



Max Planck **Graduate Center** 
mit der Johannes Gutenberg-Universität



Identification of subsurface parameters using wave physics

PhD Thesis

for the award of the academic degree of
Doctor rerum naturalium (Dr. rer. nat.) of the faculties:
09 - Chemistry, Pharmacy, Geography and Earth Sciences

Submitted by
Yi Zhang
Mainz, 14.03.2025

Supervisors: Prof. Luca De Siena
Prof. Markus Bachmayr
Prof. Boris J.P. Kaus

Day of the oral examination:

Abstract

The identification of coefficients in the wave equation and eikonal equation is a fundamental problem in various fields, including subsurface imaging, non-destructive testing, and medical imaging. In exploration geophysics, subsurface imaging aims to reconstruct Earth models from partial measurement. This process relies on wave physics and is formulated as an inverse problem, where an objective function minimizes the discrepancy between simulated and observed data. This thesis focuses on the identification of subsurface parameters using wave physics based forward model.

Several challenges arise in this endeavor. First, wave physics models are inherently approximate, leading to inaccuracies in the inversion process. Second, partial observations result in an underdetermined problem, making solutions highly dependent on the starting model, frequency bandwidth and illumination. Third, the oscillatory nature of seismic waves introduces difficulties in optimization, often causing convergence to local minima (cycle skipping).

To address these challenges, this thesis investigates three key topics:

- **Seismic anisotropy:** Seismic anisotropy is prevalent in field data, and neglecting it can introduce significant errors in imaging. This work introduces a new acoustic Tilted Transversely Isotropic (TTI) wave equation by linearizing the anelliptic term using a Taylor series expansion. The method is benchmarked against existing approaches and is applied to 2D and 3D reverse time migration, demonstrating its robustness and computational efficiency.
- **Regularization in inverse problems:** Geophysical inverse problems suffer from non-uniqueness, necessitating regularization techniques for stable solutions. While conventional H^1 regularization (where the L^2 -norm of the model gradient serves as the regularization) functional suppresses model sharpness, Total Variation (TV) regularization preserves interfaces but is computationally demanding. This work explores Levenberg-Marquardt type Total Variation (TV-LM), which combines Bregman iteration with a primal-dual Newton solver. Compared to existing methods such as Modified Total Variation (MTV), TV-LM is shown to be computationally more efficient and capable of achieving better minimization in eikonal equation inverse problems.
- **Mitigating cycle skipping in Full Waveform Inversion (FWI) and Joint Mi-**

gration Inversion (JMI): Cycle skipping is a significant challenge in waveform inversion, hindering effective convergence. Utilizing the search space expansion to mitigate this issue, the author introduces a time-domain JMI by Source Extension (ES-JMI) method using the acoustic wave equation parameterized in velocity and impedance. The Hessian of the data augmented wave equation is transformed from the source to the receiver side for computational tractability. ES-JMI is benchmarked against conventional JMI and JMI by Wavefield Reconstruction inversion (WR-JMI), showing improved performance. Additionally, a source-independent Extended Source waveform Inversion (ESI) approach is developed using an augmented Lagrangian framework, incorporating TV-LM and bounded constraints. Numerical tests confirm that our ESI algorithm mitigates cycle skipping and achieves source-wavelet independence compared to conventional FWI.

This thesis advances the field of subsurface imaging by developing new methods to address seismic anisotropy, improve regularization, and mitigate cycle-skipping challenges in wave physics based inversions. The proposed techniques enhance the accuracy, stability, and computational efficiency of subsurface parameter estimation, contributing to more reliable seismic imaging in exploration geophysics.

“Cogito, ergo sum.” – René Descartes

Zusammenfassung

Die Identifikation von Koeffizienten in der Wellengleichung und der Eikonalgleichung ist ein grundlegendes Problem in verschiedenen Bereichen, darunter Untergrundabbildung, zerstörungsfreie Prüfung und medizinische Bildgebung. In der Explorationsgeophysik zielt die Untergrundabbildung darauf ab, Erdmodelle aus partiellen Messungen zu rekonstruieren. Dieser Prozess basiert auf der Wellenausbreitungsphysik und wird als inverses Problem formuliert, bei dem eine Zielfunktion die Diskrepanz zwischen simulierten und beobachteten Daten minimiert. Diese Dissertation konzentriert sich auf die Identifikation von Untergrundparametern unter Verwendung eines auf Wellenausbreitung basierenden Vorwärtsmodells.

Mehrere Herausforderungen treten in diesem Zusammenhang auf. Erstens sind Modelle der Wellenausbreitung naturgemäß approximativ, was zu Ungenauigkeiten im Inversionsprozess führt. Zweitens führen partielle Beobachtungen zu einem unterbestimmten Problem, was bedeutet, dass die Lösungen stark vom Startmodell, der Frequenzbandbreite und der Beleuchtung abhängen. Drittens führt die oszillatorische Natur seismischer Wellen zu Schwierigkeiten bei der Optimierung und verursacht oft die Konvergenz zu lokalen Minima (Cycle Skipping).

Um diese Herausforderungen zu adressieren, untersucht diese Dissertation drei zentrale Themen:

- **Seismische Anisotropie:** Seismische Anisotropie ist in Felddaten weit verbreitet, und ihre Vernachlässigung kann zu erheblichen Fehlern in der Bildgebung führen. Diese Arbeit stellt eine neue akustische *Tilted Transversely Isotropic* (TTI) Wellengleichung vor, indem das anelliptische Glied mittels einer Taylor-Reihenentwicklung linearisiert wird. Die Methode wird mit bestehenden Ansätzen verglichen und auf 2D- und 3D-*Reverse Time Migration* angewendet, wobei ihre Robustheit und rechnerische Effizienz demonstriert werden.
- **Regularisierung in inversen Problemen:** Geophysikalische Inversionsprobleme leiden unter Nicht-Eindeutigkeit, weshalb Regularisierungstechniken erforderlich sind, um stabile Lösungen zu gewährleisten. Während herkömmliche H^1 -Regularisierung (bei der die L^2 -Norm des Modellgradienten als Regularisierung dient) die Modellschärfe unterdrückt, bewahrt *Total Variation* (TV)-Regularisierung Grenzflächen, ist jedoch rechnerisch aufwendig. Diese Ar-

beit untersucht die Levenberg-Marquardt-Variante der Total Variation (TV-LM), die eine Bregman-Iteration mit einem Primal-Dual-Newton-Löser kombiniert. Im Vergleich zu bestehenden Methoden wie der *Modified Total Variation* (MTV) zeigt sich, dass TV-LM rechnerisch effizienter ist und eine bessere Minimierung in inversen Problemen der Eikonalgleichung erreicht.

- **Minderung von Cycle Skipping in Full Waveform Inversion (FWI) und Joint Migration Inversion (JMI):** Cycle Skipping stellt eine erhebliche Herausforderung in der Wellenforminversion dar und behindert eine effektive Konvergenz. Zur Minderung dieses Problems führt der Autor eine zeithybride JMI-Methode durch Quellen-Erweiterung (ES-JMI) ein, die die akustische Wellengleichung verwendet und in Geschwindigkeit und Impedanz parametrisiert ist. Die Hesse-Matrix der erweiterten Wellengleichung wird von der Quellenseite auf die Empfängerseite transformiert, um die Rechenlast zu verringern. ES-JMI wird mit konventionellem JMI und JMI durch Wellenfeldrekonstruktioninversion (WR-JMI) verglichen und zeigt eine verbesserte Leistung. Darüber hinaus wird ein quellenunabhängiger *Extended Source waveform Inversion* (ESI)-Ansatz entwickelt, der ein erweitertes Lagrange-Framework verwendet, das TV-LM und beschränkte Nebenbedingungen integriert. Numerische Tests bestätigen, dass der ESI-Algorithmus Cycle Skipping mildert und die Unabhängigkeit von der Quellwellenform erreicht.

Diese Dissertation trägt zur Weiterentwicklung der Untergrundabbildung bei, indem sie neue Methoden zur Berücksichtigung seismischer Anisotropie, zur Verbesserung der Regularisierung und zur Minderung von Cycle-Skipping-Problemen in auf Wellenausbreitung basierenden Inversionen entwickelt. Die vorgeschlagenen Techniken erhöhen die Genauigkeit, Stabilität und rechnerische Effizienz der Untergrundparameterbestimmung und tragen somit zu einer zuverlässigeren seismischen Bildgebung in der Explorationsgeophysik bei.

Acknowledgements

The author is sincerely grateful to his supervisors, Prof. Luca De Siena, Prof. Markus Bachmayr, and Prof. Boris J.P. Kaus, for their invaluable support and guidance throughout this thesis. He would also like to express his sincere gratitude to Prof. Alexey Stovas for the insightful discussions on wave equation based problems. Additionally, he appreciates the assistance of Dr. Giovanni Angelo Meles and Prof. Niklas Linde during his visiting at the University of Lausanne.

Deep gratitude is extended for the financial support provided by the Mainz Institute of Multiscale Modelling and the MPG Mainz.

Lastly, the author would like to express heartfelt thanks to his family and his fiancée, Jiaxin Yu, for their unwavering support and encouragement.

Declaration

I hereby declare that I wrote the dissertation submitted without any unauthorized external assistance and used only sources acknowledged in the work. All textual passages which are appropriated verbatim or paraphrased from published and unpublished texts as well as all information obtained from oral sources are duly indicated and listed in accordance with bibliographical rules. In carrying out this research, I complied with the rules of standard scientific practice as formulated in the statutes of Johannes Gutenberg University Mainz to insure standard scientific practice.

Mainz, Germany, 14.03.2025, Yi Zhang

Signature:

Data availability statement

The codes and models used in this thesis are available upon request from the author. The author's contact details can be found in CV. These materials were provided to the examiners during the thesis review process.

Contents

1	Introduction	1
1.1	Review	2
1.2	Outline	8
2	Problem formulation	11
2.1	Wave equations	11
2.1.1	Kinematics	12
2.1.2	VTI media	14
2.1.3	Isotropic P- and S-wave equations	16
2.1.4	Eikonal equation	16
2.2	Fundamentals of waveform inversion	17
3	Wavefield simulation and reverse time migration for acoustic TTI media	23
3.1	Introduction	24
3.2	Theory	27
3.2.1	A 2D acoustic VTI medium	27
3.2.2	An approximation to the anellipticity term	30
3.2.3	An acoustic TTI medium	32
3.2.4	The true 2D acoustic TTI solution	34
3.3	Numerical examples	35

3.3.1	A VTI kinematics benchmark	35
3.3.2	A wavefield extrapolator comparison to the low-rank approach	37
3.3.3	A wavefield test for the Greenhorn shale	39
3.3.4	A wedge model	39
3.3.5	An interference test	40
3.3.6	An application to 2D RTM	41
3.3.7	A 3D example	45
3.4	Discussions	46
3.5	Conclusions	50
4	Iterative total variation regularization for first-break travel time inversion	51
4.1	Introduction	52
4.2	Method	55
4.2.1	Total variation regularization of inverse problems	55
4.2.2	Bregman distance	56
4.2.3	Levenberg–Marquardt-type linearization	56
4.2.4	A primal-dual Newton’s method	57
4.2.5	Eikonal equation	58
4.2.6	Numerical implementation	59
4.3	Other methods	63
4.4	Numerical tests	63
4.4.1	Perturbation test	64
4.4.2	A toy model	64
4.4.3	Marmousi model	69
4.4.4	2004 BP model	70
4.5	Conclusions	72

5	Simultaneous velocity and reflectivity estimation using extended source waveform inversion	77
5.1	Introduction	78
5.2	Forward problem	81
5.3	All-at-once method	82
5.4	Reduced objective	83
5.5	Source extension	83
5.5.1	Data augmented wavefield	84
5.5.2	Model parameters update	86
5.5.3	Algorithm	88
5.6	Connections to band-pass filter	89
5.7	Numerical tests	92
5.7.1	Marmousi	92
5.7.2	2004 BP	95
5.8	Discussions	97
5.9	Conclusions	102
6	Source independent extended source waveform inversion based on augmented Lagrangian	105
6.1	Introduction	106
6.2	Conventional FWI	108
6.3	Source extension	109
6.4	Augmented Lagrangian and TV regularization	110
6.4.1	Optimization over s and $\mathbf{u}$	112
6.4.2	Optimization over $\mathbf{v}$	112
6.5	Numerical tests	116
6.5.1	Marmousi	116

CONTENTS

6.5.2	2004 BP	123
6.6	Conclusions	128
7	Conclusions	129
8	Outlook	133
8.1	Comparisons for different TV methods	133
8.2	Applications to real data	134
8.3	Medical imaging	134
	References	149
	List of figures	156
	List of tables	157
	List of acronyms	159
	List of symbols	161
	Appendix	163
A	Coefficients for acoustic TTI wave equation	163
B	An adjoint approach for the linearized eikonal equation	164
C	Adjoint wave equation	166
D	Velocity and impedance gradients for the Lagrangian	167
E	From data augmented wavefield to source extension	168
F	Receiver side Hessian	169
CV		171

Chapter 1

Introduction

A fundamental goal of geophysics is to reconstruct high-resolution subsurface models using seismic waves. This process leverages surface-recorded seismic data to estimate key parameters such as wave speed, attenuation, anisotropy, and density (Figure 1.1). Achieving this requires simulating seismic wave propagation, solving inverse problems, and applying seismic imaging techniques. While primarily used for georesource exploration (Calvert, 2005; Jain & deFigueiredo, 1982) and environmental monitoring (Caspari et al., 2015; Pevzner, Caspari, Gurevich, Dance, & Cinar, 2015), these methods also find applications in medical imaging (Guasch, Calderon Agudo, Tang, Nachev, & Warner, 2020), nondestructive material testing (Rossato et al., 2023), and sonar technologies (Sibley et al., 2023).

