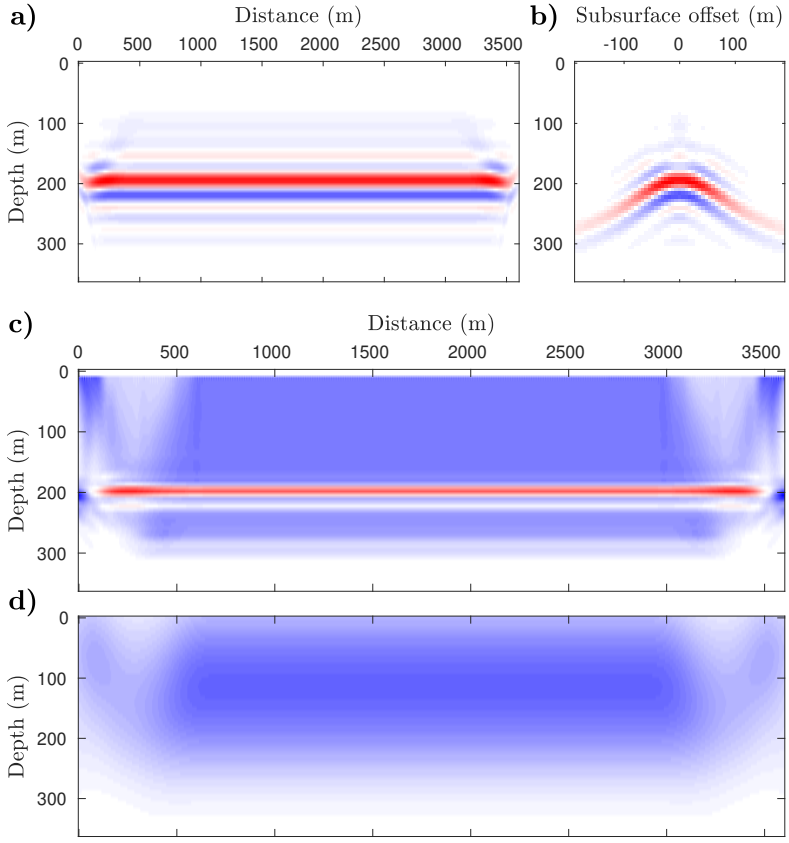

Figure 2.6: Same as for Figure 2.5, but for the inverse case. In addition, the tomographic gradients are displayed for c) $\beta = 0$ and d) $\beta = -3/2$, from Chauris and Cocher (2017).

We conclude from this section that illumination should be properly compensated during the construction of an image (inner-loop). Uneven illumination would otherwise be related to an incorrect macro-model. The iterative LSM scheme is clearly more expensive and leads to fundamental unstable macro-model updates (Liu et al., 2014; Huang, 2016; Cocher et al., 2017). The use of the approximate inverse within the iterative scheme speeds up the convergence but does not remove the instabilities (Hou and Symes, 2016a; Chauris and Cocher, 2018; Farshad and Chauris, 2021a), except in one solution proposed by Cocher et al. (2017). The second option is simply the application of the approximate inverse without any iteration. The third interesting possibility is the matching filter (Guitton, 2017) that requires the additional application of the forward and adjoint operators, which is not the case for the IVA approach. In the future, it would

be interesting to compare the three approaches on the same data set.

2.4.2 Expected convex shape

Still in this section on the “Focusing principle", we now consider the shape of different objective functions. Convexity is an important property to ensure that a gradient-based approach – the only tractable solution – leads to the correct solution. So far, we have mainly concentrated on the Differential Semblance Optimization (DSO) approach to estimate the quality of the focusing. The reason is simple: Stolk and Symes (2003) demonstrated that among all possible objective functions, the differential semblance one is the only one that ensures a smooth behavior associated to a smooth perturbation in the macro-model, at least in a high-frequency perspective. In other words, all stationary points are global minima, at least on some simplified velocity models as demonstrated by Symes (1998a). As explained in Appendix 2.12, the subsurface offset DSO approach is the equivalent of the time-domain slope tomography (Billette and Lambaré, 1998). The residual move-out shape depends on the choice of the extension (Montel and Lambaré, 2019). The total travel time and the two slope components are automatically fitted through the migration part. The correct macro-model is obtained when the two rays start from the same physical position.

On the other side, the DSO approach is sensitive to noise, especially coherent noise, enhanced by a multiplication by h or by the application of the derivative ∂_H. The replacement of the adjoint by the approximate inverse leads to a more robust scheme, but coherence noise may harm the quality of the tomographic update (Chauris et al., 2002b; Mulder and ten Kroode, 2002). The recommendation is to analyze the adjoint source and possibly propose some intermediate processing steps as the filtering of events with unexpected structures in the CIGs. This is indeed case-dependent and more research is needed for a more robust approach.

The DSO approach is thus delicate: on one side, this is the recommended strategy to ensure a convex objective function; on the other side, it is very sensitive to coherent noise. As a consequence, alternatives have been proposed. The most common ones are based on the maximization of the semblance (stack power) (Chavent and Jacewitz, 1995; Zhang and Shan, 2013; Plessix et al., 2000; Soubaras and Gratacos, 2007). As it may suffer from for cycle-skipping issues (Chauris and Noble, 2001), Zhang and Shan (2013) proposed an intermediate partial stack power approach. Symes and Carazzone (1991) explained how to combine the DSO approach with the image power. We may also cite the correlation approach (Van Leeuwen and Mulder, 2010) or more recently the one developed by Tang et al. (2020) and developed in the next section, with arguments on the shape of the macro-model gradient.

2.4.3 Gradient artifacts

Vyas and Tang (2010); Fei and Williamson (2010); Li et al. (2017) illustrated that the DSO tomographic gradient contains strong "artifacts", especially in the case of discontinuous interfaces (Figure 2.7). Instead of an expected smooth macro-model gradient, oscillations are present. They are not really artifacts in the sense that the gradient is properly derived and implemented. Different strategies have been proposed to remove them. Fei and Williamson (2010) proposed to add a $\pi/2$ phase rotation during the imaging condition, but the modified gradient is not a gradient of an objective function anymore (Shen and Symes, 2015). One may note that the reflector still has an imprint (negative red values, Figure 2.6c) on the macro-model update that cannot be removed by smoothing the gradient. Chauris and Cocher (2017) proposed to introduce an additional term inside the definition of the objective function, i.e. a multiplication by m_0^α

$$J^\alpha_{\text{MVA}} = \frac{1}{2}\left\| m_0^\alpha \, h\xi \right\|^2. \tag{2.4}$$

The standard approach, without normalization, reads $\alpha = 0$. For the other cases, the gradient consists of two terms: the classical one multiplied by $m_0^{2\alpha}$ and another one that is a local contribution. Using a 1D model, Chauris and Cocher (2017) proved that $\alpha = -3/2$ is an optimal choice. The same value is also used in 2D leading to attenuate the oscillations on the final macro-model gradients (Figure 2.6d).

For discontinuous reflectivity sections, the same effects are observed. The recommendation is to smooth the final gradient over the expected wavelength (Figure 2.8). At least two alternatives exist. Recently, Tang et al. (2020) proposed to use an Optimal Transport function (OT, Métivier et al., 2016c; Yang and Engquist, 2018), instead of the ℓ_2-norm least-squares approach. This can be easily understood in the Full Waveform Inversion context to move masses and avoid local minima. The IVA objective function is a priori convex and the need for OT is less intuitive. However, Tang et al. (2020) showed that the macro-model gradient does not contain the imprint of the reflectivity. More research is needed to better understand why and to establish a link between this approach and the one promoted by Chauris and Cocher (2017), leading to similar numerical results (Figure 2.6).

Finally, another approach to attenuate the imprint of the reflectors in the macro-model update would be to introduce the tomographic Hessian but this still remains to be demonstrated. One could also think at introducing a picking stage in a Residual moveout-based MVA, leading to smooth macro-model updates (Zhang and Biondi, 2013). Picking is a non linear process, useful to remove unwanted events. Then MVA consists of optimizing the moveout parameters as in Al-Yahya (1989). One may question what happens in complex zone with low illumination and without clear coherency.

We have reviewed that illumination is an important factor during the construction of the image, leading to distorted tomographic gradient. We have identified three most

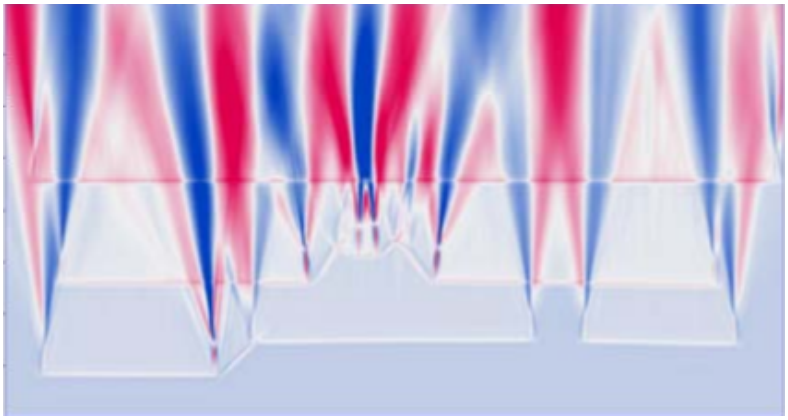

Figure 2.7: "Gradient artifacts", with the standard MVA approach in the case of discontinuous reflectors, from Fei and Williamson (2010). The horizontal and vertical axis refer to the z and x coordinates.

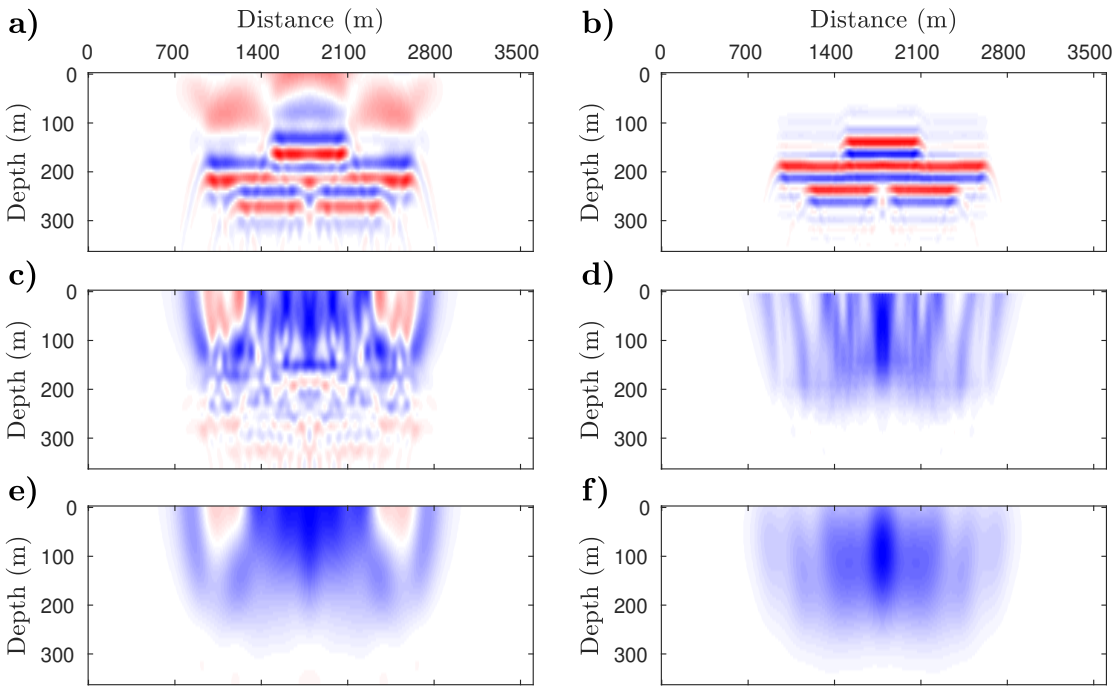

Figure 2.8: Discontinuous multilayer case, with the migration (a, c, and e) and pseudoinverse (b, d, and f) approaches, for reflectivity section (a and b); macro-model update (c and d); and the same update after Gaussian smoothing over the expected wavelength (e and f), from Chauris and Cocher (2017).

promising approaches (Optimal Transport, multiplication by m_0^α and intermediate picking step) to be compared in a future work. We now discuss how MVA could extend beyond its conventional frame.

2.5 Extensions to a more complex physics

The two main aspects addressed in the section are the assumptions about the physics of wave propagation and the Born approximation.

2.5.1 Beyond the constant-density acoustic case

In a large majority of the MVA publications, the underlying assumption is a constant-density acoustic wave equation. The first reason is certainly related to the computational aspects (see the section on "Implementation aspects") and the number of wavefields to be considered for the forward but adjoint terms, both during the construction of the image but also for the gradient derivation. A simple example shown by Farshad and Chauris (2020b) indicates that amplitude variations along the reflector due to density contrasts leads to apparently defocused energy in the CIGs even for the correct macro-model (Figure 2.9). This aspect is essential for the applications on real data.

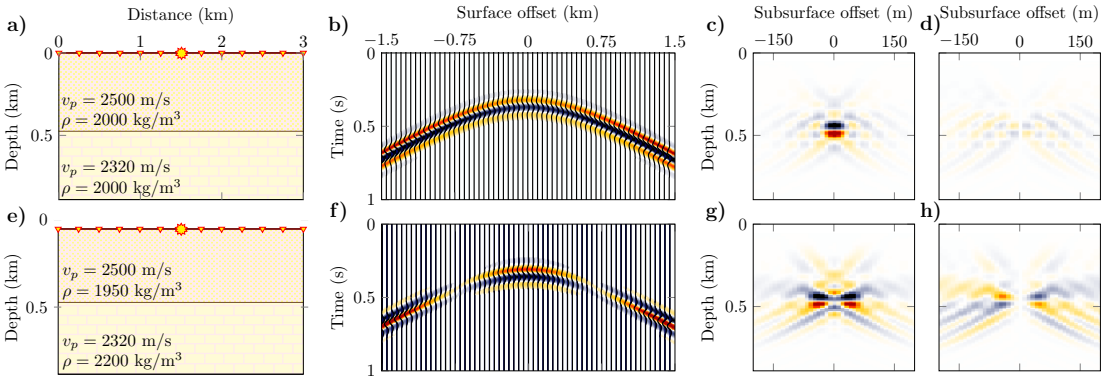

Figure 2.9: A two layered Earth model a) without density variations and the corresponding b) shot gather, inverted c) ξ and d)$(h/h_{max})\xi$. (e,f,g,h) Same as (a,b,c,d) but with density variations in the model, from Farshad and Chauris (2020b).

The first extension is straightforward: for a more complex physics, simply replace the Green's functions by their proper expressions. For example, Feng (2018) proposed an elastic PP-PS MVA scheme with applications on real data. Li et al. (2016) explained how to implement an anisotropic MVA (VTI and TTI cases) with some rock physics constraints as a regularization term. Shen et al. (2018) formulated a combined velocity and attenuation (MVA and MQA) analysis: the quality factor Q indeed has an impact of the

focusing of the migrated images and potentially may be estimated in the image domain. Regularization plays an essential role for a possible decoupling between parameters (Li et al., 2016).

As discussed in the previous section, a proper compensation for illumination is needed for a meaningful macro-model update. The iterative LSM approach being out of scope for computational and instability reasons, the Hessian-related matching filter and the direct inverse remain the two main possibilities. In the first case, the forward and adjoint operators should be once more applied. In the second case, an approximate inverse should be developed beyond the constant-density case.

We discuss here the extension to variable density and elasticity for a fixed macro-model. The first step was derived by Zhang et al. (2014b), showing that the opening angle θ at the image point is needed for the variable-density case. This is related to the radiation patterns. Qin and Lambaré (2016) proposed to obtain the angle through ray tracing. More recently, Dafni and Symes (2018) introduced a two-step process. In the first phase, $\xi(\mathbf{x}, \mathbf{h})$ is computed as for the constant-density case. Then, a Radon transform is applied to obtain angle-domain CIG $\mathcal{R}(\mathbf{x}, \theta)$. By selecting two different angles, the physical perturbations (i.e. without extension) are retrieved by projecting the Radon transform into $\delta I_p(\mathbf{x})$ and $\delta\rho(\mathbf{x})$, the perturbed impedance and density sections. For a more robust approach, Farshad and Chauris (2020a) proposed to consider all traces within the angle-domain CIGs. This approach, based on a weighted least-squares objective function, is a generalization of the work of Zhang et al. (2014b) and Dafni and Symes (2018). For each spatial position $\mathbf{x}$, the reconstruction of the physical parameters is obtained by solving a 2×2 linear system in 2D. Later, Farshad and Chauris (2021c) improved the quality of the inversion by incorporating sparsity-promoting ℓ_1-norm regularization using a shrinkage-thresholding algorithm. In all cases, a proper inverse means that the modeling from the physical parameters should lead to the original data.

The main interest of the method proposed by Farshad and Chauris (2020a) is that the imaging process is split into three parts, summarized in Algorithm 2. Only the first part ("approximate inverse ") contains wave-equation terms. It is thus the most expensive part by far. The Radon transform is only applied once. The last part ("Physical decomposition") is a bivariate curve fitting problem in which additional constraints can be introduced, typically sparsity constraints (Farshad and Chauris, 2021c). This phase is extremely fast and different iterative optimization can be tested. The sparsity-promoted approach with ℓ_1-norm constraint on each reconstructed physical parameter is for example essential in the case of sparse acquisition (Figure 2.10). In both cases, the data fit is correct (Figure 2.11), but only when constraints are added, the model parameters are well recovered (Figure 2.10).

The same strategy (application of approximate inverse Born operator, Radon transform and physical decomposition) holds for the extension to elasticity (Farshad and Chauris, 2021b). In this case, by enlarging the applicability from acoustic (two parameters) to elastic medium (three parameters), the final step "Physical decomposition"

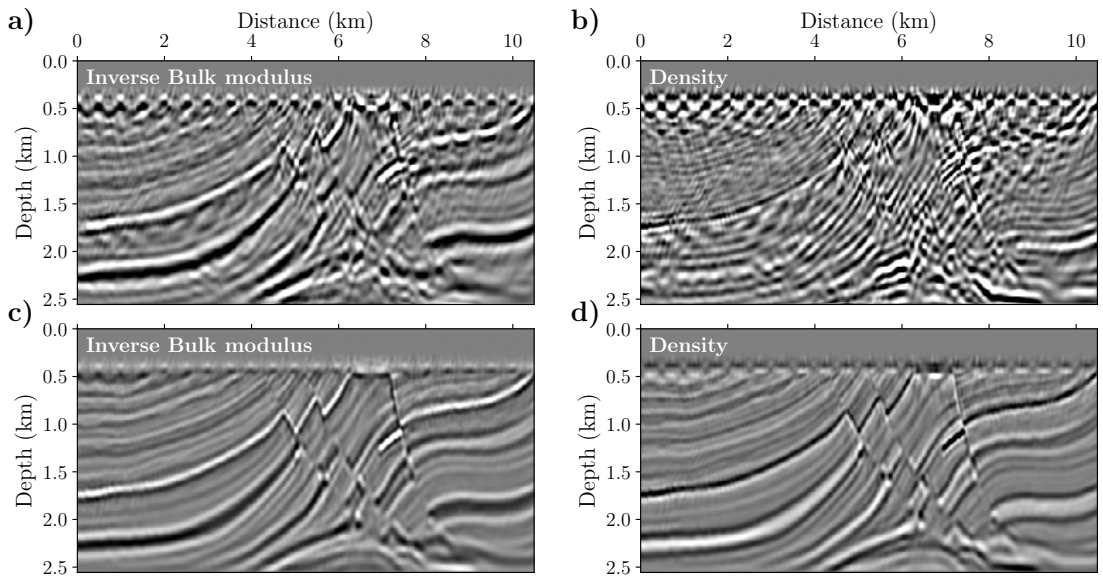

Figure 2.10: Comparison between the inverted parameters corresponding to sparse shots (sourse spacing is 432 m) for Marmousi2 model. Inverted a) inverse Bulk modulus and b) density parameters without regularization and inverted c) inverse Bulk modulus and d) density parameters with ℓ_1-norm regularization, from Farshad and Chauris (2021c).

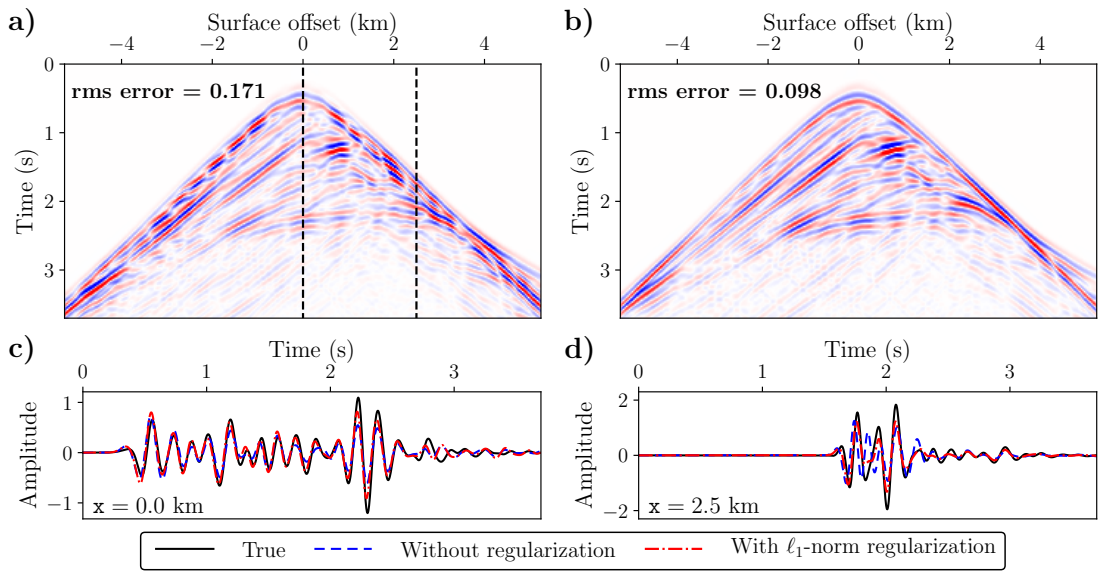

Figure 2.11: The reconstructed shot via the inverted physical parameters obtained a) without regularization and b) with ℓ_1-norm regularization. The dashed lines in (a) correspond to the extracted traces at the surface offsets c) 0.0 km and d) 2.5 km, from Farshad and Chauris (2021c).

Algorithm 2: Weighted Least-Squares approach for variable-density pseudoinverse Born inversion.

1 **Pseudoinverse: for** *each* $\mathbf{s}$ *position* **do**
2 construct $\xi_{\text{inv}}(\mathbf{x}, \mathbf{h})$
3 **Radon transform: for** *each surface position* **do**
4 construct $\mathcal{R}(\mathbf{x}, \theta)$
5 **Physical decomposition: for** *each* $\mathbf{x}$ *position* **do**
6 construct $\delta I_p(\mathbf{x}), \delta\rho(\mathbf{x})$ (+ sparsity constraints)

(Algorithm 2) becomes a trivariate curve-fitting problem. As for the variable-density case, the most computationally expensive step is the elastic approximate inverse Born operator constructing $\xi(\mathbf{x}, \mathbf{h})$. The conclusions derived for elastic inversion are consistent with the ones known for FWI: impedances are better reconstructed than the density (Köhn et al., 2012), and larger surface offsets provides better resolution (Farshad and Chauris, 2021b). One still needs to couple this recent approximate inverse within the tomographic loop. The derivation of the macro-model gradient should incorporate the new elements of the approximate inverse.

2.5.2 Beyond the single scattering approximation

The second aspect for a more complex physics is related to the tentatives to go beyond the Born approximation (single scattering). There are at least three possibilities: (1) introduction of transmitted waves, (2) introduction of multiples and (3) towards a fully non-linear modeling.

2.5.2.1 MVA with transmitted waves

In parallel and independently, Shen (2013) and Lameloise et al. (2013) proposed a way to incorporate transmitted waves within the MVA context. Let remind that MVA approaches are based on scale separation between the macro-model and the model perturbation associated to reflected waves. Instead of back-projecting the reflected wavefield, the full wavefield is used. The image is then formed as the cross-correlation between the source and the modified receiver wavefields. The new reflectivity image differs from the classical one associated with reflectors for which events are localized at a particular depth. Here, transmitted waves are mapped in the image domain from the surface to the maximum penetration depths (Shen, 2013; Lameloise et al., 2013; Lameloise and Chauris, 2016; Zhou, 2020). A subsurface offset h is introduced during the imaging condition. Within CIGs, it appears that the energy is sensitive to the quality of the

macro-model used for migration (Figure 2.12). This offers, as for the classical way, to update the macro-model, with the same J_{MVA} objective function.

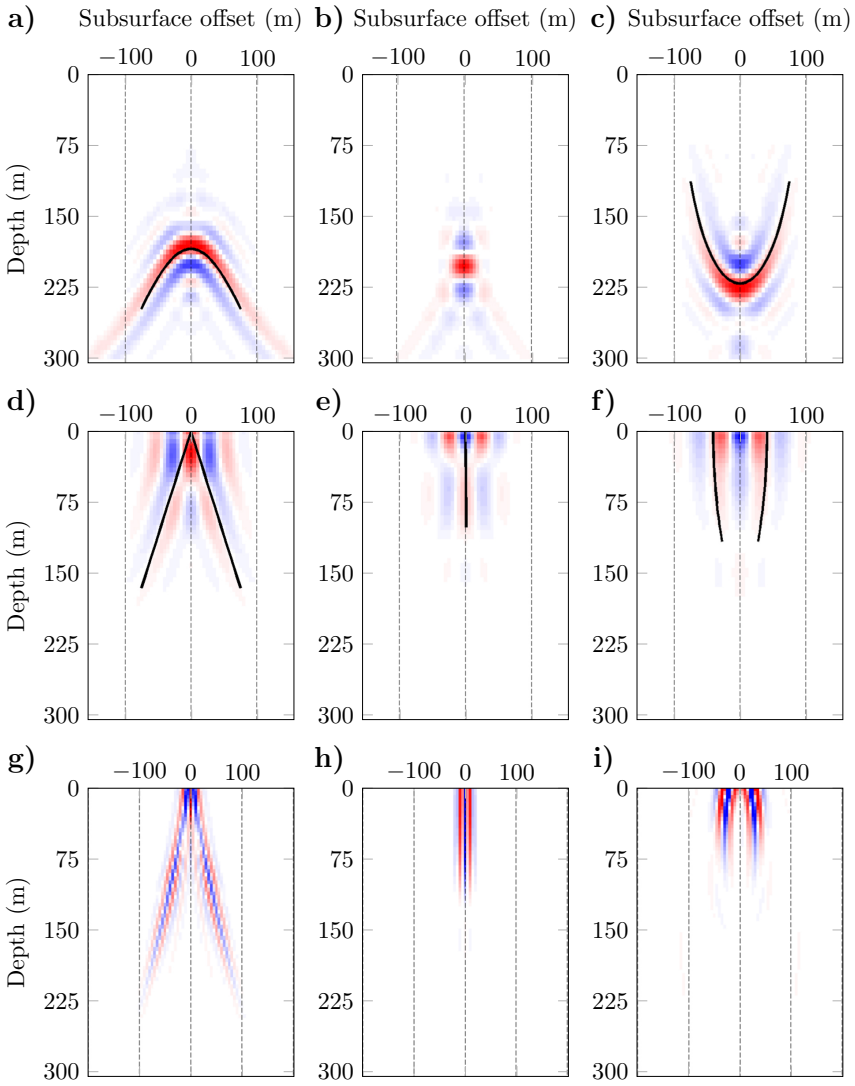

Figure 2.12: Extension beyond reflected waves: CIG for reflected waves (a, b, and c), body waves traveling along the surface (d, e, and f) and transmitted waves (g, h, and i), for an under (a, d, and g), exact (b, e, and h) and over-estimated (c, f, and i) macro-model, from Lameloise and Chauris (2016).

Such an approach was demonstrated on surface data (Shen, 2013) and cross-well data configuration (Figure 2.13, Lameloise and Chauris, 2016). The total field is dominated by transmitted waves. In the initial model, most of the energy in the CIGs is shifted towards positive h, meaning that the velocity values are under-estimated. After

optimization, energy focused around $h = 0$.

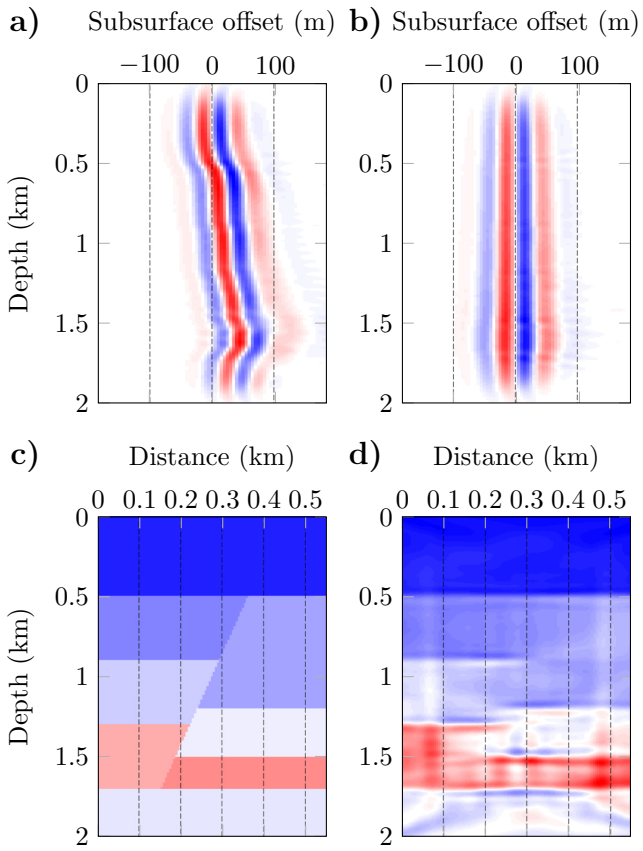

Figure 2.13: MVA cross-well configuration with the full wavefield. a) CIG in the initial homogeneous model, b) CIG in the final model, c) exact velocity model and d) inverted model. In the final model, energy is localized around $h = 0$, from Lameloise and Chauris (2016).

The recommendation is to first consider the full wavefield or the transmitted field and to later introduce the reflected part. There is a priori no need for splitting the data into transmitted and reflections. However, for a more stable reflection MVA approach, one would rely on the IVA approach. Such weights would be small for the transmission part, as they contain a term in $\cos\theta$, with θ close to $\pi/2$ (Zhou, 2020).

2.5.2.2 In the presence of multiples

Multiples are considered as coherent noise for the traditional MVA approaches as they are not modeled under the first-order Born approximation. At least a few approaches have been proposed in the MVA context to incorporate them (Mulder and van Leeuwen,

2008; Staal and Verschuur, 2012; Cocher et al., 2018; Soubaras and Gratacos, 2018). The main element is to consider a modeling tool able to generate them. Then, the reflectivity section is defined as the one minimizing the inner-loop objective function. The standard J_{MVA} is used, once the imprint of the multiples is removed on the reflectivity section.

Multiples may also offer a more robust MVA scheme as the illumination provided by surface multiples is different from the one by the reflections. For example, multiples help for subsalt imaging (Liu et al., 2011). For 3D marine acquisition, the coarse sampling in the cross-line direction can be compensated by multiples (Long et al., 2013). However, this has not been really yet demonstrated in the MVA context. The reason behind this is that the derivation of the reflectivity section to remove the imprint of the multiples is obtained through an iterative process, intending to fit the data (Brown and Guitton, 2005; Verschuur and Berkhout, 2011; Wong et al., 2014; Zhang et al., 2014a; Tu and Herrmann, 2015). As for the primary case, events with little energy are created at large subsurface offsets for an optimal fit. As a consequence, such events are enhanced in the MVA adjoint source due to the multiplication by h^2, but they do not necessarily correspond to real events. This leads to local minima and instabilities in the velocity updates (Cocher et al., 2018). For the future, an interesting research possibility is the Marchenko approach encountering for all internal multiple reflections (Wapenaar et al., 2014; Singh et al., 2015). The Marchenko imaging condition does not generate cross-talk while being sensitive to the macro-model needed for the estimation of the direct arrival. One may ask if the fundamental instabilities highlighted by Huang (2016) will also be present in the iterative Marchenko approach.

2.5.2.3 Fully non-linear modeling

Finally we discuss a non-linear data fitting method to construct the image before evaluating the quality of the focusing (Symes, 2008b). MVA can be seen as a waveform inversion approach: at the convergence, there is a good fit to the data (fitting approach) in a physical model if the optimal focusing is obtained (no need for extension). The standard MVA combines a linear process (migration part) with a non-linear update (tomographic part). First ideas have been proposed by Symes (2008b) for a fully non-linear imaging approach. Let determine an extended model m and data d

$$J[m] = \frac{1}{2}\left\|A_{\mathrm{n}}m\right\|^2, \tag{2.5}$$

$$S(d) = \left\{m : \left\|F(m) - d\right\|^2 \leq \left\|d\right\|^2\right\}, \tag{2.6}$$

where $S(d)$ are the admissible possibly extended models m for which there is a good data fit. The minimization of the first objective function ensures that m is a physical

model. It is currently only tractable in layered models and there is a need for parameterizing the feasible models, probably along the same line as the one proposed by Chavent et al. (1994); Plessix et al. (1995); Clément et al. (2001) for the Migration-Based Travel Time Tomography approach.

2.6 Implementation aspects

The last section on the limitations of MVA is devoted to the numerical aspects. We first analyze the number of wavefields to be computed for the tomographic gradient. We then discuss different strategies to reduce the computational cost, from a target-oriented MVA to sparse CIGs and finally to a frequency / subsurface-offset continuation.

Number of wavefields

The FWI gradient is obtained as the correlation between the source forward wavefield with the receiver residual wavefield (Virieux and Operto, 2009; Fichtner, 2010). For the MVA approach, the first step consists of computing the reflectivity. This is the same process as the first iteration in FWI (Lailly, 1983; Tarantola, 1984). Here is a summary of the algorithm, with details in Appendix 2.10.

Algorithm 3: MVA gradient derivation – approach 1

1 Set $\xi = 0$
2 **Reflectivity section: for** *each shot position* **do**
3 compute S_0 and R_0
4 add the $S_0 \otimes R_0$ to ξ
5 Set the gradient to 0
6 **MVA gradient: for** *each shot position* **do**
7 computed μ_s and μ_r
8 add the $\mu_s \otimes S_0 + \mu_r \otimes R_0$ to the gradient

The first approach (Algorithm 3) is a direct translation of the adjoint state gradient derivation. For the reflectivity section, 2 wavefields S_0 and R_0 are needed. They are used in combination with ξ as a new source term to obtain μ_r and μ_s, respectively. The tomographic update has two contributions from the correlations $\mu_s \otimes S_0$ and $\mu_r \otimes R_0$. The practical implementation is slightly different (Algorithm 4): in the second phase, the quantities S_0 and R_0 are not available, even if they were computed during the first phase, except if there are stored on disk. It means that they should be once more computed (line 7 in Algorithm 4). The reason is that ξ should have the contribution of all shots before starting the MVA gradient evaluation. As a summary, FWI requires 2 wavefields (S_0 and

Algorithm 4: MVA gradient derivation – approach 2

```
Set ξ = 0
Reflectivity section: for each shot position do
    compute S_0 and R_0
    add the S_0 ⊗ R_0 to ξ
Set the gradient to 0
MVA gradient: for each shot position do
    compute S_0 and R_0
    compute μ_s and μ_r
    add the μ_s ⊗ S_0 + μ_r ⊗ R_0 to the gradient
```

R_0), while MVA requires 4 (S_0, R_0, μ_s and μ_r), knowing that S_0 and R_0 are computed twice: the MVA cost is thus three times more expensive than the FWI cost. We do not even consider the cost for the extended image condition.

Among the different survey-sinking extensions to handle the triplicated wavefields, the common angle gathers are more expensive than the subsurface offset case. The subsurface extension needs to correlate the forward wavefield with the backward wavefield at different spatial positions. For efficiency reasons, these space-lags are usually a integer multiple of the original spatial grid. Subsurface offsets appears to have an easier implementation than the common angle case, either computed as a slant stack of the subsurface-offset gathers (Prucha et al., 1999; Rickett and Sava, 2002; Sava and Fomel, 2003) or obtained directly after decomposing the wavefield into different directions (Xu et al., 2011; Dickens and Winbow, 2011; Zhang, 2014; Wang et al., 2016b). Still, the subsurface offset case remains very expensive in practice.

A number of strategies have been defined to reduce the total computational and memory costs (Duveneck, 2013; van Leeuwen et al., 2015; Yang and Sava, 2015). We split the advanced computational strategies into three main classes: limiting the number of input shots at each iteration, developing a target-oriented MVA including sparse CIGs, and finally exploring a frequency / subsurface offset continuation approach.

2.6.1 Reducing the number of shots

2.6.1.1 Plane wave CIGs

A first possibility to reduce the total computational cost is to rely on the plane wave CIGs instead of the more conventional shot-based approach (Duquet and Lailly, 2006; Liu et al., 2006). This is interesting when the number of required plane waves is smaller than the number of shots: there is a balance between efficiency and spatial aliasing in CIGs when the number of plane waves is insufficient (Guo and Schuster, 2017; Feng,

2018).

2.6.1.2 Random shot selection

Another strategy widely used in the FWI context is to randomly select shot gathers during the iterative process (Romero et al., 2000; Krebs et al., 2009; Ben-Hadj-Ali et al., 2011; Schuster et al., 2011; Castellanos et al., 2014). The same approach does not directly apply in the MVA context for the following reason (Li et al., 2016): once an extended image has been obtained, it is possible to select some random shots, relying on non-linear tomographic iterations to compensate for the decimated tomographic gradient. However, as pointed out by Li et al. (2016), the same approach is not applicable within the inner loop, i.e. during the construction of the extended image. In that case, the cross-talks in the images lead to unusable tomographic gradient. As a consequence, the shot decimation could only divide the total computational cost by a factor two (large decimation on the tomographic gradient but dense shots for the construction of the extended image, Algorithm 3). For Algorithm 4, the largest factor would be three (large decimation for the second phase where S_0, R_0, μ_s and μ_r would be computed at sparse shot locations). For more efficient schemes with a decimation also during the image construction, more research is needed, for example to design a proper processing task such that the incomplete ξ section still leads to an interesting tomographic gradient update.

2.6.1.3 Exploding reflector technique

Along the same line, Guerra and Biondi (2011) proposed to use the prestack exploding reflector technique to limit the computational cost. However, cross-talks are generated from different reflectors during the modeling phase, imposing in practice non realistic sparsity assumptions.

The first strategies have been developed to reduce the size of the input data, in practice the number of shots. Now, we consider alternatives to limit the image size, while delivering a stable tomographic update.

2.6.2 Target-oriented MVA

The two main strategies consist of reducing the zone of interest and of constructing focusing panels at sparse locations.

2.6.2.1 Targeted area

Tang and Biondi (2011, 2013); Zhang and Biondi (2013) have explored a target-oriented MVA version. The main important assumption is that the macro-model is known outside of the target zone. Then, new seismic data are computed at the border of that zone with a

redatuming technique. This is performed only once. The standard MVA technique then applies, with the advantage that the propagation time is largely reduced for the different wavefields. The imaging condition is also limited to the target zone.

2.6.2.2 From dense CIGs to CIPs

Alternatively, CIGs can be computed at some sparse locations (Yang and Sava, 2015). The reason is that the tomographic MVA part aims at determining the macro-model, usually a smooth model. A rule of thumb is that the distance between two successive CIGs should be twice smaller that the typical horizontal wavelength characterizing the macro-model (Li and Chauris, 2018).

An even more advanced strategy is to consider Common Image Points (CIPs) instead of Common Image Gathers (Sava and Vasconcelos, 2009; Diaz and Sava, 2017). The extended image is only computed at some sparse locations in depth determined from the image characteristics (Cullison and Sava, 2011; Yang and Sava, 2015). As the total computational cost and memory requirements are significantly reduced, especially in 3D, this offers the possibility to consider multi-lag extensions, combining subsurface and time-lag aspects.

Both approaches (sparse CIGs and CIPs) largely reduce the memory needed to store the focusing panels / points and the computational cost associated with the imaging condition, but the different wavefields still need to be computed over the entire zone of interest, even if only sparse locations are later selected within that zone.

Frequency / subsurface offset continuation

Finally, an interesting multi-scale approach was proposed by Fu and Symes (2017). During the first tomographic iterations, the resolution does not need to be very fine and low-frequency data may be used, with a coarse sampling of the spatial grid and along the extension axis, while considering a large maximum 2D extension $e_{\max}$ for a possible incorrect macro-model. Progressively, the frequency content is increased, with a wave propagation on a finer grid, while $e_{\max}$ is reduced as images become more focused. The criterion is to preserve the data match in the time domain at each iteration. Fu and Symes (2017) successfully evaluated this approach in the case of the subsurface offset extension.

2.7 Discussion

We have discussed the important practical aspect that focusing is not necessarily obtained for the correct macro-model. This is essentially the case when the illumination of the subsurface is not properly compensated during the imaging part (inner-loop), before

evaluating the tomographic gradient. The recent introduction of the approximate inverse has two advantages over the iterative least-squares migration: a faster approach and a more stable tomographic update. More research is still needed to determine a good compromise between the convexity of the objective function and its robustness. We discuss here three main aspects to be further developed in the future, in particular the resolution provided by the MVA approaches, the tomographic Hessian and the extension to multi-parameter imaging. We suggest a number of possibilities to be developed, mainly as intermediate processing tasks.

2.7.1 Resolution

The analysis of resolution is well known in the FWI context (Virieux and Operto, 2009; Alkhalifah and Plessix, 2014). There is a need in the future to perform a similar analysis for MVA approaches. In the Inversion Velocity Analysis context, the inverted model perturbation ξ has meaningful amplitudes and is a model perturbation δm to be added to the reference model as $m = m_0 + \delta m$ (Li and Chauris, 2018; Zhou et al., 2020b). We retrieve the famous "inversion" as a combination of "migration" and "tomography" by Mora (1989). We give here an illustration on the Marmousi model. The inverted model after IVA iterations (Figure 2.14a) can be summed with the associated model perturbation (Figure 2.14b), leading to the final model (Figure 2.14c), in which the events in CIGs are focused (Figure 2.14d). Probably higher resolution could be obtained with a focus on diffractions as in Sava et al. (2005) or with the multiples as in FWI: more research could couple the Marchenko imaging scheme (Wapenaar et al., 2014) to Velocity Analysis as in Diaz et al. (2016).

The role of frequencies has never been studied as far as we know. More precisely, high enough frequencies are required to focus events, but the quality of the low-frequency part does not seem to be a determinant aspect. More research is needed to determine the optimal balance between the central processed data frequency and the maximal distance between CIGs or CIPs. Along the same line, specific processing steps should be developed to reduce the number of shots used to construct the image (inner-loop), while preserving the quality of the macro-model update (outer loop). Anti-aliasing filters should be derived in the extended domain (Abma et al., 1999).

With the use of wave-equation-based operators, there is no guarantee that the macro-model remains smooth during the tomographic update. This is however the implicit assumption during the construction of the MVA two-loop process. More research should be conducted to better determine the impact of non-smooth reference model in the analysis of the focusing.

Finally, for the robustness of the scheme with respect to coherent noise, two strategies could be combined: in both cases, the objective would be to attenuate the impact of the coherent noise, either in the data domain before migration or in the depth domain after migration but before the tomographic update.

2.7.2 Tomographic Hessian

The use of Newton approaches has been applied in the context of FWI (Pratt et al., 1998; Plessix and Mulder, 2004; Valenciano et al., 2006; Almomin and Biondi, 2013; Tang, 2009) and to a much less extent to MVA (Soubaras and Gratacos, 2017). The introduction of the second-order derivatives of the objective function can be obtained indirectly in a quasi-Newton approach (Byrd et al., 1995) or explicitly through, for example, the Truncated Newton approach (Métivier et al., 2013).

As indicated before, the inner loop objective function has a Hessian close to the identity in the case the approximate inverse is used. But it does not mean that the outer loop behaves the same way. Formally, it is possible to derive a second-order adjoint system for the tomographic part. In practice, the number of additional wavefields to be computed would become eight: with the same notations as for Algorithms 3 and 4, two wavefields S_0 and R_0 are needed for the reflectivity, two more for the tomographic gradient (μ_s and μ_r), and four more for the tomographic Hessian (λ_S, λ_R, λ_{μ_s} and λ_{μ_r}), as adjoint variables for the definition of S_0, R_0, μ_s and μ_r. Eight wavefields in total are the correct number if they can be stored on disk. Otherwise and because of the intermediate summations over sources, the total number would be $2 + 4 + 8 = 14$, respectively (S_0, R_0), then (S_0, R_0, μ_s, μ_r) and finally $(S_0, R_0, \mu_s, \mu_r, \lambda_s, \lambda_r, \lambda_{\mu_s}\lambda_{\mu_r})$. In the FWI case, the equivalent number would be $1 + 2 + 4 = 7$ (one for the objective function, two for the gradient, and four for the Hessian), but it reduces to four only because it is possible to compute the contribution of a given shot to the gradient and to the Hessian independently of other shots: this is not the case for MVA.

2.7.3 Multi-parameter imaging

Multi-parameter imaging beyond the constant-density acoustic case is an essential aspect for real data applications. This has been largely developed in the FWI case (Prieux et al., 2013; Virieux and Operto, 2009). In the MVA context, we indicated in the recent advances section the extension of the direct inverse to variable density and elasticity. But there is still a need to incorporate these schemes within the tomographic loop (Weibull and Arntsen, 2014; Wang et al., 2019).

The possible cross-talk between parameters also leads to the estimation of the uncertainties, possibly within a Bayesian approach (Jin and Madariaga, 1993, 1994; Docherty et al., 1997; Jin and Beydoun, 2000). Another possible extension is the time-lapse MVA context, in the case where the macro-model is changing over time (Shragge and Lumley, 2013).

2.8 Conclusions

Migration Velocity Analysis aims at determining the optimal subsurface model parameters by splitting the model into two parts: the extended image and the macro-model. During iterations, the initially large extension for an easy data fit, is progressively reduced to the physical domain at the convergence for the optimal macro-model. The theoretical recommended approach (subsurface-offset extension to handle triplicated wavefields, coupled to the Differential Semblance Optimization for a convex objective function) is also the one that is very sensitive to coherence noise. We have analyzed three main aspects (assumed optimal focusing for the correct macro-model, simple physics assumptions and implementation issues), currently strongly limiting the applications at an industrial scale. We have also reviewed some remedies (introduction of the approximate inverse as the reference image, more robust ways of evaluating the quality of the focusing at some sparse locations). As far as we know, no existing code with published results contains all these aspects. More research is needed for more suited preprocessing steps applied to the focusing panels before evaluating its quality to deliver more robust schemes with respect to limited acquisitions, coherent noise and shot decimation. In a longer term, incorporating the full wavefield within the analysis is essential.

2.9 Acknowledgements

The authors are grateful to many people for fruitful interactions and discussions around Migration Velocity Analysis, in particular to T. Alkhalifah, F. Audebert, R. Brossier, E. Cocher, D. Donno, C.-A. Lameloise, G. Lambaré, H. Jiang, Y. Li, C. A. Martins de Assis, L. Métivier, W. Mulder, M. Noble, S. Operto, R.-É. Plessix, P. Podvin, B. Salomons, B. Sun, W. Symes, F. ten Kroode, J. Virieux, P. Williamson and T. Zhou among others. A very special thank to W. Symes for inspiring new ideas.

2.10 Appendix I: Gradient derivation

The objective of the first appendix is to explain how to compute the gradient of the MVA objective function with respect to the macro-model m_0. We rely on the adjoint state technique (Plessix, 2006a). This consists of enlarging the original objective function with the constraints indicating how the reflectivity is constructed. We refer to Weibull and Arntsen (2013); Chauris and Cocher (2017); Li and Chauris (2018) for more details. For simplicity here, we consider the case $\alpha = 0$ (equation 2.4). Finally, we also give some elements for the IVA case when the optimal reflectivity section is obtained iteratively.

With the standard adjoint and inverse schemes

Let J_{MVA} be the extended objective function defined as

$$\begin{aligned} J_{\mathrm{MVA}} &= \frac{1}{2}\left\| A_{\mathrm{n}}\xi \right\|^2 \\ &- < \lambda, \xi - D_i \int \mathrm{d}\mathbf{s}\mathrm{d}\omega S_0^*(\mathbf{s},\mathbf{x}-\mathbf{h},\omega) R_0(\mathbf{s},\mathbf{x}+\mathbf{h},\omega) >_{\mathbf{x},\mathbf{h}} \\ &- < \mu_s(\mathbf{s},\mathbf{x},\omega), \mathcal{L}_0 S_0 - S_s(\mathbf{s},\mathbf{x},\omega) >_{\mathbf{s},\mathbf{x},\omega} \\ &- < \mu_r(\mathbf{s},\mathbf{x},\omega), \mathcal{L}_0^* R_0 - R_s(\mathbf{s},\mathbf{x},\omega) >_{\mathbf{s},\mathbf{x},\omega}, \end{aligned} \tag{2.7}$$

with

$$S_s(\mathbf{s},\mathbf{x},\omega) = K_i(\omega)\Omega^\dagger(\omega) D_s G_0(\mathbf{s},\mathbf{x},\omega) \tag{2.8}$$

$$R_s(\mathbf{s},\mathbf{x},\omega) = K_i(\omega)\Omega^\dagger(\omega) \int \mathrm{d}\mathbf{r} D_d d(\mathbf{s},\mathbf{r},\omega) D_r G_0(\mathbf{r},\mathbf{x},\omega) \tag{2.9}$$

or equivalently

$$\mathcal{L}_0 S_s(\mathbf{s},\mathbf{x},\omega) = K_i(\omega)\Omega^\dagger(\omega) D_s \delta(\mathbf{x}-\mathbf{s}) \tag{2.10}$$

$$\mathcal{L}_0^* R_s(\mathbf{s},\mathbf{x},\omega) = K_i(\omega)\Omega^\dagger(\omega) \int \mathrm{d}\mathbf{r} D_d d(\mathbf{s},\mathbf{r},\omega) D_r \delta(\mathbf{x}-\mathbf{r}) \tag{2.11}$$

The $\mathcal{L}_0$ and δ are the wave propagation operator and Dirac distribution, respectively. The $*$ denotes the adjoint and ω the angular frequency. The details of the different quantities (K_i, D_s, D_r, D_d) are provided in Appendix 2.11. They are essentially time and spatial derivatives to obtain the standard image and the approximate inverse image. In the formulation, the adjoint state variables λ, μ_s and μ_s are defined such that the state variables ξ, S_0 and R_0 satisfy the state equations. The macro-model m_0 (squared of the slowness) is contained in the $\mathcal{L}_0$ operator and for a constant-density acoustic wave-equation, we have $\partial\mathcal{L}_0/\partial m_0 = (i\omega)^2$. The final gradient thus reads

$$\frac{\partial J_{\mathrm{MVA}}}{\partial m_0} = - < \mu_s, (i\omega)^2 S_0 >_{\mathbf{s},\omega} - < \mu_r, (i\omega)^2 R_0) >_{\mathbf{s},\omega} \tag{2.12}$$

The gradient depends on the adjoint variables μ_s and μ_r. The derivative of the objective function with respect to ξ leads to the adjoint source

$$\lambda = A_{\mathrm{n}}^t A_{\mathrm{n}} \xi. \tag{2.13}$$

The derivative with respect to S_0 and R_0 yields

$$\mathcal{L}_0^* \mu_s(\mathbf{s},\mathbf{x},\omega) = \int \mathrm{d}\mathbf{x}\mathbf{h} D_i \lambda(\mathbf{x}+\mathbf{h},\mathbf{h}) R_0^*(\mathbf{s},\mathbf{x}+\mathbf{2h},\omega) \tag{2.14}$$

$$\mathcal{L}_0 \mu_r(\mathbf{s},\mathbf{x},\omega) = \int \mathrm{d}\mathbf{x}\mathbf{h} D_i \lambda(\mathbf{x}-\mathbf{h},\mathbf{h}) R_0^*(\mathbf{s},\mathbf{x}-\mathbf{2h},\omega) \tag{2.15}$$

With an iterative scheme

When the migrated section is obtained through an iterative process before evaluating the quality of the focusing, two approaches exist for deriving the tomographic gradient. In the first case, the adjoint method is applied iteratively to all the intermediate steps. In the second case, one supposes the data fit convergence. This is the recommended approach as it simplifies the derivation of the adjoint. We refer to Cocher (2017) for more details and a comparison between the two strategies. It is also possible to introduce the approximate inverse as a preconditioner to accelerate the least-squares migration (Hou and Symes, 2016b).

2.11 Appendix II: Approximate inverse

We provide the detailed expressions for different approximate inverses. We use the same notations to compare the four final expressions, both for the modeling, adjoint and inverse parts. In other words, we rewrite some original equations provided by the different authors. For the approximate inverse, we consider here the case where the macro-model is locally invariant for simpler expressions. The modeling reads

$$\begin{aligned} F[\xi](\mathbf{s},\mathbf{r},\omega) &= K_m(\omega)\Omega(\omega) \int \mathrm{d}\mathbf{x}\mathrm{d}\mathbf{h} G_0(\mathbf{s},\mathbf{x}-\mathbf{h},\omega) \\ &\quad D_m \xi(\mathbf{x},\mathbf{h}) G_0(\mathbf{s},\mathbf{x}+\mathbf{h},\omega), \qquad (2.16) \\ F^T[d](\mathbf{x},\mathbf{h}) &= D_m^T(\mathbf{x}) \int \mathrm{d}\mathbf{s}\mathrm{d}\mathbf{r}\mathrm{d}\omega K_m^T(\omega)\Omega^*(\omega) \\ &\quad G_0^*(\mathbf{s},\mathbf{x}-\mathbf{h},\omega) d(\mathbf{s},\mathbf{r},\omega) G_0^*(\mathbf{s},\mathbf{x}+\mathbf{h},\omega), \qquad (2.17) \\ F^\dagger[d](\mathbf{x},\mathbf{h}) &= D_i(\mathbf{x}) \int \mathrm{d}\mathbf{s}\mathrm{d}\mathbf{r}\mathrm{d}\omega K_i(\omega)\Omega^\dagger(\omega) \\ &\quad D_s G_0^*(\mathbf{s},\mathbf{x}-\mathbf{h},\omega) D_d d(\mathbf{s},\mathbf{r},\omega) D_r G_0^*(\mathbf{s},\mathbf{x}+\mathbf{h},\omega), \qquad (2.18) \end{aligned}$$

where $*$ is the complex conjugate and $\dagger$ the inverse. For the modeling part, $K_m = -(i\omega)^2$ and $D_m = 1$. The other quantities are defined in Table 2.1. The inverted source is obtained as $\Omega^\dagger = \Omega^*/\|\Omega\|^2$. Note that both the adjoint (equation 2.17) and the inverse

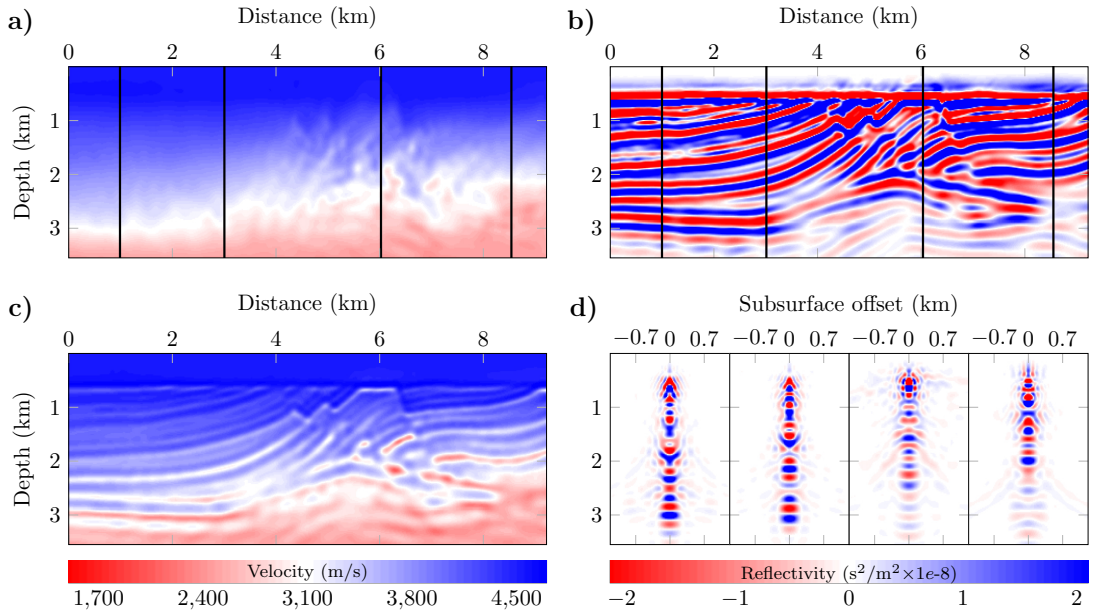

Figure 2.14: Final macro-model in m/s (a), model perturbation in s^2/m^2 (b) and total model in m/s (c), as well as the CIGs (d) for the Marmousi case, from Zhou et al. (2020b)

Table 2.1: Review of the approximate inverse operator expressions for different approaches (Chauris and Cocher, 2018).

Strategy	**A**	**B**	**C**	**D**
Ref.	ten Kroode (2012)	Hou and Symes (2015)	Hou and Symes (2017)	Chauris and Cocher (2017)
Modeling	Kirchhoff	Born	Born	Born
Dim.	3D	2D	2D	2D
$\mathcal{D}_i$	$-32/\pi/v_0^2$	$8v_0^2\sqrt{\nabla_{x,z}^2\nabla_{h,z}^2}$	$-16v_0^2\partial_z$	$-16v_0^2\partial_z$
$\mathcal{K}_i$	$(i\omega)$	$1/(i\omega)^2$	$1/(i\omega)$	$1/(i\omega)$
$\mathcal{D}_s$	∂_{s_z}	1	1	∂_{s_z}
$\mathcal{D}_d$	1	$\partial_{s_z}\partial_{r_z}$	$\partial_{s_z}\partial_{r_z}$	1
$\mathcal{D}_r$	∂_{r_z}	1	1	∂_{r_z}

operators (equation 2.18) have very similar expressions. This is an essential aspect for the implementation.

In the A–strategy, the modeling is different as initially based on the Kirchhoff integral (ten Kroode, 2012) instead of the Born modeling (strategies B to D). To remove the dependency over the scattering angle at the image point arising in the Kirchhoff representation, the author considers the linear relationship between ξ and the data. Only this approach is 3D (all others are 2D). In the D–strategy, the model perturbation was originally expressed in terms of squared slowness perturbation. Here, the expressions are converted to velocity perturbation for consistency. In approaches A, C, and D, the product of all columns (Table 2.1) are the same, except that in A, with an additional factor π coming from the 3D formulation. Note that the expression in D_i for the B–approach can be evaluated in the 2D Fourier domain by applying successive multiplications by $\sqrt{k_z^2 + k_x^2}$ and $\sqrt{k_z^2 + k_h^2}$ in the proper domains. As analyzed in Chauris and Cocher (2018), the ratio between D_i in B and D_i in C and D is approximately equal to $2\omega/(v_0 k_z)$ for limited surface offset data. It means that a vertical derivative in z in the image domain is replaced by a time derivative in the data domain. The factor $2/v_0$ could be understood as the time to depth conversion in a simple case. Note that from a practical point of view, Guitton et al. (2007) proposed to add this vertical derivative to remove some migration artifacts. Later, Weibull and Arntsen (2013); Wang et al. (2019) introduced the same term in the MVA context: this can be seen as the first step towards Inversion Velocity Analysis (IVA).

The derivatives with respect to the vertical position of the source and receiver in D_s, D_d and D_r are either applied to the Green's function computed in v_0 (strategies A and D) or to the data (strategies B and C) for which the authors supposed that the velocity model is exact in the near-surface. The vertical derivatives can be computed with dipole sources, through the $(\omega - k)$ relationship (equation 31 in ten Kroode (2012)) or using free surface multiples (Hou and Symes, 2015).

We conclude from this analysis that approaches A, C and D are almost identical. Hou and Symes (2017) compared B (more time-consuming because of the D_i factor) and C on the Marmousi model and concluded that the main differences come from non-reflecting waves such as diving waves for which the formalism is not designed.

Triplicated wavefields are properly handled in the subsurface extended domain (Stolk and Symes, 2004). Only ten Kroode (2012) really addresses this issue in the inversion formalism. Numerical examples tend to confirm the conclusion (Hou and Symes, 2015). However, diving waves and grazing waves (i.e. arriving with a tangent direction at the surface) cannot be reconstructed from the inversion formula. A simple way is to analyze the D_i, D_s and D_r for example in the D–approach: in a high frequency asymptotic sense and for $h = 0$, the application of ∂z to the two Green's functions acts as a multiplication by $\cos\theta \cos\gamma$, where θ and γ are the half opening and dip angles. Similarly, the application of ∂s_z includes a multiplication by $\cos\beta_s$, where β_s is the emerging angle at the source. The cosine values (associated with the half opening angle) are small for div-

ing and grazing rays, preventing from a stable mapping between the data and extended image domains.

The D_i term is more complicated in theory and one should consider the velocity v_0 at $\mathbf{x}-\mathbf{h}$ and $\mathbf{x}+\mathbf{h}$ (equation 51 in ten Kroode (2012) for A; equation A–34 in Hou and Symes (2015) for B; equation A–7 in Hou and Symes (2017) for C; equation 7 in Chauris and Cocher (2017) for D).

$$D_i = \frac{1}{2}\left(\frac{v_0(\mathbf{x}-\mathbf{h})}{v_0(\mathbf{x}+\mathbf{h})}+1\right)\partial_z + \frac{1}{2}\left(\frac{v_0(\mathbf{x}-\mathbf{h})}{v_0(\mathbf{x}+\mathbf{h})}-1\right)\partial_{h_z} \tag{2.19}$$

For the currently published implementations, the D_i expressions from Table 2.1 have been used. There is in principle a need from computing the derivative with respect to the vertical h_z variable for $h_z = 0$. Alali et al. (2020) were able to implement the full expression for the D_i value, but it did not result in significant modifications on the numerical investigations.

2.12 Appendix III: Equivalences between data and image domains

In this third Appendix, we analyze the possible equivalences between data-domain and image-domain strategies (Chauris et al., 2017). Data-domain approaches are typically travel time tomography (Bishop et al., 1985), slope tomography (Billette and Lambaré, 1998; Alerini et al., 2007; Lambaré, 2008), or for wave-equation-based methods Migration-Based Travel Time (MBTT, Plessix et al. (1995); Clément et al. (2001)) and the equivalent approaches Reflection Waveform Inversion (RWI, Hicks and Pratt, 2001; Brossier et al., 2015; Zhou et al., 2015). We refer to Yao et al. (2020) for a review on RWI.

The important point is the notion of kinematic invariants, already introduced in Figure 2.1 (Chauris et al., 2002a; Guillaume et al., 2008; Tavakoli et al., 2017; Sambolian et al., 2021) (Figure 2.15). We provide here more quantitative results. A locally coherent event in the data domain is in 2D characterized by the source and receiver position, a two-way travel time, and two slopes (Billette and Lambaré, 1998). In the image domain, the equivalent is a (x, z) position and two dips, along the x and surface-offset directions. Chauris et al. (2002a); Guillaume et al. (2008) have explained how to alternate between the two domains.

We first provide a summary of the results, for different schemes, including subsurface offset, here a new aspect (Table 2.2). Montel and Lambaré (2019) indicated that the residual move-out depends on the choice of the extension. We then add more details for quantitative expressions in the angle and subsurface offset extensions (second part of this Appendix). In all cases, the source and receiver positions, as well as the total

Table 2.2: Kinematic invariants. The fitted quantity is zero by construction of the migration scheme. The mid-point slope p_m and subsurface-offset slope p_H are defined as $p_m = p_s + p_r$ and $p_H = p_s - p_r$, respectively. The weights α_s and α_r are specified in the text. For the subsurface offset case, the fit to the two slopes are always satisfied, but possibly from different subsurface locations.

Migration scheme	Fitted quantity	Data residual
shot	$p_r - p_r^* = 0$	$p_s - p_s^*$
receiver	$p_s - p_s^* = 0$	$p_r - p_r^*$
surface offset	$p_m - p_m^* = 0$	$p_H - p_h^*$
angle	$\alpha_s(p_s - p_s^*) +$	$\alpha_s(p_s - p_s^*) -$
	$\alpha_r(p_r - p_r^*) = 0$	$\alpha_r(p_r - p_r^*)$
subsurface offset	$p_s - p_s^* = 0$	unique image point $\mathbf{x}$
	$p_r - p_r^* = 0$	

travel time are automatically fitted. For the common shot migration scheme, a summation is performed over all receivers: this ensures that the derivative of the travel time with respect to the receiver position p_r fits the slope in the data p_r^*. On the source side, the source slope p_s may differ from the observed data source slope p_s^*. Chauris et al. (2002a) show that the data residual is proportional to the moveout residual in the CIGs defined along the source axis $\partial z/\partial s$. The similar expression holds for the common receiver case. For the offset case, the mid-point slope $p_m = p_s + p_r$ is satisfied, while the fit on the surface-offset slope $p_H = p_s - p_r$ depends on the macro-model. For the angle and subsurface-offset cases, similar but more complex relationships exist. In practice, the weights in front of the p_s and p_r terms are not constant and depend on ray quantities (Table 2.2). Finally, for the subsurface offset and only in this case, the two slopes are fitted at the source and receiver sides, but the associated rays may start from two different positions, shifted by $2\mathbf{h}$ (Table 2.2).

Based on these results, Chauris et al. (2017) among others proposed a wave-equation approach formulated in the data domain. The new objective function reads $1/2||FA_{\text{n}}\xi_{\text{inv}}||^2$: the Born modeling operator is applied to the reflectivity section multiplied by the annihilator. The first results indicate less imprint of the reflectivity on the tomographic gradient (Chauris et al., 2017). Diaz and Sava (2017) propose to alternate between the image- and data-domains.

We now detail the derivations (Table 2.2). We extend the approach by Chauris et al. (2002a) from the shot and offset cases to the angle and subsurface cases. The derivations

are performed in 2D.

Common angle migration

Let consider a single locally event in the observed data, characterized by the source and receiver positions (s, r), a travel time $t(s, r)$ and two observed slopes $(\partial t/\partial s, \partial t/\partial r) = (p_s^*, p_r^*)$ along the source and receiver axis, respectively. In the image domain, it is associated to a reflection event at position (x, z), with a dip γ and an opening angle θ, or by two angles $(\theta_s, \theta_r) = (\gamma + \theta, \gamma - \theta)$ defined at the scatter position and pointing towards the source and receiver positions. The ray-based travel times are denoted by τ_s and τ_r.

By definition, the total travel time should fit the observed time

$$\tau_s(x, z, \theta_s) + \tau_r(x, z, \theta_r) = t(s(x, z, \theta_s), r(x, z, \theta_r)) \tag{2.20}$$

As there is an integration over all dip angles γ for a fixed scattering angle θ, the derivation of equation 2.20 with respect to γ and for a fixed θ yields

$$\frac{\partial \tau_s}{\partial \theta_s} + \frac{\partial \tau_r}{\partial \theta_r} = \frac{\partial s}{\partial \theta_s}\frac{\partial t}{\partial s} + \frac{\partial r}{\partial \theta_r}\frac{\partial t}{\partial r} \tag{2.21}$$

By definition of the horizontal components of the slopes at the surface (p_s, p_r), we finally have

$$\alpha_s p_s + \alpha_r p_r = \alpha_s p_s^* + \alpha_r p_r^*, \tag{2.22}$$

with $\alpha_s = \partial r/\partial \theta_s$ and $\alpha_r = \partial r/\partial \theta_r$. To evaluate the residual slope in the CIG, one needs to differentiate equation 2.20 for a fixed x position (i.e. value specifying the CIG location). The expression can by simplified using equation 2.22, leading to

$$\frac{\partial z}{\partial \theta} = \frac{\alpha_s(p_s^* - p_s) - \alpha_r(p_r^* - p_r)}{\frac{\partial \tau_s}{\partial z} + \frac{\partial \tau_r}{\partial z} - \frac{\partial \tau_s}{\partial \theta_s} - \frac{\partial \tau_r}{\partial \theta_r}} \tag{2.23}$$

The numerator also equals $2\alpha_s(p_s^* - p_s) = -2\alpha_r(p_r^* - p_r)$. The denominator acts as a weight in the objective function (Table 2.2).

Subsurface migration

We process along the same line for the subsurface case. Here, the travel time conditions reads

$$\tau(s, x - h, z) + \tau(r, x + h, z) = t(s, r), \tag{2.24}$$

where τ is the travel time between (x, z) and s or r. Equation 2.24 should be valid for any source and receiver positions, yielding

$$\frac{\partial \tau}{\partial s} = \frac{\partial t}{\partial s}, \tag{2.25}$$

$$\frac{\partial \tau}{\partial r} = \frac{\partial t}{\partial r}. \tag{2.26}$$

We differentiate equation 2.24 with respect to (s, z, h) for a fixed x value and as before simplify the expression, leading to

$$\frac{\partial z}{\partial h} = \frac{\frac{\partial \tau}{\partial x}(s, x-h, z) - \frac{\partial \tau}{\partial x}(r, x+h, z)}{\frac{\partial \tau}{\partial z}(s, x-h, z) + \frac{\partial \tau}{\partial z}(r, x+h, z)}, \tag{2.27}$$

$$= \frac{\frac{\sin\theta_s}{v(x-h,z)} - \frac{\sin\theta_r}{v(x+h,z)}}{\frac{\cos\theta_s}{v(x-h,z)} + \frac{\cos\theta_r}{v(x+h,z)}}. \tag{2.28}$$

In the case of a laterally invariant velocity model v, the derivative simply reads

$$\frac{\partial z}{\partial h} = \tan\theta. \tag{2.29}$$

Here, the observed slopes (p_s^*, p_r^*) are automatically fitted, but possibly for rays starting from different positions $(x-h, z)$ and $(x+h, z)$. A physical condition is obtained for $h = 0$.

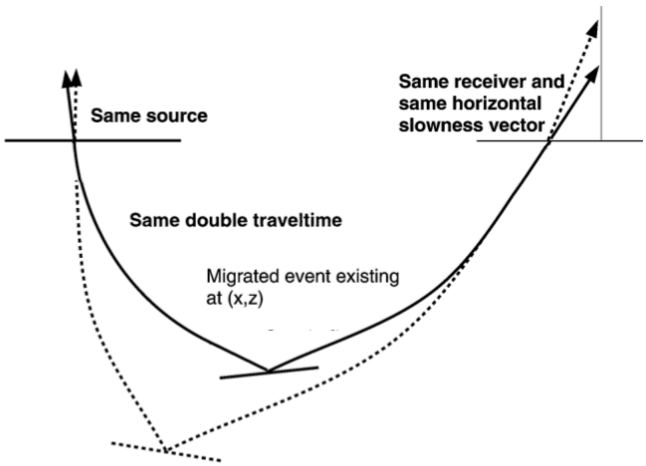

Figure 2.15: Kinematic invariants for a common shot migration in possibly two different macro-models, adapted from Chauris et al. (2002a).

Chapter 3

Extension to variable-density acoustic media

Contents

3.1 Abstract **77**
3.2 Introduction **77**
3.3 Preliminaries **79**
3.3.1 Variable density acoustic Born modeling 80
3.3.2 Pseudoinverse Born modeling 82
3.3.3 Angle domain CIGs 87
3.4 Variable density pseudoinverse modeling **89**
3.4.1 Two-trace method 89
3.4.2 Weighted least-squares (WLS) method 89
3.4.3 Taylor expansion 91
3.5 Numerical experiments **92**
3.5.1 Simple model 92
3.5.2 Marmousi2 model 102
3.6 Discussion **105**
3.7 Conclusions **112**
3.8 Acknowledgments **112**

Résumé

Ce chapitre comprend l'article intitulé « From constant- to variable-density inverse extended Born modeling », publié dans la revue *Geophysics* (Farshad and Chauris, 2020a). Pour une imagerie quantitative, la migration itérative (« iterative least-squares reserve time migration » en anglais) est l'approche préconisée. L'existence d'un inverse de l'opérateur de modélisation réduirait considérablement le nombre d'itérations nécessaires. Dans le contexte du modèle étendu, un tel pseudoinverse existe et est construit comme une version pondérée de l'opérateur adjoint. Il prend en compte la déconvolution, le facteur d'amplitude géométrique et l'illumination irrégulière. L'application du pseudoinverse suppose un milieu acoustique à densité constante, ce qui est une forte limitation pour les applications pratiques. Afin de considérer des perturbations de densité, nous proposons et investissons deux approches. La première est une généralisation d'une étude récente qui propose de retrouver les perturbations acoustiques à partir de la réponse (qui dépend de l'angle d'imagerie) de l'opérateur de modélisation. Cette nouvelle version est basée sur une fonction objective au sens des moindres carrés. La méthode non seulement apporte des résultats plus robustes, mais elle offre aussi la possibilité d'inclure des contraintes pour réduire les couplages entre paramètres. Nous proposons aussi une seconde approche avec une expansion de Taylor qui ne demande pas de transformée de Radon. Les exemples numériques sur les modèles simples et sur Marmousi2, avec des champs de vitesse corrects et incorrects, montre l'efficacité de la première méthode. L'approche de Taylor contient beaucoup trop d'artefacts pour être applicable en pratique.

From constant- to variable–density inverse extended Born modeling
Milad Farshad and Hervé Chauris
Geophysics, 2020, **85**(4), pages S217–S232
doi: 10.1190/geo2019-0489.1

3.1 Abstract

For quantitative seismic imaging, iterative least-squares reverse time migration is the recommended approach. The existence of an inverse of the forward modeling operator would considerably reduce the number of required iterations. In the context of the extended model, such a pseudo-inverse does exist and is built as a weighted version of the adjoint that accounts for the deconvolution, geometrical spreading and uneven illumination. The application of the pseudoinverse Born modeling is based on constant density acoustic media, which is a limiting factor for practical applications. To consider density perturbations, we propose and investigate two approaches. The first one is a generalization of a recent study proposing to recover acoustic perturbations from the angle-dependent response of the pseudoinverse Born modeling operator. This new version is based on weighted least-squares objective function. The method not only provides more robust results, but also offers the flexibility to include constraints in the objective function in order to reduce the parameters cross-talk. We also propose an alternative approach based on Taylor expansion that does not require any Radon transform. Numerical examples based on a simple and the Marmousi2 models using correct and incorrect background models for the variable density pseudoinverse Born modeling, verify the effectiveness of the weighted least-squares method when compared with the other two approaches. The Taylor expansion approach appears to contain too many artifacts for a successful applicability.

3.2 Introduction

Least-squares reverse time migration (LSRTM) has been proposed to overcome the shortcomings of standard reverse time migration (RTM) due to the imperfections of the seismic data resulting from significant aliasing, noise, irregular source and receiver sampling and finite recording aperture (Baysal et al., 1983; Nemeth et al., 1999; Dai et al., 2011; Zeng et al., 2014; Zhang and Schuster, 2014; Zhang et al., 2015b). In parallel, a time/space shift extended imaging condition has been proposed to potentially decouple the data fitting in LSRTM from the choice of the velocity model (Sava and Fomel, 2006; Symes, 2008b; Yang et al., 2019a). Mathematically speaking, LSRTM is trying to iteratively provide a better approximation to the inverse of the forward Born

modeling operator as required in theory, rather than applying the adjoint of the Born modeling operator to the seismic data as done in RTM. Two aspects are currently under development: the estimation of multiparameters beyond the constant density acoustic case and of efficient preconditioners for a fast LSRTM.

With the deployment of multi-component seismic data, there has been an increasing interest in extending LSRTM to multiparameter imaging in both acoustic and elastic media to provide a better description of wave propagation. Typically, density and attenuation are additional parameters included in the inversion framework (Dutta and Schuster, 2014; Yang et al., 2016b; Sun et al., 2018).

The main drawback of LSRTM is that it should be solved iteratively. Since each iteration requires the application of modeling (demigration) and adjoint (migration) operators, the computational expense of LSRTM can be considerable. Several strategies such as multi-source approach with random or linear-phase encoding (Dai et al., 2012; Xue et al., 2016) have been proposed to increase the efficiency of LSRTM.

In the context of subsurface offset extended RTM, recently, different explicit pseudoinverse expressions for the Kirchhoff modeling operator (ten Kroode, 2012) and for the Born modeling operators (Hou and Symes, 2015, 2017; Chauris and Cocher, 2017) have been proposed. The pseudoinverse operator is an alternative to the adjoint operator and provides quantitative properties within a single iteration. Although the derivations are performed under the high-frequency approximation (ray-theory), the final formulas do not contain any ray quantities but only time and spatial derivatives. It appears that the technique is very similar to the standard migration scheme, with only additional weights in the imaging operator (Hou and Symes, 2015, 2017; Chauris and Cocher, 2017). These new operators are explicit and simple in terms of implementation. They can also be used as pre-conditioners (a way to speed up the resolution of the inverse problem) for example for the full waveform inversion and migration velocity analysis (Chauris and Cocher, 2017; Hou and Symes, 2018; Li and Chauris, 2018). The theory of pseudoinverse Born modeling is established based on the constant-density acoustic wave equation. However in reality, the density of the Earth is inhomogeneous. Moreover, the amplitude of reflected seismic waves is mostly affected by acoustic impedance contrasts, the product of velocity with density. Therefore, if the P-wave velocity is the only variable parameter in the pseudoinverse Born modeling operator while density varies in reality, the inverted parameter will not be correctly estimated.

Very recently, Dafni and Symes (2018) extended the method proposed by Hou and Symes (2017) to variable density acoustic media. Their method consisted of two steps: application of pseudoinverse Born operator (Hou and Symes, 2017), and inversion of two parameters using two traces from the angle-dependent response of the result obtained from the first step. We refer to their method as the "two-trace" approach in this paper. Originally, this idea was proposed by Zhang et al. (2014b) in the context of marine acquisition. They delineate the impedance and velocity perturbations from near-angle and far-angle traces of angle-dependent response of amplitude-preserving RTM,

respectively (Zhang et al., 2014b). Compared to constant density acoustic, variable density inverse RTM provides a better description of the wave propagation and generates more accurate images. Moreover, the resulting models can further be used to study AVO effects that play an important role in lithology analysis and fluid discrimination.

In this paper, we propose a generalization of the two-trace method proposed by Dafni and Symes (2018) (least-squares method) as well as an alternative approach (Taylor expansion). We show that the final results obtained via the two-trace method slightly depends on the choice of angles for inversion. The least-squares method is based on the use of all traces in the angle domain, while the Taylor approach is based on the Taylor expansion of the Radon transform around zero angle. We discuss this method with respect to the dependency on the maximum surface offset and on artifacts in image domain. For all approaches, the starting point (first step) is a pseudoinverse Born operator either proposed by Hou and Symes (2015), Hou and Symes (2017) or Chauris and Cocher (2017). Here, in order to have a fair comparison between the two-trace approaches and the proposed extensions, we use the pseudoinverse operator proposed by Chauris and Cocher (2017) as the first step. We parameterize the subsurface via the inverse of the bulk modulus and density and invert these parameters from the angle-dependent response of the pseudoinverse Born modeling operator. We compare the three approaches by evaluating the data misfit and inverted parameters reconstruction. We also analyze the sensitivity of the methods to incorrect background models, which is not addressed in Dafni and Symes (2018). In all these approaches, the key factor determining an accurate match in image and shot domains is the pseudoinverse modeling of constant density acoustic media; their extension is easy to variable density acoustic media. The work should be understood as follows: the derivation of a proper preconditioner in the variable density acoustic case should enlarge the applicability of LSRTM, by reducing the number of required iterations.

The paper is organized as follows: we first review the preliminaries required for the extension to variable density, namely the variable density acoustic Born modeling, the adjoint and the pseudoinverse constant density Born modeling, as well as the Radon transform. Then, we explain how to extend the constant density pseudoinverse Born operator to variable density acoustic media. We present synthetic examples to compare and discuss the three approaches. Then, we apply the preferred method on the variable density Marmousi2 model. Finally, we discuss the prospects for further development of the pseudoinverse Born modeling.

3.3 Preliminaries

We give here a brief review of the concepts and formulas of the variable density acoustic wave equation, the adjoint and the pseudoinverse operators for the Born modeling, and the Radon transform.

3.3.1 Variable density acoustic Born modeling

The acoustic Earth model is parameterized with two parameters (at each point), involving P-wave velocity, V_p, and density, ρ, or their combinations, for instance, P-wave impedance, $I_p = \rho V_p$, or inverse of bulk modulus, $\beta = \frac{1}{\rho V_p^2}$. Each model parameter, for example $\mathbf{m}$, can be considered as the sum of the background model $\mathbf{m_0}$, controlling the kinematics of the wave propagation, and the model perturbation $\delta\mathbf{m}$, creating new types of waves and reflections, where both depends on the spatial coordinates $\mathbf{x} = (x, z)$ (Symes, 2008b). By definition,

$$\mathbf{m}(\mathbf{x}) = \mathbf{m}_0(\mathbf{x}) + \delta\mathbf{m}(\mathbf{x}), \tag{3.1}$$

and, under the Born approximation, we suppose that $\delta\mathbf{m}(\mathbf{x}) \ll \mathbf{m}_0(\mathbf{x})$. The definition of the perturbation model ($\delta\mathbf{m}$) can be extended to depend on more degrees of freedom. The most recent conventional choice for the extension is the subsurface offset ($\mathbf{h}$), introduced as an offset between the sunken sources and receivers by Claerbout (1985). Here, we only consider the horizontal subsurface extension as $\mathbf{h} = (h, 0)$ for the 2-D case. By using this approach, the dimension of the model space becomes the same as the data space (Table 3.1) allowing to compensate for errors in the background model (Sava and Fomel, 2006; Symes, 2008b). The mathematical expression between physical $\mathbf{m}(\mathbf{x})$ and extended domains $\mathbf{m}(\mathbf{x}, \mathbf{h})$ can be simply defined as

$$\begin{aligned} \mathbf{m}(\mathbf{x}, \mathbf{h}) &= \mathbf{m}(\mathbf{x})\delta(\mathbf{h}), \\ \mathbf{m}(\mathbf{x}) &= \int \mathrm{d}\mathbf{h}\, \mathbf{m}(\mathbf{x}, \mathbf{h}). \end{aligned} \tag{3.2}$$

Equation 3.2 is only valid when the background model is correct. For well-focused noise-free data, extracting the image at $\mathbf{h} = 0$ m could be reasonably accurate. However, it is more robust to sum over the $\mathbf{h}$-axis. In other words, one might expect the estimated physical image so obtained to be less sensitive to incoherent or numerical noise (Hou and Symes, 2015).

Table 3.1: Dimension of the data and model domains. s and r are the source and receiver coordinates; t is the time; x , y and z are the spatial coordinates; and h is the subsurface offset.

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

We also denote the source and receiver positions as $\mathbf{s}$ and $\mathbf{r}$, respectively. We parameterize the subsurface via the inverse of the bulk modulus and density. The solution of the 2-D acoustic scattered wavefield under the Born approximation can be expressed by introducing the reference Green's function G_0 in a given background model $\mathbf{m}_0$. An integral operator expression for the extended Born modeling operator $\mathcal{B}$ (Symes, 2008b) can be written as an integral over all scatter positions:

$$d(\mathbf{s},\mathbf{r},\omega) = \mathcal{B}[\mathbf{m_0}]\delta\mathbf{m}(\mathbf{s},\mathbf{r},\omega) = \Omega(\omega)\int \mathrm{d}\mathbf{x}\,\mathbf{F}_{(\beta,\rho)}(\mathbf{x},\mathbf{h},\omega;\mathbf{s},\mathbf{r})\delta\mathbf{m}_{(\beta,\rho)}(\mathbf{x},\mathbf{h}), \quad (3.3)$$

where ω is the angular frequency, $\Omega(\omega)$ is the source spectrum, $\mathbf{F}_{(\beta,\rho)}$ is the extended modeling vector given as (Symes, 2008b):

$$\mathbf{F}_{(\beta,\rho)}(\mathbf{x},\mathbf{h},\omega;\mathbf{s},\mathbf{r}) = \begin{bmatrix} -(i\omega)^2 G_0(\mathbf{s},\mathbf{x}-\mathbf{h},\omega)G_0(\mathbf{x}+\mathbf{h},\mathbf{r},\omega) \\ \\ \frac{1}{\rho_0^2}\nabla G_0(\mathbf{s},\mathbf{x}-\mathbf{h},\omega)\cdot\nabla G_0(\mathbf{x}+\mathbf{h},\mathbf{r},\omega) \end{bmatrix}^T, \quad (3.4)$$

$\delta\mathbf{m}_{(\beta,\rho)}$ is the model vector given as:

$$\delta\mathbf{m}_{(\beta,\rho)}(\mathbf{x},\mathbf{h}) = \begin{bmatrix} \delta\beta(\mathbf{x},\mathbf{h}) \\ \\ \delta\rho(\mathbf{x},\mathbf{h}) \end{bmatrix}, \quad (3.5)$$

and T denotes the transpose operator. The composition $\mathcal{B}[\mathbf{m_0}]\delta\mathbf{m}$ denotes computed data d in $(\mathbf{m_0},\delta\mathbf{m})$. Since different parameter classes have different physical units and nature, they can have different influence on the data. Note that the influence of one parameter on the data depends on the other parameters involved in the subsurface parameterization. Generally, this is referred to as parameter cross-talk, meaning that parameters are more or less coupled (Virieux and Operto, 2009). The diffraction or radiation pattern can give some insight into influence of the parameterization in the data as a function of the diffraction angle (Operto et al., 2013). The analytical expression for the diffraction pattern can be derived in the framework of the Ray+Born approximation (Forgues and Lambaré, 1997). Thus, under the high-frequency approximation, the modeling vector $\mathbf{F}_{(\beta,\rho)}$ (equation 3.4) can be rewritten as:

$$\mathbf{F}_{(\beta,\rho)}(\mathbf{x},\mathbf{h},\omega;\mathbf{s},\mathbf{r}) = \begin{bmatrix} -(i\omega)^2 G_0(\mathbf{s},\mathbf{x}-\mathbf{h},\omega)G_0(\mathbf{x}+\mathbf{h},\mathbf{r},\omega) \\ \\ \frac{\beta_0(i\omega)^2}{\rho_0} G_0(\mathbf{s},\mathbf{x}-\mathbf{h},\omega)G_0(\mathbf{x}+\mathbf{h},\mathbf{r},\omega)\cos(2\gamma) \end{bmatrix}^T \quad (3.6)$$

$$= (i\omega)^2 G_0(\mathbf{s}, \mathbf{x}-\mathbf{h}, \omega) G_0(\mathbf{x}+\mathbf{h}, \mathbf{r}, \omega) \left[-1 \quad \frac{\beta_0}{\rho_0}\cos(2\gamma)\right],$$

where γ denotes the diffraction angle. Figure 3.1 shows the graphic representation of diffraction pattern for $\mathbf{m}_{(\beta,\rho)}$ as a function of the diffraction angle. As expected from equation 3.6, the amplitude of scattered wavefield by β perturbation in the (β, ρ) parameterization shows an isotropic pattern (red line in Figure 3.1), while the amplitude by ρ perturbation shows an increasingly decreasing pattern from the small γ to the wider ones (blue line in Figure 3.1).

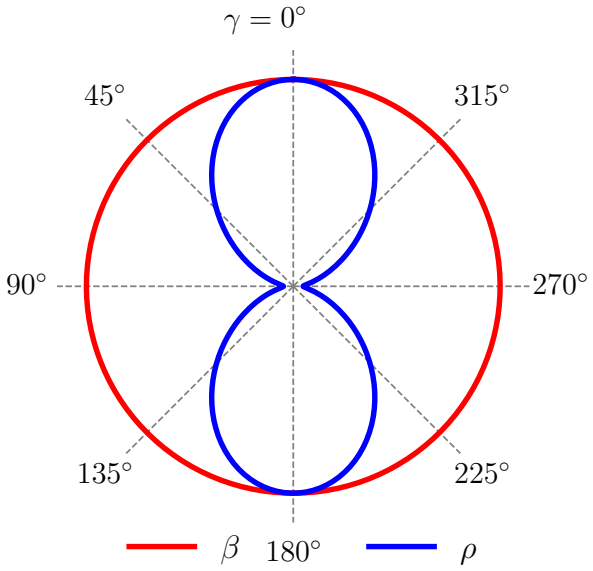

Figure 3.1: The analytical diffraction pattern for an acoustic medium parameterized by (β, ρ).

Equation 3.6 provides a linear relationship between the data and the $(\delta\beta, \delta\rho)$ perturbations. In the next section, we review the pseudoinverse operator providing optimal $\delta\beta(\mathbf{x}, \mathbf{h})$ from observed data in the case when $\delta\rho = 0$. Then we introduce the Radon transform to handle angle γ appearing in equation 3.6 when $\delta\rho \neq 0$.

3.3.2 Pseudoinverse Born modeling

RTM is one of the most powerful seismic imaging methods in complex geology. The classical form of the RTM operator can be expressed as the correlation between the forward and back-propagation of source and receiver wavefields, respectively (Tarantola, 1984). By considering constant density acoustic media ($\delta\rho = 0$), for a specific background model (β_0), migration is introduced by minimizing the misfit between the

observed (d^{obs}) and computed data ($d(\xi)$) at each shot position as:

$$J_0(\xi) = \frac{1}{2}\left\|d(\xi) - d^{obs}\right\|^2, \tag{3.7}$$

where $\xi = \delta\beta$ is the extended reflectivity and $\|\,.\,\|^2$ denotes the ℓ_2-norm. The operator of the migration for determining ξ can be defined by deriving the gradient of equation 3.7 with respect to ξ as $(\partial J_0/\partial\xi)|_{\xi=0}$ (Lailly, 1983; Tarantola, 1984), yielding

$$\begin{aligned}\xi &= \mathcal{B}^T(d^{obs})(\mathbf{x},\mathbf{h}) \\ &= -\int \mathrm{d}\mathbf{s}\,\mathrm{d}\mathbf{r}\,\mathrm{d}\omega\,(i\omega)^2\Omega^*(\omega)G_0^*(\mathbf{s},\mathbf{x}-\mathbf{h},\omega)d^{obs}(\mathbf{s},\mathbf{r},\omega)G_0^*(\mathbf{x}+\mathbf{h},\mathbf{r},\omega),\end{aligned} \tag{3.8}$$

where * denotes the complex conjugate. RTM (equation 3.8) is the adjoint of the Born modeling operator ($\mathcal{B}^T$), whereas inversion ($\mathcal{B}^\dagger$) is its asymptotic inverse. Here, for pseudoinverse Born modeling, we considered the method proposed by Chauris and Cocher (2017) which is close to the one presented in Hou and Symes (2017). Both derivations determine an pseudoinverse $\mathcal{B}^\dagger$ of the extended Born modeling operator $\mathcal{B}$ such that $\mathcal{B}^\dagger\mathcal{B} \approx I$, where I is the identity operator. However, Hou and Symes (2017) directly apply the stationary phase approximation on $\mathcal{B}^\dagger\mathcal{B}$, whereas Chauris and Cocher (2017) use a linearization of the phase of $\mathcal{B}^\dagger\mathcal{B}$, leading to different final formulations for $\mathcal{B}^\dagger$ even if both formulations are valid in an asymptotic sense (Chauris and Cocher, 2018). Following the work by Chauris and Cocher (2017), the pseudoinverse formula for the extended Born modeling operator in a constant density acoustic media can be written as:

$$\begin{aligned}\xi &= \mathcal{B}^\dagger(d^{obs})(\mathbf{x},\mathbf{h}) \\ &\simeq 32\frac{\beta_0}{\rho_0^3}\partial_z\int \mathrm{d}\mathbf{s}\,\mathrm{d}\mathbf{r}\,\mathrm{d}\omega\,\frac{\Omega^\dagger(\omega)}{(i\omega)}\partial_{s_z}G_0^*(\mathbf{s},\mathbf{x}-\mathbf{h},\omega)d^{obs}(\mathbf{s},\mathbf{r},\omega)\partial_{r_z}G_0^*(\mathbf{x}+\mathbf{h},\mathbf{r},\omega),\end{aligned} \tag{3.9}$$

where $\Omega^\dagger = \frac{\Omega^*}{\|\Omega\|^2}$ is the inverse of the seismic source. In terms of implementation, equation 3.9 is close to the standard migration algorithm (equation 3.8), with three main modifications: (1) applying vertical derivative with respect to source and receiver positions to the Green's functions; (2) using the inverse version of the source wavelet instead of adjoint version; (3) applying a first-order integration in time before cross-correlation instead of a second-order derivative and finally a vertical derivative to the result of the cross-correlation. These modifications result in applying the following weights, respectively: (1) cosines of take-off angles at the sources and receivers positions; (2) deconvolution of the source wavelet; (3) cosines of the half-opening angle at the image point. A larger weight is given to small scattering angles and short surface offsets (Chauris and Cocher, 2017). In constant density acoustic media, if the investigated background model is correct, the energy focuses around the zero-subsurface offset. Otherwise, the

extended domain allows to compensate for errors in the background velocity model by defocusing the energy in the extended reflectivity. Even in this case, it is still possible to reconstruct the observed data from the extended reflectivity (Symes, 2008b). For example, Figure 3.2a shows the observed data for a single layer Earth model in a constant density acoustic media. The corresponding common-image gather (CIG) for correct ($v = 2400$ m/s) and incorrect ($v = 2700$ m/s) background models are also shown in Figure 3.2b and 3.2c, respectively. It is worth noting that the high (low) velocity results in upward (downward) curvature in CIG domain, as shown in Figure 3.2c. To evaluate the effect of the extended domain, we first reconstruct the shot for the correct and incorrect background models in the extended domain (Figure 3.3a and 3.3b). The phase and amplitude are correctly retrieved for both cases. Then, we reconstruct the data for an incorrect background model in the physical domain by simply summing CIG over all $\mathbf{h}$ values (see equation 3.2, Figure 3.3c). Careful examination of the extracted traces indicates that the phase is not well retrieved (Figure 3.3c).