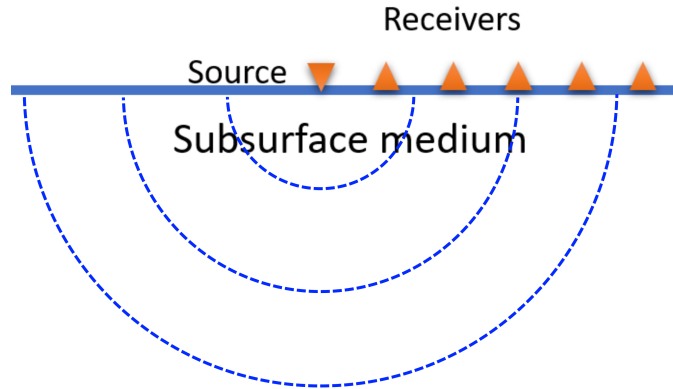


Figure 1.1: A schematic of the seismic experiment. To determine the properties of the subsurface medium, a seismic source is activated at the surface, generating waves that propagate through the subsurface. These waves are recorded by receivers at the surface, providing data that can be used to compute the subsurface parameters.

Conventional subsurface imaging follows a piecemeal approach, involving multiple subproblems such as preprocessing, demultiple, velocity analysis, migration, and post-processing. The lack of integration between these steps can introduce various

issues, including velocity errors, residual multiples, and migration artifacts, all of which degrade the quality of the subsurface image (Wei et al., 2023). In contrast, the inversion framework, formulated as a Partial Differential Equation (PDE) - constrained optimization problem, provides a more systematic approach by directly minimizing the discrepancy between simulated and observed data, leading to more accurate solutions.

At the core of subsurface parameter estimation lies the wave equation and its derivatives such as eikonal equation and transport equation, which govern the physics of the seismic wave propagation. The standard workflow of subsurface parameter estimation involves solving a forward problem, where state variables such as wavefield and travel time are simulated based on an assumed Earth model (e.g., Figure 1.2), and an inverse problem, where model parameters are iteratively updated to minimize discrepancies between observed and simulated data. Forward modeling typically involves solving PDEs such as the wave equation for waveform modeling (Figure 1.3a-c), the eikonal equation for travel time estimation (Figure 1.3d), and the radiative transfer equation for wave energy transport. By sampling the observed data at the receiver locations (Figure 1.4), one can compare the simulated observation with the field data.

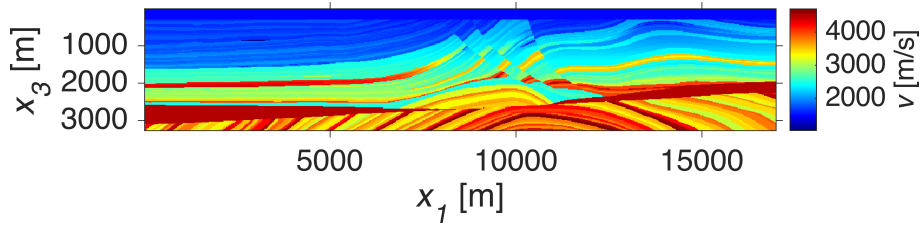


Figure 1.2: A synthetic subsurface acoustic model v defined in a 2D space $[x_1 \ x_3]^T$ created by Irons (1988) – Marmousi model.

Two major challenges arise in the forward problem. First, selecting an appropriate wave equation is crucial to accurately describe the underlying physical phenomena. Second, solving PDEs, particularly in large-scale settings, is computationally demanding.

The inverse problem, on the other hand, presents additional difficulties. Due to the limited and surface-only nature of seismic observations, the inversion is inherently ill-posed, leading to non-uniqueness (e.g., cycle skipping, Figure 1.5) and sensitivity to noise. Overcoming these challenges requires advanced modeling techniques, efficient numerical solvers, and robust regularization strategies.

1.1 Review

Seismic wave propagation in the Earth is governed by wave equations (Aki & Richards, 2002; Carcione, 1990, 2007). The subsurface often exhibits seismic anisotropy,

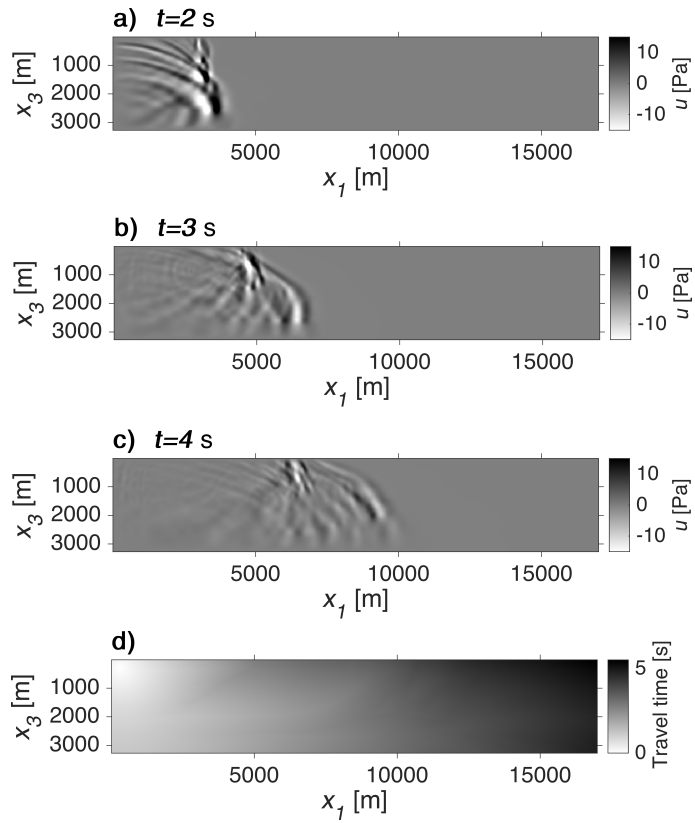


Figure 1.3: Simulations based on the model in Figure 1.2. The source is placed at $[300, 281.25]^T$ m. 292 receivers are evenly placed at $x_1 \in [300, 16669]$ m, $x_3 = 281.25$ m. (a),(b) and (c) are the acoustic wavefields at different time steps, simulated by the acoustic wave equation. (d) is the travel time obtained by the eikonal equation.

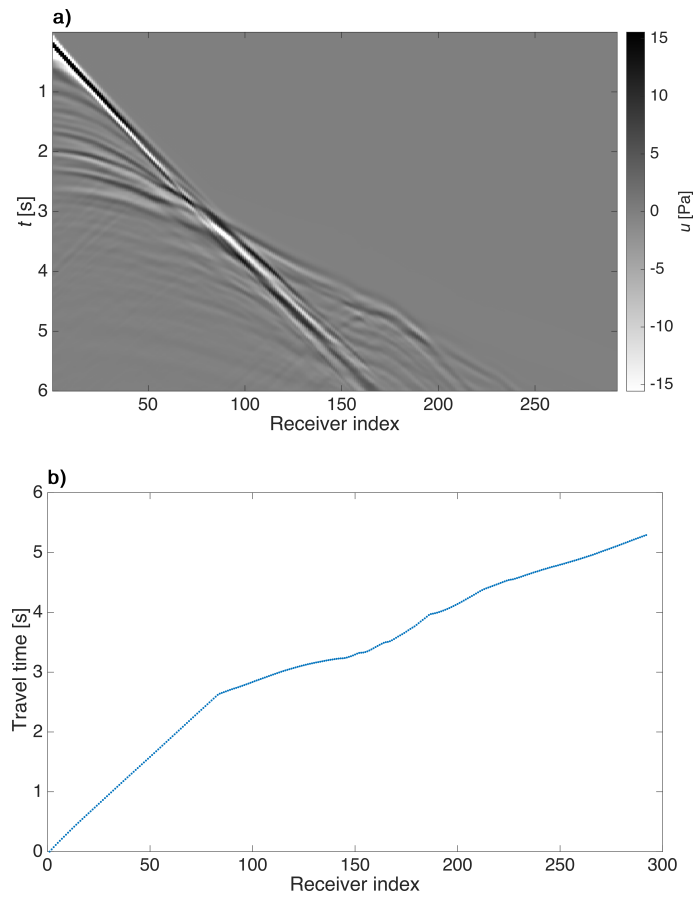


Figure 1.4: Observed data sampled for the numerical experiment at Figure 1.3, (a) for the wavefield and (b) for the travel time.

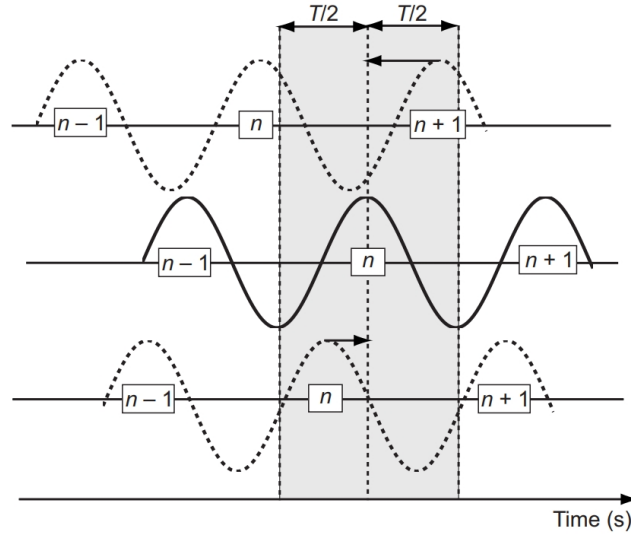


Figure 1.5: A schematic illustrating cycle skipping in Full Waveform Inversion (FWI) (Virieux & Operto, 2009). Cycle skipping is a primary cause of non-uniqueness in FWI. The upper dashed line represents the simulated data, which deviates by more than half of the period, $T/2$, from the recorded data (shown by the middle solid line). In this case, FWI mismatches cycle n to cycle $n - 1$. The bottom dashed line represents another simulated data set, which deviates by less than $T/2$ from the observed data. Here, FWI successfully matches the two signals. Cycle skipping is affected by the initial model and frequency band.

where wave velocity varies with direction due to intrinsic material properties or structural alignments (Stovas, 2017; Thomsen, 1986). Additionally, seismic attenuation influences wave propagation, further complicating wave behavior (De Siena, Pezzo, & Bianco, 2011; De Siena, Thomas, & Aster, 2014). To account for these effects, researchers have developed anisotropic wave equations (Alkhalifah, 2000, 2002) and viscoelastic wave equations (Carcione, 1990; Robertsson, Blanch, & Symes, 1994). Poroelasticity models provide further refinement by incorporating fluid-solid interactions within the Earth’s porous media (Parra, 1997; D. Yang & Zhang, 2002). These advanced models improve the accuracy of waveform simulations, aiding in a more precise characterization of subsurface properties.