In variable density acoustic media, the energy of the incident wave is partitioned at each boundary based on the contrasts in velocity and density properties across the boundary. The important aspect to note is that reflection amplitude for each boundary depends on the angle of incidence. This phenomenon is well known as amplitude variation with angle (AVA) (Chopra and Castagna, 2005). An example of observed data for a single layer Earth model in a variable density acoustic media and its reflection coefficient as a function of angle for the interface is shown Figure 3.4a and 3.4b, respectively. The velocity and density pairs for above and below the interface are (2500 m/s, 1950 kg/m^3) and (2320 m/s, 2200 kg/m^3), respectively. In such media, using equation 3.9 for inverting reflectivity will defocus the energy in extended domain to compensate the effect of AVA, even with correct background model, as can be seen in Figure 3.4c. Although it may still reconstruct the observed data, there would be no physical sense of velocity or density perturbation since the energy is defocused (Figure 3.4c).

As we showed in equation 3.6, the combination of the ray theory and the Born approximation provides quantitative estimations of the multiparameter inversion. It means that the inverted $\xi(\mathbf{x}, \mathbf{h})$ in the extended domain (equation 3.9), can be decomposed into two physical parameters $(\delta\beta(\mathbf{x}), \delta\rho(\mathbf{x}))$ based on the diffraction pattern of the specific parameterization. The equation for this relation can be written as:

$$\xi_\beta(\mathbf{x}, \mathbf{h}) \cong \delta\beta(\mathbf{x})\delta(\mathbf{h}) - \frac{\beta_0}{\rho_0}\cos(2\gamma)\delta\rho(\mathbf{x})\delta(\mathbf{h}), \tag{3.10}$$

where $\delta()$ is the Dirac delta function. In the case $\delta\rho(\mathbf{x}) = 0$, the $\delta\beta(\mathbf{x})$ can be reconstructed from $\xi_\beta(\mathbf{x}, \mathbf{h})$ by simply summing over all $\mathbf{h}$ values ($\delta\beta(\mathbf{x}) = \int \mathrm{d}\mathbf{h}\, \xi_\beta(\mathbf{x}, \mathbf{h})$). For variable density, based on the fact that inversion of $\delta\rho(\mathbf{x})$ in equation 3.10 is angle-dependent, a transformation of subsurface offset to scattering angle is required (Dafni and Symes, 2018). In the next section, we review how to transform the extended offset domain CIG to angle domain CIG via the Radon transform.

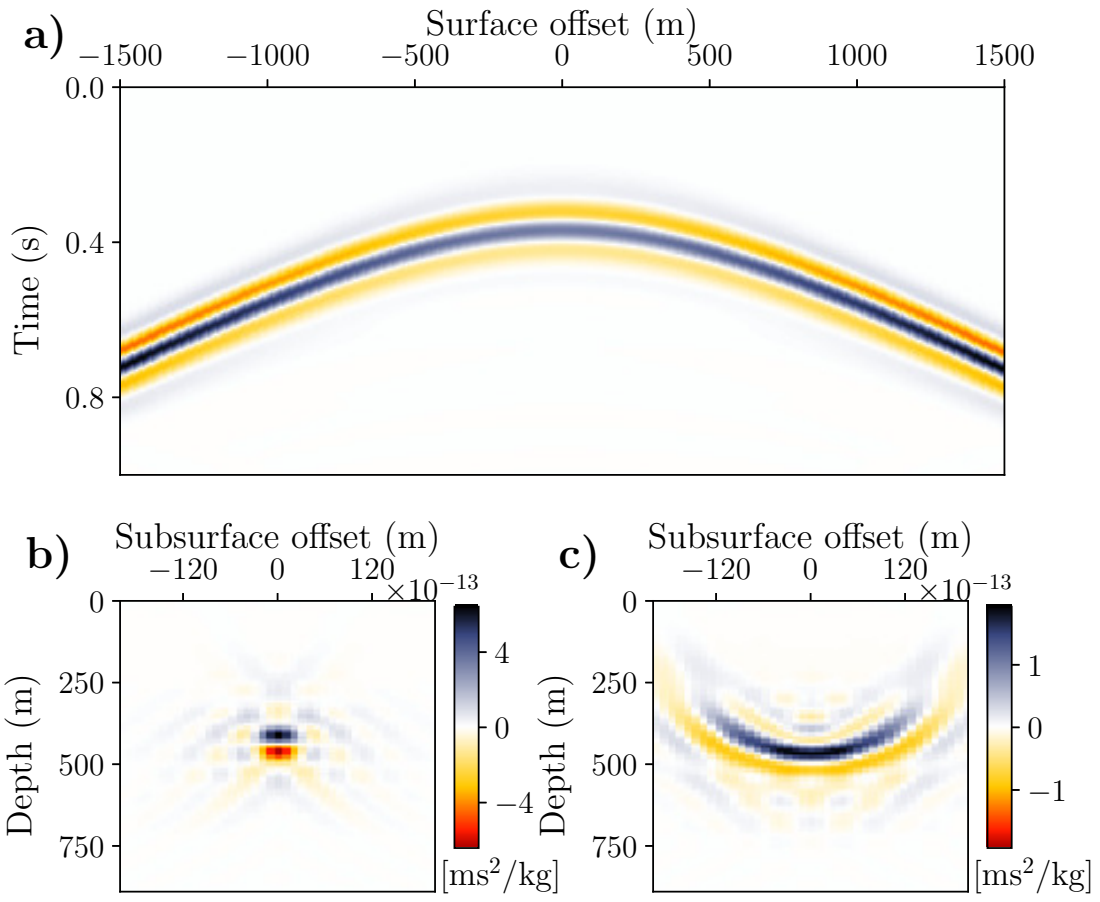

Figure 3.2: Constant density acoustic media. a) Observed data for a single layer Earth model and the corresponding CIG inverted via pseudoinverse Born modeling operator (equation 3.9) with b) correct and c) incorrect background model (high velocity).

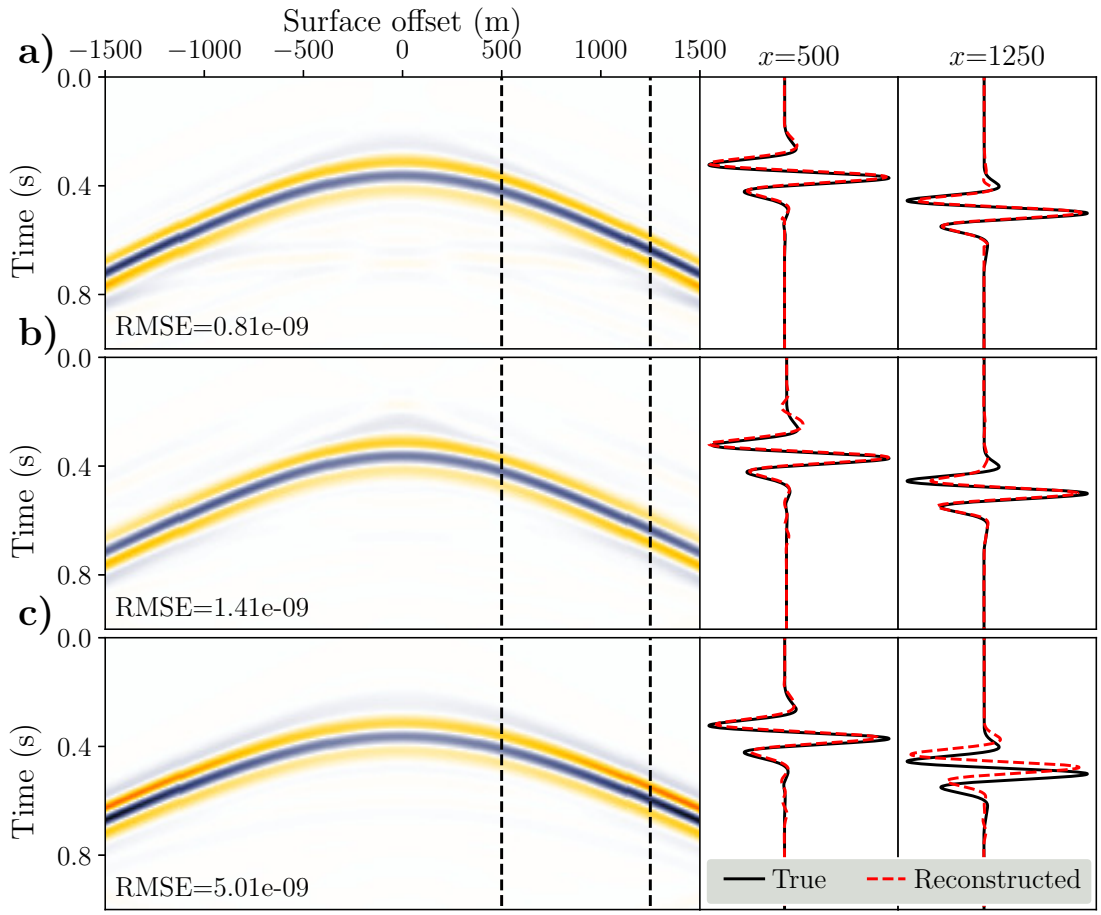

Figure 3.3: Left: reconstructed shots with a) correct velocity extended domain, b) incorrect velocity extended domain and c) incorrect velocity physical domain. The RMS error between synthetic and observed data is written on each panel. right: extracted traces for near and far offsets. Shots are plotted in the same scale.

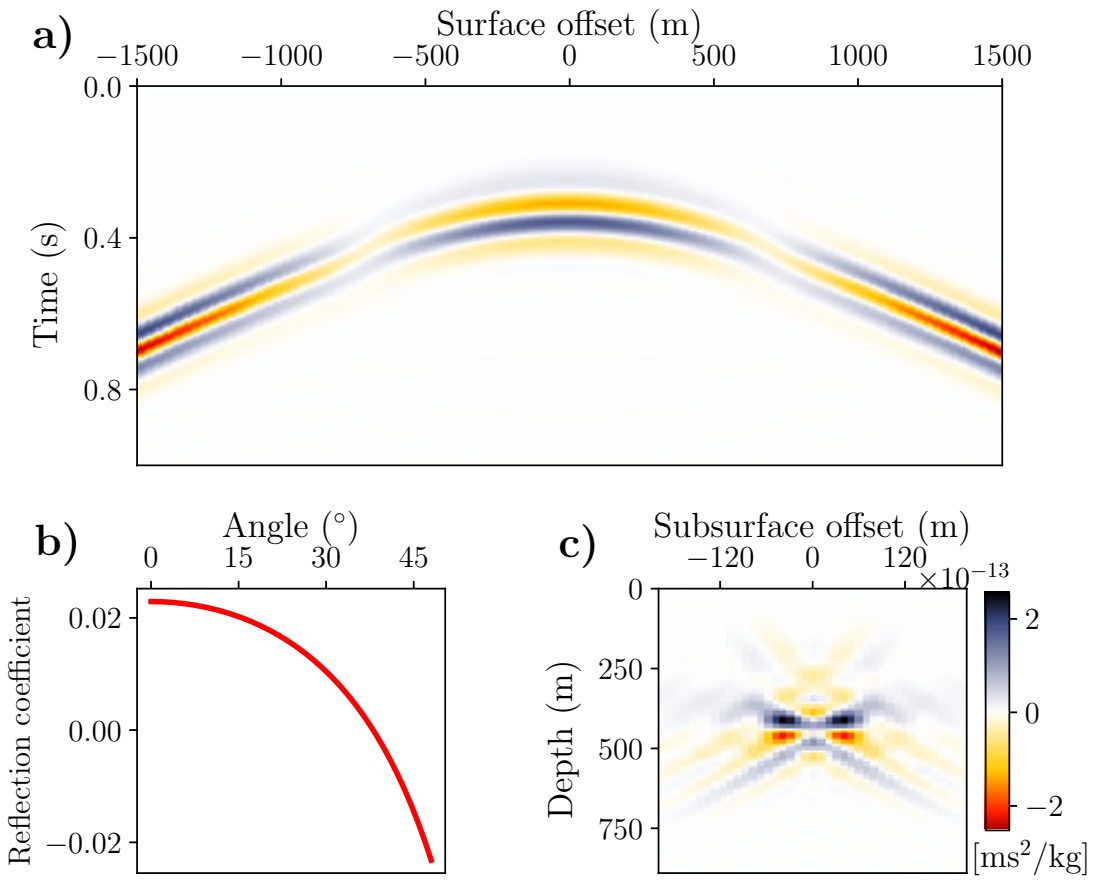

Figure 3.4: Variable density acoustic media. a) Observed data for a single layer Earth model and the corresponding b) AVA for the interface, and c) CIG inverted via pseudoinverse Born modeling operator (equation 3.9) with correct background model.

3.3.3 Angle domain CIGs

Angle gathers are the main ingredient of the AVO/AVA analysis which can give us reliable estimates of the Earth parameters, such as P-wave velocity (V_p), S-wave velocity (V_s), density (ρ), or different combinations of them. These parameters can further be used to provide information about reservoir parameters, namely, lithology, porosity and fluid content (Castagna and Smith, 1994; Chopra and Castagna, 2005). Angle gathers can be obtained using wave-equation techniques either based on wavefield methods (de Bruin et al., 1990; Sava and Fomel, 2003; Biondi and Symes, 2004; Sava and Vlad, 2011; Sava and Alkhalifah, 2013; Dafni and Symes, 2016) or ray methods (Brandsberg-Dahl et al., 2003). The fact that the wavefield methods can accurately image complex geologic structures comparing to ray-based methods makes wavefield methods superior

to ray methods. The wavefield methods can be applied either for shot-profile migration, shot-geophone migration, or prestack images after migration. Since in the first two methods the angle gathers are evaluated from wavefields prior to imaging, they are referred as "data-space methods". The angle gathers obtained from these methods are a function of offset ray parameter. Sava and Fomel (2003) showed that angle gathers can also be obtained from migrated images with a process which is completely detached from migration. The main advantages of their method are that the angle gathers are produced as a function of the reflection angle, which is not the case in data-space methods, and also with much less computational cost compared to data-space methods. This process is based on performing the Radon transform (slant stack integral) on extended CIGs. The Radon transform formula can be written in 2D as:

$$\mathcal{R}_\xi(x, z, \gamma) = \int \mathrm{d}h\, \xi(x, z + h\tan\gamma, h), \tag{3.11}$$

where $\tan\gamma$ is the trajectory of integration. Figure 3.5 illustrates the integration path for different point positions in a CIG panel. Note that it is commonly assumed that the RTM and the ray-based angle (γ in equation 3.6 and 3.11) are the same. Montel and Lambaré (2011) showed that this is not necessarily the case for incorrect background models.

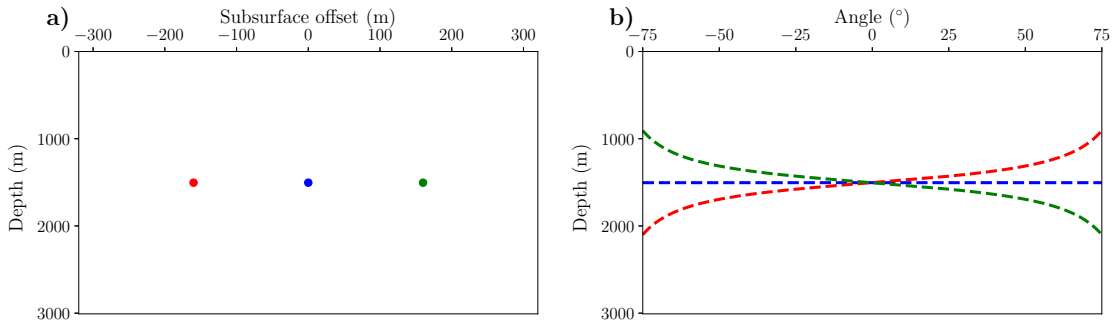

Figure 3.5: a) a CIG with three nonzero points and b) its Radon transform showing the path of integration for different points.

We now have the ingredients to take into account variable density, i.e. to determine $(\delta\beta(\mathbf{x}), \delta\rho(\mathbf{x}))$ from $\xi_\beta(\mathbf{x}, \mathbf{h})$ (equation 3.10). As the inversion formula is derived in data domain (equation 3.7), we do not expect a unique solution, but the reconstructed data from $(\delta\beta(\mathbf{x}), \delta\rho(\mathbf{x}))$ should match the observed data. In other word, we may have different combinations of $\delta\beta(\mathbf{x})$ and $\delta\rho(\mathbf{x})$, leading to approximately the same data fit. Note that $\xi_\beta(\mathbf{x}, \mathbf{h})$ is defined in the extended domain, whereas $(\delta\beta(\mathbf{x}), \delta\rho(\mathbf{x}))$ are defined in the physical domain.

3.4 Variable density pseudoinverse modeling

Here, we first review the two-trace method proposed by Dafni and Symes (2018). Then, we derive two new schemes for variable density pseudoinverse Born modeling. The first one is based on the generalization of the two-trace method, referred to as weighted least-squares (WLS), and the second one is based on the Taylor expansion of the Radon transform around $\gamma = 0^\circ$. As detailed below, the latter does not require any forward/inverse Radon transform, with a straightforward implementation. We explain its derivation and discuss its applicability.

3.4.1 Two-trace method

Recently, a multiparameter inversion method for bulk modulus and density perturbation was proposed by Dafni and Symes (2018) based on the analysis of two traces within the Radon domain. They used a similar formula as equation 3.10, to invert the perturbations from the angle-dependent response of the pseudoinverse. By applying the Radon transform on ξ (equation 3.10):

$$\mathcal{R}_\xi(\mathbf{x},\gamma) = \delta\beta(\mathbf{x}) - \frac{\beta_0}{\rho_0}\cos(2\gamma)\delta\rho(\mathbf{x}), \tag{3.12}$$

they propose to choose two traces in angle domain, as:

$$a(\mathbf{x}) = \mathcal{R}_\xi(\mathbf{x},0) = \delta\beta(\mathbf{x}) - \frac{\beta_0}{\rho_0}\delta\rho(\mathbf{x}), \tag{3.13}$$

$$b(\mathbf{x}) = \mathcal{R}_\xi(\mathbf{x},\gamma') = \delta\beta(\mathbf{x}) - \frac{\beta_0}{\rho_0}\cos(2\gamma')\delta\rho(\mathbf{x}), \tag{3.14}$$

and then calculate $\delta\beta$ and $\delta\rho$ by a system of two linear equations as:

$$\begin{aligned}\delta\rho(\mathbf{x}) &= \frac{\rho_0}{\beta_0}\frac{b(\mathbf{x})-a(\mathbf{x})}{1-\cos(2\gamma')},\\ \delta\beta(\mathbf{x}) &= \frac{b(\mathbf{x})-\cos(2\gamma')a(\mathbf{x})}{1-\cos(2\gamma')}.\end{aligned} \tag{3.15}$$

We note that inversion results for this method is angle dependent, meaning that the results might slightly change by choosing different value for γ'.

3.4.2 Weighted least-squares (WLS) method

In this paper, to get more robust results, we generalize the above approach to consider all traces in the Radon domain. Thus, we consider the whole AVA response rather than simple expansions in terms of two traces. The objective function for any given

$\mathbf{x} = (x, z)$ for optimal $\delta\beta(\mathbf{x})$ and $\delta\rho(\mathbf{x})$ can be defined as the least-squares differences between the computed and observed $\mathcal{R}_\xi$ over the all angles as:

$$J_{\delta\beta,\delta\rho} = \frac{1}{2}\left\|(\delta\beta(\mathbf{x}) - \frac{\beta_0}{\rho_0}\cos(2\gamma)\delta\rho(\mathbf{x})) - \mathcal{R}_\xi(\mathbf{x},\gamma)\right\|_W^2, \tag{3.16}$$

where $W(\mathbf{x},\gamma)$ is a weighting mask defined as:

$$W(\mathbf{x},\gamma) = \begin{cases} 1, & \text{if } |\gamma| \leq \alpha\tan^{-1}(\frac{x_{max}}{z}) \\ 0, & \text{otherwise} \end{cases}. \tag{3.17}$$

We define this mask based on the acquisition geometry to remove the artifacts in the angle domain at large angles. Thus, for each depth, we only consider the angles that would indeed be recorded. This is strictly valid for homogeneous models. In equation 3.17, α is a value close to 1, which is essential to avoid the artifacts due to the finite sampling of the offset axis. In practice, there is no need to have a very precise definition of α. We indeed choose this parameter to guarantee that we mainly include the specular events in the angle-domain CIG. Figure 3.6 shows the weighting mask for $x_{max} = 2250$ m, $\alpha = 0.85$ and $\alpha = 1$. Outside of these boundaries, the angle domain only contains the artifacts corresponding to the artifacts in the CIG domain at large subsurface offset values.

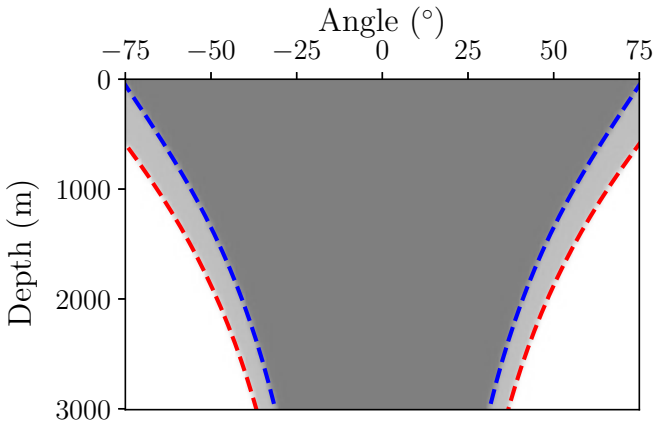

Figure 3.6: Weighting mask $W(\mathbf{x},\gamma)$. The red and blue dashed lines defines the W for $\alpha = 1$ and $\alpha = 0.85$, respectively.

To derive the optimal $\delta\beta(\mathbf{x})$ and $\delta\rho(\mathbf{x})$, we compute the gradient of the objective function J (equation 3.16) with respect to the model parameters $\delta\beta(\mathbf{x})$ and $\delta\rho(\mathbf{x})$ as:

$$\frac{\partial J}{\partial\delta\beta} = \int \mathrm{d}\gamma W(\mathbf{x},\gamma)\Big(\delta\beta(\mathbf{x}) - \frac{\beta_0}{\rho_0}\cos(2\gamma)\delta\rho(\mathbf{x}) - \mathcal{R}_\xi(\mathbf{x},\gamma)\Big) = 0, \tag{3.18}$$

$$\frac{\partial J}{\partial \delta\rho} = -\int \mathrm{d}\gamma W(\mathbf{x},\gamma)\frac{\beta_0}{\rho_0}\cos(2\gamma)\Big(\delta\beta(\mathbf{x}) - \frac{\beta_0}{\rho_0}\cos(2\gamma)\delta\rho(\mathbf{x}) - \mathcal{R}_\xi(\mathbf{x},\gamma)\Big) = 0.$$

Equation 3.18 can hence be written in a matrix formulation as:

$$\begin{bmatrix} \int \mathrm{d}\gamma W(\mathbf{x},\gamma) & -\int \mathrm{d}\gamma W(\mathbf{x},\gamma)\frac{\beta_0}{\rho_0}\cos(2\gamma) \\ -\int \mathrm{d}\gamma W(\mathbf{x},\gamma)\frac{\beta_0}{\rho_0}\cos(2\gamma) & \int \mathrm{d}\gamma W(\mathbf{x},\gamma)(\frac{\beta_0}{\rho_0})^2\cos^2(2\gamma) \end{bmatrix} \begin{bmatrix} \delta\beta(\mathbf{x}) \\ \delta\rho(\mathbf{x}) \end{bmatrix} = \begin{bmatrix} \int \mathrm{d}\gamma W(\mathbf{x},\gamma)\mathcal{R}_\xi(\mathbf{x},\gamma) \\ -\int \mathrm{d}\gamma W(\mathbf{x},\gamma)\frac{\beta_0}{\rho_0}\cos(2\gamma)\mathcal{R}_\xi(\mathbf{x},\gamma) \end{bmatrix}. \quad (3.19)$$

Solving for equation 3.19 leads to the determination of an optimal $\delta\beta(\mathbf{x})$ and $\delta\rho(\mathbf{x})$ in physical domain.

The key ingredient of the two-trace and the WLS methods is the Radon transform (equation 3.12-3.16). We now propose an alternative approach avoiding the application of the Radon transform.

3.4.3 Taylor expansion

Here, we propose an alternative approach based on Taylor expansion along the angle γ of the Radon transform around $\gamma = 0°$. The Taylor series provides a sum of the terms which will approximate the function. Thus, the Taylor expansion of the Radon transform (equation 3.11) evaluated around $\gamma = 0°$ can be written as:

$$\mathcal{R}_\xi(\mathbf{x},\gamma) \simeq \mathcal{R}_\xi(\mathbf{x},0) + \gamma\frac{\partial\mathcal{R}_\xi}{\partial\gamma}\bigg|_{\gamma=0} + \frac{1}{2}\gamma^2\frac{\partial^2\mathcal{R}_\xi}{\partial\gamma^2}\bigg|_{\gamma=0}, \quad (3.20)$$

where the first term is

$$\mathcal{R}_\xi(\mathbf{x},0) = \int \mathrm{d}h\,\xi(x,z,h), \quad (3.21)$$

the second term is

$$\frac{\partial\mathcal{R}_\xi}{\partial\gamma}\bigg|_{\gamma=0} = \int \mathrm{d}h\,h\frac{\partial\xi}{\partial z}(x,z,h), \quad (3.22)$$

and the third term is

$$\frac{\partial^2\mathcal{R}_\xi}{\partial\gamma^2}\bigg|_{\gamma=0} = \int \mathrm{d}h\,h^2\frac{\partial^2\xi}{\partial z^2}(x,z,h). \quad (3.23)$$

The first-order Taylor expansion of the $\cos(2\gamma)$ evaluated around $\gamma = 0^\circ$ can be written as

$$\cos(2\gamma) \simeq 1 - 2\gamma^2. \tag{3.24}$$

Note that the second term in equation 3.20 consists of the integration of the vertical derivative of $\xi(x, z, h)$ multiplied with h including negative and positive values. This leads the second term to be negligible compared to the first and third terms (equation 3.21 and 3.23, respectively). Substituting equation 3.20 and 3.24 into equation 3.15 gives:

$$\begin{aligned}
\delta\rho(\mathbf{x}) &\simeq \frac{\rho_0}{4\beta_0}\frac{\partial^2 \mathcal{R}_\xi}{\partial \gamma^2}\bigg|_{\gamma=0} = \frac{\rho_0}{4\beta_0}\int \mathrm{d}h\, h^2 \frac{\partial^2 \xi}{\partial z^2}(x, z, h), \\
\delta\beta(\mathbf{x}) &\simeq \mathcal{R}_\xi(\mathbf{x}, 0) + \frac{1}{4}\frac{\partial^2 \mathcal{R}_\xi}{\partial \gamma^2}\bigg|_{\gamma=0} = \int \mathrm{d}h \Big[\xi(x, z, h) + \frac{h^2}{4}\frac{\partial^2 \xi}{\partial z^2}(x, z, h)\Big].
\end{aligned} \tag{3.25}$$

The advantage of this formulation is that it does not require any application of the Radon transform. As can be seen, the $\delta\rho(\mathbf{x})$ only depends on second order derivative of the $\xi(x, z, h)$ weighted with h^2. This makes it sensitive to the artifacts in CIG domain at large subsurface offset values. Moreover, the Taylor expansion of the Radon transform is applied around $\gamma = 0^\circ$ meaning it only works for short-offset acquisitions. However, in seismic imaging we prefer long-offset acquisitions to be able to reconstruct the large and intermediate wavelengths of wavefield.

3.5 Numerical experiments

We present several numerical examples from simulated data to compare the performance of the different methods in term of accuracy, i.e., the ability for data reconstruction. For the first example, we used the same model used by Dafni and Symes (2018). We analyze the results for correct and incorrect background models for this example. In the second example, we test the application of the proposed method on a variable density Marmousi2 dataset.

3.5.1 Simple model

This model consists of four horizontal interfaces, which each of them showing different type of AVA signature. The exact β and ρ models are shown in Figure 3.7a and 3.7b. The reflection coefficient corresponding to each event is also shown in Figure 3.7c: a strong AVA effect is expected for the first and second interfaces. The sign of the amplitude for the first event changes with angle from negative to positive and it vanishes around $\gamma = 20^\circ$. For the second event, the amplitude starts from zero at $\gamma = 0^\circ$ and decreases

with angle. The model is discretized on a 189×189 grid with 16 m spacing in vertical and horizontal directions. The source wavelet for the simulation is a Ricker signal centered at 4.6 Hz (maximum frequency is 11.5 Hz). Each shot is recorded on 189 channels for 3.0 s with a 3.5 ms time interval. By considering correct background models (Figure 3.8a and 3.8b), the inverted ξ and its Radon transform are shown in Figure 3.8c-d. As expected, even by considering correct background models and inversion method, the energy in the CIG domain is defocused due to AVO/AVA effects (mainly the shallowest event). As illustrated in Figure 3.5, energy at different lags in the offset-domain CIGs leads to different integral paths in the angle domain CIGs. This leads to have more amplitude distortions at higher angles in angle domain CIGs.

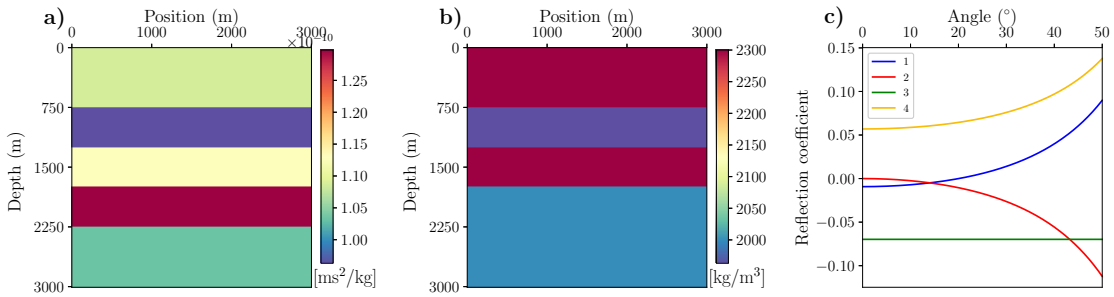

Figure 3.7: Exact a) β , b) ρ and c) reflection coefficient as a function of angle for each interface corresponding to model 1.

First, we apply the extended Born modeling operator on the inverted physical models to re-simulate the data. The observed and computed shots, and the extracted traces for near/far offsets for each shot are calculated (Figure 3.9). The colorbar scale for reconstructed data is the same as the one for the observed data. There is an accurate match between observed and computed shots both in terms of phase and amplitudes. The Root Mean Squared Error (RMSE) between observed and reconstructed shots via each method is also written on each panel showing the superiority of the WLS method. As can be seen in the extracted traces, there is a mismatch for the Taylor-expansion method, especially in the shallow part of the data containing larger angle information. This indeed is caused by the approximation of the Radon transform around $\gamma = 0°$. It is interesting to compare the observed and reconstructed data based on short-offset acquisition where the Taylor expansion should be more applicable. Therefore, we apply the modeling operator on the same model but considering a short-offset acquisition only (Figure 3.10). There is indeed better match in both shallow and deep parts. Note that the RMSE between the observed and reconstructed shots in the short-offset acquisition is 3 times less than the long-offset acquisition for the Taylor expansion method.

To further verify the effectiveness of the different methods, we compare the inverted parameters. First, we invert the β and ρ perturbations via the two-trace method for different angles (Figure 3.11). As expected, the quality of the inverted parameters may

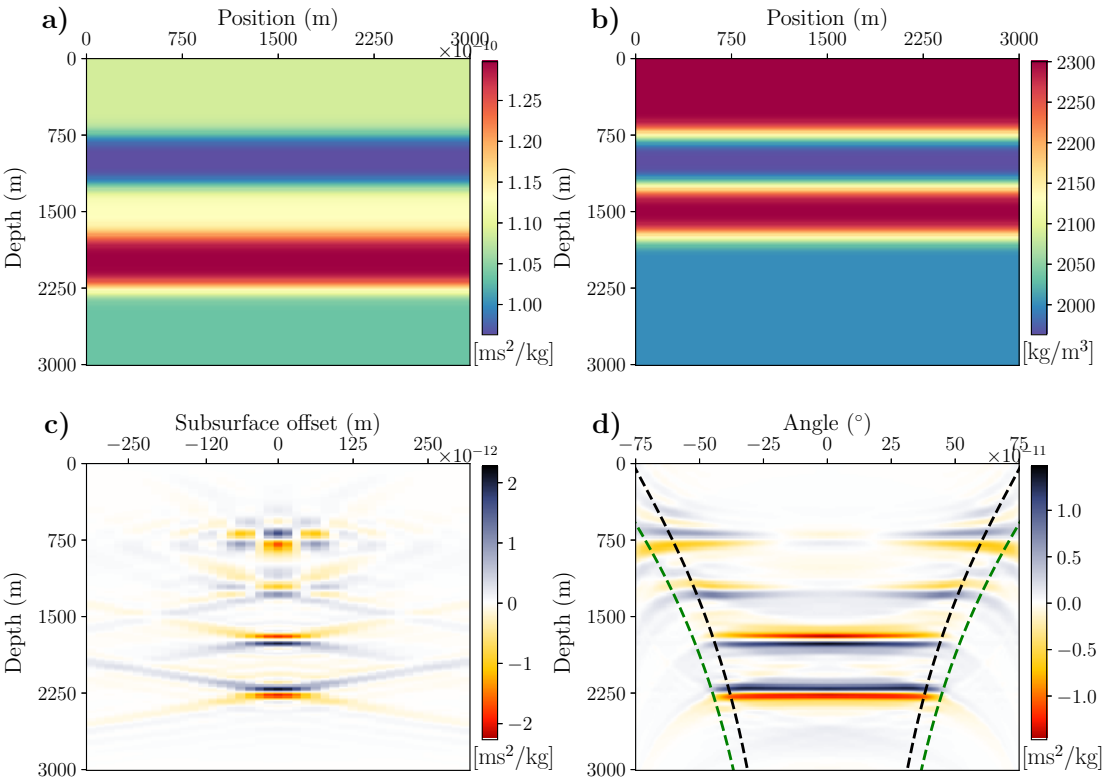

Figure 3.8: The initial correct background a) β_0 and b) ρ_0 models used for pseudoinverse Born modeling. Inverted c) ξ and its d) $\mathcal{R}_\xi$ at position $x = 1500$ m corresponding to Figure 3.7. The green and black dashed lines in (d) corresponds to the limit for acquired angles and the used ones in inversion, respectively.

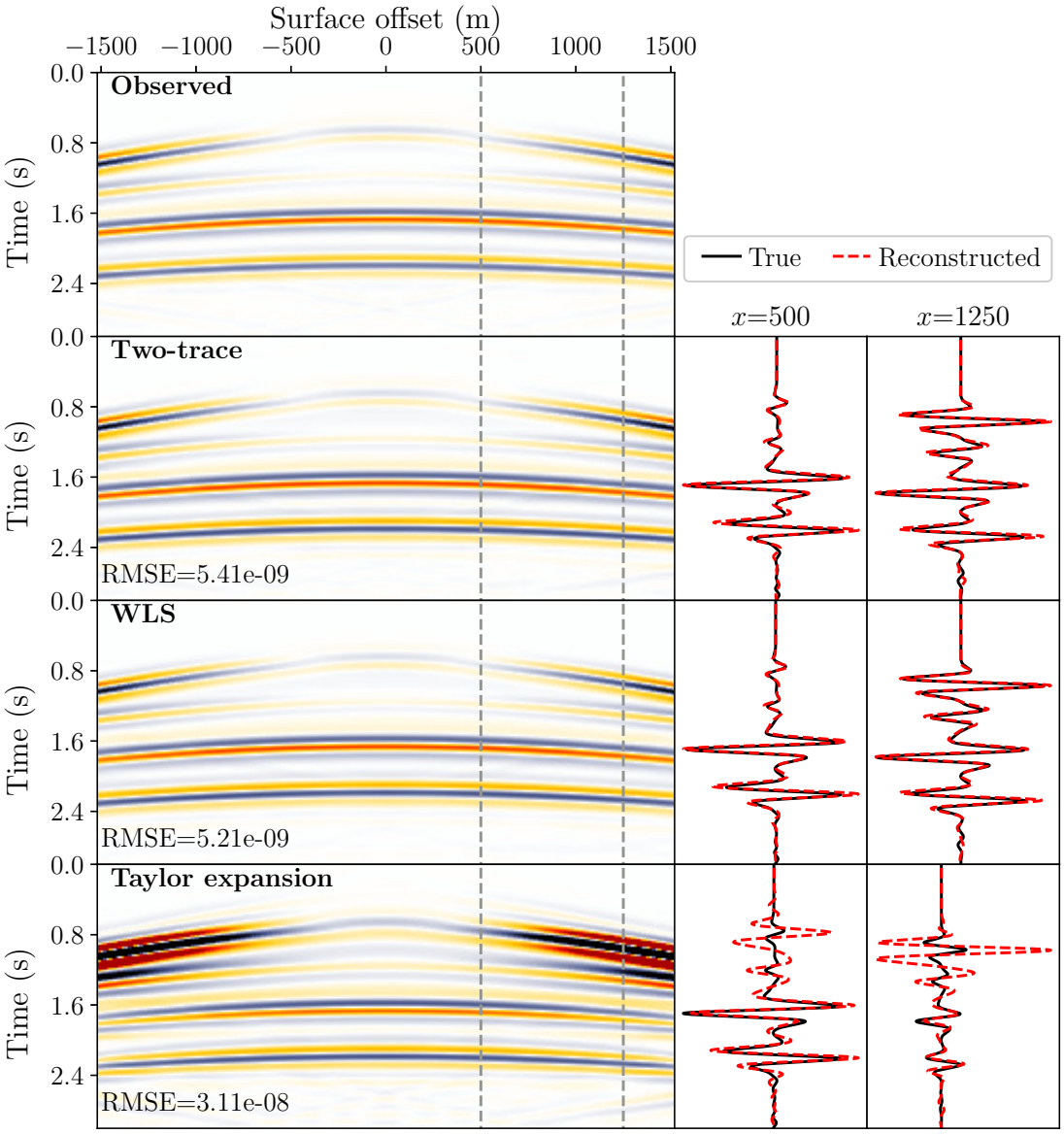

Figure 3.9: Left: observed and reconstructed shots obtained via different methods corresponding to correct background models, right: extracted traces. The RMS error between synthetic and observed data is written on each panel. Shots are plotted in the same scale.

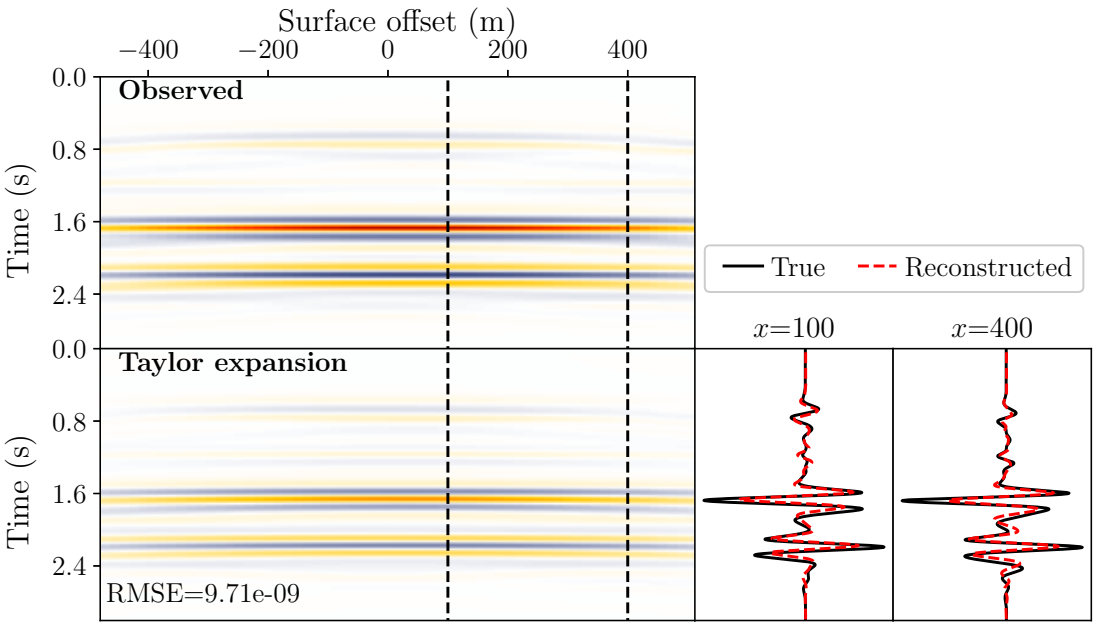

Figure 3.10: Left: observed and reconstructed shot obtained via Taylor expansion method corresponding to short-offset acquisition, right: extracted traces. Shots are plotted in the same scale.

depend on the angle chosen in equation 3.15. To better discriminate between different angles, the RMSE between true and inverted parameters is calculated for each angle (Figure 3.12a). The gray dashed line shows the maximum available angle for the last interface. The error bound for the WLS method is also added to Figure 3.12 for comparison. It is seen that the error bound of both parameters for the WLS method is lower than the two-trace method in most of the angles. Moreover, the error for the two-trace method is lower for the mid-range angles comparing to low/high-range angles. We also compare the RMSE in the shot domain for different angles of the two-trace and the WLS methods in Figure 3.12b. As can be seen, the error bound in the shot domain for the WLS method is also lower than the two-trace method in most of the angles. The error bound of the Taylor expansion is not included since it is much more larger than the other methods. We also compare the inversion results for β and ρ via the two-trace method ($\gamma' = 25°$) and the other two methods (Figure 3.13). Although both two-trace and WLS methods obtained a good match between inverted and true values, we note that the WLS method has slightly fewer oscillations in both inverted β and ρ parameters. In contrast, the Taylor expansion inverts the parameters with a numerous oscillations. This is consistent with the fact that weighting the second derivative of ξ in equation 3.23 with h^2, makes the method really sensitive to the artifacts in the CIG panel. As expected from parameters cross-talk in the radiation pattern of Figure 3.1, for the fourth event there is a leakage

between the two inverted parameters, meaning that the inversion is not unique. In other words, by adding another parameter like density into account, the ill-posedness of the inverse problem is increased since more degrees of freedom are considered. Because the migrated image mainly comes from the analysis of the reflected waves, and also to reduce the leakage between two parameters, we propose to calculate the impedance perturbation (δI_p) by non-linear re-parameterization of the inverted $\delta\beta/\beta_0$ and $\delta\rho/\rho_0$ as (Bharadwaj et al., 2018)

$$\frac{\delta I_p}{I_{p_0}} = \sqrt{\frac{\frac{\delta\rho}{\rho_0}+1}{\frac{\delta\beta}{\beta_0}+1}} - 1, \tag{3.26}$$

for different methods (Figure 3.13). The alternative could be to use the linear relationship between impedance, velocity and density perturbations. It is seen that the amplitude of the impedance perturbation is much more better estimated compared to other parameters, meaning that the coupling effect between $\delta\beta$ and $\delta\rho$ is compensated.

Although the system of equation 3.25 for the Taylor expansion is well driven, the fact that it is highly sensitive to the maximum surface offset and artifacts in the CIG domain, makes it inapplicable in sense of seismic application. Consequently, we continue the numerical experiments using the two-trace and the WLS methods.

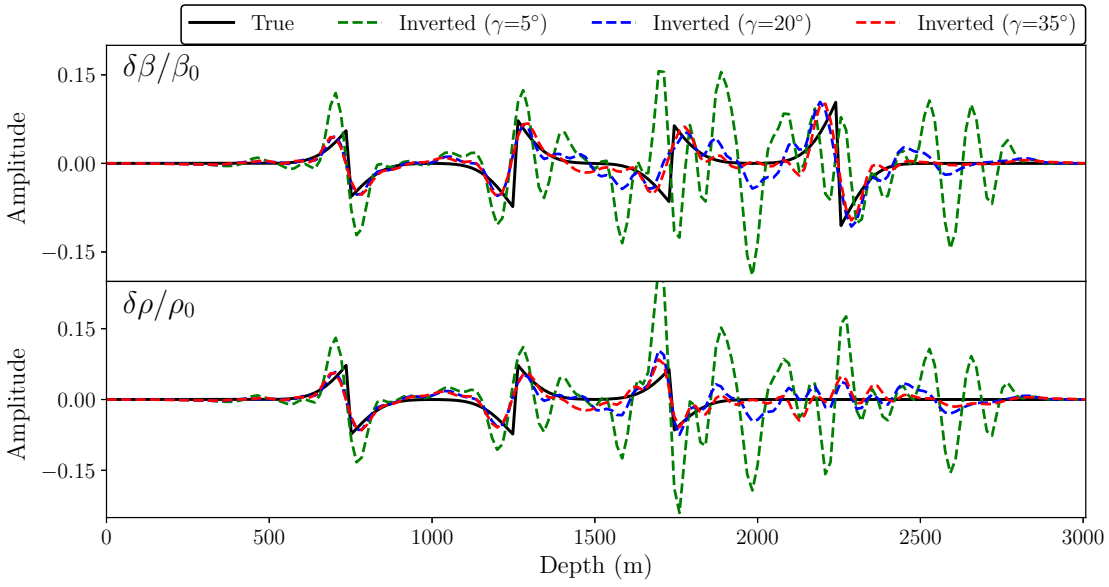

Figure 3.11: Inversion results via two-trace method for different angles.

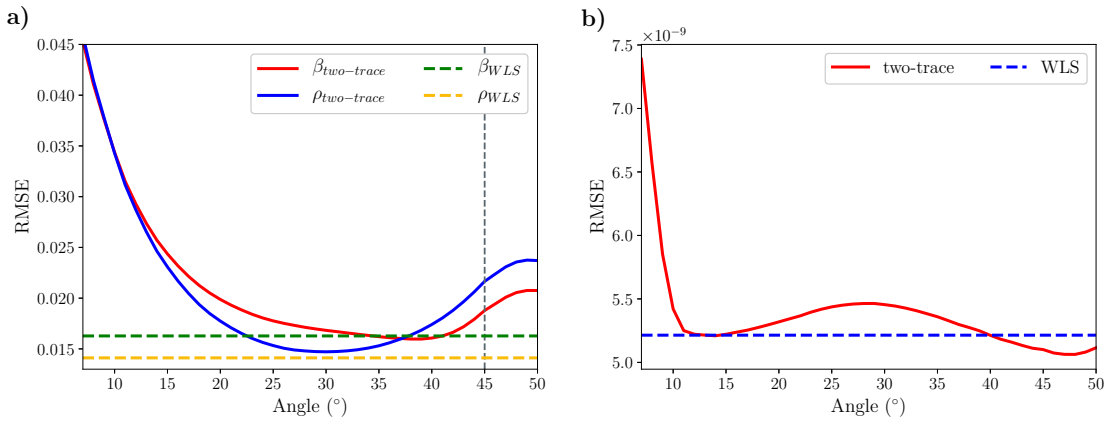

Figure 3.12: RMSE corresponding to a) inverted β and ρ and b) shot gathers for different angles via two-trace method comparing to WLS method. The gray dashed line in (a) corresponds to maximum available angle for the 4th event.

3.5.1.1 Sensitivity to background models

We now investigate the sensitivity of the methods to background models. We use the same observed data as before, but with constant velocity ($V_{p0} = 2400$ m/s) and density ($\rho_o = 2100$ kg/m^3) models for inversion. The calculated CIG and its Radon transform are shown in Figure 3.14. As expected, an upward curvature is observed because of the too-high velocity for each event. As the CIG contains migration smiles rather than focused points, the integration of Radon transform in CIG domain will not produce the flattened AVA response. We check the quality of the inversion by comparing computed data and observed data (Figure 3.15). The colorbar scale for reconstructed data is the same as the one for the observed data. Note that the reconstructed data is modeled via physical inverted parameters in an incorrect model of constant velocity and density. The largest misfit is related to the first event at large offsets. The extracted traces for two-trace method show more misfit comparing to the WLS method in both near and far offsets, specially for the first event. The RMSE between observed and reconstructed shot via each method is also written on each panel showing the superiority of the WLS method. The fact that the remodeled data after inversion nicely matched the observed data in terms of phases and amplitudes, demonstrates that the multiparameter inversion indeed provides an inverse, even in an incorrect model and modeling in physical domain. Furthermore, this provides evidence that our method can be coupled to velocity analysis. More research, beyond the scope of this paper, is required to investigate the coupling effects of multiparameter inversion to velocity analysis. We also compare the true and the inverted parameters obtained via different methods (Figure 3.16). As expected, there is a shift in every inverted parameter related to using high-velocity background model.

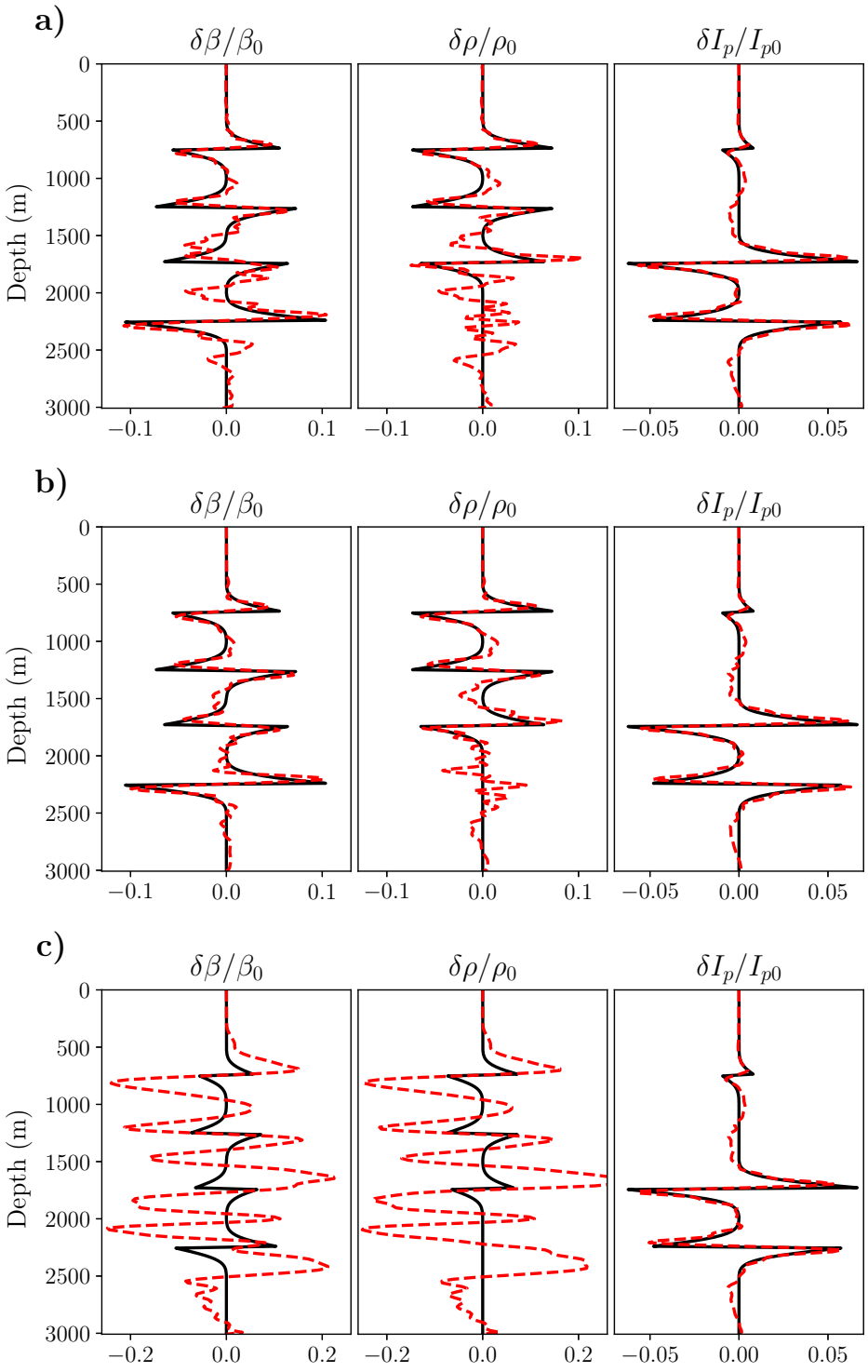

Figure 3.13: Comparison of inversion results for β, ρ and I_p obtained via a) two-trace, b) WLS and c) Taylor expansion corresponding to Figure 3.8. The solid black and dashed red lines correspond to the true and inverted values, respectively.

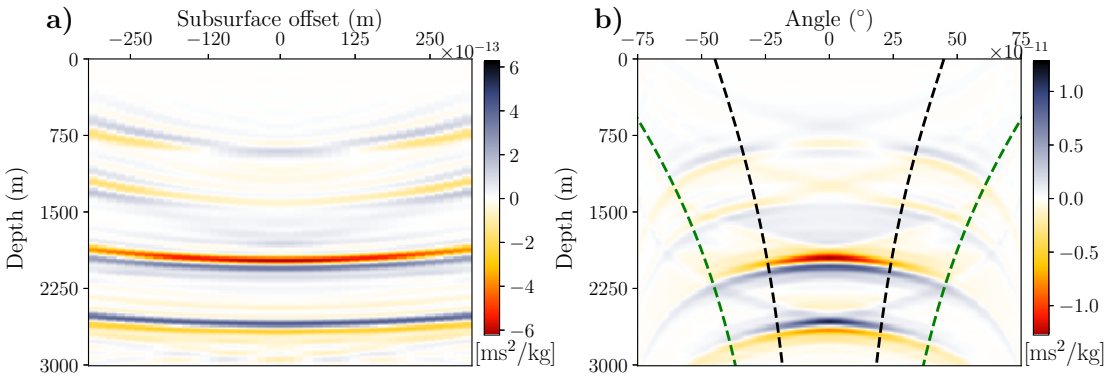

Figure 3.14: Inverted a) ξ and its b) $\mathcal{R}_\xi$ at position $x = 1500$ m corresponding to Figure 3.7. The green and black dashed lines in (b) corresponds to the limit for acquired angles and the used ones in inversion, respectively.

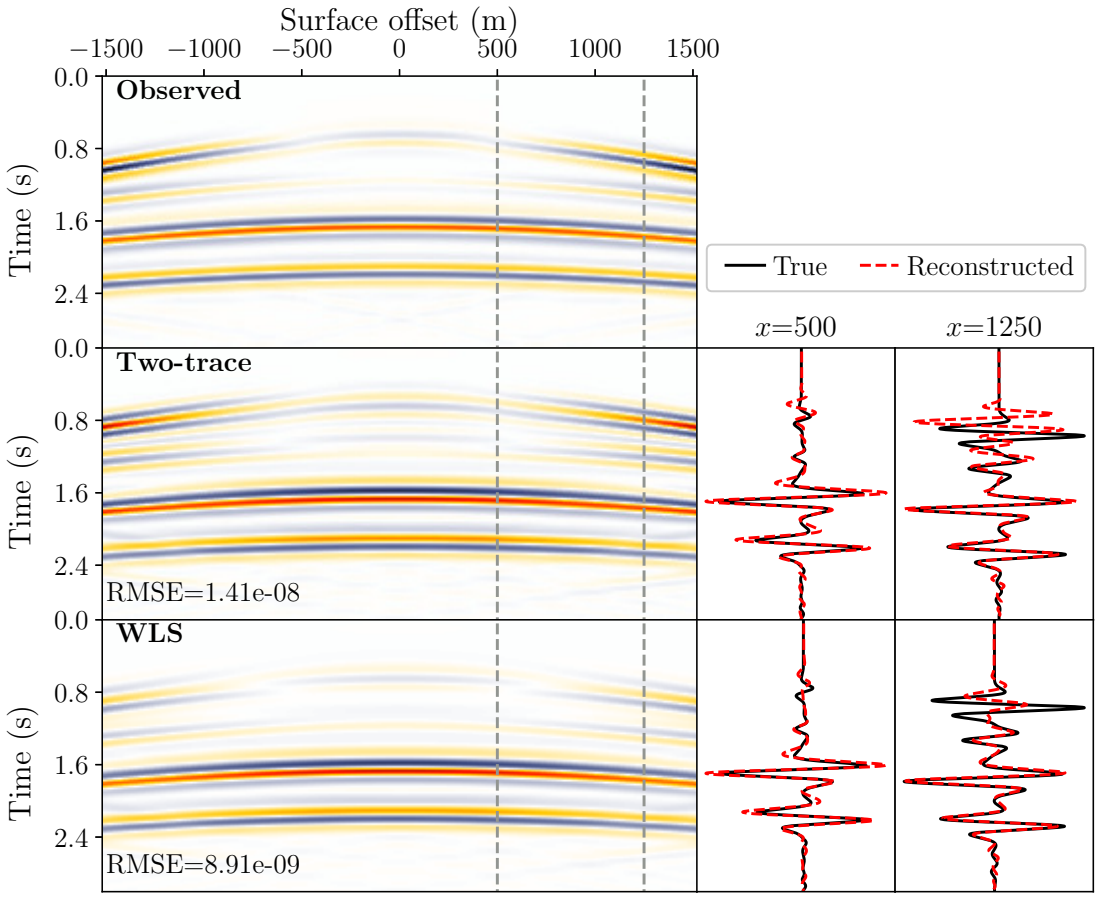

Figure 3.15: Left: observed and reconstructed shots obtained via different methods corresponding to incorrect background models, right: extracted traces. Shots are plotted in the same scale.

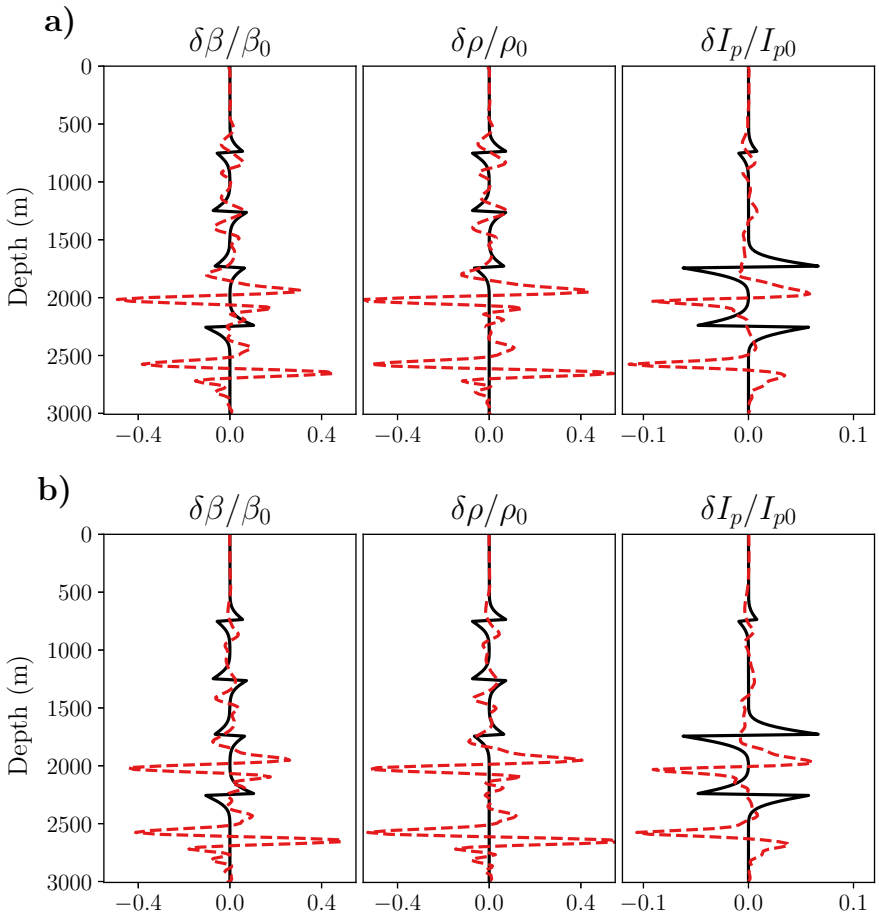

Figure 3.16: Comparison of inversion results for β, ρ and I_p obtained via a) two-trace and b) WLS corresponding to Figure 3.14. The solid black and dashed red lines correspond to the true and inverted values, respectively.

3.5.2 Marmousi2 model

We now consider a more realistic application on the modified Marmousi2 model as a benchmark test. The original Marmousi2 has 17 km width and 3.5 km depth. To reduce the computational cost, we extract the middle part of the Marmousi2 including 10.5 km width and 2.5 km depth, and also to better image the deeper part, we replace the water layer with a higher density layer as in Yang et al. (2016a) and Chen and Sacchi (2018). The exact modified models of velocity (m/s) and density (kg/m^3) of Marmousi2 are shown in Figure 3.17. The model is discretized on a 214×875 grid with 12 m spacing in z and x directions. The synthetic data correspond to a fixed acquisition geometry (stationary-receivers) with a source spacing of 36 m and receiver spacing of 12 m. We use an explosive source, represented by a Ricker wavelet centered at 4.76 Hz (maximum

frequency is 11.9 Hz). The recording time is 3.7 s, with a time interval of 1.32 ms.

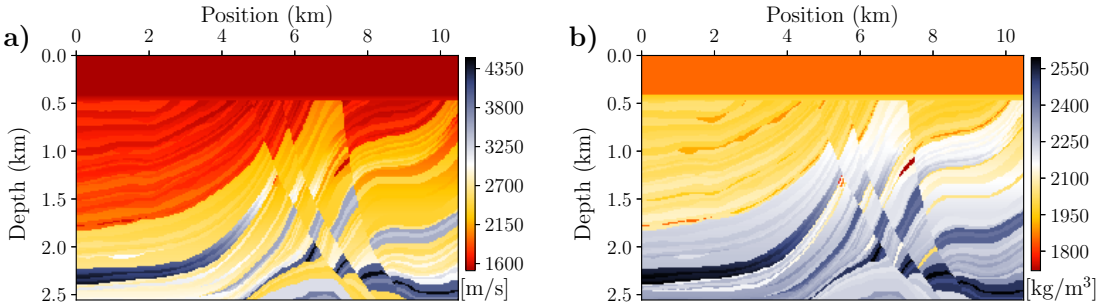

Figure 3.17: The exact a) velocity and b) density corresponding to the Marmousi2 model.

Here, we investigate the inversion for incorrect and correct background models. For our incorrect background models, we use a laterally-homogeneous velocity/density-gradient models (Figure 3.18a and 3.18b), and for our correct background models, we smooth the true velocity/density models with a 2D Gaussian filter of 60 m length in both direction (Figure 3.18c and 3.18d). The inverted ξ obtained via pseudoinverse operator for the incorrect and correct background models for a shot in the middle ($x_s = 5.2$ km) is shown in Figure 3.19a and 3.19c, and their angle-domain responses via the Radon transform are also shown in Figure 3.19b and 3.19d, respectively. The green and black lines in these figures show the application of the mask for $\alpha = 1$ and $\alpha = 0.4$ (Figure 3.18b), and $\alpha = 0.7$ (Figure 3.18d). First, we compare the data fit in the shot domain for the observed shot at position 5.2 km (Figure 3.20). The reconstructed shots are produced by applying the Born modeling on the physical inverted parameters obtained via WLS method in the incorrect background models (Figure 3.20b) as well as the correct background models (Figure 3.20c). The colorbar scale for reconstructed data is the same as the one for the observed data. To evaluate the reliability, the extracted traces for different offsets are also shown in Figure 3.20. It is seen that the shot corresponding to the incorrect background models has a correct phase and a satisfactory amplitude match, especially for the short offsets and shallow part. Looking at the incorrect velocity model (Figure 3.18a) implies that almost in all depths and locations higher velocities are considered as an incorrect model comparing to the correct velocity model (Figure 3.18c). Similar to the results in Figure 3.16, this leads to a shift and defocusing in the inverted parameters. Therefore, the error in the deeper part is related to the limited acquisition. For the correct background models, once more, there is an almost perfect match of the phase and amplitude validating the application of a pseudoinverse Born operator. Obviously, we would also expect small differences because of the geologic complexity of the model, which can be reduced further by LSM iterations.

We also compare the results of inverted physical parameters. The exact perturba-

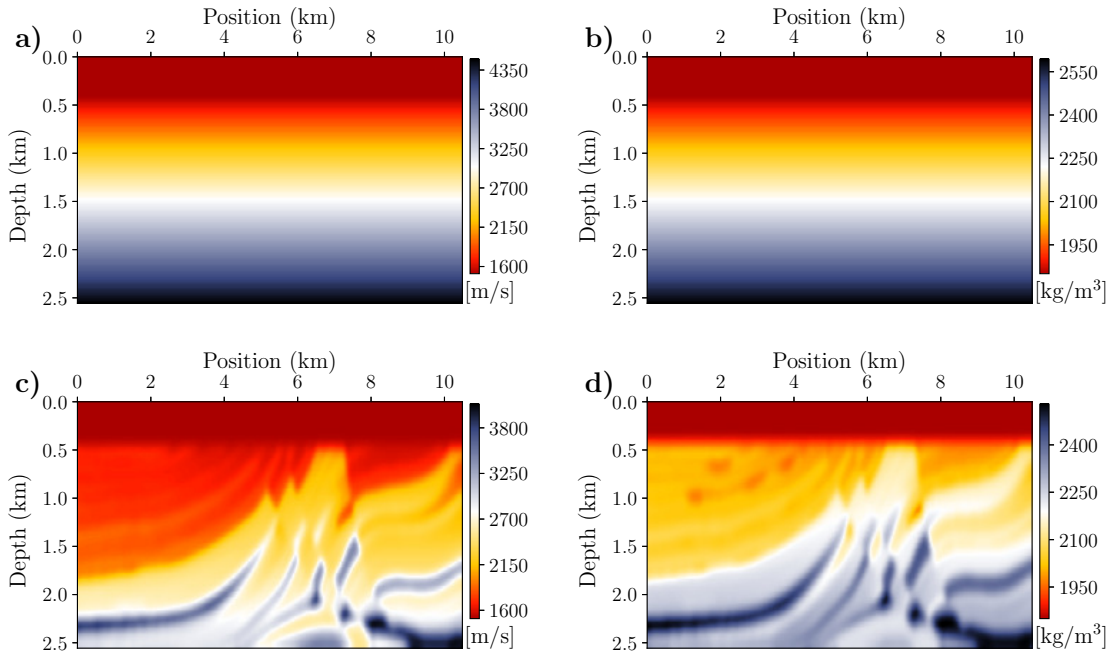

Figure 3.18: The initial incorrect background a) velocity and b) density model and the initial correct background c) velocity and d) density model.

tion models and the inverted perturbation models for incorrect and correct background models are respectively shown in the first, second and third row of Figure 3.21. As mentioned before, it is evident that there is a shift in the inverted parameters for the incorrect background models (Figure 3.21d, 3.21e and 3.21f). The extracted traces at different positions for further comparison between the exact and inverted parameters in correct background models are shown in Figure 3.22. Although the phase information is accurately preserved, the amplitudes of the $\delta\beta$ and $\delta\rho$ have small differences and show a small leakage between two parameters. Nevertheless, the amplitudes of the δI_p is well recovered, as seen in the previous test.

These different numerical tests prove that the inversion formulas developed for the 2D variable density acoustic media are indeed inverse instead of adjoint: almost perfect fit to the observed data even in incorrect background models. Moreover, through numerical experiments, the WLS method showed better and more robust results compared to the two-trace method, in both the image and shot domains. More interestingly, the results for impedance perturbation obtained by combination of other inverted parameters, is in better accordance with the true perturbations and mitigates the ill-posedness of the inverse problem.

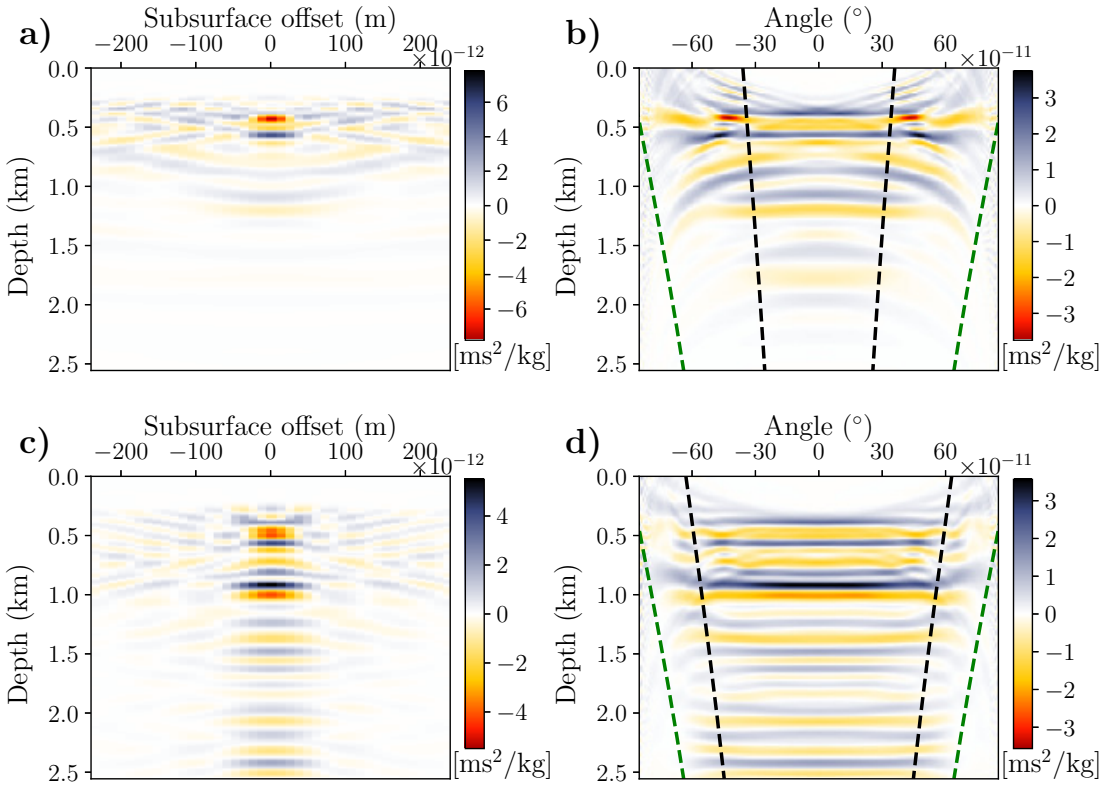

Figure 3.19: Inverted a) ξ and its b) $\mathcal{R}_\xi$ corresponding to incorrect background model and inverted c) ξ and its d) $\mathcal{R}_\xi$ corresponding to correct background model at position $x = 5.2$ km. The green and black dashed lines in (b) and (d) corresponds to the limit for acquired angles and the used ones in inversion, respectively.

3.6 Discussion

We proposed two new approaches to reconstruct the density perturbations by generalization of the two-trace method as well as the Taylor expansion of the Radon transform. The latter was able to reconstruct the observed data only for short-offset acquisitions. This is too limiting in practice, as seismic imaging requires handling long-offset data. Consequently, the proposed method in this paper is the WLS method, which achieved better results for both correct and incorrect background models. An interesting property of such a generalization is that it gives us the necessary flexibility to add the regularization term to the objective function (equation 3.16), leading to a constrained optimization problem. The objective is to mitigate the ill-posedness resulting from parameters cross-talk based on the radiation pattern (Figure 3.1) or to add a priori information.

Recently, Qin and Lambaré (2016) proposed an approach to jointly invert the velocity and density in preserved-amplitude Full Waveform Inversion (FWI). The strategies

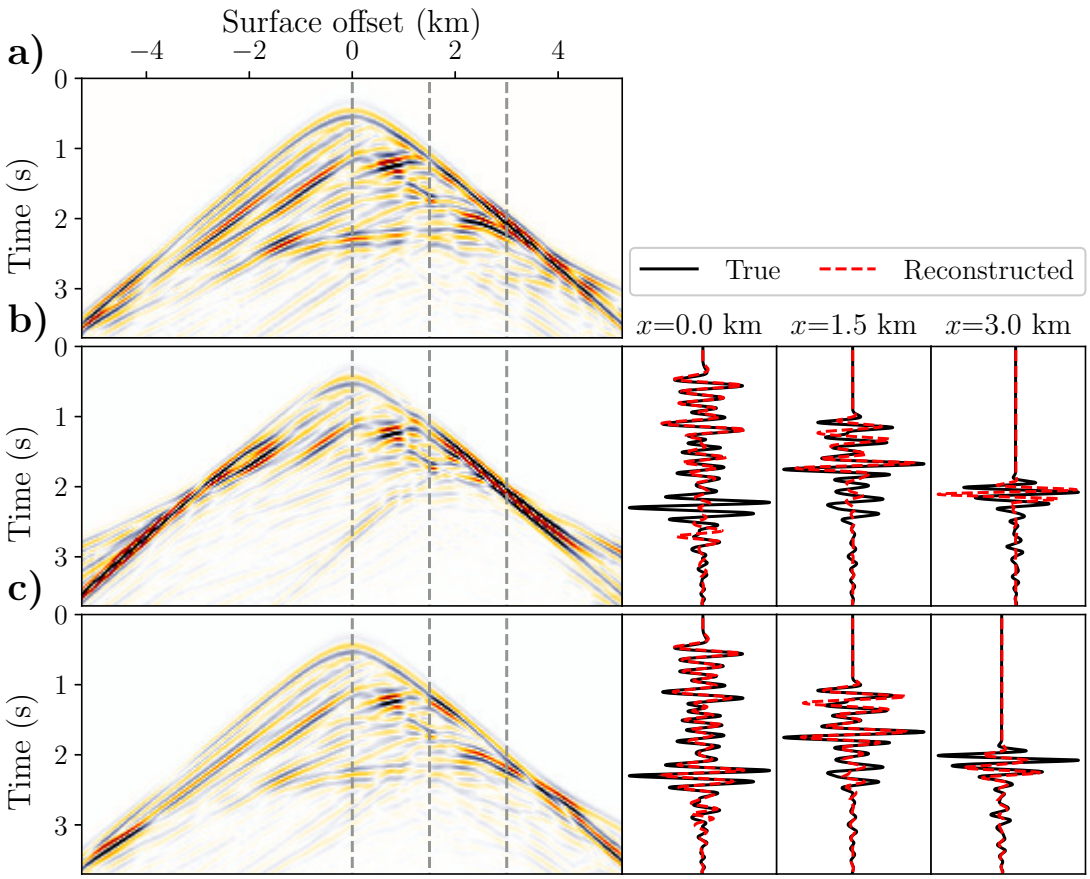

Figure 3.20: Shot gathers of a) observed data and reconstructed data for the shot at position 5.2 km with b) incorrect background models and c) correct background models. The extracted traces at different offsets are shown in left of each shot. Shots are plotted in the same scale.

developed here have two main differences with the one proposed by Qin and Lambaré (2016). First, our inversion is based on the extended domain (subsurface offset), which is not the case in Qin and Lambaré (2016). Second and more importantly, we calculate the diffraction angle (γ) either with the Radon transform or the Taylor expansion, whereas they calculate it by applying a tomographic ray tracing approach. In practice, they apply an iterative process to possibly reduce the approximation errors introduced in the estimation of the angles, which is not necessarily consistent with the wave equation-based approach.

We investigate the effect of the choice of the parameterization in the inversion. We decompose the estimated ξ_β from the simple model (Figure 3.8) to two different parameters based on the corresponding diffraction patterns (Table 3.2) in equations 3.16

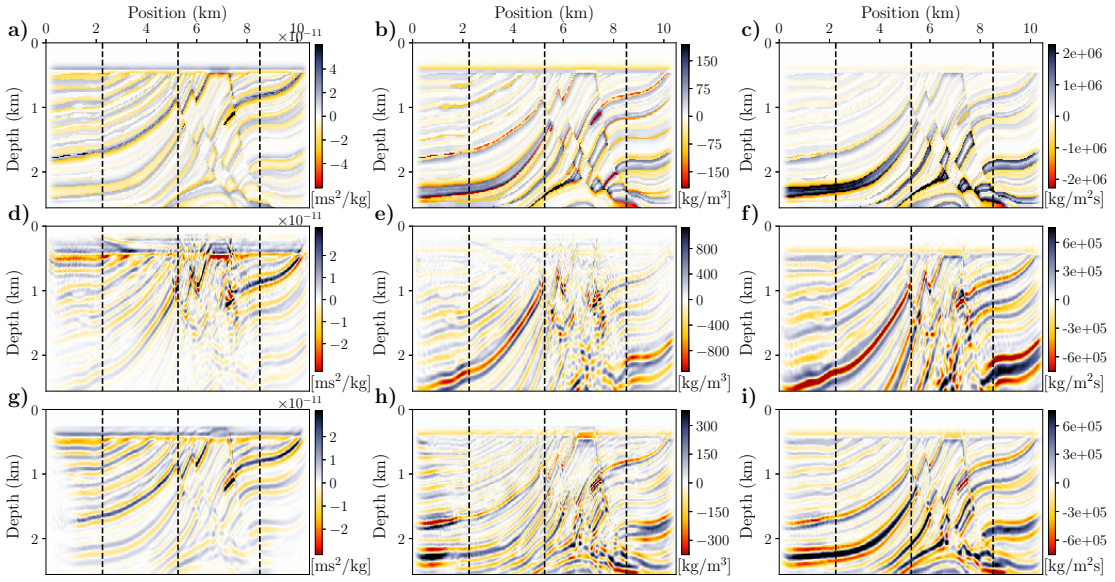

Figure 3.21: The exact perturbation model for a) β, b) ρ and c) I_p, and the inverted perturbation model for d) β, e) ρ and f) I_p corresponding to incorrect background models and g) β, h) ρ and i) I_p corresponding to correct background models. The dashed line corresponds to extracted traces.

and 3.19 (Figure 3.23). The third parameter in each subplot of Figure 3.23 is inferred from the combination of the first two parameters. The good agreement between the results of different parameterizations suggests that the choice of parameterization does not change the final results in our work. The conclusion differs in the case of FWI. In nonlinear imaging approaches such as FWI, the choice of parameterization is not neutral, meaning that the final results depend on parameterization class (Tarantola, 1984; Prieux et al., 2013); whereas migration, a linear operator by definition, does not suffer from this aspect.

Table 3.2: The different diffraction patterns for different parameterization classes.

Parameterization	First parameter	Second parameter
(β, ρ)	-1	$\cos(2\gamma)$
(I_p, ρ)	2	$-2\sin^2(\gamma)$
(V_p, ρ)	2	$2\cos^2(\gamma)$
(V_p, I_p)	$2\sin^2(\gamma)$	$2\cos^2(\gamma)$

In the case of an incorrect background model, a mismatch of data fit in the shallow

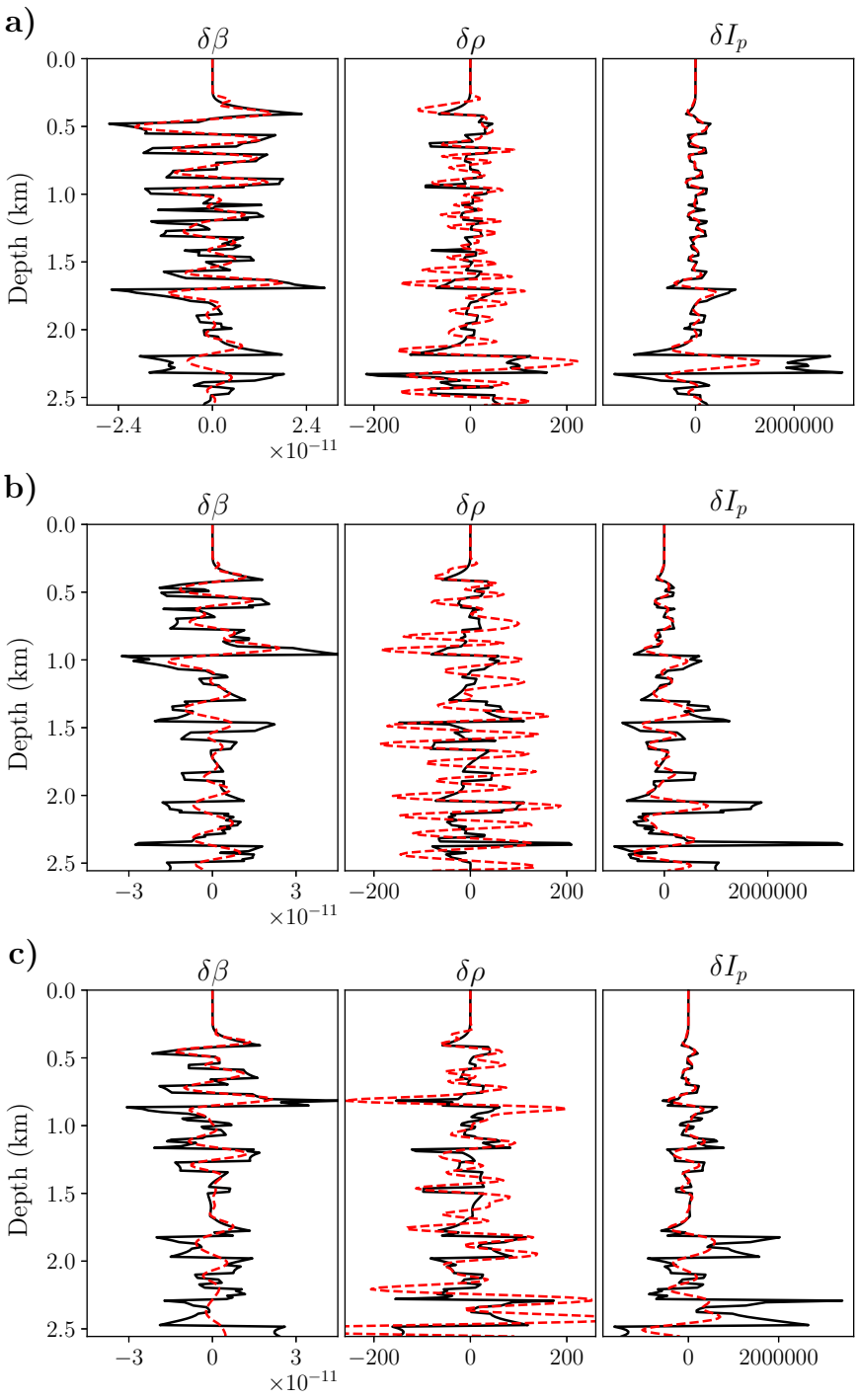

Figure 3.22: The extracted traces from perturbation models (Figure 3.21) for a) 2.2 km, b) 5.2 km and c) 8.5 km. The solid black and dashed red lines correspond to the true and inverted values, respectively.

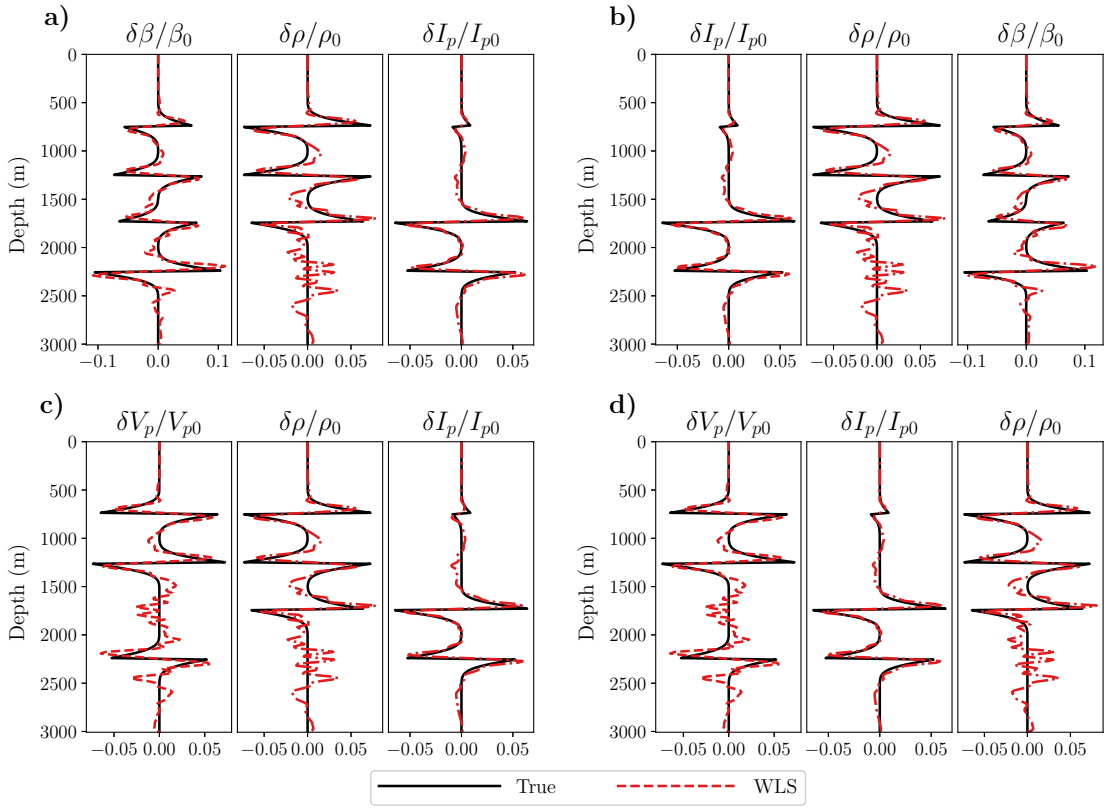

Figure 3.23: The inverted parameters for different parameterization classes as a) (β, ρ), b) (I_p, ρ), c) (V_p, ρ) and d) (V_p, I_p). The third parameter in each panel is inferred by combination of the first two parameters.

part was observed for the simple model (especially for the first event), whereas it was observed in the deeper part for the Marmousi2 model. The reasons for these observations are different. In the case of the simple model, the incorrect background models are constant velocity/density models which are very different from the correct background models, leading to extremely defocused energy in the inverted ξ (Figure 3.14a). Meanwhile, the incorrect background models for Marmousi2 are laterally homogeneous velocity/density-gradient models that are closer to the correct background models. This is also noticeable by comparing the defocused energy of ξ for different models (Figure 3.14a and 3.19c). Accordingly, the shallow part in Marmousi2 will have better data fit as the incorrect background model is close to correct one, which is not the case in simple model. To analyze the effect of the subsurface offset for the mismatch of the data fit in the deeper part of the Marmousi2, we run an additional test by doubling this parameter (Figure 3.24). We note that the mismatch in the deeper part of the reconstructed shot (Figure 3.24c) is more or less the same as before (Figure 3.20b), meaning that this error is not caused by truncation effects in the CIG domain. As already mentioned in the numerical experiments, using higher background velocity model leads to a downward shift in the inverted reflectivity (Figures 3.16 and 3.21). Thus, if the recording depth is not enough to image the reflectivity, a mismatch in the deeper part of the reconstructed shot would indeed be expected, which is the case for the Marmousi2.

In term of implementation, the variable density pseudoinverse Born modeling consists of two operators: the pseudoinverse Born modeling and the forward Radon transform operators. The Radon transform has a long history of applications in seismic processing, for instance, velocity analysis (Thorson and Claerbout, 1985), multiple attenuation (Hampson, 1986), NMO-free stacking (Gholami, 2017) and AVO-preserved processing (Farshad et al., 2018; Gholami and Farshad, 2019b). Based on the path of integration in the Radon transform, many effective methods for rapid evaluation of the traditional Radon transform have been proposed (Hu et al., 2013; Nikitin et al., 2017; Gholami and Sacchi, 2017; Gholami and Zand, 2017; Gholami and Farshad, 2019a). Here, since the size of each slice in the CIG domain is constant, it is possible to explicitly construct the matrix for the Radon transformation. The application of this matrix for the Radon transformation has a computational complexity of $O(N_x N_z N_h N_\gamma)$, whereas wave-equation based operators have computational complexity of $O(N_x N_z N_s N_t)$, given that the numbers of samples for z, h, x, γ, t and sources are N_z, N_h, N_x, N_γ, N_t and N_s. Note that in the extended domain N_h cross-correlations should be also performed. Several techniques such as computing the CIG only at specific image points (Yang and Sava, 2015) or computing the CIG with only a random choice of traces (van Leeuwen et al., 2015) have been proposed to reduce the computational burden of cross-correlations, but not for the propagation. As in practice $N_s \times N_t$ is much larger than $N_h \times N_\gamma$, the main computational burden of the variable density pseudoinverse Born modeling is due to the modeling operators and it remains of the same order as that of the constant density. It is also worth noticing that there is no need to apply the inverse of the Radon trans-

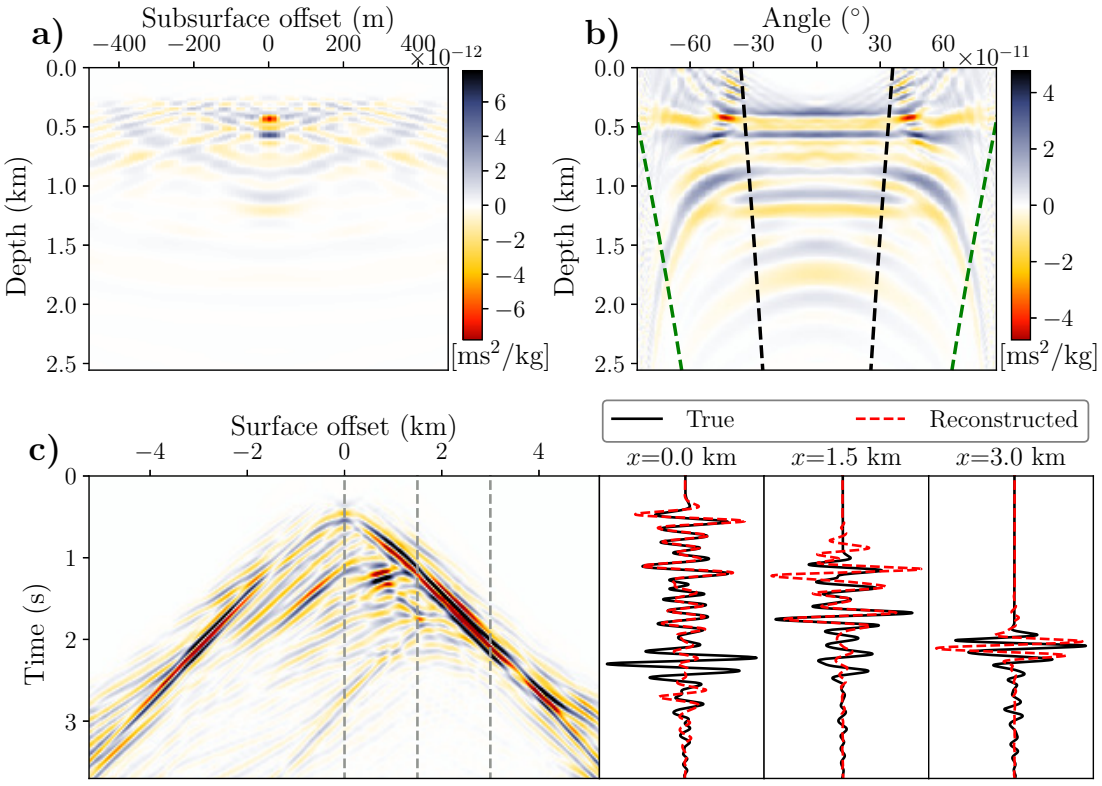

Figure 3.24: Inverted a) ξ and its b) $\mathcal{R}_\xi$ corresponding to incorrect background model and 480 m extension of subsurface offset. c) Reconstructed data for the shot at position 5.2 km with incorrect background models. The extracted traces at different offsets are shown in left on the shot gather. The green and black dashed lines in (b) corresponds to the limit for acquired angles and the used ones in inversion, respectively.

form here. Otherwise, the cost would be different as the minimization should be solved iteratively.

3.7 Conclusions

In this paper, we have proposed an efficient weighted least-squares approach to extend the constant density pseudoinverse Born modeling to variable density acoustic media. This is a generalization of the two-trace method. It is here based on the whole AVA response in the angle domain. We have also proposed another approach based on the Taylor expansion of the Radon transform, which does not require application of the Radon transform. Numerical experiments proves that the latter is not applicable for quantitative seismic imaging because of the noise enhancement at large subsurface offsets, whereas the weighted least-squares method is very promising and provides robust results when compared with the other approaches. We conclude that such a generalization also provides the flexibility to include more constraints in the inversion. Future work will consist of including regularization terms in the least-squares objective function, coupling multiparameter inversion to velocity analysis, and also extending the pseudoinverse Born modeling operator beyond the acoustic case.

3.8 Acknowledgments

The authors would like to thank editor Jeffrey Shragge, assistant editor Antoine Guitton, associate editor Elita Yunyue Li, and five anonymous reviewers for their constructive comments and suggestions

Chapter 4

Extension to variable-density elastic media

Contents

4.1 Abstract **115**
4.2 Introduction **115**
4.3 Methodology **118**
4.3.1 Elastic Born modeling 118
4.3.2 Reverse time migration 119
4.3.3 Elastic pseudoinverse Born operator 120
4.4 Numerical experiments **125**
4.4.1 Simple model 125
4.4.2 Dipping layer model 134
4.4.3 Marmousi2 model 142
4.5 Discussion **151**
4.6 Conclusion **153**
4.7 Acknowledgments **153**
4.8 Appendix I: Gradient derivation for elastic WLS method **154**
4.9 Appendix II: Acoustic WLS method **154**

Résumé

Ce chapitre comprend l'article intitulé « From acoustic to elastic inverse extended constant- to variable-density inverse extended Born modeling », publié dans la revue *Geophysics* (Farshad and Chauris, 2021b).
La migration élastique itérative (« elastic least-squares reverse time migration » en anglais) est la technique la plus avancée pour obtenir des images quantitatives du sous-sol. Les applications demandent beaucoup d'itérations (alternances modélisation / migration). Pour accélérer la convergence, plusieurs pseudoinverses ont été proposés : ils apportent une image quantitative en une seule itération, pour un coup similaire à celui de la « reverse time migration ». Cependant, ces approches supposent un cadre acoustique, ce qui conduit à des prédictions d'amplitudes inexactes et à oublier les effets des ondes S. Pour résoudre ce problème, nous étendons le pseudoinverse du cadre acoustique au cadre élastique pour prendre en compte les amplitudes élastiques des réflexions PP et pour estimer la densité et les impédances P et S. Nous restreignons l'extension au cadre marin, avec l'enregistrement de la pression à la surface. Dans un premier temps, nous remplaçons les fonctions de Green acoustiques par leur version élastique, sans modifier la structure originale du pseudoinverse. Ensuite, nous appliquons une transformée de Radon pour avoir une réponse qui dépend de l'angle. Enfin, nous inversons au sens des moindres carrés les paramètres physiques. Au travers d'expériences numériques, nous illustrons les conséquences de l'approximation acoustique sur des données élastiques, qui conduisent à des inversions biaisées et à l'apparition de réflexions artificielles. Nous démontrons que notre méthode peut inverser simultanément les paramètres élastiques en présence de structures décorrélées, d'un mauvais macro-modèle et de bruit Gaussian.

From acoustic to elastic inverse extended Born modeling: a first insight in the marine environment
Milad Farshad and Hervé Chauris
Geophysics, 2021, *doi: 10.1190/geo2020-0916.1*

4.1 Abstract

Elastic least-squares reverse time migration is the state-of-the-art linear imaging technique to retrieve high-resolution quantitative subsurface images. A successful application requires many migration/modeling cycles. To accelerate the convergence rate, various pseudoinverse Born operators have been proposed, providing quantitative results within a single iteration, while having roughly the same computational cost as reverse time migration. However, these are based on the acoustic approximation, leading to possible inaccurate amplitude predictions as well as the ignorance of S-wave effects. To solve this problem, we extend the pseudoinverse Born operator from acoustic to elastic media to account for the elastic amplitudes of PP reflections and provide an estimate of physical density, P- and S-wave impedance models. We restrict the extension to marine environment, with the recording of pressure waves at the receiver positions. Firstly, we replace the acoustic Green's functions by their elastic version, without modifying the structure of the original pseudoinverse Born operator. We then apply a Radon transform to the results of the first step to calculate the angle-dependent response. Finally, we simultaneously invert for the physical parameters using a weighted least-squares method. Through numerical experiments, we first illustrate the consequences of acoustic approximation on elastic data, leading to inaccurate parameter inversion as well as to artificial reflector inclusion. Then we demonstrate that our method can simultaneously invert for elastic parameters in the presence of complex uncorrelated structures, inaccurate background models, and Gaussian noisy data.

4.2 Introduction

Reverse time migration (RTM) is the most popular imaging method for pre-stack depth migration as it has outstanding imaging capabilities such as accounting for extreme lateral velocity variations without any dip limitations, and the ability to consider the full wavefield (Baysal et al., 1983). By definition, RTM operator is the adjoint of the forward Born modeling operator, accounting only for the kinematics (traveltimes) but not the dynamics (amplitudes). In practice, the migrated image using RTM operator may be deteriorated by different factors such as low-frequency noise, band limitation of the

source wavelet, limited recording aperture, and aliasing due to under-sampled acquisition geometry (Nemeth et al., 1999; Mulder and Plessix, 2004).

To mitigate these problems, least-squares reverse time migration (LSRTM), also known as linearized seismic inversion, was proposed (Nemeth et al., 1999; Dai et al., 2012). LSRTM differs from RTM by its iterative data fitting nature. LSRTM is formulated based on the assumption that an accurate initial velocity model is available. To partially decouple the data fitting in LSRTM from the choice of the velocity model, an offset/time lag extended imaging condition has been proposed (Symes, 2008b; Yang et al., 2019a). Most LSRTM algorithms are based on pure acoustic approximation (constant-density). With the development of multi-component data, LSRTM has been extended to multiparameter imaging problems accounting for more parameter classes, e.g., density (Zhang et al., 2014b; Yang et al., 2016a), anisotropy (Qu et al., 2017), attenuation (Dutta and Schuster, 2014), elasticity (Chen and Sacchi, 2017; Feng and Schuster, 2017; Sun et al., 2018). Compared to acoustic LSRTM, the elastic version estimates more accurate subsurface images as it better honors the wave equation. Moreover, interpreting S-wave impedance along with P-wave impedance and density can intensively help in lithology/fluid discrimination and hydrocarbon detection (Caldwell, 1999; Russell et al., 2003). However, several key challenges remain. For instance, simultaneous inversion of two or more subsurface parameters leads to parameter crosstalk (Virieux and Operto, 2009; Operto et al., 2013; Prieux et al., 2013; Sun et al., 2018). As density is the most difficult parameter to be estimated, most elastic LSRTM methods are built on a constant or known density model. Although density does not play the same key role as P- and S-wave velocities, its contribution to seismic wave amplitudes cannot be ignored (Prieux et al., 2013; Sun et al., 2018; Chen and Sacchi, 2019).

The main drawback of LSRTM, either acoustic or elastic, is its expensive computational cost classically proportional to the number of sources and to the number of iterations. In recent years, researchers have worked on either reducing the computational cost per iteration or accelerating the convergence rate of LSRTM; for instance, by conditioning the linear system using approximate Hessian (Shin et al., 2001; Rickett, 2003; Symes, 2008a; Herrmann et al., 2009; Chen and Sacchi, 2017). In addition, many studies recast RTM in the general framework of linearized seismic inverse problem theory. This led to asymptotic approximations of the inverse operator allowing one to obtain quantitative results in a single iteration (ten Kroode, 2012; Hou and Symes, 2015, 2017; Chauris and Cocher, 2017). Assuming a purely acoustic medium, the pseudoinverse Born operators are built by simple modifications of the subsurface-offset extended RTM operator. Although the derivations are performed under the high-frequency approximation assuming the absence of multiple raypaths, the final formulas do not contain any ray-related computations and appear to be accurate even in the presence of caustics (ten Kroode, 2012; Hou and Symes, 2015; Chauris and Cocher, 2017). The effectiveness of pseudoinverse Born operator as a preconditioner has also been investigated for LSRTM (Hou and Symes, 2016a; Chauris and Cocher, 2018), migration velocity analysis (MVA)

(Chauris and Cocher, 2017; Cocher et al., 2018; Li and Chauris, 2018), and full waveform inversion (FWI) (Hou and Symes, 2016b), all assuming constant-density acoustic media.

Recently, Dafni and Symes (2018) and Farshad and Chauris (2020a) proposed two approaches to extend the applicability of these pseudoinverse Born operators to variable density acoustics. The former inverted two parameters by choosing two traces from the angle-dependent response of the pseudoinverse Born operator, and solving a system of two linear equations (Dafni and Symes, 2018; Zhang et al., 2014b). The latter proposed a generalization of the former by considering all traces in the angle domain using an efficient weighted least-squares (WLS) approach. In both approaches, the Radon transform is used to estimate the scattering angle. Later, Farshad and Chauris (2020c) incorporated sparsity-promoting regularization into WLS-based pseudoinverse Born inversion, which was effective in suppressing migration artifacts and crosstalk noise caused by sparse compressed acquisition. However, in the case where elastic effects are not negligible, these pseudoinverse Born operators may fail in recovering accurate subsurface models with the choice of acoustic instead of elastic physics. Considering an acoustic imaging technique to migrate elastic data sets, the estimated image may suffer from incorrect amplitudes and artificial reflectors associated to S-waves even in marine data sets (Barnes and Charara, 2009; He and Plessix, 2017; Ren et al., 2017; Mora and Wu, 2018; Solano et al., 2020b).

In this article, we extend the applicability of pseudoinverse Born operator from acoustic to elastic media. The proposed method simultaneously inverts for density, P- and S-wave impedance perturbations. The algorithm consists of three main steps: first application of the elastic pseudoinverse Born operator to obtain subsurface-offset common image gathers, then the Radon transform to convert the subsurface offset to the imaging angle, and finally simultaneously estimation of the elastic parameters from the angle-dependent response using the WLS strategy. In this study we consider marine towed-streamer data, i.e., sources and receivers deployed within the water layer. In this context, we only consider the PP-PP and PS-SP wave modes in the inversion process. The new proposed elastic pseudoinverse Born operator is evaluated here within a single iteration, which is a step towards accelerating the convergence rate of the elastic LSRTM.

The rest of this paper is organized as follows. In the "Methodology" section, we first review the theory of elastic Born modeling and its adjoint. Then we present the methodology of the proposed elastic pseudoinverse Born inversion. In the "Numerical experiments" section, we first analyze the impact of ignoring S-wave impedance using a simple model. Hereafter, we use two examples "dipping layer model" and "Marmousi2 model" to validate the effectiveness and accuracy of this new method in the presence of parameter crosstalk, inaccurate migration background models, noise, and complex structures. We discuss the applications, in particular the role of the parameterization under the linear imaging assumption.

4.3 Methodology

4.3.1 Elastic Born modeling

We consider an elastic Earth model parameterized with 3 parameters at each point, namely $\beta_p = 1/\rho(V_p^2 - V_s^2)$, $\beta_s = 1/\rho V_s^2$ and ρ, where ρ, V_p and V_s denote density, P- and S-wave velocities, respectively. Assuming an isotropic Earth, a first-order hyperbolic elastodynamic system for 2-D $P-SV$ waves in the frequency domain can be described by the following differential system (Brossier et al., 2008):

$$\begin{aligned}
-i\omega\rho V_x &= \frac{\partial(T_1+T_2)}{\partial x} + \frac{\partial T_3}{\partial z} + \rho F_x,\\
-i\omega\rho V_z &= \frac{\partial T_3}{\partial x} + \frac{\partial(T_1-T_2)}{\partial z} + \rho F_z,\\
-i\omega\beta_p T_1 &= \frac{\partial V_x}{\partial x} + \frac{\partial V_z}{\partial z} - i\omega\beta_p T_1^0,\\
-i\omega\beta_s T_2 &= \frac{\partial V_x}{\partial x} - \frac{\partial V_z}{\partial z} - i\omega\beta_s T_2^0,\\
-i\omega\beta_s T_3 &= \frac{\partial V_z}{\partial x} + \frac{\partial V_x}{\partial z} - i\omega\beta_s T_3^0,
\end{aligned} \tag{4.1}$$

where $T = (T_1, T_2, T_3) = [(\sigma_{xx}+\sigma_{zz})/2, (\sigma_{xx}-\sigma_{zz})/2, \sigma_{xz}]$, and ω is the angular frequency. Source terms are either applied stress vectors (T_1^0, T_3^0, T_3^0) or point forces (F_x, F_z) as introduced in system of equations 4.1. This pseudo-conservative formulation is convenient for seismic imaging techniques as all the physical parameters are on the left-hand side. This system of equations can be written in a matrix form as follows

$$(-i\omega)\begin{pmatrix}\rho\mathbf{I} & \mathbf{0}\\ \mathbf{0} & \mathbf{B}\end{pmatrix}\mathbf{u} - \begin{pmatrix}\mathbf{0} & \mathbf{D}\\ \mathbf{C} & \mathbf{0}\end{pmatrix}\mathbf{u} = \mathbf{f}, \tag{4.2}$$

where

$$\begin{aligned}
\mathbf{u} = \begin{pmatrix}\mathbf{v}\\ \mathbf{t}\end{pmatrix}, \quad \mathbf{v} = \begin{pmatrix}V_x\\ V_z\end{pmatrix}, \quad \mathbf{f} = \begin{pmatrix}\mathbf{f_F}\\ \mathbf{f_T}\end{pmatrix}, \quad \mathbf{f_F} = \begin{pmatrix}\rho F_x\\ \rho F_z\end{pmatrix},\\
\mathbf{B} = \begin{pmatrix}\beta_p & 0 & 0\\ 0 & \beta_s & 0\\ 0 & 0 & \beta_s\end{pmatrix}, \quad \mathbf{t} = \begin{pmatrix}T_1\\ T_2\\ T_3\end{pmatrix},
\end{aligned} \tag{4.3}$$

$$\mathbf{f_T} = \begin{pmatrix}-i\omega\beta_p T_1^0\\ -i\omega\beta_s T_2^0\\ -i\omega\beta_s T_3^0\end{pmatrix}, \quad \mathbf{C} = \begin{pmatrix}\frac{\partial}{\partial x} & \frac{\partial}{\partial z}\\ \frac{\partial}{\partial x} & -\frac{\partial}{\partial z}\\ \frac{\partial}{\partial z} & \frac{\partial}{\partial x}\end{pmatrix}, \quad \mathbf{D} = \begin{pmatrix}\frac{\partial}{\partial x} & \frac{\partial}{\partial z}\\ \frac{\partial}{\partial x} & -\frac{\partial}{\partial z}\\ \frac{\partial}{\partial z} & \frac{\partial}{\partial x}\end{pmatrix}^*,$$

with $\mathbf{I}$ and $*$ denoting the identity matrix and transpose operator, respectively. Note that the dependence on spatial coordinates $\mathbf{x} = (x, z)$ and angular frequency ω of variables are dropped to make the notations compact but it is clear that $\mathbf{t} = \mathbf{t}(\mathbf{x}, \omega)$, $\rho = \rho(\mathbf{x})$, etc. Under the Born approximation, the model parameters are split into a smooth part (e.g. ρ_0) and a singular or perturbation part (e.g. $\delta\rho$):

$$\begin{aligned} \rho &= \rho_0 + \delta\rho, \\ \beta_p &= \beta_{p_0} + \delta\beta_p, \\ \beta_s &= \beta_{s_0} + \delta\beta_s. \end{aligned} \tag{4.4}$$

The same decomposition holds for the wavefield

$$\mathbf{u} = \mathbf{u}_0 + \delta\mathbf{u}. \tag{4.5}$$

Substituting equations 4.4 and 4.5 into equation 4.2, subtracting equation 4.2, and dropping higher order terms leads to the Born approximation

$$(-i\omega)\begin{pmatrix} \rho_0\mathbf{I} & \mathbf{0} \\ \mathbf{0} & \mathbf{B_0} \end{pmatrix}\delta\mathbf{u} - \begin{pmatrix} \mathbf{0} & \mathbf{D} \\ \mathbf{C} & \mathbf{0} \end{pmatrix}\delta\mathbf{u} = (i\omega)\begin{pmatrix} \delta\rho\mathbf{I} & \mathbf{0} \\ \mathbf{0} & \delta\mathbf{B} \end{pmatrix}\mathbf{u}_0, \tag{4.6}$$

where $\delta\mathbf{u}$ is the scattered wavefield and $\mathbf{u}_0$ denotes the incident wavefield. The data $\delta\mathbf{d}$ is obtained by sampling the scattered wavefield $\delta\mathbf{u}$. By defining $\delta\mathbf{m}$ as the vector of perturbed model parameters, the scattered wavefield $\delta\mathbf{u}$ associated to $\delta\mathbf{m}$ can be expressed in terms of the causal Green's function in a given background model. The standard elastic Born modeling operator $\mathcal{L}$ can be expressed in abstract form as (Symes, 2008b):

$$\delta\mathbf{d} = \mathcal{L}\delta\mathbf{m}. \tag{4.7}$$

The Born operator $\mathcal{L}$ can be generalized by introducing an extra dimension in the perturbation model, which could be a time and/or horizontal/vertical space shift (Sava and Fomel, 2006; Symes, 2008b). Here we consider extension by non-physical horizontal subsurface offset $\mathbf{h} = (h, 0)$, which is essentially the offset between the sunken source and sunken receiver in the concept of survey-sinking imaging condition (Claerbout, 1985; Symes, 2008b). The extended Born modeling operator computes the traces for offset $\mathbf{h}$ by using the extended $\delta\mathbf{m}(\mathbf{x}, \mathbf{h})$. The extension is essential for the derivation of the inverse Born operator, as the model size should precisely match the data size (Beylkin, 1985; ten Kroode, 2012; Chauris and Cocher, 2017). Hereafter, we refer to $\mathcal{L}$ and $\delta\mathbf{m}$ as extended Born modeling operator and extended model parameters, respectively.

4.3.2 Reverse time migration

By definition, reverse time migration is the adjoint of the Born modeling operator. An extended migration image $\delta\mathbf{m}(\mathbf{x}, \mathbf{h})$ can be estimated by applying the adjoint of the

extended forward Born modeling operator to the observed data (Tarantola, 1984; Claerbout, 1992):

$$\delta\mathbf{m}^{mig} = \mathcal{L}^*\delta\mathbf{d}. \tag{4.8}$$

Although the adjoint operator $\mathcal{L}^*$ deals with kinematics accurately, it cannot retrieve the amplitudes of the model perturbations correctly. In this case, the migrated image using adjoint can be considered as a blurred version of the subsurface model perturbations (Nemeth et al., 1999):

$$\delta\mathbf{m}^{mig} = \mathbf{H}\delta\mathbf{m}, \tag{4.9}$$

where $\mathbf{H} = \mathcal{L}^*\mathcal{L}$ denotes the Hessian matrix. In terms of computational cost and memory requirements, explicit probing of the structure of the inverse Hessian matrix is prohibitively expensive. Traditionally, the effect of Hessian is taken into account using an iterative approach, namely LSRTM (Nemeth et al., 1999; Dai et al., 2012). However, this approach is also computationally expensive as the cost of one iteration of LSRTM is nearly twice that of RTM, which further increases linearly with the number of iterations.

In the acoustic context, an interesting cost-effective alternative is to replace the adjoint RTM operator with its pseudoinverse expression. Remarkably, the pseudoinverse Born operator only consists of additional weights in RTM operator with relatively low implementation costs (ten Kroode, 2012; Hou and Symes, 2015; Chauris and Cocher, 2017). We investigate here the extension to elasticity.

4.3.3 Elastic pseudoinverse Born operator

Originally, the wave-equation-based pseudoinverse Born operator was introduced for pure acoustic media (ten Kroode, 2012; Hou and Symes, 2015; Chauris and Cocher, 2017). Later, Dafni and Symes (2018) and Farshad and Chauris (2020a) extended these pseudoinverse Born operators to variable-density acoustics. Their main motivation was: the effect of density variations is observed as defocused energy in estimated common image gathers (CIGs) even in presence of correct background models (Farshad and Chauris, 2020a,b). The pseudoinverse Born operator indeed compensates for the amplitude variation with offset (AVO) effects in the data domain by smearing the energy in the CIG domain. The extension from constant- to variable-density acoustic pseudoinverse Born operator was obtained by first incorporating density in the reference Green's functions, then getting the imaging angle through the Radon transform, and finally decomposing into physical velocity and density perturbations. We follow here the same strategy to extend the applicability of the acoustic pseudoinverse Born operator to elastic media. We split our method into three main steps as follows.

Step 1: Elastic Born inversion

We start by following the work of Chauris and Cocher (2017) and Farshad and Chauris (2020a). The pseudoinverse Born operator $\mathcal{L}^{\dagger}$ considering constant-density and constant-S-wave velocity elastic media can be written as

$$\begin{aligned}\delta m &= \mathcal{L}^{\dagger}(\delta d^{obs})(\mathbf{x},\mathbf{h}) \\ &\simeq 32\frac{\beta_{p_0}}{\rho_0^3}\partial_z \int \mathrm{d}\mathbf{x_s}\,\mathrm{d}\mathbf{x_r}\,\mathrm{d}\omega\,\frac{\Omega^{\dagger}(\omega)}{(i\omega)}\partial_{s_z}G_0^*(\mathbf{x_s},\mathbf{x}-\mathbf{h},\omega)\delta d^{obs}(\mathbf{x_s},\mathbf{x_r},\omega)\partial_{r_z}G_0^*(\mathbf{x}+\mathbf{h},\mathbf{x_r},\omega),\end{aligned} \tag{4.10}$$

where $\delta m(\mathbf{x},\mathbf{h})$ is the extended reflectivity, * denotes the complex conjugate and $\Omega^{\dagger}$ is the inverse of the seismic wavelet. The derivation of equation 4.10 is performed under the high-frequency approximation such that the reconstructed data $d^{rec} = \mathcal{L}\mathcal{L}^{\dagger}d^{obs}$ equals d^{obs}. The vertical derivatives with respect to source and receiver positions to the Green's functions have dynamic effects: their combination with the Beylkin determinant (Jacobian of transformation between subsurface parameters and acquisition surface coordinates) arising in linear phase approximation of the normal operator $\mathcal{L}^*\mathcal{L}$ cancel the geometrical spreading present in the propagating fields. This formula is valid even in the presence of incorrect background models and strictly in the absence of turning waves. It is derived assuming infinite acquisition geometries and the absence of multiple raypaths, though in practice it can resolve subsurface models accurately within the migration aperture even in the presence of caustics (ten Kroode, 2012; Hou and Symes, 2015; Chauris and Cocher, 2017).

The terms $G_0^*(\mathbf{x_s},\mathbf{x}-\mathbf{h},\omega)$ and $G_0^*(\mathbf{x}+\mathbf{h},\mathbf{x_r},\omega)$ are the elastic Green's functions in the background models β_{p_0}, β_{s_0} and ρ_0, from the source position $\mathbf{x_s}$ to imaging point $\mathbf{x}-\mathbf{h}$ and from imaging point $\mathbf{x}+\mathbf{h}$ to the receiver position $\mathbf{x_r}$, respectively. Equation 4.10 is not exactly the same as the one proposed by Hou and Symes (2015), Chauris and Cocher (2017), and Farshad and Chauris (2020a). The first difference is in the definition of β_p, here being $1/\rho(V_p^2 - V_s^2)$, providing the flexibility to include V_{s0} in the inversion process. The second and more important aspect is that the Green's functions here are calculated using the elastic Born modeling engine (equation 4.6) to accurately simulate the kinematics and amplitudes of the P- and S-waves.

As mentioned, the pseudoinverse Born operator was originally proposed for pure acoustic media. The progressing wave approximation (Courant and Hilbert, 1962) was based on acoustic Green's function (Hou and Symes, 2015). Here, in order to not violate this assumption, we consider streamer dataset, limiting the observed data δd^{obs} to the pressure wavefield T_1 (equation 4.1). In practice, we solve the system of 6 equations in elastic wave propagation to simulate $\delta\mathbf{u} = \begin{bmatrix}\delta V_x & \delta V_z & \delta T_1 & \delta T_2 & \delta T_3\end{bmatrix}^*$ (equation 4.6) but the observed data are only in the δT_1 component (pressure wavefield in a marine environment). Acoustic inversion has been favored mainly for their lower computational

cost, particularly in the case of marine towed-streamer data where the waves must convert to pressure wavefield in the water layer and elastic effects in the recorded data are generally assumed to be small. However, researchers have demonstrated that in presence of sharp-contrast interfaces, such as boundaries between sediments and salt or carbonates, seismic data is affected by significant mode conversions between P- and S-waves. In this case, elastic inversion is highly recommended as it gives results that are more reliable images of the subsurface (Thiel et al., 2019; Marjanović et al., 2019; Dickens et al., 2020; Solano et al., 2020a).

Step 2: Transforming the subsurface offset to scattering angle

The inverted $\delta m(\mathbf{x}, \mathbf{h})$ in the extended domain (equation 4.10) can be decomposed into three physical parameters based on the diffraction pattern of the specific parameterization (Zhang et al., 2014b; Dafni and Symes, 2018; Farshad and Chauris, 2020a). Here, as the concept is seismic migration resolving only the short-wavelength part of the model (perturbations), we choose density (ρ), P-wave (I_p) and S-wave impedances (I_s) for subsurface parameterization (Tarantola, 1986). However, we have the flexibility to easily change the parameterization class, as explained in the "Discussion" section. By considering the physical model parameters $\delta\mathbf{p}(\mathbf{x})$ as

$$\delta\mathbf{p}(\mathbf{x}) = \begin{bmatrix} \zeta_{I_p}(\mathbf{x}) & \zeta_{I_s}(\mathbf{x}) & \zeta_{\rho}(\mathbf{x}) \end{bmatrix}^*, \tag{4.11}$$

where $\zeta_a = \delta a / a_0$, and the corresponding diffraction pattern for P-P wave mode (Figure 4.1) (Forgues and Lambaré, 1997) as

$$\mathfrak{D}(\theta) = \begin{bmatrix} 2 & -4C_0 \sin^2(2\theta) & 2C_0 \sin^2(2\theta) - 2\sin^2(\theta) \end{bmatrix}, \tag{4.12}$$

where $C_0 = (V_{s_0}/V_{p_0})^2$, the decomposing relationship between extended and physical parameters becomes

$$\frac{\delta m}{\beta_{p_0}}(\mathbf{x}, \mathbf{h}) \cong \mathfrak{D}(\theta)\delta\mathbf{p}(\mathbf{x})\delta(\mathbf{h}), \tag{4.13}$$

where $\delta()$ is the Dirac delta function and θ denotes the scattering angle. The right side of equation 4.13 is obtained under the ray + Born approximation (Forgues and Lambaré, 1997; Farshad and Chauris, 2020a). The terms related to θ (equation 4.12) are not simple multiplications but act as pseudo-differential operators (Dafni and Symes, 2018). Since the left side equation 4.13 is subsurface-offset dependent, and the right side angle-dependent, a conversion of subsurface offset $\mathbf{h} = (h, 0)$ to scattering angle θ is required. This conversion can be done by implementing the Radon transform (slant-stack integral) on extended δm (Sava and Fomel, 2003). The application of 2D Radon

transform on the left side of equation 4.13 yields

$$\mathcal{R}_{\delta m}(x, z, \theta) = \frac{1}{\beta_{p0}} \int \mathrm{d}h \, \delta m(x, z + h \tan\theta, h), \tag{4.14}$$

where $\mathcal{R}_{\delta m}(\mathbf{x}, \theta)$ denotes the corresponding angle-domain CIG (Radon transform). Note that the algorithm only requires the application of the forward Radon transform.

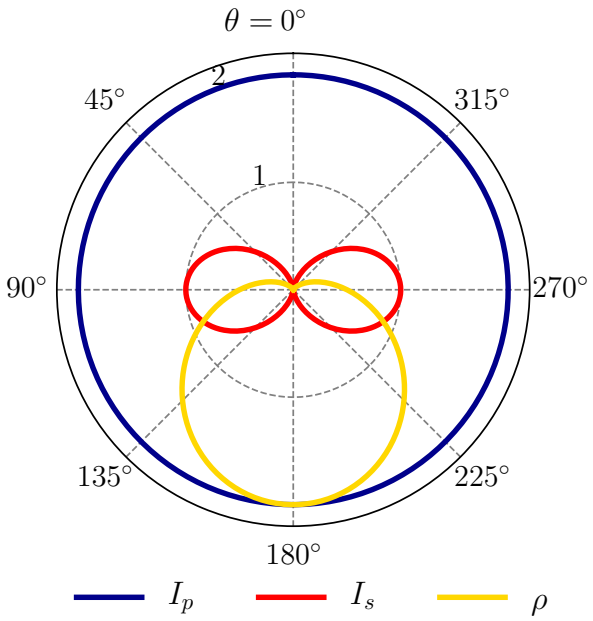

Figure 4.1: P- to P-wave elastic amplitude diffraction pattern.

Step 3: Weighted least-squares method

Following the work of Farshad and Chauris (2020a), we now define a weighted least-squares (WLS) objective function to simultaneously invert for different parameters as

$$J_{\zeta_{I_p},\zeta_{I_s},\zeta_\rho} = \tag{4.15}$$
$$\frac{1}{2} \left\| \left(2\zeta_{I_p}(\mathbf{x}) - 4C_0 \sin^2(2\theta)\zeta_{I_s}(\mathbf{x}) + 2(C_0 \sin^2(2\theta) - \sin^2(\theta))\zeta_\rho(\mathbf{x}) \right) - \mathcal{R}_{\delta m}(\mathbf{x}, \theta) \right\|_W^2,$$

where $W(\mathbf{x}, \theta)$ is a weighting mask defined as:

$$W(\mathbf{x}, \theta) = \begin{cases} 1, & \text{if } |\theta| \leq \alpha \tan^{-1}(\frac{x_{max}}{z}) \\ 0, & \text{otherwise} \end{cases}, \tag{4.16}$$

where x_{max} denotes the maximum surface offset. The role of this weighting mask, controlled by parameter α, is to mainly include specular energy and remove artifacts at large angles due to finite sampling of the surface offset (Sava and Fomel, 2003). In practice, this is a conservative approach that does not require a precise selection of this parameter and can be chosen visually or based on a homogeneous model assumption.

The optimal ζ_{I_p}, ζ_{I_s} and ζ_ρ can be obtained by computing the gradient of the objective function J (equation 4.15) with respect to the model parameters (see Appendix 4.8 for more details). Abstractly, the elastic WLS method can be written in matrix and functional form:

$$\mathcal{A}_W\{\mathbf{M}\}\delta\mathbf{p} = \mathcal{A}\{\mathbf{\Phi}\}, \tag{4.17}$$

where

$$\mathbf{M} = \tag{4.18}$$
$$\begin{bmatrix} 4 & -8C_0\sin^2(2\theta) & 4(C_0\sin^2(2\theta)-\sin^2(\theta)) \\ -8C_0\sin^2(2\theta) & 16C_0^2\sin^4(2\theta) & -8(C_0^2\sin^4(2\theta)-C_0\sin^2(2\theta)\sin^2(\theta)) \\ 4(C_0\sin^2(2\theta)-\sin^2(\theta)) & -8(C_0^2\sin^4(2\theta)-C_0\sin^2(2\theta)\sin^2(\theta)) & 4(C_0\sin^2(2\theta)-\sin^2(\theta))^2 \end{bmatrix},$$

$$\mathbf{\Phi} = \begin{bmatrix} 2\mathcal{R}_{\delta m}(\mathbf{x},\theta) \\ -4C_0\sin^2(2\theta)\mathcal{R}_{\delta m}(\mathbf{x},\theta) \\ 2(C_0\sin^2(2\theta)-\sin^2(\theta))\mathcal{R}_{\delta m}(\mathbf{x},\theta) \end{bmatrix}, \tag{4.19}$$

and $\mathcal{A}_W$ is an integral operator with kernel W acting on input $\mathbf{f}$ as:

$$\mathcal{A}_W\{\mathbf{f}\} = \int \mathrm{d}\theta\, W(\mathbf{x},\theta)\mathbf{f}(\mathbf{x},\theta). \tag{4.20}$$

By definition, the elastic WLS method (equation 4.17) leads to a separable trivariate curve-fitting problem, meaning that the optimization can be applied for each spatial coordinate separately. Note that the elastic WLS is obtained by a simple modification of the acoustic WLS, with three unknowns instead of two (Farshad and Chauris, 2020a). We also recall the acoustic WLS method for parameter class (ζ_{I_p},ζ_ρ) in Appendix 4.9, to compare the results of acoustic and elastic Born inversion in presence of S-wave velocity effect in the first numerical experiment.

4.4 Numerical experiments

In this section, we assess the viability of the elastic pseudoinverse Born inversion through carefully designed numerical experiments. For all examples, the modeling engine for the solution of the 2-D elastic wave equation (equations 4.2 and 4.6) is staggered-grid finite-difference with eighth-order accuracy in space and second-order accuracy in time implemented with perfectly matched layers (PML) absorbing boundary conditions (Virieux, 1986). For all numerical experiments, we consider a fixed acquisition geometry (stationary receivers), and Ricker wavelet as the source signature emitting only P-waves (marine environment). The details regarding each model discretization and their acquisition geometry (e.g. maximum frequency, source/receiver spacing, etc) are outlined in Table 4.1. To get more insight into the resolution of the inverted images, we calculate the normalized correlation coefficient (NCC) defined as

$$\mathrm{NCC}(\%) = \frac{\sum_{\mathbf{x}} \zeta_{true}(\mathbf{x})\zeta_{inv}(\mathbf{x})}{\sqrt{\sum_{\mathbf{x}} \zeta^2_{true}(\mathbf{x})}\sqrt{\sum_{\mathbf{x}} \zeta^2_{inv}(\mathbf{x})}} \times 100, \tag{4.21}$$

where ζ_{true} and ζ_{inv} are the true and inverted model parameters, respectively. The NCC score is normalized between 0 and 100; the higher the score, the higher the correlation is. We also evaluate the quality of the reconstructed data by calculating the root-mean-squared (rms) error defined as

$$\text{rms error} = \sqrt{\sum_{i=1}^{K} \frac{(\kappa_i - \widetilde{\kappa}_i)^2}{K}}, \tag{4.22}$$

where κ and $\widetilde{\kappa}$ are the observed and resimulated data.

We start with a simple model to first investigate the effect of S-waves in both acoustic and elastic Born inversion results. Then, a dipping layer model with uncorrelated structure is analyzed to verify the effectiveness of the proposed method using correct and incorrect background models. Finally, the Marmousi2 model is used as a benchmark of complex geology. We also assess the sensitivity of the proposed method to random noise using the Marmousi2 model.

4.4.1 Simple model

We start with a simple model containing uncorrelated impedance-density horizontal structures (Figure 4.2). This example is used to compare the performance of acoustic and elastic WLS-based pseudoinverse Born inversion in both image (true-amplitude parameter inversion) and data domains (data reconstruction). Note that the second interface in the P-wave impedance model, the third interface in the S-wave impedance model, and both second and third interfaces in the density model are missing in the other models. The true reflectivity models are obtained by smoothing the exact models with a 2D

Table 4.1: Main parameter configuration for different numerical experiments.

	Simple model	Dipping layer model	Marmousi2 model
Model grids $(n_x \times n_z)$	351×145	540×194	575×175
Grid intervals (m)	20	17.5	17.5
Source intervals (m)	60	52.5	52.5
Receiver intervals (m)	20	17.5	17.5
Number of sources	117	180	192
Number of receivers per source	351	540	575
Time sampling rate (ms)	5.5	4	2
Max. recording time (s)	4.4	6.1	5.6
Wavelet max. freq. (Hz)	14	20	20

Gaussian filter of 120 m length in both directions and then subtracting them from the exact models (Figure 4.3).

Assuming an acoustic approximation, P- to S-converted waves are not modeled, meaning that the amplitude of the P-waves is also inaccurate. We illustrate this in Figure 4.4, by comparing elastic and variable-density acoustic wave propagation for this model. We note the main differences are in AVO of the reflected P-waves (shown by red arrows in Figure 4.4) and converted P-S-P waves generated in elastic propagation (shown by yellow arrows in Figure 4.4). From now, we only consider the elastic observed data. The details regarding the modeling engines for each Figure for this experiment are outlined in Table 4.2. We apply acoustic and elastic pseudoinverse Born operator on the elastic observed data. The estimated CIGs and their angle-domain responses are shown in Figure 4.5. Note that the specular events in both elastic and acoustic subsurface-offset and angle-domain CIGs have the same shapes but different amplitudes (Figures 4.5e and 4.5f).

Before comparing the inverted parameters, we first check the quality of inversion by comparing the observed and reconstructed data in the middle position $x_s = 3.5$ km (Figure 4.6). There is an accurate match between the observed and reconstructed shots concerning amplitudes and phase for both elastic and acoustic inversion. For the acoustic case, this is indeed expected since whatever the input, the inversion scheme (we mean acoustic pseudoinverse Born operator followed by acoustic modeling) accurately reconstructs the observed data (Chauris and Cocher, 2017). However, the quality of the estimated parameters depends on the consistency between the assumed physics in the

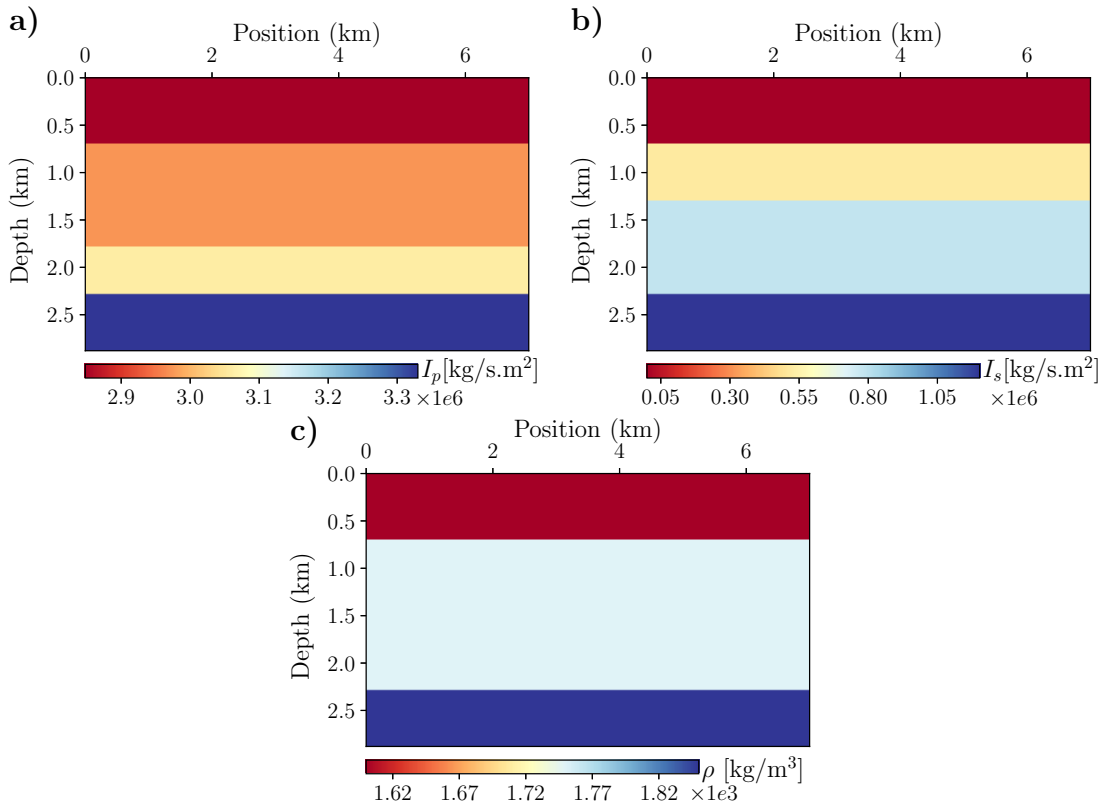

Figure 4.2: Simple model experiment: exact a) P-impedance, b) S-impedance and c) density models.

Table 4.2: Modeling engine for each Figure corresponding to simple model experiment.

	Figure number							
	4.4a	4.4b	4.5a-b	4.5c-d	4.6a	4.6b	4.7a-c	4.7d-e
Input obs. data	elastic	acoustic	elastic	elastic	elastic	elastic	elastic	elastic
Imaging scheme	–	–	elastic	acoustic	–	–	elastic	acoustic
Reconst. data	–	–	–	–	elastic	acoustic	–	–

inversion scheme and observed data. In the case of the elastic inversion, the data reconstruction ability is a new result as it implies that the elastic pseudoinverse Born operator is indeed valid. We now compare the inverted parameters (Figure 4.7). The estimated parameters using elastic WLS show a precise correlation to the true models, albeit small oscillations at the second and third reflector positions ($z = 1.3$ and $z = 1.8$ km) are

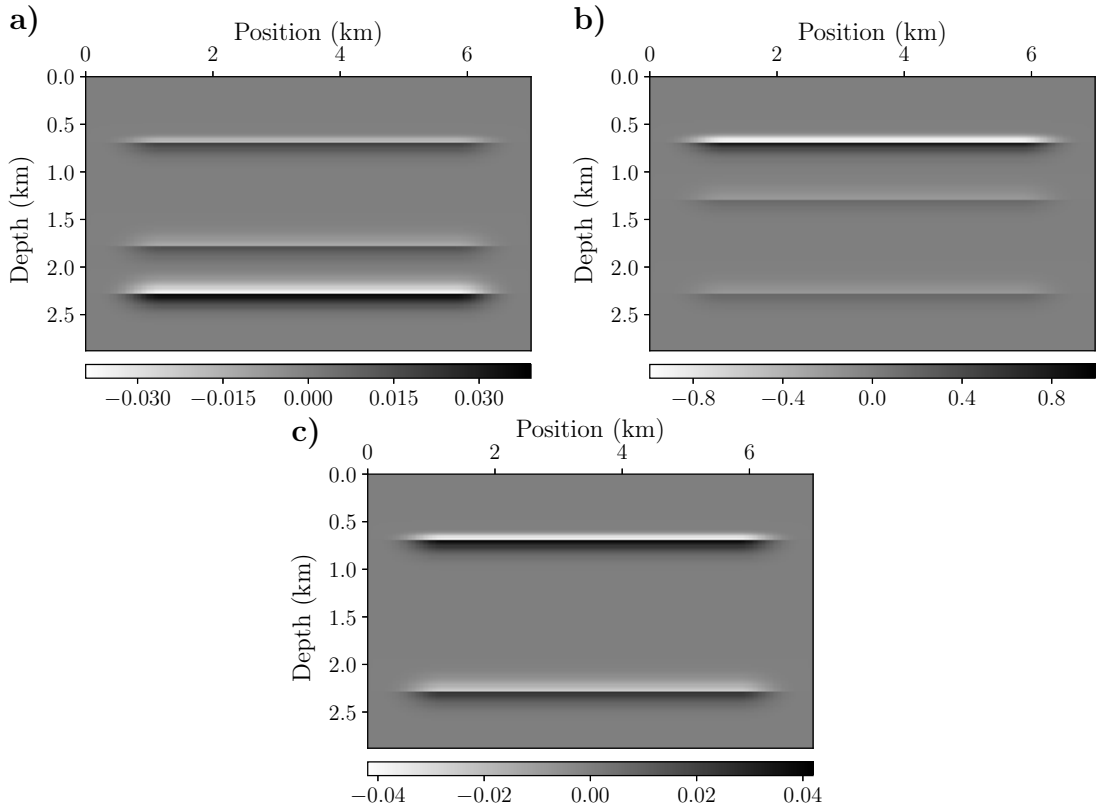

Figure 4.3: Simple model experiment: true perturbation models of a) P-impedance, b) S-impedance and c) density.

observed in the estimated density model (Figure 4.7b). The NCC scores written on each panel also confirm the superiority of the elastic inversion. The inverted parameters using acoustic WLS contain artificial reflectors especially for the second and third reflector positions of the density model (Figure 4.7e). We extract a vertical log from each panel to further quantitatively compare the inversion results (Figure 4.8). Although both methods provide good P-wave impedance reconstruction, the acoustic WLS compensates for S-wave effects by adjusting amplitudes and including artificial reflectors in the estimated density model. This behavior is indeed due to the limitation of the acoustic inversion trying to match the P-S mode converted energy and AVO effects (Barnes and Charara, 2009; Mora and Wu, 2018). Note that the amplitudes of the S-wave impedance are also accurately retrieved using the elastic WLS method (Figure 4.8).

We conclude from this experiment that the inconsistency between modeling and inverse, we mean utilizing acoustic inversion using elastic data, leads to not only artificial reflectors but also to incorrect amplitudes in the estimated images. In the next two numerical experiments, we only investigate the performance of the elastic pseudoinverse

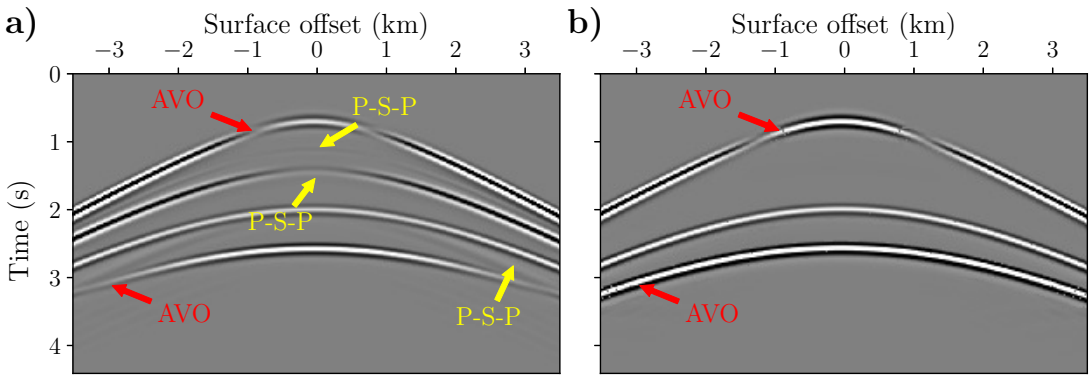

Figure 4.4: Elastic and acoustic shot-domain comparisons for the simple model experiment: a) elastic observed shot versus b) acoustic observed shot. The yellow arrows indicate the main differences between the elastic and acoustic simulations (AVO effects or converted waves).

Born operator using a more complex dipping layer and the Marmousi2 models. We also discuss in more details which parameters are better inverted.

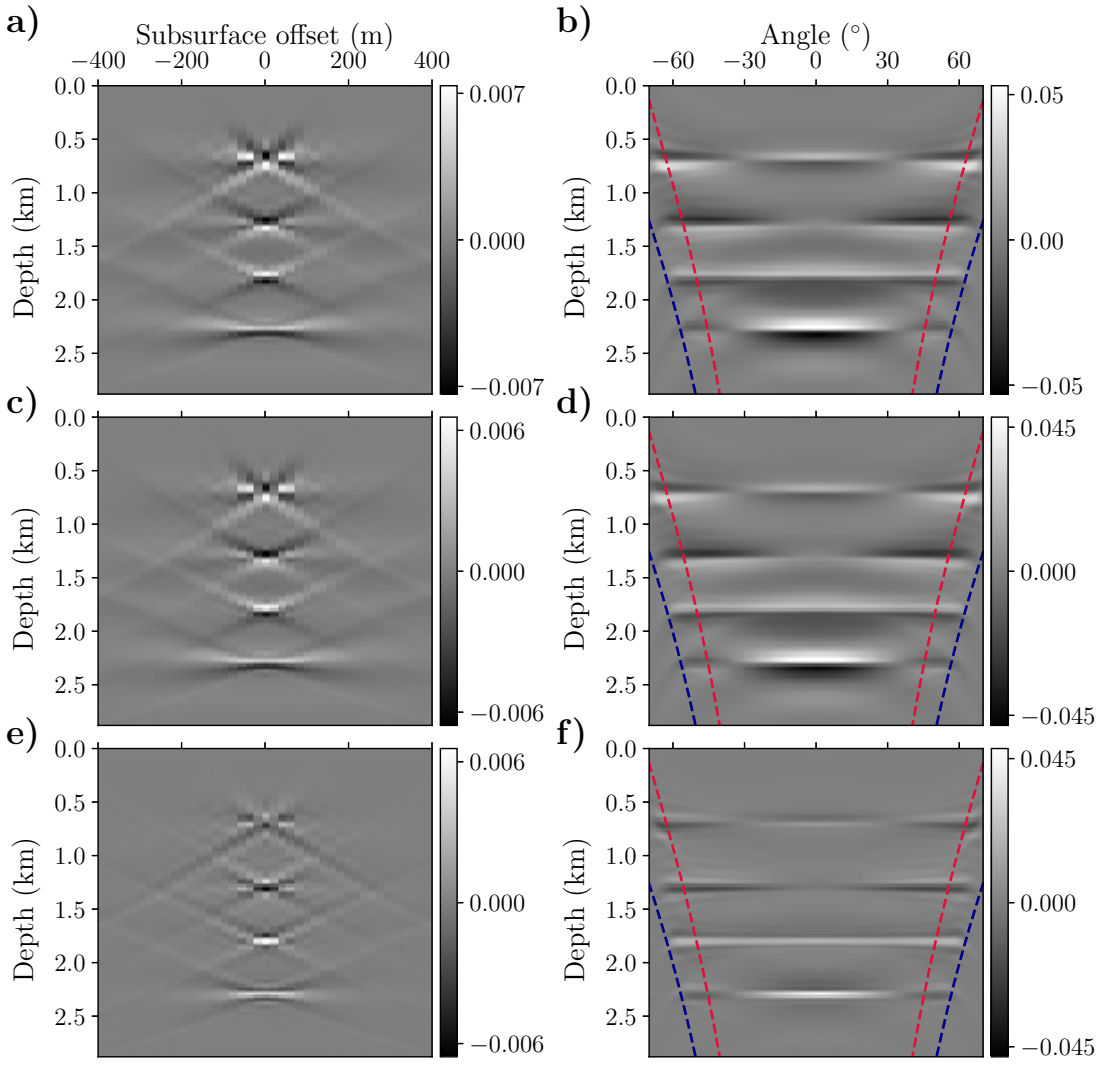

Figure 4.5: Simple model experiment using the elastic observed data: a) inverted δm using elastic pseudoinverse Born operator and b) its Radon transform, c) inverted δm using acoustic pseudoinverse Born operator and d) its Radon transform. The difference between (a,c) and (b,d) are shown in (e) and (f), respectively. The blue and red lines in (b,d,f) correspond to the theoretical limit for acquired angles and the practical ones used in inversion, respectively.

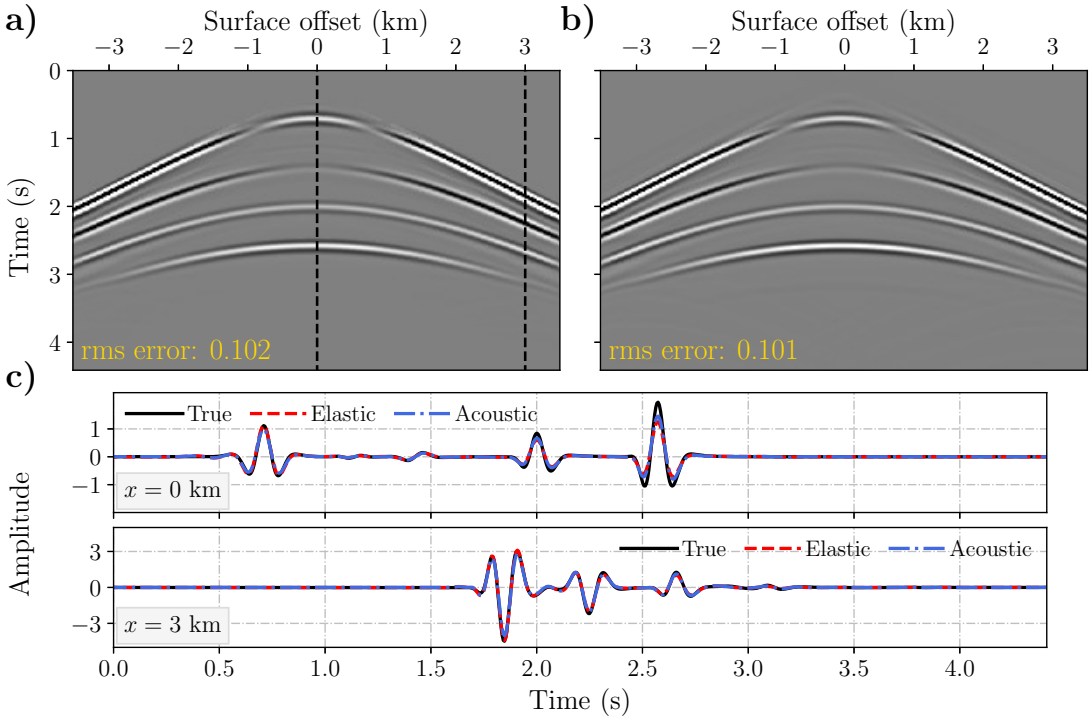

Figure 4.6: Elastic and acoustic reconstructed shot comparisons for the simple model experiment: a) elastic reconstructed shot versus b) acoustic reconstructed shot. Traces are extracted from two positions (indicated as black dashed line in (a)) for surface offset c) 0 and 3 km. Shots are plotted at the same scale as the observed shot (Figure 4.4a).

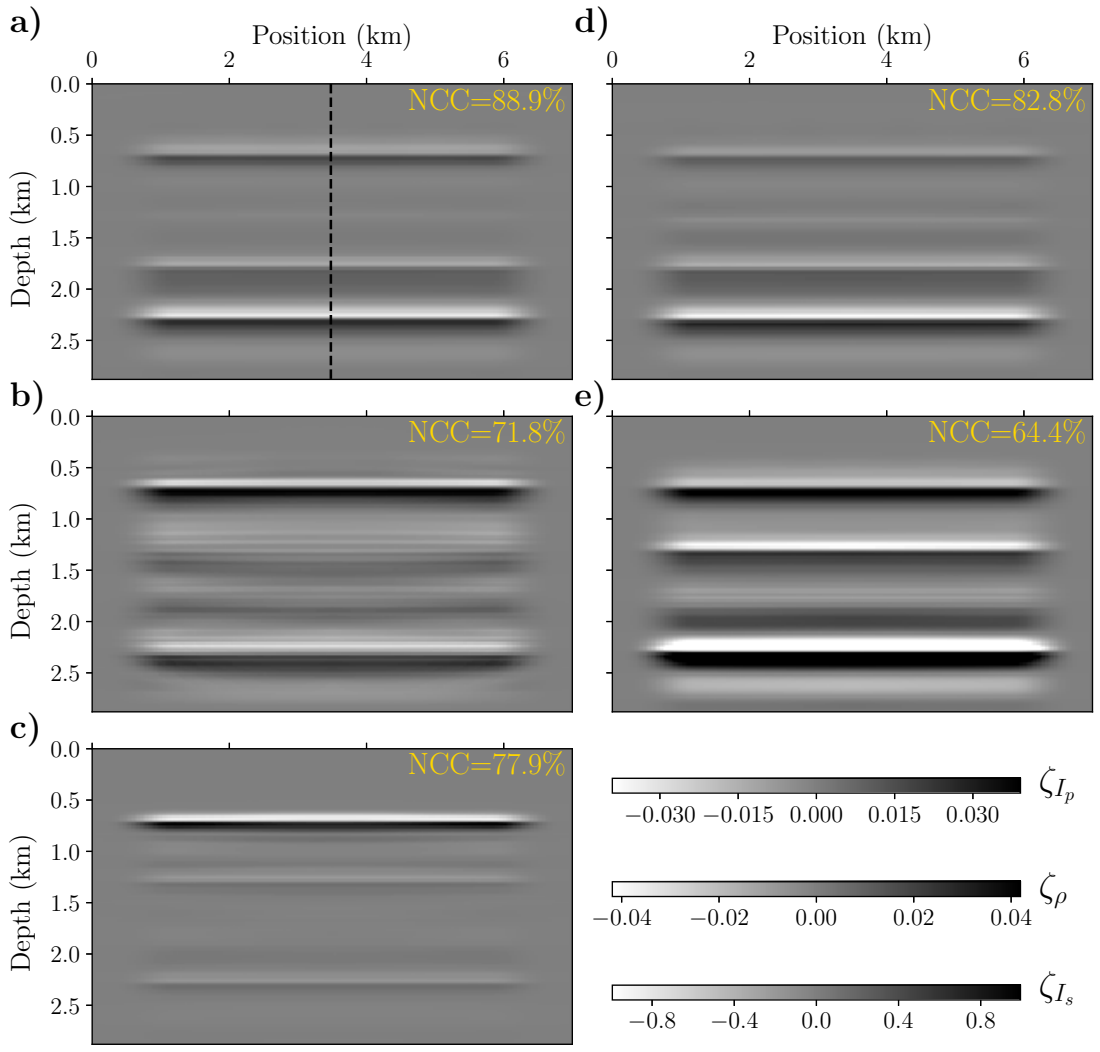

Figure 4.7: Elastic and acoustic inverted image-domain comparisons for the simple model experiment: inverted a) ζ_{I_p}, b) ζ_ρ and c) ζ_{I_s} using elastic WLS, d) ζ_{I_p} and e) ζ_ρ using acoustic WLS.

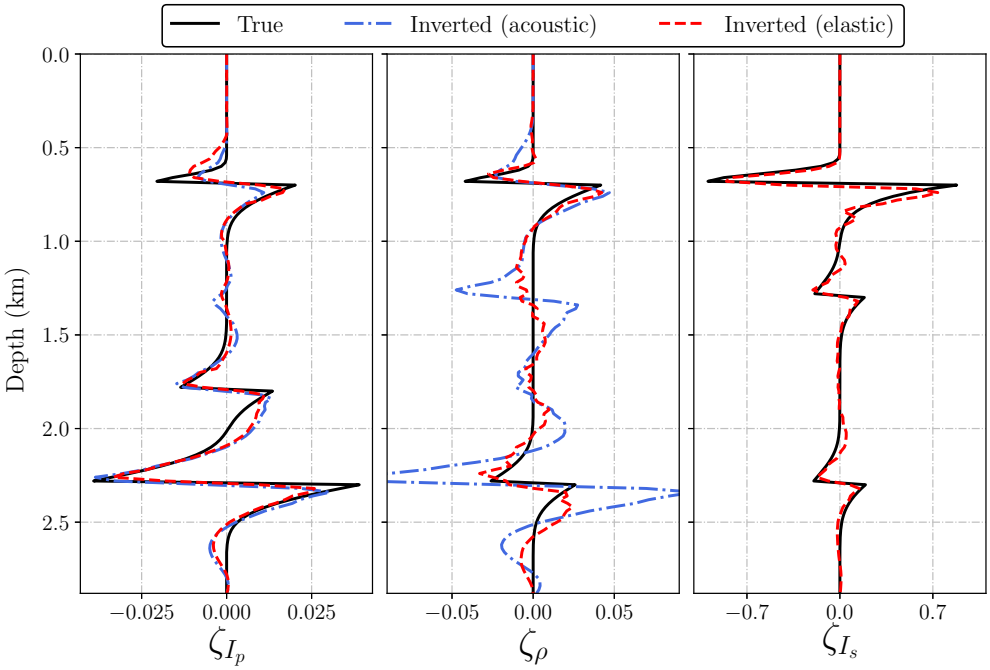

Figure 4.8: Vertical logs extracted at $x = 3.4$ km (Figure 4.7) comparing the true models (black) and estimated ones using acoustic WLS (blue) and elastic WLS (red). The blue line for ζ_{I_s} is not displayed as it is not reconstructed.

4.4.2 Dipping layer model

We consider a more realistic dipping layer model (Figure 4.9) to investigate the sensitivity of the elastic pseudoinverse Born operator to uncorrelated structures and to the accuracy of the background models. The main targets in this model are gas- and water-charged sand channels (indicated by arrows in Figure 4.9c). By uncorrelated, me mean that these channels are easy to see from the P-wave impedance and density models, but not from the S-wave impedance model. Fluid-property discrimination is indeed possible by analyzing all three elastic parameters (Russell et al., 2003).

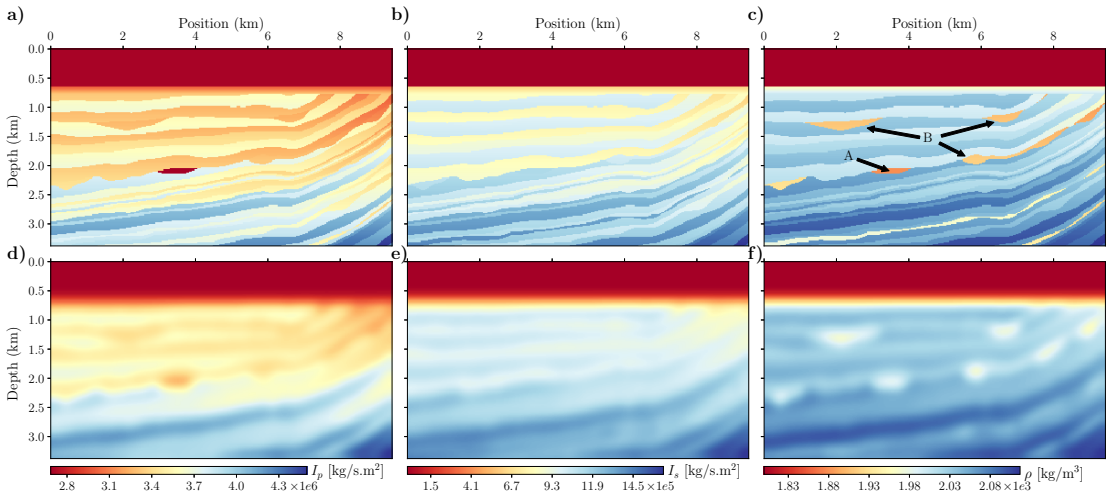

Figure 4.9: Dipping layer model experiment: exact models of a) P-impedance, b) S-impedance and c) density, and background models of d) P-impedance, e) S-impedance and f) density. A and B in (c) indicate gas-charged sand channel and water-wet sand channels, respectively.

In the first test, we consider accurate background models obtained by 2D Gaussian smoothing of the exact models over 7 grid points (Figure 4.9d-f). We start with steps 1 and 2 in our algorithm, estimating subsurface-offset and angle-domain CIGs (Figure 4.10). We first compare the data (elastic) reconstruction for a shot located at 4.7 km (Figure 4.11). The reconstructed shot is plotted with the same colorbar scale as the observed shot. The near- and far- offset extracted traces illustrate an accurate match in terms of amplitudes and traveltimes (Figure 4.11c). Finally, we compare the inverted parameters inferred from the elastic WLS method (Figure 4.12). Note the colorbar scales for the inverted panels are the same as the exact ones for amplitude consistency. At first glance, we observe that our method produces clear subsurface images, with P-wave impedance as the best-reconstructed parameter. Both gas- and water-charged channels are easily detectable. However, some artifacts around the gas-charged sand channel are observed in the S-wave impedance image, which is indeed caused by parameter crosstalk

(yellow arrow in Figure 4.12e). For a more detailed comparison, we extract two vertical logs at the gas- and water- charged sand channels locations (indicated by dashed lines in Figure 4.12a-c) and add them to the background models (Figure 4.13). The elastic WLS method indeed managed to accurately reconstruct all parameters. It is interesting to note that the crosstalk artifact in S-wave impedance at the location of the gas channel (Figure 4.13a) is observed as small oscillations with an acceptable level. Once more, the density is the less constrained parameter.

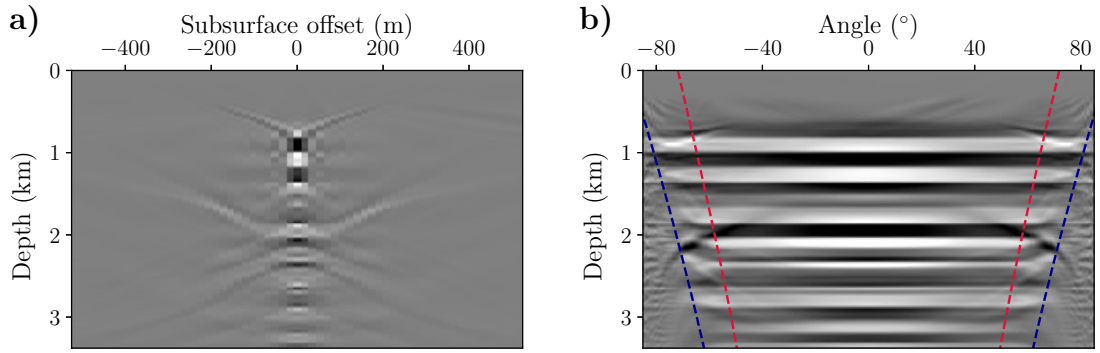

Figure 4.10: Dipping layer model experiment. a) Inverted δm and b) its Radon transform. The blue and red lines in (b) correspond to the limit for acquired angles and the used ones in inversion, respectively.

4.4.2.1 Sensitivity analysis to background models

In real data applications, the errors in migration background models are unavoidable (Yang et al., 2019a). Thus in the second test on the dipping layer model, we study the sensitivity of our proposed method to the reference macro-model. To increase the reliability of the results, we perform the elastic inversion for background models with different error values up to 8% obtained by scaling the accurate background models with a factor between 0.92 and 1.00. In each test, the error is included only in either V_{p0}, V_{s0} or ρ_0 (background model), while the other two models are assumed to be accurate. The evaluation of data misfit and NCC values for image reconstruction with respect to the errors in each background model are shown in Figure 4.14. The data reconstruction indicates a very stable behavior for all inaccurate background models (Figure 4.14a). The NCC curves for each reconstructed image show a large drop by increasing error in the P-wave velocity background model (Figure 4.14a), whereas these curves are rather stable for different error values in the S-wave velocity (Figure 4.14c) and density background models (Figure 4.14d).

To further analyze the results, we display the subsurface-offset- and angle-domain CIGs corresponding to 8% error in the P-wave velocity background model (Figure 4.15).

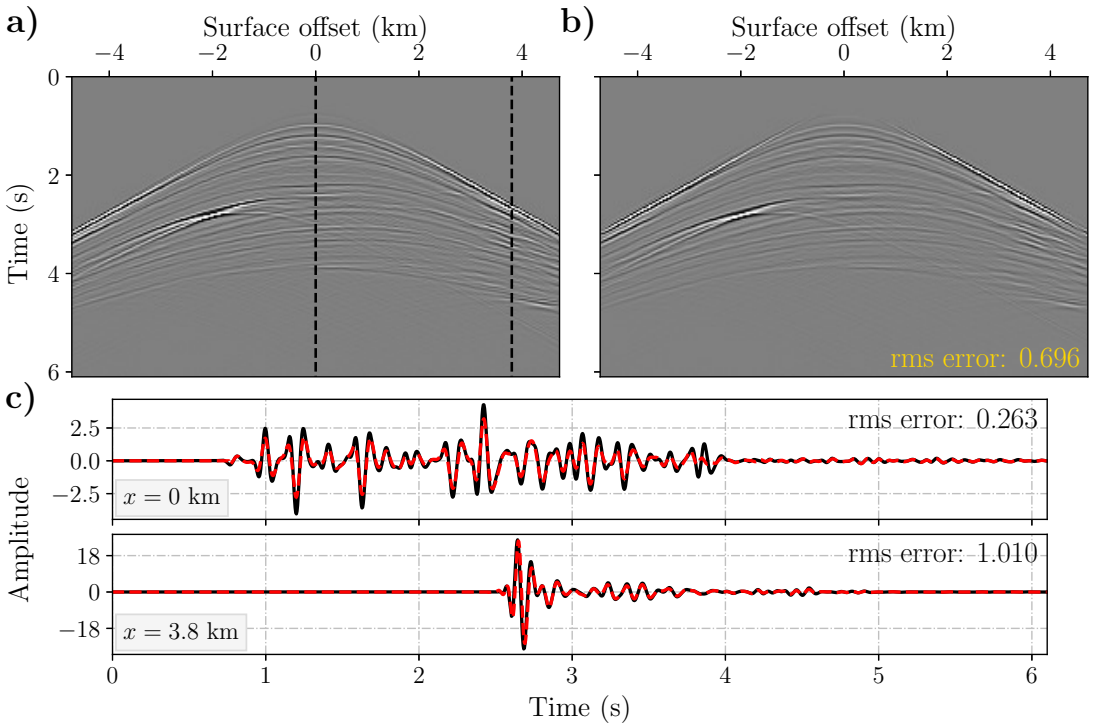

Figure 4.11: The a) observed and b) reconstructed shots for the dipping layer model experiment. Shots are plotted in the same scale. Traces are extracted from two positions (indicated as black dashed lines in (a)) for surface offset c) 0 and 3.8 km. The solid black and dashed red lines correspond to the true and reconstructed values, respectively.

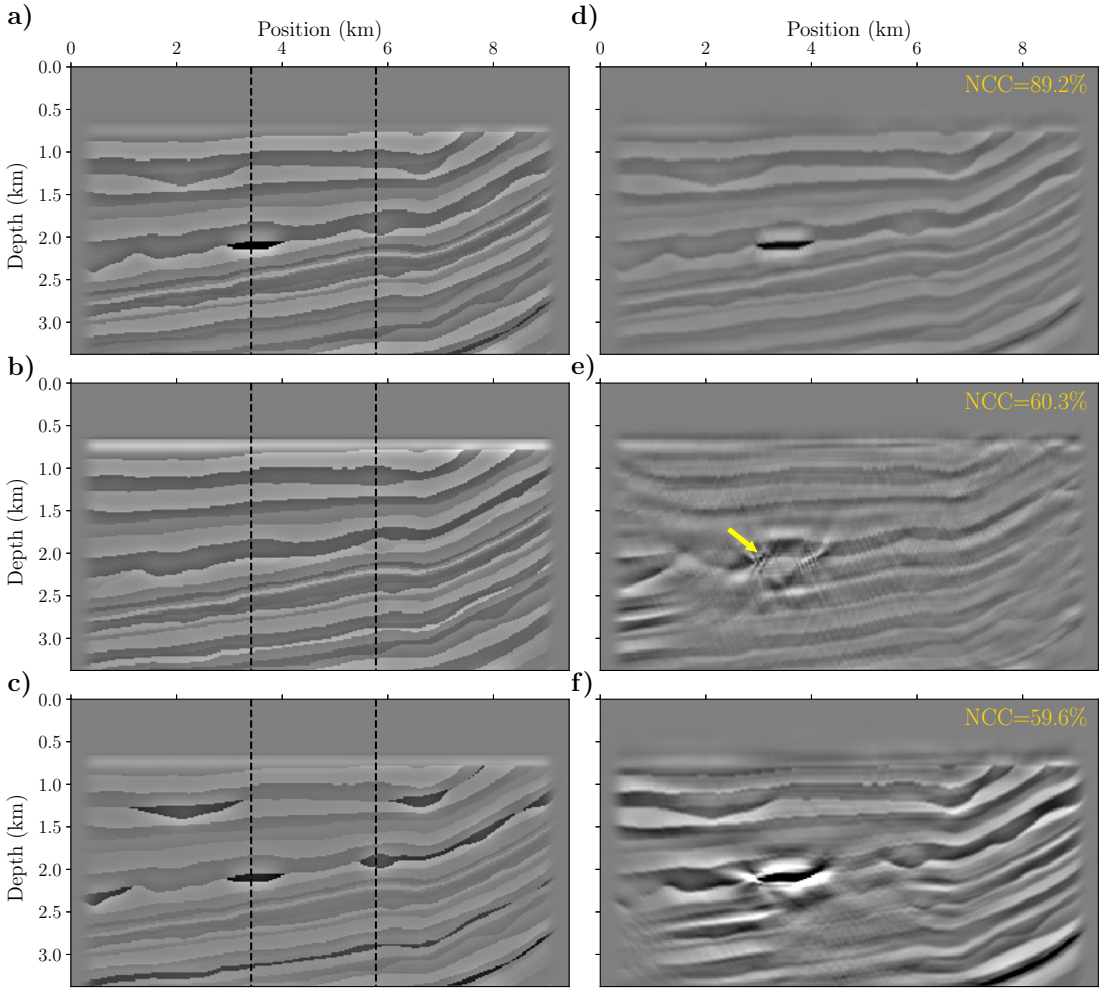

Figure 4.12: Exact and inverted image-domain comparisons for the dipping layer model experiment: exact a) ζ_{I_p}, b) ζ_{I_s} and c) ζ_ρ, and inverted d) ζ_{I_p}, e) ζ_{I_s} and f) ζ_ρ. The dashed lines in (a,b,c) indicate the position of the extracted traces. Inverted models are plotted in the same scale as the exact models.

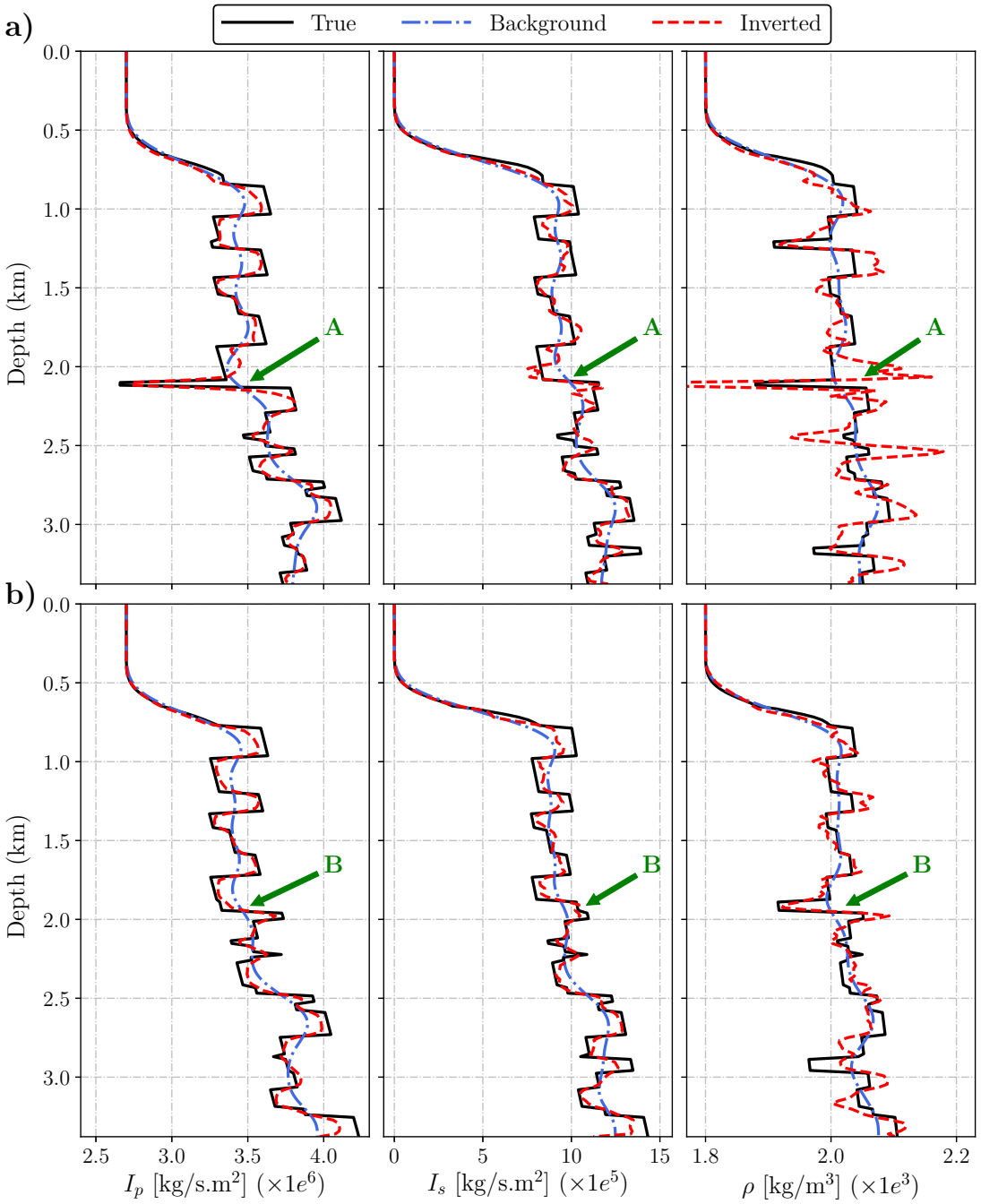

Figure 4.13: Vertical logs passing through the two target areas from inverted dipping layer model (Figure 4.12) for a) 3.4 km and b) 5.8 km. A and B arrows in (a) and (b) indicate the gas charged sand channel and water wet sand channels, respectively.

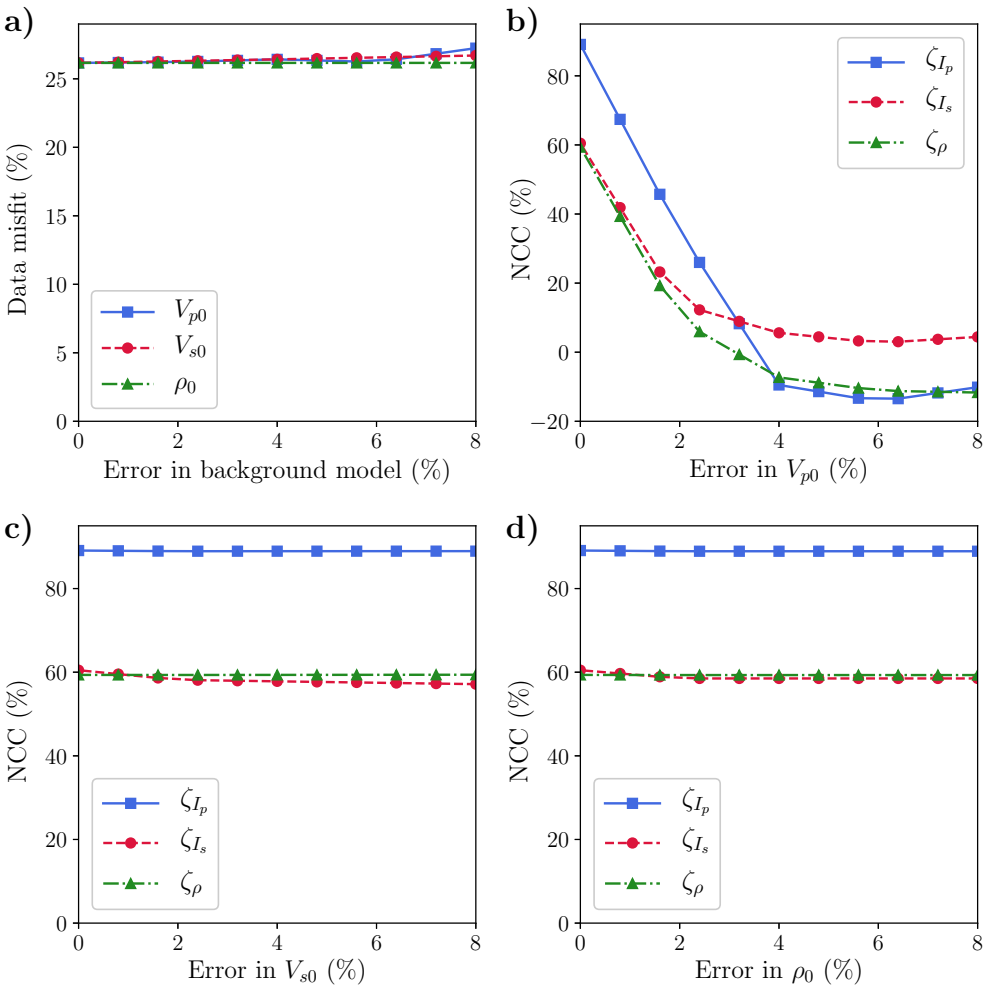

Figure 4.14: Evaluation of a) data misfit with respect to error in different background models, and NCC in image domain with respect to error in b) P-wave velocity background model, c) S-wave velocity background model and d) density background model corresponding to dipping layer model.

Note the downward (upward) curvature of the events in the subsurface offset- (angle-) domain CIG due to the too-low velocity background model. We now analyze the inverted parameters estimated via the elastic WLS method (Figure 4.16). Although crude background models were used, the inverted models are fairly well reconstructed in terms of structure. However, the quality of the S-wave impedance reconstruction is slightly lower, with more parameter crosstalks observed on the S-wave impedance model. Careful examination indicates an upward shift in the inverted parameters consistent with too low background velocity models. Therefore the drop in the NCC values for an inac-

curate P-wave velocity model is inaccurate is expected as the kinematics are mainly governed by the P-wave velocity model.

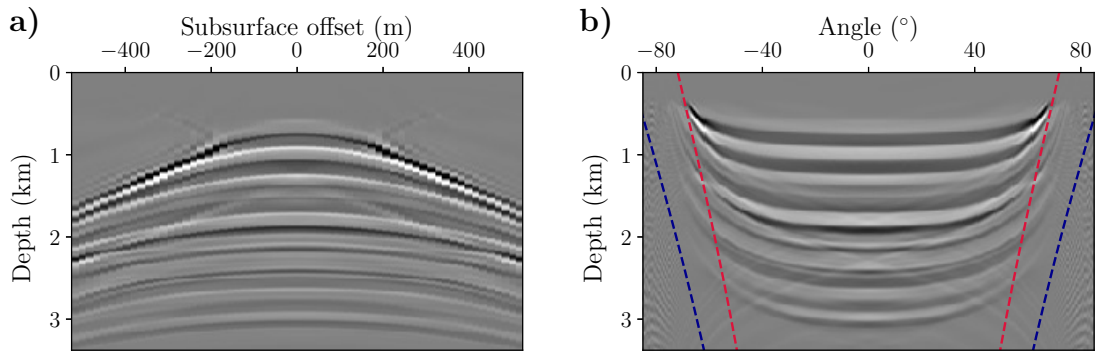

Figure 4.15: Same as Figure 4.10, but for inaccurate P-wave velocity background model (8% error): a) inverted δm and b) its Radon transform.

By noting the ability in reconstructing the observed data even in the presence of inaccurate background models, we conclude here that the proposed elastic inversion scheme can be coupled to the inner loop of MVA to accelerate the convergence rate (Chauris and Cocher, 2017). Updating the background models in the context of elastic imaging is beyond the scope of this paper.

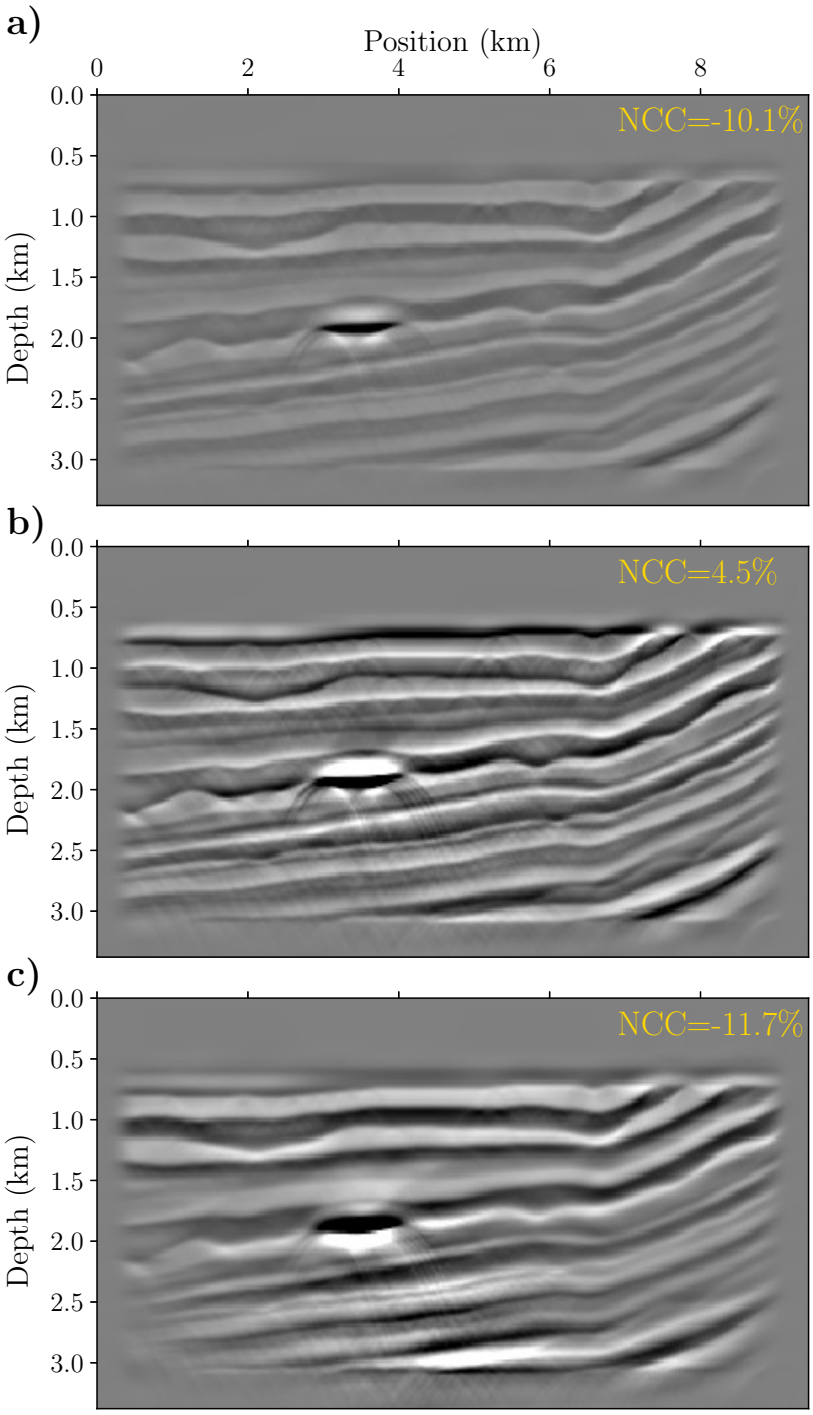

Figure 4.16: Same as Figure 4.12, but for inaccurate P-wave velocity background model (8% error): inverted a) ζ_{I_p}, b) ζ_{I_s} and c) ζ_ρ. Inverted model are plotted in the same scale as the exact models.

4.4.3 Marmousi2 model

To further assess the robustness of our method in the presence of random noise, we conduct two experiments on a portion of Marmousi2 model (Martin et al., 2006). This model is representative of a realistic complex structure with spatially uncorrelated elastic parameters, i.e., the density, P- and S-wave impedance models have independent structure and are not scaled versions of each other (Figure 4.17a-c). The model also contains a potential hydrocarbon reservoir (indicated as gas sand on Figure 4.17). As for the previous examples, the background models are obtained by performing 2D Gaussian smoothing of 125 m length on the exact models in both directions (Figure 4.17d-f). It is also worth noticing that gas sand is not detectable in the background models.

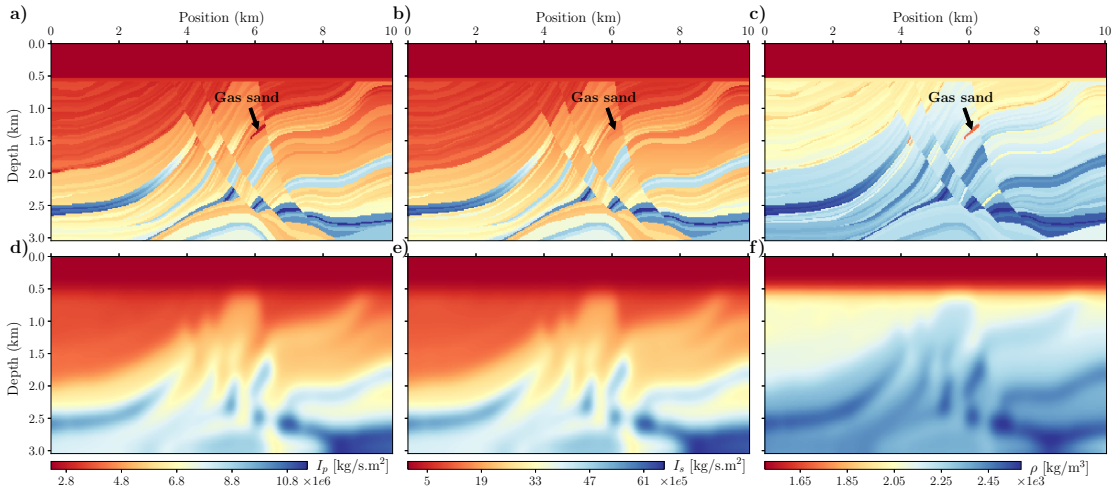

Figure 4.17: Marmousi2 model experiment: exact models of a) P-impedance, b) S-impedance and c) density, and background models of d) P-impedance, e) S-impedance and f) density.

In the first experiment, we consider noise-free observed data (Figure 4.18a). Similar to the previous numerical experiments, we first examine the data reconstruction (Figure 4.18b). Once more, the extracted traces confirm that the reconstructed shot from inverted models predicts data very close to the observed shot (Figure 4.18c). The underestimation and difference in data reconstruction, especially at large offsets and complex models, could result from the energy out of the asymptotic framework and finite acquisition geometry (ten Kroode, 2012; Hou and Symes, 2017). The quality of the fitting can be also affected by the stabilizing parameter ϵ in source wavelet inversion using $\Omega^\dagger = \Omega^*/(\|\Omega\|^2 + \epsilon)$ in the frequency domain. The same amplitude behavior was also observed by Hou and Symes (2015) and Chauris and Cocher (2017). The image-domain comparison between true and inverted parameters via the elastic WLS method is shown in Figure 4.19. As for the two first examples, the NCC scores corresponding to

each parameter (written on each panel of Figure 4.19) rank P-wave impedance, S-wave impedance and density from best-inverted parameter to worst, respectively. To assess the quantitative nature of inversion, we extract two vertical logs (indicated by dashed lines in Figure 4.19a-c) and sum them with the background models. The position of the gas sand is indicated by the green arrow in Figure 4.20b. As can be seen, P- and S-wave impedances are nicely reconstructed while density has a more oscillatory nature, being in line with the NCC scores (Figure 4.20).Chauris2017from

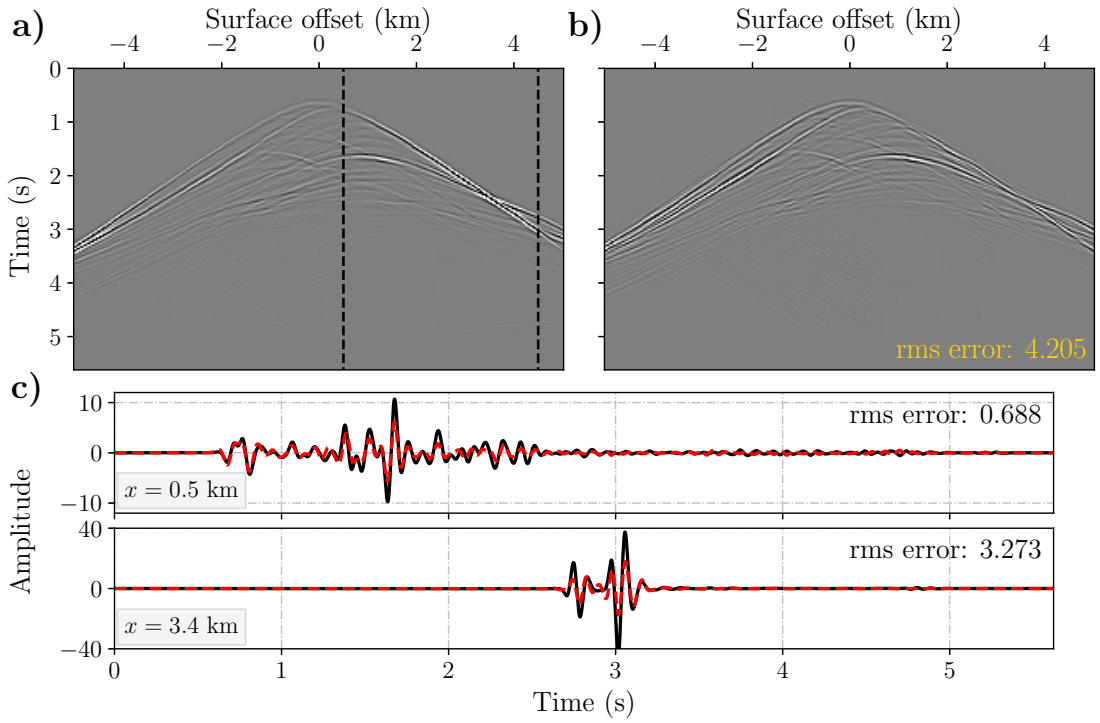

Figure 4.18: The a) observed and b) reconstructed shots for the Marmousi2 model experiment. Shots are plotted in the same scale. Traces are extracted at two positions (indicated as black dashed lines in (a)) for surface offset c) 0.5 and 4.5 km. The solid black and dashed red lines correspond to the true and reconstructed values, respectively.

To further analyze the resolution of the inverted images, we carry out an eigenvalue (λ) decomposition of the 3×3 matrix $\mathbf{M}$ (equation 4.18). The first (λ_1), second (λ_2), and third (λ_3) eigenvalues are mainly related to ζ_{I_p}, ζ_{I_s}, and ζ_ρ. We calculate the radio between 2nd and 1st (λ_2/λ_1), and 3rd and 1st eigenvalues (λ_3/λ_1) (Figure 4.21). The structure of the resolution depends on the structure of the background and the reflectivity models. As illustrated, a higher resolution is expected in the shallow part of the models (higher λ_2/λ_1 ratio in the shallow part). Moreover, λ_3/λ_1 has a considerable lower ratio compared to λ_2/λ_1, confirming the fact that density is more difficult to reconstruct even

in the shallow part of the model. It is worth to mention that the same resolution analysis using eigenvalues has been investigated for interparameter trade-off quantification in FWI (Plessix and Cao, 2011; Pan et al., 2018).

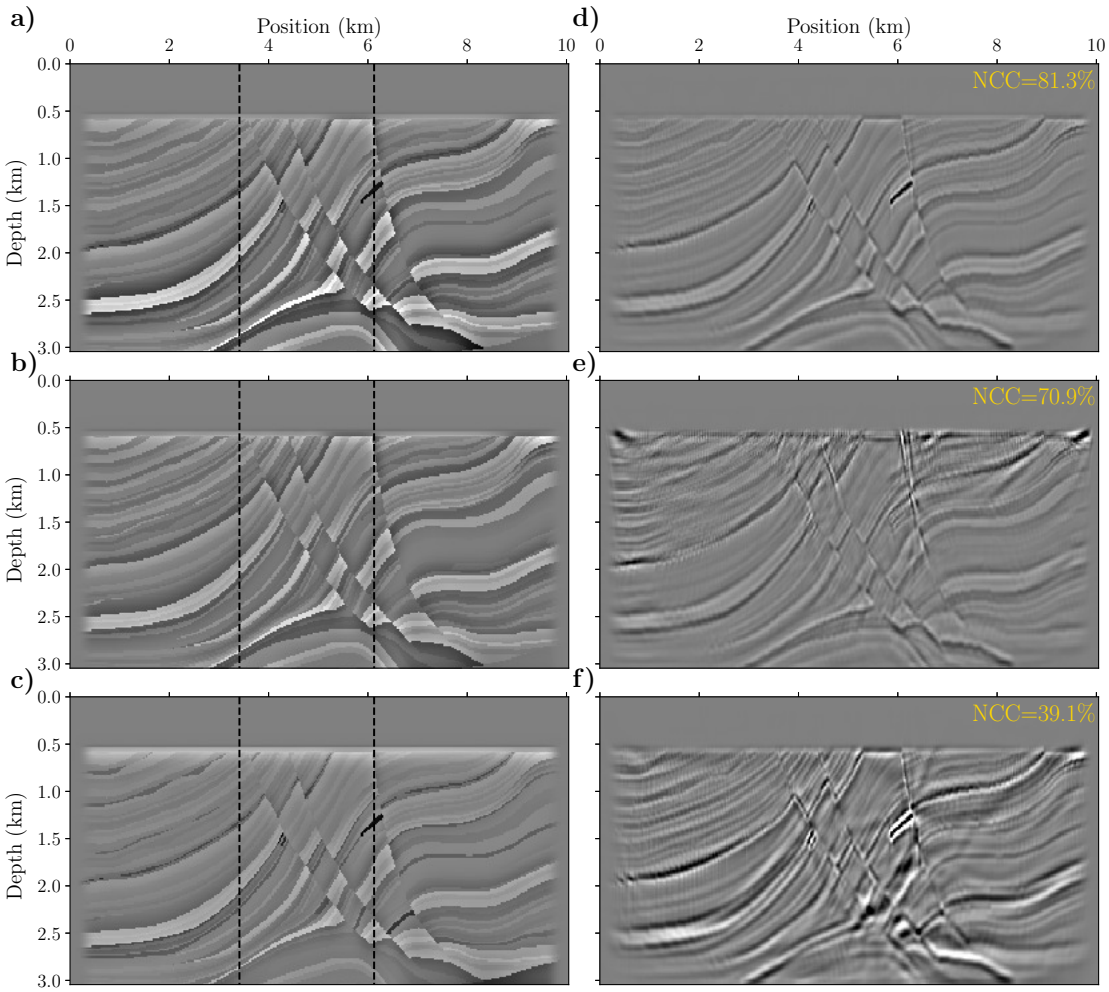

Figure 4.19: Exact and inverted image-domain comparisons for the Marmousi2 model experiment: exact a) ζ_{I_p}, b) ζ_{I_s} and c) ζ_ρ, and inverted d) ζ_{I_p}, e) ζ_{I_s} and f) ζ_ρ. The dashed lines in (a,b,c) indicate the position of the extracted traces. Inverted models are plotted in the same scale as the exact models.