In addition to the wave equation, ray theory, known for its computational efficiency, is widely used in geophysics to model the forward problem by focusing on specific seismic phases based on the infinite frequency limit of seismic waves (Fomel, 1997; Julian & Gubbins, 1977; Rueger, 1993). Ray-based methods are particularly useful for initial model estimations, serving as a precursor to more detailed wave equation-based inversions.

Forward models, essential to inversions, aid in estimating subsurface parameters by solving constrained optimization problems. Due to the nonlinear relationship between travel time or waveform and material properties, iterative refinement is necessary to achieve more accurate estimations. In geophysical inversions, the objective is to determine subsurface parameters by solving these constrained optimization

problems, also known as inversions.

In geophysical inversions, the constraint is typically a relevant physical equation implemented through a Lagrangian function (Haber & Ascher, 2000). Ideally, each iteration involves solving the first-order optimality condition, known as the Karush-Kuhn-Tucker (KKT) condition. This process requires the simultaneous updating of the forward solution, the adjoint solution, and the model parameters within a single iteration. However, the dimensionality of the KKT condition is prohibitively large, making this direct approach computationally impractical. To address this challenge, a popular solution is the reduced-space method (e.g., Plessix, 2006). In this approach, the KKT condition is decomposed and solved sequentially. Specifically, the forward PDE, the adjoint PDE, and the gradient problem are tackled in succession. This iterative process significantly reduces computational complexity while maintaining the integrity of the inversion process.

The reduced-space approach is widely employed in geophysics for waveform inversion (Leung & Qian, 2006; Plessix, 2006; Tromp, Tape, & Liu, 2004) and travel time inversion (Taillandier, Noble, Chauris, & Calandra, 2009). This method utilizes the adjoint variables, also known as Lagrangian multipliers, along with the state variables to compute the model update direction. In practice, a line search is conducted to determine the optimal step length for each iteration. A more sophisticated class of methods involves Newton techniques, which incorporate the Hessian matrix to achieve faster convergence and improved results. However, these methods are computationally intensive and are primarily used for solving frequency-domain wave equation problems (Gholami, 2024; Métivier, Brossier, Virieux, & Operto, 2013; Pratt, 1999) and eikonal equation problems (Aki & Lee, 1976). For large-scale inverse problems, the L-BFGS method, a quasi-Newton approach, is preferred. This method bypasses the need for direct Hessian calculation by using a limited number of recent gradients to approximate the Hessian’s diagonal, making it feasible for most geophysical inversions.

For waveform inversion, recent studies have shown that the reduced-space method is highly non-convex and requires a well-chosen starting model. To address this limitation, van Leeuwen and Herrmann (2013) proposed a physics-penalized cost function in addition to the data misfit. This approach extends the search space to both the state variables and model parameters, in contrast to the single-parameter search (only the model parameters) in the reduced-space method. This formulation is known as Wavefield Reconstruction Inversion (WRI).

The key difference between WRI and the conventional reduced-space approach is that model updates in WRI are performed using the data-augmented wavefield rather than the wavefield solved solely from the model. WRI can be formulated and solved in three ways:

- **Alternating updates:** This approach iterates between updating the data-augmented wavefield and the model parameters (van Leeuwen & Herrmann, 2013). While simple to implement, it typically requires many iterations to

achieve a good minimizer due to the basic optimization scheme. The Hessian for the model update is exactly sparse and can be accumulated in a running sum.

- **Reduced WRI objective function:** Here, the data-augmented wavefield is projected onto the model parameters (van Leeuwen & Herrmann, 2016). The resulting Hessian is exactly dense but allows for a natural sparse approximation. The model update can be efficiently solved using quasi-Newton or Gauss–Newton methods.
- **Augmented Lagrangian formulation:** This method is currently the most accurate way to solve the WRI problem (Aghamiry, Gholami, & Operto, 2019a). The Hessian remains exactly sparse and can be stored in a running sum. However, a major drawback is its high storage demand for dual variables. As a result, this approach is not suitable for time-domain implementations, where storing time-domain dual variables for each source is prohibitively expensive.

The WRI problem can be solved in either the time domain or the frequency domain. The key difference lies in the data-augmented wave equation, as solving it in the time domain is not straightforward. Previous studies have explored approximate solutions to the data-augmented wave equation in the time domain (Aghamiry, Gholami, & Operto, 2019b, 2020b; Z.-C. Li, Lin, Zhang, Li, & Yu, 2018; Métivier et al., 2013).

Source extension is a more general formulation of WRI, where an annihilator is utilized as the weighting function for the physics penalty (G. Huang, Nammour, & Symes, 2018b). A key advantage of Extended Source waveform Inversion (ESI) is that it can be made source independent (G. Huang, Nammour, & Symes, 2018a). This represents a significant advancement, as it eliminates the need to estimate the source from data before inversion—an aspect that is typically highly dependent on the initial model (Virieux & Operto, 2009).

Beyond wave speed models, seismic reflectivity is crucial for identifying interfaces and estimating volumes. Traditional approaches first estimate wave speed and then relocate reflectors using migration. Common migration techniques include Kirchhoff migration (Calandra, Baina, Hanitzsch, & Rousseau, 2001), one-way wave equation methods (Pedersen, Ursin, & Helgesen, 2010), reverse time migration (RTM) (Z.-Z. Zhou, Howard, & Mifflin, 2011), and least-squares RTM (LS-RTM) (Mu, Huang, Yang, Guo, & Guo, 2020). Recent advancements focus on simultaneously updating both wave speed and reflectivity (Whitmore, Ramos-Martinez, Yang, & Valenciano, 2020; Y. Yang, Ramos-Martinez, Whitmore, Huang, & Chemingui, 2021a, 2021b; Y. Yang, Ramos-Martinez, Whitmore, Valenciano, & Chemingui, 2020). In this case, greater attention is given to the coupling between the background velocity and reflectivity.

1.2 Outline

The primary objective of this thesis is to investigate seismic anisotropy, develop improved total variation regularization techniques, and enhance inversion strategies for more accurate parameter identification.

Chapter 2 formulates the foundational forward problems, covering the wave equation, seismic anisotropy, and the eikonal equation. Additionally, it provides an introductory overview of waveform inversion.

In Chapter 3, the author presents **Wavefield simulation and reverse time migration for acoustic TTI media**. The motivation for performing this work stems from a new finding by Stovas, Alkhalifah, and Waheed (2020), where the researchers proposed a general approach to approximate pure P-waves in anisotropic media. Specifically, by setting the S-wave velocity to zero in all propagating directions within the dispersion relation, pure P-wave information can be extracted. However, the wave equation in Stovas et al. (2020) is not straightforward to solve using numerical methods.

This chapter introduces a new algorithm for approximating the acoustic wavefield in Tilted Transversely Isotropic (TTI) media. In pure acoustic TTI media, the space and wavenumber components are typically coupled within the wavefield extrapolator, making wavefield simulation particularly challenging. To obtain a pure P-wave without any S-wave artifacts, the S-wave velocity in the Christoffel equation must be set to zero in all propagating directions. However, the associated wave equation becomes intractable to solve due to the mixed space/wavenumber coupling and the presence of high-order derivatives. The author proposes an analytical decoupling approach by employing a Taylor series expansion to linearize the anellipticity term, effectively decoupling the space-dependent and wavenumber-dependent variables within the wavefield extrapolator. The resulting wave equation can be solved using a hybrid finite-difference/pseudospectral method. The proposed method is benchmarked against classical acoustic approximations for TTI media (Alkhalifah, 2000; Chu, Macy, & Anno, 2011; Pestana, Ursin, & Stoffa, 2012; Stovas, 2020; L. Zhang & Ohen, 2022). It has been demonstrated that this approach provides competitive accuracy for wavefield simulation while reducing computational cost.

Chapter 4 discusses **Iterative total variation regularization for first-break travel time inversion**. In geophysics, H^1 regularization is widely used to stabilize the inversion, where the L^2 -norm of the gradient serves as the regularization term. However, this penalizes sharp interfaces. On the other hand, Total Variation (TV) regularization is well-known for preserving sharp interfaces, but it is numerically expensive due to the non-differentiability of the subgradient. In geophysics, split-type TV methods, such as Modified Total Variation (MTV) (Jiang & Zhang, 2017; Lin & Huang, 2015a, 2015b) and split-Bregman (Aghamiry et al., 2019a), are widely adopted due to their ability to decompose the original inversion

problem with TV regularization into two simpler subproblems, making them computationally efficient. Specifically, in the case of MTV, the subproblems involve solving a Tikhonov-regularized problem and an image denoising problem. The denoising problem can be tackled with various TV denoising algorithms, such as the Chambolle method (Chambolle, 2004) or the split-Bregman method (Goldstein & Osher, 2009). The split-Bregman method, in particular, can be directly applied to geophysical inversion problems, as demonstrated here (Aghamiry et al., 2019a). While split-type methods simplify computations, they introduce additional weighting factors that complicate parameter tuning. To overcome this issue, the author proposes a new algorithm that requires less parameter tuning.

This chapter addresses the travel time inversion problem using a new TV regularization. In this work, the author employs a smoothed TV-norm, replacing the TV regularization term with the Bregman distance. A Levenberg-Marquardt-type linearization is applied to the forward problem (Bachmayr & Burger, 2009), and the resulting algorithm is named TV with Levenberg-Marquardt-type linearization (TV-LM) in this thesis. The primal-dual Newton’s method is then used to solve the first-order optimality conditions. Numerical tests show that TV-LM converges to better minimizers compared to L-BFGS (Leung & Qian, 2006; Taillandier et al., 2009) applied to the least-squares functional with H^1 regularization (Rawlinson & Sambridge, 2003), as well as to the widely used modified total variation (MTV, Y. Huang, Ng, & Wen, 2008; Jiang & Zhang, 2018). The author demonstrates that TV-LM is more accurate and efficient in reconstructing the subsurface model for the eikonal equation problem.

Chapter 5 introduces **Simultaneous velocity and reflectivity estimation using extended source waveform inversion**. A major issue in waveform inversion is the local minimum problem caused by cycle skipping. This typically occurs when a crude initial iterate is provided or when low-frequency data components are missing. Wavefield Reconstruction Inversion (WRI) can mitigate this local minimum issue by enlarging the search space (van Leeuwen & Herrmann, 2013). However, a significant challenge for time-domain WRI implementation is the expensive evaluation of the data-augmented wave equation. In recent years, several studies have explored ways to approximate the time-domain data-augmented wave equation (Aghamiry, Gholami, & Operto, 2020a; Gholami, Aghamiry, & Operto, 2022; Rizzuti, Louboutin, Wang, Daskalakis, & Herrmann, 2019; C. Wang, Yingst, Farmer, & Leveille, 2016).

Simultaneous velocity and reflectivity inversion, also known as Joint Migration Inversion (JMI), is similarly affected by the local minimum issue. However, few studies have investigated JMI within the extended search space. The source extension (G. Huang et al., 2018a; P. Yang & Zhou, 2024), a location-weighted version of WRI, reduces the inversion’s vulnerability to cycle skipping. The author employs source extension to velocity-impedance parameterized acoustic wave equation. The resulting algorithm is named Joint Migration Inversion by Extended Source waveform inversion (ES-JMI) as it can reconstructs velocity and reflectivity simultaneously. In benchmarks, the author shows that ES-JMI outperforms JMI by Wavefield Reconstruction (WR-JMI) and conventional Full Waveform Inversion (FWI) followed

by Least-Squares Reverse Time Migration (LS-RTM) in terms of wave speed and reflectivity reconstruction.

Chapter 6 introduces **Source independent extended source waveform inversion based on augmented Lagrangian**. Recent studies on source extension have successfully explored the possibility of source signature independent inversion (G. Huang et al., 2018a). However, these studies are limited to the quasi-Newton method for model updates. A more sophisticated optimization method for the ESI problem is the Augmented Lagrangian (AL) (Aghamiry et al., 2019a; Gholami, Aghamiry, & Operto, 2021; Gholami et al., 2022).

In this chapter, the author combines three key ingredients for waveform inversion: source extension for the cost function, a reformulation of the source extension problem into an augmented Lagrangian framework, and TV-LM regularization. The proposed method does not require prior knowledge of the source wavelet. The augmented Lagrangian approach enables a fixed physics penalty size, thus avoiding tedious tuning. Meanwhile, the TV-LM regularization makes the algorithm suitable for reconstructing piecewise-constant models, such as high contrast media. Numerical tests demonstrate that the proposed method can reconstruct accurate subsurface models without any input of the source wavelet, and it is more reliable than conventional FWI.