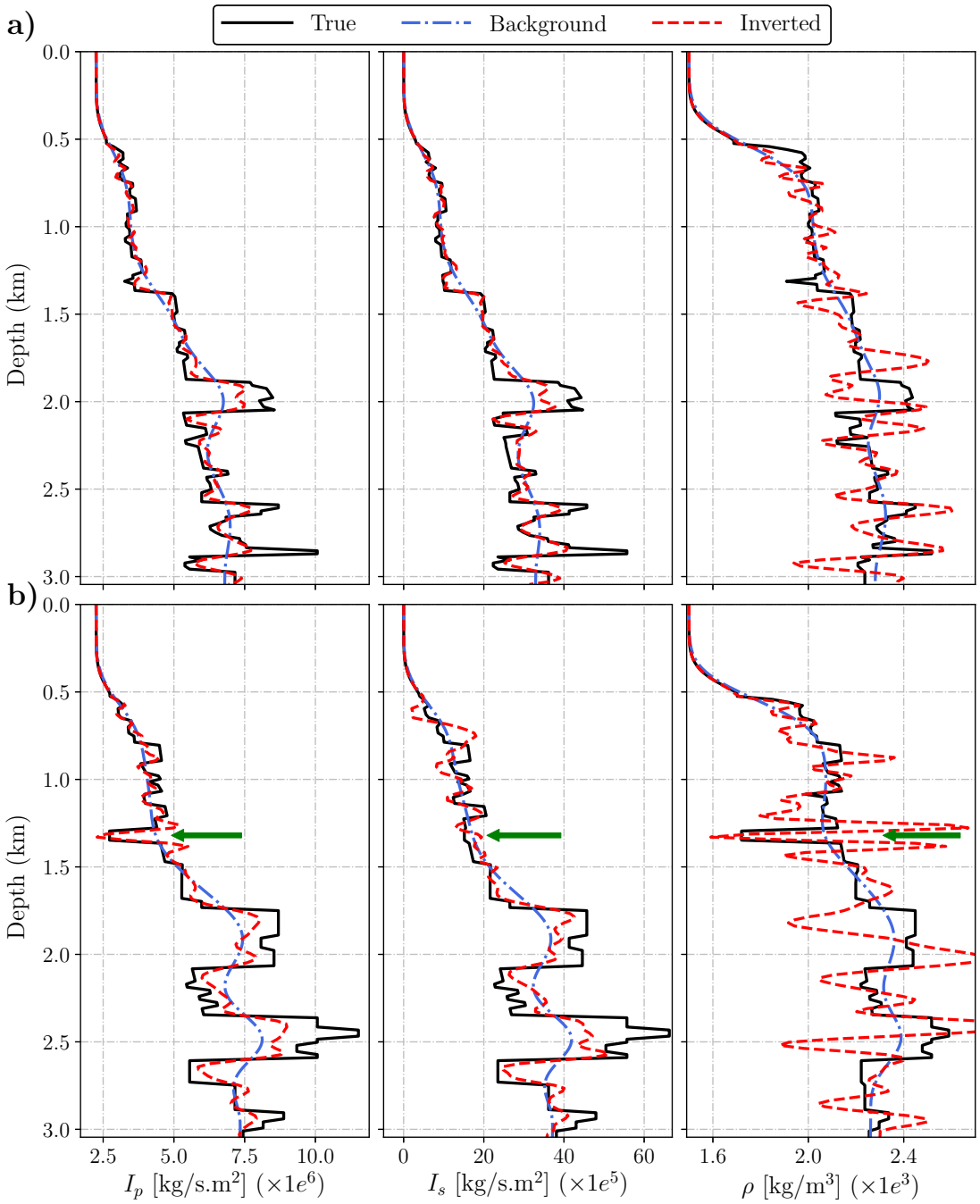

Figure 4.20: Vertical logs from inverted Marmousi2 model (Figure 4.19) for a) 3.4 km and b) 6.1 km. The green arrow in (b) indicates the position of the gas sand.

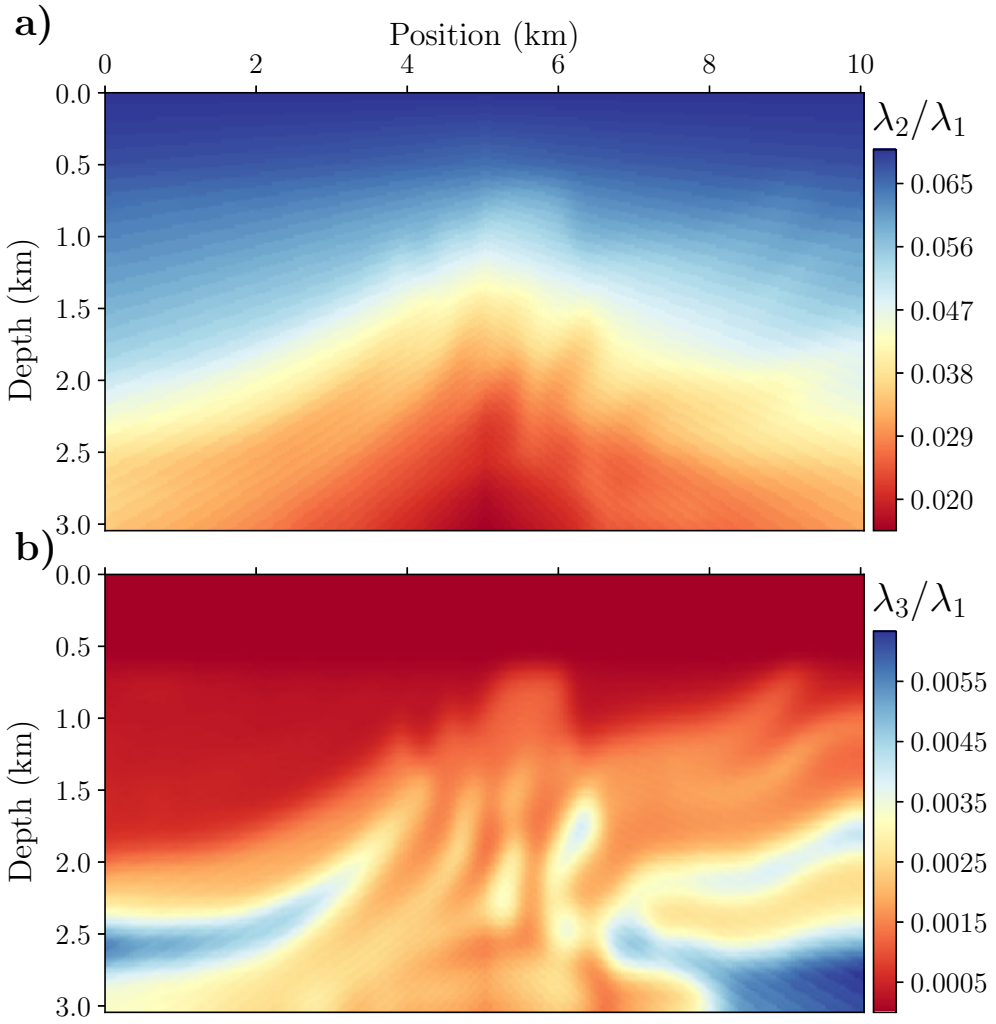

Figure 4.21: Radio of a) 2nd/1st and b) 3rd/1st eigenvalues of the matrix M at each spatial coordinate corresponding to the Marmousi2 model.

4.4.3.1 Sensitivity analysis to white noise

The existence of noise in field data is inevitable. Therefore, in the second experiment on Marmousi2, we assess the robustness of our method to Gaussian noise effect by performing three inversion tests using different noise levels. The Gaussian noise not only can degrade the quality of the RTM images (Ren et al., 2017) but also can have a different effects on different subsurface parameters in least-squares imaging (Brossier et al., 2010). In order to produce noisy observed data, we first generate Gaussian noise gathers with a signal-to-noise ratio (SNR) of -1, -5, and -10 dB, and then filter them at the maximum frequency of the applied wavelet before adding to the previously observed

data (Figure 4.22a-c). In addition, we mute the noise before the first arrivals. The negative sign in SNR indicates that the noise power is greater than the signal power.

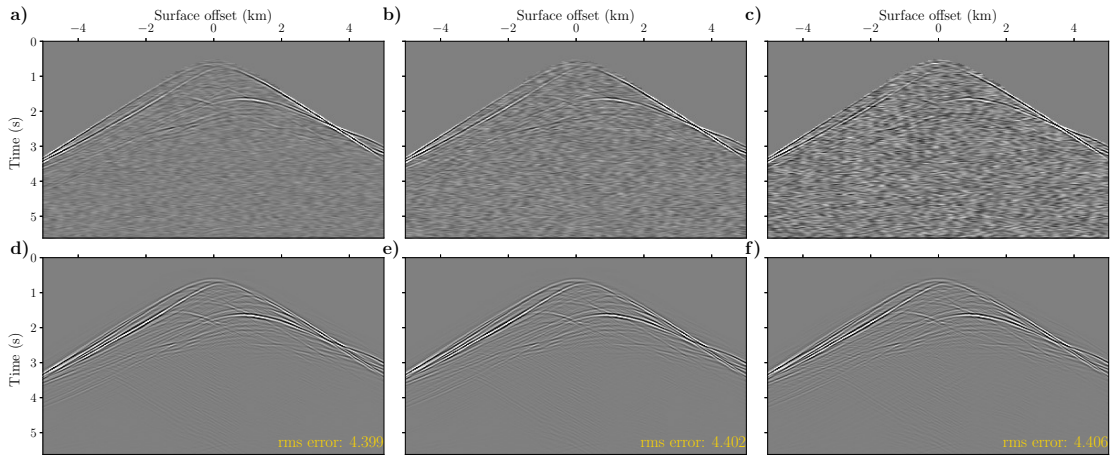

Figure 4.22: The Marmousi2 noisy observed shot with a) SNR -1 dB, b) SNR -5 dB and c) SNR -10 dB, and the reconstructed shots corresponding to noisy data with d) SNR -1 dB, e) SNR -5 dB and f) SNR -10 dB. Shots are plotted in the same scale as the noiseless observed shot (Figure 4.18a).

The comparisons of the reconstructed shots show that the synthetic data modeled from the inverted parameters are almost the same as the original data (Figure 4.22e-f). For details, we extract one trace ($x = 0.5$ km) from each panel of Figure 4.22 to further demonstrate the accuracy of phase and amplitudes (Figure 4.23). It is interesting that although the pseudoinverse Born operator is derived such that $d^{rec} \approx \mathcal{L}\mathcal{L}^{\dagger} d^{obs}$, it can nicely discriminate between the signal and the noise within a single iteration, even in the case that the signal is buried in the noise (Figure 4.23). The important aspect is that the noise is not coherent from a trace to another. The recovered three elastic parameters corresponding to different noise levels are shown in Figure 4.24. The effect of noise is observed as unwanted dots and small oscillations on inverted panels. Comparing NCC scores reveals that the degradation in quality is rather small. To get more insight into the effect of noise on the inverted images, we calculate localized NCC values for each model within a window of 0.5 km length in both directions (Figure 4.25). Although a detailed look reveals small degradation in quality by increasing noise, the localized NCC images are almost identical to each other. We once more verify that the worst-reconstructed parameter is the density.

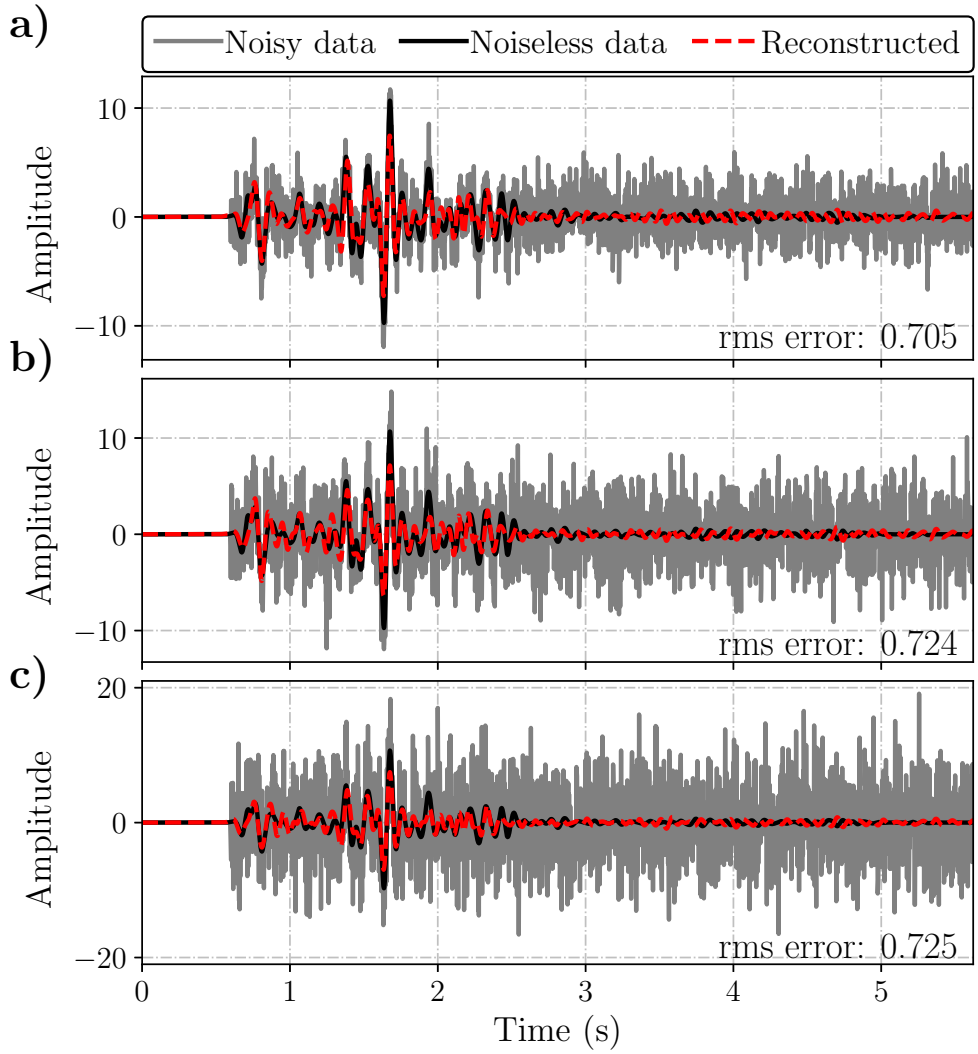

Figure 4.23: The extracted traces at $x = 0.5$ km from Figure 4.22 for a) SNR -1 dB, b) SNR -5 dB and c) SNR -10 dB.

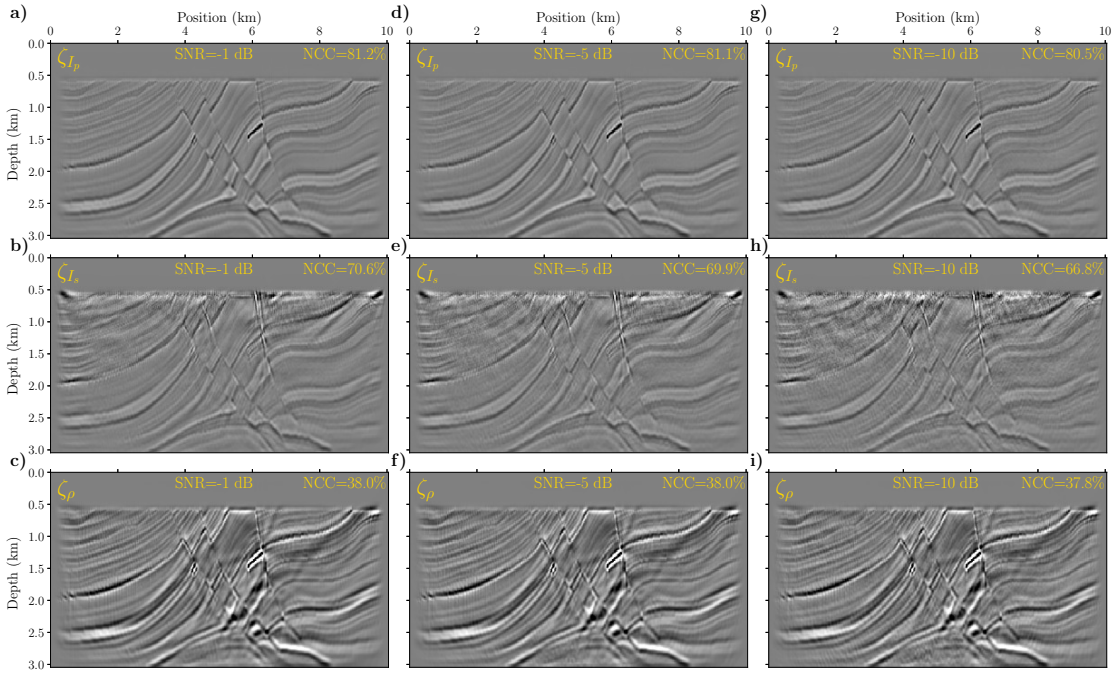

Figure 4.24: The inverted image-domain comparisons for the Marmousi2 model using noisy data: inverted (a,d,g) ζ_{I_p}, (b,e,h) ζ_{I_s} and (c,f,i) ζ_{ρ} using noisy data with (a,b,c) SNR -1 dB, (d,e,f) SNR -5 dB and (g,h,i) SNR -10 dB.

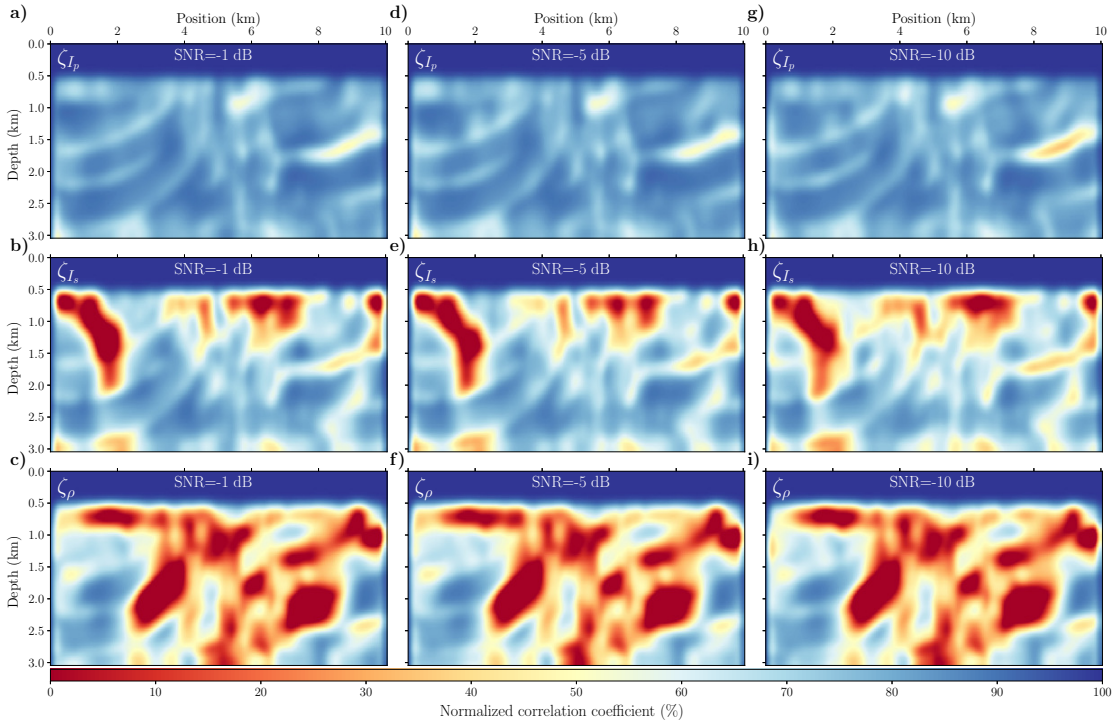

Figure 4.25: The localized NCC values corresponding to Figure 4.24: for (a,d,g) ζ_{I_p}, (b,e,h) ζ_{I_s} and (c,f,i) ζ_{ρ} using noisy data with (a,b,c) SNR -1 dB, (d,e,f) SNR -5 dB and (g,h,i) SNR -10 dB.

4.5 Discussion

We have extended the acoustic pseudoinverse Born operator to elastic media. The elastic and acoustic pseudoinverse Born operators are similar in structure. The main differences are elastic versus acoustic Green's functions in the extended pseudoinverse formula, and then different decomposition in the WLS method (three parameters versus two parameters). The main limitation is the assumption of marine towed-streamer data. The extension to land data (non zero V_s in the shallow part of the model) remains a research topic for future study. It would include the analysis of multi-component data, whereas here only the T_1 component is used and related to the pressure wavefield in the marine environment. The wavefields would have to be separated in advance and associated with different radiation patterns.

Another possible way to estimate the quantitative elastic parameters is based on AVO/AVA inversion schemes which also requires measuring the reflectivity as a function of incident angle (Chopra and Castagna, 2005). These schemes, based on Zoeppritz equations or their approximations, impose some limitations. They assume a convolutional Earth model which not applicable to seismic data at large offsets. They are only valid for the plane-wave propagation with relatively small angles of incidence in isotropic layered media with no geometrical-spreading term, no scattering/transmission loss, no mode converted wave, no multiples, no effect of thin layering and they are not always applicable to structurally complex models (Mosher et al., 1996; Mallick, 2001). These assumptions, however, are not met in practical applications and require sophisticated data processing steps. The proposed method in this article not only does not have any limitation on the maximum offset/angle but also better estimates model parameters using a larger maximum offset/angle. It uses the numerical solutions to the elastic wave equation to construct the offset-/angle-domain CIGs and compensates for the effect of wavelet, geometrical spreading, and uneven illumination, thus it can be categorized as a true-amplitude migration/inversion technique.

The elastic pseudoinverse Born inversion proposed here is a new step toward accelerating the convergence rate of other multi-parameter imaging methods such as LSRTM, MVA, FWI, and reflection waveform inversion (RWI). The application as a preconditioner for LSRTM is straightforward as it is defined for this case (Hou and Symes, 2016a), whereas it requires more investigation for MVA, FWI, and RWI. The main difference between MVA and FWI/RWI lies in the definition of the objective function, being image-domain driven for MVA and data-domain driven for FWI/RWI. In case of acoustic FWI, several authors showed that the pseudoinverse Born operator can indeed provide higher resolution subsurface images, but cannot necessarily mitigate the well-known cycle-skipping issue (Métivier et al., 2015; Hou and Symes, 2016b). The concept of MVA/RWI, though, is rather different, being a nested optimization procedure with updating the true-amplitude reflectivity in the inner loop (migration mode) and updating the background model in the outer loop (tomographic mode) (Xu et al., 2012a). One im-

portant aspect observed here is that the proposed multi-parameter pseudoinverse Born inversion can reconstruct the data even using inaccurate background models. Therefore, this provides evidence that our method can be coupled to the inner loop of MVA and possibly RWI. The curvature of an event in the angle-domain CIG is indeed the bridge between migration and velocity updates in MVA (Biondi and Symes, 2004).

In the concept of multi-parameter imaging, a key problem that comes to the fore is the parameter cross-talk. In general, seismic imaging methods are formulated as least-squares inverse problems and solved by optimization algorithms (Virieux and Operto, 2009). In this case, the least-squares iterations fully resolve the Hessian matrix and remove the cross-talks artifacts (Chen and Sacchi, 2017; Feng and Schuster, 2017). The reader will have noticed that there is an amplitude leakage between the estimated images by the pseudoinverse Born operator (Figures 4.12 and 4.19) which means further least-squares iterations are required. The most important further step is implementing the proposed method as a preconditioner either by defining as an operator within the Preconditioned Conjugate Gradient algorithm or as weighted norms within the weighted Conjugate Gradient algorithm (Hou and Symes, 2016a). More research, beyond the scope of this paper, is required to investigate the role of the pseudoinverse Born operator as a preconditioner in reducing parameter cross-talks.

We now discuss the choice of the parametrization. Generally, impedance-density class is considered in the least-squares inversion of seismic reflection data (Tarantola, 1986). The other possible parameterization classes are Lamé parameters or velocities. In the case of variable-density acoustic Born inversion, Farshad and Chauris (2020a) concluded that P-wave impedance is the best-reconstructed parameter. In the context of elasticity, the conclusions derived here are consistent with the ones known for FWI: impedances are better reconstructed than the density (Köhn et al., 2012; Ren et al., 2017), and larger surface offsets provides better resolution. This can be also noticed from the radiation pattern (Figure 4.1): the density only radiates energy at very large angles, which means that in the reflection data, there is only a little information from density. In this sense, the density does not influence the estimated impedance models, and the imaging result for density is less constrained and reliable. The most computationally expensive step of the proposed method is the elastic pseudoinverse Born operator calculating $\delta m(\mathbf{x}, \mathbf{h})$ (equation 4.10). The computational cost of the second and third steps, namely the Radon transform (equation 4.14) and the WLS method (equation 4.17), is negligible as it is detached from the imaging step that contains wave-equation based operators. It is also worth mentioning that the choice of parametrization is only important in the third step decomposing the extended reflectivity $\delta m(\mathbf{x}, \mathbf{h})$ to three elastic parameters. Thus the algorithm has the flexibility to easily exploit the other parameterization classes. In case of variable-density acoustics, Farshad and Chauris (2020a) investigated the choice of parameterization in WLS method. By using non-linear reparameterization, they demonstrated that there is a good agreement between different parameter classes suggesting that the choice of parameterization does not change the fi-

nal results in this scheme. The conclusion differs in the case of FWI (nonlinear imaging approach) (Virieux and Operto, 2009), and maybe in the linearized waveform inversion methods that invert for parameters through the wavefields and imaging condition (Sun et al., 2018). The reason is that we consider here a linear imaging process in the derivation of the extended images.

Finally, we discuss the impact of regularization. Incorporating a priori information about the medium to be imaged, such as well logs or statistical characteristics, can indeed alleviate the risk of overfitting (Asnaashari et al., 2013). Since the WLS objective function is set up as an inverse problem, it has the flexibility to easily incorporate regularization terms. It is important to note that this step does not involve the expensive computation of wavefields, but only the resolution of 3×3 matrix systems for each image point. Recently, Farshad and Chauris (2020c) implemented sparsity-promoting regularization terms in the acoustic WLS objective function. The regularization can indeed further suppress migration artifacts caused by parameters cross-talk, noise, and incomplete data.

4.6 Conclusion

We have developed an elastic pseudoinverse Born inversion approach that can simultaneously retrieve P- and S-wave impedances and density to a less extend. Quantitative results were obtained within a single iteration. Numerical experiments proved that the proposed method is very promising and provides robust results in terms of data reconstruction and quantitative parameter estimation even in presence of uncorrelated structures, inaccurate migration background models, and noisy data. The new pseudoinverse Born operator can be used as a preconditioner. Therefore, it offers the possibility of enlarging the applicability of iterative imaging methods by enhancing the convergence rate. Future work will explore coupling multi-parameter Born inversion to other iterative imaging techniques such as least-squares migration, migration velocity analysis, and full/reflection waveform inversion as well as investigating the possibility to extend the pseudoinverse Born operator to anisotropic media.

4.7 Acknowledgments

We would like to thank assistant editor Arthur C. H. Cheng, anonymous associate editor, reviewer Ettore Biondi, and two anonymous reviewers for their valuable suggestions and comments that improved the manuscript.

4.8 Appendix I: Gradient derivation for elastic WLS method

In this appendix, we detail the derivation of the gradient of the elastic WLS objective function (equation 4.15) with respect to the model parameters as

$$\frac{\partial J}{\partial \zeta_{I_p}} = \int \mathrm{d}\theta\, 2W(\mathbf{x},\theta) \quad \Big(2\zeta_{I_p}(\mathbf{x}) - 4C_0\sin^2(2\theta)\zeta_{I_s}(\mathbf{x}) + 2(C_0\sin^2(2\theta) - \sin^2(\theta))\zeta_\rho(\mathbf{x}) - \mathcal{R}_{\delta m}(\mathbf{x},\theta)\Big), \tag{4.23}$$

$$\frac{\partial J}{\partial \zeta_{I_s}} = \int -\mathrm{d}\theta\, 4W(\mathbf{x},\theta)\, C_0\sin^2(2\theta) \quad \Big(2\zeta_{I_p}(\mathbf{x}) - 4C_0\sin^2(2\theta)\zeta_{I_s}(\mathbf{x}) + 2(C_0\sin^2(2\theta) - \sin^2(\theta))\zeta_\rho(\mathbf{x}) - \mathcal{R}_{\delta m}(\mathbf{x},\theta)\Big), \tag{4.24}$$

and

$$\frac{\partial J}{\partial \zeta_{\rho}} = \int \mathrm{d}\theta\, 2W(\mathbf{x},\theta)(C_0\sin^2(2\theta) - \sin^2(\theta)) \quad \Big(2\zeta_{I_p}(\mathbf{x}) - 4C_0\sin^2(2\theta)\zeta_{I_s}(\mathbf{x}) + 2(C_0\sin^2(2\theta) - \sin^2(\theta))\zeta_\rho(\mathbf{x}) - \mathcal{R}_{\delta m}(\mathbf{x},\theta)\Big). \tag{4.25}$$

Subsequently, the optimal ζ_{I_p}, ζ_{I_s} and ζ_ρ parameters are obtained where the gradients in equations 4.23, 4.24 and 4.25 vanish.

4.9 Appendix II: Acoustic WLS method

Here we review the main steps of the WLS-based variable-density acoustic pseudoinverse Born inversion proposed by Farshad and Chauris (2020a). The strategy is the same as the elastic case described in this article. The first step is the application of the pseudoinverse Born operator. It is clear that in the acoustic case, the Green's functions in equation 4.10 are calculated using an acoustic Born modeling engine, i.e. $V_s = 0$ and $\beta_p = 1/\rho V_p^2$. We consider (I_p, ρ) parameterization, so the diffraction pattern becomes

$$\mathfrak{D}(\theta) = \begin{bmatrix} 2 & -2\sin^2(\theta) \end{bmatrix}. \tag{4.26}$$

The second step, namely the Radon transform (equation 4.14), remains the same in both acoustic and elastic case. Finally, the acoustic WLS objective function can be defined as

$$J_{\zeta_{I_p},\zeta_\rho} = \frac{1}{2}\left\| \Big(2\zeta_{I_p}(\mathbf{x}) - 2\sin^2(\theta)\zeta_\rho(\mathbf{x})\Big) - \mathcal{R}_{\delta m}(\mathbf{x},\theta)\right\|_W^2. \tag{4.27}$$

In a similar way, the optimal ζ_{I_p} and ζ_ρ can be obtained by setting the gradient $\partial J/\partial(\zeta_{I_p}, \zeta_\rho) = 0$ (Farshad and Chauris, 2020a). Subsequently, the matrices $\mathbf{M}$ and $\mathbf{\Phi}$ corresponding to equation 4.17 reduces to

$$\mathbf{M} = \begin{bmatrix} 4 & -4\sin^2(\theta) \\ -4\sin^2(\theta) & 4\sin^4(\theta) \end{bmatrix}, \tag{4.28}$$

and

$$\mathbf{\Phi} = \begin{bmatrix} 2\mathcal{R}_{\delta m}(\mathbf{x}, \theta) \\ -2\sin^2(\theta)\mathcal{R}_{\delta m}(\mathbf{x}, \theta) \end{bmatrix}. \tag{4.29}$$

Accordingly, this is a separable bivariate curve-fitting problem, implying an only dimensional difference between acoustic (2×2) and elastic (3×3) WLS methods.

Chapter 5

Sparsity-promoting regularization

Contents

5.1 Abstract **159**
5.2 Introduction **159**
5.3 Theory **162**
5.3.1 Forward Born modeling 162
5.3.2 Pseudoinverse Born operator 163
5.3.3 Extension to variable density 164
5.4 Numerical experiments **167**
5.4.1 Discontinuous interfaces 168
5.4.2 The Marmousi2 model 172
5.4.3 2004 BP salt model 179
5.5 Discussion **187**
5.6 Conclusion **189**
5.7 Acknowledgments **190**

Résumé

Ce chapitre comprend l'article intitulé « Sparsity-promoting multiparameter pseudoinverse Born inversion in acoustic media », publié dans la revue *Geophysics* (Farshad and Chauris, 2021c).
La migration itérative (« least-squares reverse-time migration » en anglais) est la méthode de choix pour l'imagerie sismique quantitative. Le plus gros inconvénient est que l'approche demande beaucoup de cycles de migration/modélisation. La convergence de la migration itérative peut être accélérée avec un préconditioner adapté. Dans le contexte du domaine étendu et d'un milieu acoustique à densité variable, le pseudoinverse est le préconditioner recommandé. Il permet d'obtenir des résultats quantitatifs en une seule itération. Cette méthode fonctionne en deux étapes : application du pseudoinverse puis inversion des deux paramètres après transformée de Radon et résolution d'un petit problème inverse. Comme attendu, des artefacts se produisent à cause de l'acquisition limitée. Nous présentons une inversion vitesse – densité, avec des contraintes ℓ_1 sur chacun des paramètres pour supprimer les artefacts. L'algorithme de « shrinkage-thresholding » est utilisé. Dans une approche classique, les contraintes ℓ_1 affecteraient tout le processus d'imagerie. Ici, comme l'imagerie est séparée en deux étapes, seulement la partie avec la transformée de Radon est modifiée ; elle n'implique pas de calculs coûteux de champ d'onde. Au travers d'expériences numériques, nous vérifions la robustesse de l'approche vis-à-vis de différents artefacts comme les couplages entre les paramètres, les interfaces tronquées, les géométries d'acquisition avec peu de points de tir, les données bruitées et les forts contrastes de vitesse.

Sparsity-promoting multiparameter pseudoinverse Born inversion in acoustic media
Milad Farshad and Hervé Chauris
Geophysics, 2021, **86**(3), pages S205–S220
doi: 10.1190/geo2020-0527.1

5.1 Abstract

Least-squares reverse-time migration has become the method of choice for quantitative seismic imaging. The main drawback of such scheme is that it requires many migration/modeling cycles. The convergence of least-squares reverse-time migration can be accelerated by using a suitable preconditioner. In the context of extended domain in a variable density acoustic media, the pseudoinverse Born operator is the recommended preconditioner, providing quantitative results within a single iteration. This method consists of two steps: application of the pseudoinverse Born operator, and inversion of two parameters using an efficient weighted least-squares approach based on the Radon transform. As expected, cross-talk artifacts are generated in the second step due to limited acquisition. We present a variable density pseudoinverse Born operator constrained with the ℓ_1-norm for each model parameter to suppress the artifacts. The fast iterative shrinkage-thresholding algorithm is used to carry out the optimization problem. In classical iterative least-squares migration, the ℓ_1-norm constraints would affect the whole imaging process. As the imaging method is split into two steps, only the Radon transform part is modified, where no wave-based operators are involved. Through numerical experiments, we verify the robustness of the proposed method against different migration artifacts including the parameter cross-talk, interfaces with abrupt truncations, sparse shot acquisition geometry, noisy data and high contrast complex structures.

5.2 Introduction

Seismic migration algorithms seek to provide a subsurface image that accurately characterizes the Earth structure with reliable physical properties. Typically, a standard migration operator defined as the adjoint of the forward modeling operator, i.e. the adjoint Born modeling operator, only accounts for the kinematic but not the dynamic effects. In practice, various factors such as noise, aliasing, limited recording aperture, narrow frequency bandwidth and under-sampled acquisition geometry, contribute to the non-quantitative nature of the adjoint operator, resulting in low-resolution images containing artifacts (Baysal et al., 1983; Nemeth et al., 1999).

To partially remedy these problems, least-squares migration (LSM), also known as linearized seismic inversion, was proposed: migration is reformulated as a linear inverse problem with the objective to minimize the differences between the simulated and the observed data (Nemeth et al., 1999). LSM has been applied either using ray-based Kirchhoff migration (Nemeth et al., 1999; Duquet et al., 2000), one-way wave-equation migration (Kühl and Sacchi, 2003; Wang and Sacchi, 2007), or two-way wave-equation migration also known as reverse-time migration (RTM) (Dai et al., 2012; Zeng et al., 2014; Zhang and Schuster, 2014; Zhang et al., 2015b). Note that the quality of the final image obtained via LSM also depends on the initial velocity model (Symes, 2008b).

The most significant drawback of LSM is its computational cost as it is solved iteratively. Thus, for a fixed number of shots, LSM is $2N$ times (modeling and adjoint) more expensive than the conventional migration, where N is the number of iterations required for LSM. Different techniques have been developed to possibly accelerate the convergence rate of LSM, such as multi-source approach with either random-phase encoding or plane-wave encoding (Liu et al., 2006; Dai et al., 2012; Xue et al., 2016), and scaling the linear system with proper preconditioners such as the approximate inverse normal operator, also known as Hessian operator (Shin et al., 2001; Rickett, 2003; Herrmann et al., 2009; Huang et al., 2016).

Recently, in the context of extended-domain methods, different pseudoinverse operators have been proposed via asymptotic analysis providing quantitative estimations within a single iteration (ten Kroode, 2012; Hou and Symes, 2015, 2017; Chauris and Cocher, 2017). Although these operators are obtained under high-frequency approximation, they do not contain any ray quantities in their final formula (Hou and Symes, 2015, 2017; Chauris and Cocher, 2017). The pseudoinverse formulas are strictly valid in the absence of turning waves, and they are derived assuming infinite acquisition geometries. In practice, the latter is not really restrictive as they provide accurate results within the migration aperture, whereas including turning waves would require introducing of vertical subsurface offsets (Biondi and Shan, 2002; Biondi and Symes, 2004). For that reason, these pseudoinverse operators are often used as a preconditioner to accelerate the convergence rate (Hou and Symes, 2016a; Li and Chauris, 2018). In practice, the other more important limitation is the restriction to constant-density acoustic media.

In the oil and gas industry, density as well as velocity play an important role in seismic interpretation steps. Very recently, two approaches have been proposed by Dafni and Symes (2018) and Farshad and Chauris (2020a) to extend the applicability of the pseudoinverse Born operator to variable density acoustic media. The first step consists in applying the same pseudoinverse Born operator as the one for the constant-density acoustic case. In the second phase, the Radon transform is applied to map the imaging angle. The work of Dafni and Symes (2018) shares the same idea as Zhang et al. (2014b), who proposed to invert the acoustic parameters by using two traces (two-trace method) from the angle-domain common-image gathers (CIGs) estimated either by LSRTM or by application of the pseudoinverse Born operator. Farshad and Chauris

(2020a) show that the selection of the two traces is important, and proposed a generalization by taking into account all traces in the angle-domain CIGs using an efficient weighted least-squares (WLS) approach. This generalization not only yielded more robust results compared to the two-trace method, but also provided the flexibility to include more constraints through the formulation as an inverse problem (Farshad and Chauris, 2020a). As expected, the multi-parameter inversion suffers from cross-talk artifacts (Virieux and Operto, 2009; Operto et al., 2013; Prieux et al., 2013). To mitigate this problem, some a priori information about the model can be included in the inversion process as a regularization term (Gholami and Siahkoohi, 2010; Asnaashari et al., 2013).

In recent years many regularization techniques such as ℓ_0-norm (Gholami and Farshad, 2019b), ℓ_1-norm (Herrmann and Li, 2012; Wu et al., 2016), Tikhonov (Tikhonov et al., 2013; Asnaashari et al., 2013), total-variation (Guitton, 2012; Aghamiry et al., 2019a; Zand et al., 2020), hybrid Tikhonov and total-variation (Gholami and Hosseini, 2013; Lin and Lianjie, 2015; Aghamiry et al., 2018, 2019b; Ren and Li, 2020), and shaping regularization (Fomel, 2007; Xue et al., 2016; Yao et al., 2019) have been successfully applied in many areas of signal processing and seismic imaging. These regularization terms enforce some a priori assumptions about the shape and statistical characteristics of the subsurface. For example, a subsurface that as expected to contain smoothly varying components (e.g. background model), is characterized by a normal prior and therefore smooth regularization techniques are employed. On the other hand, a subsurface that is expected to contain sparse structure (e.g. reflectivity model), is characterized by a long-tail prior thus requiring sparse regularization techniques (Polson and Sokolov, 2019).

In this paper, we propose to include the ℓ_1-norm to each model parameter as the regularization terms in the second phase of the WLS-based pseudoinverse Born operator, namely within the Radon transform step. The objective is to stabilize the inverse problem. More precisely, the role of regularization terms is to mitigate noise and artifacts in the images to finally obtain sparsity-promoted images even in presence of highly decimated data (Herrmann and Li, 2012; Wu et al., 2016; Aghamiry et al., 2019b). We assess the robustness of the proposed method against various migration artifacts associated to parameter cross-talk, abrupt truncation of interfaces, random noise, sparse shot acquisition and complex structures with possibly large contrasts. This work should be understood as a new step for the applicability of multi-parameter LSRTM. By applicability, we mean the derivation of proper preconditioners and the possibility to get quantitative LSRTM results within a few iterations only.

The outline of this paper is as follows. First, we review the extended domain Born modeling in a constant density acoustic Earth, its corresponding pseudoinverse, and the extension to variable density media. Next, we add sparsity constraints to the variable density pseudoinverse Born operator. Then, we further demonstrate the effectiveness of the multi-parameter inversion with or without sparsity constraints on a discontinuous interfaces, Marmousi2 and 2004 BP models (Billette and Brandsberg-Dahl, 2004).

5.3 Theory

In the following, we first briefly review the concepts and formulas of the forward/pseudoinverse Born operators in a constant density acoustic media. Then, we review the extension of the pseudoinverse operator to a variable density acoustic media using the WLS method. Finally, we combine the regularization terms with WLS objective function. The two most important aspects are: (1) the total imaging process is split into two parts where only the first part contains the wavefield simulations; and (2) the second part is also set up as an inverse problem, providing the possibility to introduce constraints.

5.3.1 Forward Born modeling

The basic element of the linearized seismic imaging technique is the Born approximation. Under this approximation, the exact model $m(\mathbf{x})$ in 2D at position $\mathbf{x} = (x, z)$ is the sum of a background model $m_0(\mathbf{x})$ and a perturbation model $\delta m(\mathbf{x})$ as (Symes, 2008b):

$$m(\mathbf{x}) = m_0(\mathbf{x}) + \delta m(\mathbf{x}), \tag{5.1}$$

where m_0 controls the kinematics of the wave propagation while δm generates the diffracted/reflected waves. Here, we consider a 2D constant density acoustic Earth model parameterized with $m = \beta$ as the inverse of the Bulk modulus $\beta = 1/\rho v^2$, where v and ρ denote the velocity and density at spatial position $\mathbf{x}$, respectively. Under the Born approximation, the recorded data wavefield $d(\mathbf{x_s}, \mathbf{x_r}, \omega)$ at the receiver position $\mathbf{x_r} = (x_r, z_r)$ from a source positioned at $\mathbf{x_s} = (x_s, z_s)$ can be expressed in an integral form as:

$$\begin{aligned} d(\mathbf{x_s}, \mathbf{x_r}, \omega) &= L_{[m_0, \delta m]}(\mathbf{x_s}, \mathbf{x_r}, \omega) \\ &= -(i\omega)^2 \Omega(\omega) \int \mathrm{d}\mathbf{x}\, G_0(\mathbf{x_s}, \mathbf{x}, \omega) \delta m(\mathbf{x}) G_0(\mathbf{x}, \mathbf{x_r}, \omega), \end{aligned} \tag{5.2}$$

where ω is the angular frequency, L is the Born modeling operator, Ω is the source spectrum, and G_0 is the reference Green's function derived in a given smooth model m_0. The imaging condition can be generalized by introducing a redundancy parameter, which could be a time and/or space shift (Sava and Fomel, 2006; Symes, 2008b). Here, we consider a horizontal shift referred to as subsurface offset $\mathbf{h} = (h, 0)$. This extension allows to handle the errors in the background models by matching the image dimension (x, z, h) as same as the data dimension (s, r, t) (Symes, 2008b). In this case, the forward Born modeling operator in the extended domain can be written as (Symes, 2008b):

$$\begin{aligned} d(\mathbf{x_s}, \mathbf{x_r}, \omega) &= L_{[m_0, \delta m]}(\mathbf{x_s}, \mathbf{x_r}, \omega) \\ &= -(i\omega)^2 \Omega(\omega) \int \mathrm{d}\mathbf{x}\, \mathrm{d}\mathbf{h}\, G_0(\mathbf{x_s}, \mathbf{x} - \mathbf{h}, \omega) \delta m(\mathbf{x}, \mathbf{h}) G_0(\mathbf{x} + \mathbf{h}, \mathbf{x_r}, \omega). \end{aligned} \tag{5.3}$$

The difference between classic (equation 5.2) and extended Born modeling (equation 5.3) is the extra loop over $\mathbf{h}$. In this case, the physical model $\delta m(\mathbf{x})$ is extended to a non-physical model $\delta m(\mathbf{x}, \mathbf{h})$ where the receiver and source Green's functions are evaluated at $(\mathbf{x} + \mathbf{h})$ and $(\mathbf{x} - \mathbf{h})$ positions, respectively. In the next section, we review how to derive a pseudoinverse formula for equation 5.3.

5.3.2 Pseudoinverse Born operator

The classic migration operator in a specific background model m_0 is formulated as (Claerbout, 1992):

$$\delta m^{mig} = L^T d^{obs}, \tag{5.4}$$

where δm^{mig} is the migrated image, d^{obs} is the observed data, and ^{T} denotes the transpose operator, leading L^T to be the adjoint of the Born operator. The adjoint Born operator accounts for the traveltime of the observed data, but not for the amplitude. The true amplitude image of the subsurface (δm) can be iteratively estimated by minimizing a misfit function:

$$J_0(\delta m) = \frac{1}{2} \left\| d(\delta m) - d^{obs} \right\|_2, \tag{5.5}$$

where $\|.\|_p$ denotes the ℓ_p-norm. The optimal solution is obtained when the gradient $\partial J_0(\delta m)/\partial \delta m$ is equal to zero, such that (Lailly, 1983; Tarantola, 1984):

$$\delta m = (L^T L)^{-1} \delta m^{mig}. \tag{5.6}$$

Here, $L^T L$ is the Hessian matrix. Since the Born operator is not unitary, the Hessian matrix is not equal to the Identity matrix. Therefore, the migrated δm^{mig} is considered as a Hessian-blurred version of the optimal δm (Schuster, 1993; Nemeth et al., 1999; Chavent and Plessix, 1999). The two possibilities to obtain the true amplitude image δm are: (1) iteratively minimizing equation 5.5, i.e., least-squares migration (LeBras and Clayton, 1988; Nemeth et al., 1999), (2) replacing the adjoint (L^T) by its pseudoinverse $(L^\dagger = (L^T L)^{-1} L^T)$ and directly applying it to d^{obs}. The latter achieves quantitative properties within a single iteration (ten Kroode, 2012; Hou and Symes, 2015, 2017; Chauris and Cocher, 2017). The derivation of the pseudoinverse is performed such that the application of the $LL^\dagger$ on any data, accurately reconstructs the data, even in an incorrect background model. Here, we follow the work by Chauris and Cocher (2017), proposing the pseudoinverse extended Born operator in a constant density acoustic media as:

$$\delta m \quad = \quad L^\dagger (d^{obs})(\mathbf{x}, \mathbf{h}) \tag{5.7}$$

$$\simeq 32\frac{\beta_0}{\rho_0^3}\partial_z \int \mathrm{d}\mathbf{x_s}\,\mathrm{d}\mathbf{x_r}\,\mathrm{d}\omega\,\frac{\Omega^\dagger(\omega)}{(i\omega)}\partial_{s_z}G_0^*(\mathbf{x_s},\mathbf{x}-\mathbf{h},\omega)d^{obs}(\mathbf{x_s},\mathbf{x_r},\omega)\partial_{r_z}G_0^*(\mathbf{x}+\mathbf{h},\mathbf{x_r},\omega),$$

where * denotes the complex conjugate and $\Omega^\dagger$ is the inverse of the seismic wavelet. This formula is an asymptotic inverse and is only valid in the extended domain for constant density acoustic media.

In a more realistic acoustic medium, the density variations cannot be ignored. In such a case, the velocity variations govern the amplitude and kinematics of the reflections, whereas the density variations are mainly related to amplitude of the reflections (Virieux and Operto, 2009). Farshad and Chauris (2020a) showed that if the data (d^{obs}) contain density variations, the application of equation 5.7 results in energy distortion in the CIGs, even if the investigated background models are correct (see Farshad and Chauris, 2020a, their Figure 4). In the next section, we describe how equation 5.7 can be extended to variable density acoustic media.

5.3.3 Extension to variable density

In the case of variable density, the inverted reflectivity model $\delta m(\mathbf{x},h)$ from equation 5.7 can be decomposed into two different parameters such that (Dafni and Symes, 2018):

$$\frac{\delta m}{\beta_0}(\mathbf{x},h) \cong \mathcal{D}_\beta\frac{\delta\beta}{\beta_0}(\mathbf{x})\delta(h)+\mathcal{D}_\rho\frac{\delta\rho}{\rho_0}(\mathbf{x})\delta(h), \tag{5.8}$$

where $\delta()$ is the Dirac delta distribution, $\mathcal{D}_\beta = 1$ and $\mathcal{D}_\rho = -\cos(2\theta)$ denote the diffraction patterns of β and ρ (Figure 5.1), respectively (Forgues, 1996). Note that $\delta\beta$ and $\delta\rho$ are physical parameters (they only depend on $\mathbf{x}$, and not h), while δm is said to be extended. In equation 5.8, the $\mathcal{D}_\rho = -\cos 2\theta$ term is not a simple multiplication but acts as a pseudo-differential operator (Dafni and Symes, 2018). The parameter class (β,ρ) in equation 5.8 can be replaced by any acoustic parameter, i.e. P-wave velocity V_p or P-wave impedance I_p, by choosing the corresponding proper diffraction pattern $\mathcal{D}$. The only remaining ingredient to decompose δm into $\delta\beta$ and $\delta\rho$ is the estimation of θ.

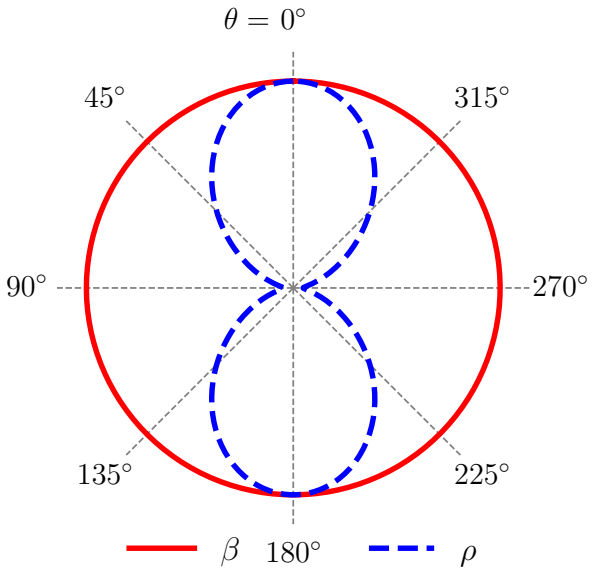

Figure 5.1: The diffraction pattern for $(\beta - \rho)$ parameterization.

Qin and Lambaré (2016) proposed to estimate θ by using an iterative tomographic ray tracing approach. Farshad and Chauris (2020a) proposed an efficient least-squares approach as a generalization of the method proposed by Dafni and Symes (2018), by extending the pseudoinverse Born operator from constant- to variable-density acoustic media. This method, referred to as WLS, is based on two steps. In the first step, we apply the Radon transform, which results in transforming the offset-domain to the angle-domain CIGs $\mathcal{R}_{\delta m}(\mathbf{x}, \theta)$ such that (Sava and Fomel, 2003)

$$\mathcal{R}_{\delta m}(\mathbf{x}, \theta) = \frac{1}{\beta_0} \int \mathrm{d}h \, \delta m(x, z + h \tan\theta, h), \tag{5.9}$$

where θ is the diffraction angle. This gives the angle θ needed to split the inverted δm into two different parameters (equation 5.8). We refer to ten Kroode (2012) for a detailed discussion regarding the relationship between the angle θ at the imaging point (ray-based quantity) and in the Radon transform. In the last step, we minimize a cost function defined as the least-squares differences between the observed and the computed angle-domain response. The WLS method solves the following optimization problem (Farshad and Chauris, 2020a):

$$\min_{\delta\beta, \delta\rho} \left(\frac{1}{2} \left\| \mathcal{R}_{\delta m}(\mathbf{x}, \theta) - \left(\frac{\delta\beta}{\beta_0}(\mathbf{x}) - \frac{\delta\rho}{\rho_0}(\mathbf{x}) \cos(2\theta) \right) \right\|_W^2 \right), \tag{5.10}$$

where $W(\mathbf{x}, \theta)$ is a weighting mask defined as:

$$W(\mathbf{x}, \theta) = \begin{cases} 1, & \text{if } |\theta| \leq \alpha \tan^{-1}(\frac{x_{max}}{z}) \\ 0, & \text{otherwise} \end{cases}, \tag{5.11}$$

and x_{max} denotes the maximum surface offset. Based on the value of α, the objective of this weighting mask is to remove the artifacts in the angle domain which are due to the limited acquisition geometry (Sava and Fomel, 2003; Farshad and Chauris, 2020a). This definition is strictly valid for homogeneous models. A short discussion on the choice of this parameter is provided below. By definition, the WLS method (equation 5.10) is a separable bivariate curve fitting problem: the optimization can be applied for each spatial coordinate separately. It should be noted that the estimated $\delta\beta(\mathbf{x})$ and $\delta\rho(\mathbf{x})$ are non-extended physical parameters, whereas $\delta m(\mathbf{x}, h)$ and $\mathcal{R}_{\delta m}(\mathbf{x}, \theta)$ are said to be extended by non-physical parameters h and θ. For simplicity, the solution in a least-squares sense can be written as $\mathbf{G}\delta\mathbf{u} = \mathbf{R}$, where

$$\mathbf{G} = \begin{bmatrix} \int \mathrm{d}\theta\, W(\mathbf{x}, \theta) & -\int \mathrm{d}\theta\, W(\mathbf{x}, \theta) \cos(2\theta) \\ -\int \mathrm{d}\theta\, W(\mathbf{x}, \theta) \cos(2\theta) & \int \mathrm{d}\theta\, W(\mathbf{x}, \theta) \cos^2(2\theta) \end{bmatrix}, \tag{5.12}$$

$$\delta\mathbf{u} = \begin{bmatrix} \frac{\delta\beta}{\beta_0}(\mathbf{x}) \\ \frac{\delta\rho}{\rho_0}(\mathbf{x}) \end{bmatrix}, \tag{5.13}$$

and

$$\mathbf{R} = \begin{bmatrix} \int \mathrm{d}\theta\, W(\mathbf{x}, \theta) \mathcal{R}_{\delta m}(\mathbf{x}, \theta) \\ -\int \mathrm{d}\theta\, W(\mathbf{x}, \theta) \cos(2\theta) \mathcal{R}_{\delta m}(\mathbf{x}, \theta) \end{bmatrix}. \tag{5.14}$$

It is well known that adding more parameters into the inversion increases the ill-posedness of the inverse problem (Virieux and Operto, 2009). Consequently, a leakage between inverted parameters is expected, which results in non-physically meaningful subsurface images (Dafni and Symes, 2018; Farshad and Chauris, 2020a). Moreover, incomplete illumination, aliasing artifacts, undersampled or noisy data can also lead to subsurface images that are degraded in quality due to overfitting of the noise. Therefore, it is essential to incorporate some regularization terms into the inversion process allowing more accurate representations of the subsurface reflectivity model. Since the

reflectivity model is sparse (Wu et al., 2016; Zhang et al., 2017), we propose to add independently ℓ_1-norm constraints to each inverted parameter as regularization terms in the WLS method (equation 5.10). Since the ℓ_1-norm regularization penalizes small coefficients more severely than large ones, it will enforce sparse subsurface images. Therefore, the constrained optimization problem becomes

$$\min_{\delta\beta,\delta\rho}\left(\frac{1}{2}\left\|\mathcal{R}_{\delta m}(\mathbf{x},\theta)-\left(\frac{\delta\beta}{\beta_0}(\mathbf{x})-\frac{\delta\rho}{\rho_0}(\mathbf{x})\cos(2\theta)\right)\right\|_W^2+\lambda_1\left\|\frac{\delta\beta}{\beta_0}(\mathbf{x})\right\|_1+\lambda_2\left\|\frac{\delta\rho}{\rho_0}(\mathbf{x})\right\|_1\right) \quad (5.15)$$

where $\lambda_1, \lambda_2 > 0$ are the trade-off parameters governing the balance between the data misfit and regularization terms. We provide some guidelines regarding how to tune these trade-off parameters below. We refer to our method as the regularized WLS (RWLS) method, which is a sparsity-promoting separable bivariate curve fitting problem. The RWLS method has an interesting interpretation: among the two parameters $\delta\beta(\mathbf{x})$ and $\delta\rho(\mathbf{x})$, the optimization selects the one that has stronger effect in predicting the data and thus this can reduce the cross-talk artifacts. In order to carry out the optimization problem described in equation 5.15, we use a popular sparse solver based on the soft thresholding method called fast iterative shrinkage-thresholding algorithm (FISTA; Beck and Teboulle, 2009), which is the generalized version of ISTA (Daubechies et al., 2004). Note that the only step in the RWLS method that contains wavefield propagation is the application of the pseudoinverse Born operator, which is only applied once to obtain δm (equation 5.7). The other steps include single application of the Radon transform (equation 5.9), and the iterative FISTA.

The proposed RWLS method is summarized in Algorithm 5, where γ is the appropriate step size for the iterative scheme. For each spatial position in the physical domain, $\mathbf{G}$ and $\mathbf{R}$ are precomputed 2×2 matrices. The stopping criterion in this algorithm is defined by the maximum number of iterations. The FISTA algorithm mainly consists of three steps: (1) calculation of the gradient direction (line 9 in Algorithm 5), (2) application of soft-threshold (shrinkage) operator (lines 10 and 11 in Algorithm 5), which is basically equivalent to shifting $\delta\beta$ and $\delta\rho$ towards zero with λ_1 and λ_2 units, (3) updating $\delta\mathbf{u}^{k+1}$ using a specific linear combination of the previous two $\delta\mathbf{u}^k$ and $\delta\mathbf{u}^{k-1}$ (lines 12 and 13 in Algorithm 5). Notice that the last step generates a momentum that accelerates the convergence rate. This is indeed the difference between ISTA and FISTA.

5.4 Numerical experiments

In this section we assess the performance of the ℓ_1-regularized pseudoinverse Born operator through different numerical experiments. For all these examples, the modeling engine for the solution of the 2-D acoustic wave equation is the finite-difference method implemented with perfectly matched layers (PMLs) absorbing boundary conditions. In

Algorithm 5: RWLS-based psuedoinverse Born operator by FISTA

Input: d^{obs}, λ_1, λ_2
Initialization
  $\delta m = L^{\dagger}(d^{obs})$
  $\mathbf{R} \leftarrow$ calculate $\mathcal{R}_{\delta m}$
  $\delta \mathbf{u}^1 = \delta \mathbf{u}^0 = 0$
  $t^1 = 1$
  $\gamma \geq$ (maximum eigenvalue of $\mathbf{G}^T\mathbf{G}$)
for $k = 1$ **to** *maximum iteration* **do**
  $\delta \mathbf{u}^k = \delta \mathbf{u}^k - \frac{1}{\gamma}\mathbf{G}^T(\mathbf{G}\delta \mathbf{u}^k - \mathbf{R})$
  $\frac{\delta\beta}{\beta_0} = \frac{\delta\beta/\beta_0}{|\delta\beta/\beta_0|}\max(\frac{\delta\beta}{\beta_0} - \lambda_1, 0)$
  $\frac{\delta\rho}{\rho_0} = \frac{\delta\rho/\rho_0}{|\delta\rho/\rho_0|}\max(\frac{\delta\rho}{\rho_0} - \lambda_2, 0)$
  $t^{k+1} = \frac{1+\sqrt{1+4t^{k^2}}}{2}$
  $\delta \mathbf{u}^{k+1} = \delta \mathbf{u}^k + \left(\frac{t^k - 1}{t^{k+1}}\right)(\delta \mathbf{u}^k - \delta \mathbf{u}^{k-1})$

all numerical examples, we apply a smooth taper on a few shallow samples of the inverted δm to attenuate the undesirable numerical imprint of the source positions.

We start with a model containing discontinuous interfaces. This example is used to assess the effectiveness of the RWLS method in reducing the amplitude leakage due to the parameter cross-talk addressed by Dafni and Symes (2018) and Farshad and Chauris (2020a), and also spurious oscillations related to the truncated interfaces (Fei and Williamson, 2010; Shen and Symes, 2015; Chauris and Cocher, 2017). Then, the Marmousi2 model is used as a benchmark of complex geology. To check the effectiveness of the method with decimated acquisition geometries, the inversion is performed with dense and decimated shots. Finally, the 2004 BP salt model is investigated to assess the sensitivity of the RWLS method to random noise, and its potential to image high contrast complex structures.

5.4.1 Discontinuous interfaces

The first example is inspired by Shen and Symes (2015) and Chauris and Cocher (2017): it consists of four interfaces with a limited lateral extension in the velocity model (Figure 5.2a), and only one truncated interface in the density model (Figure 5.2b). The model is discretized into 170 (vertical direction) by 301 (horizontal direction) grid points with a 16 m grid interval. A Ricker wavelet centered at 4.84 Hz (maximum frequency is 12.1 Hz) is used as a source, and the sources are excited at the surface from 0 to 4.8 km with a 37.5 m interval. The receivers for each shot are positioned at every grid position

and record for 2.6 s with a 3.8 ms time interval.

We start by applying the pseudoinverse extended Born operator (equation 5.7) and the Radon transform (Figure 5.3). Unwanted oscillations around the reflector positions are visible in both the subsurface offset- and angle-domain CIGs, due to the abrupt truncation of the interfaces. The true perturbation models (Figure 5.4a and 5.4b) are compared to the inverted parameters estimated via the WLS (Figure 5.4c and 5.4d) and the RWLS methods (Figure 5.4e and 5.4f). We note two types of artifacts in the results obtained via the WLS method: an amplitude leakage due to the parameter cross-talk, and oscillations around the edges of the interfaces due to the abrupt truncation (yellow ellipses in Figure 5.4c). In contrast, the regularization terms in the RWLS method lead to more robust and sparse results compared to the WLS method. Traces at different positions are extracted for a closer examination (Figure 5.5). The artifacts and the oscillations are efficiently removed with the RWLS method.

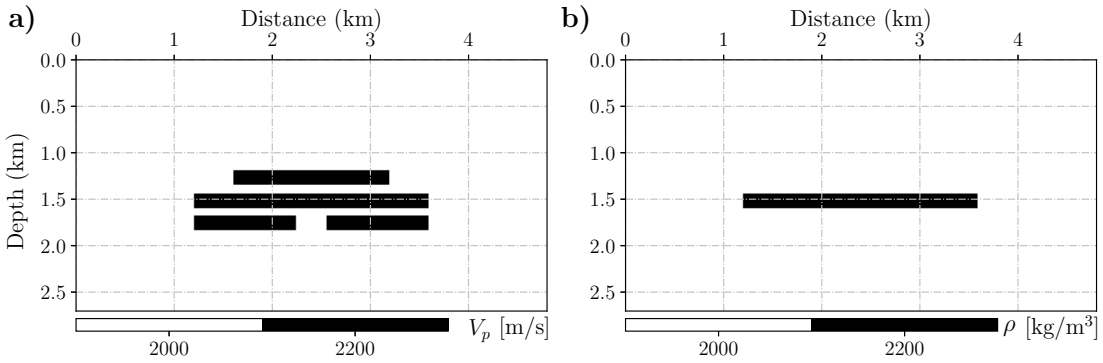

Figure 5.2: The exact a) V_p (m/s) and b) ρ (kg/m^3) models corresponding to discontinuous interfaces.

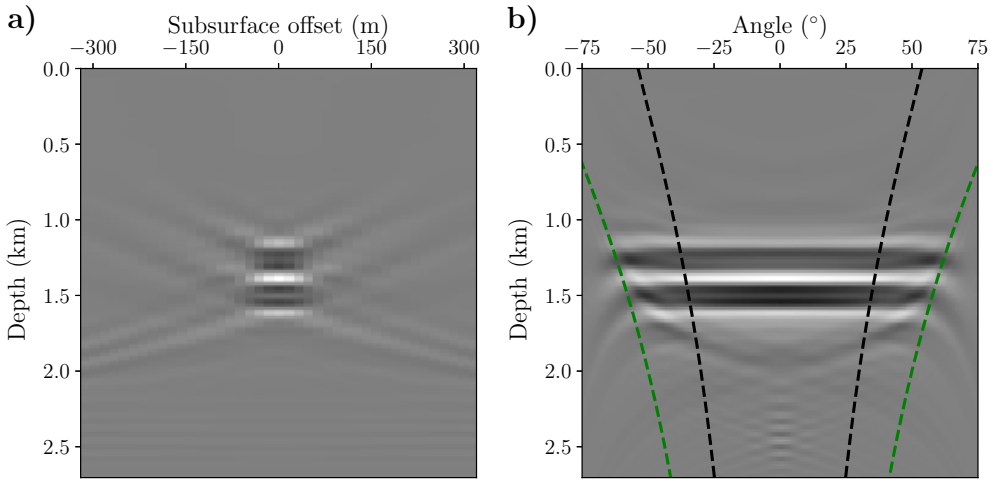

Figure 5.3: The a) offset-domain CIG (δm) and b) its angle-domain response corresponding to Figure 5.2. The green and black dashed lines define the W for $\alpha = 1.0$ and $\alpha = 0.6$, respectively.

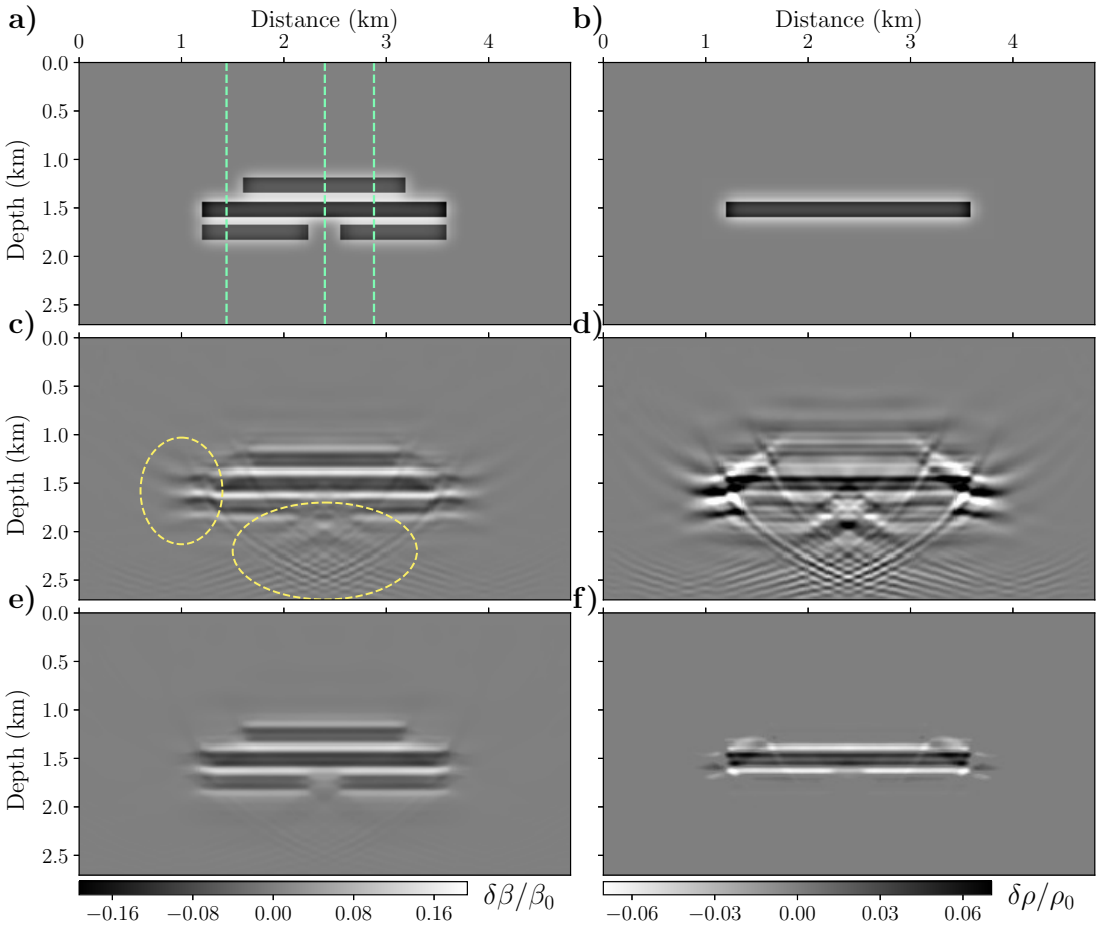

Figure 5.4: Comparison between the inverted parameters for discontinuous interfaces. Exact a) $\delta\beta/\beta_0$ and b) $\delta\rho/\rho_0$. Inverted c) $\delta\beta/\beta_0$ and d) $\delta\rho/\rho_0$ via the WLS method, and inverted e) $\delta\beta/\beta_0$ and f) $\delta\rho/\rho_0$ via the RWLS method. The green dashed lines in (a) indicate the positions of the extracted traces. The truncation artifacts are circled by the yellow ellipses in (c).

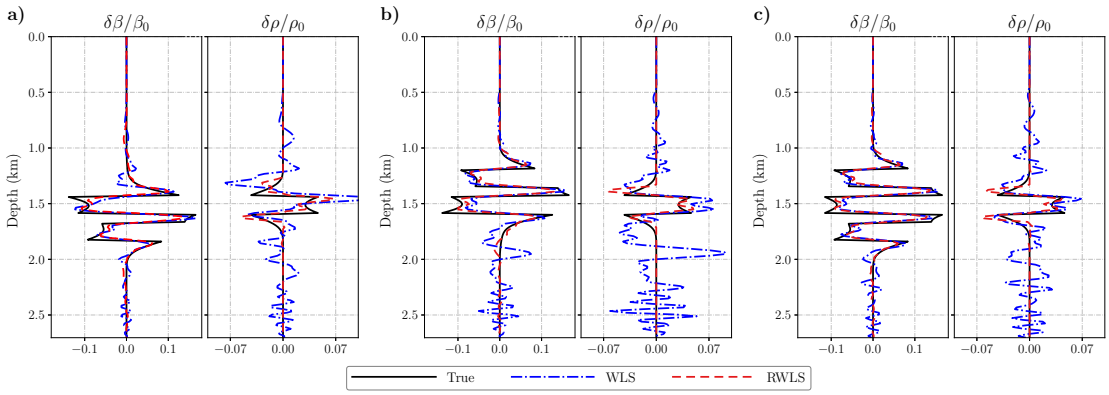

Figure 5.5: Comparison between the extracted traces from Figure 5.4 at a) x $= 1.4$ km, b) x $= 2.4$ km and c) x $= 2.9$ km.

5.4.2 The Marmousi2 model

As a benchmark for a structurally complex model, we use the Marmousi2 model composed of 214 (vertical direction) by 875 (horizontal direction) discretization points with a 12 m grid interval. To reduce the computational cost, only the middle part of the Marmousi2 model is used (Figure 5.6a and 5.6b). The perturbation model is obtained by subtracting the background model, obtained by applying a 60 m long 2D Gaussian filter on the exact model in both directions, from the exact model (Figure 5.6c and 5.6d). The wavelet for our simulation is a Ricker wavelet centered at 4.76 Hz (maximum frequency is 11.9 Hz). The recording time is 3.7 s with a 1.32 ms time interval.

5.4.2.1 Dense shots

In this first experiment, we use 292 shots located at the surface and evenly distributed, with a source interval of 36 m and a fixed receiver grid. We perform the WLS and RWLS methods to invert the perturbations. The number of iterations of FISTA for the RWLS method is set to 100. We first compare the observed data (Figure 5.7a) with the reconstructed data estimated with the WLS (Figure 5.7b) and the RWLS methods (Figure 5.7c). Two traces at two different offsets are extracted from each panel (Figure 5.7d). To assess the accuracy of our inversion schemes, the root-mean-square error (RMSE) is measured as

$$\mathrm{RMSE}(\kappa) = \sqrt{\sum_{i=1}^{K} \frac{(\kappa_i - \widetilde{\kappa_i})^2}{K}}, \tag{5.16}$$

where κ and $\widetilde{\kappa}$ are the observed and estimated data. We find that the RWLS method yields a RMSE around 25% lower than the WLS method.

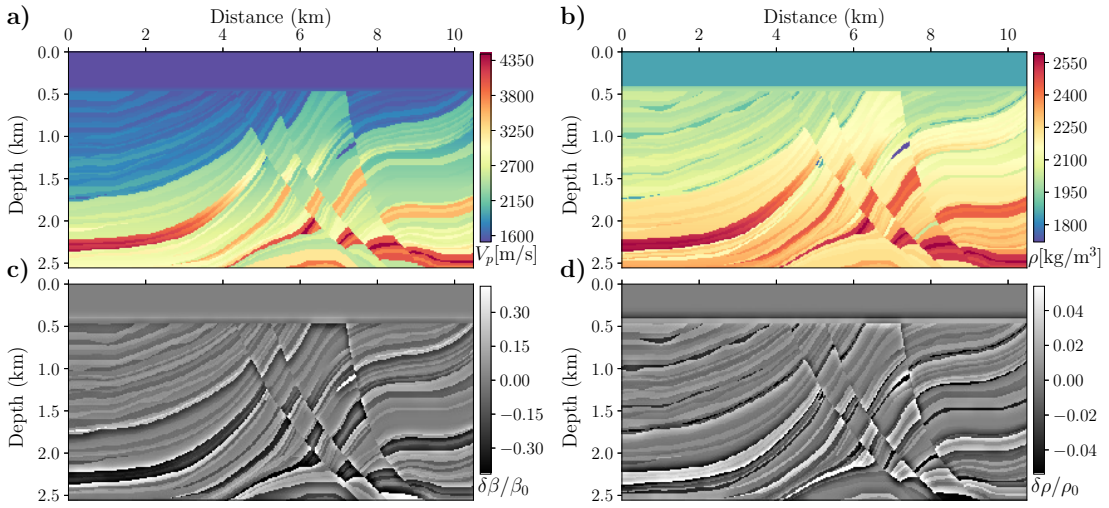

Figure 5.6: The exact a) V_p (m/s), b) ρ (kg/m^3), c) $\delta\beta/\beta_0$ and d) $\delta\rho/\rho_0$ models corresponding to the Marmousi2 model.

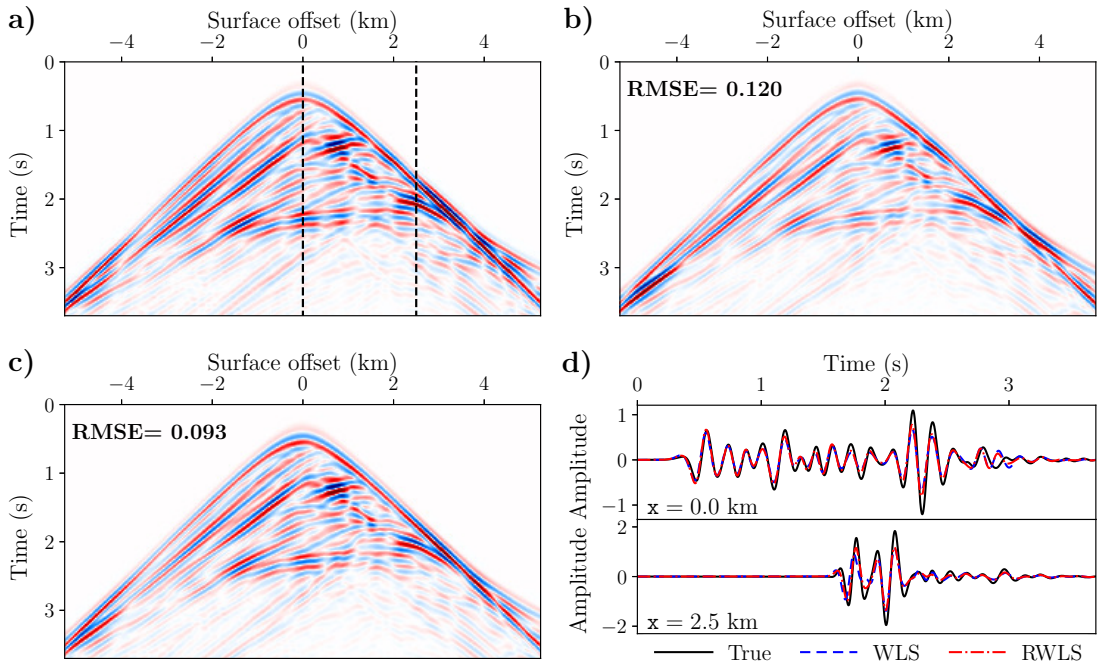

Figure 5.7: The Marmousi2 dense shots experiment. The a) observed data and reconstructed data for the shot at position 5.2 km via the b) WLS and c) RWLS methods. The dashed lines in (a) correspond to the d) extracted traces. Shots are plotted at the same scale.

Next, we compare the true perturbation models (Figure 5.6c and 5.6d) with the estimated ones (Figure 5.8). Note the scale discrepancy between the exact (Figure 5.6) and inverted (Figure 5.8) parameters, which reveals an underestimation in $\delta\beta$ and an overestimation in $\delta\rho$. The WLS models include low-frequency oscillations and energy distortions, especially for density (Figure 5.8b). This apparent higher resolution in Figure 5.8b may be misinterpreted as thin layers. The regularization clearly attenuates these artifacts (Figure 5.8d). In order to get more insight into the quality of these images, we calculate the normalized correlation coefficient (NCC) percentage as

$$\mathrm{NCC}(\%) = \frac{\sum_{\mathbf{x}} m_{true}(\mathbf{x}) m_{est}(\mathbf{x})}{\sqrt{\sum_{\mathbf{x}} m^2_{true}(\mathbf{x})}\sqrt{\sum_{\mathbf{x}} m^2_{est}(\mathbf{x})}} \times 100, \tag{5.17}$$

where m_{true} and m_{est} are the true and estimated models, respectively. The NCC scoring is easy to understand: the higher the value, the higher the correlation (maximum value is 100%). The NCC values corresponding to each method are written on each panel of Figure 5.8, showing the superiority of the RWLS method. To evaluate the qualitative reliability of the final images, we extract two traces at different positions ($x = 4.5$ and 7 km) from each panel of Figure 5.8, and sum them with the background models (Figure 5.9). The RMSE values between true and inverted traces are also calculated and written on each panel (Figure 5.9). Clearly, the RWLS method provides more accurate models.

5.4.2.2 Decimated shots

In this second experiment, we use 25 shots evenly spaced every 432 m. Note that the shot numbers are 12 times fewer than the previous experiment. The total computational cost is reduced by the same factor, as the wave-based operator (equation 5.7) dominates the computational burden. The number of FISTA iterations is set to be the same as with the first experiment (100 iterations). The reconstructed shots by using the inverted parameters are shown in Figure 5.10. Similar to the previous experiment, the RWLS method yields a lower error for the data reconstruction. In the image domain, the effect of the decimated shots are clearly visible on the inverted perturbations with the WLS method as discontinuous reflectors and distorted energy (Figure 5.11a and 5.11b). Contrary to the previous experiment, the structure of the Marmousi2 model is not visible, especially in the deep and central parts of β and ρ. However, the RWLS method in spite of the decimated shots obtains a reasonable level of accuracy and correlation, while reducing the artifacts (Figure 5.11c and 5.11d). The profiles extracted at $x = 4.5$ and 7 km confirm the quantitative reliability of the RWLS method (Figure 5.12). We further examine the performance of these methods by comparing wavenumber spectra (Figure 5.13). The color scales used for the panels corresponding to the WLS and RWLS methods are the same as the one used for the true models. The repetitive vertical lines in the wavenumber components obtained by the WLS method (indicated by red arrows in Figure 5.13b and

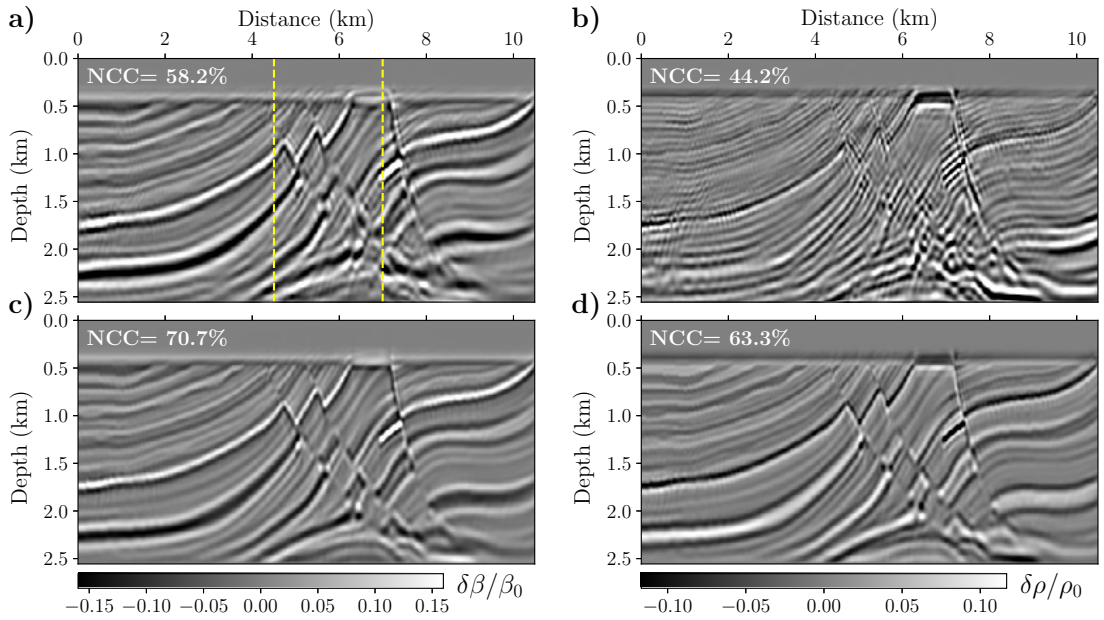

Figure 5.8: Comparison between the inverted parameters corresponding to dense shots (292 shots) for Marmousi2 model. Inverted a) $\delta\beta/\beta_0$ and b) $\delta\rho/\rho_0$ vis the WLS method, c) $\delta\beta/\beta_0$ and d) $\delta\rho/\rho_0$ via the RWLS method. The dashed lines in (a) correspond to the extracted traces.

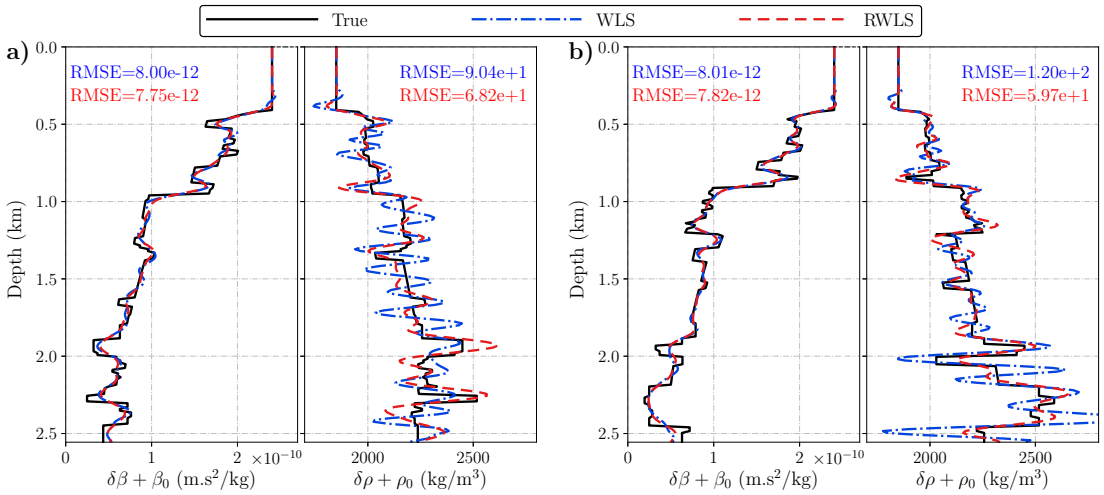

Figure 5.9: Comparison between the extracted inverted parameters corresponding to dense shots (292 shots) for Marmousi2 model. Inverted $\delta\beta + \beta_0$ and $\delta\rho + \rho_0$ for a) x $= 4.5$ km and b) x $= 7.0$ km.

5.13e) correspond to strong aliasing artifacts, and are associated to the decimated shots (Figure 5.11a and 5.11b). Additional artifacts (indicated by green box in Figure 5.13e) contaminate the spectrum obtained by the WLS method. On the contrary, the RWLS method significantly mitigates these problems, and provides wavenumber components closer to the true models (Figure 5.13c and 5.13f).

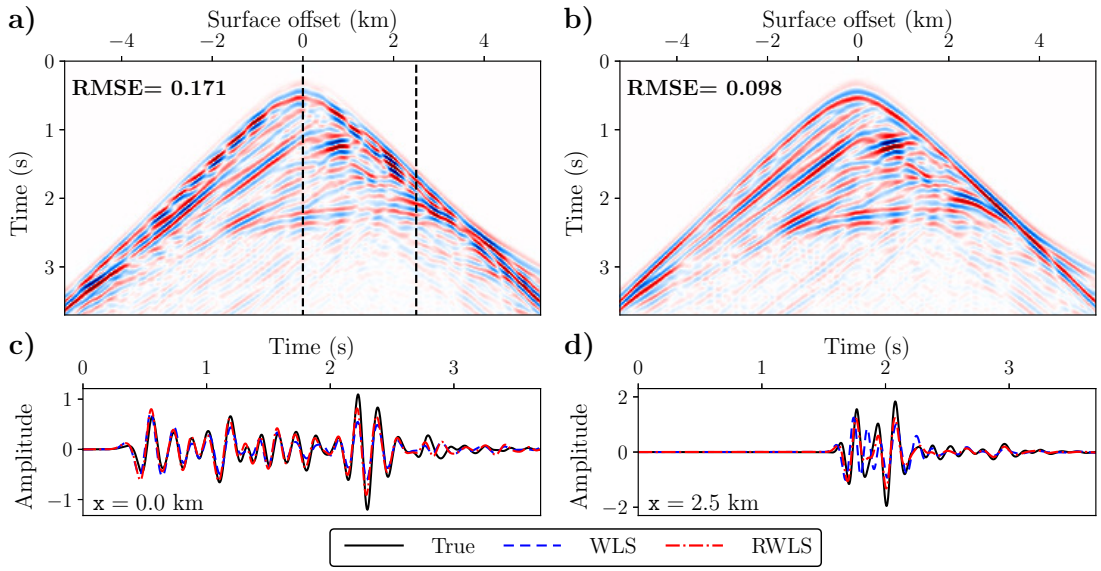

Figure 5.10: Same as for Figure 5.7, but for the decimated shots experiment. The reconstructed data for the shot at position 5.2 km via the a) WLS and b) RWLS methods. The dashed lines in (a) correspond to the extracted traces at surface offset c) 0.0 km and d) 2.5 km. Shots are plotted at the same scale as the observed shot (Figure 5.7a).

To increase the reliability of the results, this last experiment is performed for different source intervals, and the RMSE and NCC values are calculated in the data and image domains, respectively. As can be seen from the RMSE and NCC curves in Figure 5.14, the RWLS method always performs better than the WLS method. The data reconstruction shows a very stable behavior for the RWLS method, whereas an abrupt change for the WLS method is observed when the source spacing is above 360 m (gray dashed line in Figure 5.14b). This distance is about the propagated mean wavelength (330 m) in the shallow part of the model. The same behavior can be also observed for the NCC in the image domain (Figure 5.14a), which is quite stable for the RWLS method, or follows a decreasing trend with the WLS method. We draw the same dashed line as in Figure 5.14a. However, the change in the decreasing trend is not as pronounced as the one in the data domain. For this specific example, we first recommend to not go beyond a source spacing of around 330 m using the WLS method. But in all cases, we recommend to include the regularization: it achieves better results in both the image and data

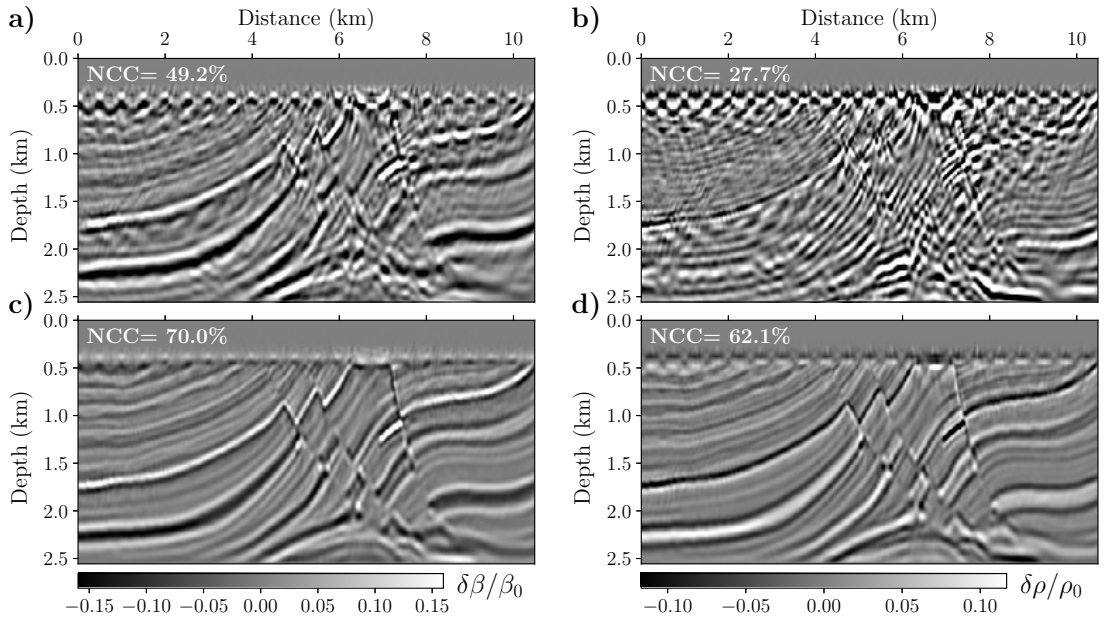

Figure 5.11: Same as for Figure 5.8, but for the decimated shots experiment. Inverted a) $\delta\beta/\beta_0$ and b) $\delta\rho/\rho_0$ via the WLS method, c) $\delta\beta/\beta_0$ and d) $\delta\rho/\rho_0$ via the RWLS method.

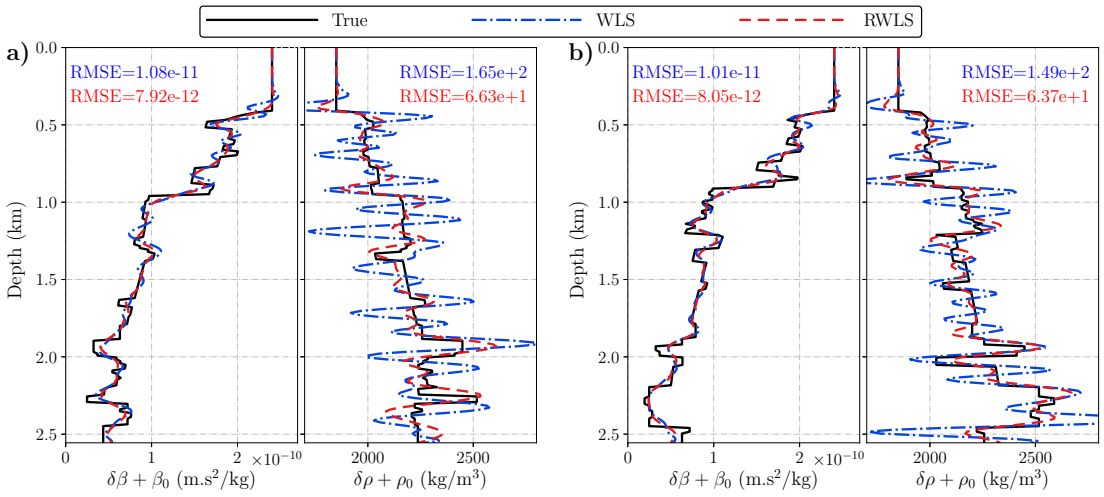

Figure 5.12: Same as for Figure 5.9, but for the decimated shots experiment. Inverted $\delta\beta + \beta_0$ and $\delta\rho + \rho_0$ for a) x $= 4.5$ km and b) x $= 7.0$ km.

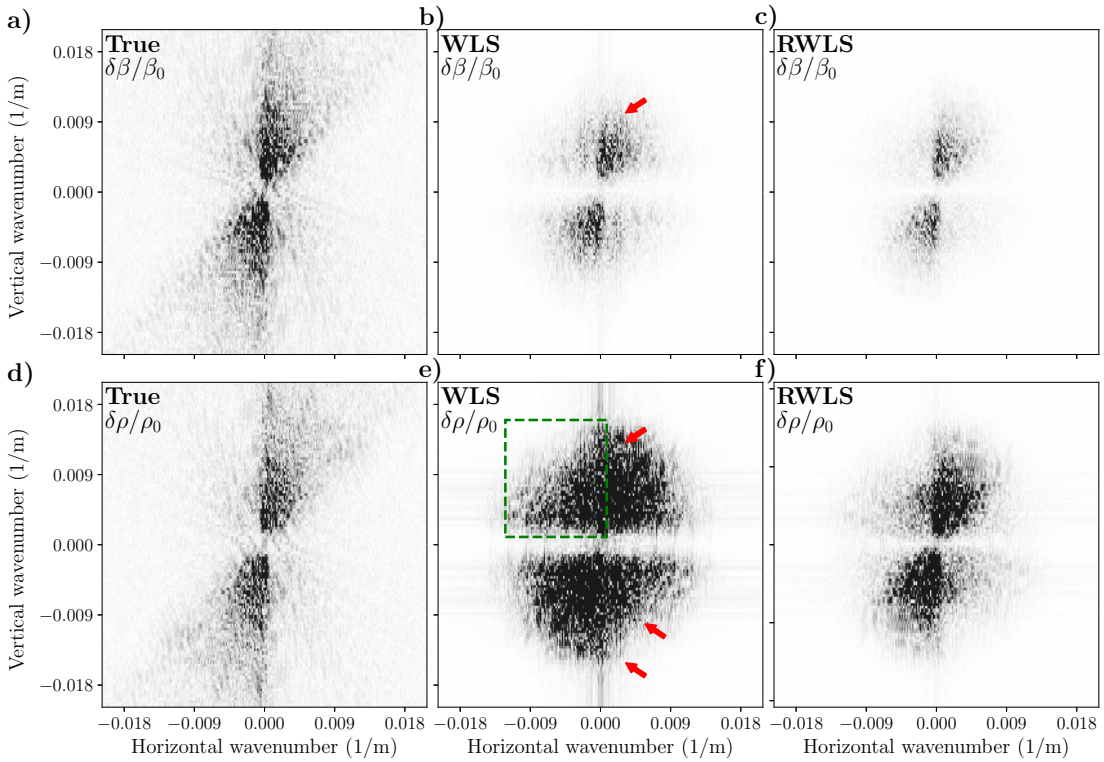

Figure 5.13: The Marmousi2 decimated shots experiment. Wavenumber spectra of $\delta\beta/\beta_0$ corresponding to a) true, b) WLS and c) RWLS, and $\delta\rho/\rho_0$ corresponding to d) true, e) WLS and f) RWLS. Panels are plotted on the same amplitude scale as (a) and (d). The box and arrows highlight the observed artifacts.

domains, and it can potentially deal with a source distance that is larger than the one allowed with the WLS method.

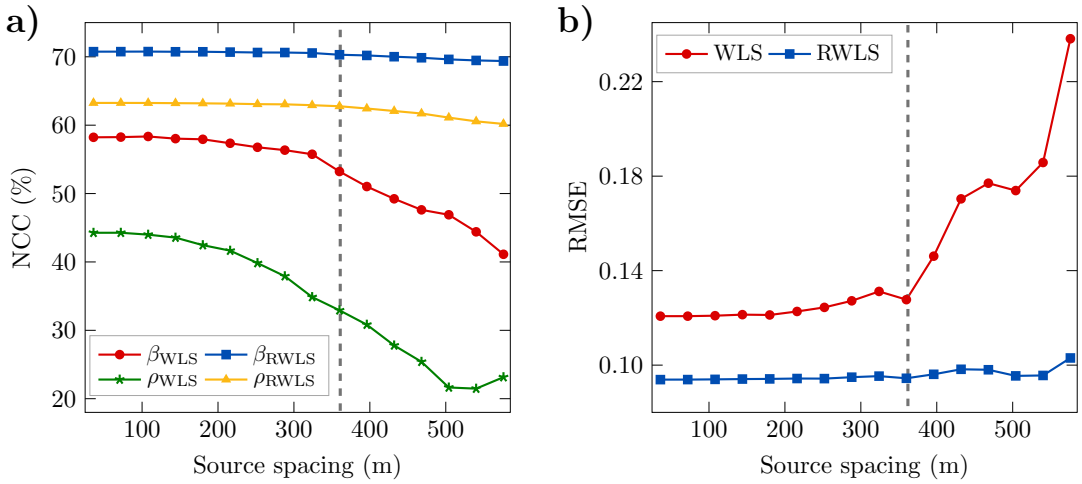

Figure 5.14: Evaluation of a) NCC in image domain and b) RMSE in data domain with respect to the source spacing corresponding to Marmousi2 model. The gray dashed line in (b) indicates the abrupt change in behavior of the WLS method, and the one in (a) is deduced from the one in (b).

5.4.3 2004 BP salt model

In this section, we assess the robustness of the proposed method against random noise and with models with high contrasts. As a benchmark, we consider the 2004 BP salt model (Billette and Brandsberg-Dahl, 2004). The left part of this model is representative of the geology found in the Gulf of Mexico. This model consists of a simple velocity background model with a complex salt body. The density model is constructed by scaling the reflectivity model obtained from real data stacks (Billette and Brandsberg-Dahl, 2004). We notice that the velocity and density/reflectivity models are highly uncorrelated as the density model has more complex trends compared to the ones found in the velocity model. Laboratory measurements show that in general, changes in the density across an interface create changes in velocity (Gardner et al., 1974b; Garia et al., 2019). Thus, we modify the exact velocity model based on the density model while the characteristics of the salt is preserved (Figure 5.15a and 5.15b). We infer the structure of the velocity model from density model by using Gardner's law (Gardner et al., 1974b). Consequently, the modified velocity model consists of a complex velocity background with a complex salt body which follows the density, hence the reflectivity trend. The inverse problem is ill-posed as an additional degree of freedom (density) is taken into account (Virieux and Operto, 2009).

The model is discretized on a 308×787 grid with a 12 m spacing in vertical and horizontal directions. A Ricker wavelet centered at 4.96 Hz (maximum frequency is 12.4 Hz) is used as a source. The shots are evenly spaced at a 36 m interval with a fixed receiver grid. The true background models, obtained by application of a 60 m long 2D Gaussian window on the true models, are used for the pseudoinverse Born operator. The true perturbation models are obtained by taking the difference between the original and the true background models (Figure 5.15c and 5.15d).

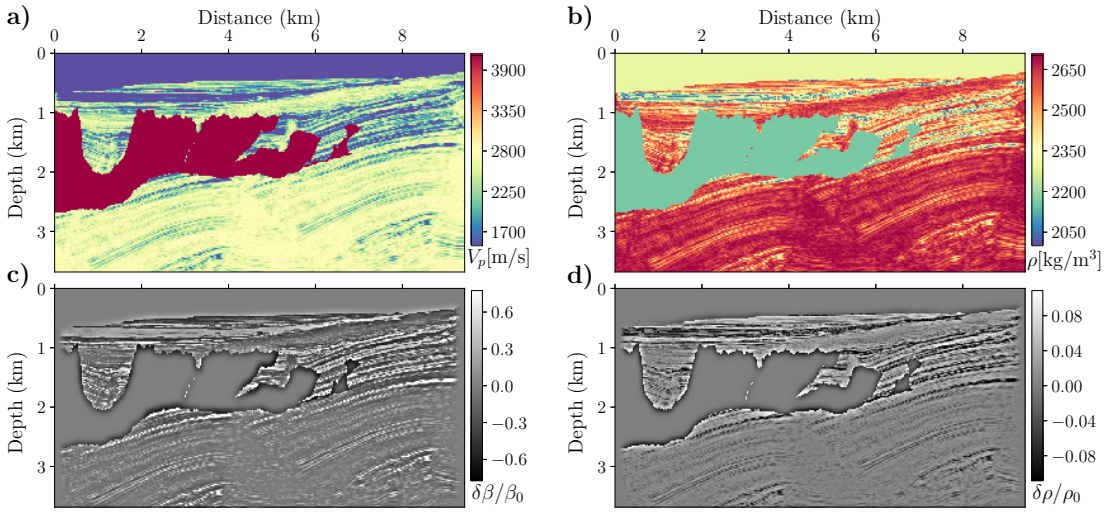

Figure 5.15: The exact a) V_p (m/s), b) ρ (kg/m^3), c) $\delta\beta/\beta_0$ and d) $\delta\rho/\rho_0$ corresponding to the modified 2004 BP salt model.

To assess the resilience of the proposed method to noise, we generate Gaussian random noise with a predefined signal-to-noise (SNR) value. Then, we filter this noise at the maximum frequency of the applied wavelet. The SNR value in decibels (dB) is defined as

$$\mathrm{SNR} = 10 \log_{10} \left(\frac{P_{signal}}{P_{noise}} \right), \tag{5.18}$$

where P is the average power. Here, we add noise with a SNR of -8 dB to the observed data. The negative sign indicates that the noise power is greater than the signal power. We mute the noise before the main event. This is not necessarily easy at far offsets. We first compare the data reconstruction in the middle ($x_s = 4.7$ km) using the noiseless and noisy observed data (Figure 5.16a and 5.16d). The shots are reconstructed using the inverted models obtained via the WLS and RWLS methods corresponding to the noiseless (Figure 5.16b and 5.16c) and noisy data (Figure 5.16e and 5.16f). The number of FISTA iterations is still set to 100 for the RWLS method. The results from the noiseless and noisy data are quite similar in the data domain. For a quantitative evaluation, RMSE

between the observed and reconstructed data is written on each panel. In addition, the central trace from each panel is extracted and NCC is also calculated for each method (Figure 5.17). These values show the superiority of the RWLS method in both noiseless and noisy cases.

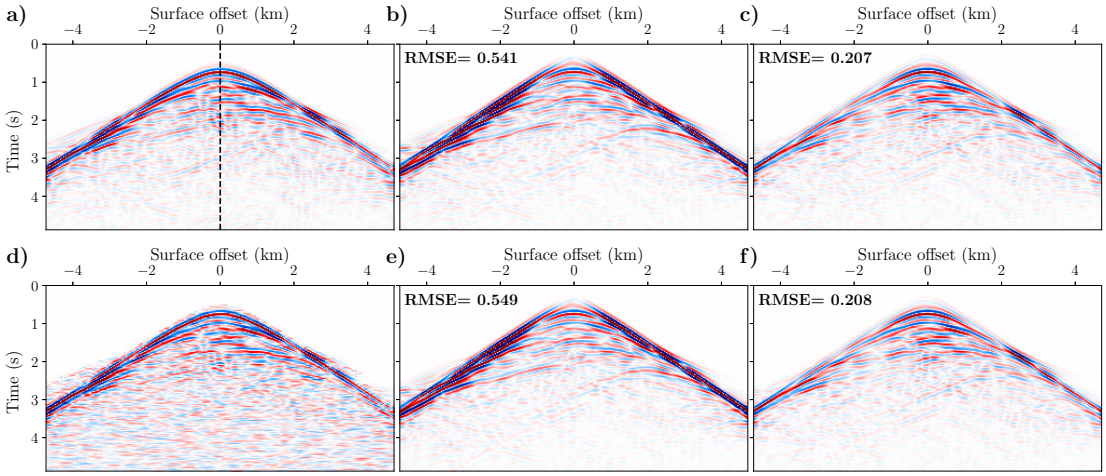

Figure 5.16: The 2004 BP salt model. The a) observed and constructed data via the b) WLS and c) RWLS methods corresponding to the noiseless data. (d), (e), (f) same as (a), (b), (c) corresponding to the noisy data ($\mathrm{SNR} = -8$ dB). The dashed line in (a) corresponds to the extracted trace in Figure 5.17. Shots are plotted at the same scale.

Now, we compare the inverted parameters estimated with the WLS and RWLS methods corresponding to the noiseless (Figure 5.18) and noisy data (Figure 5.19). Similar to the Marmousi2 experiment, we notice an underestimation in $\delta\beta$ and an overestimation in $\delta\rho$. Clearly, the RWLS method provides images closer to the true solutions in both cases. Comparing the $\delta\beta/\beta_0$ images corresponding to the noiseless and noisy cases (Figure 5.18a and 5.19a) using the WLS method reveals that the data noise leads to energy distortion, blurred boundaries of the salt and unclear structures beneath the salt; whereas the image obtained via the RWLS method is nearly the same as the one in the noiseless case (Figure 5.18c and 5.19c). These parts are marked by yellow dashed boxes in Figure 5.19a and 5.19c. It is also observed that the WLS method fails to achieve reasonable density images (Figure 5.18b and 5.19b). We extract two traces at different positions ($x = 3.8$ and 7 km) from each panel of Figure 5.18 and 5.19, and sum them with the background models to qualitatively evaluate the final images (Figure 5.20). The RMSE values between true and inverted traces are also calculated and written on each panel (Figure 5.20). The results further confirm that the RWLS method provides more reliable models, with a significant jump of density quality compared to the case where regularization is not used. It is also interesting to compare histograms of the model parameters estimated by the WLS and RWLS methods for the noisy data (Figure 5.21). In

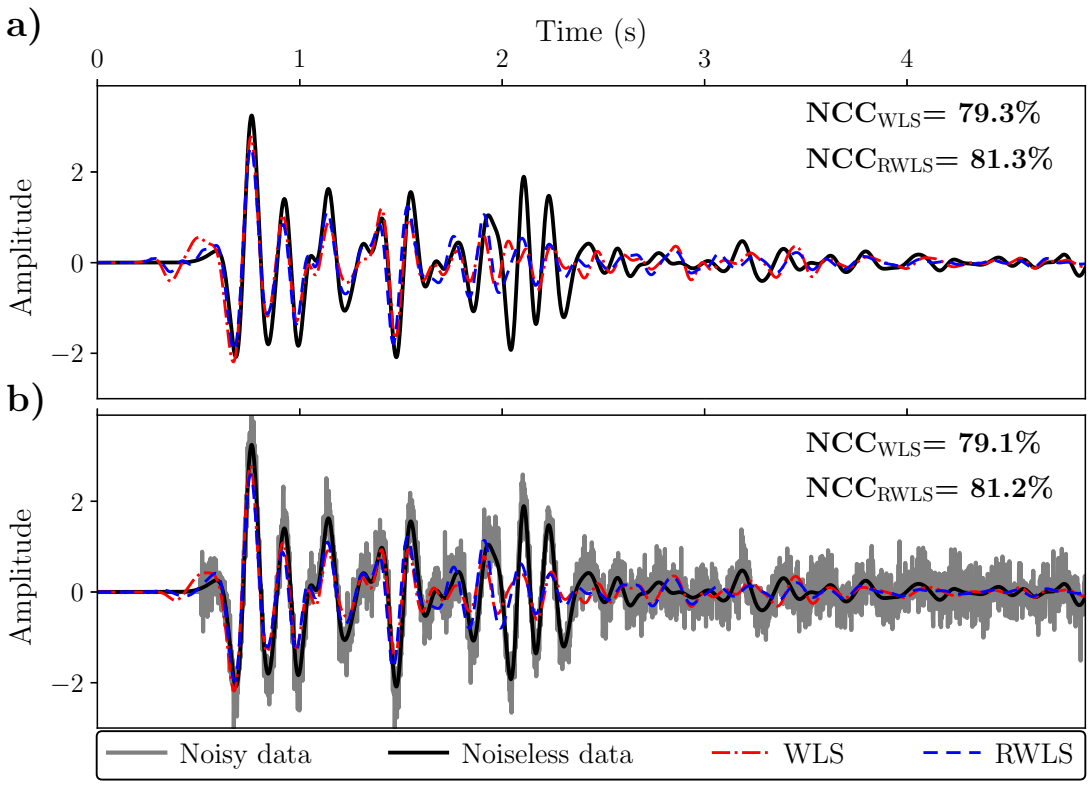

Figure 5.17: The 2004 BP salt model. The zero-offset traces from shots in Figure 5.16 for the a) noiseless and b) noisy data. The NCC values corresponding to each trace are written on the panels.

the context of full-waveform inversion, Aghamiry et al. (2020b) plot these probability densities to assess the ability of hybrid Tikhonov and total variation regularizations in reconstructing the model parameters. The probability densities corresponding to the true models illustrate that the true models are sparse (Figure 5.21a and 5.21b). The probability densities corresponding to the WLS method (Figure 5.21c and 5.21d) illustrate long-tailed prior, which further support the choice of the ℓ_1-norm regularization for this problem (Polson and Sokolov, 2019). The probability densities corresponding to the RWLS method (Figure 5.21e and 5.21f) are similar to the true models and decay faster than the ones of WLS. The latter is indeed consistent with the fact that the ℓ_1-norm regularization forces the model parameters to be sparse.

To evaluate the reliability of the results, this experiment is performed for different SNR values, and the RMSE and NCC values are calculated in data and image domains, respectively. Again, the RWLS method outperforms the WLS method in both image (Figure 5.22a) and data domains (Figure 5.22b). As expected, the NCC corresponding to each parameter in the image domain shows a decreasing trend (Figure 5.22a), whereas

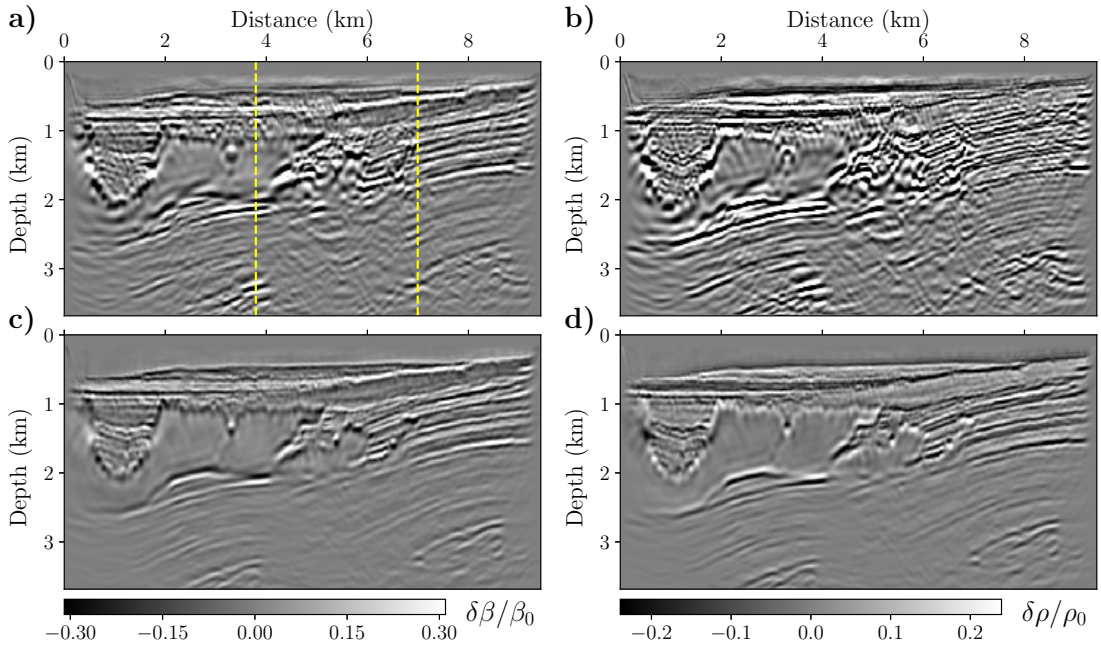

Figure 5.18: The inverted parameters corresponding to the noiseless 2004 BP salt model. Inverted a) $\delta\beta/\beta_0$ and b) $\delta\rho/\rho_0$ via the WLS method, and inverted c) $\delta\beta/\beta_0$ and d) $\delta\rho/\rho_0$ via the RWLS method. The dashed lines in (a) correspond to the extracted traces.

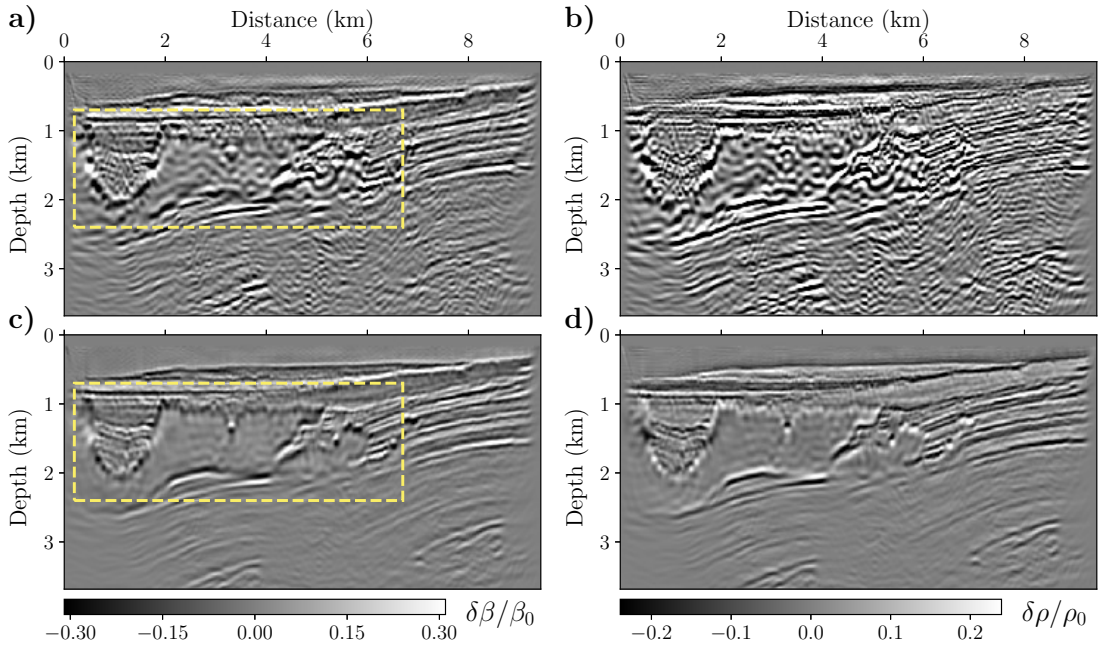

Figure 5.19: The inverted parameters corresponding to the noisy 2004 BP salt model. Inverted a) $\delta\beta/\beta_0$ and b) $\delta\rho/\rho_0$ via the WLS method, and inverted c) $\delta\beta/\beta_0$ and d) $\delta\rho/\rho_0$ via the RWLS method. The yellow dashed boxes indicate the salt boundary.

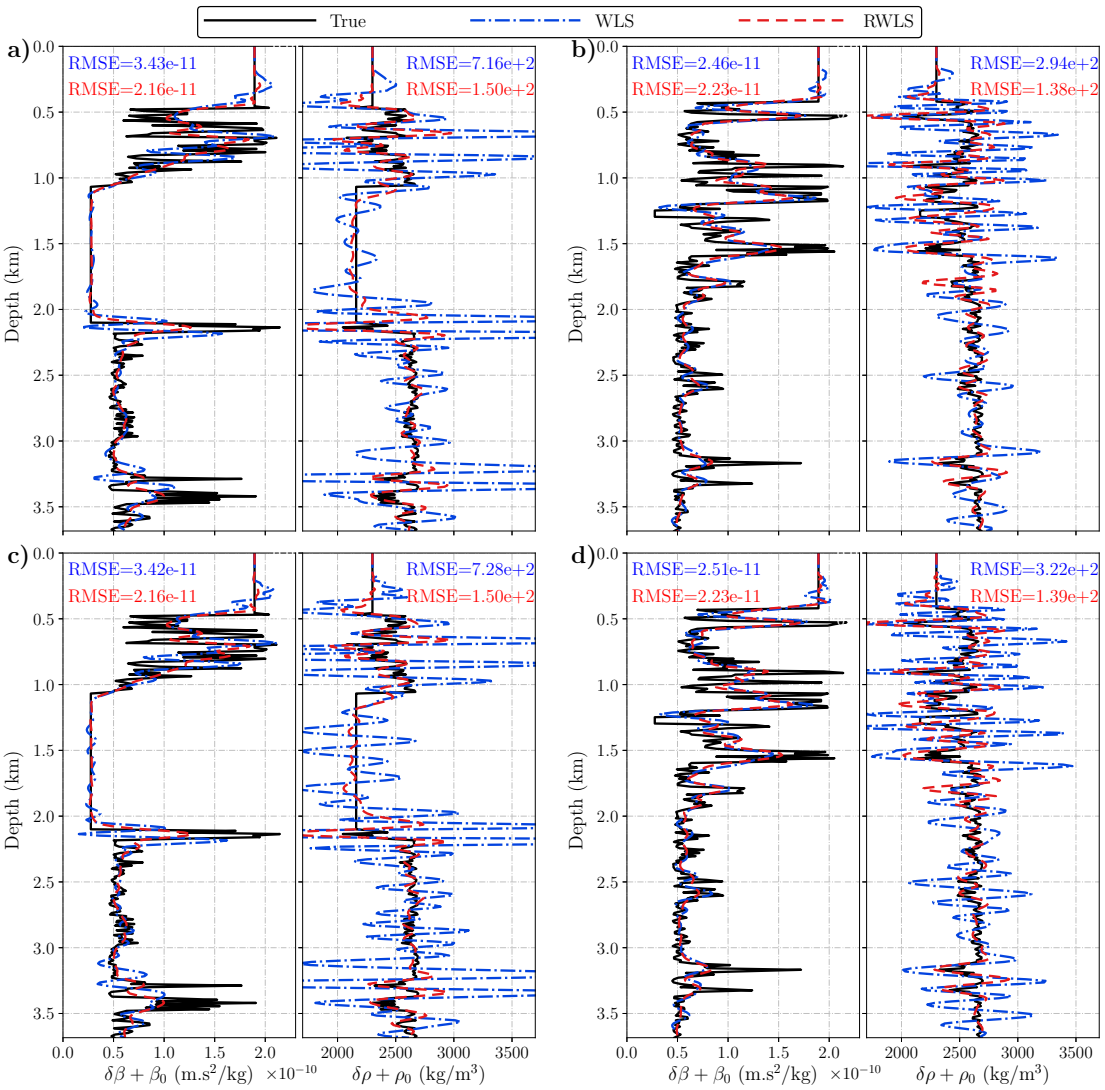

Figure 5.20: The final inverted parameters corresponding to the (a,b) noiseless and (c,d) noisy 2004 BP salt model. Inverted $\delta\beta + \beta_0$ and $\delta\rho + \rho_0$ for (a,c) x $= 3.8$ km and (b,d) x $= 7.0$ km.

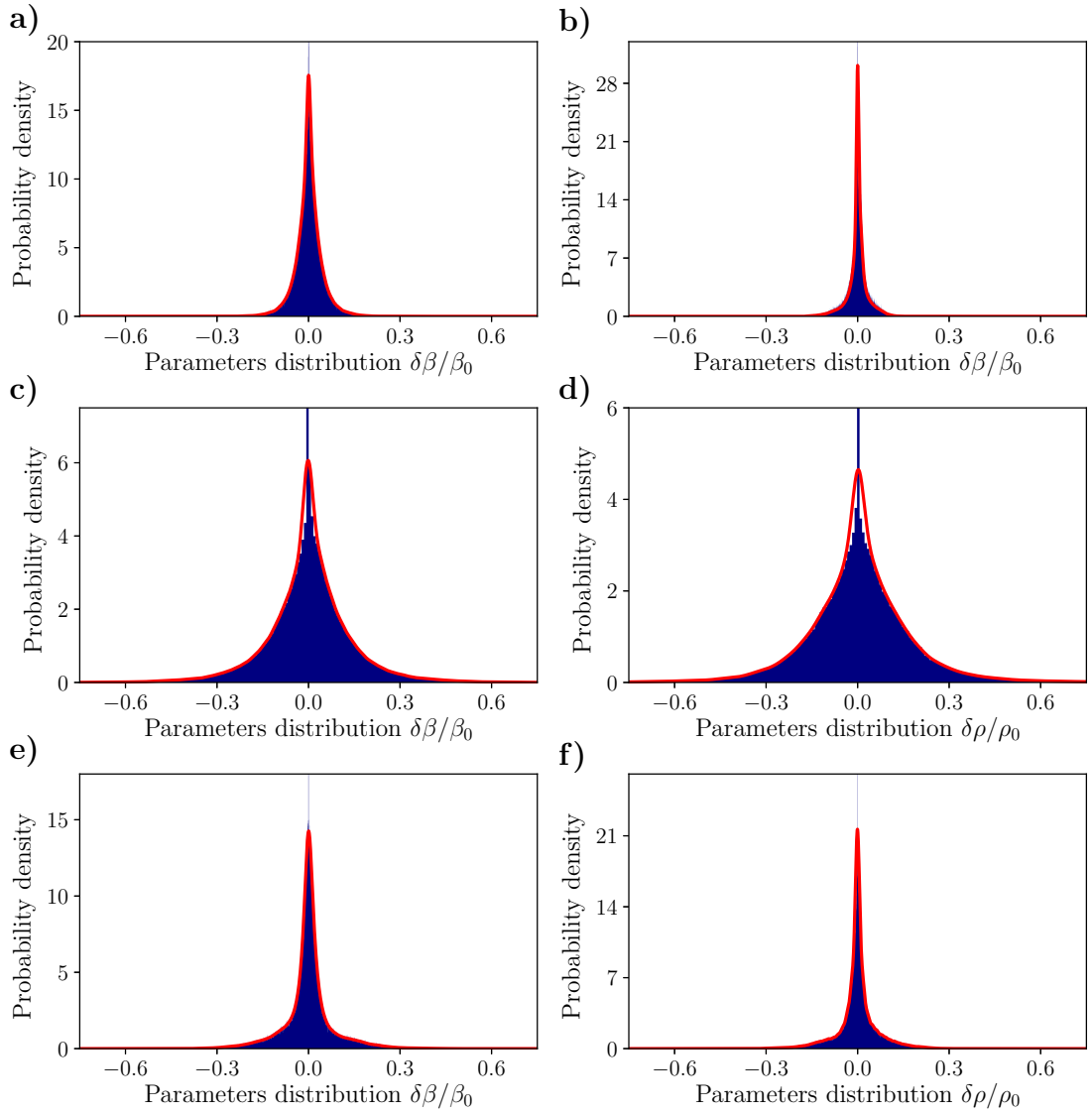

Figure 5.21: Comparison between the histogram of the true model components and the inverted ones (Figure 5.19). a) $\delta\beta/\beta_0$ and b) $\delta\rho/\rho_0$ corresponding to the true models, c) $\delta\beta/\beta_0$ and d) $\delta\rho/\rho_0$ inverted via the WLS method, e) $\delta\beta/\beta_0$ and f) $\delta\rho/\rho_0$ inverted via the RWLS method. The red lines indicate the probability density functions fitted to the histograms.

the RMSE in the data domain shows a increasing trend (Figure 5.22b). Although these trends have a higher absolute slope values for the WLS method, no abrupt change in the behavior of the RWLS method is observed.

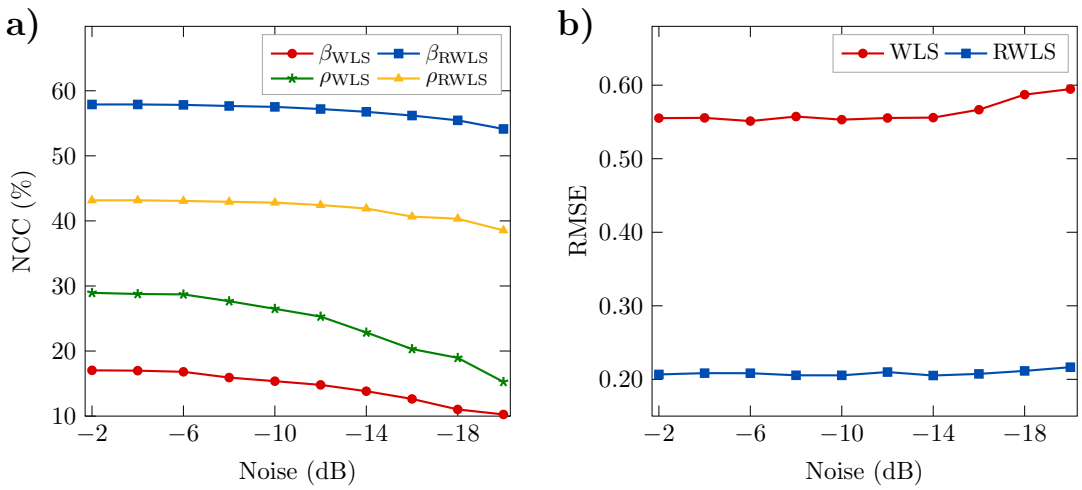

Figure 5.22: Evaluation of a) NCC in image domain and b) RMSE in data domain with respect to noise corresponding to 2004 BP salt model.

5.5 Discussion

We have implemented a ℓ_1-regularization in the weighted least-squares (WLS) method recently introduced by Farshad and Chauris (2020a) to extend the constant density pseudoinverse Born operator to variable density acoustic media. To solve the new constrained optimization problem, we have used FISTA as an appropriate sparse solver (Beck and Teboulle, 2009).

We first discuss the choice of parameter α for the weighting mask (equation 5.11). As mentioned before, the good of this mask is to remove artifacts in the angle domain (Farshad and Chauris, 2020a). For the Marmousi2 and the 2004 BP salt models, the W is applied with $\alpha = 0.48$. The application of this mask for these two models at different positions is investigated (Figure 5.23). In practice, there is no need for a precise selection of α. We choose this parameter visually, to guarantee that we are on the safe side of the angle-domain CIG which does not contain strong artifacts related to the finite sampling (Figure 5.23). This is indeed a conservative approach. Otherwise the risk would be including non-useful (artifacts) information. A possible alternative is choosing this value manually for each CIG position.

With decimated shots, the application of the WLS method and the pseudoinverse Born operator produced artifacts and gave a blurred image of the subsurface. In contrary, adding the regularization term efficiently reduced the acquisition footprint and produced a high-resolution image of the subsurface. Another alternative is to use the LSM method (Nemeth et al., 1999), which requires the application of the adjoint and forward modeling operators at each iteration of the LSM. Although the method is robust, it has a higher computational burden compared to the RWLS method. This comes

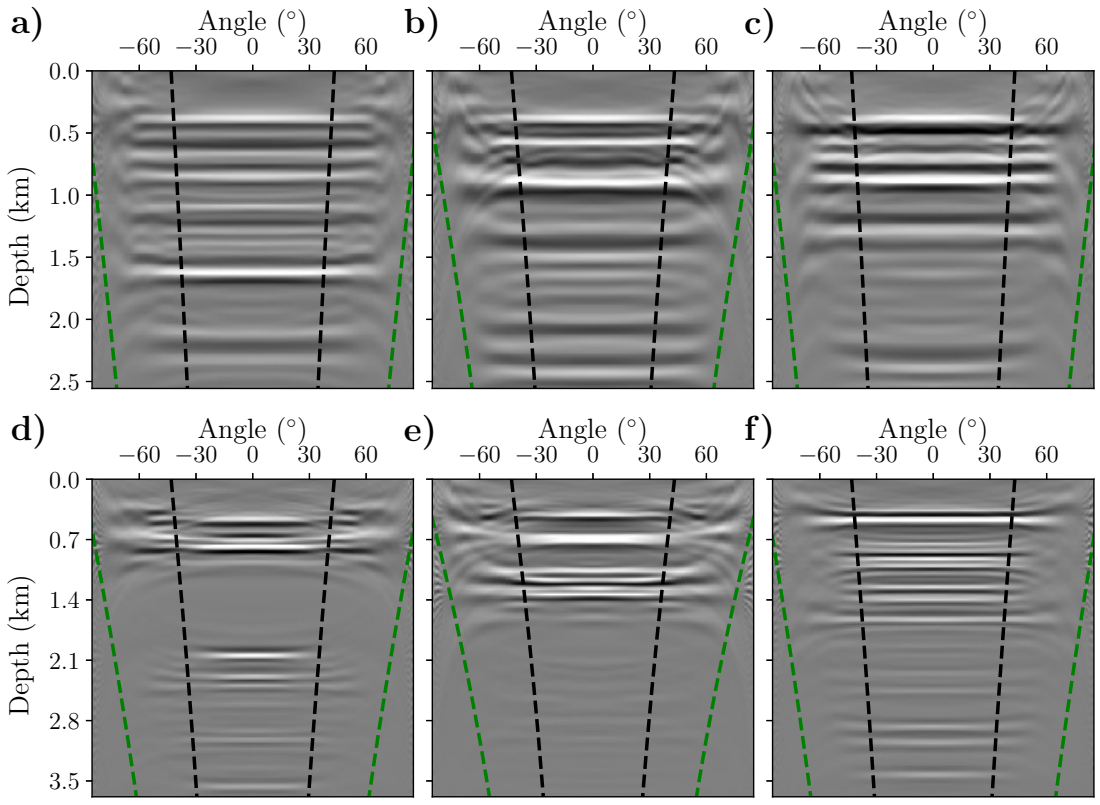

Figure 5.23: The angle-domain CIGs positioned at a) 2.6 km, b) 5.2 km and c) 7.8 km corresponding to the Marmousi2 model, and at d) 2.6 km, e) 5.2 km and f) 7.8 km corresponding to the 2004 BP salt model. The green and black dashed lines define the W for $\alpha = 1$ and $\alpha = 0.48$, respectively.

from the fact that the RWLS method is completely decoupled from the wave-equation based operators, including a single application of the pseudoinverse operator ($\delta m(\mathbf{x}, \mathbf{h})$), a single application of the Radon transform ($\mathcal{R}_{\delta m}(\mathbf{x}, \gamma)$), and iterative FISTA. Thus, the RWLS method can be considered as a postprocessing step after imaging, leading to a more computationally efficient scheme. To get a better insight on the efficiency, we compare the run-time for the dense shots experiment of the Marmousi2 example using 20 processors. The run-times for the pseudoinverse Born operator, the WLS and RWLS methods are 5898, 4 and 43 s, respectively. Note that in this example, the computational costs of the WLS and RWLS methods are $< 1.0\%$ of the cost of the pseudoinverse Born operator. Therefore, the RWLS method provides high-resolution images requiring less computational complexity than LSM.

Finally, we discuss the choice of λ_1 and λ_2 coefficients. First, these are only needed in the final inexpensive FISTA part. In practice, we estimate λ_1 and λ_2 by balancing

between the sparsity of the coefficients and the data fitting. This requires a parameter tuning step to find the suitable λ. We plot in Figure 5.24 the Pareto curves, here in 2D, between the coefficients λ_1, λ_2 and the corresponding cost function J (equation 5.15) for the Marmousi2 model. This method is known as L-hypersurface, the generalization of L-curve method (Belge et al., 1998). This curve exhibits a typical "L" shape with distinctive turning points. The optimal regularization parameter λ corresponds to the corner of the "L", recognized as the turning points (shown by circle in Figure 5.24) (Belge et al., 2002; Brezinski et al., 2003). The cost function reaches the minimum value when λ_1 and λ_2 are relatively small (i.e. around $\lambda_1 = \lambda_2 = 10^{-5}$ for this example). Referring to equation 5.15, it is clear that for $\lambda_1 = \lambda_2 = 0$, the RWLS method reduces to the WLS method. In this case, the regularization terms will not affect the output. By defining higher values (i.e. around 1 for this example), the solution will be completely sparse, containing only important information in the data. We note that the L-hypersurface in this example has two plateaus. A pure 1D analysis for a fixed λ_1 (λ_2) would have led to possibly another choice for λ_2 (λ_1). As discussed before, the computational cost of the RWLS method is negligible. This gives the flexibility to easily draw the L-hypersurface for parameter tuning in the corresponding example. This is a significant advantage over the common regularization techniques applied in seismic imaging methods such as LSM (Anagaw and Sacchi, 2012; Dutta et al., 2017; Ren and Li, 2020) and full-waveform inversion (Askan and Bielak, 2008; Aghamiry et al., 2018, 2019a). This step can be carried out either by trial and error, generalized cross-validation (Craven and Wahba, 1978), or χ^2 methods (Tarantola, 2005), which are indeed computationally expensive and cumbersome for large-scale inverse problems.

5.6 Conclusion

In this paper, to improve the quality of the inverted parameters by the weighted least-squares approach for variable density pseudoinverse Born operator, we proposed to include regularization terms. We considered the ℓ_1-norm of each parameter as the regularization terms. Through numerical experiments, we have demonstrated that the proposed method can effectively suppress artifacts caused by parameter cross-talk, abrupt truncation of interfaces, incomplete data, noisy data and large contrast complex geology. Therefore, it can provide better data reconstruction and higher subsurface images. Future work will explore the use of the multi-parameter pseudoinverse Born operator as a suitable preconditioner within the frame work of least-squares migration, and extension of pseudoinverse Born operator beyond the acoustic approximation.

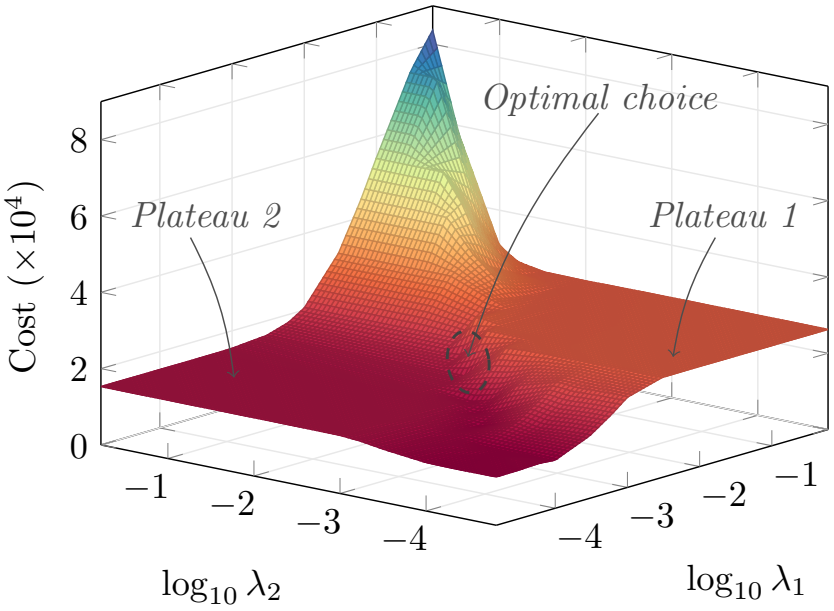

Figure 5.24: The evaluation of the cost function (equation 5.15) with respect to λ_1 and λ_2 for the Marmousi2 model.

5.7 Acknowledgments

We would like to thank assistant editor A. Guitton, associate editor A. Baumstein, and three anonymous reviewers for their valuable comments and suggestions that helped to improve the manuscript.

Chapter 6

Preconditioning least-squres reverse time migration

Contents

6.1 Abstract **193**
6.2 Introduction **193**
6.3 Reverse time migration **197**
6.4 Least-squares reverse time migration **200**
6.5 Preconditioning with pseudoinverse Born operator **200**
6.6 Numerical experiments **202**
6.6.1 Simple model 203
6.6.2 Dipping layer model 211
6.6.3 Marmousi2 model 219
6.7 Discussion **226**
6.8 Conclusion **230**
6.9 Appendix I: Derivation of Hessian **230**
6.10 Appendix II: Alternative approach **231**

Résumé

Ce chapitre comprend l'article intitulé « Accelerating the multi-parameter least-squares reverse time migration using an appropriate preconditioner », publié dans la revue *Computational Geosciences* (Farshad and Chauris, 2021a).
La migration itérative (« least-squares reverse-time migration » en anglais) est connue pour être la technique d'imagerie linéaire recommandée pour les structures complexes du sous-sol. Dans le cas d'un milieu acoustique à densité constante, les itérations ont le potentiel pour compenser pour les artefacts causés par les fréquences finies des sources sismiques, l'acquisition limitée, l'illumination irrégulière et le couplage entre les paramètres. Le principal inconvénient de l'approche itérative est le coût numérique élevé avec le calcul des opérateurs de migration et de modélisation à chaque itération. Pour accélérer la convergence, nous proposons d'utiliser le pseudoinverse à densité variable comme préconditioner. Notre schéma d'imagerie comprend deux étapes. Nous construisons dans un premier temps une image étendue avec les amplitudes préservées, par une approche de Gradient Conjugué, avec et sans préconditioner. Ensuite, après application d'une transformée de Radon 2D, nous estimons simultanément les paramètres physiques avec une approche aux moindres carrés. La seconde étape ne comprend pas de terme de propagation des ondes. Au travers d'expériences numériques, nous montrons que le schéma proposé non seulement réduit le nombre d'itérations requises pour converger, mais aussi augmente de manière significative la qualité des images inversées, même en présence de couplages forts entre les paramètres et de macro-modèles inexacts. C'est confirmé par l'analyse du Hessian multi-paramètre obtenu pour un modèle de taille limité.

Accelerating the multi-parameter least-squares reverse time migration using an appropriate preconditioner
Milad Farshad and Hervé Chauris
Computational Geosciences, 2021, *doi: 10.1007/s10596-021-10089-4*

6.1 Abstract

Least-squares reverse time migration has proven to be the state-of-the-art for linear imaging technique of complex subsurface structures. Assuming a variable-density acoustic medium, least-squares iterations have the potential to compensate for the artifacts caused by finite frequency of the seismic source, limited acquisition aperture, uneven illumination and parameter cross-talk. The main drawback of such an iterative imaging scheme is the substantial computational expense induced by additional modeling/adjoint steps at each iteration. To accelerate the convergence rate, we propose to leverage the variable-density pseudoinverse extended Born operator as a preconditioner. Our imaging scheme consists of two main steps. We first construct a true-amplitude extended image through Conjugate Gradient iterations with/without preconditioning. Then, by applying a 2D Radon transform, we simultaneously estimate the physical parameters from the angle-domain response using a weighted least-squares method. The second step does not involve wave propagation terms. Through numerical experiments, we show that the proposed preconditioning scheme not only largely reduces the required number of iterations to achieve a given data misfit but also significantly increases the quality of the inverted images even in presence of strong parameter cross-talk and inaccurate migration background models. This is further confirmed by analyzing the shape of the multi-parameter Hessian obtained on a model with limited size.

6.2 Introduction

In exploration geophysics, seismic experiments are commonly conducted for subsurface imaging and hydrocarbon detection. In a usual seismic acquisition, an active source and an array of receivers are positioned at the surface (Sheriff and Geldart, 1995). As an example, we construct a simple model with three layers (Figure 6.1a) and the corresponding seismic data collected from receivers (Figure 6.1b). The seismic data contains many traces recording the pressure field for the same shot, as a function of time and the distance from the source (also referred to as surface offset). Generally, the reflected seismic events demonstrate a hyperbolic shape with possible amplitude variations along this hyperbola due to velocity and density variations (Sheriff and Geldart, 1995).

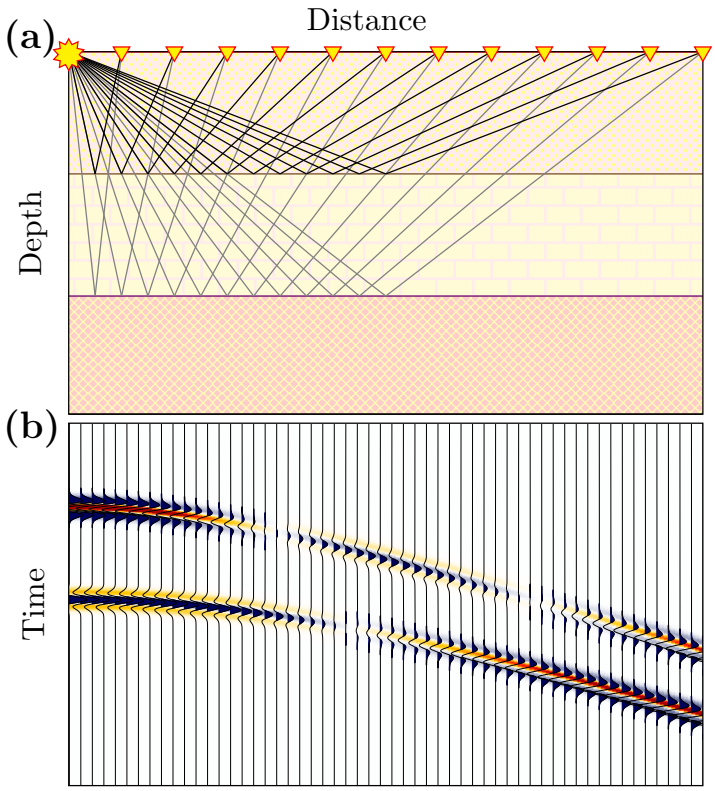

Figure 6.1: a) Typical surface 2D acquisition setting. The *star* and the *triangles* indicate the seismic source and the receiver positions, respectively. The lines emerged from the source represent the ray path of the primary reflections to illustrate the wave propagation. b) The reflected wavefield recorded by the receivers.

Under the weak-scattering assumption, the model parameters m can be decomposed as (Symes, 1995)

$$m = m_0 + \delta m, \tag{6.1}$$

where m_0 denotes the background model controlling the kinematics of the wave propagation (long wavelengths), and δm is the reflectivity model creating the diffracted/reflected waves (short wavelengths). Migration is a linear imaging method that reconstructs the reflectivity model δm in a given background model m_0. Among different linear imaging algorithms, reverse time migration (RTM) has become the preferred approach for seismic imaging in complex geologic structures (Baysal et al., 1983). Classically, RTM can be formulated as the adjoint of the linearized forward modeling operator, i.e., adjoint Born operator. Such scheme can only produce qualitative structural images, but no proper amplitude information. In practice, the quality of the migrated image using RTM operator may be considerably degraded by several causes such as noise, band limitation of the source wavelet, limited recording aperture, and aliasing due to coarsely spaced ac-

quisition geometry (Nemeth et al., 1999; Etgen et al., 2009; Mulder and Plessix, 2004; Chauris, 2019).

Least-squares reverse time migration (LSRTM), also known as linearized seismic inversion, can be adapted to suppress migration artifacts via iterative methods (Lailly, 1983; Nemeth et al., 1999; Zeng et al., 2014). Although LSRTM provides superior results compared to RTM, it is computationally extensive and relies on many iterations involving migrations and modeling operators. Despite the evolution of computer technologies, reducing the overall computational LSRTM cost is an important aspect. Moreover, in the context of multi-component seismology, extension to multi-parameter imaging is an active area of research. This article presents an approach for significantly accelerating the convergence rate of multi-parameter LSRTM.

The necessity for realistic wave propagation phenomena has prompted researchers to extend pure acoustic LSRTM to account for parameter classes other than the P-wave velocity, such as density (Yang et al., 2016a; Chen and Sacchi, 2019), attenuation (Dutta and Schuster, 2014; Sun et al., 2016; Guo and McMechan, 2018), S-wave velocity (Ren et al., 2017; Chen and Sacchi, 2019), and anisotropy (Qu et al., 2017). Among different subsurface parameters, density is one of the key elements for rock property analysis and lithologic interpretation. However, estimating a good quality density image is a very challenging task (Tarantola, 1986; Köhn et al., 2012). In the concept of multi-parameter imaging, an issue that comes to the fore is the parameter cross-talk (for example imprint of density on the velocity section). In practice, the diffraction patterns of different parameters overlap within certain illumination angles, which makes multi-parameter imaging a highly ill-posed problem especially with limited surface acquisition (Virieux and Operto, 2009).

The second topic related to the computational burden, can be satisfied either by lowering the computational cost per iteration or by accelerating the convergence rate of LSRTM. Concerning the first kind, researchers proposed to reduce the cost either by blending shot gathers using random-phase/plane-wave encoding methods (Romero et al., 2000; Liu et al., 2006; Schuster et al., 2011), constraining the estimated image by regularization techniques (Lin and Lianjie, 2015; Xue et al., 2016), or compressed sensing and stochastic optimization (van Leeuwen et al., 2011; Herrmann and Li, 2012; Zand et al., 2020). The second kind of techniques is based on scaling the linear system by the inverse of the Hessian matrix. In practice, constructing and then calculating the inverse of the Hessian matrix demands extensive computations. Thus the Hessian is usually approximated by only its diagonal components, possibly surrounded by a limited number of off-diagonals (Shin et al., 2001; Rickett, 2003; Symes, 2008a; Herrmann et al., 2009; Chen and Sacchi, 2017). The method proposed in this article incorporates the Hessian information in the context of multi-parameter imaging.

In recent years, several pseudoinverse Born operators, often referred to as true-amplitude migration, have been developed by simple modifications of the extended RTM operator (ten Kroode, 2012; Hou and Symes, 2015, 2017; Chauris and Cocher,

2017). By extended, we mean that the modeling and adjoint operators define a mapping between the data and image domains. These pseudoinverse operators asymptotically are inverse of the Hessian but without any ray quantities in the final formulation. The application of these operators as a preconditioner has been studied for LSRTM (Hou and Symes, 2016a; Chauris and Cocher, 2018), migration velocity analysis (Chauris and Cocher, 2017; Cocher et al., 2018; Li and Chauris, 2018), and the non-linear Full waveform inversion (FWI) (Hou and Symes, 2016b). This framework was originally developed for constant-density acoustic media, and recently has been extended to account for density variations (Dafni and Symes, 2018; Farshad and Chauris, 2020a,c). However, Dafni and Symes (2018) and Farshad and Chauris (2020a) also illustrated that the variable-density pseudoinverse Born operator, being only an approximate inverse, suffers from parameter cross-talks. Implementing variable-density pseudoinverse Born operator as a preconditioner for LSRTM has not been exploited yet (Hou and Symes, 2016a).

In this article, we implement the pseudoinverse Born operator proposed by Farshad and Chauris (2020a) for preconditioning the variable-density LSRTM algorithm. The main motivation of this combination (pseudoinverse Born + LSRTM) is to highly accelerate the convergence rate while reducing the parameter cross-talks. The total imaging process implemented here (with/without preconditioner) consists of two main steps. First, we perform iterative migration: from the observed data, we obtain "true-amplitude extended images". Then, we estimate two physical acoustic parameters (impedance and density perturbations) from the angle-dependent response of the extended image using an efficient weighted least-squares (WLS) approach. The second step, namely the WLS method, is a generalization of the work of Zhang et al. (2014b) in the sense that parameters are inverted using all traces in the angle-domain rather than two traces. Using a small-sized model, we illustrate the multi-parameter Hessian matrix without and with preconditioner. Furthermore, we numerically assess the capability of the proposed method in mitigating parameter cross-talk while accelerating the convergence rate.

The rest of the paper is organized as follows. First, we review the theory of multi-parameter RTM and the solution of LSRTM with the Conjugate Gradient (CG) method through Sections 6.3 and 6.4. Next, we explain how to implement the pseudoinverse Born operator as a preconditioner using the preconditioned Conjugate Gradient (PCG) method. In Section 6.6, we present three numerical examples to verify the effectiveness of the proposed method in the presence of parameter cross-talk, inaccurate migration background models, and complex structures. Finally, we discuss the structure of the multi-parameter Hessian to further analyze the numerical experiments.

6.3 Reverse time migration

The 2D acoustic Earth model can be parameterized in terms of two parameters on each spatial coordinate $\mathbf{x} = (x, z)$, involving P-wave velocity (V_p) and density (ρ) or their combination, namely P-wave impedance ($I_p = \rho V_p$) or inverse of the bulk modulus ($\beta = 1/\rho V_p^2$). Under the Born approximation, the model parameters are split into a smooth part (background model, e.g. β_0) and a singular part (perturbation model, e.g. $\delta\beta$). The forward Born modeling operator $\mathcal{L}$ can be expressed as

$$\begin{aligned} d(\mathbf{s},\mathbf{r},\omega) &= \mathcal{L}(\delta\beta,\delta\rho)(\mathbf{s},\mathbf{r},\omega) \\ &= \Omega(\omega)\int \mathrm{d}\mathbf{x} \begin{bmatrix} -(i\omega)^2 G_0(\mathbf{s},\mathbf{x},\omega)\delta\beta(\mathbf{x})G_0(\mathbf{x},\mathbf{r},\omega) \\ \rho_0^{-2}\nabla G_0(\mathbf{s},\mathbf{x},\omega)\delta\rho(\mathbf{x})\cdot\nabla G_0(\mathbf{x},\mathbf{r},\omega) \end{bmatrix}, \end{aligned} \tag{6.2}$$

where ω is the angular frequency, $\Omega(\omega)$ is the source signature, and d denotes the single-scattered reflection data. The terms $G_0(\mathbf{s},\mathbf{x},\omega)$ and $G_0(\mathbf{x},\mathbf{r},\omega)$ are the variable-density acoustic Green's functions in the background models, from the source position $\mathbf{s} = (x_s, z_s)$ to imaging point $\mathbf{x}$, and from $\mathbf{x}$ to the receiver position $\mathbf{r} = (x_r, z_r)$.

A standard migrated image $\delta m^{mig}(\mathbf{x})$ (same unit as $\delta\beta$: m.s^2/kg) can be estimated by applying RTM operator, i.e. adjoint Born operator, to the observed data (Baysal et al., 1983):

$$\delta m^{mig} = \mathcal{L}^T d^{obs}, \tag{6.3}$$

where ^{T} denotes the adjoint operator.

We generalize the forward/adjoint Born operator by introducing a redundancy parameter, involving nonphysical subsurface offset or time-lag extension (Sava and Fomel, 2006; Symes, 2008b). Here we consider extension by horizontal subsurface offset $\mathbf{h} = (h, 0)$, defined as sunken sources and receivers in Claerbout's survey-sinking concept (Figure 6.2) (Claerbout, 1985). This is indeed essential in the derivation of the pseudoinverse Born operator to have the same data and model dimensions (Beylkin, 1985; ten Kroode, 2012). Moreover, such modification offers the flexibility to easily extend RTM operator to variable-density acoustic medium. Accordingly, the extended migration image is formulated as

$$\begin{aligned} \delta m^{mig} &= \mathcal{L}^T(d^{obs})(\mathbf{x},\mathbf{h}) \\ &= -\int \mathrm{d}\mathbf{s}\,\mathrm{d}\mathbf{r}\,\mathrm{d}\omega\,(i\omega)^2\Omega^*(\omega)G_0^*(\mathbf{s},\mathbf{x}-\mathbf{h},\omega)d^{obs}(\mathbf{s},\mathbf{r},\omega)G_0^*(\mathbf{x}+\mathbf{h},\mathbf{r},\omega), \end{aligned} \tag{6.4}$$

where * denotes the complex conjugate. For each couple $(\mathbf{x},\mathbf{h})$, δm^{mig} is a scalar quantity (result of wavefield cross-correlation). In equation 6.4, the acoustic Green's

functions in background models β_0 and ρ_0 are evaluated from the source position $\mathbf{s}$ to imaging point $\mathbf{x}-\mathbf{h}$ and from imaging point $\mathbf{x}+\mathbf{h}$ to the receiver position $\mathbf{r}$.

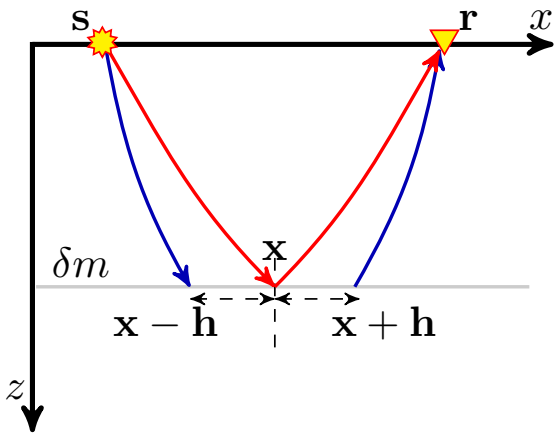

Figure 6.2: Sketch of the subsurface offset extension. The red and blue arrows indicate the interaction between the source and receiver wavefields for the non-extended and extended modeling operators, respectively (Claerbout, 1985).

Under the ray+Born approximation, the extended image can be simultaneously decomposed into two physical parameters based on the diffraction patterns (Zhang et al., 2014b; Dafni and Symes, 2018; Farshad and Chauris, 2020a). Since the aim of RTM is to resolve only the short-wavelength part of the model, we choose impedance (I_p) and density (ρ) for parameterization (Tarantola, 1986). The decomposing relationship reads

$$\frac{\delta m^{mig}}{\beta_0}(\mathbf{x},h) \cong \mathcal{D}_{I_p}\zeta_{I_p}(\mathbf{x})\delta(h) + \mathcal{D}_{\rho}\zeta_{\rho}(\mathbf{x})\delta(h), \tag{6.5}$$

where ζ is the relative model perturbation defined as

$$\zeta_a = \frac{\delta a}{a_0}. \tag{6.6}$$

In equation 6.5, $\delta()$ denotes the Dirac delta function, θ is the scattering angle, $\mathcal{D}_{I_p} = -2$ and $\mathcal{D}_{\rho} = 2\sin^2(\theta)$ are the radiation patterns of I_p and ρ, respectively (Forgues, 1996). Note that the left side of equation 6.5 is $\mathbf{h}$-dependent, whereas the right side θ-dependent (due to $\mathcal{D}_{\rho}$). Thus a conversion of subsurface offset $\mathbf{h}$ to scattering angle θ is required. This can be efficiently obtained by the application of the 2D linear Radon transform (slant-stack integral) on the extended δm^{mig} as (Sava and Fomel, 2003)

$$\psi = \mathcal{R}(\delta m^{mig})(\mathbf{x},\theta) = \int \mathrm{d}h\, \delta m^{mig}(x, z + h\tan\theta, h), \tag{6.7}$$

where $\psi(\mathbf{x},\theta)$ denotes the angle-domain response of the extended reflectivity integrated along the $h\tan\theta$ direction.

Recently, Farshad and Chauris (2020a) proposed an efficient approach, namely WLS, to retrieve optimal physical parameters from $\psi(\mathbf{x},\theta)$ by defining a new objective function reading as

$$J_{WLS}(\zeta_{I_p},\zeta_\rho) = \frac{1}{2}\left\| 2\Big(-\zeta_{I_p}(\mathbf{x}) + \zeta_\rho(\mathbf{x})\sin^2(\theta)\Big) - \psi(\mathbf{x},\theta)\right\|_W^2 . \tag{6.8}$$

The term $W(\mathbf{x},\theta)$ is a weighted space defined as

$$W(\mathbf{x},\theta) = \begin{cases} 1, & \text{if } |\theta| \leq \alpha\tan^{-1}(\frac{x_{max}}{z}) \\ 0, & \text{otherwise} \end{cases} , \tag{6.9}$$

where x_{max} is the maximum surface offset, and $\alpha < 1$ is a scaling parameter to suppress the artifacts at large angles in $\psi(\mathbf{x},\theta)$ which are due to the limited acquisition geometry. The reader is referred to Farshad and Chauris (2020a) and Farshad and Chauris (2021c) for a more detailed discussion regarding $W(\mathbf{x},\theta)$ and the choice of parameter α. The optimal physical perturbation models can be obtained by setting the gradient of equation 6.8 equal to zero. Abstractly, for the model parameters defined as $\delta\mathbf{u} = \begin{bmatrix}\zeta_{I_p}(\mathbf{x}) & \zeta_\rho(\mathbf{x})\end{bmatrix}^T$, the WLS method can be written in a matrix form as $\mathbf{G}\delta\mathbf{u} = \mathbf{R}$, where

$$\mathbf{G} = \begin{bmatrix} 4\int \mathrm{d}\theta\, W(\mathbf{x},\theta) & -4\int \mathrm{d}\theta\, W(\mathbf{x},\theta)\sin^2(\theta) \\ -4\int \mathrm{d}\theta\, W(\mathbf{x},\theta)\sin^2(\theta) & 4\int \mathrm{d}\theta\, W(\mathbf{x},\theta)\sin^4(\theta) \end{bmatrix} , \tag{6.10}$$

and

$$\mathbf{R} = \begin{bmatrix} -2\int \mathrm{d}\theta\, W(\mathbf{x},\theta)\psi(\mathbf{x},\theta) \\ 2\int \mathrm{d}\theta\, W(\mathbf{x},\theta)\sin^2(\theta)\psi(\mathbf{x},\theta) \end{bmatrix} . \tag{6.11}$$

Since δm^{mig} is obtained by RTM operator (simply adjoint), the estimated $\delta\mathbf{u}$ suffers from various undesirable artifacts, amplitude errors as well as parameters cross-talk due to non-linearity of the multi-parameter inverse problem (Nemeth et al., 1999; Virieux and Operto, 2009). For high-quality subsurface images with accurate amplitudes, δm^{mig} should be replaced by true-amplitude δm obtained by the inverse operator. We now review how to derive it through an iterative process.

6.4 Least-squares reverse time migration

LSRTM aims to solve true-amplitude δm by minimizing a least-squares objective function

$$J(\delta m) = \frac{1}{2} \left\| \mathcal{L}\delta m - d^{obs} \right\|_2^2 . \tag{6.12}$$

The optimal solution for equation 6.12 satisfies $\partial J(\delta m)/\partial \delta m = 0$, leading to the normal equation (Tarantola, 1984)

$$(\mathcal{L}^T \mathcal{L})\delta m = \delta m^{mig}, \tag{6.13}$$

where $\mathcal{L}^T\mathcal{L}$ is the Hessian operator, the second-order partial derivatives of the objective function (equation 6.12) with respect to the model parameters. The optimal model δm can be recovered from δm^{mig} by removing the effect of the Hessian matrix (Schuster, 1993). The construction of the Hessian matrix in an explicit way is prohibitively expensive. Traditionally, the effect of Hessian matrix is removed by iteratively minimizing J (equation 6.12) (Nemeth et al., 1999). To carry out the least-squares problem (equation 6.13), we use the standard Conjugate Gradient (CG) algorithm (Nocedal and Wright, 2006). The CG method has the highest convergence rate among the other first-order iterative approaches such as gradient decent, singular-value or QR-decomposition. The algorithm of CG-LSRTM is summarized in Algorithm 6. In the first iteration ($i = 0$), the update direction (p_0, line 10 in Algorithm 6) is the steepest descent direction, i.e. gradient direction ($-r_0$, lines 4 and 5 in Algorithm 6), along the minimizer κ (line 9 in Algorithm 6). In the next iterations, each direction p_{i+1} (line 13 in Algorithm 6) is chosen to be a linear combination of the current gradient direction r_{i+1} and the previous direction p_i, where the parameter η_{i+1} (line 12 in Algorithm 6) satisfies the conjugacy between p_{i+1} and p_i. Note that the operators $\mathcal{L}$ and $\mathcal{L}^T$ (line 8 in Algorithm 6) are applied in two successive paths (first $\mathcal{L}$ and then $\mathcal{L}^T$).

The main drawback of CG-LSRTM is a slow convergence rate due to the ill-conditioned normal operator $\mathcal{L}^T\mathcal{L}$. In this article, we tackle this issue by introducing pseudoinverse Born operator as a preconditioner to accelerate the convergence rate of CG-LSRTM.

6.5 Preconditioning with pseudoinverse Born operator

Originally, the wave-equation-based pseudoinverse Born operators were introduced for pure acoustic media (Hou and Symes, 2015; Chauris and Cocher, 2017). Recently and in parallel, Dafni and Symes (2018) and Farshad and Chauris (2020a) enlarged the applicability of these pseudoinverse Born operators to variable-density acoustics. Following the works of Chauris and Cocher (2017) and Farshad and Chauris (2020a), the pseudoinverse Born operator $L^\dagger$ reads as

$$\delta m^{inv} = \mathcal{L}^\dagger(d^{obs})(\mathbf{x}, \mathbf{h}) \tag{6.14}$$

Algorithm 6: LSRTM by CG algorithm

Input: d^{obs}
Initialization
  $\delta m_0 \leftarrow 0$
  $r_0 \leftarrow \mathcal{L}^T(\mathcal{L}\delta m_0 - d^{obs}) = -\mathcal{L}^T d^{obs}$
  $p_0 \leftarrow -r_0$
  $i \leftarrow 0$
CG: while *stopping criteria not satisfied* **do**
  $q_i \leftarrow \mathcal{L}^T\mathcal{L}p_i$
  $\kappa_i \leftarrow \langle r_i, r_i\rangle / \langle p_i, q_i\rangle$
  $\delta m_{i+1} \leftarrow \delta m_i + \kappa_i p_i$
  $r_{i+1} \leftarrow r_i + \kappa_i q_i$
  $\eta_{i+1} \leftarrow \langle r_{i+1}, r_{i+1}\rangle / \langle r_i, r_i\rangle$
  $p_{i+1} \leftarrow -r_{i+1} + \eta_{i+1} p_i$
  $i \leftarrow i + 1$
Physical decomposition: for *each* $\mathbf{x}$ *position* **do**
  $\psi \leftarrow \mathcal{R}(\delta m)$
  for *each depth position* **do**
    Construct $\mathbf{G}$, $\mathbf{R}$
    $\delta\mathbf{u} \leftarrow \mathbf{G}^{-1}\mathbf{R}$

$$\simeq \quad 32\frac{\beta_0}{\rho_0^3}\partial_z \int \mathrm{d}\mathbf{s}\,\mathrm{d}\mathbf{r}\,\mathrm{d}\omega\, \frac{\Omega^\dagger(\omega)}{(i\omega)}\partial_{s_z} G_0^*(\mathbf{s}, \mathbf{x}-\mathbf{h}, \omega) d^{obs}(\mathbf{s}, \mathbf{r}, \omega)\partial_{r_z} G_0^*(\mathbf{x}+\mathbf{h}, \mathbf{r}, \omega),$$

where † denotes the pseudoinverse operator. Comparing to the adjoint (equation 6.4), the main differences in the inversion formula are: (1) applying a normalization factor $32\beta_0/\rho_0^3$ and a vertical derivative at the image points, (2) replacing the adjoint source wavelet with its inverse version, (3) applying a first-order integration in time instead of a second-order derivative in imaging condition, and (4) applying vertical derivatives with respect to source and receiver positions to the Green's functions. These modifications have dynamic effects: larger weights to short offsets and small dips (Chauris and Cocher, 2017).

To incorporate $\mathcal{L}^\dagger$ as a preconditioner in LSRTM, we perform a change of variable in equation 6.12 from δm to $\delta\widetilde{m}$ as (Nocedal and Wright, 2006)

$$\delta m = \mathcal{L}^\dagger \delta\widetilde{m}, \tag{6.15}$$

and thus the objective function defined by equation 6.12 transforms to

$$\widetilde{J}(\delta\widetilde{m}) = \frac{1}{2}\left\|\mathcal{L}\mathcal{L}^\dagger\delta\widetilde{m} - d^{obs}\right\|_2^2 = J(\delta m). \tag{6.16}$$

We consequently end up with the modified normal equation

$$\mathcal{L}^{\dagger^T}(\mathcal{L}^T\mathcal{L})\mathcal{L}^\dagger\delta\widetilde{m} = \mathcal{L}^{\dagger^T}\mathcal{L}^T d^{obs}, \tag{6.17}$$

where $\mathcal{L}^{\dagger^T}$ (adjoint of the pseudoinverse Born operator) is defined as (Cocher, 2017)

$$\begin{aligned}\mathcal{L}^{\dagger^T}(\delta m)(\mathbf{s},\mathbf{r},\omega) \quad &\simeq \\ -32\frac{\Omega^{\dagger^*}(\omega)}{(i\omega)}\int \mathrm{d}\mathbf{x}\,\mathrm{d}\mathbf{h}\,\partial_{s_z}G_0(\mathbf{s},\mathbf{x}-\mathbf{h},\omega)\frac{\beta_0}{\rho_0^3}\partial_z\delta m(\mathbf{x},\mathbf{h})\partial_{r_z}G_0(\mathbf{x}+\mathbf{h},\mathbf{r},\omega).\end{aligned} \tag{6.18}$$

Comparing $\mathcal{L}^\dagger$ and $\mathcal{L}$ recalls that the pseudoinverse Born operator is a weighted version of RTM operator. We notice the same weights in $\mathcal{L}^{\dagger^T}$. Therefore, this can also be interpreted as a weighted forward Born modeling operator that emphasizes small scattering angles and short surface offsets.

It was demonstrated that the pseudoinverse $\mathcal{L}^\dagger$ satisfies $\mathcal{L}^\dagger\mathcal{L} \approx \mathcal{I}$ in an asymptotic sense (ten Kroode, 2012; Hou and Symes, 2015, 2017; Chauris and Cocher, 2017). Subsequently, the left-side of equation 6.17 interestingly cancel out each other as:

$$\mathcal{L}^{\dagger^T}(\mathcal{L}^T\mathcal{L})\mathcal{L}^\dagger = (\mathcal{L}\mathcal{L}^\dagger)^T(\mathcal{L}\mathcal{L}^\dagger) \approx \mathcal{I}, \tag{6.19}$$

implying that the convergence rate of equation 6.17 is expected to be faster than the one corresponding to equation 6.13. In practice, it is not necessary to change the variable via equation 6.15 explicitly, but rather defining $\mathcal{L}^\dagger\mathcal{L}^{\dagger^T}$ as a preconditioner. The algorithm is summarized in Algorithm 7. The main difference between the CG and PCG methods is in the application of the preconditioner (lines 5 and 13 in Algorithm 7). Thus the application of WLS method remains the same for both CG and PCG methods.

6.6 Numerical experiments

In this section, we compare the performance of LSRTM with CG and PCG methods using three examples. We start with a simple model containing sparse layers to first investigate the role of preconditioner in attenuating parameter cross-talks. Then, a dipping layer model with dense structure is analyzed to verify the effectiveness of the proposed method using correct and incorrect background models. Finally, we use the Marmousi2 model as a benchmark of complex geology.

To generate synthetic data, we perform a time-domain staggered-grid finite-difference modeling scheme with eighth-order accuracy in space and second-order accuracy in time (Virieux, 1986) implemented with perfectly matched layers (PML) absorbing boundary conditions (Bérenger, 1994) . We consider a fixed acquisition geometry, i.e. stationary receivers, and a Ricker wavelet as the source wavelet for all numerical experiments. Table 6.1 outlines the model discretization and acquisition geometry configurations for

Algorithm 7: Preconditioned LSRTM by CG algorithm

Input: d^{obs}
Initialization
  $\delta m_0 \leftarrow 0$
  $r_0 \leftarrow \mathcal{L}^T(\mathcal{L}\delta m_0 - d^{obs}) = -\mathcal{L}^T d^{obs}$
  $y_0 \leftarrow \mathcal{L}^\dagger \mathcal{L}^{\dagger^T} r_0$
  $p_0 \leftarrow -y_0$
  $i \leftarrow 0$
PCG: while *stopping criteria not satisfied* **do**
  $q_i \leftarrow \mathcal{L}^T \mathcal{L} p_i$
  $\kappa_i \leftarrow \langle r_i, y_i \rangle / \langle p_i, q_i \rangle$
  $\delta m_{i+1} \leftarrow \delta m_i + \kappa_i p_i$
  $r_{i+1} \leftarrow r_i + \kappa_i q_i$
  $y_{i+1} \leftarrow \mathcal{L}^\dagger \mathcal{L}^{\dagger^T} r_{i+1}$
  $\eta_{i+1} \leftarrow \langle r_{i+1}, y_{i+1} \rangle / \langle r_i, y_i \rangle$
  $p_{i+1} \leftarrow -y_{i+1} + \eta_{i+1} p_i$
  $i \leftarrow i + 1$
Physical decomposition: for *each* $\mathbf{x}$ *position* **do**
  $\psi \leftarrow \mathcal{R}(\delta m)$
  for *each depth position* **do**
    Construct $\mathbf{G}$, $\mathbf{R}$
    $\delta \mathbf{u} \leftarrow \mathbf{G}^{-1}\mathbf{R}$

each example. The stopping criterion for the CG method is a predefined number of iterations, whereas, for the PCG method, the number of iterations is determined such that the optimization achieves approximately the same data misfit as the CG method. Table 6.2 details the number of iterations and the final computational time (in hours) for each numerical example demonstrating that CG-LSRTM requires at least twice more iterations (thus computational time) than its preconditioned version. To evaluate the quality of the inverted images, we calculate the normalized correlation coefficient (NCC) between the true and inverted model parameters.

6.6.1 Simple model

We start with a simple model containing four layers in the impedance model and three layers in the density model (Figure 6.3a, b). The background models are obtained by smoothing the exact models with a 2D Gaussian filter of 60 m length along the two dimensions (Figure 6.3c, d).

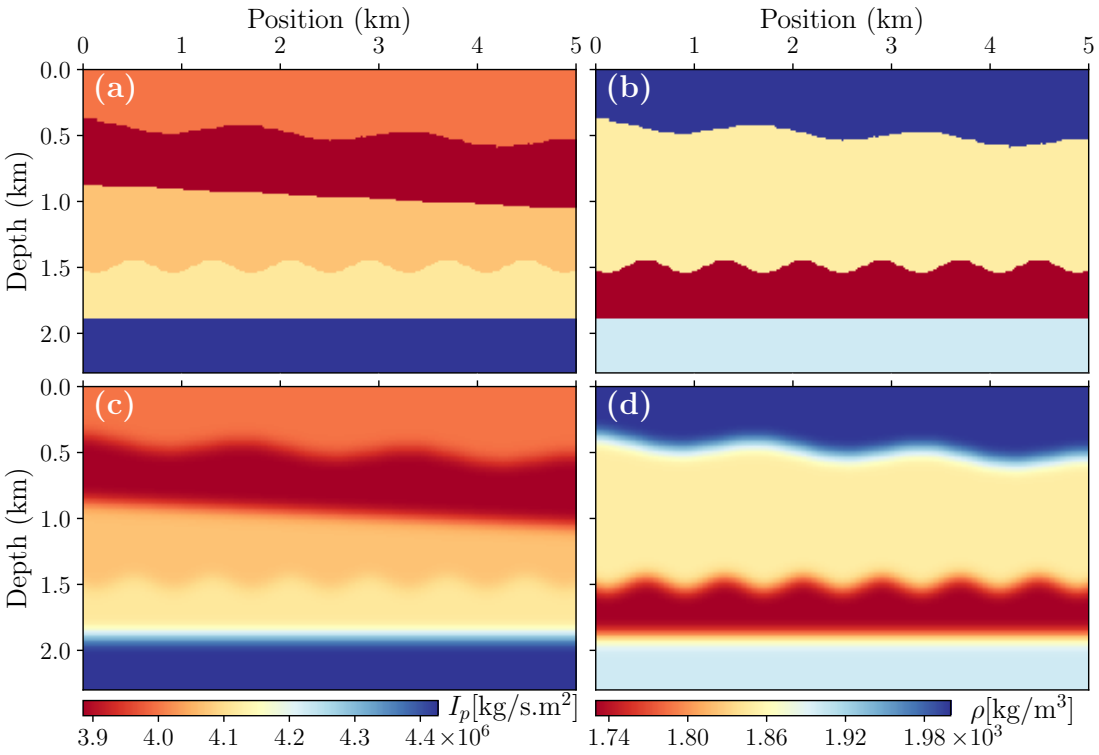

Figure 6.3: Simple model experiment: the exact a) P-impedance and b) density models, and background c) P-impedance and d) density models.

Table 6.1: Main parameter configuration for the different numerical experiments.

	Simple model	Dipping layer model	Marmousi2 model
Model grids ($n_x \times n_z$)	501×231	677×272	558×199
Grid intervals (m)	10	12.5	17.5
Source intervals (m)	30	50	52.5
Receiver intervals (m)	10	12.5	17.5
Number of sources	167	170	186
Number of receivers	501	677	558
Time sampling rate (ms)	2.1	2.8	2
Max. recording time (s)	2.9	5.2	6
Wavelet max. freq. (Hz)	30	20	20

Table 6.2: Comparison of the number of iterations and total computational time (in hours) for CG- and PCG-LSRTM.

	CG-LSRTM		**PCG-LSRTM**	
Numerical exp.	Iter. num.	Total time	Iter. num.	Total time
Simple model	60	73.5	15	28.8
Dipping layer	40	63.1	15	37.0
Marmousi2	40	93.8	20	73.6

We first compare the inverted parameters for CG (60 iterations) and PCG (15 iterations) (Figure 6.4). The main structure is recovered. The density image obtained with CG-LSRTM exhibits cross-talk artifacts for the second reflector (Figure 6.4d), whereas these artifacts are efficiently reduced in the PCG-LSRTM image (Figure 6.4f). The NCC values written on each panel also confirm the superiority of the PCG-LSRTM. To get more insight into the evaluation of the inversion through iterations, we display the inverted impedance (Figure 6.5) and density (Figure 6.6) images for 1, 15 and 30 iterations of CG-LSRTM, and 1, 5, and 10 iterations of PCG-LSRTM. Despite being in the early iterations, the NCC values exhibit a significant jump in quality for the PCG-LSRTM approach. This is more noticeable in the density images, as the preconditioner manages to mitigate the parameter cross-talk from the first iteration (Figure 6.6d). But to fully resolve the subsurface and remove the imprint of the impedance on the density model

(Figure 6.6f), more PCG iterations are required which indeed further supports our motivation to implement pseudoinverse Born operator in an iterative fashion (PCG-LSRTM).

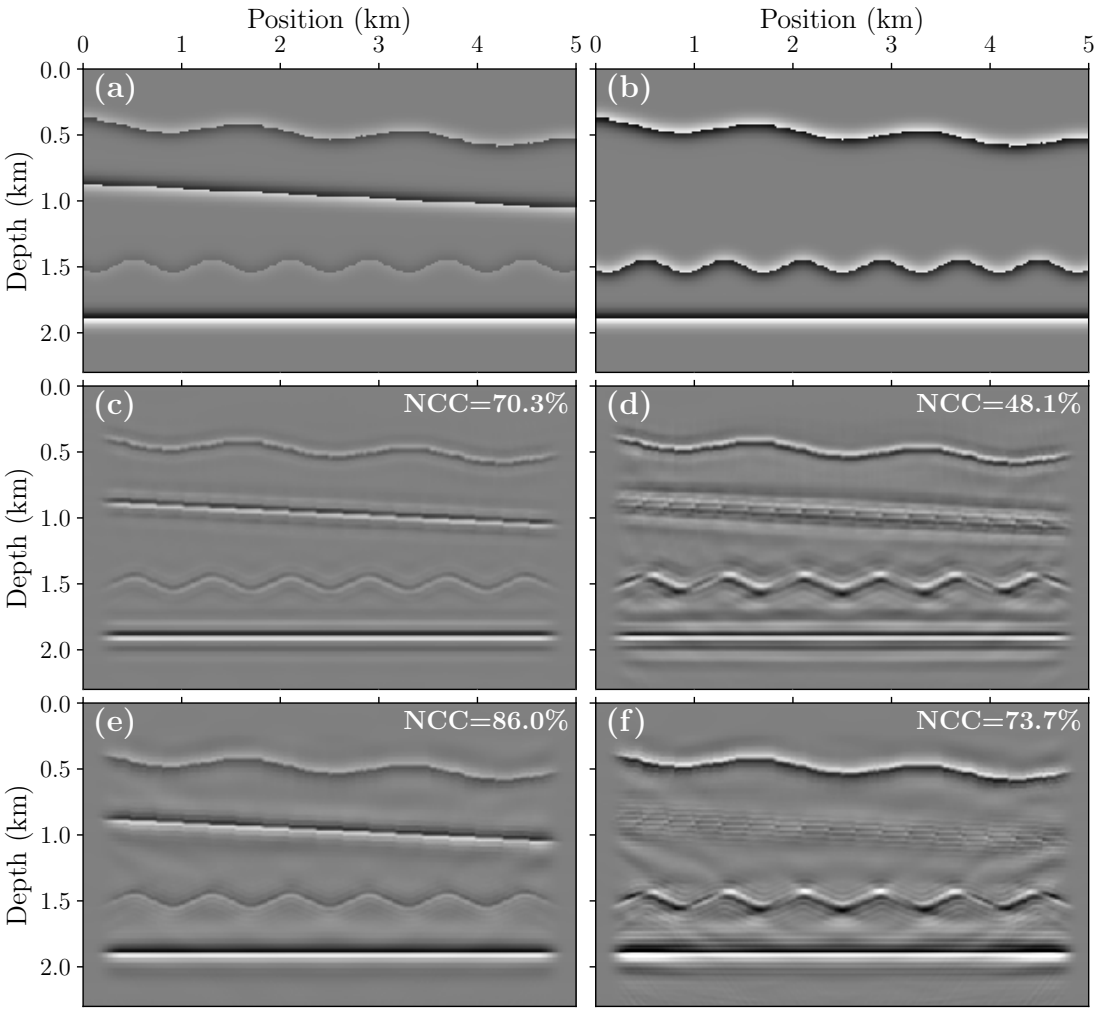

Figure 6.4: Relative model perturbation ζ (equation 6.6) for the exact a) impedance ζ_{I_p} and b) density ζ_ρ, and inverted c) ζ_{I_p} and d) ζ_ρ after 60 iterations of CG-LSRTM, inverted e) ζ_{I_p} and f) ζ_ρ after 15 iterations of PCG-LSRTM. Each inverted model is plotted with the same scale as for the exact model.

To quantitatively compare the ability of CG- and PCG-LSRTM to preserve the amplitude variations with angle (AVA) behavior in the observed wavefield, we compare the amplitude of the angle-domain CIGs extracted for the first two interfaces with the computed reflection coefficient using the Zoeppritz equations (Figure 6.7). Clearly, the AVA responses estimated by both methods closely follow the ones predicted by the Zoeppritz equations, particularly at near angles (up to 35°). At far angles (above 35°), the

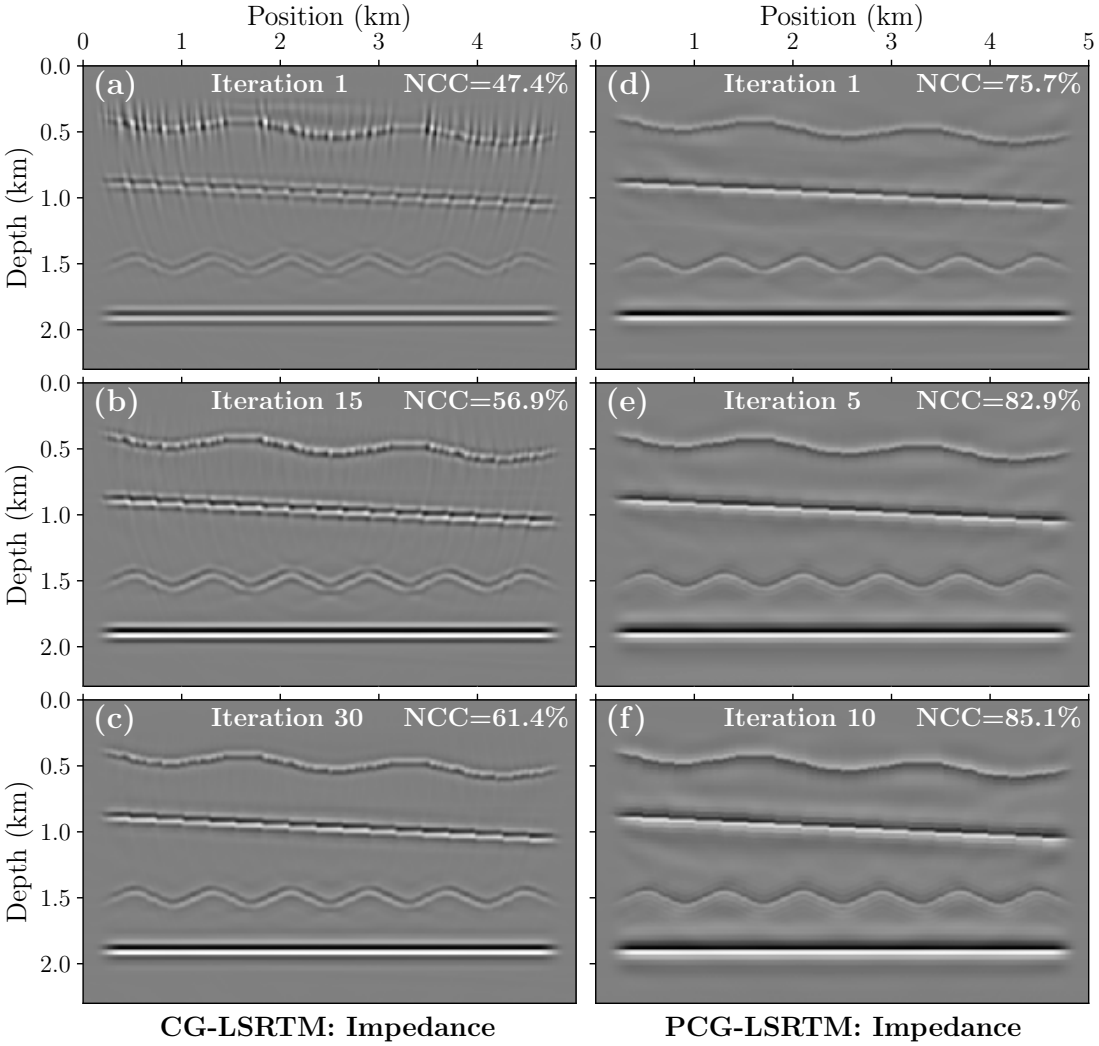

Figure 6.5: Comparison of inverted impedance perturbation ζ_{I_p} images after a) 1, b) 15 and c) 30 iterations of CG-LSRTM, d) 1, e) 5 and f) 10 iterations of PCG-LSRTM.

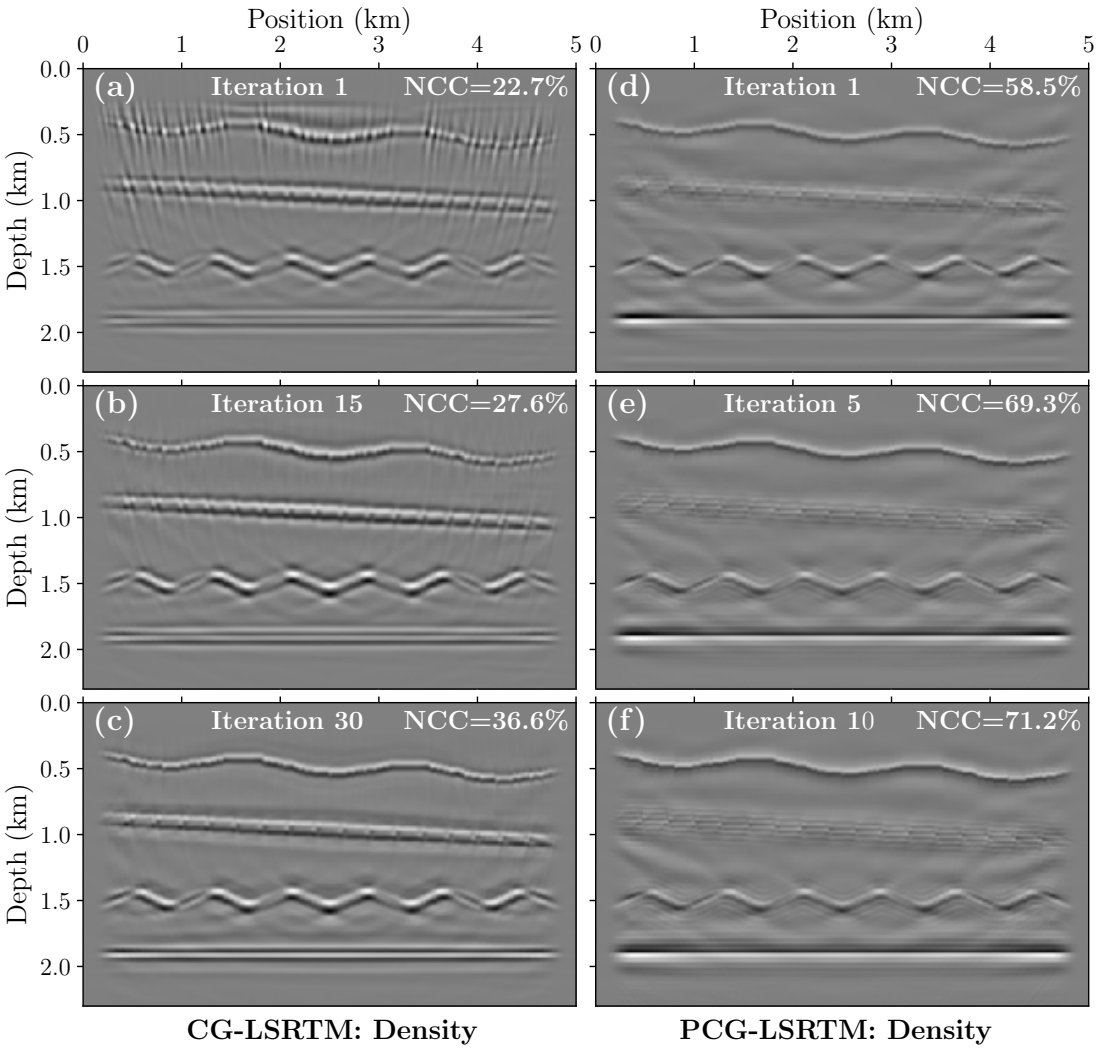

Figure 6.6: Same as Figure 6.5, but for the density model. Inverted density perturbation ζ_ρ images after a) 1, b) 15 and c) 30 iterations of CG-LSRTM, d) 1, e) 5 and f) 10 iterations of PCG-LSRTM.

AVA response estimated by CG-LSRTM starts to diverge from the theoretical Zoeppritz response, whereas PCG-LSRTM manages to improve the measured AVA behavior by better matching the Zoeppritz response. This further emphasizes the impact of the preconditioner in constructing true-amplitude subsurface images as the WLS method is the same for both methods. Finally, we plot the convergence history for relative data misfit (Figure 6.8a) and image reconstruction (Figure 6.8b, c), both exhibiting a remarkable acceleration of PCG- over CG-LSRTM.

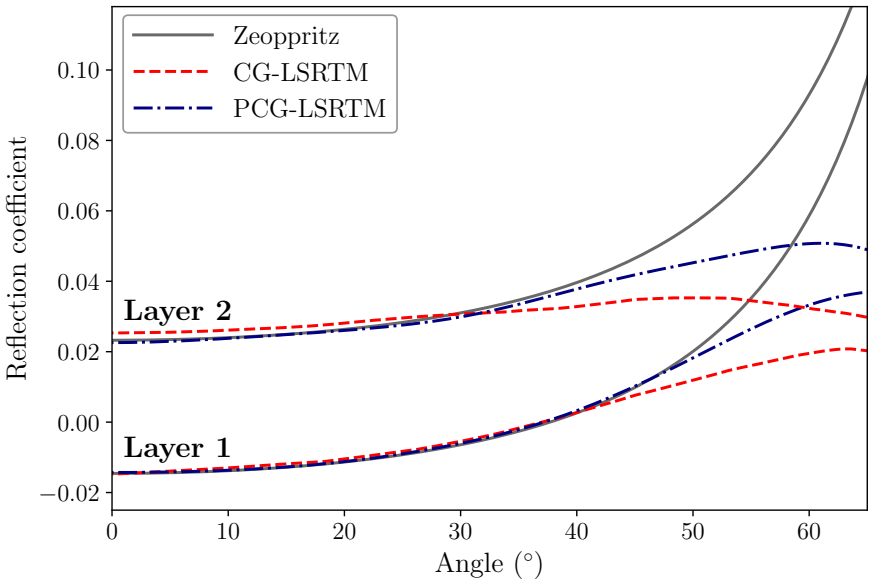

Figure 6.7: Comparison of AVA response for the first two layers of Figure 6.4 after 60 iterations of CG-LSRTM and 15 iterations of PCG-LSRTM.