Chapter 2

Problem formulation

This chapters formulates several problems mathematically, including wave equation, eikonal equation, anisotropy and basics of waveform inversions.

2.1 Wave equations

The elastic wave equation is a combination of the generalized Hook's law and the equation of motion.

In this thesis, the 3D Cartesian coordinate with right-hand rule is applied. The three principal coordinates are $x = [x_1 \ x_2 \ x_3]^T \in \mathbb{R}^3$.

In this case, the generalized Hook's law is given by,

$$\sigma = C\epsilon, \quad (2.1)$$

where σ is the vector composed of stress tensor elements; C is the stiffness matrix using the Voigt notation and the forth-order stiffness tensor is given by c_{ijel} , where $i, j, e, l \in \{1, 2, 3\}$ (Chapter 1.1.4, Tsvankin, 2012). To convert the 4th-order tensor c to the stiffness matrix C , the following convention is applied: $11 \rightarrow 1$, $22 \rightarrow 2$, $33 \rightarrow 3$, $23 \rightarrow 4$, $13 \rightarrow 5$, and $12 \rightarrow 6$. ϵ is the strain. For a 3D problem, $\sigma = [\sigma_{11} \ \sigma_{22} \ \sigma_{33} \ \sigma_{23} \ \sigma_{13} \ \sigma_{12}]^T$, $C \in \mathbb{R}^{6 \times 6}$ and $\epsilon = [\epsilon_{11} \ \epsilon_{22} \ \epsilon_{33} \ \epsilon_{23} \ \epsilon_{13} \ \epsilon_{12}]^T$. For 2D in 1-3 plane, $\sigma^{(2D)} = [\sigma_{11} \ \sigma_{33} \ \sigma_{13}]^T$, $C^{(2D)} \in \mathbb{R}^{3 \times 3}$ and $\epsilon^{(2D)} = [\epsilon_{11} \ \epsilon_{33} \ \epsilon_{13}]^T$.

The strain is defined by (Poplavko, 2019)

$$\epsilon_{ij} = \frac{1}{2} \left(\frac{\partial u_{d,i}}{\partial x_j} + \frac{\partial u_{d,j}}{\partial x_i} \right), \quad (2.2)$$

where $u_d = [u_{d,1} \ u_{d,2} \ u_{d,3}]^T$ is the displacement vector.

The equation of motion can be cast as (Poplavko, 2019),

$$\rho \frac{\partial V_i}{\partial t} = \frac{\partial \sigma_{ij}}{\partial x_j} , \quad (2.3)$$

where ρ is the density; $V = [V_1 \ V_2 \ V_3]^T$ is the particle velocity vector; t is the time. The Einstein notation is applied to equation 2.3.

Take the time derivatives for equation 2.1 and recognize,

$$\frac{\partial u_{d,i}}{\partial t} = V_i . \quad (2.4)$$

The time derivative for the Hook's law in 3D is given by,

$$\frac{\partial}{\partial t} \begin{bmatrix} \sigma_{11} \\ \sigma_{22} \\ \sigma_{33} \\ \sigma_{23} \\ \sigma_{13} \\ \sigma_{12} \end{bmatrix} = C \frac{\partial}{\partial t} \begin{bmatrix} \frac{\partial V_1}{\partial x_1} \\ \frac{\partial V_2}{\partial x_2} \\ \frac{\partial V_3}{\partial x_3} \\ \frac{1}{2} \left(\frac{\partial V_2}{\partial x_3} + \frac{\partial V_3}{\partial x_2} \right) \\ \frac{1}{2} \left(\frac{\partial V_1}{\partial x_3} + \frac{\partial V_3}{\partial x_1} \right) \\ \frac{1}{2} \left(\frac{\partial V_1}{\partial x_2} + \frac{\partial V_2}{\partial x_1} \right) \end{bmatrix} . \quad (2.5)$$

Now one can solve the system combining equations 2.3 and 2.5 to update V and σ . This is named velocity-stress formulation of the wave equation.

Furthermore, one can eliminate σ given in the system by equations 2.3 and 2.5 and change V to u_d by equation 2.4, the wave equation in the displacement form can be obtained,

$$\rho \frac{\partial^2 u_{d,i}}{\partial t^2} = \frac{\partial}{\partial x_j} c_{ijel} \frac{\partial u_{d,e}}{\partial x_l} + s_i , \quad (2.6)$$

where s_i is the source term in the i -direction.

2.1.1 Kinematics

In this subsection, the author demonstrates how to compute the phase velocity for the wave equations.

Consider an isotropic stiffness matrix,

$$C^{(\text{iso})} = \begin{bmatrix} \lambda_{\text{iso}} + 2\mu_{\text{iso}} & \lambda_{\text{iso}} & \lambda_{\text{iso}} & 0 & 0 & 0 \\ \lambda_{\text{iso}} & \lambda_{\text{iso}} + 2\mu_{\text{iso}} & \lambda_{\text{iso}} & 0 & 0 & 0 \\ \lambda_{\text{iso}} & \lambda_{\text{iso}} & \lambda_{\text{iso}} + 2\mu_{\text{iso}} & 0 & 0 & 0 \\ 0 & 0 & 0 & \mu_{\text{iso}} & 0 & 0 \\ 0 & 0 & 0 & 0 & \mu_{\text{iso}} & 0 \\ 0 & 0 & 0 & 0 & 0 & \mu_{\text{iso}} \end{bmatrix} , \quad (2.7)$$

and the corresponding wave equation can be obtained from equation 2.6,

$$\rho \frac{\partial^2 u_{d,i}}{\partial t^2} = (\lambda_{\text{iso}} + \mu_{\text{iso}}) \frac{\partial}{\partial x_i} \frac{\partial u_{d,j}}{\partial x_j} + \mu_{\text{iso}} \frac{\partial}{\partial x_j} \frac{\partial u_{d,i}}{\partial x_j} , \quad (2.8)$$

where the weak variation for Lamé constants λ_{iso} and μ_{iso} is assumed for simplicity. The source term is ignored in equation 2.8 as it does not affect the phase velocity.

Let us define the Fourier transform in both time and space domains,

$$u_d(\omega, k) = \mathbb{F}(u_d(t, x)) = \frac{1}{(2\pi)^{N_{\text{dim}}/2}} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} u_d(t, x) \exp(-i\omega t + i k \cdot x) dt dx , \quad (2.9)$$

where N_{dim} is the number of dimensions; ω is the angular frequency; k is the angular wave vector; $i = \sqrt{-1}$.

In 3D, $k = [k_1 \ k_2 \ k_3]^T$ and $N_{\text{dim}} = 3$. Now we use the spherical coordinate with dip angle ϕ , azimuth angle θ and wavenumber $\hat{k}$, the following relationships can be obtained,

$$k_1 = \hat{k} \sin \theta \cos \phi , \quad (2.10)$$

$$k_2 = \hat{k} \sin \theta \sin \phi , \quad (2.11)$$

$$k_3 = \hat{k} \cos \theta . \quad (2.12)$$

Then equation 2.8 can be recast in (ω, k) domain,

$$A u_d(\omega, k) = 0 , \quad (2.13)$$

where $A = A^{(\text{iso})}$ for isotropic media, and

$$\begin{aligned} A^{(\text{iso})} &= \begin{bmatrix} A_{11}^{(\text{iso})} & A_{12}^{(\text{iso})} & A_{13}^{(\text{iso})} \\ A_{21}^{(\text{iso})} & A_{22}^{(\text{iso})} & A_{23}^{(\text{iso})} \\ A_{31}^{(\text{iso})} & A_{32}^{(\text{iso})} & A_{33}^{(\text{iso})} \end{bmatrix} , \\ A_{11}^{(\text{iso})} &= -\rho\omega^2 + (\lambda_{\text{iso}} + 2\mu_{\text{iso}})k_1^2 + \mu_{\text{iso}}k_2^2 + \mu_{\text{iso}}k_3^2 , \\ A_{12}^{(\text{iso})} &= (\lambda_{\text{iso}} + \mu_{\text{iso}})k_1k_2 , \\ A_{13}^{(\text{iso})} &= (\lambda_{\text{iso}} + \mu_{\text{iso}})k_1k_3 , \\ A_{21}^{(\text{iso})} &= -(\lambda_{\text{iso}} + \mu_{\text{iso}})k_1k_2 , \\ A_{22}^{(\text{iso})} &= -\rho\omega^2 + \mu_{\text{iso}}k_1^2 + (\lambda_{\text{iso}} + 2\mu_{\text{iso}})k_2^2 + \mu_{\text{iso}}k_3^2 , \\ A_{23}^{(\text{iso})} &= (\lambda_{\text{iso}} + \mu_{\text{iso}})k_2k_3 , \\ A_{31}^{(\text{iso})} &= (\lambda_{\text{iso}} + \mu_{\text{iso}})k_1k_3 , \\ A_{32}^{(\text{iso})} &= (\lambda_{\text{iso}} + \mu_{\text{iso}})k_2k_3 , \\ A_{33}^{(\text{iso})} &= -\rho\omega^2 + \mu_{\text{iso}}k_1^2 + \mu_{\text{iso}}k_2^2 + (\lambda_{\text{iso}} + 2\mu_{\text{iso}})k_3^2 . \end{aligned} \quad (2.14)$$

Equation 2.13 has a non-zero solution for $u_d(\omega, k)$, which implies

$$\det(A) = 0 . \quad (2.15)$$

Equation 2.15 is a sextic equation for $\hat{k}$, and its solution is independent of θ and ϕ ,

$$\hat{k} = \omega \sqrt{\frac{\rho}{\lambda_{\text{iso}} + 2\mu_{\text{iso}}}} \text{ (P - wave)}, \quad \omega \sqrt{\frac{\rho}{\mu_{\text{iso}}}} \text{ (SV - wave)}, \quad \omega \sqrt{\frac{\rho}{\mu_{\text{iso}}}} \text{ (SH - wave)} . \quad (2.16)$$

Only the positive solution for the wavenumber is kept. For additional properties of the three waves, readers are referred to Chapter 1.2 of Tsvankin (2012).

According to the definition of the phase velocity,

$$v = \frac{\omega}{\text{Real}(\hat{k})} \quad (2.17)$$

one can obtain the phase velocity for the three waves,

$$v = \sqrt{\frac{\lambda_{\text{iso}} + 2\mu_{\text{iso}}}{\rho}} \text{ (P-wave)}, \quad \sqrt{\frac{\mu_{\text{iso}}}{\rho}} \text{ (SV-wave)}, \quad \sqrt{\frac{\mu_{\text{iso}}}{\rho}} \text{ (SH-wave)} . \quad (2.18)$$

As the phase velocities are not dependent on ϕ and θ , those velocities are isotropic.

2.1.2 VTI media

Seismic anisotropy happens when the phase velocities for the three waves are angle dependent. Common anisotropy includes VTI (Vertically Transverse Isotropy), TTI (Tilted Transverse Isotropy), ORT (Orthorhombic Transverse Isotropy), Monoclinic and Triclinic (Tsvankin, Gaiser, Grechka, van der Baan, & Thomsen, 2010).