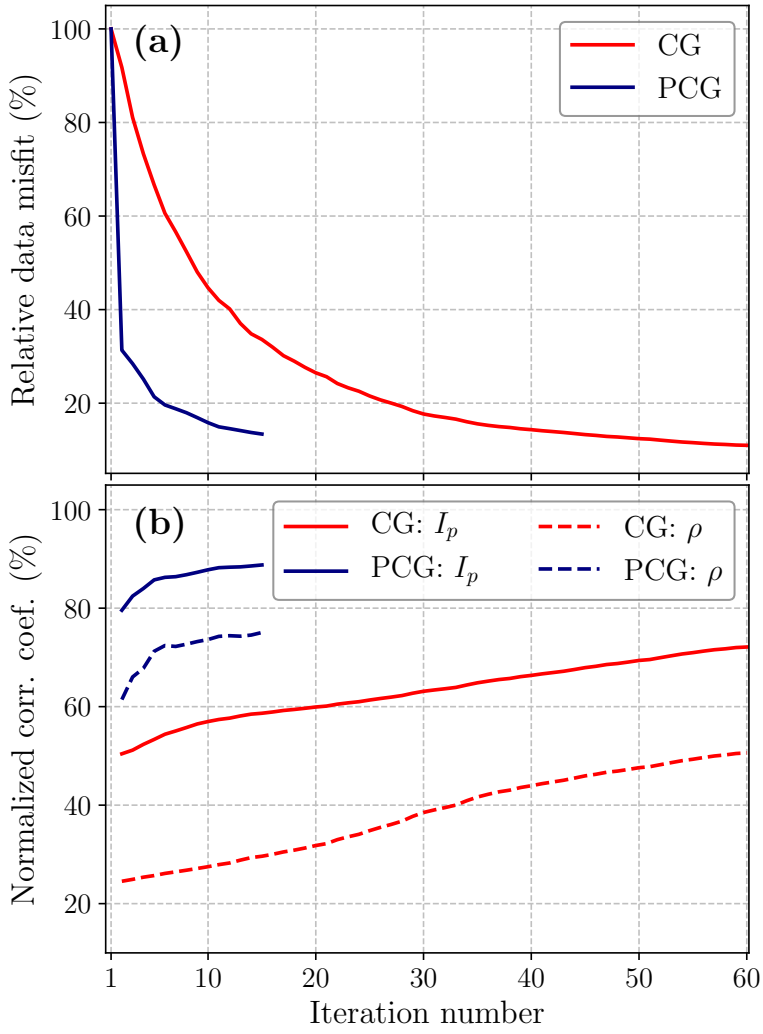

Figure 6.8: Evaluation of a) relative data misfit, and b) normalized correlation coefficients of P-impedance and density models with respect to the iteration number.

6.6.2 Dipping layer model

Our second experiment is based on a dipping layer model consisting of dense structures (Figure 6.9a, b). The main target in this model is a gas-charged sand channel (indicated by black arrow in Figure 6.9a, b), with local velocity and density perturbations. In the first test, we consider accurate background models built by applying a 2D Gaussian filter (125 m long in both directions) to the exact models (Figure 6.9c, d). The resultant inverted images after 40 CG-LSRTM iterations and 15 PCG-LSRTM iterations are illustrated in Figure 6.10. As for the previous case, this figure clearly shows that the preconditioning is effective in reducing artifacts and providing more accurate subsurface images. As a quantitative quality control, we extract vertical profiles across the gas-charged sand channel and add them to the background models (Figure 6.11). We also calculate the root-mean-square errors (RMSE) between the true and inverted parameters (written on Figure 6.11). Clearly, PCG-LSRTM manages to better reconstruct the anomaly than CG-LSRTM.

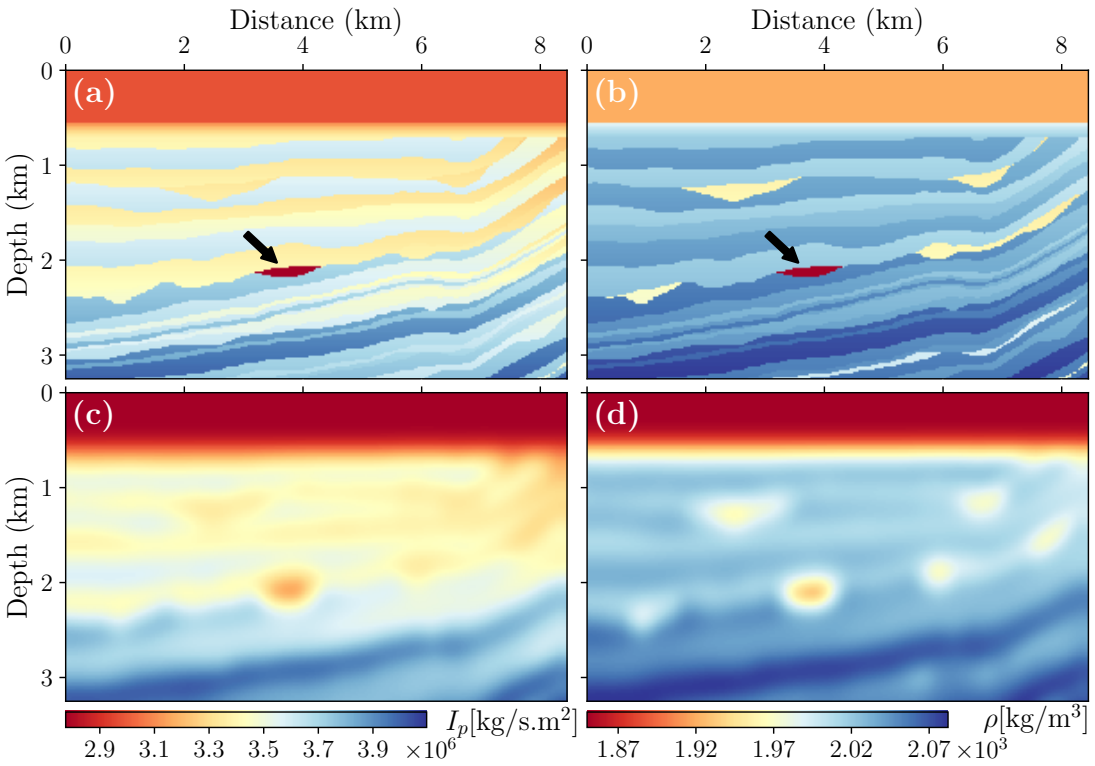

Figure 6.9: Dipping layer model experiment: exact models of a) P-impedance and b) density, and background models of c) P-impedance and d) density. The black arrows indicate the position of the gas-charged sand channel.

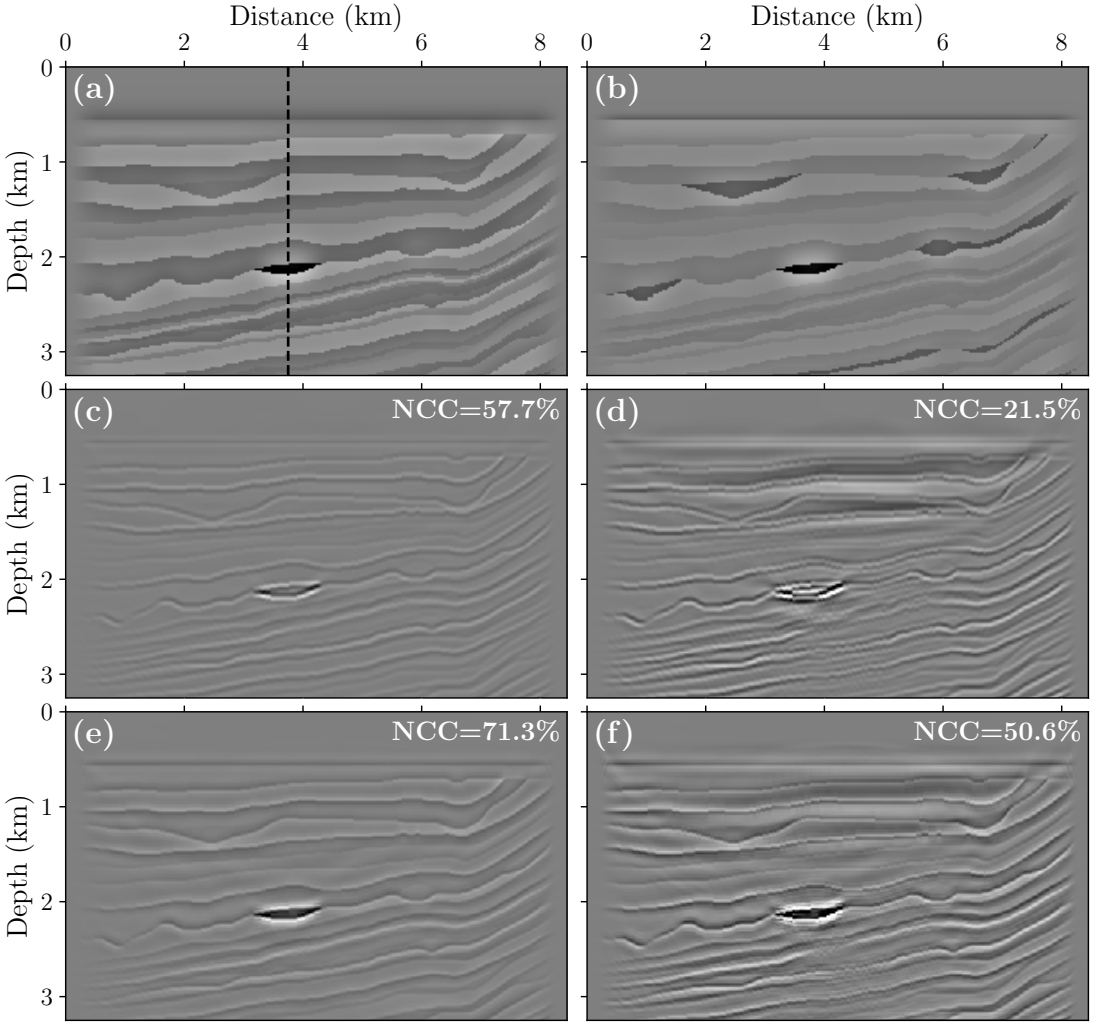

Figure 6.10: The exact model perturbations a) ζ_{I_p} and b) ζ_ρ, and inverted c) ζ_{I_p} and d) ζ_ρ via 40 iterations of CG-LSRTM, inverted e) ζ_{I_p} and f) ζ_ρ via 15 iterations of PCG-LSRTM. Each inverted model is plotted with the same scale as for the exact model. The dashed line in (a) indicates the position of the extracted trace.

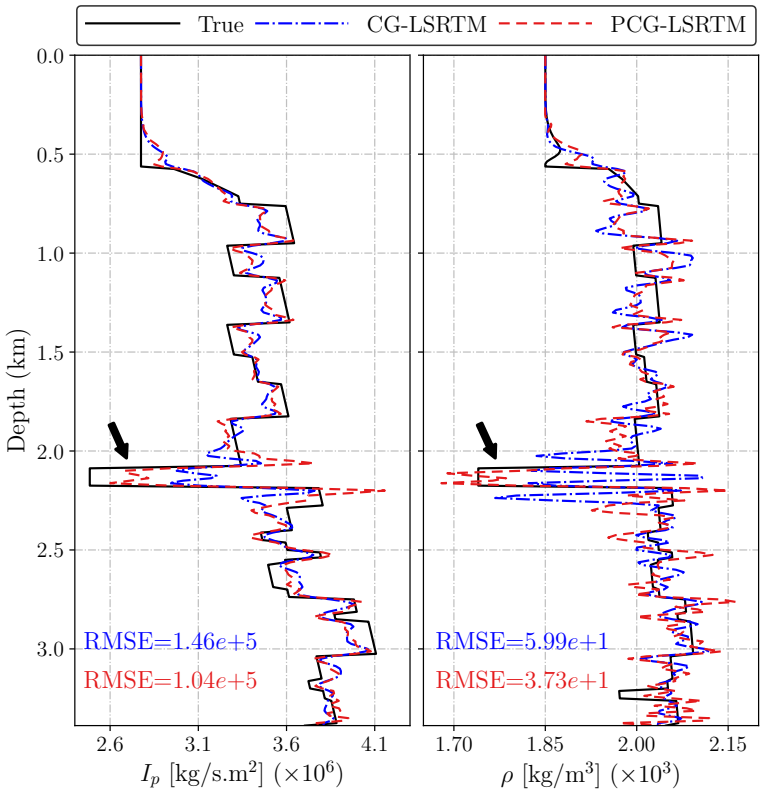

Figure 6.11: Dipping layer model experiment: extracted vertical profiles at $\mathtt{x} = 3.75$ km from inverted P-impedance (left) and density (right) models. The black arrow indicates the gas-charged sand channel.

Figure 6.12 illustrates the relative data misfit and the NCC convergence curves as a function of iteration. Clearly, the proposed PCG-LSRTM has a substantially faster convergence rate in terms of data and image reconstruction.

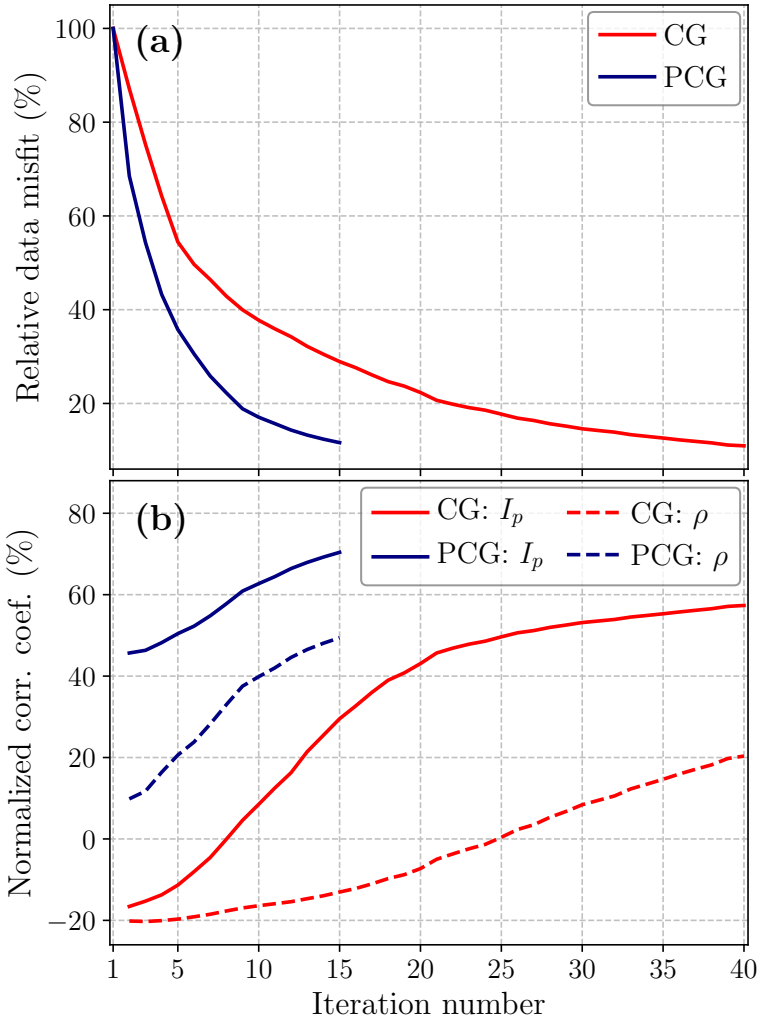

Figure 6.12: Dipping layer model experiment: evaluation of a) relative data misfit, and b) normalized correlation coefficients of P-impedance and density models with respect to the iteration number.

6.6.2.1 Sensitivity to background models error

In field data applications, inaccuracies in the migration background models are inevitable (Yang et al., 2019a). Assuming an erroneous velocity model, the effectiveness

of the pseudoinverse Born operator as a preconditioner has already been investigated for constant-density acoustic medium (Hou and Symes, 2016a). Here, we evaluate the robustness of proposed PCG-LSRTM in the presence of inaccurate impedance and density background models.

The incorrect background models are built by highly smoothing the exact models with a 2D Gaussian filter (375 m long, Figure 6.13). To get an insight in the quality of the background models, we first compare the subsurface offset- and angle-domain CIGs estimated by PCG-LSRTM (Figure 6.14) with the one obtained in our previous test (accurate background models). For the case of accurate background models, as expected, the specular energies are focused around the zero offset on the subsurface offset-domain CIG (Figure 6.14a), and produce flat events on the angle-domain CIG (Figure 6.14b). In the case of inaccurate background models, these energies exhibit a downward curvature in the subsurface offset-domain CIG (Figure 6.14c) and an upward curvature in the angle-domain CIG (Figure 6.14d), implying that the background velocity model is lower than the exact one. This is the principle for migration velocity analysis (Symes, 2008b). Same as the previous test, we compare the inverted images after 40 CG-LSRTM iterations and 15 PCG-LSRTM iterations (Figure 6.15). Although the errors in the background models degrade the quality of the inversion (lower NCC), again PCG-LSRTM outperforms CG-LSRTM in both the inverted P-impedance and density images. The convergence history plots (Figure 6.16) also illustrate the effectiveness of our preconditioner in presence of erroneous background models.

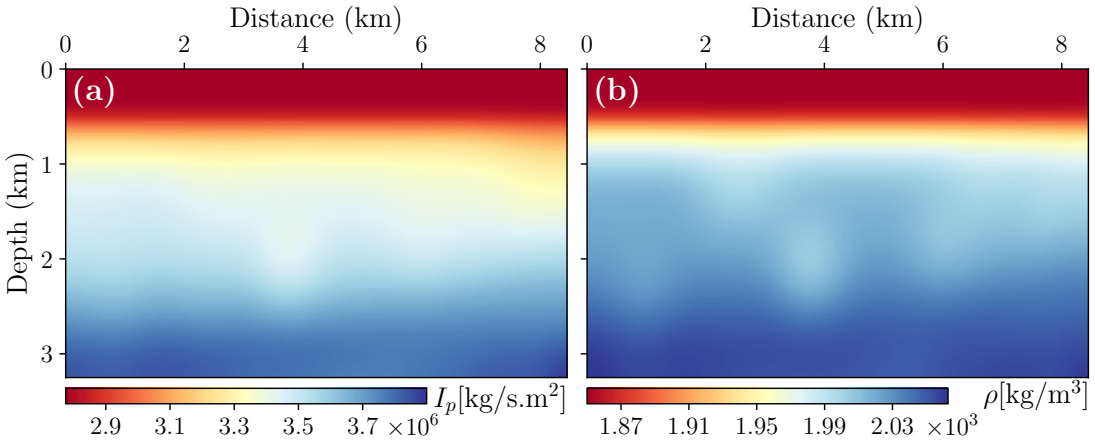

Figure 6.13: Dipping layer model experiment: inaccurate migration background models for a) P-impedance and b) density.

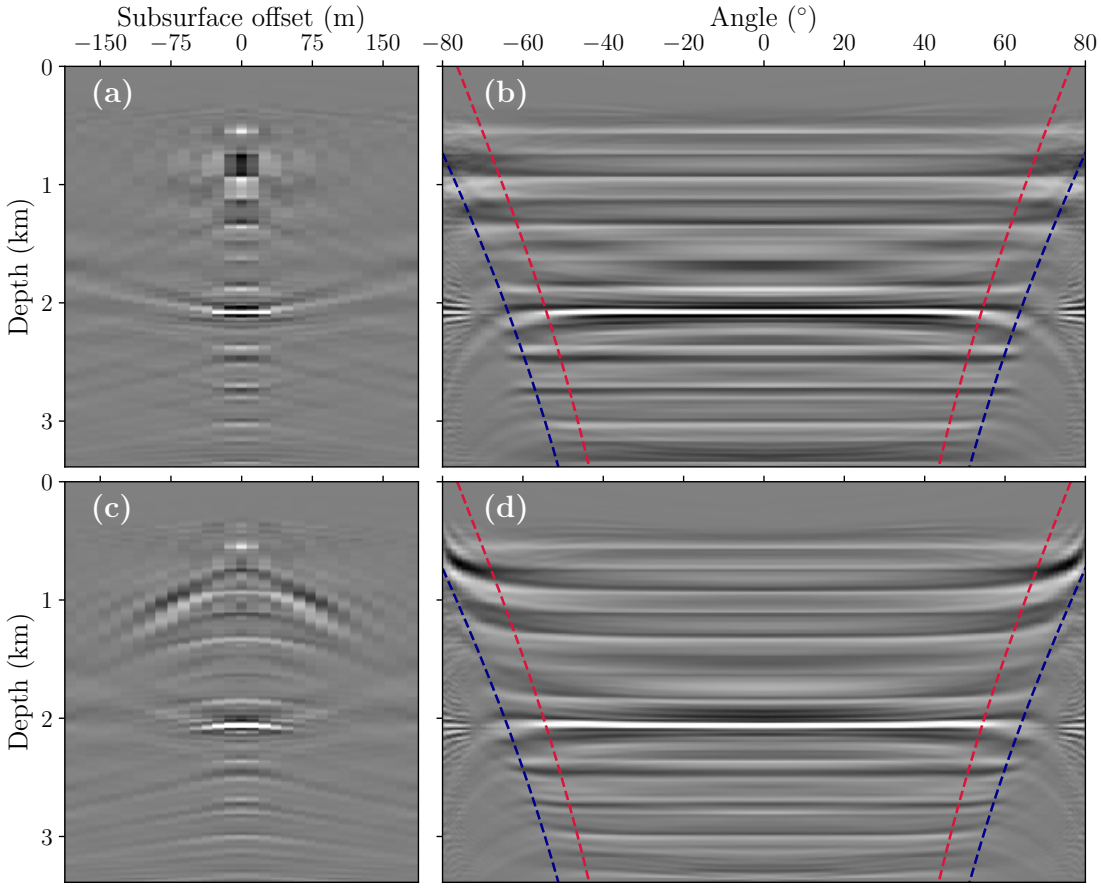

Figure 6.14: Dipping model experiment: the estimated (a,c) subsurface offset-domain CIG and (b,d) its angle-domain CIG after 15 iterations of PCG-LSRTM assuming (a,b) accurate and (c,d) inaccurate background models. The blue and red lines in the angle-domain CIGs correspond to the theoretical limit for acquired angles ($\alpha = 1$, equation 6.9) and the practical ones used in inversion ($\alpha = 0.85$), respectively.

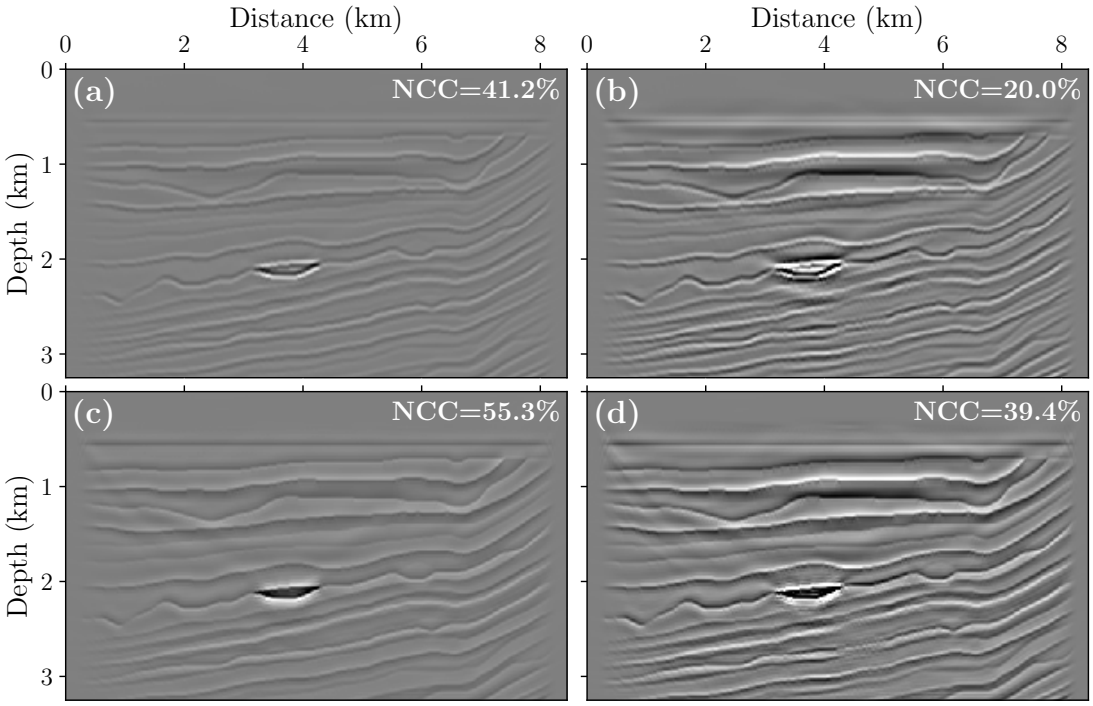

Figure 6.15: Same as Figure 6.10, but for inaccurate background models: inverted a) ζ_{I_p} and b) ζ_ρ via 40 iterations of CG-LSRTM, inverted c) ζ_{I_p} and d) ζ_ρ via 15 iterations of PCG-LSRTM. Each inverted model is plotted with the same scale as for the exact model.

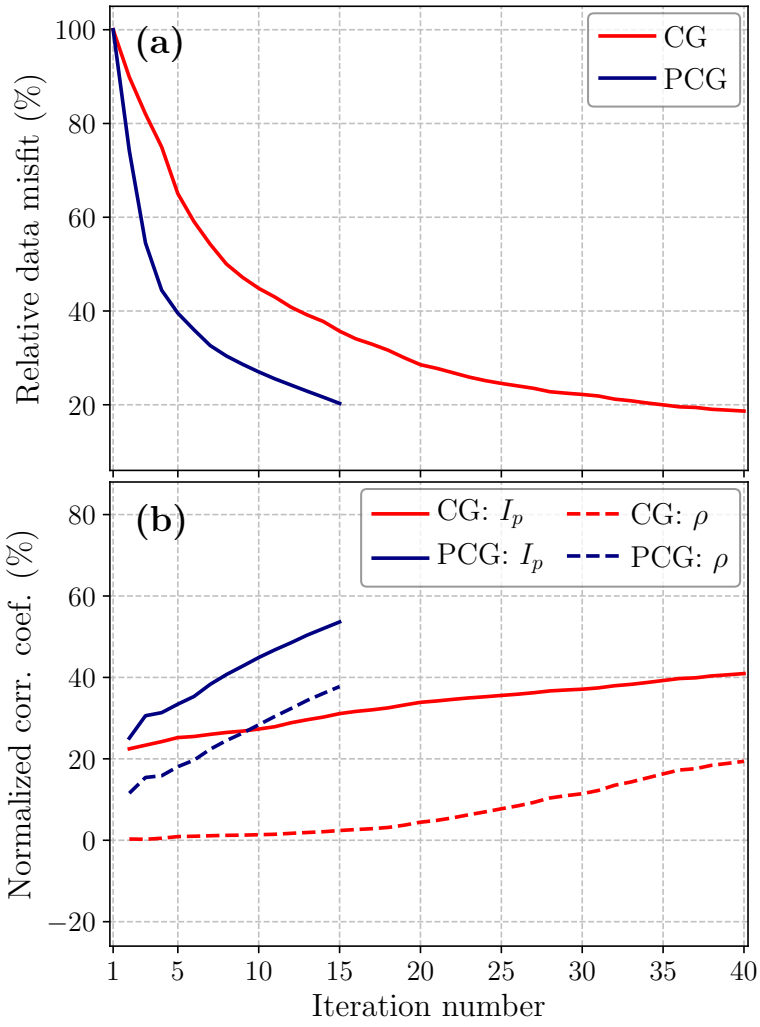

Figure 6.16: Same as Figure 6.12, but for inaccurate background models: evaluation of a) relative data misfit, and b) normalized correlation coefficients of P-impedance and density models with respect to the iteration number.

6.6.3 Marmousi2 model

In our last numerical experiment, we consider a more realistic model Marmousi2 as a benchmark for migration/inversion methods (Martin et al., 2006). To reduce the computational cost, we conduct our test only on the central target of this model (Figure 6.17). The background models are derived from smoothing the exact models using a 175 m long 2D Gaussian kernel (Figure 6.17c, d).

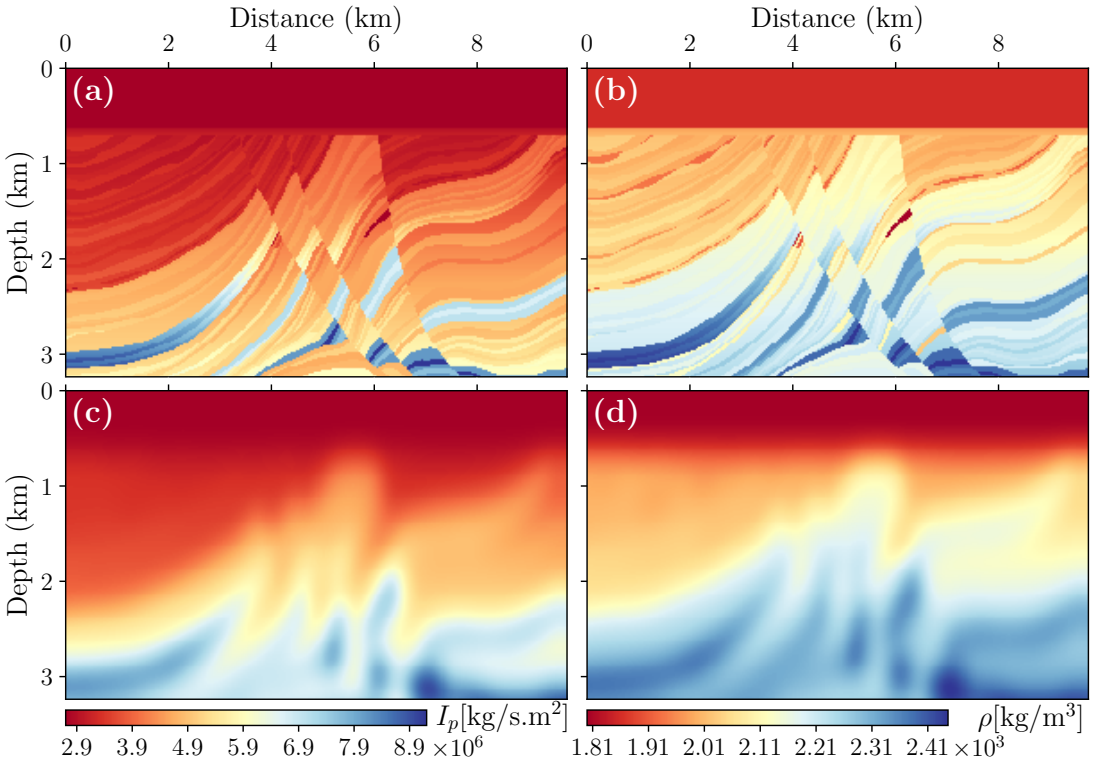

Figure 6.17: The Marmousi2 experiment: exact models of a) P-impedance and b) density, and background models of c) P-impedance and d) density.

We first compare the angle-domain CIGs estimated by the application of the Radon transform (Figure 6.18). We highlight the main differences with arrows in Figure 6.18 showing that the angle gathers associated with PCG-LSRTM illustrate better correlation compared to the ones for CG-LSRTM. We now compare inverted images with 40 CG-LSRTM iterations and 20 PCG-LSRTM iterations (Figure 6.19). The color scales used for the inverted images are the same as the one used for the true models. Clearly, the quality of the inverted images is remarkably improved with preconditioning compared with the ones obtained without preconditioning. To get more insight into the areas of the discrepancies between the true and inverted models, we calculate localized NCC values

within a 0.4 km long 2D window (Figure 6.20). At first glance, we notice that the inverted P-impedance models have a higher correlation compared to the inverted density models. A detailed look reveals that PCG-LSRTM achieves better correlation in most of both the P-impedance and density models (Figure 6.20c, d). To further evaluate the quality of the inverted parameters, we compare 2D wavenumber spectra (Figure 6.21). The spectra obtained with CG-LSRTM clearly present a gap in both the low- and high wavenumbers (Figure 6.21b, e). On the other hand, the preconditioner tends to fill these gaps by broadening the frequency band (Figure 6.21c, f). As a final evaluation of the inversion performance of both methods, we plot the convergence curves demonstrating a faster convergence rate for PCG-LSRTM in both the data and image domains (Figure 6.22).

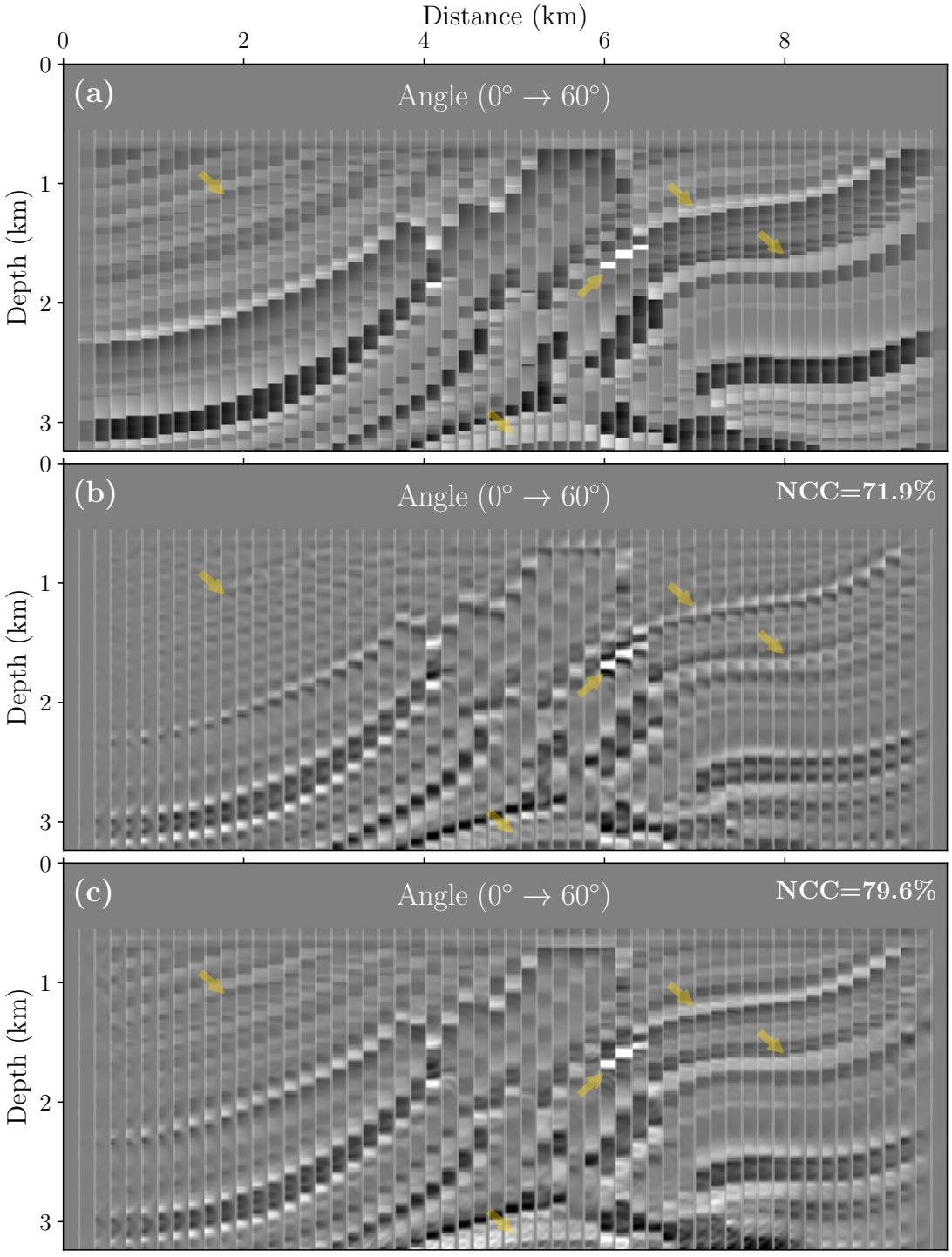

Figure 6.18: The Marmousi2 experiment: a) true angle-domain CIGs and the ones built by application of the Radon transform using inverted model via b) 40 CG-LSRTM iterations and c) 20 PCG-LSRTM iterations. The two successive vertical lines (from left to right) indicate angles of range 0 to 60 degrees. The arrows highlight the main differences.

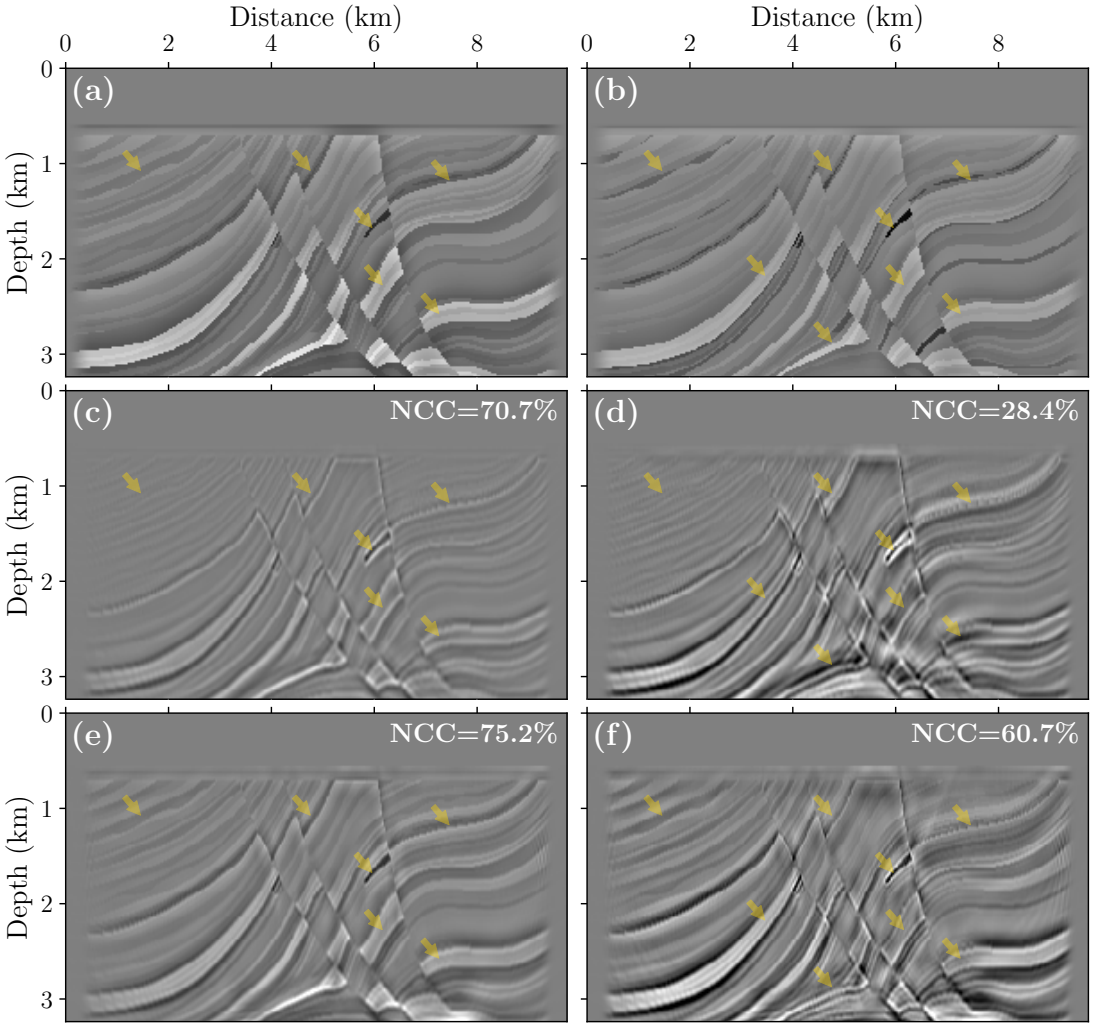

Figure 6.19: The exact model perturbations a) ζ_{I_p} and b) ζ_ρ, and inverted c) ζ_{I_p} and d) ζ_ρ via 40 CG-LSRTM iterations, inverted e) ζ_{I_p} and f) ζ_ρ via 20 PCG-LSRTM iterations. The arrows highlight the main differences.

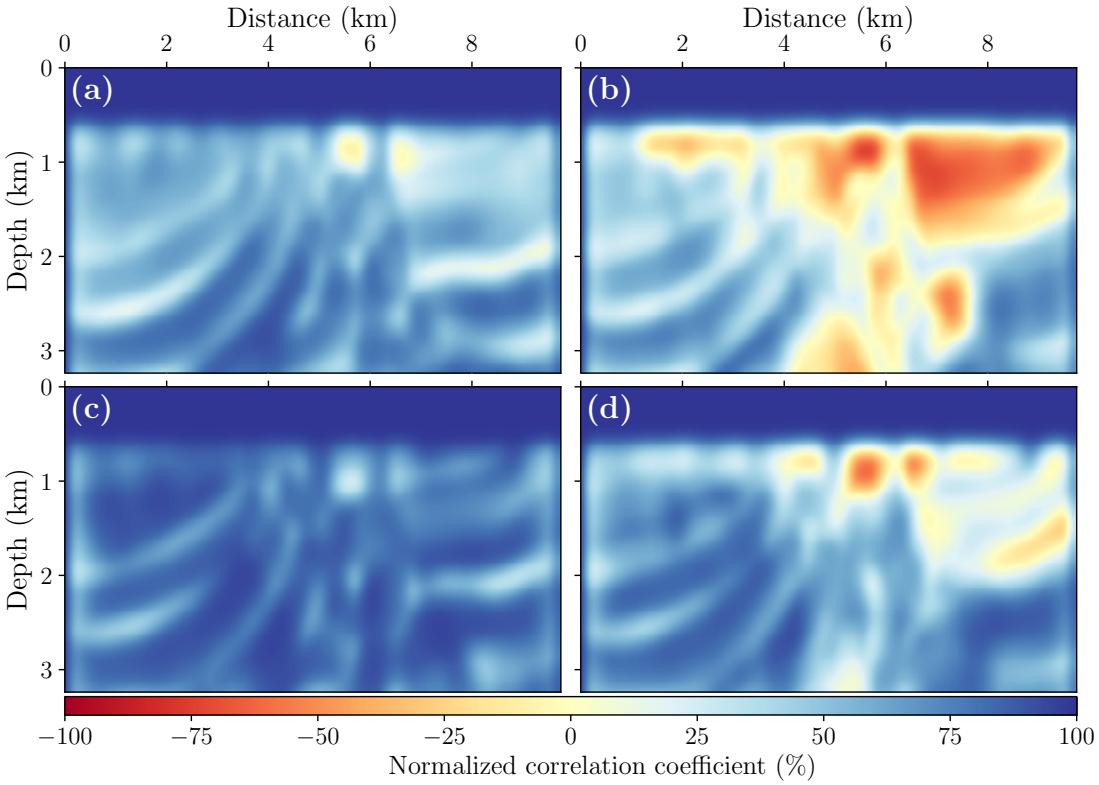

Figure 6.20: The localized NCC values corresponding to Figure 6.19: for (a,c) ζ_{I_p} and (b,d) ζ_ρ via (a,b) CG-LSRTM and (c,d) PCG-LSRTM.

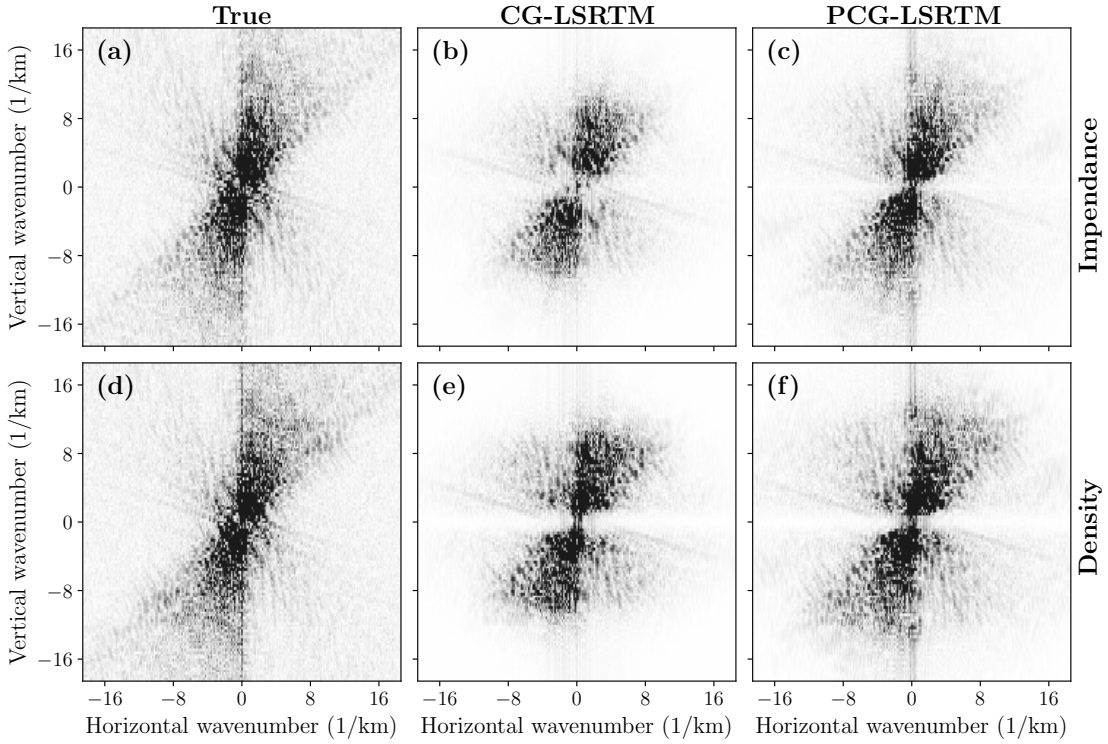

Figure 6.21: The Marmousi2 experiment: wavenumber spectra of impedance corresponding to a) true, b) CG-LSRTM and c) PCG-LSRTM, and density corresponding to d) true, e) CG-LSRTM and f) PCG-LSRTM. All panels are plotted with the same amplitude scale.

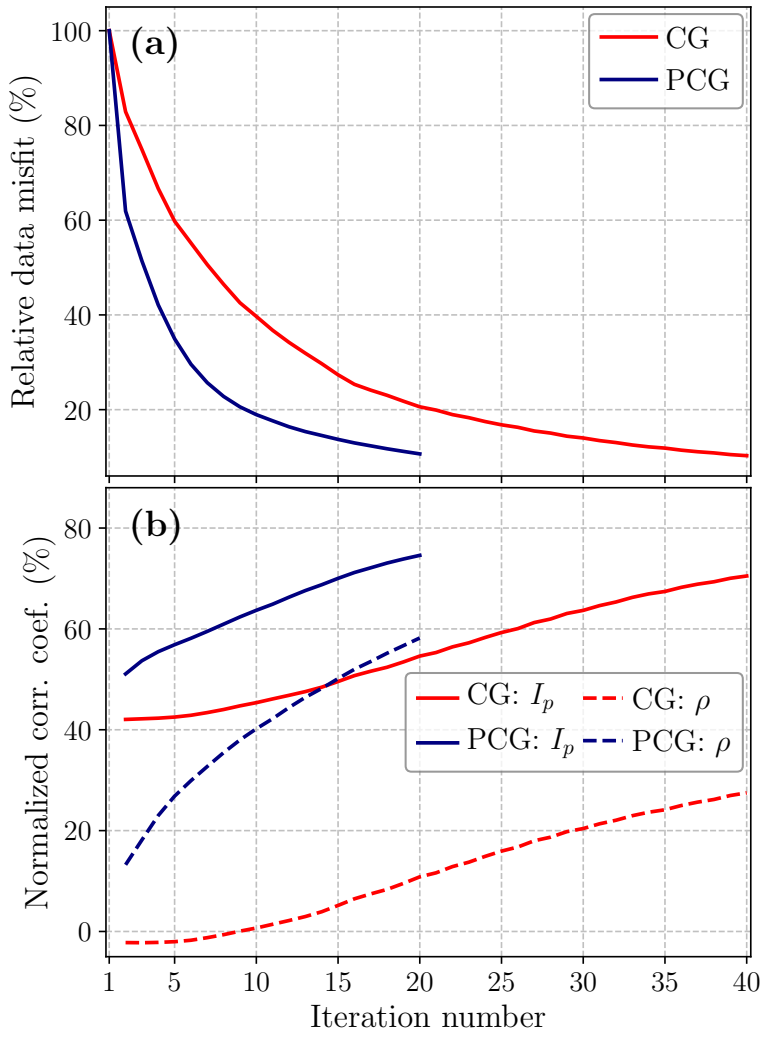

Figure 6.22: The Marmousi2 experiment: evaluation of a) relative data misfit, and b) normalized correlation coefficients of P-impedance and density models with respect to the iteration number.

6.7 Discussion

We have proposed a method to accelerate the convergence rate of LSRTM while mitigating cross-talk artifacts. We have used the pseudoinverse Born operator to preconditioning the variable-density LSRTM. The WLS method is remained the same for CG- and PCG-LSRTM. In this section, we discuss several interesting aspects of this subject that merit additional investigation, including the structure of the multi-parameter Hessian and possible ways to reduce the computational cost.

In the concept of constant-density acoustics, Biondi et al. (2019) demonstrated that the extended image space is able to preserve the elastic behavior of the primary reflections even when these events are acoustically migrated using LSRTM. They illustrated that the AVA response of the extended LSRTM estimated through the Radon transform nicely matches the theoretical Zoeppritz response up to a scaling factor. Therefore, our proposed method in this article not only extends the applicability of Biondi et al. (2019) from constant- to variable-density acoustics but also accelerates its convergence rate and provides physical subsurface parameters.

An interesting link can be made from the WLS method to AVO inversion methods utilizing surface offsets. The latter, based on Zoeppritz equations or their approximations, impose some limitations. Assuming a convolutional isotropic layered Earth model, they are only valid for relatively small angles of incidence ignoring geometrical spreading, scattering/transmission loss, wave mode conversion, multiples and thin layering effect (Mallick, 2001). Moreover, they only hold for relatively small elastic-property contrasts across an interface for the plane-wave propagation. These assumptions, though, are not met in realistic applications and need sophisticated data processing steps, particularly for structurally complex models (Gholami et al., 2018). The WLS method, on the other hand, not only does not have any restriction on the maximum offset/angle but also better estimates model parameters utilizing a larger maximum offset/angle. In this case, combination of LSRTM and WLS methods can be categorized as a true-amplitude migration/inversion scheme providing quantitative subsurface parameters. Note that the accuracy of the quantitative parameter estimation for both WLS and AVO inversion methods depends on the quality of the background models (long wavelengths). More research is needed to effectively compare these methods.

As noted through all numerical experiments, a significant difference in image quality (NCC) was observed between CG- and PCG-LSRTM approaches at each iteration. A careful reader will have noticed that although the final data misfits for both methods might be approximately equal, the quality of the inverted parameters via PCG-LSRTM are much higher than the ones estimated by CG-LSRTM. The problem is formulated as a least-squares misfit in the data domain (equation 6.12). The preconditioner also plays a role on the reconstructed model side. The pseudoinverse Born operator in PCG-LSRTM accounts for the finite frequency nature of the source wavelet, geometric spreading, uneven illumination, and at least to some extend, parameter cross-talks. The CG-LSRTM

compensates for these points through CG iterations. Consequently, any incorrect projection of migration artifacts onto the estimated images, especially the ones related to parameter cross-talk, will include features that will be very difficult to mitigate at later iterations (Virieux et al., 2017). Although reducing data misfit in later CG iterations is very small, these are essential to further improve the updated model for small details.

In the concept of multi-parameter imaging, the Hessian is a blockwise matrix (Operto et al., 2013). For variable-density acoustic case, it has two diagonal blocks and two off-diagonal blocks. Each block is a $(n_z n_x \times n_z n_x)$ square matrix where n_z and n_x denote the number of the grids in the vertical and horizontal directions, respectively. These blocks have a banded nature due to finite frequency effects. The diagonal elements of the diagonal blocks mainly account for geometric spreading and illumination compensation. The off-diagonal blocks have mathematical capability to reduce the parameter cross-talks (Métivier et al., 2014; Jun et al., 2015; Hagen et al., 2018). The full Hessian matrix can be expressed as

$$\mathbf{H} = \begin{bmatrix} \mathbb{H}_{I_p,I_p} & \mathbb{H}_{I_p,\rho} \\ \mathbb{H}_{\rho,I_p} & \mathbb{H}_{\rho,\rho} \end{bmatrix}, \tag{6.20}$$

where $\mathbb{H}_{i,j}$ describes the effect of perturbation in the parameter type i on the parameter type j. Here, we prefer to calling $\mathbf{H}$ approximate Hessian, since the way we construct it for multi-parameter imaging proposed in this article does not lead to the classical Hessian. This is because the construction of $\mathbb{H}_{i,j}$ requires the application of two processes: first extended image obtained with $\mathcal{L}^T\mathcal{L}$, and then WLS method (see Appendix 6.9 for more details). In practice, although the Hessian matrix $\mathbf{H}$ is computationally intractable, we construct this approximate Hessian explicitly for a relatively small-sized homogeneous model ($n_z = n_x = 47$) using $\mathcal{L}^T\mathcal{L}$ (without preconditioner) and $(\mathcal{L}^\dagger\mathcal{L}^{\dagger^T})(\mathcal{L}^T\mathcal{L})$ (with preconditioner) to get more insight on the effects of this preconditioning (Figure 6.23). For each grid point, we perturb one parameter by 0.05% while keeping the other one constant. This is known as the point spread function technique. A 30 Hz (maximum frequency) Ricker wavelet is used for modeling. The receivers are deployed on each grid point along the top surface of the model.

As expected, the Hessian without preconditioner is a band diagonal matrix due to finite-frequency effects, and the amplitude on the diagonal elements decreases due to the geometrical spreading (Figure 6.23a). On the other hand, the Hessian with preconditioner manages to reduce the effect of the wavelet and compensates for the geometrical spreading (Figure 6.23b). To further confirm this, we plot three elements of the diagonal subblocks of the Hessian matrices in the distance-depth domain (Figure 6.24). As illustrated, the specular energy is focused and the migration smiles are suppressed by using the preconditioner (Figure 6.24c, d). We also notice that the density reconstruction contains more oscillations. To investigate the parameter cross-talk, we normalize each

submatrix by its maximum value (Table 6.3). Comparing the true perturbation with the normalized value (Table 6.3) reveals that the preconditioner approximately holds equation 6.19 and can mitigate the cross-talk artifacts. We further analyze this by calculating the ratio $\mathbb{H}_{I_p,I_p}/\mathbb{H}_{I_p,\rho}$ and $\mathbb{H}_{\rho,\rho}/\mathbb{H}_{\rho,I_p}$ (Table 6.3). To reduce the parameter cross-talks, these ratios should diverge to infinity. The ratio values in Table 6.3 show that the Hessian with preconditioner is approximately 10 times better scaled than the one without preconditioner, and thus can better reduce the parameter cross-talks. This element, mitigating parameter cross-talks, was not present in the original definition of the pseudoinverse Born operator. In the context of FWI, researches have also shown promising performance on mitigating parameter cross-talks by incorporating the multi-parameter Hessian (Operto et al., 2013; Métivier et al., 2014; Jun et al., 2015).

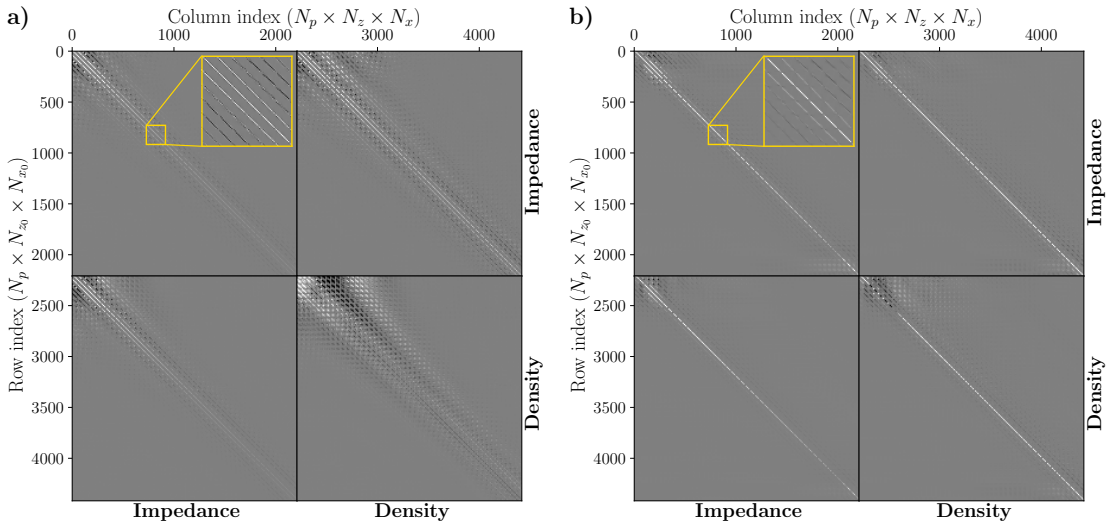

Figure 6.23: The multi-parameter approximated Hessian: a) without preconditioner and b) with preconditioner. The yellow boxes delineate a close-up structure of the Hessian.

Table 6.3: Maximum value of each block of the Hessian matrix.

	True model			**Without precon.**			**With precon.**		
Param.	Imp.	Den.	Ratio	Imp.	Den.	Ratio	Imp.	Den.	Ratio
Imp.	2.0e+3	0	∞	6.1e+17	2.5e+14	2.4e+3	1.7e+3	6.0$e-2$	2.8e+4
Den.	0	1.0e+0	∞	6.0e+16	3.0e+13	5.0e-4	1.6e+2	6.5$e-1$	4.0e-3

We now discuss how to reduce the computational cost. In terms of implementation, the high computational burden of $\mathcal{L}^\dagger\mathcal{L}^{\dagger T}$ is an unattractive aspect of the proposed

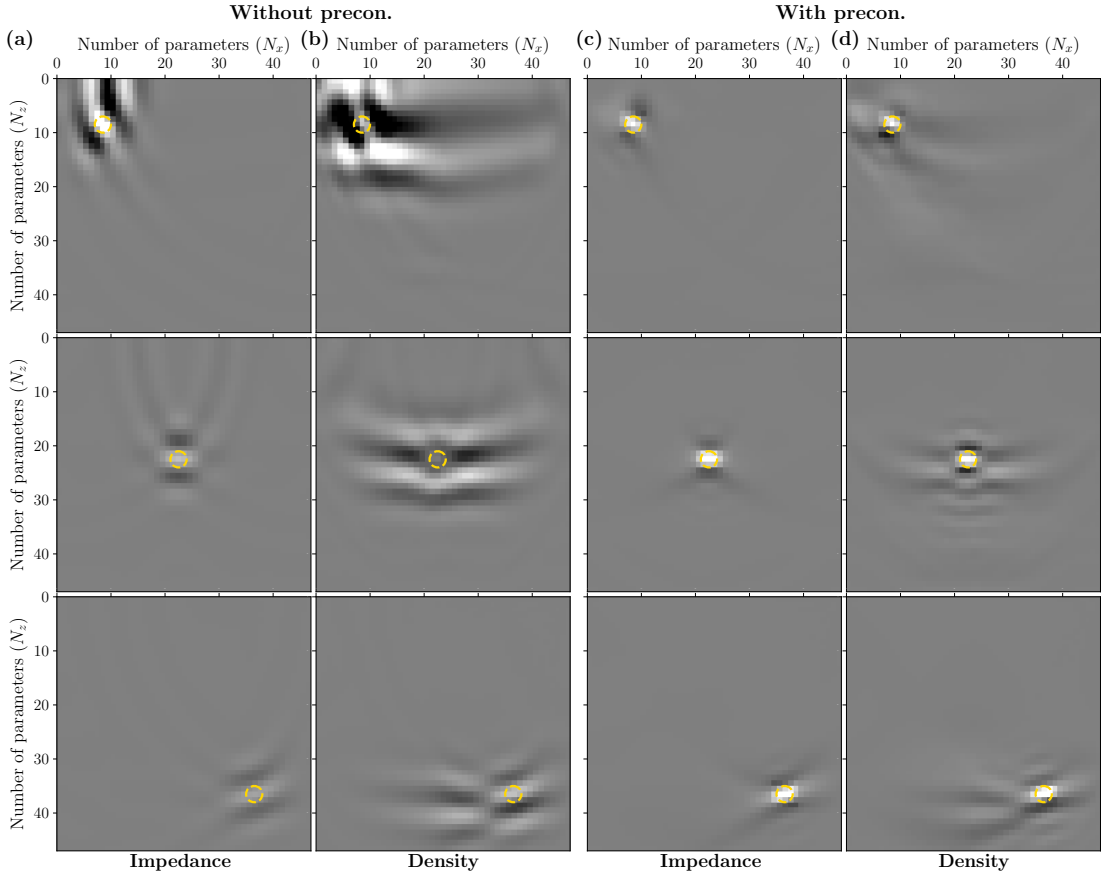

Figure 6.24: One element of the (a,c) impedance-impedance and (b,d) density-density block matrix of the Hessian in distance-depth domain corresponding to (a,b) without preconditioner and (c,d) with preconditioner. The first, second, and third rows correspond to the shallow, middle and deep part of the model, respectively. The yellow dashed circles denote the position of the perturbation.

preconditioning scheme. In the constant-density acoustic case, the pseudoinverse Born operators are constructed either by stationary phase approximation (Hou and Symes, 2015) or linearization of the phase of the normal operator (Chauris and Cocher, 2017). The former describes the pseudoinverse Born operator as a combination of the adjoint operator $\mathcal{L}^T$ with two weighting norms defined in the data and model spaces. Although the final formulas of both approaches are almost identical (Chauris and Cocher, 2018), expressing the pseudoinverse with such weighted norms has the flexibility of easy coupling to the standard CG iterations as weighted inner products (Hou and Symes, 2016a). For the reader's convenience, we include this scheme in Appendix 6.10. The other possible way to reduce the computational burden is to adopt a random subsurface-offset sampling procedure to construct the subsurface offset-domain CIGs (van Leeuwen et al., 2015).

6.8 Conclusion

We have implemented the pseudoinverse Born operator as a preconditioner for variable-density least-squares reverse time migration. The pseudoinverse operator asymptotically accounts for multiparameter Hessian, explaining limited bandwidth effects, geometrical spreading, poor illumination, and parameter cross-talks. Although the proposed preconditioning scheme has higher computational burden in each Conjugate Gradient iteration, numerical experiments clearly prove the efficiency of the proposed algorithm even in presence of parameter cross-talk and inaccurate background models: for almost equal data misfit, standard least-squares reverse time migration requires at least twice more iterations than the preconditioned version, while the quality of the images obtained with the former is much lower than the one with the latter. Future work will explore coupling multi-parameter pseudoinverse Born operator to other iterative imaging techniques such as migration velocity analysis and full waveform inversion.

6.9 Appendix I: Derivation of Hessian

The exact definition of the normal equation reads as (Pratt et al., 1998)

$$\int \mathrm{d}\mathbf{x}_0 \begin{bmatrix} \mathbb{H}_{I_p,I_p}(\mathbf{x},\mathbf{x}_0) & \mathbb{H}_{I_p,\rho}(\mathbf{x},\mathbf{x}_0) \\ \mathbb{H}_{\rho,I_p}(\mathbf{x},\mathbf{x}_0) & \mathbb{H}_{\rho,\rho}(\mathbf{x},\mathbf{x}_0) \end{bmatrix} \begin{bmatrix} \zeta_{I_p}(\mathbf{x}_0) \\ \zeta_{\rho}(\mathbf{x}_0) \end{bmatrix} = \begin{bmatrix} \zeta_{I_p}^{mig}(\mathbf{x}) \\ \zeta_{\rho}^{mig}(\mathbf{x}) \end{bmatrix}, \tag{6.21}$$

where the diagonal blocks of the Hessian matrix are

$$\mathbb{H}_{I_p,I_p}(\mathbf{x},\mathbf{x}_0) = \frac{\partial^2 J(\zeta_{I_p},\zeta_\rho)}{\partial\zeta_{I_p}(\mathbf{x})\partial\zeta_{I_p}(\mathbf{x}_0)}, \quad \mathbb{H}_{\rho,\rho}(\mathbf{x},\mathbf{x}_0) = \frac{\partial^2 J(\zeta_{I_p},\zeta_\rho)}{\partial\zeta_{\rho}(\mathbf{x})\partial\zeta_{\rho}(\mathbf{x}_0)}, \tag{6.22}$$

and the non-diagonal blocks are

$$\mathbb{H}_{I_p,\rho}(\mathbf{x},\mathbf{x}_0)=\frac{\partial^2 J(\zeta_{I_p},\zeta_\rho)}{\partial\zeta_{I_p}(\mathbf{x})\partial\zeta_\rho(\mathbf{x}_0)},\quad \mathbb{H}_{\rho,I_p}(\mathbf{x},\mathbf{x}_0)=\frac{\partial^2 J(\zeta_{I_p},\zeta_\rho)}{\partial\zeta_\rho(\mathbf{x})\partial\zeta_{I_p}(\mathbf{x}_0)}. \tag{6.23}$$

Following the Schwarz's theorem, the Hessian is a symmetric matrix. The approaches described in this article cannot provide the structure of the exact Hessian in equations 6.22 and 6.23 as it consists of a two-step approach (see Algorithms 6 and 7: Conjugate Gradient followed by weighted least-squares projection). To represent the Hessian at each grid point, we first generate the response to a point scatterer either in ζ_{I_p} or ζ_ρ (application of $\mathcal{L}$). Then, we migrate the resulting data to generate a subsurface offset-domain CIG (application of $\mathcal{L}^T$), and finally apply the WLS method to estimate the corresponding point spread function in terms of ζ_{I_p} and ζ_ρ. The preconditioner is this scheme can be applied to the subsurface offset-domain CIG (application of $\mathcal{L}^\dagger\mathcal{L}^{\dagger^T}$) before the WLS step. Consequently, these approaches do not follow the Schwarz's theorem, and thus cannot provide exactly a symmetric Hessian matrix, but rather approximately the Hessian.

6.10 Appendix II: Alternative approach

Hou and Symes (2015) expressed the pseudoinverse Born formula as a modification of RTM operator using two weighting operators in model and data spaces as

$$\mathcal{L}^\dagger_{Hou}=W_{model}\mathcal{L}^T W_{data}, \tag{6.24}$$

where

$$W_{model}=32\frac{\beta_0}{\rho_0^3}\partial_z,\quad W_{data}=I_t(I_t\partial_{s_z})(I_t\partial_{r_z}), \tag{6.25}$$

and I_t denotes the causal indefinite time integration. Note that the weights in equations 6.14 and 6.25 are exactly the same. We refer to Chauris and Cocher (2018) for a more detailed comparison.

To accelerate the convergence rate of the extended LSRTM in a constant-density acoustic medium, Hou and Symes (2016a) reformulated the CG algorithm by replacing Euclidean norms with new weighted norms in data ($\mathscr{D}$) and model ($\mathscr{M}$) spaces using W_{model} and W_{data} as

$$\langle\delta m_1,\delta m_2\rangle_{\mathscr{M}} = \langle\delta m_1, W_{model}\delta m_2\rangle, \tag{6.26}$$

$$\langle\delta d_1,\delta d_2\rangle_{\mathscr{D}} = \langle\delta d_1, W_{data}\delta d_2\rangle. \tag{6.27}$$

The algorithm of weighted CG (WCG) for variable-density LSRTM is summarized in Algorithm 8. In terms of implementation, as the application of these weight operators are relatively cheap (steps 8, 9, and 12 in Algorithm 8), the computational cost of the WCG algorithm is negligibly higher than the standard CG (Algorithm 6).

Algorithm 8: LSRTM by WCG algorithm

Input: d^{obs}
Initialization
  $\delta m_0 \leftarrow 0$
  $r_0 \leftarrow \mathcal{L}^{\dagger}_{Hou}(\mathcal{L}\delta m_0 - d^{obs}) = -\mathcal{L}^{\dagger}_{Hou} d^{obs}$
  $p_0 \leftarrow -r_0$
  $i \leftarrow 0$
WCG: while *stopping criteria not satisfied* **do**
  $q_i \leftarrow \mathcal{L}^{\dagger}_{Hou}\mathcal{L}p_i$
  $\kappa_i \leftarrow \langle r_i, r_i\rangle_{\mathscr{M}} / \langle \mathcal{L}p_i, \mathcal{L}p_i\rangle_{\mathscr{D}}$
  $\delta m_{i+1} \leftarrow \delta m_i + \kappa_i p_i$
  $r_{i+1} \leftarrow r_i + \kappa_i q_i$
  $\eta_{i+1} \leftarrow \langle r_{i+1}, r_{i+1}\rangle_{\mathscr{M}} / \langle r_i, r_i\rangle_{\mathscr{M}}$
  $p_{i+1} \leftarrow -r_{i+1} + \eta_{i+1} p_i$
  $i \leftarrow i + 1$
Physical decomposition: for *each* $\mathbf{x}$ *position* **do**
  $\psi \leftarrow \mathcal{R}(\delta m)$
  for *each depth position* **do**
    Construct $\mathbf{G}$, $\mathbf{R}$
    $\delta \mathbf{u} \leftarrow \mathbf{G}^{-1}\mathbf{R}$

Chapter 7

A case study from the Eastern Nankai Trough

Contents

7.1 Abstract **236**
7.2 Introduction **236**
7.3 Geological background, survey description and seismic data **238**
7.3.1 Geological background 238
7.3.2 Survey description and Seismic data 239
7.4 Processing **239**
7.4.1 Velocity-stack inversion 241
7.4.2 Automatic velocity picking 243
7.4.3 Normal-moveout-free stacking 244
7.4.4 Interval-velocity building and time to depth conversion . . . 245
7.4.5 Data filtering and wavelet estimation 246
7.5 Review of multi-parameter pseudoinverse Born inversion **247**
7.5.1 Application of pseudoinverse Born operator 248
7.5.2 Transforming the subsurface offset to the diffraction angle . 249
7.5.3 Solving a weighted least-squares objective function 251
7.6 Results and validation **253**
7.7 Discussion **264**
7.8 Conclusion **265**
7.9 Acknowledgments **265**

7.10 Appendix: Gradient derivation for WLS **265**

Résumé

Ce chapitre comprend l'article intitulé « The importance of including density in multi-parameter asymptotic linearized direct waveform inversion: a case study from the Eastern Nankai Trough », publié dans la revue *Geophysical Journal International* (Farshad et al., 2021).
La migration élastique itérative (« elastic least-squares reverse time migration » en anglais) est la technique linéarisée la plus avancée pour obtenir des images quantitatives du sous-sol. Le principal inconvénient d'une telle approche itérative est le coût significatif associé au calcul des opérateurs de modélisation et d'adjoint au cours des itérations. Dans le contexte du domaine étendu, une alternative intéressante est l'approche directe au sens asymptotique qui permet d'avoir des résultats quantitatifs en une seule itération. Cette approche a d'abord été proposée pour les milieux acoustiques à densité constante, puis récemment étendue au cas de densité variable. Cette dernière comprend deux étapes supplémentaires : construction d'une réponse qui dépend de l'angle d'imagerie, puis résolution d'un petit système linéaire pour l'estimation des paramètres physiques. Pour examiner l'importance de prendre en compte les variations de densité, nous comparons les techniques d'inversion sans et avec densité variable sur un jeu de données réelles 3D (Eastern Nankai Trough, Japon). Les résultats d'inversion confirment l'efficacité de l'approche pour estimer de manière quantitative les paramètres en une seule itération. Le cas de la densité variable donne des images du sous-sol qui (1) ont une meilleure résolution, et (2) reconstruisent mieux les données, même si ici les données ne contiennent pas assez de grands offsets pour un découplage entre perturbations de vitesse et de densité.

The importance of including density in multi-parameter asymptotic linearized direct waveform inversion: a case study from the Eastern Nankai Trough
Milad Farshad, Hervé Chauris and Mark Noble
Geophysical Journal International, 2021, *doi: 10.1093/gji/ggab405*

7.1 Abstract

Iterative least-squares reverse time migration (LSRTM) is the state-of-the-art linearized waveform inversion method to obtain quantitative subsurface parameters. The main drawback of such an iterative imaging scheme is the significant computational expense of many modeling/adjoint cycles through iterations. In the context of the extended domain, an interesting alternative to LSRTM is the asymptotic linearized direct waveform inversion, providing quantitative results with only a single iteration. This approach was first proposed for constant-density acoustics and recently extended to the variable-density case. The former is based on the application of the asymptotic inverse Born operator, whereas the latter has two more extra steps: building an angle-dependent response of the asymptotic inverse Born operator and then solving a weighted least-squares approach for simultaneous inversion of two acoustic parameters. To examine the importance of accounting for density variations, we compare constant- and variable-density linearized direct waveform inversion techniques applied to a marine real dataset from the Eastern Nankai Trough, offshore Japan. The inversion results confirm the efficiency of the asymptotic linearized direct waveform inversion in estimating quantitative parameters within a single iteration. The variable-density direct inversion yields subsurface images that (1) exhibits a superior resolution and (2) better reconstructs the field data than does the constant-density approach, even if the dataset does not contain large enough surface offset to fully decompose velocity and density perturbations.

7.2 Introduction

Full-waveform inversion (FWI) is an advanced nonlinear imaging method that provides an accurate subsurface model by nonlinear minimization of the difference between the observed and the synthetic data (Tarantola, 1984; Virieux and Operto, 2009). Under the weak-scattering assumption, the subsurface model can be decomposed into two components: the background model controlling the kinematics of the wave propagation and the reflectivity model generating the refracted and reflected waves. Reverse time migration (RTM) is the state-of-the-art linear imaging method to estimate a reflectivity model in a given background model (Baysal et al., 1983). A standard RTM operator is formulated as the adjoint of the linearized forward modeling operator, i.e., adjoint

Born operator. Such operator can only produce qualitative structural images (kinematics) without any quantitative amplitude information (dynamics). In practice, the quality of the migrated image may be significantly deteriorated by various causes such as noise, narrow frequency bandwidth, limited acquisition aperture and irregular source-receiver distributions (Baysal et al., 1983; Nemeth et al., 1999).

Least-squares reverse time migration (LSRTM), also known as linearized waveform inversion, overcomes the shortcomings of RTM and retrieves high-resolution quantitative reflectivity images (Nemeth et al., 1999; Dai et al., 2011; Chen and Sacchi, 2017). LSRTM is posed as an optimization problem, with the objective to derive a reflectivity model that minimizes the misfit between the observed and reconstructed reflection data. A successful application of LSRTM requires a kinematically accurate background velocity model. Generally, iterative methods are deployed to carry out the optimization problem. As each iteration requires the computation of two (forward and adjoint) numerical simulations of the seismic wave equation per shot, the computational expense of LSRTM can be a significant bottleneck. Thus in recent years, either reducing the computational cost per iteration or accelerating the convergence rate of LSRTM has been a matter of development (Romero et al., 2000; Shin et al., 2001; Herrmann and Li, 2012; Lin and Lianjie, 2015; Zand et al., 2020).

An interesting alternative to iterative migration scheme is the direct inversion, which can be achieved by replacing the adjoint migration operator, with its pseudoinverse version. The direct inverse was initially proposed for ray-based migration (Beylkin, 1985; Bleistein, 1987), and recently developed for RTM (ten Kroode, 2012; Hou and Symes, 2015, 2017; Chauris and Cocher, 2017), enabling one to get quantitative results in a single iteration. It accounts for the finite frequency nature of the source wavelet, geometric spreading and uneven illumination. The pseudoinverse formulas are derived assuming infinite acquisition geometries and are only valid in the absence of turning waves (ten Kroode, 2012). In practice, the former is not a limiting factor as they produce accurate results within the migration aperture, whereas including turning waves would require adding vertical subsurface offsets (Biondi and Shan, 2002). Originally, the theory of asymptotic linearized direct waveform inversion, also known as pseudoinverse Born inversion, was established for pure acoustic media. Through synthetic numerical experiments, these pseudoinverse Born operators have illustrated very promising results in terms of quantitative imaging within a single iteration, and also accelerating the convergence rate of LSRTM (Hou and Symes, 2016a), Migration velocity analysis (MVA) (Chauris and Cocher, 2017), and FWI (Hou and Symes, 2016b) by acting as a proper preconditioner. Once again, all were limited to constant-density acoustics.

The importance taking into account density in the inversion process has been frequently highlighted in the concepts of LSRTM (Yang et al., 2016a; Sun et al., 2018; Chen and Sacchi, 2019) and FWI (Przebindowska et al., 2012; Guitton, 2014; Qin and Lambaré, 2016). Although density does not affect the kinematics of wave propagation, its impact on the amplitudes of the seismic reflection data is inevitable. Lately, Dafni

and Symes (2018) and Farshad and Chauris (2020a) extended the applicability of pseudoinverse Born operators to variable-density acoustics. Farshad and Chauris (2020a) illustrated that the density variations in the observed data result in defocused energy in the common-image gather (CIG), which can be misinterpreted as an incorrect background velocity model. The same observation was made by Zhou et al. (2020a), who proposed the guidelines for the user parameters to be used in constant-density acoustic MVA. The effectiveness of the pseudoinverse Born operator either in single- or multi-parameter inversion has not been examined when dealing with real datasets. The main contribution of this article is twofold. Firstly, we validate the performance of pseudoinverse Born inversion with a case study devoted to the imaging of the Eastern Nankai Trough, offshore Japan. In this region many seismic surveys have been acquired in order to characterize methane hydrates where density can play an important role (Fujii et al., 2015). Thus secondly, we investigate the importance of considering density variation in quantitative parameter estimation. In this study, impedance is the main investigated parameter. We discuss if density is only an additional parameter for a better data fit or if it may have a physical interpretation.

The rest of this paper is organized as follows. We start with a brief introduction of the study area, survey description, and the presentation of the seismic data. After that, we describe our integrated processing workflow followed by a review of multi-parameter pseudoinverse Born inversion. Finally, we present the results corresponding to constant and variable-density acoustics.

7.3 Geological background, survey description and seismic data

7.3.1 Geological background

In the Nankai Trough, extending about 700 km from the Suruga Trough to the northern end of the KyushuPalau ridge, bottom simulating reflectors (BSRs) are widely observed on the seismic reflection data (Ashi et al., 2002; Nouzé et al., 2004). A BSR, occurring in the upper few hundred meters of marine sediments, is a strong reflection with reverse polarity with respect to the seafloor reflection and is generally interpreted to mark the base of the gas hydrate stability zone (Hyndman and Spence, 1992). The phase reversal of BSR is due to the distinct impedance contrast between gas hydrates above and gas-saturated sediments below. The impedance contrast across the BSR is mainly caused by a sharp decrease in P- and S-wave velocity while density remains almost constant (Ojha and Sain, 2007; Fujii et al., 2015). A continuous BSR is usually used to get a first-order indication of the possible existence of gas hydrate, attaining great interest during the last decades. Gas hydrates not only contain one of the major energy resources, but also may have a significant impact on the climate and environment (Ojha and Sain, 2009).

7.3.2 Survey description and Seismic data

This seismic data was acquired during the SFJ cruise on N/O Nadir in the framework of the French-Japanese program in the year 2000 (Figure 7.1) (Nouzé et al., 2004). The dataset consists of 390 shot points with a shot spacing of 25 m, the minimum and maximum offsets are 225 and 1462.5 m, respectively with a receiver spacing of 12.5 m. The source was composed of an array of mini G-I guns, with a total volume of 151 inch^3 at a 2 m depth, the streamer was towed at a 15 m immersion. The recording length is 4 s, with a sampling interval of 1 ms. Figure 7.2 shows a shot gather recorded on a towed-streamer geometry. The main events are highlighted in this image.

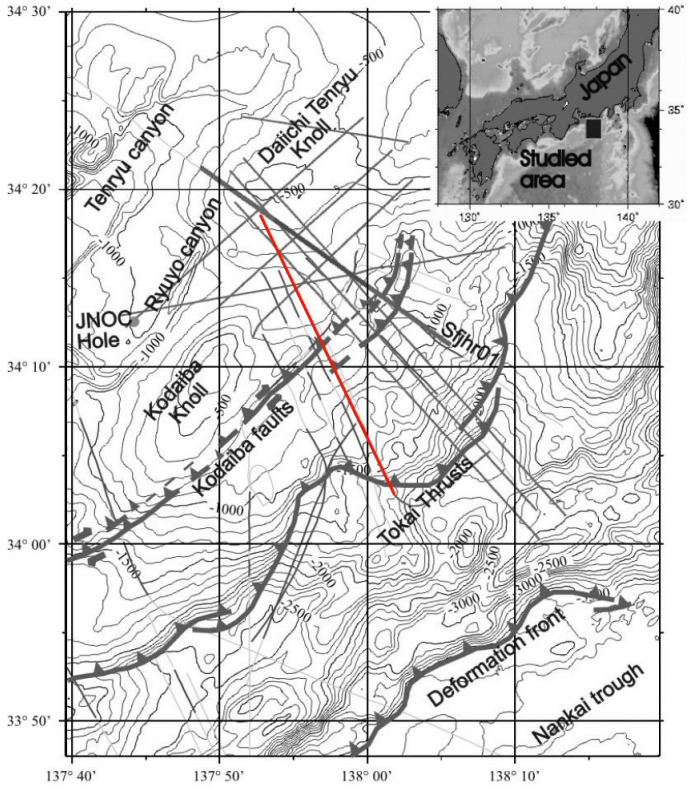

Figure 7.1: Location map of the studied area. The red line indicates the investigated seismic line in this study (Nouzé et al., 2004).

7.4 Processing

As a first step, we mute the direct arrivals and events after 1.6 s to remove first order multiples. Thus, the observed data in our processing and imaging steps mainly contains

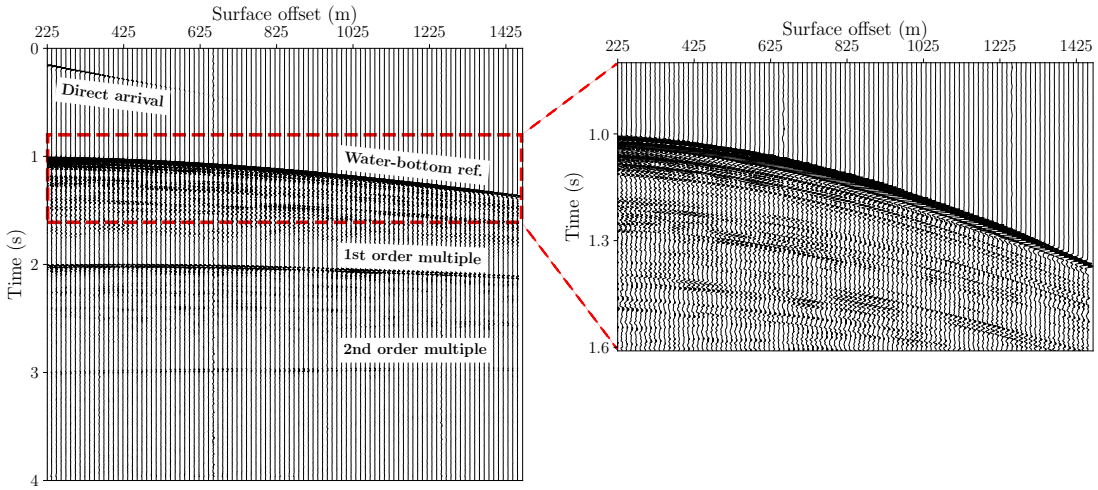

Figure 7.2: An example of observed common-shot gather.

reflected waves. We then sort the data as common-mid-point gathers (CMPs) to conduct the velocity analysis. In the next steps, we only consider the part of the model with full-fold CMPs (Figure 7.3). In a typical scenario, one obtains a velocity model using a classical semblance-based velocity analysis followed by a tedious manual velocity picking step. The manual picking of peak values in the velocity panels is not only a time-consuming process but also challenging in some cases as the peak values may extensively vary in both horizontal and vertical directions (Neidell and Taner, 1971; Claerbout and Green, 2010). In the following processing steps, we first focus on building a reliable velocity-stack section using sparse inversion and automatic velocity picking algorithms. We then proceed by constructing a stacked section, data filtering, and wavelet estimation.

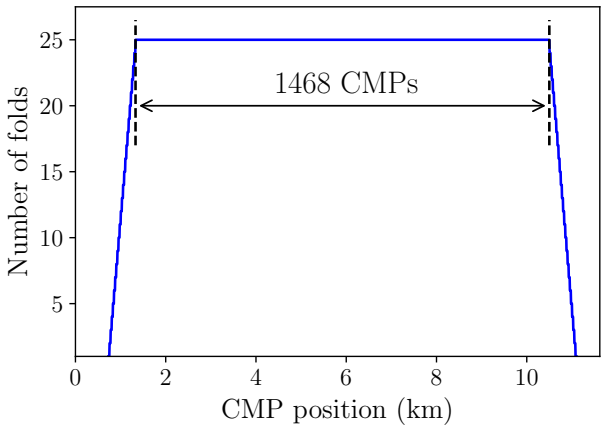

Figure 7.3: The CMP fold coverage.

7.4.1 Velocity-stack inversion

For simple geological structures, seismic reflection data in the CMP domain mainly show hyperbolic shapes (Thorson and Claerbout, 1985). Hyperbolic Radon transform (HRT) is a common and powerful tool in seismic data processing that permits to model seismic reflections via a superposition of hyperbolas as (Trad et al., 2003)

$$d(t,x) = \iint \mathrm{d}\tau \mathrm{d}p \; m(\tau,p)\delta(\tau - \sqrt{t^2 - p^2x^2}). \tag{7.1}$$

For a given mid-point position in CMP domain, $d(t,x)$ denotes the seismic reflection data as a function of time t and offset x, $m(\tau,p)$ is the Radon plane as a function of intercept time τ and slowness p, and δ denotes the Dirac delta function. Accordingly, adjoint HRT reads

$$m(\tau,p) = \iint \mathrm{d}t\mathrm{d}x \; d(t,x)\delta(t - \sqrt{\tau^2 + p^2x^2}). \tag{7.2}$$

The conventional velocity analysis method utilizes equation (7.2) within a time window to measure the coherency of trace signal. Semblance, the most common coherence measure, has a poor resolution velocity spectrum which affects the velocity-picking procedure (Neidell and Taner, 1971). To increase the resolution of the velocity spectrum, HRT is often formulated as an inverse problem with possible sparsity constraints (Thorson and Claerbout, 1985). Here, we perform a high-resolution (sparse) solution to the following optimization problem

$$\arg\min_{\mathbf{m}} \Big(\frac{1}{2}\big\|\mathbf{d} - \mathbf{R}\mathbf{m}\big\|_2^2 + \lambda\big\|\mathbf{m}\big\|_1\Big), \tag{7.3}$$

where $\mathbf{d}$ and $\mathbf{m}$ denote the data and model vectors, $\mathbf{R}$ is the forward HRT operator, $\|\|_p$ denotes the ℓ_p norm, and λ is the fixed trade-off parameter balancing between data fit and sparseness of the model. In practice, we estimate λ by plotting the L-curve between the first and second terms in equation 7.3 (Farshad and Chauris, 2021c). We perform a fast iterative shrinkage-thresholding algorithm (FISTA) to carry out the optimization problem (Beck and Teboulle, 2009). Although some efficient methods for rapid application of HRT have been proposed (Nikitin et al., 2017; Gholami and Farshad, 2019a), we explicitly construct the matrix $\mathbf{R}$ as the size of each CMP gather is relatively small here. Figure 7.4 shows three different CMP gathers and their corresponding sparse Radon planes generated by 800 iterations of FISTA with a total runtime of 10 sec (per CMP). In the following, we describe how to perform automatic velocity picking through the high-resolution Radon planes.

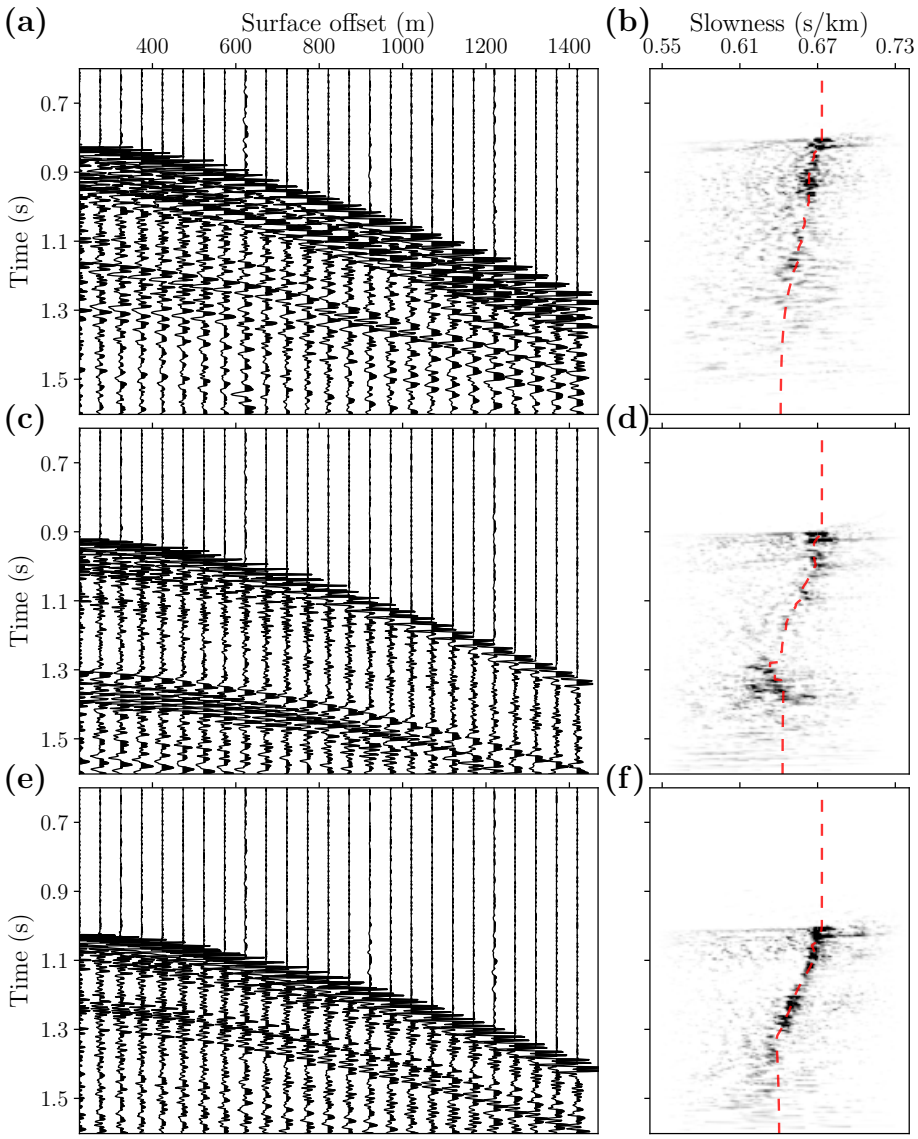

Figure 7.4: Velocity-stack inversion: (a,c,e) CMP gathers and (b,d,f) their corresponding absolute values of sparse Radon transform plane. The red dashed curves in (b,d,f) show the slowness model calculated by equation (7.4).

7.4.2 Automatic velocity picking

After the velocity-stack inversion, we conduct an automatic velocity picking process by using the sparse Radon coefficients. Claerbout and Green (2010) proposed to estimated the root-mean squared (RMS) or stacking velocity model $V_{rms}(\tau)$ at each CMP location

by a weighted average of all slowness ranges as

$$V_{rms}^{-1}(\tau) = \frac{\sum_p p\left|m(\tau,p)\right|}{\sum_p \left|m(\tau,p)\right|} \tag{7.4}$$

where the absolute value of the Radon coefficients $m(\tau, p)$ serve as the weighting function. To stabilize the velocity picking and increase the accuracy, we consider the Radon coefficients with a threshold above 10% of the maximum absolute value at each CMP position (Gholami, 2017). The red dashed lines in Figures 7.4b, d and f show the picked slowness via equation (7.4). Note that the slowness value in the shallow part of the model is set to the water slowness (0.673 s/km = 1.485 km/s, the area above the blue dash-dotted dashed line in Figure 7.5a). The area in white in Figure 7.5a represents the points that are larger than the penalty (10% of the maximum absolute value of $m(\tau, p)$ for a given CMP position) and thus are used in equation (7.4). A detailed look at the later times of the CMP gathers plus their Radon planes (Figure 7.4) reveal that no clear reflections are available at the deeper part of the model (approximately after 1.4 s), and the recorded data at that times are mainly artifacts which are indeed due to the free-gas in the area. Thus we also mute the picked velocity values in this region (the area below the red dashed line in Figure 7.5a). The velocity values in black areas in Figure 7.5a are estimated by interpolating/extrapolating from the neighboring available samples (Figure 7.5b).

7.4.3 Normal-moveout-free stacking

The traditional workflow to construct a stacked section involves Normal MoveOut (NMO) correction. This process results in undesirable frequency distortion called NMO stretching, particularly for shallow events and at large offsets, (Sheriff and Geldart, 1995). To construct a NMO-free stacked section, we project the Radon coefficients $m(\tau, p)$ to the time axis τ (Figure 7.6) (Gholami, 2017). In other words, as the Radon coefficients $m(\tau, p)$ are obtained by stacking the reflectors along their true slowness value and then putting the obtained energy in the corresponding τ_0 (equation 7.2), the final NMO-free stacked trace at each CMP location can be obtained by simply summing $m(\tau, p)$ along the p-axis. The sparseness of the Radon coefficients at each CMP position has a significant impact on the quality of the constructed stacked section. As $m(\tau, p)$ is very sparse, the quality of the resultant stacked section is remarkably high. The two reflectors showing the characteristic properties of BSRs are highlighted by arrows in Figure 7.6. By analyzing the temperature and pressure conditions at the seafloor, Foucher et al. (2002) showed that the upper BSR most likely corresponds to the present-day base of methane hydrate stability zone whereas the lower BSR, namely double BSR (DBSR), is a residual BSR. In other words, they suggested that the base of the hydrate stability zone migrated due to tectonic uplift or sea bottom warming (Nouzé et al., 2004). The attenuated reflection amplitude and energy distortion in the left part (before 2.6 km) and also the lower

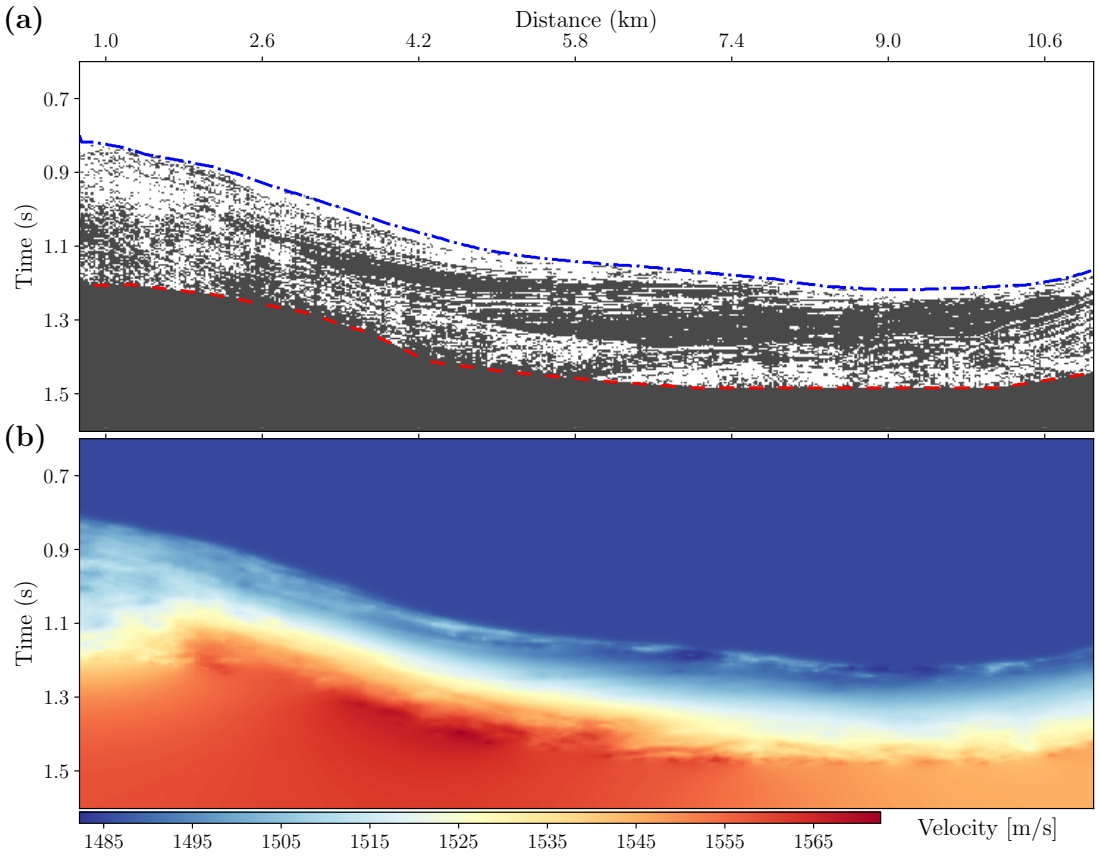

Figure 7.5: a) The area shown in white are those points whose velocity value was calculated. b) RMS velocity model.

part (below 1.4 s) of the stacked section could be caused by the free gas existing in the area.