The author will give an example of Vertically Transversely Isotropic (VTI) media. The stiffness matrix for a VTI medium is given by

$$C^{(\text{vti})} = \begin{bmatrix} C_{11} & C_{11} - 2C_{66} & C_{13} & 0 & 0 & 0 \\ C_{11} - 2C_{66} & C_{11} & C_{13} & 0 & 0 & 0 \\ C_{13} & C_{13} & C_{33} & 0 & 0 & 0 \\ 0 & 0 & 0 & C_{55} & 0 & 0 \\ 0 & 0 & 0 & 0 & C_{55} & 0 \\ 0 & 0 & 0 & 0 & 0 & C_{66} \end{bmatrix} . \quad (2.19)$$

Following the same procedure for the isotropic medium phase velocity, we can obtain

A in equation 2.13 but for VTI media is given by

$$\begin{aligned}
 A^{(\text{vti})} &= \begin{bmatrix} A_{11}^{(\text{vti})} & A_{12}^{(\text{vti})} & A_{13}^{(\text{vti})} \\ A_{21}^{(\text{vti})} & A_{22}^{(\text{vti})} & A_{23}^{(\text{vti})} \\ A_{31}^{(\text{vti})} & A_{32}^{(\text{vti})} & A_{33}^{(\text{vti})} \end{bmatrix}, \\
 A_{11}^{(\text{vti})} &= -\rho\omega^2 + C_{11}k_1^2 + C_{66}k_2^2 + C_{55}k_3^2, \\
 A_{12}^{(\text{vti})} &= (C_{11} - C_{66})k_1k_2, \\
 A_{13}^{(\text{vti})} &= (C_{13} + C_{55})k_1k_3, \\
 A_{21}^{(\text{vti})} &= -(C_{11} - C_{66})k_1k_2, \\
 A_{22}^{(\text{vti})} &= -\rho\omega^2 + C_{66}k_1^2 + C_{11}k_2^2 + C_{55}k_3^2, \\
 A_{23}^{(\text{vti})} &= (C_{13} + C_{55})k_2k_3, \\
 A_{31}^{(\text{vti})} &= (C_{13} + C_{55})k_1k_3, \\
 A_{32}^{(\text{vti})} &= (C_{13} + C_{55})k_2k_3, \\
 A_{33}^{(\text{vti})} &= -\rho\omega^2 + C_{55}k_1^2 + C_{55}k_2^2 + C_{33}k_3^2.
 \end{aligned} \tag{2.20}$$

Similarly, the phase velocity for the three waves are given by (Tsvankin, 2012),

$$v = \sqrt{\frac{(C_{11} + C_{55})\sin^2\theta + (C_{33} + C_{55})\cos^2\theta \pm a}{2\rho}} \tag{2.21}$$

$$\text{(plus for qP-wave and minus for qSV-wave),} \tag{2.22}$$

$$\sqrt{\frac{C_{55}\cos^2\theta + C_{66}\sin^2\theta}{\rho}} \text{ (qSH-wave),} \tag{2.23}$$

$$a = \sqrt{((C_{11} - C_{55})\sin^2\theta - (C_{33} - C_{55})\cos^2\theta)^2 + 4(C_{13} + C_{55})^2\sin^2\theta\cos^2\theta}. \tag{2.24}$$

The phase velocities are independent of the azimuth ϕ but depend on the dip angle θ . In a VTI medium, pure P- and S-waves exist only along specific propagation directions. Therefore, the three wave modes are referred to as Quasi P (qP), Quasi SV (qSV), and Quasi SH (qSH). For further details, readers are referred to Chapter 1.1.2 of Tsvankin (2012). Additionally, P- and SV-waves are intrinsically coupled in a VTI medium, making their separation impossible using Helmholtz decomposition.

In seismic applications, the pure acoustic wave equation is widely used due to its lower computational cost and reduced parameter trade-offs in inversion compared to its elastic counterpart. Chapter 3 provides a detailed explanation of acoustic approximations for a more general case – Tilted Transversely Isotropic (TTI) media—and analyzes their applications in reverse time migration.

2.1.3 Isotropic P- and S-wave equations

In this part, the author introduces how to derive isotropic P- and S-wave equations.

The vector form of equation 2.8 is given by (ignoring the source term),

$$\rho \frac{\partial^2 u_d}{\partial t^2} = (\lambda_{\text{iso}} + 2\mu_{\text{iso}}) \nabla(\nabla \cdot u_d) - \mu_{\text{iso}} \nabla \times (\nabla \times u_d) . \quad (2.25)$$

Take divergence on both sides for equation 2.25 and define P-wave scalar potential $u_p = \nabla \cdot u_d$,

$$\frac{\partial^2 u_p}{\partial t^2} = \frac{\lambda_{\text{iso}} + 2\mu_{\text{iso}}}{\rho} \nabla^2 p = v_p^2 \nabla^2 p , \quad (2.26)$$

which is the P-wave equation and it is a scalar equation. $v_p = \sqrt{\frac{\lambda_{\text{iso}} + 2\mu_{\text{iso}}}{\rho}}$ is the P-wave velocity.

Take the curl on both sides for equation 2.25 and define vector S-wave potential $u_s = \nabla \times u_d$,

$$\frac{\partial^2 u_s}{\partial t^2} = \frac{\mu_{\text{iso}}}{\rho} \nabla^2 u_s = v_s^2 \nabla^2 u_s , \quad (2.27)$$

which is the S-wave equation and this is a vector equation. $v_s = \sqrt{\frac{\mu_{\text{iso}}}{\rho}}$ is the S-wave velocity.

2.1.4 Eikonal equation

Eikonal equation is a PDE for first-break travel time computation. In this thesis, the isotropic eikonal equation is considered.

Consider a plane wave solution,

$$u_p(x, t) = u_{p0}(x) \exp(-i\omega(t - t_0(x))) , \quad (2.28)$$

where u_{p0} is the amplitude function; t_0 is a phase function that describes the distribution of a constant phase.

Inset equation 2.28 into equation 2.26,

$$\nabla^2 u_{p0} + 2i\omega \nabla u_{p0} \cdot \nabla t_0 + i\omega u_{p0} \nabla^2 t_0 - \omega^2 u_{p0} \nabla t_0 \cdot \nabla t_0 + \omega^2 \frac{u_{p0}}{v_p^2} = 0 . \quad (2.29)$$

Assume $\omega \rightarrow \infty$, the eikonal equation can be obtained through the real part,

$$\nabla t_0 \cdot \nabla t_0 = \frac{1}{v_p^2} . \quad (2.30)$$

Equation 2.30 can be used to compute the travel time t_0 .

And the transport equation can be obtained in the imaginary part,

$$2\nabla u_{p0} \cdot \nabla t_0 + u_{p0} \nabla^2 t_0 = 0 . \quad (2.31)$$

Equation 2.31 can be used to compute amplitude distribution u_{p0} with given travel time t_0 .

The forward problem of Chapter 3 is based on the eikonal equation.

2.2 Fundamentals of waveform inversion

This section provides an example of implementing waveform inversion for the time domain acoustic wave equation (equation 2.26).

Consider an N_{dim} -dimensional space, $\Omega \subseteq \mathbb{R}^{N_{\text{dim}}}$. The author uses u as the state variable of the forward problem. For the acoustic wave equation (equation 2.26), $u = u_p$. The following objective function is to be minimized in the waveform inversion:

$$E(v, u) = \int_{-\infty}^{\infty} \int_{\Omega_r} (u(x, t) - u_{\text{obs}}(x, t))^2 dx dt \text{ s.t. } u(x, t) - f(v) = 0 , \quad (2.32)$$

where u_{obs} denotes the observed wavefield, Ω_r represents the receiver location.

The acoustic wave equation with a source term can be written as

$$\frac{1}{v(x)^2} \frac{\partial^2 u(x, t)}{\partial t^2} - \nabla^2 u = \delta(x - x_s) s(t) , \quad (2.33)$$

where $v = v_p$, $\delta(\cdot)$ is the Dirac delta function, x_s is the source location, and s is the source function.

Equation 2.33 has the following initial condition,

$$u(x, 0) = 0, x \in \Omega , \quad (2.34)$$

$$\frac{\partial u(x, 0)}{\partial t} = 0, x \in \Omega , \quad (2.35)$$

$$u(x, t) = 0, x \in \partial\Omega, 0 \leq t \leq T , \quad (2.36)$$

$$\nabla u(x, t) = 0, x \in \partial\Omega, 0 \leq t \leq T , \quad (2.37)$$

where the wavefield is simulated until time $t = T$.

Equation 2.37 can be written in a discretized form as:

$$\mathbf{A}(\mathbf{v})\mathbf{u} = \mathbf{b}_{\text{obs}} , \quad (2.38)$$

where $\mathbf{A}$ is the discrete version of $\frac{1}{v^2} \frac{\partial^2}{\partial t^2} - \nabla^2$, $\mathbf{u}$ is the discrete version of u , and $\mathbf{b}_{\text{obs}}$ is the discrete version of $\delta(x - x_s)s(t)$.

The waveform inversion solves a minimization problem. The corresponding objective function in the discrete form of equation 2.32 is written as,

$$E(\mathbf{v}, \mathbf{u}) = \|\mathbf{R}\mathbf{u} - \mathbf{d}_{\text{obs}}\|_2^2 \text{ s.t. } \mathbf{A}(\mathbf{v})\mathbf{u} - \mathbf{b}_{\text{obs}} = \mathbf{0} , \quad (2.39)$$

where $\mathbf{R}$ is the sampling operator, $\mathbf{u}$ is the vector composed of the discrete values of u , and $\mathbf{d}_{\text{obs}}$ is u_{obs} sampled at receiver locations.

The Lagrangian for the minimization of the objective function in equation 2.32 is given by,

$$\min_{\mathbf{v}, \mathbf{u}, \mathbf{u}_a} L(\mathbf{v}, \mathbf{u}, \mathbf{u}_a), \quad L = \frac{1}{2} \|\mathbf{R}\mathbf{u} - \mathbf{d}_{\text{obs}}\|_2^2 - \mathbf{u}_a^* (\mathbf{A}(\mathbf{v})\mathbf{u} - \mathbf{b}_{\text{obs}}) , \quad (2.40)$$

where $\mathbf{u}_a$ is the discrete adjoint wavefield, and $(\cdot)^*$ denotes the conjugate transpose.

The optimality condition for equation 2.40 requires,

$$\nabla_{\mathbf{v}} L = 0 , \quad (2.41)$$

$$\nabla_{\mathbf{u}} L = 0 , \quad (2.42)$$

$$\nabla_{\mathbf{u}_a} L = 0 . \quad (2.43)$$

Equation 2.43 gives equation 2.38, with the continuous form given in equation 2.33.

Equations 2.41 - 2.43 are the Karush–Kuhn–Tucker (KKT) conditions for the Lagrangian in equation 2.40. Haber and Ascher (2000) updates the $(\mathbf{v}, \mathbf{u}, \mathbf{u}_a)$ system using Newton's method. However, this is rather expensive to solve, as the matrix system is prohibitively large for real applications. In practice, the KKT conditions are updated successively: equation 2.43 for $\mathbf{u}$, equation 2.42 for $\mathbf{u}_a$, and equation 2.41 for $\mathbf{v}$.

The continuous form of the Lagrangian can be written as

$$L(v, u, u_a) = \int_0^T \int_{\Omega_r} \frac{1}{2} (u(x, t) - u_{\text{obs}}(x, t))^2 dx dt - \int_0^T \int_{\Omega} u_a(x, t) \left(\frac{1}{v(x)^2} \frac{\partial^2 u(x, t)}{\partial t^2} - \nabla^2 u(x, t) \right) dx dt \quad (2.44)$$

$$\begin{aligned} &= \int_0^T \int_{\Omega_r} (u - u_{\text{obs}})^2 dx dt \\ &\quad - \left(\int_{\Omega} \frac{1}{v^2} \frac{\partial u}{\partial t} u_a \Big|_0^T dx - \int_{\Omega} \frac{1}{v^2} \frac{\partial u_a}{\partial t} u \Big|_0^T dx + \int_0^T \int_{\Omega} \frac{1}{v^2} u \frac{\partial^2 u_a}{\partial t^2} dx dt \right. \\ &\quad - \int_0^T \int_{\partial\Omega} u_a \nabla u \cdot n_b dx dt \\ &\quad + \int_0^T \int_{\partial\Omega} u \nabla u_a \cdot n_b dx dt \\ &\quad \left. - \int_0^T \int_{\Omega} u \nabla^2 u_a dx dt \right) . \end{aligned} \quad (2.45)$$

where the integrate by parts is applied; u_a is the adjoint wavefield in the continous form, and n_b is the outward unit normal vector for the boundary of Ω .

With the following condition for u_a

$$u_a(x, T) = 0, x \in \Omega , \quad (2.46)$$

$$\frac{\partial u_a(x, T)}{\partial t} = 0, x \in \Omega , \quad (2.47)$$

$$u_a(x, t) = 0, x \in \partial\Omega, 0 \leq t \leq T , \quad (2.48)$$

$$\nabla u_a(x, t) = 0, x \in \partial\Omega, 0 \leq t \leq T . \quad (2.49)$$

equation 2.49 should run backwards in time.

Equation 2.45 can be simplified as

$$L(v, u, u_a) = \int_0^T \int_{\Omega_r} (u - u_{\text{obs}})^2 dx dt - \int_0^T \int_{\Omega} \left(\frac{1}{v^2} u \frac{\partial^2 u_a}{\partial t^2} - u \nabla^2 u_a \right) dx dt . \quad (2.50)$$

Take the variation of u in equation 2.45 and compute the difference between the following two terms,

$$\begin{aligned} L(v, u + \delta u, u_a) - L(v, u, u_a) &= \int_0^T \int_{\Omega_r} \delta u (u - u_{\text{obs}}) dx dt \\ &\quad - \int_0^T \int_{\Omega} \delta u \left(\frac{1}{v^2} \frac{\partial^2 u_a}{\partial t^2} - \nabla^2 u_a \right) dx dt . \end{aligned} \quad (2.51)$$

Thus,

$$\frac{\partial L}{\partial u} = \int_0^T \int_{\Omega_r} (u - u_{\text{obs}})^2 dx dt - \int_0^T \int_{\Omega} \left(\frac{1}{v^2} \frac{\partial^2 u_a}{\partial t^2} - \nabla^2 u_a \right) dx dt . \quad (2.52)$$

Equation 2.52 indicates:

$$\frac{1}{v^2} \frac{\partial^2 u_a}{\partial t^2} - \nabla^2 u_a = u - u_{\text{obs}}, x \in \Omega_r , \quad (2.53)$$

$$\frac{1}{v^2} \frac{\partial^2 u_a}{\partial t^2} - \nabla^2 u_a = 0, x \notin \Omega_r . \quad (2.54)$$

Take the variation of v in equation 2.45 and compute the difference between the following two terms,

$$\begin{aligned} L(v + \delta v, u, u_a) - L(v, u, u_a) &= - \int_0^T u_a(x, t) \\ &\quad \left(\frac{1}{(v(x) + \delta v(x))^2} \frac{\partial^2 u(x, t)}{\partial t^2} - \frac{1}{v(x)^2} \frac{\partial^2 u(x, t)}{\partial t^2} \right) dt \\ &\approx \int_0^T \frac{2\delta v(x)}{v^3} \frac{\partial^2 u(x, t)}{\partial t^2} u_a(x, t) dt . \end{aligned} \quad (2.55)$$

Equation 2.55 indicates

$$\frac{\partial L}{\partial v} = \int_0^T \frac{2}{v^3} \frac{\partial^2 u(x, t)}{\partial t^2} u_a(x, t) dt . \quad (2.56)$$

Now, the waveform inversion can be solved in a successive manner: solve the forward problem (equation 2.33) for u ; solve the adjoint equation (equations 2.53 and 2.54) for u_a ; solve the model update (equation 2.56) for the update of v , which can be expressed as

$$v_{i+1} = v_i - \alpha_{\text{step}} H_v \frac{\partial L}{\partial v} , \quad (2.57)$$

where i denotes the iteration, α_{step} is the step length, and H_v is the Hessian, which can be approximated using the quasi-Newton method.

The successive update for the three variables corresponds to the reduced space approach. The waveform u depends on v in the inversion process; thus, the cost function in equation 2.32 now becomes

$$E(\mathbf{v}) = \|\mathbf{R}\mathbf{u}(\mathbf{v}) - \mathbf{d}_{\text{obs}}\|_2^2 \text{ s.t. } \mathbf{A}(\mathbf{v})\mathbf{u} - \mathbf{b}_{\text{obs}} = \mathbf{0} . \quad (2.58)$$

However, as the problem is solved only in the v space, rather than in the (v, u) space as in equation 2.32, it suffers from cycle skipping when a crude initial iterate

is supplied or when the data frequency is high. Chapters 5 and 6 will discuss further details on how to mitigate this issue using the extended space approach.

To enforce regularity constraints on v , one can add a regularization term to the inverse problem, for example, in equation 2.32,

$$E(v, u) = \int_{-\infty}^{\infty} \int_{\Omega_r} (u(x, t) - u_{obs}(x, t))^2 + \alpha_{reg} h_{reg}(v) \, dx \, dt \quad (2.59)$$

$$\text{s.t. } u(x, t) - f(v) = 0 \, , \quad (2.60)$$

where α_{reg} controls the strength of the regularization, and $h_{reg}(v)$ is the regularization term.

There are various choices for h_{reg} . For instance, $h_{reg}(v) = \int_{\Omega} (v - v_{pri})^2 \, dx$ corresponds to a prior regularization with the reference model v_{pri} . The term $h_{reg}(v) = \int_{\Omega} |\nabla v|^2 \, dx$ represents the H^1 regularization, which is suitable for constructing smooth models. Meanwhile, $h_{reg}(v) = \int_{\Omega} |\nabla v| \, dx$ corresponds to total variation (TV) regularization, which is well-suited for constructing piecewise constant models. Chapter 4 discusses H^1 and TV regularizations applied to the eikonal equation. Chapter 6 explains how to combine model extension and TV regularization for waveform inversion problem.

Chapter 3

Wavefield simulation and reverse time migration for acoustic TTI media

The work presented in this chapter was published at: Zhang, Y., Stovas, A., and De Siena, L. (2024). Wavefield simulation and reverse time migration for acoustic TTI media. *GEOPHYSICS*, 1-56.

Y.Z. formulated the research idea, derived the equations, developed the code, conducted validation, drafted the manuscript, and handled revisions. A.S. verified the mathematical derivations and results, and reviewed the manuscript. L.D.S. supervised the research, reviewed the manuscript, and contributed to its improvement.

Abstract

Seismic imaging necessitates precisely separating P and S waves to mitigate undesirable crosstalk between them. Failure to properly handle this crosstalk can lead to distorted or misinterpreted images. In an elastic anisotropic medium, the polarization of a P wave is not necessarily parallel to its propagating direction, making the separation of P and S waves more difficult than in the case of an elastic isotropic medium. A pure acoustic Transversely Isotropic (TI) wave equation involves a mixed-domain problem. This usually means the space dependencies and wavenumbers are coupled in the wavefield extrapolator. An efficient solution to the wavefield requires a separation between the space dependencies and wavenumbers in the extrapolator. There are two primary approaches: (1) an analytical separation of the mixed-domain term; and (2) a numerical solution to the mixed-domain term, such as low-rank and Poynting vector methods. In this study, the acoustic approximation is derived by setting the S-wave phase velocity to zero in all phase directions. An analytical separation is developed in this study by employing a Taylor series to linearize

the anellipticity term, thereby decoupling the space-dependent and wavenumber-dependent variables within the wavefield extrapolator. The kinematics analysis demonstrates that our method has competitive phase velocity accuracy compared with existing methods. Moreover, the corresponding TI acoustic wave equation can be efficiently solved by the hybrid finite-difference/pseudospectral method. In a 2D case, it only requires one fast Fourier transform, two inverse Fourier transforms, and a few additional spatial differentiations at each time step. Remarkably, our wavefield solution remains independent of the wavefield heterogeneity (as observed in the low-rank approach) and is immune to wavefield interference (as encountered in the Poynting vector method). The developed acoustic approximation ensures the absence of any S-wave artifacts with relatively cheap computation. Numerical examples indicate our acoustic tilted transversely isotropic wave equation is reliable for accurate wavefield modeling and reverse time migration.

3.1 Introduction

The acoustic wave equation is extensively utilized in geophysics for tasks such as modeling, migration, and inversion, primarily to mitigate issues related to crosstalk between P- and S-waves and parameter tradeoff (Q. Li, Zhou, Zhang, Chen, & Sheng, 2015; Pestana et al., 2012; J. Yang & Zhu, 2018). Crosstalk refers to the situation in which the P- and S-wave components of both source and receiver wavefields are not entirely separated, resulting in interference between them. The challenge of parameter tradeoff arises because many inverse problems are ill-posed or underdetermined, demanding a delicate equilibrium between these parameters. Enhancing the estimation of one parameter often comes at the cost of the accuracy of others. The inverse problem for the elastic wave equation is inherently more intricate compared to the acoustic counterpart, making the elastic inverse problem more susceptible to issues related to parameter tradeoffs. The combination of crosstalk and parameter tradeoff can lead to the generation of artifacts, noise, and distortions in migrated images, particularly in anisotropic media, for which wave propagation is direction dependent (in the Christoffel equation, the wave speed and the polarization are functions of the propagating direction). In addition, the acoustic wavefield is generally cheaper to solve than the elastic one. To obtain the acoustic wavefield, the separation of P- and S-waves in an elastic medium is necessary. Anisotropy is pervasive across Earth media and ignoring it can lead to severe distortions and loss of resolution in the final model (Thomsen, 2002). The polarization of a qP-wave (P-wave to remain concise) is not necessarily parallel to the propagating direction in an anisotropic medium. Thus, differently from an isotropic medium, the anisotropic P-wavefield cannot be obtained by applying a Helmholtz decomposition to the elastic wavefield (Tsvankin, 2012).

The pure P-wave dispersion relation for anisotropic media introduces complexity as the space dependency and wavenumbers are coupled, making direct solutions of the resulting wave equation usually computationally expensive. Consequently,

there are two general approaches to solving the pure acoustic wavefield efficiently: (1) Analytical methods involve formulating an acoustic wave equation with reduced complexity by separating space dependency and wavenumbers analytically (Alkhalifah, 1998a, 2000; Chu et al., 2011); (2) Alternatively, specialized numerical methods can be employed to separate space and wavenumbers. Two of the most popular numerical methods are low-rank approximations (Cheng & Fomel, 2014; Engquist & Ying, 2009; Song & Alkhalifah, 2013) and the Poynting vector method (Xu & Zhou, 2014; L. Zhang & Ohen, 2022).

Under given approximations, the acoustic wave equation can be obtained analytically from the anisotropic elastic wave equation. For instance, by setting the vertical S-wave phase velocity to zero in the Vertical Transversely Isotropic (VTI) elastic dispersion relation, Alkhalifah (1998a) and Alkhalifah (2000) obtained an acoustic VTI wave equation with fourth-order derivatives in both time and space. However, this form produces S-wave artifacts because the order of the characteristic equation from the Christoffel equation is four; thus, it contains two wave modes corresponding to P-wave and S-wave artifacts. In practical applications, one always needs to taper the anellipticity η in the source region to suppress the S-wave artifacts. Following this, H. Zhou, Zhang, and Bloor (2006) extend the work to Tilt Transversely Isotropic (TTI) media, but it still contains S-wave artifacts and is not numerically stable with fast change of the tilt angle. To address the instability issue, Y. Zhang, Zhang, and Zhang (2011) proposed a stable acoustic TTI equation, derived from Hooke's law and the equation of motion, but it still contains S-wave artifacts. Another approach to obtain the pure P-wave information is to expand the term of the square root in the P-wave dispersion using Taylor series (Chu et al., 2011), which is free from S-wave artifacts. However, to make a feasible numerical solution (separation between space dependencies and wavenumbers), one additional assumption has to be made: Horizontal P-wave phase velocity v_{ph} and vertical P-wave phase velocity v_{pz} satisfy $v_{ph} \approx v_{pz}$. Pestana et al. (2012) used the same approach for 2D VTI reverse time migration (RTM), with a weaker assumption $\frac{v_{ph}^2 - v_{sz}^2}{v_{pz}^2 - v_{sz}^2} \approx \text{constant}$, where v_{sz} is the vertical S-wave phase velocity. A better approach to obtain the pure P-wave equation in TI media is proposed by Stovas et al. (2020). The idea is to set S-wave velocity in all the propagating directions to be zero in the coupled P-SV dispersion relations. A similar approach is also used here (Xu, Stovas, Alkhalifah, & Mikada, 2020). While also free from any S-wave artifacts, the mixed space-wavenumber anellipticity term burdens computation. B. Li and Stovas (2021a) propose an acoustic wave equation for TI media that has a simpler form of the anellipticity term, with space and wavenumbers decoupled. This is done by manipulating the S-wave velocity such that the material properties and the wavenumbers are no longer coupled in the anellipticity term. However, the S-wave velocity is not ensured zero in all the propagating directions and this approximation accuracy might be dependent on material parameters. This method is also used by B. Li and Stovas (2021b) for the acoustic approximation for orthorhombic media.