7.4.4 Interval-velocity building and time to depth conversion

To get an initial velocity model (background model) for our imaging step, we construct the interval velocity v_{int} using the Dix formula (Dix, 1952). To stabilize the inversion, we smooth out the oscillations in the RMS velocity model before the Dix conversion. By assuming that the lateral velocity variations are negligible and reflectors are mainly horizontal (Figure 7.5), we perform a time to depth conversion using vertical rays (Figure 7.7). The interval velocity model reveals that the layers above the BSR exhibit a higher velocity (1750-1810 m/s) compared to the marine sediments (1650-1700 m/s), which can be associated with gas hydrate presence. On the other hand, the layers below the BSR show a low velocity (1550-1600 m/s) and can be related to free gas presence

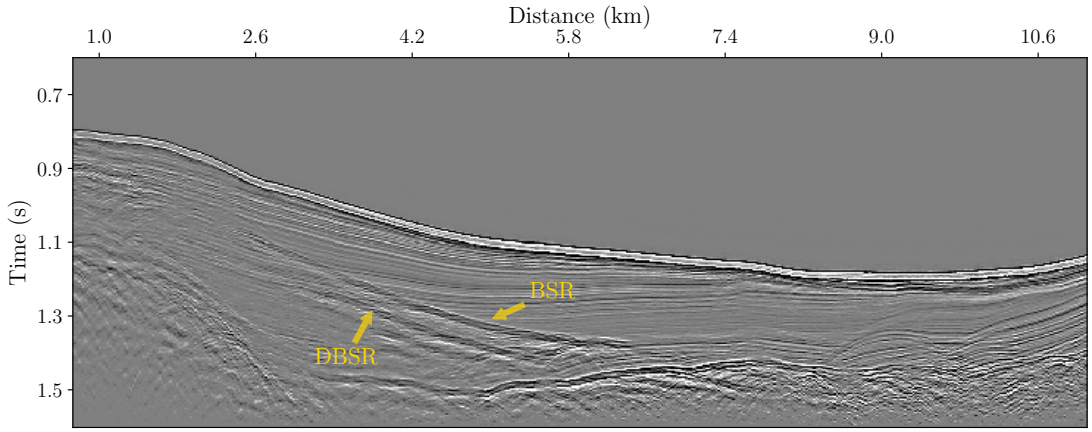

Figure 7.6: The normal-moveout-free stacked section.

(Hyndman and Spence, 1992). We will further investigate the quality of the inverted velocity model through the imaging step.

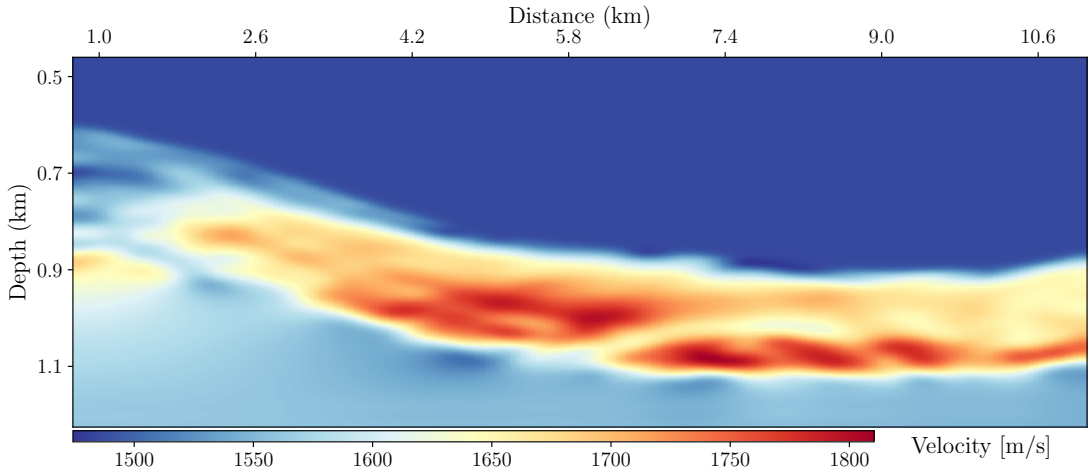

Figure 7.7: The interval velocity section obtained with Dix formula.

7.4.5 Data filtering and wavelet estimation

Generally, the computational cost of the wave-equation-based imaging methods is expensive at high frequencies, and it increases cubically (in a 2D model) with increasing frequency. The frequency spectrum of the recorded data ranges from 2 to 220 Hz (Figure 7.8). The 50 Hz notches in the recorded frequency spectrum are related to the ghost wavefields (Amundsen et al., 2013). To reduce the computational cost and also mitigate

the receivers notches, we apply a zero-phase sine-squared filter to pass a band of frequencies from 2 Hz to 45 Hz (Figure 7.9). To estimate the seismic wavelet, we extract two traces from water-bottom reflection at near and far offsets of three different shots positioned at 3.4, 5.6 and 8.1 km (Figure 7.10). The extracted traces from different shot positions and different offsets are highly correlated, and illustrate that the Ricker wavelet can represent a decent estimation of the seismic wavelet (Figure 7.10). The alternative would be using linear least-squares waveform inversion of the direct waves in the frequency domain to estimate the seismic wavelet (Pratt, 1999).

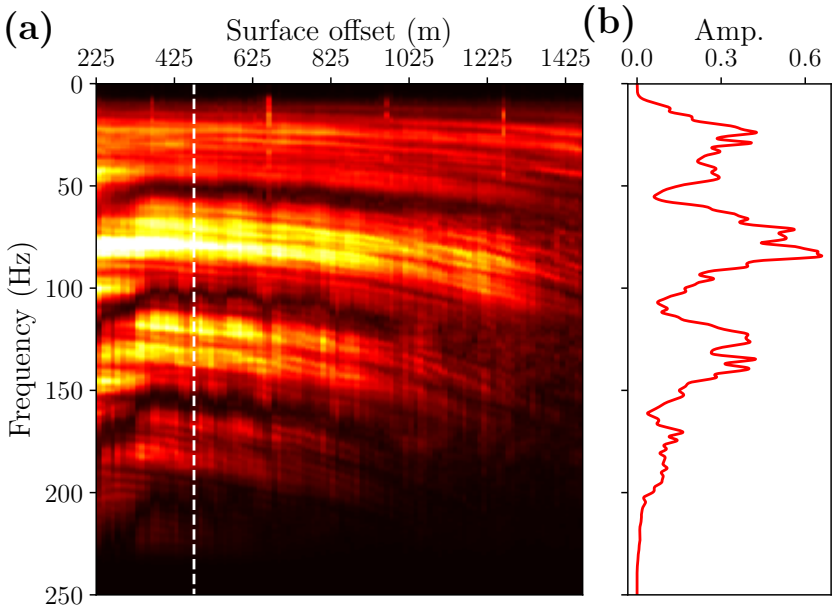

Figure 7.8: The a) frequency spectrum of the original shot illustrated in Figure 7.2, and b) one extracted trace at the location illustrated by dashed line in (a).

7.5 Review of multi-parameter pseudoinverse Born inversion

Under the Born approximation, the model parameters (e.g. a) on each spatial coordinate $\mathbf{x} = (x, z)$, are split into a smooth part (background model, e.g. a_0) and a singular part (perturbation model, e.g. δa). The pseudoinverse formula for RTM is expressed in the concept of the extended model domain, such that the dimensions for model and data spaces are equivalent (ten Kroode, 2012; Hou and Symes, 2015; Chauris and Cocher, 2017). The extended domain is defined by introducing a redundancy parameter, involving nonphysical subsurface offset or time-lag extension (Sava and Fomel, 2006;

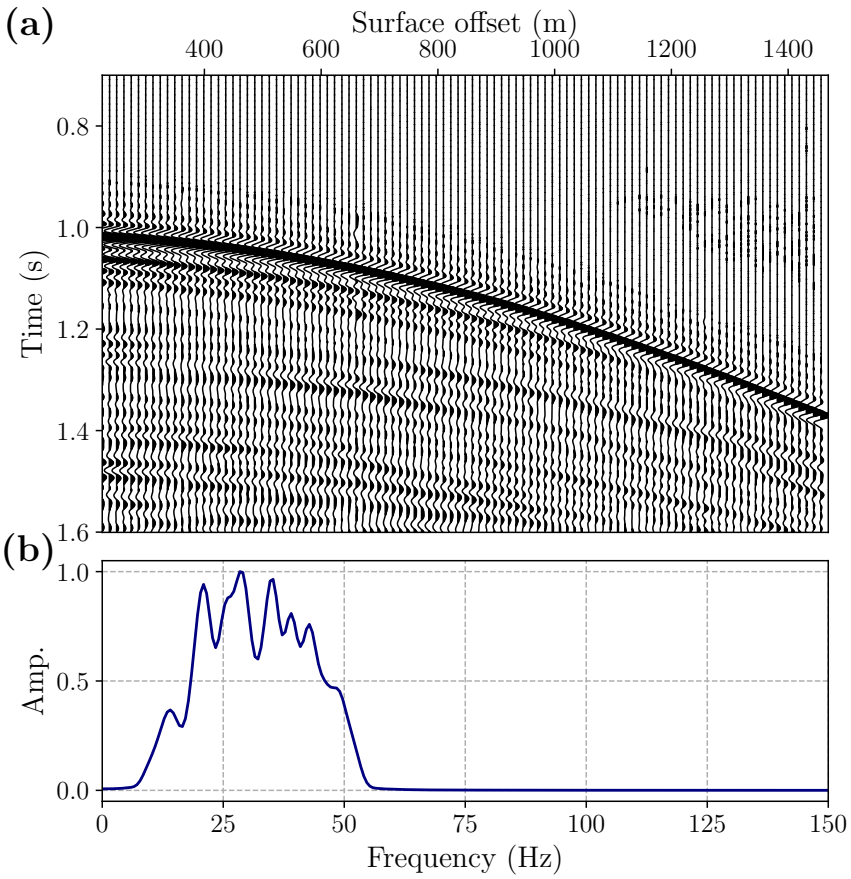

Figure 7.9: The a) shot gather and b) its spectrum after low-pass filtering with cut frequency of 45 Hz.

Symes, 2008b). Here we consider the psuedoinverse Born operator extended by horizontal subsurface offset $\mathbf{h} = (h, 0)$, defined as sunken sources and receivers in Claerbout's survey-sinking concept (Claerbout, 1985; Chauris and Cocher, 2017; Farshad and Chauris, 2020a). The variable-density pseudoinverse Born inversion can be summarized in three main steps as follows (Farshad and Chauris, 2020a).

7.5.1 Application of pseudoinverse Born operator

Following the work of Farshad and Chauris (2020a), the pseudoinverse Born operator $\mathcal{L}^\dagger$ in smoothly varying background models β_0 and ρ_0 can be written as

$$\begin{aligned} \xi &= \mathcal{L}^\dagger(d^{obs})(\mathbf{x}, \mathbf{h}) \\ &\simeq 32\frac{\beta_0}{\rho_0^3}\partial_z \int \mathrm{d}\mathbf{x_s}\,\mathrm{d}\mathbf{x_r}\,\mathrm{d}\omega\, \frac{\Omega^\dagger(\omega)}{(i\omega)}\partial_{s_z} G_0^*(\mathbf{x_s}, \mathbf{x}-\mathbf{h}, \omega) d^{obs}(\mathbf{x_s}, \mathbf{x_r}, \omega)\partial_{r_z} G_0^*(\mathbf{x}+\mathbf{h}, \mathbf{x_r}, \omega), \end{aligned} \tag{7.5}$$

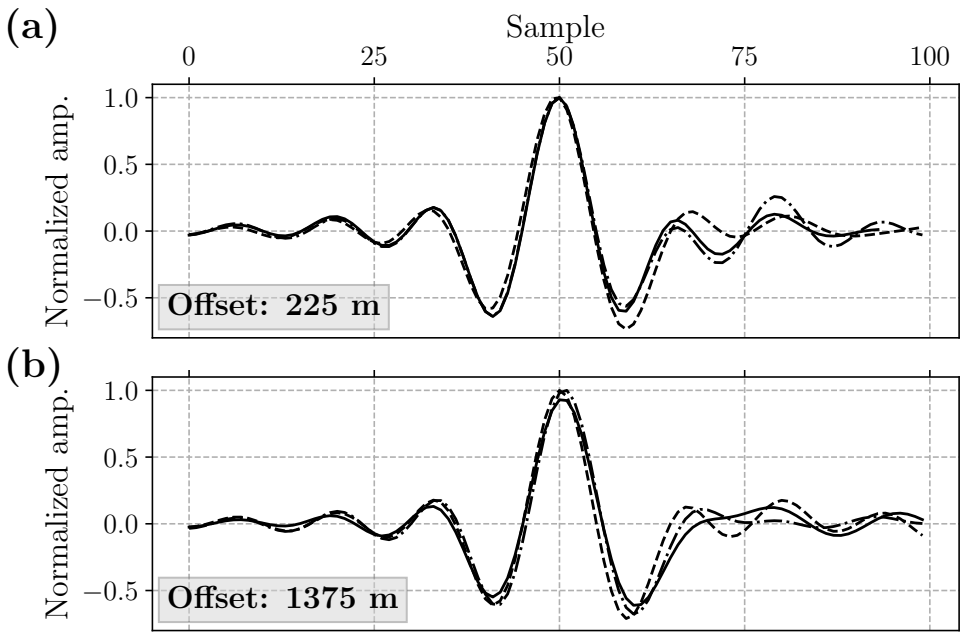

Figure 7.10: The extracted traces around the water bottom reflection from the filtered shot (Figure 7.9) for offsets a) 225 m and b) 1175 m. The solid, dashed and dot-dashed lines correspond to the shots positioned at 3.4 km, 5.6 km and 8.1 km, respectively.

where * is the complex conjugate, ω denotes the angular frequency, and $\Omega^\dagger$ is the inverse of the seismic source. The extended reflectivity $\xi(\mathbf{x}, \mathbf{h})$ has the same unit as the inverse of the bulk modulus $\beta = 1/\rho V^2$, where ρ denotes the density model. The Green's functions $G_0(\mathbf{x_s}, \mathbf{x} - \mathbf{h}, \omega)$ and $G_0(\mathbf{x} + \mathbf{h}, \mathbf{x_r}, \omega)$ are evaluated in the background models β_0 and ρ_0, from the source position $\mathbf{x_s}$ to imaging point $\mathbf{x} - \mathbf{h}$ and from imaging point $\mathbf{x} + \mathbf{h}$ to the receiver position $\mathbf{x_r}$, respectively. The derivation of equation (7.5) is performed under the high-frequency approximation such that $\mathcal{L}\mathcal{L}^\dagger d^{obs} \approx d^{obs}$. It accounts for geometrical spreading, source signature inversion, and uneven illumination. Although the derivation is performed under the asymptotic approximation of the Green's functions, the final formula does not contain any ray quantity and it tends to be valid even in presence of triplicated wavefields (ten Kroode, 2012; Hou and Symes, 2015; Chauris and Cocher, 2017).

7.5.2 Transforming the subsurface offset to the diffraction angle

The extended image ξ can be decomposed into various parameters such that each parameter is determined in the physical domain (non-extended) (Dafni and Symes, 2018; Farshad and Chauris, 2020a). Here, since the pseudoinverse Born inversion is a linear

imaging method resolving only the short-wavelength part of the model (reflectivity), we parameterize the subsurface by P-wave impedance I_p and density ρ (Tarantola, 1986). The decomposing relationship reads as

$$\frac{\xi}{\beta_0}(\mathbf{x},\mathbf{h}) \cong 2\frac{\delta I_p}{I_{p0}}(\mathbf{x})\delta(\mathbf{h}) - 2\sin^2(\theta)\frac{\delta\rho}{\rho_0}(\mathbf{x})\delta(\mathbf{h}), \tag{7.6}$$

where θ denotes the diffraction angle, and the weights 2 and $-2\sin^2(\theta)$ in front of δI_p and $\delta\rho$ are the diffraction patterns (Figure 7.11) obtained under the ray+Born approximation (Forgues and Lambaré, 1997). Using this parameterization, the wavefield scattered by δI_p has an isotropic pattern (solid blue line in Figure 7.11), whereas $\delta\rho$ generates scattering at wide-to-intermediate θ only (dashed red line in Figure 7.11). Here, as the maximum offset of the investigated dataset is limited, impedance is the main sought parameter. In the case of constant-density acoustics ($\delta\rho(\mathbf{x}) = 0$), the impedance perturbation $\delta I_p(\mathbf{x})$ can be reconstructed from the extended parameter $\xi(\mathbf{x},\mathbf{h})$ by simply summing over all subsurface offset values as

$$\delta I_p(\mathbf{x}) = \frac{I_{p0}}{2\beta_0}\int \mathrm{d}\mathbf{h}\,\xi(\mathbf{x},\mathbf{h}). \tag{7.7}$$

In the case of variable-density acoustics ($\delta\rho(\mathbf{x}) \neq 0$), the subsurface offset $\mathbf{h}$ on the left side of equation (7.6) should be transformed to the diffraction angle by utilizing a 2D Radon transform (slant-stack integral) defined as (Sava and Fomel, 2003)

$$\psi(x,z,\theta) = \frac{1}{\beta_0}\int \mathrm{d}h\,\xi(x, z+h\tan\theta, h), \tag{7.8}$$

where $\psi(\mathbf{x},\theta)$ denotes the corresponding angle-domain CIG. Note that equation (7.8) is similar to equation (7.1) in spirit: both are slant-stack integrals but with different integration paths. Note that the algorithm does not require the application of the adjoint of the Radon transform (equation 7.8).

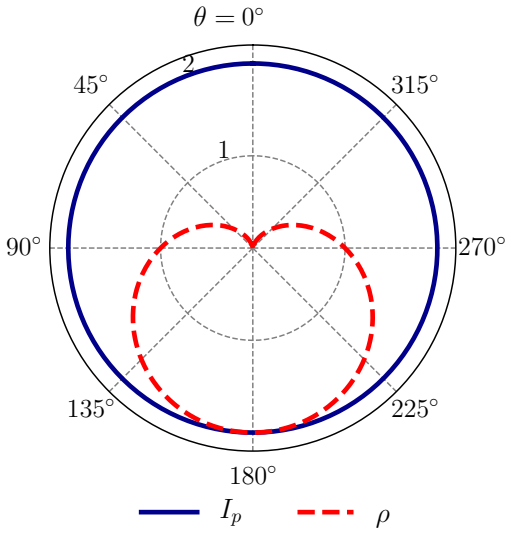

Figure 7.11: The analytical diffraction pattern for a variable-density acoustic medium parameterized by impedance and density.

7.5.3 Solving a weighted least-squares objective function

To obtain optimal $\delta I_p(\mathbf{x})$ and $\delta\rho(\mathbf{x})$ parameters, Farshad and Chauris (2020a) proposed to define a weighted linear least-squares problem as

$$\min_{\delta I_p, \delta\rho} \left(\frac{1}{2} \left\| 2\left(\frac{\delta I_p}{I_{p0}}(\mathbf{x}) - \frac{\delta\rho}{\rho_0}(\mathbf{x}) \sin^2(\theta) \right) - \psi(\mathbf{x}, \theta) \right\|_W^2 \right) \tag{7.9}$$

where the weighted space defined as

$$W(\mathbf{x}, \theta) = \begin{cases} 1, & \text{if } |\theta| \leq \alpha \tan^{-1}\left(\frac{x_{max}}{z}\right) \\ 0, & \text{otherwise} \end{cases} . \tag{7.10}$$

Assuming a homogeneous Earth model, the weighted space $W(\mathbf{x}, \theta)$ is designed based on acquisition geometry (using depth z and maximum surface offset x_{max}) to only consider specular energy in the inversion process (Farshad and Chauris, 2020a). In this definition, parameter α ensures that the angle-domain does not contain strong artifacts related to the finite sampling (Farshad and Chauris, 2021c). The optimal δI_p and $\delta\rho$ can be obtained by solving the WLS problem (equation 7.9) with respect to the model parameters (see Appendix 7.10 for more details). The WLS problem also has the flexibility to include the regularization terms (Farshad and Chauris, 2021c).

The main differences between the constant- and variable-density pseudoinverse Born inversion are in the application of the Radon transform (equation 7.8) and solving the

WLS method, as they are not required in the constant-density case. By assuming $\delta\rho = 0$, the physical reflectivity is estimated by simply summing $\xi(\mathbf{x}, \mathbf{h})$ along the subsurface offset $\mathbf{h}$ using equation (7.7). In the next section, we evaluate the effect of neglecting density variations in pseudoinverse Born inversion by mainly analyzing the data fit and impedance reconstruction.

7.6 Results and validation

In this section, we first discuss how to build impedance and density background models for our imaging scheme. We then assess the performance of constant-density and variable-density pseudoinverse Born inversions in terms of reconstructing common-shot/common-offset sections and estimating subsurface parameters. Regarding data reconstruction, the two key factors that dictates the quality of the inversion are the amplitude and the phase of the reconstructed data. To evaluate the amplitude and also impact of the phase, we calculate the root-mean-squared (RMS) error defined as

$$\text{RMS error} = \sqrt{\sum_{i=1}^{K} \frac{(\kappa_i - \widetilde{\kappa}_i)^2}{K}}, \tag{7.11}$$

where κ and $\widetilde{\kappa}$ are the observed and resimulated data, and to get more insight into the quality of the phase we calculate the normalized correlation coefficient (NCC) defined as

$$\text{NCC}(\%) = \frac{\sum_{\mathbf{x}} \kappa(\mathbf{x})\widetilde{\kappa}(\mathbf{x})}{\sqrt{\sum_{\mathbf{x}} \kappa^2(\mathbf{x})}\sqrt{\sum_{\mathbf{x}} \widetilde{\kappa}^2(\mathbf{x})}} \times 100. \tag{7.12}$$

The higher the NCC score, the higher the correlation is (the highest score is $100\% = 1$).

To construct the density model below the water bottom (sediments), we use an empirical relationship (Ojha and Sain, 2014)

$$\rho = 1.3813 V_p^{0.5083}. \tag{7.13}$$

The background models are built by applying a 50 m long 2D Gaussian filer in both directions (Figure 7.12). We start by performing the pseudoinverse Born operator and the Radon transform. This step is the same for both constant- and variable-density acoustic Born inversion. The estimated CIGs and their angle-domain responses for three different positions are shown in Figure 7.13. The focused energy around zero subsurface offset (Figure 7.13a) and nearly flat events in the angle-domain CIGs (Figure 7.13b) prove that our initial velocity model is kinematically accurate. Note that the events in the offset-domain CIGs are mainly concentrated in one side of the subsurface offset (Figure 7.13a) as the acquisition geometry is end-on spread.

Before comparing the inverted parameters, we first perform the Born modeling operator on the inverted physical parameters to resimulate the data (Figure 7.14). The dashed yellow boxes in Figure 7.14 denote some examples of discrepancy between the AVO responses. The RMS and NCC values written on each panel show the superiority of the variable-density inversion. The extracted traces at near and far offsets also illustrate that the variable-density inversion accurately retrieves the phase and amplitudes

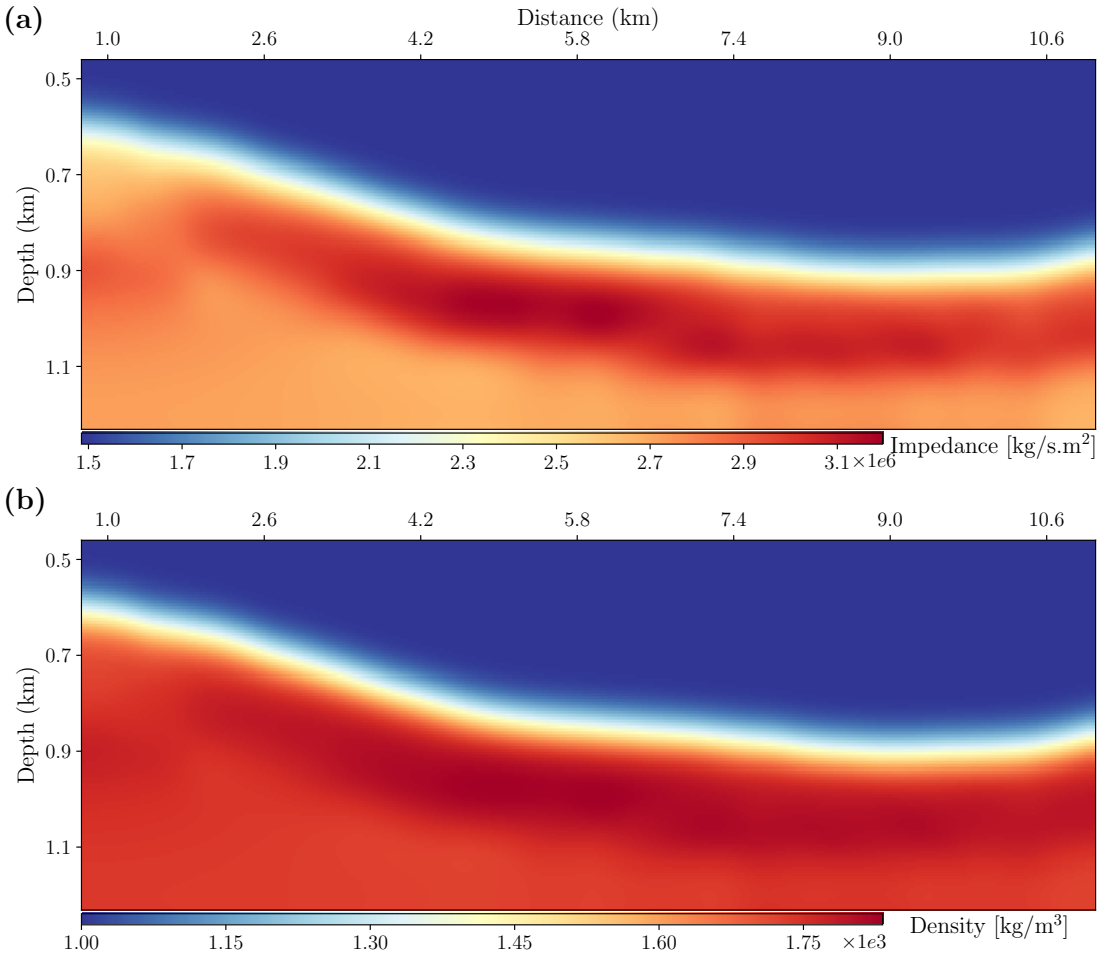

Figure 7.12: The a) impedance and b) density background models.

of the observed data (Figure 7.15). To further analyze the AVO responses, we construct two common-offset sections for 287.5 and 756.25 m (Figure 7.16). The NCC and RMS values written on each panel show that the variable-density case reconstructed the common-offset gathers with smaller error and higher correlation. The residual wavefield for each common-offset gather clearly shows that the variable-density approach allows for a better amplitude fit of the reflections (Figure 7.17). The remaining amplitude/phase mismatch in the variable-density case could be due to inaccuracies in the migration background models or energy beyond the acoustic assumption (Figures 7.17c and d). To get more insight into the areas of the discrepancies, we calculate localized NCC values for each reconstructed common-offset section within a window of 11 grid points in both directions (Figure 7.18). In the case of constant-density case, most of the uncorrelated energies are observed around the sea bottom reflector (Figure 7.18a)

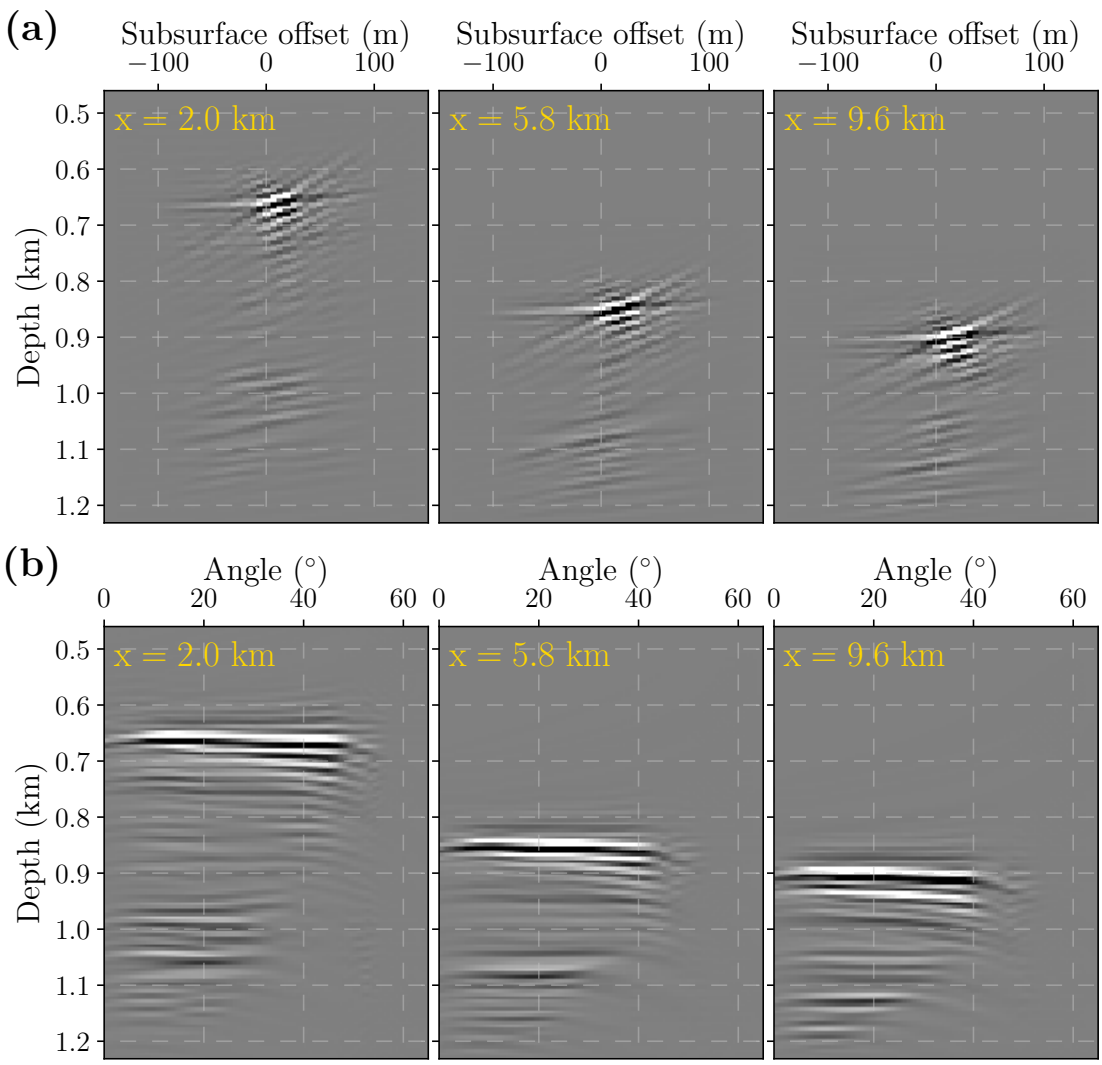

Figure 7.13: The extracted a) subsurface-offset domain and b) angle-domain CIG. The position of each CIG is written on each panel.

and in the sediments at the left part of the model in which free gas might be trapped (Figure 7.18b). This is indeed due to ignoring the density variations that cause AVO effects in these areas. On the other hand, the variable-density inversion manages to construct highly correlated common-offset sections in the most parts of the subsurface (Figures 7.18c and d).

We now compare the inverted physical parameters (Figure 7.19). In both constant- and variable-density cases, the sharp drop in P-impedance at the BSR is in good agreement with the expected physics due to the hydrated sediments above and gas-saturated

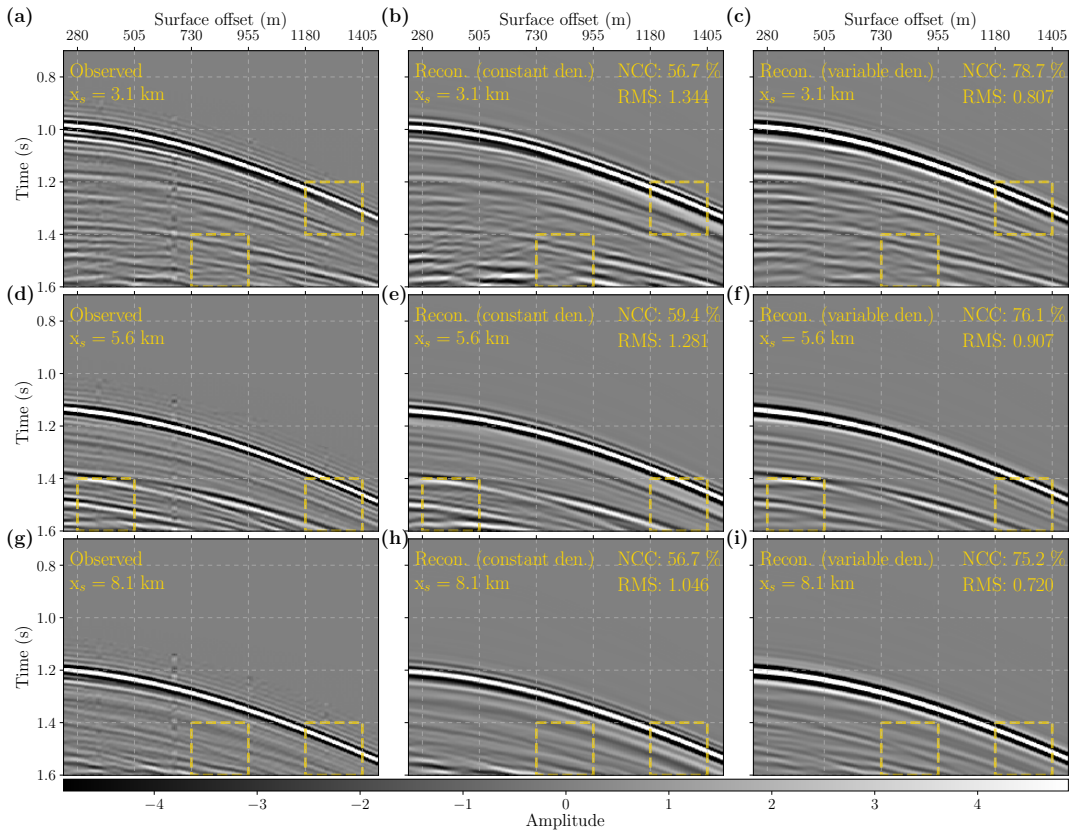

Figure 7.14: The (a,d,g) observed and (b,c,e,f,h,i) reconstructed data assuming (b,e,h) constant-density and (c,f,i) variable-density acoustics for the shots positioned at (a,b,c) 3.1 km, (d,e,f) 5.6 km and (g,h,i) 8.1 km. The NCC and RMS values are written on each panel. The dashed yellow boxes denote the areas with main differences in AVO responses at each shot position.

sediments below the BSR (Ojha and Sain, 2007; Fujii et al., 2015). However, the P-impedance image obtained with constant-density pseudoinverse Born inversion contains some high-frequency noise and artifacts particularly in the sediments containing free gas (below 1.1 km). To further compare, we plot the magnified views of the yellow boxes in Figure 7.19 (Figure 7.20). We also detect the main structures in this portion of the subsurface. It is visible that the constant-density impedance model contains artifacts (dotted green lines in Figure 7.20c), and distinguishing the reflectors is rather challenging due to the phase rotation (Figure 7.20a). The latter leads to uncertainty in the selection and interpretation of the fault dip direction (Figure 7.20c). On the other hand, the variable-density impedance model provides more continuously imaged reflectors providing a structurally precise correlation with the stacked section (Figures 7.20b and d).

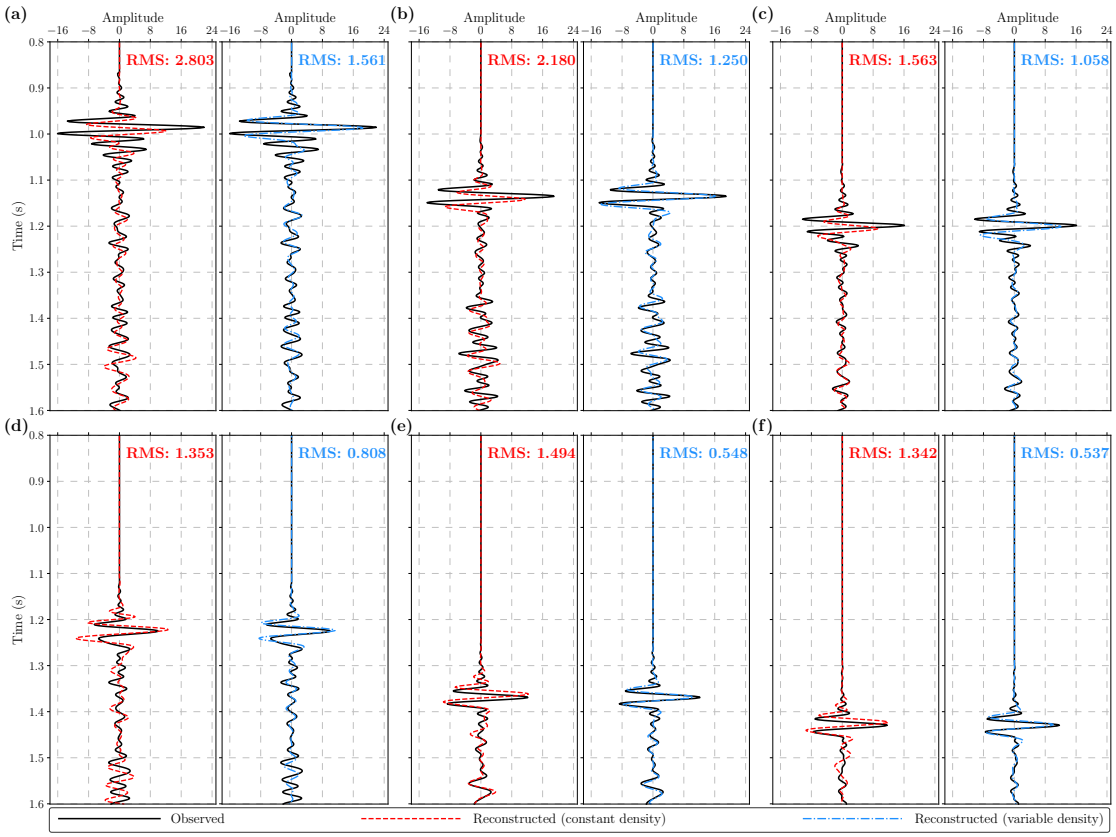

Figure 7.15: The extracted traces from the observed and reconstructed shots (Figure 7.14) at (a,b,c) 225 m and (d,e,f) 1180 m for shots positioned at (a,d) 3.1 km, (b,e) 5.6 km and (c,f) 8.1 km. The RMS values between the observed and reconstructed traces are written on each panel.

We now analyze the estimated density model (Figure 7.19c). In contrast with the inverted impedance model, the density along the BSR is either constant (dashed yellow arrows in Figure 7.19c) or slightly decreased (solid red arrow in Figure 7.19c). The former is consistent with the nature of BSR: the density of gas-hydrates-bearing sediments and free-gas saturated sediments across the BSR remains almost the same as that of the background density (Ojha and Sain, 2007). The small oscillation and drop in density across the BSR can be an imprint of the impedance model due to parameter cross-talks. Further LS iteration might require to fully remove the parameter cross-talks (Virieux and Operto, 2009). We conclude here that although the estimated density model might contain some uncertainty at some level, neglecting it in the inversion process may lead to artifacts in the estimated impedance images, and thus in data reconstruction.

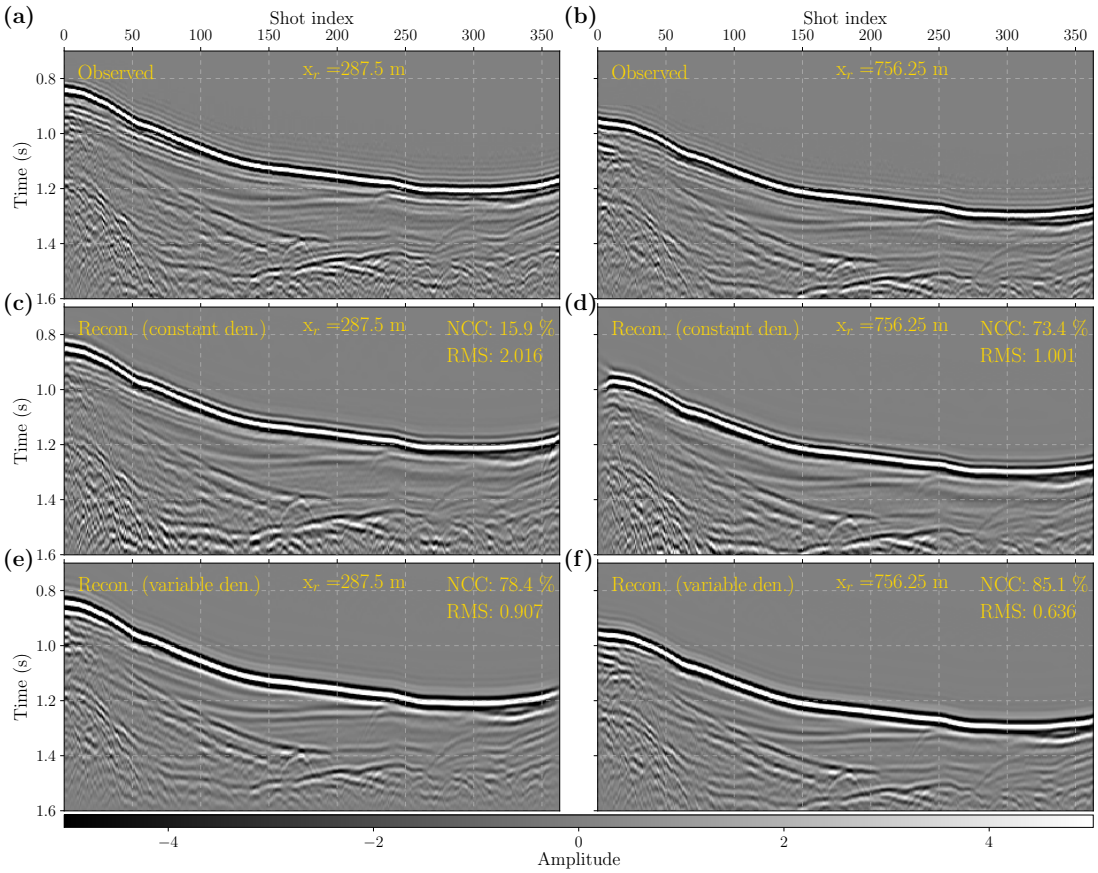

Figure 7.16: The common-offset sections comparison: (a,b) the observed common-offset sections versus the reconstructed ones assuming (c,d) constant-density and (e,f) variable-density acoustics for offset (a,c,e) 287.5 m and (b,d,f) 756.25 m. The NCC and RMS values are written on each panel.

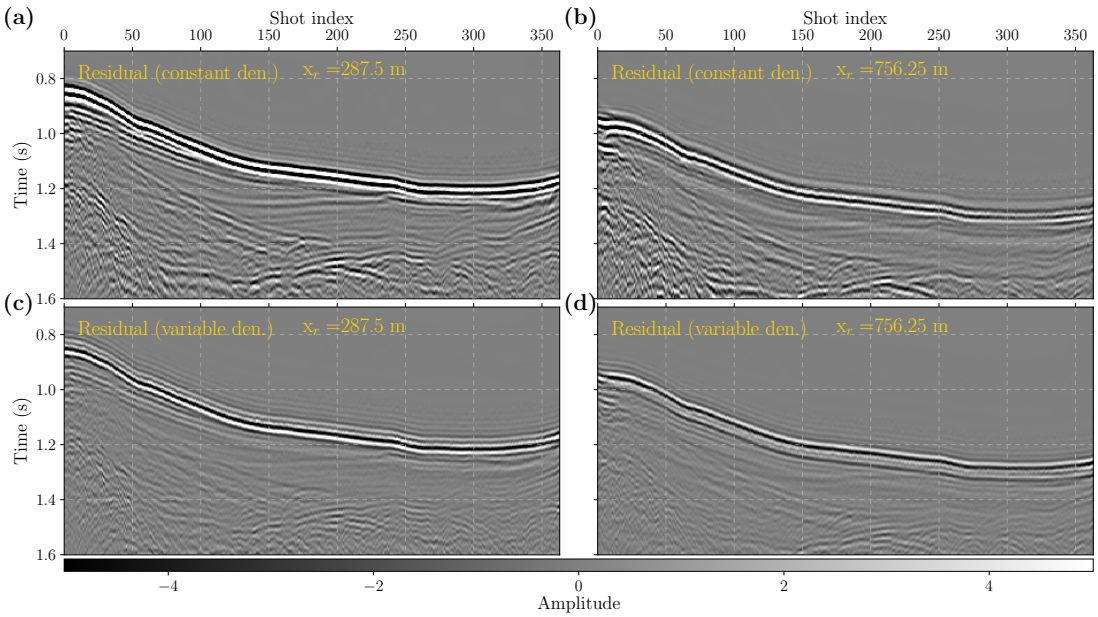

Figure 7.17: The residual common-offset sections comparison: the difference between the observed and reconstructed common-offset sections (Figure 7.16) assuming (a,b) constant-density and (c,d) variable-density acoustics for offset (a,c) 287.5 m and (c,d) 756.25 m.

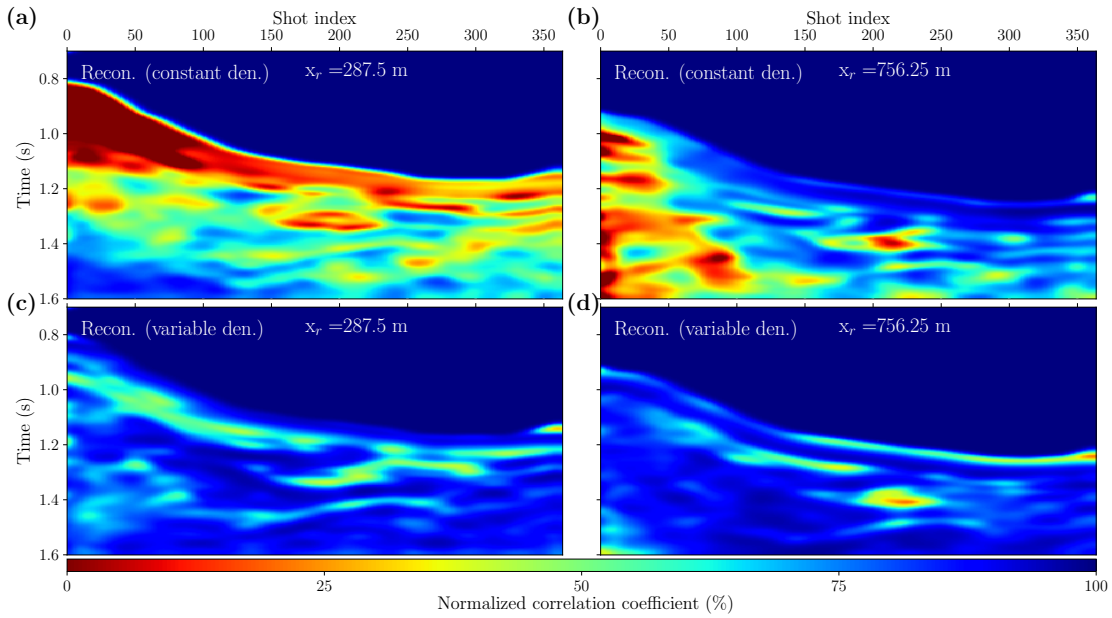

Figure 7.18: The localized NCC values corresponding to Figure 7.16: NCC panels for reconstructed common-offset sections assuming (a,b) constant-density and (c,d) variable-density acoustics for offset (a,c) 287.5 m and (c,d) 756.25 m.

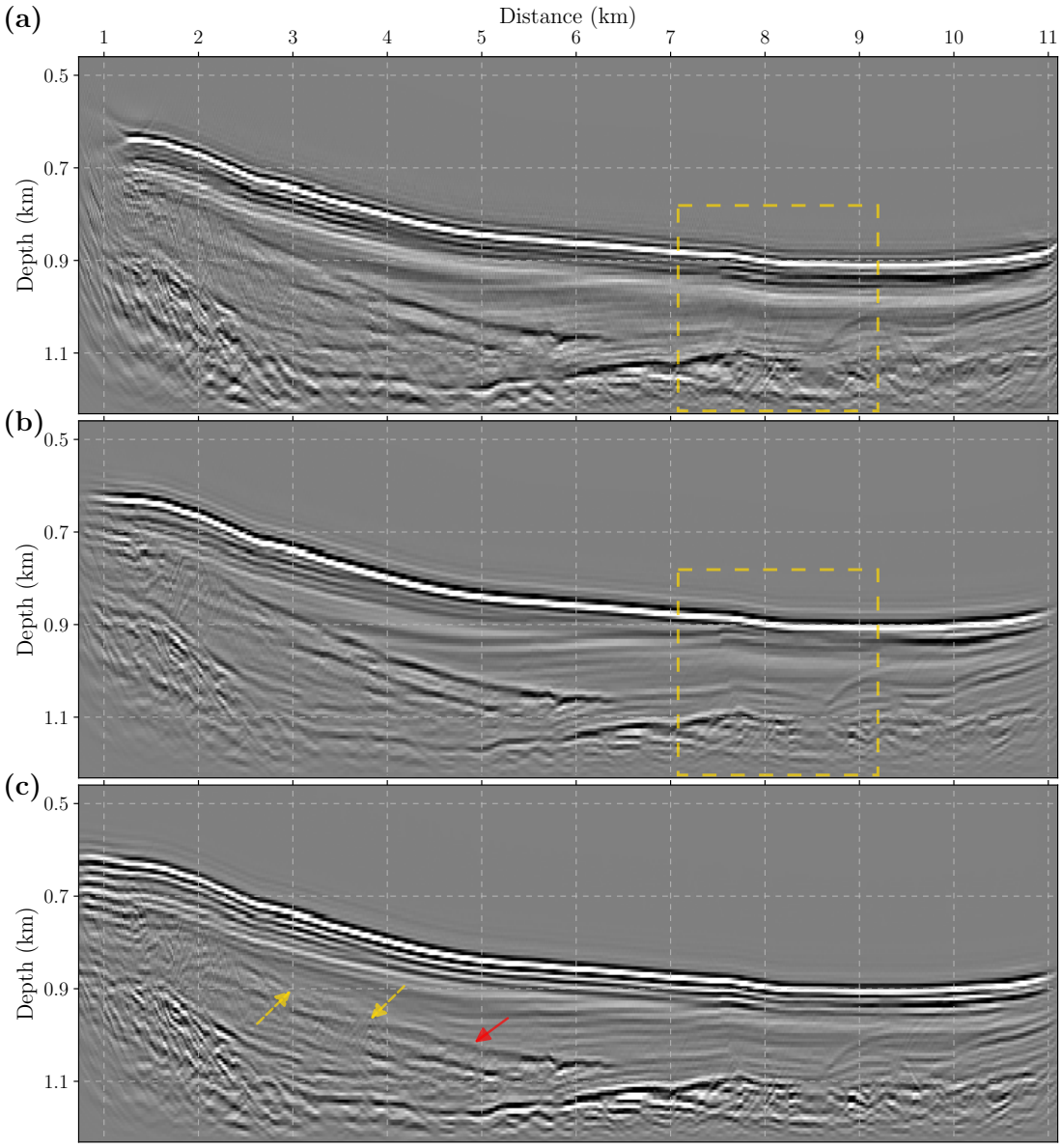

Figure 7.19: The inverted a) δI_p assuming constant-density acoustics, and b) δI_p and c) $\delta\rho$ assuming variable-density acoustics. The dashed yellow boxes point to the areas for the magnified views.

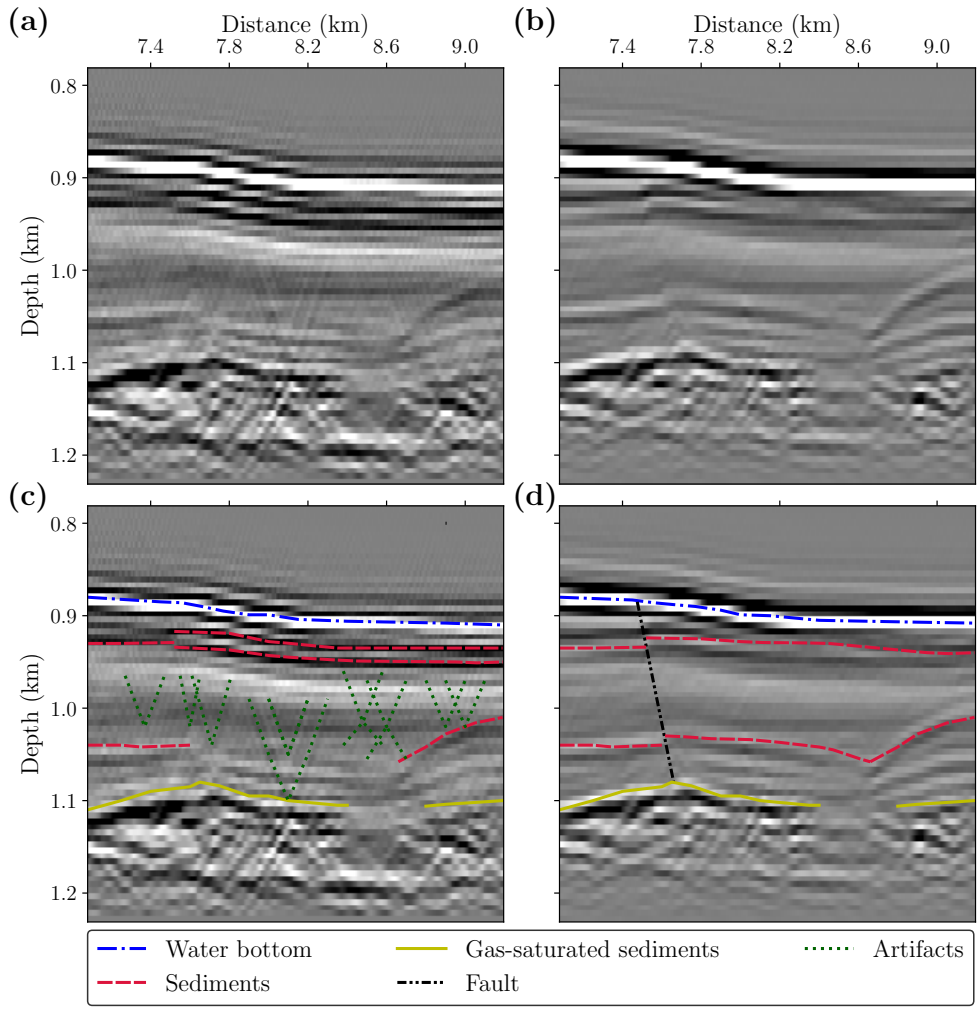

Figure 7.20: Magnified views of the inverted impedance models corresponding to a) constant-density and b) variable-density acoustics, and the main structures detected overlaid on the impedance section for c) constant-density and d) variable-density acoustics.

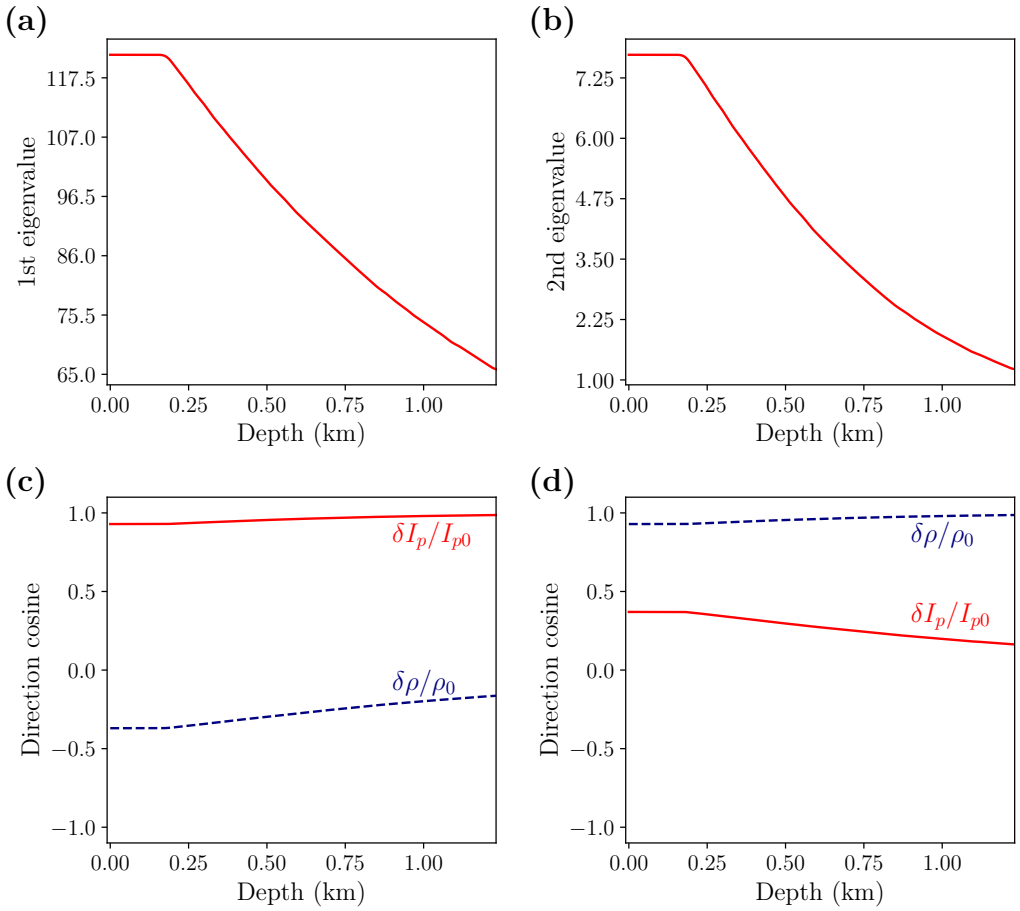

Figure 7.21: The eigenvalue/eigenvector decomposition of the matrix $\mathbf{M}$ versus depth: a) 1st eigenvalue, b) 2nd eigenvalue, c) 1st eigenvector and d) 2nd eigenvector.

7.7 Discussion

We have investigated the importance of including density variation in the asymptotic linearized direct waveform inversion using a real data. The pseudoinverse Born operator derived in such that the application of $\mathcal{L}\mathcal{L}^{\dagger}$ on any data, accurately resimulates the data even in presence of incorrect background models (Hou and Symes, 2015; Chauris and Cocher, 2017). The extended domain $\mathbf{h}$ is a key component here, as it captures the effect of incorrect background models. The specular energy is expected to be focused around $\mathbf{h} = 0$ for correct background models, though, this is not the case in the presence of density variations. The extended domain indeed also accounts for the inconsistency between the physics of the observed data and the one assumed in the modeling/inversion operator Farshad and Chauris (2020a). In other words, performing the Born modeling operator $\mathcal{L}$ in correct background models on the inverted physical model $\xi(\mathbf{x}) = \int \mathrm{d}\mathbf{h}\, \xi(\mathbf{x}, \mathbf{h})$ can reconstruct the data only if the density variations in the data are negligible. Otherwise, it is required to decompose the extended $\xi(\mathbf{x}, \mathbf{h})$ in terms of two physical acoustic parameters and then perform the modeling operator.

In the concept of multi-parameter imaging, a problem that comes to the fore is the parameter cross-talk, resulting from the trade-off effects between different parameters (for example imprint of density on the impedance section or vice versa). In practice, the similarity of the diffraction patterns of different parameters for certain illumination angles is responsible for parameter cross-talk. This makes multi-parameter imaging a highly ill-posed problem especially with a limited surface acquisition (Virieux and Operto, 2009). To mitigate parameter cross-talk artifact one can either implement pseudoinverse Born operator as a preconditioner in LSRTM and let LS iteration remove the artifacts (Chen and Sacchi, 2019), or implement regularization techniques within the WLS method to reduce the ill-posedness of the inverse problem (Farshad and Chauris, 2021c).

The inverted impedance δI_p and density $\delta\rho$ parameters using asymptotic direct waveform inversion are true-amplitude up to a scaling factor. This scaling is related to the accuracy of the wavelet estimation/inversion. In presence of an accurate seismic wavelet, the inverted parameters can be related to rock physics (Farshad and Chauris, 2020a, 2021c). Unfortunately, there is no well data to confirm the inverted values. The radiation pattern (Figure 7.11) shows that the density parameter only radiates energy at very large angles. The investigated dataset here does not contain large enough surface offset, meaning that the estimated density model might contain some uncertainty. To get more insight into the resolution of the inverted parameters, we carry out an eigenvalue and eigenvector decomposition of the 2×2 matrix $\mathbf{M}$ (equation 7.16, Figure 7.21). The components of eigenvectors, known as direction cosine, associated with the eigenvalue define the relative significance of the model parameter (Plessix and Cao, 2011). The first (Figure 7.21c) and second (Figure 7.21d) eigenvectors are mainly related to the impedance and density parameters, respectively. The first eigenvalue (Figure 7.21a) is

much larger than the second eigenvalue (Figure 7.21b). Therefore, the most reliable information in the inversion process is related to the impedance model (higher eigenvalues with larger weight in the eigenvector). As illustrated, a higher resolution is expected in the shallow part of the models (lower eigenvalues in the shallow part, Figures 7.21a and b). Note that the depth of interest in this dataset approximately starts from 0.6 km. In this case, one can refer to the inverted density as a pseudo-density model, used as a proxy to increase the degree of freedom in the inversion process and soak up some of the artifacts in the impedance model (Borisov et al., 2014; Operto and Miniussi, 2018).

7.8 Conclusion

In this study, we have demonstrated an application of variable-density asymptotic linearized direct waveform inversion to a case study from the Eastern Nankai Trough. The multi-parameter pseudoinverse Born operator can estimate high-resolution physical images with better continuity which accurately resimulates the field data. The analysis illustrated that neglecting the density parameter may produce artifacts in the image domain leading to false AVO response in the data domain. Future studies will explore the applicability of multi-parameter asymptotic linearized direct waveform inversion as a suitable preconditioner for other iterative imaging techniques such as migration velocity analysis and full-waveform inversion.

7.9 Acknowledgments

We would like to thank scientific editor Michal Malinowski, reviewer Serge Sambolian, and anonymous reviewer for their valuable suggestions and comments that improved the manuscript.

7.10 Appendix: Gradient derivation for WLS

We explain here how to compute the gradient of the WLS problem (equation 7.9) with respect to the model parameters as

$$\frac{\partial J^{WLS}}{\partial \delta I_p / I_{p0}} = 2 \int \mathrm{d}\theta \, W(\mathbf{x}, \theta) \Big(\frac{\delta I_p}{I_{p0}}(\mathbf{x}) - \frac{\delta \rho}{\rho_0}(\mathbf{x}) \sin^2(\theta) - \psi(\mathbf{x}, \theta) \Big), \tag{7.14}$$

and

$$\frac{\partial J^{WLS}}{\partial \delta \rho / \rho_0} = -2 \int \mathrm{d}\theta \, W(\mathbf{x}, \theta) \sin^2(\theta) \Big(\frac{\delta I_p}{I_{p0}}(\mathbf{x}) - \frac{\delta \rho}{\rho_0}(\mathbf{x}) \sin^2(\theta) - \psi(\mathbf{x}, \theta) \Big). \tag{7.15}$$

Subsequently, the optimal physical parameters are estimated where the gradients in equation (7.14) and (7.15) are equal to zero. Abstractly, the WLS method in matrix and functional form reads as $\mathbf{M}\delta\mathbf{u} = \mathbf{\Phi}$ where

$$\mathbf{M} = \begin{bmatrix} 2\int \mathrm{d}\theta\, W(\mathbf{x},\theta) & -2\int \mathrm{d}\theta\, W(\mathbf{x},\theta)\sin^2(\theta) \\ -2\int \mathrm{d}\theta\, W(\mathbf{x},\theta)\sin^2(\theta) & 2\int \mathrm{d}\theta\, W(\mathbf{x},\theta)\sin^4(\theta) \end{bmatrix}, \tag{7.16}$$

$$\delta\mathbf{u} = \begin{bmatrix} \frac{\delta I_p}{I_{p0}}(\mathbf{x}) \\ \frac{\delta \rho}{\rho_0}(\mathbf{x}) \end{bmatrix}, \tag{7.17}$$

and

$$\mathbf{\Phi} = \begin{bmatrix} 2\int \mathrm{d}\theta\, W(\mathbf{x},\theta)\psi(\mathbf{x},\theta) \\ -2\int \mathrm{d}\theta\, W(\mathbf{x},\theta)\sin^2(\theta)\psi(\mathbf{x},\theta) \end{bmatrix}. \tag{7.18}$$

By definition, the WLS method leads to a separable bivariate curve-fitting problem, meaning that the optimization can be applied for each spatial coordinate separately. Note that the matrix $\mathbf{M}$ in this article is required to be calculated only once, as the acquisition configuration (maximum offset) is constant.

Chapter 8

Conclusions and Perspectives

Contents

8.1 Conclusion **269**

8.1.1 Beyond constant density 269

8.1.2 Iterative approaches 270

8.1.3 A case study from the Eastern Nankai Trough 271

8.2 Perspectives **272**

8.2.1 Extension to anisotropic media 272

8.2.2 Extension to multi-component seismic data 274

8.2.3 Coupling multi-parameter direct inverse with migration velocity analysis 275

8.2.4 Coupling multi-parameter direct inverse with reflection waveform inversion 276

8.2.5 Extension to 3D 276

Résumé

Dans cette thèse, j'ai développé un schéma d'inversion direct pour la résolution du problème d'imagerie sismique linéarisé. L'inverse direct fait référence au pseudoinverse. Il permet d'avoir des résultats quantitatifs en une seule itération. La formulation originale pour l'inverse direct a été proposée pour des milieux acoustiques à densité constante. Une description plus réaliste de la subsurface demande une densité variable et la prise en compte des effets élastiques.

Je me suis principalement intéressé à quatre problématiques : (1) pour prendre en compte une physique plus élaborée, j'ai étendu l'inverse direct aux cas de densité variable et élastiques ; (2) pour supprimer les artefacts de migration et améliorer la qualité de l'inversion, j'ai incorporé une régularisation ℓ_1 pour l'inverse direct à densité variable ; (3) pour accélérer la convergence de la migration itérative, j'ai implémenté l'inverse direct en tant que préconditioner ; et finalement (4) pour valider les résultats, j'ai appliqué l'inverse directe à densité variable sur un jeu de données réelles (Eastern Nankai Trough).

Le principal intérêt du schéma d'imagerie est que le processus d'imagerie est séparé en deux parties : imagerie + décomposition en paramètres physiques. Seulement la première partie demande de calculer les champs d'onde. C'est donc la partie la plus coûteuse de loin. La seconde partie est un problème de décomposition, avec la possibilité d'introduire des contraintes. Comme perspectives, je propose de développer l'approche directe pour les milieux anisotropes, de regarder les données multi-composantes, de coupler l'inverse direct avec l'Analyse de Vitesse par Migration ou avec l'approche RWI (« Reflection Waveform Inversion ») et enfin de voir comment l'approche peut être applicable en 3D.

8.1 Conclusion

In this thesis, I have investigated a direct inversion scheme for the resolution of the seismic linear inverse problem. The direct inverse also referred to as pseudoinverse Born inversion, provides quantitative results within a single iteration. The original formulation for the direct inverse was proposed assuming a constant-density acoustic wave-equation approximation. A more realistic description of the subsurface introduces density and elastic effects. These parameters have a significant impact on the reflection amplitudes and should be taken into account.

In this thesis, I have mainly investigated these four issues: (1) to account for better physics, I have extended the direct inversion method to variable-density acoustic and variable-density elastic media; (2) to suppress migration artifacts and improve the quality of the inversion, I have incorporated ℓ_1-norm regularization to the variable-density direct inverse; (3) to accelerate the convergence rate of iterative migration, I have implemented direct inverse as a preconditioner; and finally (4) to validate the results, I have applied the variable-density direct inverse to a marine real dataset from the Eastern Nankai Trough.

The main interest in my developed true-amplitude imaging scheme is that the imaging process is split into two main parts: Imaging + Physical decomposition. Only the first part contains wave-equation terms. It is thus the most computationally expensive part by far. The second part ("Physical decomposition") is a multivariate curve fitting problem in which additional constraints can be introduced.

8.1.1 Beyond constant density

Extension to variable-density acoustic

In Chapter 3, I have shown that the density variations distort the energy in the CIGs, which would be misinterpreted as an incorrect background model. Thus I have extended the applicability of the pseudoinverse Born operator from constant-density acoustic to variable-density acoustic media. This extension consists of three main steps. The first step is the application of the same pseudoinverse Born operator as the one for the constant-density acoustic case. In the second phase, the Radon transform is applied to map the imaging angle. The work of Dafni and Symes (2018) shares the same idea as Zhang et al. (2014b), who propose to invert the acoustic parameters by using two traces (two-trace method) from the angle-domain CIGs estimated either by LSRTM or by application of the pseudoinverse Born operator. I have illustrated that the selection of the two traces is important and proposed a generalization by taking into account all traces in angle-domain CIGs using an efficient weighted least-squares (WLS) approach. This generalization not only yielded more robust results compared to the two-trace method, but it also provided the flexibility to include more constraints through the formulation

as an inverse problem. By investigating the computational complexity of the variable-density pseudoinverse Born inversion, I have highlighted that the main computational burden of the proposed method is due to the modeling operators and it remains of the same order as that of the constant density. The main limitation is the non-uniqueness of the inversion results: by adding another parameter such as density into account, the ill-posedness of the inverse problem is increased because more degrees of freedom are considered. In that case, regularization techniques are expected (section 8.1.2).

Extension to variable-density elastic

In Chapter 4, using the same strategy as described in Chapter 3, I have extended the applicability of the pseudoinverse Born operator from acoustic to elastic media. The elastic and acoustic pseudoinverse Born operators have similar structures. The main differences are elastic versus acoustic Green's functions in the pseudoinverse formula, and then different decomposition in the WLS method (three parameters versus two parameters). Using numerical experiments, I proved that the proposed method is very promising and provides robust results in terms of data reconstruction and quantitative parameter estimation even in presence of uncorrelated structures, inaccurate migration background models, and noisy data. The main limitation is the assumption of marine towed-streamer data. Through numerical examples, I ranked P-wave impedance, S-wave impedance and density from best-inverted parameter to worst. Analyzing eigenvalue decomposition also confirmed the fact that density is more difficult to reconstruct even in the shallow part of the model.

8.1.2 Iterative approaches

Sparsity-promoting regularization

In Chapter 5, to tackle the ill-posedness of the inverse problem discussed in Chapter 3, I have proposed to add independently ℓ_1-norm constraints to each inverted parameter as regularization terms in the WLS method. To carry out the optimization problem, I have used a popular sparse solver based on the soft thresholding method called the fast iterative shrinkage-thresholding algorithm (FISTA). Through numerical experiments, I have demonstrated that the proposed method can effectively suppress artifacts caused by parameter crosstalk, abrupt truncation of interfaces, incomplete data, noisy data, and large contrast complex geology. Furthermore, I discussed that the computational costs of the WLS and RWLS methods are below 1.0% of the cost of the pseudoinverse Born operator. Therefore, the RWLS method delivers high-resolution images requiring less computational complexity than least-squares migration.

Preconditioning least-squares reverse time migration

In Chapter 6, I have implemented the variable-density pseudoinverse Born operator (Chapter 3) for preconditioning LSRTM algorithm. Through numerical experiments, I have showed that the proposed preconditioning scheme not only largely reduces the required number of iterations to achieve a given data misfit but also significantly increases the quality of the inverted images even in presence of strong parameter cross-talk and inaccurate migration background models. Using a small-sized model, I have further analyzed the shape of the multi-parameter Hessian with and without preconditioner: the preconditioner not only manages to reduce the effect of the wavelet and compensate for geometrical spreading but also mitigates the parameter cross-talks.

Discussion

One would wonder what are the pros and cons of preconditioned LSRTM over regularized pseudoinverse Born operator. The computational cost of preconditioned LSRTM is much more expensive than the regularized pseudoinverse Born operator: the former requires eight wavefield propagations per iteration (2 forward + 2 adjoint of forward + 2 direct inverse + 2 adjoint of direct inverse = 8) whereas the latter requires only two wavefield propagations in total. In the case of an ideal acquired seismic data (dense source and receivers with limited noise level), preconditioned LSRTM provides more reliable results with a resolution close to the seismic wavelength by fitting seismic observed data. But sometimes seismic surveys perform only a sparse recording of the wavefields at the source/receiver positions and so the inverse problem has a huge null space, meaning that there are a lot of models that can fit the recorded data. The regularized pseudoinverse Born operator can tackle this ill-posed problem by injecting prior information via regularization to drive the inversion toward a good model everywhere. Thus my suggestion would be to first start the imaging process with the regularized pseudoinverse Born operator and then further improve the results using iterative LSRTM. An interesting alternative would be introducing the regularization techniques within the preconditioned LSRTM. To do this, I would suggest employing Bregmanized operator splitting method (Zhang et al., 2010) to solve this nonlinearly constrained minimization problem (Zand and Siahkoohi, 2021).

8.1.3 A case study from the Eastern Nankai Trough

In Chapter 7, I have investigated the importance of accounting for density variations in pseudoinverse Born inversion using a marine real dataset from the Eastern Nankai Trough, offshore Japan. The inversion results demonstrates that the variable-density pseudoinverse Born operator yields subsurface images that exhibits a superior resolution (less artifacts and more continuously imaged reflectors) and better reconstructs the field

data than does the constant-density approach. The analysis illustrated that neglecting the density parameter may produce artifacts in the image domain leading to incorrect AVO response in the data domain.

Unfortunately, there was no well data to confirm the inverted values. By analyzing eigenvalue and eigenvector decomposition, I illustrated that the limited offset in this dataset might cause some uncertainty in the inversion of density parameter. In this case, one can refer to the inverted density as a pseudo-density model, used as a proxy to increase the degree of freedom in the inversion process and recover a "cleaner" impedance model. The second limitation is that the non-linear effects (multiples) are not taken into account in the inversion process.

8.2 Perspectives

8.2.1 Extension to anisotropic media

The knowledge of the wave equation is of fundamental importance for a satisfying understanding of the wave propagation phenomena. Wave propagation in anisotropic media continues being studied by many authors (Duveneck et al., 2008; Qu et al., 2017; He et al., 2019). Assuming an acoustic anisotropy, specifically, a transversely isotropic medium with a vertical symmetry direction (VTI), Figure 8.1 illustrates that the anisotropic parameters can have a significant impact on AVO and thus on the focusing of events in CIGs. The CIGs in Figure 8.1 are obtained by applying an isotropic pseudoinverse Born operator on isotropic (Figure 8.1a) and anisotropic (Figure 8.1c) observed data.

The kinematics of P-wave in VTI media is controlled by the vertical velocity V_{P0} and Thomsen parameters ϵ and δ (Thomsen, 1986). The parameter ϵ is representative of the ratio between the vertical and horizontal P-wave velocity, while δ reflects sub-vertical anisotropy:

$$\epsilon = \frac{c_{11} - c_{33}}{2c_{33}}, \tag{8.1}$$

$$\delta = \frac{(c_{13} + c_{44})^2 - (c_{33} - c_{44})^2}{2c_{33}(c_{33} - c_{44})} \tag{8.2}$$

The anisotropic acoustic wave equation reads (Prieux, 2012)

$$\begin{aligned} -i\omega\rho V_x &= \frac{\partial T_1}{\partial x} + \frac{\partial T_2}{\partial x}, \\ -i\omega\rho V_z &= \frac{\partial T_1}{\partial z} - \frac{\partial T_2}{\partial z}, \\ -i\omega T_1 &= \frac{c_{11} + c_{13}}{2}\frac{\partial V_x}{\partial x} + \frac{c_{13} + c_{33}}{2}\frac{\partial V_z}{\partial z} - i\omega T_1^0, \end{aligned}$$

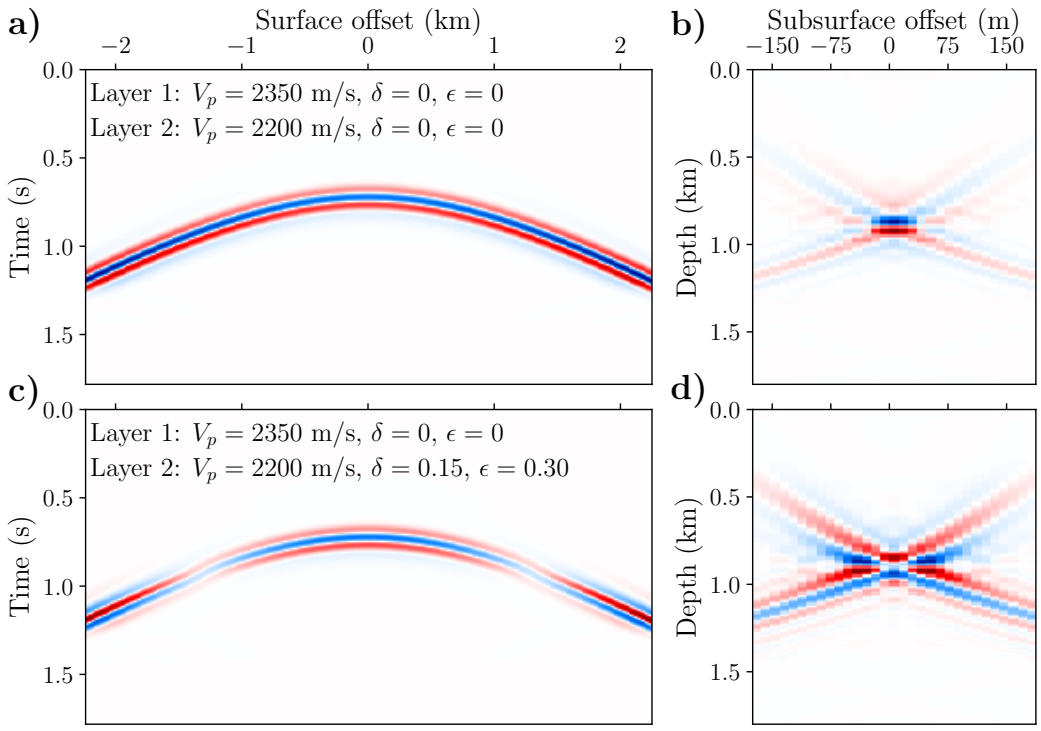

Figure 8.1: The a) observed shot assuming isotropic acoustic medium and its b) inverted CIG, and c) observed shot assuming VTI anisotropic acoustic medium and its d) inverted CIG. The inverted CIGs are obtained with isotropic pseudoinverse Born operator.

$$-i\omega T_2 \;=\; \frac{c_{11}-c_{13}}{2}\frac{\partial V_x}{\partial x} + \frac{c_{13}-c_{33}}{2}\frac{\partial V_z}{\partial z} - i\omega T_2^0, \tag{8.3}$$

In anisotropic acoustic approximation, unlike the acoustic isotropic case, T_2 is different from 0, except if the midpoint is elliptical ($\delta = \epsilon$). By linearizing around smooth background models and dropping higher order terms, the Born approximation can be written as

$$\begin{aligned}
-i\omega(\rho^0+\delta\rho)(V_x^0+\delta V_x) &= \frac{\partial({T_1}^0+\delta T_1)}{\partial x} + \frac{\partial({T_2}^0+\delta T_2)}{\partial x},\\
-i\omega(\rho^0+\delta\rho)(V_z^0+\delta V_z) &= \frac{\partial({T_1}^0+\delta T_1)}{\partial z} - \frac{\partial({T_2}^0+\delta T_2)}{\partial z},\\
-i\omega({T_1}^0+\delta T_1) &= \frac{c_{11}^0+\delta c_{11}+c_{13}^0+\delta c_{13}}{2}\frac{\partial(V_x^0+\delta V_x)}{\partial x} +\\
&\quad \frac{c_{13}^0+\delta c_{13}+c_{33}^0+\delta c_{33}}{2}\frac{\partial(V_z^0+\delta V_z)}{\partial z},\\
-i\omega({T_2}^0+\delta T_2) &= \frac{c_{11}^0+\delta c_{11}-c_{13}^0-\delta c_{13}}{2}\frac{\partial(V_x^0+\delta V_x)}{\partial x} +
\end{aligned}$$

$$\frac{c_{13}^0 + \delta c_{13} - c_{33}^0 - \delta c_{33}}{2} \frac{\partial (V_z^0 + \delta V_z)}{\partial z}, \tag{8.4}$$

and

$$\begin{aligned}
-i\omega\rho^0 \delta V_x &= \frac{\delta T_1}{\partial x} + \frac{\delta T_2}{\partial x} + i\omega\delta\rho V_x^0, \\
-i\omega\rho^0 \delta V_z &= \frac{\partial \delta T_1}{\partial z} - \frac{\delta T_2}{\partial z} + i\omega\delta\rho V_z^0, \\
-i\omega\delta T_1 &= \frac{c_{11}^0 + c_{13}^0}{2}\frac{\partial \delta V_x}{\partial x} + \frac{c_{13}^0 + c_{33}^0}{2}\frac{\partial \delta V_z}{\partial z} + \\
&\quad \frac{\delta c_{11} + \delta c_{13}}{2}\frac{\partial V_x^0}{\partial x} + \frac{\delta c_{13} + \delta c_{33}}{2}\frac{\partial V_z^0}{\partial z}, \\
-i\omega\delta T_2 &= \frac{c_{11}^0 - c_{13}^0}{2}\frac{\partial \delta V_x}{\partial x} + \frac{c_{13}^0 - c_{33}^0}{2}\frac{\partial \delta V_z}{\partial z} + \\
&\quad \frac{\delta c_{11} - \delta c_{13}}{2}\frac{\partial V_x^0}{\partial x} + \frac{\delta c_{13} - \delta c_{33}}{2}\frac{\partial V_z^0}{\partial z}.
\end{aligned} \tag{8.5}$$

The terms in red color in system of equations 8.5 denote the secondary sources that generate the scattered field. Assuming weak anisotropy, I have linearized around an isotropic smooth background (i.e., $\delta_0 = 0$, $\epsilon_0 = 0$), and I have extended the applicability of pseudoinverse Born operator to acoustic VTI medium. Thus I have only considered anisotropic effects on the reflectivity side (dynamics) while the background velocity model remains isotropic. Figure 8.2 illustrates a numerical experiment on the Hess VTI model. The RTM image (Figure 8.2d) contains strong low-frequency artifacts due to the large contrasts in the velocity model. Laplacian filter or the other high-pass filters are often used to remove the artifacts (Guitton et al., 2007). These low-frequency artifacts on the other hand are effectively reduced in the images obtained with isotropic/anisotropic pseudoinverse Born operators (Figure 8.2e and 8.2f). Comparing to RTM and isotropic pseudoinverse Born operator, the anisotropic pseudoinverse Born operator provides cleaner subsurface image with less parameter cross-talk artifacts (highlighted by arrows in Figure 8.2). It would be interesting to further investigate this extension without the weak anisotropic assumption and anisotropy not related to the reflectivity but also incorporated in the background model.

8.2.2 Extension to multi-component seismic data

In this thesis, all of the implementations are restricted to pressure wavefield (mono-component data). Over the past several years, significant advances have been made in the acquisition, processing, and analysis/interpretation of multi-component seismic data resulting in renewed enthusiasm for use of this technology. In this thesis, I have illustrated that the extended model contains the effects of different parameter variations

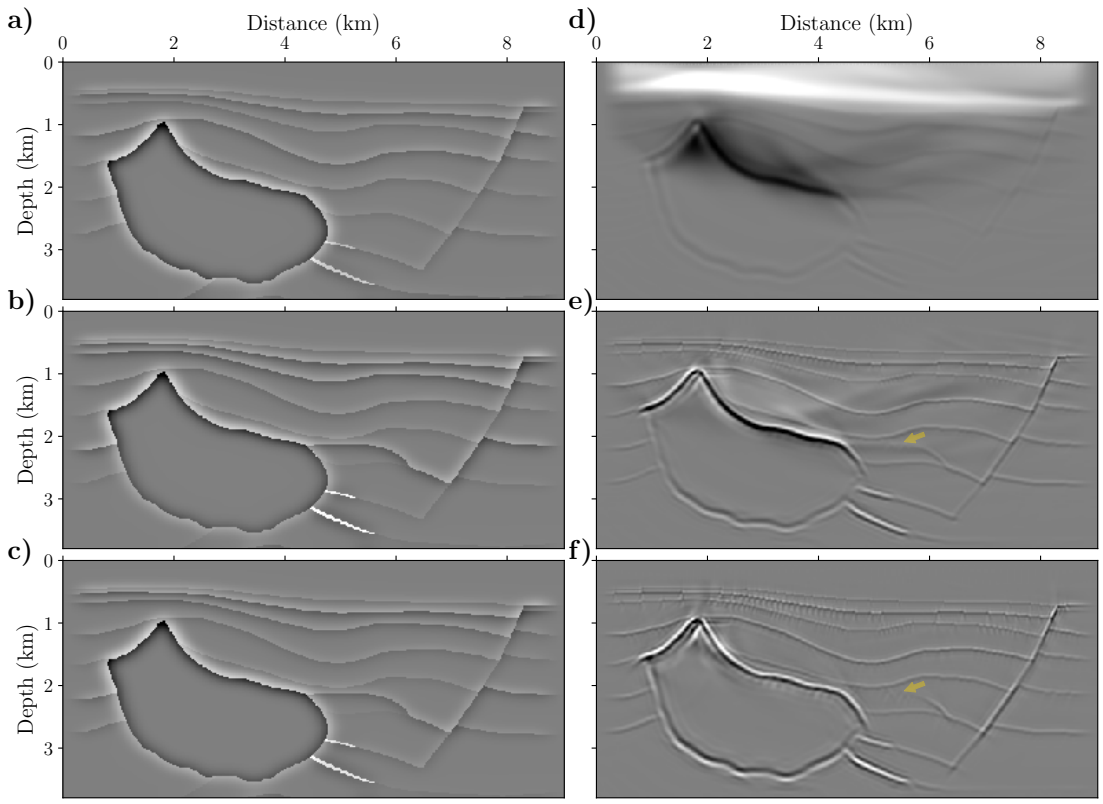

Figure 8.2: The Hess VTI model experiment: exact a) δV_p, b) δV_h and c) δV_n, and inverted δV_p image using d) RTM, e) isotropic pseudoinverse Born operator and f) anisotropic pseudoinverse Born operator. The yellow arrow in (e) and (f) denote the parameter cross-talk.

(velocity and density perturbations). Furthermore, performing the Born modeling operator on either non-physical extended model or physical models (after decomposition) can reconstruct the observed data. In the case of multi-component seismic data, an extended image can be obtained for each component of seismic data. It is still an open question whether these extended model spaces carry useful and manageable information for linear imaging scheme. The inclusion of different components of seismic data in a unified framework in the direct inversion process needs further investigations.

8.2.3 Coupling multi-parameter direct inverse with migration velocity analysis

One important feature observed through Chapters 3 and 4 is that the density variations or elastic effects distort the energy in the subsurface-offset domain CIGs, even in presence

of accurate background model. The key aspect is that the proposed multi-parameter pseudoinverse Born inversion can reconstruct the data even using inaccurate background models. Therefore, this provides evidence that these methods can be introduced in the inner loop of MVA. The curvature of an event in the angle-domain CIG is indeed the bridge between migration and velocity updates in MVA (Biondi and Symes, 2004).

8.2.4 Coupling multi-parameter direct inverse with reflection waveform inversion

Reflection waveform inversion (RWI) is a nested optimization procedure (Chavent et al., 1994; Xu et al., 2012b; Yao et al., 2020). In the inner loop, given the current background velocity, the reflectivity is iteratively updated to minimize the data misfit at zero-/short-offsets. In the outer loop, the best reflectivity derived in the inner loop is used to evaluate the misfit at larger offsets and to update the background velocity model. The inner loop of RWI is recognized as LSRTM, providing a true-amplitude reflectivity section. This makes RWI highly computationally expensive as iterative LSRTM has to be applied for each outer iteration. In the case of single-parameter inversion, there is a dimensional difference between LSRTM in the inner loop of RWI (physical domain) and the pseudoinverse Born operator (extended domain). But the multi-parameter direct inverse proposed in this thesis provides physical results (non-extended domain) and can play a key role as a preconditioner to enlarge the applicability of RWI (Algorithm 9). More research is required to investigate the role of density/impedance in updating the reflectivity (inner loop) and velocity in updating the background model (outer loop).

Algorithm 9: RWI with pseudoinverse Born operator

```
1 Outer loop: for each tomographic iteration do
2     Inner loop: for each reflectivity update do
3         Pseudoinverse Born operator
4         Radon transform
5         Physical decomposition
6     Evaluate the data misfit
7     Update the background model
```

8.2.5 Extension to 3D

In 3D, the observed data has five dimensions (s_x, s_y, r_x, r_y, t). Thus the extension of direct inverse from 2D to 3D requires two extra model space parameters, for example, vertical and horizontal subsurface offsets. New weights in 3D direct inverse should

be defined with a strategy similar to the 2D case: their combination with the Beylkin determinant (Jacobian of transformation between subsurface parameters and acquisition surface coordinates) arising in linear phase approximation of the normal operator should cancel the geometrical spreading present in the propagating fields. The main difficulty in 3D extension is its computational cost: 3D wavefield propagation, storing 3D images, and cross-correlation in imaging condition for each couple (h_x, h_y). New strategies such as computing the CIGs only at specific image points (Yang and Sava, 2015) or computing the CIGs with randomly sampled space shifts (Yang et al., 2020) should be further investigated for a more affordable 3D extension.

Bibliography

Abma, R., Sun, J. and Bernitsas, N. [1999] Antialiasing methods in Kirchhoff migration. *Geophysics*, **64**(6), 1783–1792. (cit. on p. 62.)

Aghamiry, H.S., Gholami, A. and Operto, S. [2018] Hybrid Tikhonov + total-variation regularization for imaging large–contrast media by full–waveform inversion. In: *88th Annual International Meeting, SEG, Expanded Abstracts*, 1253–1257. (cit. on pp. 161 and 189.)

Aghamiry, H.S., Gholami, A. and Operto, S. [2019a] Implementing bound constraints and total–variation regularization in extended full–waveform inversion with the alternating direction method of multiplier: application to large contrast media. *Geophysical Journal International*, **218**, 855–872. (cit. on pp. 161 and 189.)

Aghamiry, H.S., Gholami, A. and Operto, S. [2019b] On the robustness of $\ell 1$–regularized ADMM-based wavefield reconstruction inversion against compressed acquisition sampling. In: *89th Annual International Meeting, SEG, Expanded Abstracts*, 1385–1389. (cit. on p. 161.)

Aghamiry, H.S., Gholami, A. and Operto, S. [2020a] Accurate and efficient data-assimilated wavefield reconstruction in the time domain. *Geophysics*, **85**(2), A7–A12. (cit. on p. 17.)

Aghamiry, H.S., Gholami, A. and Operto, S. [2020b] Compound Regularization of Full–Waveform Inversion for Imaging Piecewise Media. *IEEE Transactions on Geoscience and Remote Sensing*, **58**, 1192–1204. (cit. on p. 182.)

Akcelik, V. [2003] *Multiscale Newton-Krylov methods for inverse acoustic wave propagation*. Ph.D. thesis, Carnegie Mellon University. (cit. on p. 10.)

Aki, K. and Richards, P.G. [1980] *Quantitative seismology*. W. H. Freeman and Co. (cit. on p. 11.)

Al-Yahya, K. [1989] Velocity analysis by iterative profile migration. *Geophysics*, **54**(6), 718–729. (cit. on pp. 16, 17, 35, and 49.)

Alali, A., Sun, B. and Alkhalifah, T. [2020] The effectiveness of a pseudo-inverse extended Born operator to handle lateral heterogeneity for imaging and velocity analysis applications. *Geophysical Prospecting*, **68**(4), 1154–1166. (cit. on p. 69.)

Albertin, U., Sava, P., Etgen, J. and Maharramov, M. [2006] Adjoint wave-equation velocity analysis. In: *76th Annual International Meeting, SEG, Expanded Abstracts*.

3345–3349. (cit. on p. 35.)

Alerini, M., Lambaré, G., Baina, R., Podvin, P. and Le Bégat, S. [2007] 2D PP / PS-Stereotomography : P- and S-waves velocities estimation from OBC data. *Geophysical Journal International*, **170**, 725–736. (cit. on p. 69.)

Alkhalifah, T. [2005] Tau-Migration and Velocity Analysis: Application to Data from the Red Sea. *Geophysical Prospecting*, **53**(5), 643–653. (cit. on p. 35.)

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 p. 62.)