The above reviews some analytical approaches to derive a pure acoustic wave equation without coupling between space and wavenumbers. Numerical approaches can

also be employed to deal with the coupled space-wavenumber problem. For instance, Song and Alkhalifah (2013) solve the acoustic wavefield for orthorhombic media using the low-rank approach directly on the P-wave wavefield extrapolator, where the original P-wave wavefield extrapolator entails a coupling between material properties (space-dependent) and wavenumbers in the mixed domain. Utilizing the low-rank approach to address the space-wavenumber problem has gained popularity in the literature (Cheng, Alkhalifah, Wu, Zou, & Wang, 2016; Cheng & Fomel, 2014; Engquist & Ying, 2009; Fomel, Ying, & Song, 2013; Wu, Alkhalifah, & Zhang, 2019). The key advantage of this approach is its ability to circumvent numerical dispersion issues arising from the Laplacian operator. However, when dealing with more intricate media, higher ranks are often necessary to achieve a sufficiently accurate approximation. This requirement arises from the fact that the rank of a matrix (wavefield extrapolator) represents its number of degrees of freedom or complexity. In more complex media, the corresponding matrix for the low-rank approximation becomes more challenging, demanding higher ranks for an adequate representation. And each rank demands one Inverse Fast Fourier Transform (iFFT) in space.

Another notable numerical method is the Poynting vector method. Because the dispersion relation contains operations among wavenumbers, Xu and Zhou (2014) proposed to transform the wave vector k to the propagation direction vector n_p to reduce the computational complexity, where $n_p = \frac{k}{|k|}$. And n_p can be computed by $n_p = -\frac{\nabla u_p}{|\nabla u_p|}$ where u_p is the pressure wavefield. Though the original papers by Xu and Zhou (2014) did not mention the term Poynting vector, we name it the Poynting vector method as it adopts the idea of Poynting vector to compute the phase angle, similar to Schaeken, Hoogerbrugge, and Verschuur (2022) for a non-reflecting wave equation. L. Zhang and Ohen (2022) used the same approach and only k in the auxiliary variable v_{sz} of the wave equation was converted to n_p . The method of the Poynting vector is efficient as only the first-order derivatives are needed to compute the phase direction vector. However, the Poynting vector is not accurate for the approximation of the phase angle when the interference is strong as mentioned in Schaeken et al. (2022).

We have mentioned two famous numerical examples to solve the mixed-domain problem in the pure-P-wave equation for TI media, the low-rank approach and the Poynting method. The low-rank method is highly accurate if we keep enough ranks. But the number of iFFTs in space is dependent on the medium heterogeneity in the low-rank approach, which is not always cheap. The Poynting vector method may suffer from interference issues, though it is generally cheap. In this study, we seek a solution that remains independent of heterogeneity and is unaffected by interference. Our investigation begins with the acoustic VTI wave equation proposed by Stovas et al. (2020). The computational cost associated with the anellipticity term in the acoustic wave equation Stovas et al. (2020) is substantial using the hybrid finite-difference/pseudospectral method due to the presence of the mixed space-wavenumber term (anellipticity term). To address this, we leverage the assumption of weak anellipticity η and approximate the anellipticity term using a Taylor series. This approximation effectively decouples the space-dependent

material properties from the wavenumbers, eliminating the need for the low-rank approach or the Poynting vector method. Subsequently, we employ a hybrid finite-difference/pseudospectral method to solve the equation (Zhan, Pestana, & Stoffa, 2013). At each time step for a 2D case, only one Fast Fourier Transform (FFT) in space, two inverse Fast Fourier Transforms (iFFTs) in space, and additional spatial differentiations by the finite-difference method are required, ensuring computational efficiency. By employing this approach, we achieve a feasible cost in computational complexity while maintaining accurate and interference-free solutions.

In Section Theory, we show how we derive the pure acoustic wave equation for VTI media first. We demonstrate that the coupled space-wavenumber term, as in the anellipticity term by Stovas et al. (2020), causes some numerical difficulties in solving the wavefield. Then the anellipticity term is approximated by the Taylor series with decoupled space and wavenumbers. We propose a new acoustic TTI wave equation based on this. In Section Numerical Examples, the kinematics error is compared first between the proposed method and the methods by Chu et al. (2011), Y. Zhang et al. (2011), Pestana et al. (2012) and Stovas et al. (2020). Then we present the wavefield extrapolator comparisons between the low-rank method and our approach. Following this, we show the wavefield benchmark, between the proposed method and those methods (Alkhalifah, 2000; L. Zhang & Ohen, 2022), for the homogeneous Greenhorn shale. We also use a wedge model and a checkerboard model with multiple sources to test the stability and accuracy of our proposed method and that of L. Zhang and Ohen (2022). Following are some applications, such as RTM for a modified 2D Hess model and wavefield simulation and RTM for a modified 3D EAGE/SEG Salt model. Following this we discuss the importance of performing the space-wavenumber separation, the advantage of our proposed equation, and the necessity of including TTI anisotropic parameters in RTM.

3.2 Theory

3.2.1 A 2D acoustic VTI medium

We recall the 2D VTI elastic velocity-stress P-SV Christoffel wave equation in x_h - x_3 plane, symmetry axis parallel to x_3 -direction, and in (ω, k) domain (Carcione, 2007),

$$-\mathbf{i}\omega u_{\text{psv}}(\omega, k) = W u_{\text{psv}}(\omega, k), \quad (3.1)$$

where

$$u_{\text{psv}} = [\sigma_{hh} \quad \sigma_{33} \quad \sigma_{h3} \quad V_h \quad V_3]^T, \quad (3.2)$$

$$W = \begin{bmatrix} 0 & 0 & 0 & C_{11}\mathbf{i}k_h & C_{13}\mathbf{i}k_3 \\ 0 & 0 & 0 & C_{13}\mathbf{i}k_h & C_{33}\mathbf{i}k_3 \\ 0 & 0 & 0 & C_{44}\mathbf{i}k_3 & C_{44}\mathbf{i}k_h \\ \frac{1}{\rho}\mathbf{i}k_h & 0 & \frac{1}{\rho}\mathbf{i}k_3 & 0 & 0 \\ 0 & \frac{1}{\rho}\mathbf{i}k_3 & \frac{1}{\rho}\mathbf{i}k_h & 0 & 0 \end{bmatrix},$$

where ρ is the material density; C_{ij} is the material stiffness; σ_{ij} is the stress; V_i is the particle velocity; $k = [k_h \ k_3]^T$ is the wavenumber; ω is the angular frequency; $\mathbf{i} = \sqrt{-1}$.

As equation 3.1 has a non-zero solution for u_{psv} , we have the following characteristic equation,

$$\det(W + \mathbf{i}\omega\mathbb{I}) = 0 . \quad (3.3)$$

where $\mathbb{I}$ is the identity matrix.

Equation 3.3 can be written in the following form,

$$f_{\text{psv}} = v^4 + b_{\text{psv}}v^2 + c_{\text{psv}} = 0 , \quad (3.4)$$

with $v = \frac{\omega}{|k|}$ being the phase velocity and $|k|^2 = k_h^2 + k_3^2$.

Equation 3.4 has two solutions regarding v^2 , corresponding to P- and SV-wave modes,

$$v^2 = \frac{-b_{\text{psv}} \pm \sqrt{b_{\text{psv}}^2 - 4c_{\text{psv}}}}{2} , \quad (3.5)$$

This is the exact solution of the phase velocities and the sign $+$ is for P-wave and $-$ for SV-wave.

Now we introduce the density normalized stiffness, $A_{ij} = \frac{C_{ij}}{\rho}$, and

$$\eta = \frac{(A_{11} - A_{44})(A_{33} - A_{44}) - (A_{13} + A_{44})^2}{2((A_{33} - A_{44})A_{44} + (A_{13} + A_{44})^2)} . \quad (3.6)$$

Then we can parameterize equation 3.4 in A_{11} , A_{33} , A_{44} and η with

$$\begin{aligned} b_{\text{psv}} &= -[(A_{11} + A_{44})n_{\text{p},h}^2 + (A_{33} + A_{44})n_{\text{p},3}^2] , \\ c_{\text{psv}} &= A_{11}A_{44}n_{\text{p},h}^4 + \left(\frac{A_{11}A_{44} + 2\eta A_{11}A_{33}}{1 + 2\eta} + A_{33}A_{44} \right) n_{\text{p},h}^2 n_{\text{p},3}^2 + A_{33}A_{44}n_{\text{p},3}^4 , \end{aligned} \quad (3.7)$$

with $n_{\text{p},h} = \cos(\theta) = \frac{k_h}{|k|}$, $n_{\text{p},3} = \sin(\theta) = \frac{k_3}{|k|}$ and θ being the counterclockwise angle between the propagating direction and x_3 -direction. One classic analytical acoustic approximation (Alkhalifah, 2000) corresponds to $A_{44} = 0$ and contains S-wave artifacts because equation 3.4 still contains two wave modes, as shown in equation 3.5. A better way to approximate is to set the phase velocity of the S-wave to be zero along all propagating directions (Stovas et al., 2020). As the P-wave phase velocity is weakly dependent on A_{44} (Alkhalifah, 1998b, 2000; Tsvankin & Thomsen, 1994), we can select $A_{44} = \tilde{A}_{44}$ such that $c_{\text{p}} = c_{\text{psv}}(A_{44} = \tilde{A}_{44}) = 0$ and get a new $b_{\text{p}} = b_{\text{psv}}(A_{44} = \tilde{A}_{44})$. By doing so, the new $f_{\text{p}} = f_{\text{psv}}(A_{44} = \tilde{A}_{44}) = 0$ only has one propagating wave mode, which is the P-wave approximation. f_{p} is now given in the following form,

$$f_{\text{p}} = v^2 + b_{\text{p}} = 0 , \quad (3.8)$$

with

$$\begin{aligned} b_p &= -A_{11}n_{p,h}^2 - A_{33}n_{p,3}^2 + D_p(n_{p,h}^2 + n_{p,3}^2) , \\ D_p &= \frac{2\eta A_{11}A_{33}n_{p,h}^2n_{p,3}^2}{(1+2\eta)A_{11}n_{p,h}^4 + (A_{11} + (1+2\eta)A_{33})n_{p,h}^2n_{p,3}^2 + (1+2\eta)A_{33}n_{p,3}^4} . \end{aligned} \quad (3.9)$$

This acoustic approximation is identical to equation 12 in Stovas et al. (2020).

With equation 3.9, a formulation of the wave equation can be derived for P-wave pressure $u_p(t, x)$

$$\frac{\partial^2 u_p}{\partial t^2} = A_{11} \frac{\partial^2 u_p}{\partial x_h^2} + A_{33} \frac{\partial^2 u_p}{\partial x_3^2} - D_p(x, n_p) \nabla^2 u_p, \quad (3.10)$$

where $x = [x_h \ x_3]^T$ is the location coordinate and $n_p = [n_{p,h} \ n_{p,3}]^T$ is the phase direction vector. As the Poynting vector represents the phase direction of the dominant wavefield (Yoon, Guo, Cai, & Wang, 2011), n_p can be approximated by the Poynting vector. The Poynting vector for the acoustic wave equation is defined by $s_p = -\frac{\partial u_p}{\partial t} \nabla u_p$. And then, $n_p \approx \frac{s_p}{|s_p|} = -\frac{\nabla u_p}{|\nabla u_p|}$. Since the sign of n_p does not affect the solution of the wavefield, Xu and Zhou (2014) used a different sign for n_p . This method is only valid when the interference is negligible. The method has also been applied to solve the pure acoustic wave equation for TTI media (L. Zhang & Ohen, 2022).

Now, we will explain why the wave equation directly obtained by equation 3.8 is not cheap to solve by the hybrid finite-difference/pseudospectral method. We multiply k^2 on both sides of equation 3.8 and assign P-wave pressure $u_p(\omega, k)$,

$$\begin{aligned} (-i\omega)^2 u_p &= -b_p(i\mathbf{k})^2 u_p \\ &= (A_{11}(i\mathbf{k}_h)^2 + A_{33}(i\mathbf{k}_3)^2 - w_p) u_p , \end{aligned} \quad (3.11)$$

where

$$w_p = \frac{2\eta A_{11}A_{33}k_h^2k_3^2(k_h^2 + k_3^2)}{(1+2\eta)A_{11}k_h^4 + (A_{11} + (1+2\eta)A_{33})k_h^2k_3^2 + (1+2\eta)A_{33}k_3^4} . \quad (3.12)$$

Then under the assumption of weak variation of the material properties, the solution for $u_p(t, x)$ is,

$$\frac{\partial^2 u_p}{\partial t^2} = A_{11} \frac{\partial^2 u_p}{\partial x_h^2} + A_{33} \frac{\partial^2 u_p}{\partial x_3^2} - \frac{1}{(2\pi)^{N_{\text{dim}}/2}} \int_{\Omega_k} w_p(x, k) \mathbb{F}(u_p)(k) \exp(i\mathbf{k} \cdot \mathbf{x}) d\mathbf{k} , \quad (3.13)$$

where $\mathbb{F}$ corresponds to the fast Fourier transform in space (FFT), Ω_k is the full space of k , and $u_p(t, k) = \mathbb{F}(u_p(t, x)) = \frac{1}{2\pi^{N_{\text{dim}}/2}} \int_{\Omega_x} u_p(t, x) \exp(-i\mathbf{k} \cdot \mathbf{x}) d\mathbf{x}$ where Ω_x is the full space of x , and N_{dim} is the number of space dimensions.