Almomin, A. and Biondi, B. [2013] Tomographic full waveform inversion (TFWI) by successive linearizations and scale separations. In: *93th Annual International Meeting, SEG, Expanded Abstracts*. 1048–105. (cit. on p. 63.)

Amundsen, L., Zhou, H., Reitan, A. and Weglein, A.B. [2013] On seismic deghosting by spatial deconvolution. *Geophysics*, **78**(6), V267–V271. (cit. on p. 246.)

Anagaw, A.Y. and Sacchi, M.D. [2012] Edge–preserving seismic imaging using the total variation method. *Journal of Geophysics and Engineering*, **9**, 138–146. (cit. on p. 189.)

Ashi, J., Tokuyama, H. and Taira, A. [2002] Distribution of methane hydrate BSRs and its implication for the prism growth in the Nankai Trough. *Marine Geology*, **187**(1-2), 177–191. (cit. on p. 238.)

Askan, A. and Bielak, J. [2008] Full anelastic waveform tomography including model uncertainty. *Bulletin of the Seismological Society of America*, **98**, 2975–2989. (cit. on p. 189.)

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. 7, 153, and 161.)

Audebert, F., Diet, J.P., Guillaume, P., Jones, I. and Zhang, X. [1997] CRP Scan: 3-D preSDM velocity analysis via zero offset tomographic inversion. In: *67th Annual International Meeting, SEG, Expanded Abstracts*. 1805–1808. (cit. on p. 35.)

Babich, V. and Buldyrev, V.S. [1989] *Asymptotic methods in short-wavelength diffraction theory*. Springer-Verlag. (cit. on p. 6.)

Barnes, C. and Charara, M. [2009] The domain of applicability of acoustic full-waveform inversion for marine seismic data. *Geophysics*, **74**(6), WCC91–WCC103. (cit. on pp. 117 and 128.)

Baysal, E., Kosloff, D.D. and Sherwood, J.W. [1983] Reverse time migration. *Geophysics*, **48**(11), 1514–1524. (cit. on pp. 19, 37, 77, 115, 159, 194, 197, 236, and 237.)

Beaudoin, G. and Ross, A.A. [2007] Field design and operation of a novel deepwater, wide-azimuth node seismic survey. *The Leading Edge*, **26**(4), 494–503. (cit. on p. 4.)

Beck, A. and Teboulle, M. [2009] A Fast Iterative Shrinkage–Thresholding Algorithm for Linear Inverse Problems. *SIAM Journal on Imaging Sciences*, **2**, 183–202. (cit. on pp. 167, 187, and 242.)

Belge, M., Kilmer, M.E. and Miller, E.L. [1998] Simultaneous multiple regularization parameter selection by means of the L-hypersurface with applications to linear inverse problems posed in the wavelet transform domain. In: *Bayesian Inference for Inverse Problems*, 3459, 328–336. (cit. on p. 189.)

Belge, M., Kilmer, M.E. and Miller, E.L. [2002] Efficient determination of multiple regularization parameters in a generalized L-curve framework. *Inverse problems*, **18**, 1161. (cit. on p. 189.)

Ben-Hadj-Ali, H., Operto, S. and Virieux, J. [2011] An efficient frequency-domain full waveform inversion method using simultaneous encoded sources. *Geophysics*, **76**(4), R109–R124. (cit. on p. 60.)

Bérenger, J.P. [1994] A perfectly matched layer for the absorption of electromagnetic waves. *Journal of computational physics*, **114**(2), 185–200. (cit. on pp. 28 and 202.)

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 p. 18.)

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), 99–108. (cit. on pp. 22, 44, 119, 197, and 237.)

Bharadwaj, P., Mulder, W. and Drijkoningen, G. [2018] A parameterization analysis for acoustic full-waveform inversion of sub-wavelength anomalies. *arXiv preprint arXiv:1804.01184*. (cit. on pp. 15 and 97.)

Billette, F. and Lambaré, G. [1998] Velocity macro-model estimation from seismic reflection data by stereotomography. *Geophysical Journal International*, **135**(2), 671–680. (cit. on pp. 38, 48, and 69.)

Billette, F.J. and Brandsberg-Dahl, S. [2004] The 2004 BP Velocity Benchmark. In: *67th Annual International Conference and Exhibition, EAGE, Extended Abstracts*, B035. (cit. on pp. 161 and 179.)

Biondi, B. and Almomin, A. [2014] Simultaneous inversion of full data bandwidth by tomographic full-waveform inversion. *Geophysics*, **79(3)**, WA129–WA140. (cit. on p. 17.)

Biondi, B. and Shan, G. [2002] Prestack imaging of overturned reflections by reverse time migration. In: *72th Annual International Meeting, SEG, Expanded Abstracts*, 1284–1287. (cit. on pp. 19, 160, and 237.)

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. 35, 38, 43, 87, 152, 160, and 276.)

Biondi, E., Meadows, M. and Biondi, B. [2019] Amplitude preserving migration through extended acoustic least-squares RTM. In: *89th Annual International Meeting, SEG, Expanded Abstracts*, 4196–4200. (cit. on pp. 39 and 226.)

Bishop, T.N., Bube, K.P., Cutler, R.T., Langan, R.T., Love, P.L., Resnick, J.R., Shuey, R.T. and Spinder, D.A. [1985] Tomographic determination of velocity and depth in

laterally varying media. *Geophysics*, **50**, 903–923. (cit. on p. 69.)

Blackburn, J., Daniels, J., Dingwall, S., Hampden-Smith, G., Leaney, S., Le Calvez, J., Nutt, L., Menkiti, H., Sanchez, A. and Schinelli, M. [2007] Borehole seismic surveys: Beyond the vertical profile. *Oilfield Review*, **19**(3), 20–35. (cit. on pp. xiii, 4, and 5.)

Bleistein, N. [1987] On the imaging of reflectors in the Earth. *Geophysics*, **52**(7), 931–942. (cit. on pp. 22, 23, 44, and 237.)

Borisov, D., Stopin, A. and Plessix, R.É. [2014] Acoustic pseudo–density full waveform inversion in the presence of hard thin beds. In: *76th EAGE Conference and Exhibition, Extended Abstracts*. We–P01–10. (cit. on p. 265.)

Born, M. [1926] Quantum mechanics of impact process. *Z. Phys*, **38**, 803–827. (cit. on p. 18.)

Boschetti, F., Dentith, M.C. and List, R.D. [1996] Inversion of seismic refraction data using genetic algorithms. *Geophysics*, **61**(6), 1715–1727. (cit. on p. 9.)

Brandsberg-Dahl, S., de Hoop, M.V. and Ursin, B. [2003] Focusing in dip and AVA compensation on scattering–angle/azimuth common image gathers. *Geophysics*, **68**(1), 232–254. (cit. on pp. 41 and 87.)

Brezinski, C., Redivo-Zaglia, M., Rodriguez, G. and Seatzu, S. [2003] Multi-parameter regularization techniques for ill-conditioned linear systems. *Numerische Mathematik*, **94**, 203–228. (cit. on p. 189.)

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 pp. 10, 11, and 15.)

Brossier, R., Operto, S. and Virieux, J. [2010] Which data residual norm for robust elastic frequency-domain full waveform inversion? *Geophysics*, **75**(3), R37–R46. (cit. on p. 146.)

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. 11, 16, and 69.)

Brossier, R., Virieux, J. and Operto, S. [2008] Parsimonious finite-volume frequency-domain method for 2-D P–SV-wave modelling. *Geophysical Journal International*, **175**, 541–559. (cit. on p. 118.)

Brown, M.P. and Guitton, A. [2005] Least-Squares Joint Imaging of Multiples and Primaries. *Geophysics*, **70**(5), S79–S89. (cit. on p. 57.)

de Bruin, C.G.M., Wapenaar, C.P.A. and Berkhout, A.J. [1990] Angle–dependent reflectivity by means of prestack migration. *Geophysics*, **55**(9), 1223–1234. (cit. on p. 87.)

Bunks, C., Saleck, F.M., Zaleski, S. and Chavent, G. [1995] Multiscale seismic waveform inversion. *Geophysics*, **60**(5), 1457–1473. (cit. on p. 15.)

Byrd, R.H., Lu, P. and Nocedal, J. [1995] A Limited Memory Algorithm for Bound Constrained Optimization. *SIAM Journal on Scientific and Statistical Computing*, **16**(5), 1190–1208. (cit. on pp. 10 and 63.)

Caldwell, J. [1999] Marine multicomponent seismology. *The Leading Edge*, **18**, 1274–1282. (cit. on p. 116.)

Castagna, J.P. and Smith, S.W. [1994] Comparison of AVO indicators: A modeling study. *Geophysics*, **59**(12), 1849–1855. (cit. on p. 87.)

Castellanos, C., Métivier, L., Operto, S., Brossier, R. and Virieux, J. [2014] Fast full waveform inversion with source encoding and second-order optimization methods. *Geophysical Journal International*, **200**(2), 720–744. (cit. on p. 60.)

Červenỳ, V., Molotkov, I.A., Molotkov, I.A. and Pšenčík, I. [1977] *Ray method in seismology*. Univerzita Karlova. (cit. on pp. 6 and 18.)

Chauris, H. [2019] Full waveform inversion7. *Seismic imaging: a practical approach*, 123–145. (cit. on pp. 14 and 195.)

Chauris, H. and Cocher, E. [2017] From migration to inversion velocity analysis. *Geophysics*, **82**(03), S207–S223. (cit. on pp. xiv, 18, 22, 23, 24, 40, 44, 45, 46, 47, 49, 50, 64, 69, 78, 79, 83, 116, 117, 119, 120, 121, 126, 140, 142, 160, 163, 168, 195, 196, 200, 201, 202, 230, 237, 247, 248, 249, and 264.)

Chauris, H. and Cocher, E. [2018] Review of different expressions for the extended Born approximate inverse operator. In: *80th EAGE Conference and Exhibition, Extended Abstracts*. WS07. (cit. on pp. xxvii, 23, 45, 47, 67, 68, 83, 116, 196, 230, and 231.)

Chauris, H., Li, Y. and Cocher, E. [2017] Image–domain versus data–domain velocity analysis based on true–amplitude subsurface extended migration. In: *74th EAGE Conference and Exhibition, Extended Abstracts*. European Association of Geoscientists & Engineers, WS09 D02. (cit. on pp. 21, 69, and 70.)

Chauris, H. and Noble, M.S. [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. 35 and 48.)

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. xv, 38, 39, 69, 70, and 73.)

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 p. 48.)

Chavent, G. [1974] Identification of functional parameters in partial differential equations. In: *Joint Automatic Control Conference*, 12. 155–156. (cit. on p. 13.)

Chavent, G., Clément, F. and Gómez, S. [1994] Automatic determination of velocities via migration-based traveltime waveform inversion: A synthetic data example. In: *64th Annual International Meeting, SEG, Expanded Abstracts*. 1179–1182. (cit. on pp. 16, 58, and 276.)

Chavent, G. and Jacewitz, C.A. [1995] Determination of background velocities by multiple migration fitting. *Geophysics*, **60**(2), 476–490. (cit. on p. 48.)

Chavent, G. and Plessix, R.É. [1999] An optimal true–amplitude least–squares prestack

depth–migration operator. *Geophysics*, **64**, 508–515. (cit. on pp. 21, 44, and 163.)
Chen, K. and Sacchi, M.D. [2017] Elastic least-squares reverse time migration via linearized elastic full-waveform inversion with pseudo-Hessian preconditioning. *Geophysics*, **82**(5), S341–S358. (cit. on pp. 116, 152, 195, and 237.)
Chen, K. and Sacchi, M.D. [2018] Should we include the density perturbation in elastic least–squares reverse time migration? In: *88th Annual International Meeting, SEG, Expanded Abstracts*. 4226–4230. (cit. on pp. 10, 22, and 102.)
Chen, K. and Sacchi, M.D. [2019] The importance of including density in elastic least–squares reverse time migration: Multiparameter crosstalk and convergence. *Geophysical Journal International*, **216**, 61–80. (cit. on pp. 116, 195, 237, and 264.)
Chopra, S. and Castagna, J.P. [2005] *AVO*. Society of Exploration Geophysicists. (cit. on pp. 84, 87, and 151.)
Claerbout, J.F. [1971] Toward a unified theory of reflector mapping. *Geophysics*, **36**(3), 467–481. (cit. on pp. 19 and 37.)
Claerbout, J.F. [1985] *Imaging the Earth's interior*. Blackwell scientific publications Oxford. (cit. on pp. xiii, xxii, 8, 9, 19, 21, 80, 119, 197, 198, and 248.)
Claerbout, J.F. [1992] *Earth soundings analysis: Processing versus inversion*. Blackwell Scientific Publications. (cit. on pp. 120 and 163.)
Claerbout, J.F. [2014] *Geophysical Image Estimation by Example*. Blackwell Scientific Publications. (cit. on p. 28.)
Claerbout, J.F. and Green, I. [2010] Basic Earth imaging. *Cambridge: Free Software Foundation*. (cit. on pp. 240 and 243.)
Clapp, M.L., Clapp, R.G. and Biondi, B.L. [2005] Regularized least-squares inversion for 3-D subsalt imaging. In: *75th Annual International Meeting, SEG, Expanded Abstracts*. 1814–1817. (cit. on pp. 20 and 44.)
Clapp, R.G. and Biondi, B. [2000] Tau domain migration velocity analysis using angle CRP gathers and geologic constraints. In: *70th Annual International Meeting, SEG, Expanded Abstracts*. 926–929. (cit. on p. 35.)
Clément, F., Chavent, G. and Gómez, S. [2001] Migration-based traveltime waveform inversion of 2D simple structures: A synthetic example. *Geophysics*, **66**(3), 845–860. (cit. on pp. 44, 58, and 69.)
Cocher, E. [2017] *Iterative Migration Velocity Analysis: extension to surface-related multiple reflections*. Ph.D. thesis, MINES ParisTech–PSL Research University. (cit. on pp. 66 and 202.)
Cocher, E., Chauris, H. and Plessix, R.É. [2017] Seismic iterative migration velocity analysis: two strategies to update the velocity model. *Computational Geosciences*, **21**, 1–22. (cit. on pp. 44 and 47.)
Cocher, E., Chauris, H. and Plessix, R.É. [2018] Toward a stable iterative migration velocity analysis scheme. *Geophysics*, **83**(5), R475–R495. (cit. on pp. 57, 117, and 196.)
Cordua, K.S., Hansen, T.M. and Mosegaard, K. [2012] Monte Carlo full-waveform inversion of crosshole GPR data using multiple-point geostatistical a priori information.

Geophysics, **77**(2), H19–H31. (cit. on p. 9.)

Courant, R. and Hilbert, D. [1962] *Methods of mathematical physics*. WileyInterscience. (cit. on p. 121.)

Crase, E., Pica, A., Noble, M., McDonald, J. and Tarantola, A. [1990] Robust elastic nonlinear waveform inversion: Application to real data. *Geophysics*, **55**(5), 527–538. (cit. on p. 10.)

Craven, P. and Wahba, G. [1978] Smoothing noisy data with spline functions. *Numerische mathematik*, **31**, 377–403. (cit. on p. 189.)

Cullison, T. and Sava, P. [2011] An image-guided method for automatically picking common-image-point gathers. In: *81th Annual International Meeting, SEG, Expanded Abstracts*. (cit. on p. 61.)

Dafni, R. and Symes, W.W. [2016] Scattering and dip angle decomposition based on subsurface offset extended wave–equation migration. *Geophysics*, **81**(3), S119–S138. (cit. on p. 87.)

Dafni, R. and Symes, W.W. [2018] Asymptotic inversion of the variable density acoustic model. In: *88th Annual International Meeting, SEG, Expanded Abstracts*. 570–574. (cit. on pp. 43, 52, 78, 79, 84, 89, 92, 117, 120, 122, 160, 164, 165, 166, 168, 196, 198, 200, 237, 249, and 269.)

Dai, W., Fowler, P. and Schuster, G.T. [2012] Multi-source least-squares reverse time migration. *Geophysical Prospecting*, **60**(4), 681–695. (cit. on pp. 20, 22, 78, 116, 120, and 160.)

Dai, W. and Schuster, G.T. [2013] Plane-wave least-squares reverse-time migration. *Geophysics*, **78**(4), S165–S177. (cit. on p. 22.)

Dai, W., Wang, X. and Schuster, G.T. [2011] Least-squares migration of multisource data with a deblurring filter. *Geophysics*, **76**(5), R135–R146. (cit. on pp. 77 and 237.)

Dai, W., Xu, Z. and Coates, R. [2015] Least-squares reverse-time migration for visco-acoustic media. In: *85th Annual International Meeting, SEG, Expanded Abstracts*. 3387–3391. (cit. on p. 22.)

Dariu, H., Garotta, R. and Granger, P. [2003] Simultaneous inversion of PP and PS wave AVO/AVA data using simulated annealing. In: *73th Annual International Meeting, SEG, Expanded Abstracts*, 120–123. (cit. on p. 9.)

Datta, D., Sen, M.K., Liu, F. and Morton, S. [2019] Full-waveform inversion of salt models using shape optimization and simulated annealing. *Geophysics*, **84**(5), R793–R804. (cit. on p. 9.)

Daubechies, I., Defrise, M. and De Mol, C. [2004] An iterative thresholding algorithm for linear inverse problems with a sparsity constraint. *Communications on Pure and Applied Mathematics: A Journal Issued by the Courant Institute of Mathematical Sciences*, **57**, 1413–1457. (cit. on p. 167.)

Deregowski, S.M. [1990] Common-offset migrations and velocity analysis. *First Break*, **8**, 225–234. (cit. on p. 37.)

Devaney, A.J. [1982] A filtered backpropagation algorithm for diffraction tomography.

Ultrasonic imaging, **4**(4), 336–350. (cit. on p. 13.)

Díaz, E. and Sava, P. [2015] Wavefield tomography using reverse time migration backscattering. *Geophysics*, **80**(1), R57–R69. (cit. on p. 43.)

Diaz, E. and Sava, P. [2017] Cascaded wavefield tomography and inversion using extended common-image-point gathers: A case study. *Geophysics*, **82**(5), S391–S401. (cit. on pp. 36, 61, and 70.)

Diaz, E., Sava, P., Snieder, R. and Singh, S. [2016] Imaging the model through the wave equation. In: *86th Annual International Meeting, SEG, Expanded Abstracts*. 3708–3713. (cit. on p. 62.)

Dickens, T., Vdovina, T., Cha, Y. and Lu, R. [2020] Acoustic and elastic FWI in complex salt environments: Learnings from a Gulf of Mexico production field. In: *82th Annual International Conference and Exhibition, EAGE, Workshop Programme*, WS02–06. (cit. on p. 122.)

Dickens, T. and Winbow, G. [2011] RTM angle gathers using Poynting vectors. In: *81th Annual International Meeting, SEG, Expanded Abstracts*. (cit. on p. 59.)

Dix, C.H. [1952] *Seismic prospecting for oil*. Harper. (cit. on p. 245.)

Docherty, P., Silva, R., Singh, S., Song, Z. and Wood, M. [1997] Migration velocity analysis using a genetic algorithm. *Geophysical Prospecting*, **45**, 865–878. (cit. on p. 63.)

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 and Exhibition, Extended Abstracts*. We–SRS2–12. (cit. on pp. 22 and 23.)

Duquet, B. and Lailly, P. [2006] Efficient 3D wave-equation migration using virtual planar sources. *Geophysics*, **71**(55), S185–S197. (cit. on p. 59.)

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 pp. 20 and 160.)

Dutta, G., Giboli, M., Agut, C., Williamson, P. and Schuster, G.T. [2017] Least–squares reverse time migration with local Radon–based preconditioning. *Geophysics*, **82**(2), S75–S84. (cit. on p. 189.)

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 pp. 22, 78, 116, and 195.)

Duveneck, E. [2013] A Pragmatic Approach for Computing Full-volume RTM Reflection Angle/azimuth Gathers. In: *75th EAGE Conference and Exhibition, Extended Abstracts*. Tu–11–01. (cit. on p. 59.)

Duveneck, E. and Bakker, P.M. [2011] Stable P-wave modeling for reverse-time migration in tilted TI media. *Geophysics*, **76**(2), S65–S75. (cit. on p. 22.)

Duveneck, E., Milcik, P., Bakker, P.M. and Perkins, C. [2008] Acoustic VTI wave equations and their application for anisotropic reverse-time migration. In: *78th Annual International Meeting, SEG, Expanded Abstracts*. 2186–2190. (cit. on pp. 22 and 272.)

Eckhardt, W. [1994] Velocity model updating using image gathers. *Geophysical Prospecting*, **42**, 975–986. (cit. on p. 35.)

Ensley, R.A. [1984] Comparison of P-and S-wave seismic data; a new method for detecting gas reservoirs. *Geophysics*, **49**(9), 1420–1431. (cit. on p. 11.)

Etgen, J., Gray, S.H. and Zhang, Y. [2009] An overview of depth imaging in exploration geophysics. *Geophysics*, **74**(6), WCA5–WCA17. (cit. on pp. 35, 37, and 195.)

Farshad, M. and Chauris, H. [2020a] From constant- to variable–density inverse extended Born modeling. *Geophysics*, **85**(4), S217–S232. (cit. on pp. 52, 76, 117, 120, 121, 122, 123, 124, 152, 154, 155, 160, 161, 164, 165, 166, 168, 187, 196, 198, 199, 200, 238, 248, 249, 251, and 264.)

Farshad, M. and Chauris, H. [2020b] Reconstruction of quantitative variable density acoustic reflectivity in the context of velocity model building. In: *82th Annual International Conference and Exhibition, EAGE, Extended Abstracts*, WS09. (cit. on pp. xiii, xv, 25, 26, 43, 51, and 120.)

Farshad, M. and Chauris, H. [2020c] Sparsity–Promoting Pseudo-Inverse Born Operator in the Presence of Density Variations: An Efficient MultiParameter Imaging Tool. In: *82th Annual International Conference and Exhibition, EAGE, Extended Abstracts*, Th–Dome5–17. (cit. on pp. 117, 153, and 196.)

Farshad, M. and Chauris, H. [2021a] Accelerating the multi-parameter least-squares reverse time migration using an appropriate preconditioner. *Computational Geosciences*, **25**, 2071–2092. (cit. on pp. 45, 47, and 192.)

Farshad, M. and Chauris, H. [2021b] From acoustic to elastic inverse extended Born modeling: a first insight in the marine environment. *Geophysics*, **86**(6), R939–R957. (cit. on pp. 52, 54, and 114.)

Farshad, M. and Chauris, H. [2021c] Sparsity–promoting multiparameter pseudoinverse Born inversion in acoustic media. *Geophysics*, **86**(3), S205–S220. (cit. on pp. xv, 52, 53, 158, 199, 242, 251, and 264.)

Farshad, M., Chauris, H. and Noble, M. [2021] The importance of including density in multi-parameter asymptotic linearized direct waveform inversion: a case study from the Eastern Nankai Trough. *Geophysical Journal International*, **228**(2), 1373–1391. (cit. on p. 235.)

Farshad, M., Gholami, A. and Mobarakeh, A.M. [2018] AVO–preserving Deconvolutive Radon transform. In: *80th Annual International Conference and Exhibition, EAGE, Extended Abstracts*. Tu–P3–07. (cit. on p. 110.)

Fei, W. and Williamson, P. [2010] On the gradient artifacts in migration velocity analysis based on differential semblance optimization. In: *80th Annual International Meeting, SEG, Expanded Abstracts*, 4071–4076. (cit. on pp. xiv, 49, 50, and 168.)

Fei, W., Williamson, P. and Khoury, A. [2009] 3-D common-azimuth wave-equation migration velocity analysis. In: *79th Annual International Meeting, SEG, Expanded Abstracts*. 2283–2287. (cit. on p. 36.)

Feng, Z. [2018] *Wave-Equation Elastic Least-Squares Migration and Migration Velocity*

Analysis. Ph.D. thesis, KAUST. (cit. on pp. 51 and 59.)

Feng, Z. and Schuster, G.T. [2017] Elastic least-squares reverse time migration. *Geophysics*, **82**(2), S143–S157. (cit. on pp. 22, 116, and 152.)

Fichtner, A. [2010] *Full seismic waveform modelling and inversion*. Springer Science & Business Media. (cit. on pp. 35 and 58.)

Fletcher, R.P., Nichols, D., Bloor, R. and Coates, R.T. [2015] Least-Squares Migration: Data Domain Versus Image Domain. In: *77th EAGE Conference and Exhibition, Extended Abstracts*. N106–06. (cit. on p. 44.)

Fomel, S. [2007] Shaping regularization in geophysical–estimation problems. *Geophysics*, **72**(2), R29–R36. (cit. on p. 161.)

Fomel, S., Berryman, J.G., Clapp, R.G. and Prucha, M. [2008] Iterative resolution estimation in least-squares Kirchhoff migration. *Geophysical prospecting*, **50**(6), 577–588. (cit. on p. 20.)

Forgues, É. [1996] *Inversion linearisée multi-paramètres via la théorie des rais*. Ph.D. thesis, Institut Français du Pétrole - University Paris VII. (cit. on pp. 164 and 198.)

Forgues, E. and Lambaré, G. [1997] Parameterization study for acoustic and elastic ray + Born inversion. *Journal of Seismic Exploration*, **6**(2-3), 253–277. (cit. on pp. 11, 15, 81, 122, and 250.)

Foucher, J.P., Nouzé, H. and Henry, P. [2002] Observation and tentative interpretation of a double BSR on the Nankai slope. *Marine Geology*, **187**(1–2), 161–175. (cit. on p. 244.)

French, W.S. [1975] Computer migration of oblique seismic reflection profiles. *Geophysics*, **40**(6), 961–980. (cit. on p. 18.)

Fu, L. and Symes, W.W. [2017] An adaptive multiscale algorithm for efficient extended waveform inversion. *Geophysics*, **82**(3), R183–R197. (cit. on p. 61.)

Fujii, T., Suzuki, K., Takayama, T., Tamaki, M., Komatsu, Y., Konno, Y., Yoneda, J., Yamamoto, K. and Nagao, J. [2015] Geological setting and characterization of a methane hydrate reservoir distributed at the first offshore production test site on the Daini-Atsumi Knoll in the eastern Nankai Trough, Japan. *Marine and Petroleum Geology*, **66**, 310–322. (cit. on pp. 238 and 256.)

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. 9.)

Gardner, G., French, W. and Matzuk, T. [1974a] Elements of migration and velocity analysis. *Geophysics*, **39**(6), 811–825. (cit. on pp. 18, 35, and 37.)

Gardner, G., Gardner, L. and Gregory, A. [1974b] Formation velocity and density–The diagnostic basics for stratigraphic traps. *Geophysics*, **39**(6), 770–780. (cit. on p. 179.)

Garia, S., Pal, A.K., Ravi, K. and Nair, A.M. [2019] A comprehensive analysis on the relationships between elastic wave velocities and petrophysical properties of sedimentary rocks based on laboratory measurements. *Journal of Petroleum Exploration and Production Technology*, **9**, 1869–1881. (cit. on p. 179.)

Gauthier, O., Virieux, J. and Tarantola, A. [1986] Two-dimensional nonlinear inversion of seismic waveforms: Numerical results. *Geophysics*, **51**(7), 1387–1403. (cit. on p. 15.)

Gholami, A. [2017] Deconvolutive Radon transform. *Geophysics*, **82**(2), V117–V125. (cit. on pp. 110 and 244.)

Gholami, A., Aghamiry, H.S. and Abbasi, M. [2018] Constrained nonlinear amplitude variation with offset inversion using Zoeppritz equations. *Geophysics*, **83**(3), R245–R255. (cit. on p. 226.)

Gholami, A. and Farshad, M. [2019a] Fast hyperbolic Radon transform using chirp–z transform. *Digital Signal Processing*, **87**, 34–42. (cit. on pp. 110 and 242.)

Gholami, A. and Farshad, M. [2019b] The Shuey–Radon transform. *Geophysics*, **84**(3), V197–V206. (cit. on pp. 110 and 161.)

Gholami, A. and Hosseini, S.M. [2013] A balanced combination of Tikhonov and total variation regularizations for reconstruction of piecewise-smooth signals. *Signal Processing*, **93**, 1945–1960. (cit. on p. 161.)

Gholami, A. and Sacchi, M.D. [2017] Time–invariant Radon transform by generalized Fourier slice theorem. *Inverse Problems & Imaging*, **11**(3), 501–519. (cit. on p. 110.)

Gholami, A. and Siahkoohi, H.R. [2010] Regularization of linear and non–linear geophysical ill–posed problems with joint sparsity constraints. *Geophysical Journal International*, **180**, 871–882. (cit. on p. 161.)

Gholami, A. and Zand, T. [2017] Fast ℓ_1–regularized Radon transforms for seismic data processing. *Digital Signal Processing*, **71**, 83–94. (cit. on p. 110.)

Gill, P.E., Murray, W. and Wright, M.H. [1981] *Practical optimization*. SIAM. (cit. on p. 9.)

Gilpatrick, R. and Fouquet, D. [1989] A user's guide to conventional VSP acquisition. *The Leading Edge*, **8**(3), 34–39. (cit. on p. 4.)

Gockenbach, M.C. and Symes, W.W. [1995] Waveform inversion for velocity: Where havethe minima gone? In: *65th Annual International Meeting, SEG, Expanded Abstracts*. 1095–1098. (cit. on p. 35.)

Guerra, C. and Biondi, B. [2011] Fast 3D migration-velocity analysis by wavefield extrapolation using the prestack exploding-reflector model. *Geophysics*, **76**(5), WB151–WB167. (cit. on pp. 36 and 60.)

Guillaume, P., Lambaré, G., Leblanc, O., Mitouard, P., Moigne, J.L., Montel, J.P., Prescott, T., Siliqi, R., Vidal, N., Zhang, X. et al. [2008] Kinematic invariants: an efficient and flexible approach for velocity model building. In: *78th Annual International Meeting, SEG, Expanded Abstracts*, 3687–3692. (cit. on p. 69.)

Guitton, A. [2004] Amplitude and kinematic corrections of migrated images for nonunitary imaging operators. *Geophysics*, **69**, 1017–1024. (cit. on p. 44.)

Guitton, A. [2012] Blocky regularization schemes for Full-Waveform Inversion. *Geophysical Prospecting*, **60**, 870–884. (cit. on p. 161.)

Guitton, A. [2014] On the velocity–density ambiguity in acoustic full–waveform inver-

sion. In: *76th EAGE Conference and Exhibition*, We–E106–03. (cit. on p. 237.)
Guitton, A. [2017] Fast 3D least-squares RTM by preconditioning with nonstationary matching filters. In: *87th Annual International Meeting, SEG, Expanded Abstracts*. 4395–4399. (cit. on pp. 43 and 47.)
Guitton, A., Kaelin, B. and Biondi, B. [2007] Least-squares attenuation of reverse-time-migration artifacts. *Geophysics*, **72**(1), S19–S23. (cit. on pp. 68 and 274.)
Guo, B. and Schuster, G.T. [2017] Wave-equation migration velocity analysis using plane-wave common-image gathers. *Geophysics*, **82**(5), S327–S340. (cit. on p. 59.)
Guo, P. and McMechan, G.A. [2018] Compensating Q effects in viscoelastic media by adjoint-based least-squares reverse time migration. *Geophysics*, **83**(2), S151–S172. (cit. on pp. 22 and 195.)
Hagen, V.S., Arntsen, B. and Raknes, E.B. [2018] A new method for constructing the Hessian in elastic full waveform inversion. In: *88th Annual International Meeting, SEG, Expanded Abstracts*, 1359–1363. (cit. on p. 227.)
Hampson, D. [1986] Inverse velocity stacking for multiple elimination. In: *56th Annual International Meeting, SEG, Expanded Abstracts*. 422–424. (cit. on p. 110.)
Hardage, B.A. [1985] Vertical seismic profiling. *The Leading Edge*, **4**(11), 59–59. (cit. on p. 4.)
He, W., Brossier, R., Métivier, L. and Plessix, R.É. [2019] Land seismic multiparameter full waveform inversion in elastic VTI media by simultaneously interpreting body waves and surface waves with an optimal transport based objective function. *Geophysical Journal International*, **219**(3), 1970–1988. (cit. on p. 272.)
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**, 1004–1024. (cit. on p. 117.)
He, W., Plessix, R.É. and Singh, S. [2018] Parametrization study of the land multiparameter VTI elastic waveform inversion. *Geophysical Journal International*, **213**(3), 1660–1672. (cit. on p. 11.)
Herrmann, F.J., Brown, C.R., Erlangga, Y.A. and Moghaddam, P.P. [2009] Curvelet-based migration preconditioning and scaling. *Geophysics*, **74**(4), A41–A46. (cit. on pp. 22, 116, 160, and 195.)
Herrmann, F.J. and Li, X. [2012] Efficient least–squares imaging with sparsity promotion and compressive sensing. *Geophysical prospecting*, **60**, 696–712. (cit. on pp. 161, 195, and 237.)
Hicks, G.J. and Pratt, R.G. [2001] Reflection waveform inversion using local descent methods: estimating attenuation and velocity over a gas-sand deposit. *Geophysics*, **66**, 598–612. (cit. on pp. 44 and 69.)
de Hoop, M.V. and Bleistein, N. [1997] Generalized Radon transform inversions for reflectivity in anisotropic elastic media. *Inverse Problems*, **13**(3), 669–690. (cit. on p. 23.)
de Hoop, M.V. and Stolk, C. [2006] Seismic inverse scattering in the downward contin-

uation approach. *Wave Motion*, **43**, 579–598. (cit. on p. 41.)

Hou, J. and Symes, W.W. [2015] An approximate inverse to the extended Born modeling operator. *Geophysics*, **80**(6), R331–R349. (cit. on pp. 18, 22, 23, 24, 44, 45, 68, 69, 78, 79, 80, 116, 120, 121, 142, 160, 163, 195, 200, 202, 230, 231, 237, 247, 249, and 264.)

Hou, J. and Symes, W.W. [2016a] Accelerating extended least–squares migration with weighted conjugate gradient iteration. *Geophysics*, **81**(4), S165–S179. (cit. on pp. 23, 45, 47, 116, 151, 152, 160, 196, 215, 230, 231, and 237.)

Hou, J. and Symes, W.W. [2016b] Approximate Gauss-Newton iteration for full-waveform inversion. In: *86th Annual International Meeting, SEG, Expanded Abstracts*, 1163–1168. (cit. on pp. 23, 45, 66, 117, 151, 196, and 237.)

Hou, J. and Symes, W.W. [2016c] Inversion velocity analysis via approximate born inversion. In: *86th Annual International Meeting, SEG, Expanded Abstracts*, 5274–5279. (cit. on p. 23.)

Hou, J. and Symes, W.W. [2017] An alternative formula for approximate extended Born inversion. *Geophysics*, **82**(1), S1–S8. (cit. on pp. 18, 22, 23, 24, 44, 45, 68, 69, 78, 79, 83, 116, 142, 160, 163, 195, 202, and 237.)

Hou, J. and Symes, W.W. [2018] Inversion velocity analysis in the subsurface–offset domain. *Geophysics*, **83**(2), R189–R200. (cit. on pp. 21, 46, and 78.)

Hu, J., Fomel, S., Demanet, L. and Ying, L. [2013] A fast butterfly algorithm for generalized Radon transforms. *Geophysics*, **78**(4), U41–U51. (cit. on p. 110.)

Hu, S., Wang, X., Liu, W., McMechan, G. and Guo, P. [2019] On frequency and resolution of viscoelastic least-squares reverse time migration. In: *81th EAGE Conference and Exhibition, Extended Abstracts*. Tu–R01–13. (cit. on p. 22.)

Hu, W., Abubakar, A., Habashy, T. and Liu, J. [2011] Preconditioned non-linear conjugate gradient method for frequency domain full-waveform seismic inversion. *Geophysical Prospecting*, **59**(3), 477–491. (cit. on p. 10.)

Huang, G., Nammour, R. and Symes, W.W. [2018a] Source-independent extended waveform inversion based on space-time source extension: Frequency-domain implementation. *Geophysics*, **83**(5), R449–R461. (cit. on p. 17.)

Huang, G., Nammour, R. and Symes, W.W. [2018b] Volume source-based extended waveform inversion. *Geophysics*, **83**(5), R369–387. (cit. on p. 17.)

Huang, Y. [2016] *Born waveform inversion in shot coordinate domain*. Ph.D. thesis, Rice University. (cit. on pp. 37, 44, 47, and 57.)

Huang, Y., Nammour, R. and Symes, W.W. [2016] Flexibly preconditioned extended least–squares migration in shot–record domain. *Geophysics*, **81**(5), S299–S315. (cit. on pp. 22 and 160.)

Hyndman, R. and Spence, G. [1992] A seismic study of methane hydrate marine bottom simulating reflectors. *Journal of Geophysical Research: Solid Earth*, **97**(B5), 6683–6698. (cit. on pp. 238 and 246.)

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. 8.)

Jin, S. and Beydoun, W. [2000] 2D multiscale non-linear velocity estimation. *Geophysical Prospecting*, **48**, 163–180. (cit. on p. 63.)

Jin, S. and Madariaga, R. [1993] Background velocity inversion with a genetic algorithm. *Geophysical research Letters*, **20**(2), 93–96. (cit. on p. 63.)

Jin, S. and Madariaga, R. [1994] Nonlinear velocity inversion by a two-step Monte Carlo method. *Geophysics*, **59**(4), 577–590. (cit. on pp. 9 and 63.)

Jin, S., Madariaga, R., Virieux, J. and Lambaré, G. [1992] Two-dimensional asymptotic iterative elastic inversion. *Geophysical Journal International*, **108**, 575–588. (cit. on p. 44.)

Joncour, F., Lambaré, G. and Svay-Lucas, J. [2011] Preserved-amplitude angle domain migration by shot-receiver wavefield continuation. *Geophysical Prospecting*, **59**, 256–268. (cit. on p. 44.)

Jun, H., Park, E. and Shin, C. [2015] Weighted pseudo-Hessian for frequency-domain elastic full waveform inversion. *Journal of Applied Geophysics*, **123**, 1–17. (cit. on pp. 227 and 228.)

Keho, T. and Beydoun, W. [1988] Paraxial ray Kirchhoff migration. *Geophysics*, **53**(12), 1540–1546. (cit. on p. 18.)

Kelly, K., Ward, R., Treitel, S. and Alford, R. [1976] Synthetic seismograms: A finite-difference approach. *Geophysics*, **41**(1), 2–27. (cit. on p. 6.)

Kern, M. and Symes, W.. [1994] Inversion of reflection seismograms by differential semblance analysis: Algorithm structure and synthetic examples. *Geophysical Prospecting*, **99**, 565–614. (cit. on pp. 35 and 37.)

Kim, Y., Shin, C., Calandra, H. and Min, D.J. [2013] An algorithm for 3D acoustic time-Laplace-Fourier-domain hybrid full waveform inversion. *Geophysics*, **78**(4), R151–R166. (cit. on p. 17.)

Kiyashchenko, D., Plessix, R.E., Kashtan, B. and Troyan, V. [2007] A modified imaging principle for true-amplitude wave-equation migration. *Geophysical Journal International*, **168**(3), 1093–1104. (cit. on p. 43.)

Köhn, D., De Nil, D., Kurzmann, A., Przebindowska, A. and Bohlen, T. [2012] On the influence of model parametrization in elastic full waveform tomography. *Geophysical Journal International*, **191**, 325–345. (cit. on pp. 54, 152, and 195.)

Krebs, J.R., Anderson, J.E., Kinkley, D., Neelamani, R., Lee, S., Baumstein, A. and Lacasse, M.D. [2009] Fast full-wavefield seismic inversion using encoded sources. *Geophysics*, **74**(6), WCC177–WCC188. (cit. on p. 60.)

ten Kroode, F. [2012] A wave–equation-based Kirchhoff operator. *Inverse Problems*, **28**(11), 115013. (cit. on pp. 18, 22, 23, 24, 44, 45, 68, 69, 78, 116, 119, 120, 121, 142, 160, 163, 165, 195, 197, 202, 237, 247, and 249.)

Kühl, H. and Sacchi, M.D. [2003] Least-squares wave-equation migration for AVP/AVA

inversion. *Geophysics*, **68**(1), 262–273. (cit. on pp. 20 and 160.)

Lafond, C.F. and Levander, A.R. [1993] Migration moveout analysis and depth focusing. *Geophysics*, **58**(1), 582–598. (cit. on p. 35.)

Lailly, P. [1983] The seismic inverse problem as a sequence of before stack migrations: Conference on Inverse Scattering, Theory and Application. In: *Society for Industrial and Applied Mathematics, Expanded Abstracts*. 206–220. (cit. on pp. 12, 13, 14, 18, 21, 58, 83, 163, and 195.)

Lambaré, G. [2008] Stereotomography. *Geophysics*, **73**(5), VE25–VE34. (cit. on pp. 38 and 69.)

Lambaré, G., Guillaume, P. and Montel, J. [2014] Recent advances in ray-based tomography. In: *76th EAGE Conference and Exhibition, Extended Abstracts*. We G103 01. (cit. on pp. xiii, 8, and 9.)

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 pp. 23 and 44.)

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, and 44.)

Lameloise, C., Chauris, H. and Donno, D. [2013] Migration Velocity Analysis with reflected and transmitted waves. In: *75th EAGE Conference and Exhibition, Extended Abstracts*. We–P01–01. (cit. on p. 54.)

Lameloise, C.A. and Chauris, H. [2016] Extension of Migration velocity Analysis to transmitted wave fields. *Geophysical Journal International*, **207**(1), 343–356. (cit. on pp. xv, 54, 55, and 56.)

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. 35 and 45.)

LeBras, R. and Clayton, R.W. [1988] An iterative inversion of back–scattered acoustic waves. *Geophysics*, **53**(4), 501–508. (cit. on pp. 20 and 163.)

Lecomte, I. and Gelius, L.J. [1998] Have a look at the resolution of prestack depth migration for any model, survey and wavefields. In: *68th Annual International Meeting, SEG, Expanded Abstracts*. 1112–1115. (cit. on p. 43.)

Lecomte, I., Lavadera, P.L., Anell, I., Buckley, S.J., Schmid, D.W. and Heeremans, M. [2015] Ray-based seismic modeling of geologic models: Understanding and analyzing seismic images efficiently. *Interpretation*, **3**(4), SAC71–SAC89. (cit. on p. 7.)

van Leeuwen, T., Aravkin, A.Y. and Herrmann, F.J. [2011] Seismic waveform inversion by stochastic optimization. *International Journal of Geophysics*, **2011**. (cit. on p. 195.)

van Leeuwen, T. and Herrmann, F.J. [2013] Mitigating local minima in full-waveform inversion by expanding the search space. *Geophysical Journal International*, **195**(1), 661–667. (cit. on p. 17.)

van Leeuwen, T. and Herrmann, F.J. [2015] A penalty method for PDE-constrained optimization in inverse problems. *Inverse Problems*, **32**(1), 015007. (cit. on p. 17.)

van Leeuwen, T., Kumar, R. and Herrmann, F. [2015] Affordable Full subsurface image volume–An Application to WEMVA. In: *77th Annual International Conference and Exhibition, EAGE, Extended Abstracts*. WS05–C01. (cit. on pp. 59, 110, and 230.)

Levander, A.R. [1988] Fourth-order finite-difference P-SV seismograms. *Geophysics*, **53**(11), 1425–1436. (cit. on p. 6.)

Li, J. and Symes, W.W. [2007] Interval velocity estimation via NMO-based differential semblance. *Geophysics*, **72**, U75–U88. (cit. on p. 35.)

Li, V., Tsvankin, I. and Alkhalifah, T. [2019] Image-domain wavefield tomography for VTI media. *Geophysics*, **84**(2), C119–C128. (cit. on p. 36.)

Li, V., Wang, H., Tsvankin, I., Díaz, E. and Alkhalifah, T. [2017] Inversion gradients for acoustic VTI wavefield tomography. *Geophysics*, **82**(4), WA55–WA65. (cit. on p. 49.)

Li, Y., Biondi, B., Clapp, R. and Nichols, D. [2014] Wave-equation migration velocity analysis for VTI models. *Geophysics*, **79**(3), WA59–WA68. (cit. on p. 35.)

Li, Y., Biondi, B., Clapp, R. and Nichols, D. [2016] Integrated VTI model building with seismic data, geologic information, and rock-physics modeling – Part 2: Field data test. *Geophysics*, **81**(5), C205–C218. (cit. on pp. 36, 51, 52, and 60.)

Li, Y. and Chauris, H. [2018] Coupling direct inversion to common–shot image–domain velocity analysis. *Geophysics*, **83**(5), R497–R514. (cit. on pp. 21, 22, 23, 61, 62, 64, 78, 117, 160, and 196.)

Lin, Y. and Lianjie, L. [2015] Least–squares reverse–time migration with modified total–variation regularization. In: *85th Annual International Meeting, SEG, Expanded Abstracts*, 4264–4269. (cit. on pp. 161, 195, and 237.)

Lines, L. and Treitel, S. [1984] A review of least-squares inversion and its application to geophysical problems. *Geophysical prospecting*, **32**(2), 159–186. (cit. on p. 10.)

Liu, F., Hanson, D.W., Whitmore, N.D., Day, R.S. and Stolt, R.H. [2006] Toward a unified analysis for source plane–wave migration. *Geophysics*, **71**(4), S129–S139. (cit. on pp. 59, 160, and 195.)

Liu, Y., Chang, X., Jin, D., He, R., Sun, H. and Zheng, Y. [2011] Reverse Time Migration of Multiples for Subsalt Imaging. *Geophysics*, **76**(5), WB209–WB216. (cit. on p. 57.)

Liu, Y., Symes, W.W. and Li, Z. [2014] Inversion velocity analysis via Differential Semblance Optimization. In: *76th EAGE Conference and Exhibition, Extended Abstracts*. Th–P02–07. (cit. on pp. 44 and 47.)

Liu, Z. [1997] An analytical approach to migration velocity analysis. *Geophysics*, **62**, 1238–1249. (cit. on p. 35.)

Liu, Z. and Bleistein, N. [1995] Migration velocity analysis: Theory and an iterative algorithm. *Geophysics*, **60**, 142–153. (cit. on p. 35.)

Long, A., Lu, S., Whitmore, D., Gleut, H.L., Jones, R., Chemingui, N. and Farouki,

M. [2013] Mitigation of the 3D Cross-Line Acquisition Footprint Using Separated Wavefield Imaging of Dual-Sensor Streamer Seismic. In: *75th EAGE Conference and Exhibition, Extended Abstracts*. Th–01–05. (cit. on p. 57.)

Luo, S. and Hale, D. [2014] Least-squares migration in the presence of velocity errors. *Geophysics*, **79**(4), S153–S161. (cit. on p. 21.)

Luo, S. and Sava, P. [2011] A deconvolution-based objective function for wave-equation inversion. In: *81th Annual International Meeting, SEG, Expanded Abstracts*. 2788–2792. (cit. on p. 17.)

Luu, K., Noble, M., Gesret, A., Belayouni, N. and Roux, P. [2017] Propagation of Velocity Uncertainties to Microseismic Locations Using a Competitive Particle Swarm Optimizer. In: *79th EAGE Conference and Exhibition, Extended Abstracts*. We–B4–10. (cit. on p. 9.)

Ma, Y. and Hale, D. [2012] Quasi-Newton full-waveform inversion with a projected Hessian matrix. *Geophysics*, **77**(5), R207–R216. (cit. on p. 10.)

Mallick, S. [2001] AVO and elastic impedance. *The leading edge*, **20**, 1094–1104. (cit. on pp. 151 and 226.)

Marjanović, M., Plessix, R.É., Stopin, A. and Singh, S.C. [2019] Elastic versus acoustic 3-D full waveform inversion at the east pacific rise 9 50' N. *Geophysical Journal International*, **216**(3), 1497–1506. (cit. on p. 122.)

Martin, G.S., Wiley, R. and Marfurt, K.J. [2006] Marmousi2: An elastic upgrade for Marmousi. *The leading edge*, **25**, 156–166. (cit. on pp. 142 and 219.)

Mazzotti, A., Bienati, N., Stucchi, E., Tognarelli, A., Aleardi, M. and Sajeva, A. [2016] Two-grid genetic algorithm full-waveform inversion. *The Leading Edge*, **35**(12), 1068–1075. (cit. on p. 9.)

McMechan, G.A. [1983] Migration by extrapolation of time-dependent boundary values. *Geophysical prospecting*, **31**(3), 413–420. (cit. on pp. 19 and 37.)

Meng, Z., Bleistein, N. and Valasek, P. [1999a] 3-D analytical migration velocity analysis II: velocity gradient estimation. In: *69th Annual International Meeting, SEG, Expanded Abstracts*. 1731–1734. (cit. on p. 35.)

Meng, Z., Bleistein, N. and Wyatt, K. [1999b] 3-D analytical migration velocity analysis I: two-step velocity estimation by reflector-normal update. In: *69th Annual International Meeting, SEG, Expanded Abstracts*. 1727–1730. (cit. on p. 35.)

Menke, W. [1984] *Geophysical data analysis: Discrete inverse theory*. Academic press. (cit. on p. 7.)

Messud, J., Poncet, R. and Lambaré, G. [2021] Optimal transport in full-waveform inversion: Analysis and practice of the multidimensional Kantorovich-Rubinstein norm. *Inverse Problems*. (cit. on p. 17.)

Métivier, L., Brossier, R., Mérigot, Q., Oudet, E. and Virieux, J. [2016a] Increasing the robustness and applicability of full-waveform inversion: An optimal transport distance strategy. *The Leading Edge*, **35**(12), 1060–1067. (cit. on p. 17.)

Métivier, L., Brossier, R., Mérigot, Q., Oudet, E. and Virieux, J. [2016b] Measuring the

misfit between seismograms using an optimal transport distance: application to full waveform inversion. *Geophysical Journal International*, **205**(1), 345–377. (cit. on p. 17.)

Métivier, L., Brossier, R., Mérigot, Q., Oudet, E. and Virieux, J. [2016c] An optimal transport approach for seismic tomography: Application to 3D full waveform inversion. *Inverse Problems*, **32**(11), 115008. (cit. on p. 49.)

Métivier, L., Brossier, R., Operto, S. and Virieux, J. [2014] Multi-parameter FWI-an illustration of the Hessian operator role for mitigating trade-off between parameter classes. In: *6th EAGE Saint Petersburg International Conference and Exhibition*, Tu–BC–01. (cit. on pp. 227 and 228.)

Métivier, L., Brossier, R. and Virieux, J. [2015] Combining asymptotic linearized inversion and full waveform inversion. *Geophysical Journal International*, **201**, 1682–1703. (cit. on pp. 23 and 151.)

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. 10 and 63.)

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.)

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. 6.)

Montel, J.P. and Lambaré, G. [2011] RTM and Kirchhoff angle domain common–image gathers for migration velocity analysis. In: *81th Annual International Meeting, SEG, Expanded Abstracts*. 3120–3124. (cit. on p. 88.)

Montel, J.P. and Lambaré, G. [2019] Kinematics of common-image gathers – Part 1: Theory. *Geophysics*, **84**(5), S437–S447. (cit. on pp. 38, 39, 48, and 69.)

Montel, J.P. and Lambaré, G. [2019] Kinematics of common-image gathers – Part 2: Tomographic ray tracing and applications. *Geophysics*, **84**(5), S501–S510. (cit. on p. 38.)

Mora, P. [1987] Nonlinear two-dimensional elastic inversion of multioffset seismic data. *Geophysics*, **52**(9), 1211–1228. (cit. on pp. 10 and 11.)

Mora, P. [1989] Inversion = migration + tomography. *Geophysics*, **54**(12), 78–101. (cit. on pp. 8 and 62.)

Mora, P. and Wu, Z. [2018] Elastic versus acoustic inversion for marine surveys. *Geophysical Journal International*, **214**, 596–622. (cit. on pp. 117 and 128.)

Mosegaard, K. and Vestergaard, P.D. [1991] A aimulated annealing approach to seismic model optimization with sparse prior information. *Geophysical Prospecting*, **39**(5), 599–611. (cit. on p. 9.)

Mosher, C.C., Keho, T.H., Weglein, A.B. and Foster, D.J. [1996] The impact of migration on AVO. *Geophysics*, **61**(6), 1603–1615. (cit. on p. 151.)

Mulder, W.A. [2014] Subsurface offset behaviour in velocity analysis with extended re-

flectivity images. *Geophysical Prospecting*, **62**(1), 17–33. (cit. on pp. 35, 39, and 43.)
Mulder, W.A. and ten Kroode, A.P.E. [2002] Automatic velocity analysis by Differential Semblance Optimization. *Geophysics*, **67**(4), 1184–1191. (cit. on pp. 35 and 48.)
Mulder, W.A. and van Leeuwen, T. [2008] Automatic migration velocity analysis and multiples. In: *78th Annual International Meeting, SEG, Expanded Abstracts*. 3128–3132. (cit. on p. 56.)
Mulder, W.A. and Plessix, R.É. [2004] A comparison between one-way and two-way wave-equation migration. *Geophysics*, **69**(6), 1491–1504. (cit. on pp. 19, 116, and 195.)
Neidell, N.S. and Taner, M.T. [1971] Semblance and other coherency measures for multichannel data. *Geophysics*, **36**(3), 482–497. (cit. on pp. 240 and 241.)
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, 21, 44, 77, 116, 120, 159, 160, 163, 187, 195, 199, 200, and 237.)
Nichols, D. [2012] Resolution in seismic inversion- Spectral gap or spectral overlap, which is harder to handle? In: *74th EAGE Conference and Exhibition, Extended Abstracts*. E12. (cit. on p. 8.)
Nikitin, V.V., Andersson, F., Carlsson, M. and Duchkov, A.A. [2017] Fast hyperbolic Radon transform represented as convolutions in log–polar coordinates. *Computers & Geosciences*, **105**, 21–33. (cit. on pp. 110 and 242.)
Nocedal, J. and Wright, S. [2006] *Numerical optimization*. Springer Science & Business Media. (cit. on pp. 9, 10, 200, and 201.)
Nolan, C. and Symes, W.W. [1997] Global solution of a linearized inverse problem for the wave equation. In: *Expanded Abstracts*. 919–952. (cit. on p. 41.)
Nouzé, H., Henry, P., Noble, M., Martin, V. and Pascal, G. [2004] Large gas hydrate accumulations on the eastern Nankai Trough inferred from new high-resolution 2-D seismic data. *Geophysical Research Letters*, **31**(13). (cit. on pp. xxiv, 238, 239, and 244.)
Ojha, M. and Sain, K. [2007] Seismic velocities and quantification of gas–hydrates from AVA modeling in the western continental margin of India. *Marine Geophysical Researches*, **28**(2), 101–107. (cit. on pp. 238, 256, and 257.)
Ojha, M. and Sain, K. [2009] Seismic attributes for identifying gas-hydrates and free-gas zones: application to the Makran accretionary prism. *Episodes*, **32**(4), 264–270. (cit. on p. 238.)
Ojha, M. and Sain, K. [2014] Velocity–porosity and velocity–density relationship for shallow sediments in the Kerala-Konkan basin of western Indian margin. *Journal of the Geological Society of India*, **84**(2), 187–191. (cit. on p. 253.)
Operto, S., Gholami, Y., Prieux, V., Ribodetti, A., Brossier, R., Métivier, 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 pp. 7, 11, 15, 81, 116, 161, 227, and 228.)

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. 23.)

Operto, S. and Miniussi, A. [2018] On the role of density and attenuation in three-dimensional multiparameter viscoacoustic VTI frequency-domain FWI: an OBC case study from the North Sea. *Geophysical Journal International*, **213**, 2037–2059. (cit. on p. 265.)

Operto, S., Virieux, J., Amestoy, P., L'Excellent, J.Y., Giraud, L. and Ali, H.B.H. [2007] 3D finite-difference frequency-domain modeling of visco-acoustic wave propagation using a massively parallel direct solver: A feasibility study. *Geophysics*, **72**(5), SM195–SM211. (cit. on p. 6.)

Operto, S., Virieux, J., Ribodetti, A. and Anderson, J.E. [2009] Finite-difference frequency-domain modeling of viscoacoustic wave propagation in 2D tilted transversely isotropic (TTI) media. *Geophysics*, **74**(5), T75–T95. (cit. on p. 11.)

Operto, S., Xu, S. and Lambaré, G. [2000] Can we quantitatively image complex structures with rays? *Geophysics*, **65**(4), 1223–1238. (cit. on p. 44.)

Pan, W., Geng, Y. and Innanen, K.A. [2018] Interparameter trade-off quantification and reduction in isotropic–elastic full-waveform inversion: synthetic experiments and Hussar land data set application. *Geophysical Journal International*, **213**, 1305–1333. (cit. on p. 144.)

Plessix, R.É. [2006a] Estimation of velocity and attenuation coefficient maps from crosswell seismic data. *Geophysics*, **71**(6), S235–S240. (cit. on pp. 40 and 64.)

Plessix, R.É. [2006b] A review of the adjoint-state method for computing the gradient of a functional with geophysical applications. *Geophysical Journal Inernational*, **167**(2), 495–503. (cit. on pp. 10, 28, and 44.)

Plessix, R.É. and Cao, Q. [2011] A parametrization study for surface seismic full waveform inversion in an acoustic vertical transversely isotropic medium. *Geophysical Journal International*, **185**, 539–556. (cit. on pp. 144 and 264.)

Plessix, R.É., Chavent, G. and De Roeck, Y. [1995] Automatic and simultaneous migration velocity analysis and waveform inversion of real data using a MBTT/WBKBJ formulation. In: *65th Annual International Meeting, SEG, Expanded Abstracts*. 1099–1101. (cit. on pp. 58 and 69.)

Plessix, R.É., Milcik, P., Rynja, H., Stopin, A., Matson, K. and Abri, S. [2013] Multiparameter full–waveform inversion: Marine and land examples. *The Leading Edge*, **32**(9), 1030–1038. (cit. on p. 10.)

Plessix, R.É. and Mulder, W.A. [2004] Frequency-domain finite-difference amplitude-preserving migration. *Geophysical Journal International*, **157**, 975–987. (cit. on pp. 44 and 63.)

Plessix, R.É., Mulder, W.A. and ten Kroode, A.P.E. [2000] Automatic cross-well tomography by semblance and differential semblance optimization: theory and gradient computation. *Geophysical Prospecting*, **48**, 913–935. (cit. on pp. 40 and 48.)

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. 6.)

Polson, N.G. and Sokolov, V. [2019] Bayesian regularization: From Tikhonov to horseshoe. *Wiley Interdisciplinary Reviews: Computational Statistics*, **11**, e1463. (cit. on pp. 161 and 182.)

Pratt, R.G. [1999] Seismic waveform inversion in the frequency domain, Part 1: Theory and verification in a physical scale model. *Geophysics*, **64**(3), 888–901. (cit. on pp. 14 and 247.)

Pratt, R.G., Shin, C. and Hick, G.J. [1998] Gauss–Newton and full Newton methods in frequency–space seismic waveform inversion. *Geophysical Journal International*, **133**(2), 341–362. (cit. on pp. 10, 14, 63, and 230.)

Pratt, R.G. and Shipp, R.M. [1999] Seismic waveform inversion in the frequency domain, Part 2: Fault delineation in sediments using crosshole data. *Geophysics*, **64**(3), 902–914. (cit. on p. 11.)

Pratt, R.G. and Worthington, M.H. [1990] Inverse theory applied to multi-source cross-hole tomography. *Geophysical Prospecting*, **38**(3), 287–310. (cit. on p. 14.)

Prieux, V. [2012] *Imagerie sismique des milieux visco-acoustiques et visco-élastiques à deux dimensions par stéréotomographie et inversion des formes d'ondes : applications au champ pétrolier de Valhall*. Ph.D. thesis, UNS - Université Nice Sophia Antipolis. (cit. on p. 272.)

Prieux, V., Brossier, R., Gholami, Y., Operto, S., Virieux, J., Barkved, O. and Kommedal, J. [2011] On the footprint of anisotropy on isotropic full waveform inversion: the Valhall case study. *Geophysical Journal International*, **187**(3), 1495–1515. (cit. on p. 11.)

Prieux, V., Brossier, R., Operto, S. and Virieux, J. [2013] 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 pp. 11, 15, 63, 107, 116, and 161.)

Prucha, M., Biondi, B. and Symes, W.W. [1999] Angle-domain common image gathers by wave-equation migration. In: *69th Annual International Meeting, SEG, Expanded Abstracts*. 824–827. (cit. on p. 59.)

Przebindowska, A., Kurzmann, A., Köhn, D. and Bohlen, T. [2012] The role of density in acoustic full waveform inversion of marine reflection seismics. In: *74th EAGE Conference and Exhibition incorporating EUROPEC 2012*, cp–293. (cit. on p. 237.)

Qin, B., Allemand, T. and Lambaré, G. [2015] Full waveform inversion using preserved amplitude reverse time migration. In: *85th Annual International Meeting, SEG, Expanded Abstracts*. 1252–1257. (cit. on p. 44.)

Qin, B. and Lambaré, G. [2016] Joint inversion of velocity and density in preserved–amplitude full–waveform inversion. In: *86th Annual International Meeting, SEG, Expanded Abstracts*. 1325–1330. (cit. on pp. 44, 52, 105, 106, 165, and 237.)

Qu, Y., Huang, J., Li, Z., Guan, Z. and Li, J. [2017] Attenuation compensation in anisotropic least-squares reverse time migration. *Geophysics*, **82**(6), S411–S423. (cit. on pp. 116, 195, and 272.)

Ren, Z. and Li, Z. [2020] Imaging of elastic seismic data by least–squares reverse time migration with weighted L2–norm multiplicative and modified total–variation regularizations. *Geophysical Prospecting*, **68**, 411–430. (cit. on pp. 161 and 189.)

Ren, Z., Liu, Y. and Sen, M.K. [2017] Least-squares reverse time migration in elastic media. *Geophysical Journal International*, **208**, 1103–1125. (cit. on pp. 117, 146, 152, and 195.)

Ribodetti, A. and Virieux, J. [2000] Asymptotic theory for imaging the attenuation factor Q. *Geophysics*, **63**(5), 1767–1778. (cit. on p. 15.)

Rickett, J.E. [2003] Illumination-based normalization for wave-equation depth migration. *Geophysics*, **68**(4), 1371–1379. (cit. on pp. 22, 44, 116, 160, and 195.)

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. 21, 37, and 59.)

Rocha, D. and Sava, P. [2018] Elastic least-squares reverse time migration using the energy norm. *Geophysics*, **83**(3), S237–S248. (cit. on p. 22.)

Rodi, W. and Mackie, R.L. [2001] Nonlinear conjugate gradients algorithm for 2-D magnetotelluric inversion. *Geophysics*, **66**(1), 174–187. (cit. on p. 10.)

Romero, L.A., Ghiglia, D.C., Ober, C.C. and Morton, S.A. [2000] Phase encoding of shot records in prestack migration. *Geophysics*, **65**(2), 426–436. (cit. on pp. 60, 195, and 237.)

Russell, B.H., Hedlin, K., Hilterman, F.J. and Lines, L.R. [2003] Fluid-property discrimination with AVO: A Biot-Gassmann perspective. *Geophysics*, **68**, 29–39. (cit. on pp. 116 and 134.)

Sambolian, S., Operto, S., Ribodetti, A. and Virieux, J. [2021] Consistent event relocation and subsurface parameters inversion through slope tomography: a variable-projection approach. *Geophysical Journal International*, **224**(3), 1956–1979. (cit. on p. 69.)

Sattlegger, J. [1975] Migration Velocity determination: part I. Philosophy. *Geophysics*, **40**(1). (cit. on p. 35.)

Sava, P. and Alkhalifah, T. [2013] Wide–azimuth angle gathers for anisotropic wave–equation migration. *Geophysical Prospecting*, **61**(1), 75–91. (cit. on p. 87.)

Sava, P. and Biondi, B. [2004] Wave-equation migration velocity analysis. I. theory. *Geophysical Prospecting*, **52**, 593–606. (cit. on pp. 35 and 37.)

Sava, P. and Fomel, S. [2006] Time-shift imaging condition in seismic migration. *Geophysics*, **71**(6), S209–S217. (cit. on pp. 21, 38, 43, 77, 80, 119, 162, 197, and 247.)

Sava, P. and Vasconcelos, I. [2009] Extended common-image-point gathers for wave-equation migration. In: *71th EAGE Conference and Exhibition, Extended Abstracts*. CP–127–00354. (cit. on p. 61.)

Sava, P. and Vasconcelos, I. [2011] Extended imaging conditions for wave-equation migration. *Geophysical Prospectings*, **59**, 35–55. (cit. on p. 37.)

Sava, P. and Vlad, I. [2011] Wide–azimuth angle gathers for wave–equation migration. *Geophysics*, **76**(3), S131–S141. (cit. on p. 87.)

Sava, P.C., Biondi, B. and Etgen, J. [2005] Wave-equation migration velocity analysis by focusing diffractions and reflections. *Geophysics*, **70**(3), U19–U27. (cit. on pp. 16 and 62.)

Sava, P.C. and Fomel, S. [2003] Angle-domain common-image gathers by wavefield continuation methods. *Geophysics*, **68**(3), 1065–1074. (cit. on pp. 21, 37, 59, 87, 88, 122, 124, 165, 166, 198, and 250.)

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. 7.)

Schleicher, J., Tygel, M. and Hubral, P. [1993] 3-D true-amplitde finite-offset migration. *Geophysics*, **58**(8), 1112–1126. (cit. on p. 44.)

Schneider, W.A. [1978] Integral formulation for migration in two and three dimensions. *Geophysics*, **43**(1), 49–76. (cit. on p. 18.)

Schuster, G.T. [1993] Least-squares cross–well migration. In: *63th Annual International Meeting, SEG, Expanded Abstracts*, 110–113. (cit. on pp. 21, 163, and 200.)

Schuster, G.T. [2011] *Seismic Imaging, Overview*. Springer Netherlands, 1121–1134. (cit. on p. 3.)

Schuster, G.T. [2017] *Seismic Inversion*. Society of Exploration Geophysicists. (cit. on pp. 3 and 7.)

Schuster, G.T., Wang, X., Huang, Y., Dai, W. and Boonyasiriwat, C. [2011] Theory of multisource crosstalk reduction by phase-encoded statics. *Geophysical Journal International*, **184**(3), 1289–1303. (cit. on pp. 60 and 195.)

Sen, M.K. and Stoffa, P.L. [2013] *Global optimization methods in geophysical inversion*. Cambridge University Press. (cit. on p. 9.)

Shabelansky, A.H., Malcolm, A.E., Fehler, M.C., Shang, X. and Rodi, W.L. [2015] Source independent full wavefield converted-phase elastic migra- tion velocity analysis. *Geophysical Journal International*, **200**, 954–968. (cit. on p. 36.)

Shan, G. and Biondi, B. [2005] Imaging overturned waves by plane-wave migration in tilted coordinates. In: *75th Annual International Meeting, SEG, Expanded Abstracts*. 969–972. (cit. on p. 43.)

Shen, P. and Calandra, H. [2005] One-way waveform inversion within the framework of adjoint state differential migration. In: *75th Annual International Meeting, SEG, Expanded Abstracts*. 1709–1712. (cit. on p. 35.)

Shen, P. and Symes, W.W. [2008] Automatic velocity analysis via shot profile migration. *Geophysics*, **73**(5), VE49–VE59. (cit. on p. 35.)

Shen, P. and Symes, W.W. [2015] Horizontal contraction in image domain for velocity inversion. *Geophysics*, **80**(3), R95–R110. (cit. on pp. 49 and 168.)

Shen, P.S. [2013] Subsurface Focusing Measurement of Diving Waves and its Application to Reflection Tomography. In: *75th EAGE Conference and Exhibition, Extended Abstracts*. Th–10–05. (cit. on pp. 54 and 55.)

Shen, Y., Biondi, B. and Clapp, R. [2018] Q-model building using one-way wave-equation migration Q analysis – Part 2: 3D field-data test. *Geophysics*, **83**(2), S111–S126. (cit. on pp. 36 and 51.)

Sheriff, R.E. [2002] *Encyclopedic dictionary of applied geophysics*, 13. Society of Exploration Geophysicists. (cit. on p. 3.)

Sheriff, R.E. and Geldart, L.P. [1995] *Exploration seismology*. Cambridge university press, 2 edn. (cit. on pp. 3, 4, 193, and 244.)

Shi, Y., Zhao, W. and Cao, H. [2007] Nonlinear process control of wave-equation inversion and its application in the detection of gas. *Geophysics*, **72**(1), R9–R18. (cit. on p. 11.)

Shin, C., Jang, S. and Min, D.J. [2001] Improved amplitude preservation for prestack depth migration by inverse scattering theory. *Geophysical Prospecting*, **49**(5), 592–606. (cit. on pp. 22, 116, 160, 195, and 237.)

Shin, C. and Min, D.J. [2006] Waveform inversion using a logarithmic wavefield. *Geophysics*, **71**(3), R31–R42. (cit. on p. 17.)

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. 14.)

Shragge, J. and Lumley, D. [2013] Time-lapse wave-equation migration velocity analysis. *Geophysics*, **78**(2), S69–S79. (cit. on p. 63.)

Singh, S., Snieder, R., Behura, J., van der Neut, J., Wapenaar, K. and Slob, E. [2015] Marchenko Imaging: Imaging with Primaries, Internal Multiples, and Free-Surface Multiples. *Geophysics*, **80**(5), S165–S174. (cit. on p. 57.)

Sirgue, L. [2006] The importance of low frequency and large offset in waveform inversion. In: *68th EAGE Conference and Exhibition incorporating SPE EUROPEC 2006*. A037. (cit. on p. 14.)

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. 14 and 15.)

Smith, W.D. [1975] The application of finite element analysis to body wave propagation problems. *Geophysical Journal International*, **42**(2), 747–768. (cit. on p. 6.)

Solano, C.P., Chang, G., Plessix, R.É., Bao, K., Stopin, A. and Wang, X. [2020a] When do we need elastic waveform inversion for velocity model building? marine and land examples. In: *82th Annual International Conference and Exhibition, EAGE, Workshop Programme*, WS02–08. (cit. on p. 122.)

Solano, C.P., Plessix, R.É., Bao, K., Perkins, C. and Kiehn, M. [2020b] Velocity model building using elastic waveform inversion on multi-Component OBN data in the Gulf of Mexico. In: *82th Annual International Conference and Exhibition, EAGE, Ex-*

tended Abstracts, Fr–Dome1–04. (cit. on p. 117.)
Soubaras, R. and Gratacos, B. [2007] Velocity model building by semblance maximization of modulated-shot gathers. *Geophysics*, **72**(5), U67–U73. (cit. on p. 48.)
Soubaras, R. and Gratacos, B. [2017] Mitigating the gradient artefacts of Migration Velocity Analysis by Gauss-Newton update. In: *79th EAGE Conference and Exhibition, Extended Abstracts*. WS09–D01. (cit. on p. 63.)
Soubaras, R. and Gratacos, B. [2018] Migration Velocity Analysis with Multiple Modeling: an Inversion Toolbox. In: *80th EAGE Conference and Exhibition, Extended Abstracts*. Th–A12–14. (cit. on p. 57.)
Staal, X.R. and Verschuur, D.J. [2012] Velocity estimation using internal multiples. In: *82th Annual International Meeting, SEG, Expanded Abstracts*. 1–5. (cit. on p. 57.)
Stanton, A. and Sacchi, M.D. [2017] Elastic least-squares one-way wave-equation migration. *Geophysics*, **82**(4), S293–S305. (cit. on p. 22.)
Stolk, C.C. and Symes, W.W. [2003] Smooth objective functionals for seismic velocity inversion. *Wave Motion*, **32**, 267–290. (cit. on pp. 35 and 48.)
Stolk, C.C. and Symes, W.W. [2004] Kinematic artifacts in prestack depth migration. *Geophysics*, **69**(2), 562–575. (cit. on pp. 38, 41, and 68.)
Stuart, G.K., Minkoff, S.E. and Pereira, F. [2019] A two-stage Markov chain Monte Carlo method for seismic inversion and uncertainty quantification. *Geophysics*, **84**(6), R1003–R1020. (cit. on p. 9.)
Sun, H. and Schuster, G.T. [2001] 2-D wavepath migration. *Geophysics*, **66**(5), 1528–1537. (cit. on p. 35.)
Sun, J., Fomel, S., Zhu, T. and Hu, J. [2016] Q–compensated least–squares reverse time migration using low–rank one–step wave extrapolation. *Geophysics*, **81**(4), S271–S279. (cit. on p. 195.)
Sun, M., Dong, L., Yang, J., Huang, C. and Liu, Y. [2018] Elastic least–squares reverse time migration with density variations. *Geophysics*, **83**(6), S533–S547. (cit. on pp. 78, 116, 153, and 237.)
Symes, W.W. [1993] A differential semblance criterion for inversion of multioffset seismic reflection data. *Journal of Geophysical Research*, **98**, 2061–2073. (cit. on p. 35.)
Symes, W.W. [1995] Mathematics of reflection seismology. *Rice University*, 1–85. (cit. on pp. 18 and 194.)
Symes, W.W. [1998a] All stationary points of differential semblance are asymtotic global minimizers: layered acoustics. Tech. rep., The Rice Inversion Project. (cit. on pp. 35 and 48.)
Symes, W.W. [1998b] High frequency asymptotics, differential semblance, and velocity estimation. In: *68th Annual International Meeting, SEG, Expanded Abstracts*. 1616–1619. (cit. on pp. 35 and 40.)
Symes, W.W. [2008a] Approximate linearized inversion by optimal scaling of prestack depth migration. *Geophysics*, **73**(2), R23–R35. (cit. on pp. 116 and 195.)
Symes, W.W. [2008b] Migration velocity analysis and waveform inversion. *Geophysical*

Prospecting, **56**(6), 765–790. (cit. on pp. 16, 17, 18, 21, 34, 35, 37, 39, 57, 77, 80, 81, 84, 116, 119, 160, 162, 197, 215, and 248.)

Symes, W.W. [2009] The seismic reflection inverse problem. *Inverse Problems*, **25**, 123008. (cit. on pp. 35 and 37.)

Symes, W.W. and Carazzone, J.J. [1991] Velocity inversion by Differential Semblance Optimization. *Geophysics*, **56**, 654–663. (cit. on p. 48.)

Tang, Y. [2009] Target-oriented wave-equation least-squares migration/inversion with phase-encoded Hessian. *Gephysics*, **74**(6), WCA95–WCA107. (cit. on p. 63.)

Tang, Y. and Biondi, B. [2011] Target-oriented wavefield tomography using synthesized Born data. *Geophysics*, **76**(5), WB191–WB207. (cit. on pp. 44 and 60.)

Tang, Y. and Biondi, B. [2013] Subsalt velocity estimation by target-oriented wave-equation migration velocity analysis: A 3D field-data example. *Geophysics*, **78**(1), U19–U29. (cit. on pp. 36 and 60.)

Tang, Y., Sun, B. and Alkhalifah, T. [2020] Wave-equation migration velocity analysis via the optimal-transport-based objective function. In: *90th Annual International Meeting, SEG, Expanded Abstracts*. 3709–3713. (cit. on pp. 48 and 49.)

Tarantola, A. [1984] Inversion of seismic reflection data in the acoustic approximation. *Geophysics*, **49**(8), 1259–1266. (cit. on pp. 7, 10, 12, 13, 14, 18, 21, 58, 82, 83, 107, 120, 163, 200, and 236.)

Tarantola, A. [1986] A strategy for nonlinear elastic inversion of seismic reflection data. *Geophysics*, **51**(10), 1893–1903. (cit. on pp. 11, 15, 122, 152, 195, 198, and 250.)

Tarantola, A. [2005] *Inverse problem theory and methods for model parameter estimation*, 89. SIAM. (cit. on pp. 3, 7, and 189.)

Tavakoli, B., Operto, S., Ribodetti, A. and Virieux, J. [2017] Slope tomography based on eikonal solvers and the adjoint-state method. *Geophysical Journal International*, **209**, 1629–1647. (cit. on p. 69.)

Thiel, N., Hertweck, T. and Bohlen, T. [2019] Comparison of acoustic and elastic full-waveform inversion of 2D towed-streamer data in the presence of salt. *Geophysical prospecting*, **67**(2), 349–361. (cit. on p. 122.)

Thierry, P., Operto, S. and Lambaré, G. [1999] Fast 2D ray-Born inversion/migration in complex media. *Geophysics*, **64**(1), 162–181. (cit. on p. 44.)

Thomsen, L. [1986] Weak elastic anisotropy. *Geophysics*, **51**(10), 1954–1966. (cit. on p. 272.)

Thorson, J.R. and Claerbout, J.F. [1985] Velocity–stack and slant–stack stochastic inversion. *Geophysics*, **50**(12), 2727–2741. (cit. on pp. 110 and 241.)

Tikhonov, A.N., Goncharsky, A., Stepanov, V. and Yagola, A.G. [2013] *Numerical methods for the solution of ill-posed problems*. Springer Science & Business Media. (cit. on p. 161.)

Trad, D., Ulrych, T. and Sacchi, M. [2003] Latest views of the sparse Radon transform. *Geophysics*, **68**(1), 386–399. (cit. on p. 241.)

Tran, K.T. and Hiltunen, D.R. [2012] Two-dimensional inversion of full waveforms us-

ing simulated annealing. *Journal of Geotechnical and Geoenvironmental Engineering*, **138**(9), 1075–1090. (cit. on p. 9.)

Tu, N. and Herrmann, F.J. [2015] Fast Imaging with Surface-Related Multiples by Sparse Inversion. *Geophysical Journal International*, **201**(1), 304–317. (cit. on p. 57.)

Valenciano, A.A., Biondi, B. and Guitton, A. [2006] Target-oriented wave-equation inversion. *Geophysics*, **71**(4), A35–A38. (cit. on pp. 44 and 63.)

Van Leeuwen, T. and Mulder, W.A. [2010] A correlation-based misfit criterion for wave-equation traveltime tomography. *Geophysical Journal International*, **182**(3), 1383–1394. (cit. on pp. 17 and 48.)

Červený, V. (Ed.) [2001] *Seismic ray tracing*. Cambridge University Press. (cit. on p. 41.)

Verschuur, D.J. and Berkhout, A.J. [2011] Seismic Migration of Blended Shot Records with Surface-Related Multiple Scattering. *Geophysics*, **76**(1), A7–A13. (cit. on p. 57.)

Versteeg, R.J. and Grau, G. (Eds.) [1991] *The Marmousi experience*. Proceedings of the 1990 EAEG workshop on Pratical Aspects of Seismic Data Inversion, Eur. Ass. Expl. Geophys. (cit. on p. 43.)

Vidale, J. [1988] Finite-difference calculation of travel times. *Bulletin of the Seismological Society of America*, **78**(6), 2062–2076. (cit. on p. 6.)

Virieux, J. [1986] P-SV wave propagation in heterogeneous media: Velocity-stress finite-difference method. *Geophysics*, **51**(4), 889–901. (cit. on pp. 6, 125, and 202.)

Virieux, J., Asnaashari, A., Brossier, R., Métivier, L., Ribodetti, A. and Zhou, W. [2017] An introduction to full waveform inversion. In: *Encyclopedia of exploration geophysics*, Society of Exploration Geophysicists, R1–1. (cit. on p. 227.)

Virieux, J., Brossier, R., Métivier, L., Operto, S. and Ribodetti, A. [2016] Direct and indirect inversions. *Journal of Seismology*, **20**(4), 1107–1121. (cit. on p. 44.)

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 p. 6.)

Virieux, J. and Operto, S. [2009] An overview of full-waveform inversion in exploration geophysics. *Geophysics*, **74**(6), WCC1–WCC26. (cit. on pp. 11, 13, 14, 15, 22, 25, 35, 58, 62, 63, 81, 116, 152, 153, 161, 164, 166, 179, 195, 199, 236, 257, and 264.)

Vyas, M. and Tang, Y. [2010] Gradients for wave-equation migration velocity analysis. In: *80th Annual International Meeting, SEG, Expanded Abstracts*. 4077–4081. (cit. on p. 49.)

Wang, B., Pann, K. and Meek, R.A. [1995] Macro velocity model estimation through model-based globally-optimized residual-curvature analysis. In: *65th Annual International Meeting, SEG, Expanded Abstracts*. 1084–1087. (cit. on p. 35.)

Wang, C., Cheng, J., Weibull, W.W. and Arntsen, B. [2019] Elastic wave-equation migration velocity analysis preconditioned through mode decoupling. *Geophysics*, **84**(3), R341–R353. (cit. on pp. 63 and 68.)

Wang, C., Yingst, D., Farmer, P. and Leveille, J. [2016a] Full-waveform inversion with

the reconstructed wavefield method. In: *SEG Technical Program Expanded Abstracts*. 1237–1241. (cit. on p. 17.)

Wang, F., Donno, D., Chauris, H., Calandra, H. and Audbert, F. [2016b] Waveform inversion based on wavefield decomposition. *Geophysics*, **81**(6), R457–R470. (cit. on p. 59.)

Wang, J. and Sacchi, M.D. [2007] High–resolution wave–equation amplitude–variation–with–ray–parameter (AVP) imaging with sparseness constraints. *Geophysics*, **72**(1), S11–S18. (cit. on p. 160.)

Wang, Y. [2008] *Seismic Inverse Q Filtering*. Wiley-Blackwell, Hoboken, New Jersey. (cit. on p. 11.)

Wapenaar, C.P.A. [1996] Inversion versus migration: A new perspective to an old discussion. *Geophysics*, **61**(3), 804–814. (cit. on p. 8.)

Wapenaar, K. and van der Neut, J. [2010] A representation for Greens function retrieval by multidimensional deconvolution. *Journal of the Acoustical Society of America*, **128**, EL366–EL371. (cit. on p. 43.)

Wapenaar, K., Thorbecke, J., van der Neut, J., Broggini, F., Slob, E. and Snieder, R. [2014] Marchenko Imaging. *Geophysics*, **79**(3), WA39–WA57. (cit. on pp. 57 and 62.)

Weibull, W.W. and Arntsen, B. [2013] Automatic velocity analysis with reverse-time migration. *Geophysics*, **78**(4), S179–S192. (cit. on pp. 35, 40, 64, and 68.)

Weibull, W.W. and Arntsen, B. [2014] Anisotropic migration velocity analysis using reverse-time migration. *Geophysics*, **79**(1), R13–R25. (cit. on pp. 35 and 63.)

Whitmore, N.D. [1983] Iterative depth migration by backward time propagation. In: *SEG Technical Program Expanded Abstracts*. 382–385. (cit. on p. 37.)

Witten, B. and Shragge, J. [2017] Microseismic image-domain velocity inversion: Marcellus Shale case study. *Geophysics*, **82**(6), KS99–KS112. (cit. on p. 36.)

Wong, M., Biondi, B. and Ronen, S. [2014] Imaging with Multiples Using Least-Squares Reverse Time Migration. *The Leading Edge*, **33**(9), 970–976. (cit. on p. 57.)

Woodward, M.J. [1992] Wave–equation tomography. *Geophysics*, **57**(1), 15–26. (cit. on p. 35.)

Woodward, M.J., Nichols, D., Zdraveva, O., Whitfield, P. and Johns, T. [2008] A decade of tomography. *Geophysics*, **73**(5), VE5–VE11. (cit. on p. 35.)

Wu, D., Yao, G., Cao, J. and Wang, Y. [2016] Least–squares RTM with L1 norm regularisation. *Journal of Geophysics and Engineering*, **13**, 666–673. (cit. on pp. 161 and 167.)

Wu, S., Wang, Y., Zheng, Y. and Chang, X. [2015] Limited-memory BFGS based least-squares pre-stack Kirchhoff depth migration. *Geophysical Journal International*, **202**(2), 738–747. (cit. on p. 10.)

Xu, S., Chauris, H., Lambaré, G. and Noble, M. [2001] Common angle image gather: a strategy for imaging complex media. *Geophysics*, **66**(6), 1877–1894. (cit. on p. 37.)

Xu, S., Wang, D., Chen, F., Lambaré, G. and Zhang, Y. [2012a] Inversion on reflected seismic wave. In: *82th Annual International Meeting, SEG, Expanded Abstracts*. (cit.

on p. 151.)

Xu, S., Wang, D., Chen, F., Zhang, Y. and Lambaré, G. [2012b] Full waveform inversion for reflected seismic data. In: *74th EAGE Conference and Exhibition, Extended Abstracts*. W024. (cit. on pp. 16 and 276.)

Xu, S., Zhang, Y. and Tang, B. [2011] 3D angle gathers from reverse time migration. *Geophysics*, **76**(2), S77–S92. (cit. on p. 59.)

Xue, Z., Chen, Y., Fomel, S. and Sun, J. [2016] Seismic imaging of incomplete data and simultaneous-source data using least-squares reverse time migration with shaping regularization. *Geophysics*, **81**(1), S11–S20. (cit. on pp. 20, 22, 78, 160, 161, and 195.)

Yang, J., Elita Li, Y., Cheng, A., Liu, Y. and Dong, L. [2019a] Least-squares reverse time migration in the presence of velocity errors. *Geophysics*, **84**(6), S567–S580. (cit. on pp. 77, 116, 135, and 214.)

Yang, J., Li, Y.E., Liu, Y. and Zong, J. [2020] Least-squares extended reverse time migration with randomly sampled space shifts. *Geophysics*, **85**(6), S357–S369. (cit. on p. 277.)

Yang, J., Liu, Y. and Dong, L. [2016a] Least–squares reverse time migration in the presence of density variations. *Geophysics*, **81**(6), S497–S509. (cit. on pp. 22, 102, 116, 195, and 237.)

Yang, J., Liu, Y. and Dong, L. [2016b] Simultaneous estimation of velocity and density in acoustic multiparameter full–waveform inversion using an improved scattering–integral approach. *Geophysics*, **81**(6), R399–R415. (cit. on p. 78.)

Yang, J., Zhu, H., McMechan, G., Zhang, H. and Zhao, Y. [2019b] Elastic least-squares reverse time migration in vertical transverse isotropic media. *Geophysics*, **84**(6), S539–S553. (cit. on p. 22.)

Yang, T. and Sava, P. [2011] Wave-equation migration velocity analysis with time-shift imaging. *Geophysical Prospecting*, **59**(4), 635–650. (cit. on pp. 21 and 38.)

Yang, T. and Sava, P. [2013] 3D image-domain wavefield tomography using time-lag extended image. In: *83th Annual International Meeting, SEG, Expanded Abstracts*. 4816–4821. (cit. on p. 36.)

Yang, T. and Sava, P. [2015] Image–domain wavefield tomography with extended common–image–point gathers. *Geophysical Prospecting*, **63**(5), 1086–1096. (cit. on pp. 43, 59, 61, 110, and 277.)

Yang, T., Shragge, J. and Sava, P.C. [2013] Illumination compensation for image-domain wavefield tomography. *Geophysics*, **78**(5), U65–U76. (cit. on pp. 43 and 44.)

Yang, Y. and Engquist, B. [2018] Analysis of optimal transport and related misfit functions in full-waveform inversion. *Geophysics*, **83**(1), A7–A12. (cit. on p. 49.)

Yao, G., da Silva, N.V. and Wu, D. [2019] Reflection–Waveform Inversion Regularized with Structure–Oriented Smoothing Shaping. *Pure and Applied Geophysics*, **176**, 5315–5335. (cit. on p. 161.)

Yao, G., Wu, D. and Wang, S.X. [2020] A review on reflection-waveform inversion.

Petroleum Science, **17**(2), 334–351. (cit. on pp. 69 and 276.)

Yilmaz, Ö. [2001] *Seismic data analysis: Processing, inversion, and interpretation of seismic data*. Society of exploration geophysicists. (cit. on pp. 3 and 37.)

Zand, T. and Siahkoohi, H.R. [2021] Sparsity-promoting least-squares reverse time migration via preconditioned Bregmanized operator splitting. *Geophysical Prospecting*, **69**(4), 727–744. (cit. on p. 271.)

Zand, T., Siahkoohi, H.R., Malcolm, A., Gholami, A. and Richardson, A. [2020] Consensus optimization of total variation–based reverse time migration. *Computational Geociences*, **24**, 1393–1407. (cit. on pp. 161, 195, and 237.)

Zeng, C., Dong, S. and Wang, B. [2014] Least-squares reverse time migration: Inversion-based imaging toward true reflectivity. *The Leading Edge*, **33**(9), 962–968. (cit. on pp. 77, 160, and 195.)

Zhang, D. and Schuster, G.T. [2014] Least-squares reverse time migration of multiples. *Geophysics*, **79**(1), S11–S21. (cit. on pp. 77 and 160.)

Zhang, Q. [2014] RTM angle gathers and Specular Filter (SF) RTM using optical flow. In: *84th Annual International Meeting, SEG, Expanded Abstracts*. 10.1190/segam2014–0792.1. (cit. on p. 59.)

Zhang, S., Luo, Y. and Schuster, G. [2015a] Shot- and angle-domain wave-equation traveltime inversion of reflection data: Synthetic and field data examples. *Geophysics*, **80**(4), S79–S92. (cit. on p. 36.)

Zhang, X., Burger, M., Bresson, X. and Osher, S. [2010] Bregmanized nonlocal regularization for deconvolution and sparse reconstruction. *SIAM Journal on Imaging Sciences*, **3**(3), 253–276. (cit. on p. 271.)

Zhang, X., Dai, N., Li, W., Wang, D. and Wang, F. [2017] A practical implementation of least–square migration. In: *87th Annual International Meeting, SEG, Expanded Abstracts*, 4529–4533. (cit. on p. 167.)

Zhang, Y. and Biondi, B. [2013] Moveout-based wave-equation migration velocity analysis. *Geophysics*, **78**(2). (cit. on pp. 49 and 60.)

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 p. 57.)

Zhang, Y., Duan, L. and Xie, Y. [2015b] A stable and practical implementation of least-squares reverse time migration. *Geophysics*, **80**(1), V23–V31. (cit. on pp. 20, 77, and 160.)

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 pp. 52, 78, 79, 116, 117, 122, 160, 196, 198, and 269.)

Zhang, Y. and Shan, G. [2013] Wave-equation migration velocity analysis using partial stack-power maximization. In: *83th Annual International Meeting, SEG, Expanded Abstracts*. 4847–4852. (cit. on p. 48.)

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. 22 and 44.)

Zhang, Y., Zhang, G. and Bleistein, N. [2003] True amplitude wave equation migration arising from true amplitude one-way wave equations. *Inverse Problems*, **19**(5), 1113–1138. (cit. on p. 22.)

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. 44 and 45.)

Zhang, Y., Zhang, H. and Zhang, G. [2011] A stable TTI reverse time migration and its implementation. *Geophysics*, **76**(3), WA3–WA11. (cit. on p. 22.)

Zhang, Z., Mei, J., Lin, F., Huang, R. and Wang, P. [2018] Correcting for salt misinterpretation with full-waveform inversion. In: *SEG Technical Program Expanded Abstracts*. 1143–1147. (cit. on p. 17.)

Zhou, T. [2020] *Migration velocity analysis: selection of user parameters and introduction of transmitted waves*. Ph.D. thesis, MINES ParisTech–PSL Research University. (cit. on pp. xiv, 25, 26, 36, 54, and 56.)

Zhou, T., Chauris, H. and Audebert, F. [2020a] Applicability of Migration Velocity Analysis on a Marine Real Data Set. In: *82th EAGE Conference and Exhibition*, Th–Dome5–05. (cit. on p. 238.)

Zhou, T., Chauris, H. and Audebert, F. [2020b] Impact of user parameters in inversion velocity analysis. *Geophysical Prospecting*, **68**(5), 1492–1508. (cit. on pp. xv, 43, 62, and 67.)

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*, **22**, 1535–1554. (cit. on pp. 11 and 69.)

Zhu, F., Huang, J. and Yu, H. [2018] Least-squares Fourier finite-difference pre-stack depth migration for VTI media. *Journal of Geophysics and Engineering*, **15**(2), 421–437. (cit. on p. 20.)

RÉSUMÉ

L'imagerie sismique est l'une des méthodes les plus pertinentes pour l'estimation des paramètres physiques (vitesse, densité, ...) depuis l'acquisition de données en surface. Avec l'hypothèse de séparation d'échelles, plusieurs méthodes d'imagerie décomposent le modèle de vitesse entre un macro-modèle lisse et un modèle de réflectivité. Le but des techniques de migration est de déterminer la réflectivité dans un macro-modèle donné. Parmi différentes solutions, la Reverse Time Migration (RTM) est devenue la méthode de choix pour les milieux complexes. Par définition, RTM est l'adjoint de l'opérateur de Born et souffre de différents artéfacts de migration. Des développements récents ont permis d'analyser la RTM avec une approche asymptotique. Ils ont conduit à une méthode directe pour inverser l'opérateur de modélisation, et apporter une solution quantitative en une seule itération. L'inverse direct suppose un milieu acoustique à densité constante, ce qui représente une limite forte pour les applications pratiques.

Dans cette thèse, j'ai d'abord étendu l'applicabilité de l'inverse direct depuis une densité constante à une densité variable et vers les milieux élastiques. Dans le cadre de l'imagerie multi-paramètres, la principale limitation est la non unicité de la solution. Pour cela, je propose d'ajouter des contraintes avec une norme ℓ_1 sur chacune des classes de paramètres. De plus, je propose d'utiliser l'inverse direct pour accélérer la convergence de la RTM multi-paramètres. Les méthodologies sont développées et analysées sur des données synthétiques 2D et sur un cas réel marin.

MOTS CLÉS

imagerie sismique, reverse time migration, multi-paramètre, inversion, régularisation

ABSTRACT

Seismic imaging is one of the most effective methods for estimating the Earth's physical parameters from seismic data. Based on the assumption of scale separation, several imaging methods split the velocity model into a smooth background model and a reflectivity model. The goal of Migration techniques is to determine the reflectivity in a given background model. Among different migration algorithms, reverse time migration (RTM) has become the method of choice in complex geologic structures. By definition, RTM is the adjoint of the linearized Born modeling operator and suffers from various migration artifacts. Recent developments recast the asymptotic inversion in the context of RTM. They determine a direct method to invert the Born modeling operator, providing quantitative results within a single iteration. The direct inverse is based on constant-density acoustic media, which is a limiting factor for practical applications.

In this thesis, I first extend the applicability of direct inverse from constant-density acoustic to variable-acoustic acoustic and elastic media. In the concept of the multi-parameter imaging, the main limitation is the non-uniqueness of the inversion results. To tackle the ill-posedness of the inverse problem, I propose to add independently ℓ_1-norm constraints to each inverted parameter as regularization terms. Furthermore, I utilize the direct inverse to accelerate the convergence rate of multi-parameter least-squares RTM. The methodologies are developed and analyzed on 2D synthetic datasets and a marine real dataset.

KEYWORDS

seismic imaging, reverse time migration, multi-parameter, inversion, regularization