The third term on the right-hand-side of equation 3.13 contains mixed (x, k) domain, and the integral operator is expensive to evaluate as generally one iFFT in space is required for each grid point.

A cheaper solution is the low-rank approximation to w and then the number of iFFTs in space is the same as the ranks. According to Song and Alkhalifah (2013), If one model exhibits serious roughness and randomness, the ranks will end up being large, thus making it computationally expensive.

3.2.2 An approximation to the anellipticity term

In this study, we choose the hybrid finite-difference/pseudospectral method. As of now, the only numerical difficulty is caused by w on the right-hand-side of equation 3.13, which is a function of coupled space (x) and wavenumber (k). We aim to decouple space-dependent material properties and wavenumbers in w_p so that

$$\begin{aligned}
 & \int_{\Omega_k} w_p(x, k) \mathbb{F}(u_p)(k) \exp(\mathbf{i}k \cdot x) dk \\
 & \approx \int_{\Omega_k} \sum_i^N w_{p,1i}(x) w_{p,2i}(k) \mathbb{F}(u_p)(k) \exp(\mathbf{i}k \cdot x) dk \\
 & = \sum_i^N w_{p,1i}(x) \int_{\Omega_k} w_{p,2i}(k) \mathbb{F}(u_p)(k) \exp(\mathbf{i}k \cdot x) dk \\
 & = (2\pi)^{N_{\text{dim}}/2} \sum_i^N w_{p,1i}(x) \mathbb{F}^{-1}(w_{p,2i}(k) \mathbb{F}(u_p)(k)) ,
 \end{aligned} \tag{3.14}$$

where $w_{p,1} = [w_{p,11} \ w_{p,12} \ \dots \ w_{p,1N}]^T$ only contains material properties, varying with space; $w_{p,2} = [w_{p,21} \ w_{p,22} \ \dots \ w_{p,2N}]^T$ only contains wavenumbers; $\mathbb{F}^{-1}$ corresponds to iFFT in space; N is the number of separated terms, as well as the number of the iFFTs required at each time step. The approximation presented in equation 3.14 is tantamount to approximating $D_p(n_{p,h}^2 + n_{p,3}^2)$ in equation 3.9 using the following expression,

$$D_p(x, n_p)(n_{p,h}^2 + n_{p,3}^2) \approx \sum_i^N D_{p,1i}(x) D_{p,2i}(n_p) . \tag{3.15}$$

In this study, we achieve the approximation in equation 3.15 using Taylor series under the assumption of weak anisotropy. This is equivalent to decoupling x and k in equation 3.14 analytically.

We construct a functional to approximate $D_p(x, n_p)(n_{p,h}^2 + n_{p,3}^2)$ in equation 3.9,

$$g(h_{p,1}, h_{p,2}, h_{p,3}) = \frac{q_4 n_{p,h}^2 n_{p,3}^2 (n_{p,h}^2 + n_{p,3}^2)}{h_{p,1} n_{p,h}^4 + h_{p,2} n_{p,h}^2 n_{p,3}^2 + h_{p,3} n_{p,3}^4} , \tag{3.16}$$

where $q_4 = 2\eta A_{11}A_{33}$ and $h_{p,1}$, $h_{p,2}$ and $h_{p,3}$ are scalars. Then

$$\begin{aligned}
 D_p(x, n_p)(n_{p,h}^2 + n_{p,3}^2) &= g(q_1, q_2, q_3) \\
 &\approx g(q_5, 2q_5, q_5) \\
 &\quad + \frac{\partial g(q_5, 2q_5, q_5)}{\partial h_{p,1}}(q_1 - q_5) + \frac{\partial g(q_5, 2q_5, q_5)}{\partial h_{p,2}}(q_2 - 2q_5) \\
 &\quad + \frac{\partial g(q_5, 2q_5, q_5)}{\partial h_{p,3}}(q_3 - q_5) \\
 &= \frac{q_4}{q_5} \frac{n_{p,h}^2 n_{p,3}^2}{n_{p,h}^2 + n_{p,3}^2} \\
 &\quad + \frac{q_4(q_5 - q_1)}{q_5^2} \frac{n_{p,h}^6 n_{p,3}^2}{(n_{p,h}^2 + n_{p,3}^2)^3} \\
 &\quad + \frac{q_4(2q_5 - q_2)}{q_5^2} \frac{n_{p,h}^4 n_{p,3}^4}{(n_{p,h}^2 + n_{p,3}^2)^3} + \frac{q_4(q_5 - q_3)}{q_5^2} \frac{n_{p,h}^2 n_{p,3}^6}{(n_{p,h}^2 + n_{p,3}^2)^3} \\
 &= \sum_i^N D_{p,1i}(x) D_{p,2i}(n_p), \tag{3.17}
 \end{aligned}$$

where $q_1 = (1 + 2\eta)A_{11}$, $q_2 = A_{11} + (1 + 2\eta)A_{33}$, $q_3 = (1 + 2\eta)A_{33}$, q_5 is a scalar to be chosen, $N = 4$, $D_{11} = \frac{q_4}{q_5}$, $D_{12} = \frac{q_4(q_5 - q_1)}{q_5^2}$, $D_{13} = \frac{q_4(2q_5 - q_2)}{q_5^2}$, $D_{14} = \frac{q_4(q_5 - q_3)}{q_5^2}$, $D_{21} = \frac{n_{p,h}^2 n_{p,3}^2}{n_{p,h}^2 + n_{p,3}^2}$, $D_{22} = \frac{n_{p,h}^6 n_{p,3}^2}{(n_{p,h}^2 + n_{p,3}^2)^3}$, $D_{23} = \frac{n_{p,h}^4 n_{p,3}^4}{(n_{p,h}^2 + n_{p,3}^2)^3}$, and $D_{24} = \frac{n_{p,h}^2 n_{p,3}^6}{(n_{p,h}^2 + n_{p,3}^2)^3}$. It is found that $q_5 = q_3$ results in a low approximation error. An additional advantage of doing this is that it cancels the term involving $(q_5 - q_3)$ on the right-hand-side of equation 3.17, which reduces the computational burden. Equation 3.17 successfully separates material properties and wave propagation directions, which means the condition of equation 3.14 is fulfilled when equation 3.17 is transformed into a wave equation. Now the P-wave kinematics with linearized anellipticity term is

$$\begin{aligned}
 v^2 &= A_{11}n_{p,h}^2 + A_{33}n_{p,3}^2 \\
 &\quad - \frac{q_4}{q_3} \frac{n_{p,h}^2 n_{p,3}^2}{n_{p,h}^2 + n_{p,3}^2} - \frac{q_4(q_3 - q_1)}{q_3^2} \frac{n_{p,h}^6 n_{p,3}^2}{(n_{p,h}^2 + n_{p,3}^2)^3} \\
 &\quad - \frac{q_4(2q_3 - q_2)}{q_3^2} \frac{n_{p,h}^4 n_{p,3}^4}{(n_{p,h}^2 + n_{p,3}^2)^3}. \tag{3.18}
 \end{aligned}$$

We can then transform equation 3.18 back to the (t, x) domain to get the corre-

sponding acoustic wave equation,

$$\begin{aligned}
 \frac{\partial^2 u_p}{\partial t^2} = & A_{11} \frac{\partial^2 u_p}{\partial x_h^2} + A_{33} \frac{\partial^2 u_p}{\partial x_3^2} \\
 & + \frac{q_4}{q_3} \frac{\partial^4}{\partial x_h^2 \partial x_3^2} F_1 \\
 & + \frac{q_4(q_3 - q_1)}{q_3^2} \frac{\partial^8}{\partial x_h^6 \partial x_3^2} F_2 \\
 & + \frac{q_4(2q_3 - q_2)}{q_3^2} \frac{\partial^8}{\partial x_h^4 \partial x_3^4} F_2 ,
 \end{aligned} \tag{3.19}$$

where $F_1 = \mathbb{F}^{-1} \left(\frac{1}{k_h^2 + k_3^2} \mathbb{F}(u_p) \right)$ and $F_2 = \mathbb{F}^{-1} \left(\frac{1}{(k_h^2 + k_3^2)^3} \mathbb{F}(u_p) \right)$.

Our proposed acoustic wave equation for a VTI media is given in equation 3.19. The spatial dependency and wavenumbers are decoupled. $q_i, i \in \{1, 2, \dots, 4\}$ only contains material properties. The wavenumber components are merged into the spatial derivatives for F_1 and F_2 . This separation is based on the preceding Taylor series, which performs competitively compared with some existing acoustic approximations for VTI media in the later accuracy analysis. It is also beneficial for the hybrid finite-difference/pseudospectral method due to the decoupled material properties and wavenumber components.

3.2.3 An acoustic TTI medium

With equation 3.18, one can approximate the acoustic kinematics in weakly anelliptic VTI media. Now we can expand it to a generalized TTI media by

$$\begin{bmatrix} n_{p,h} \\ n_{p,3} \end{bmatrix} \rightarrow \begin{bmatrix} \cos(\theta_{ti}) & -\sin(\theta_{ti}) \\ \sin(\theta_{ti}) & \cos(\theta_{ti}) \end{bmatrix} \begin{bmatrix} n_{p,h} \\ n_{p,3} \end{bmatrix} , \tag{3.20}$$

where θ_{ti} is the counterclockwise rotation angle of the symmetry axis of the TI media relative to the positive x_3 -axis.

One can apply the transform in equation 3.20 to equation 3.18 and obtain the

approximated P-wave kinematics for a TTI medium,

$$\begin{aligned}
v^2 = & (A_{11} \cos^2(\theta_{ti}) + A_{33} \sin^2(\theta_{ti})) n_{p,h}^2 \\
& + (A_{33} \sin(2\theta_{ti}) - A_{11} \sin(2\theta_{ti})) n_{p,h} n_{p,3} \\
& + (A_{33} \cos^2(\theta_{ti}) + A_{11} \sin^2(\theta_{ti})) n_{p,3}^2 \\
& + \frac{(q_4 - q_4 \cos(4\theta_{ti}))}{8q_3} \frac{n_1^4}{n_{p,h}^2 + n_{p,3}^2} \\
& + \frac{q_4 \sin(4\theta_{ti})}{2q_3} \frac{n_h^3 n_{p,3}}{n_{p,h}^2 + n_{p,3}^2} \\
& + \frac{q_4 + 3q_4 \cos(4\theta_{ti})}{4q_3} \frac{n_h^2 n_{p,3}^2}{n_{p,h}^2 + n_{p,3}^2} \\
& - \frac{q_4 \sin(4\theta_{ti})}{2q_3} \frac{n_{p,h} n_{p,3}^3}{n_{p,h}^2 + n_{p,3}^2} \\
& + \frac{q_4 - q_4 \cos(4\theta_{ti})}{8q_3} \frac{n_4^4}{n_{p,h}^2 + n_{p,3}^2} \\
& + \sum_{i=1}^9 Q_i \frac{n_{p,1}^{9-i} n_{p,3}^{i-1}}{(n_{p,h}^2 + n_{p,3}^2)^3},
\end{aligned} \tag{3.21}$$

where $Q_i, i \in \{1, 2, \dots, 9\}$ are given in Appendix A and they are only functions of space.

Equation 3.21 corresponds to a wave equation,

$$\begin{aligned}
\frac{\partial^2 u_p}{\partial t^2} = & (A_{11} \cos^2(\theta_{ti}) + A_{33} \sin^2(\theta_{ti})) \frac{\partial^2 u_p}{\partial x_h^2} \\
& + (A_{33} \sin(2\theta_{ti}) - A_{11} \sin(2\theta_{ti})) \frac{\partial^2 u_p}{\partial x_h \partial x_3} \\
& + (A_{33} \cos^2(\theta_{ti}) + A_{11} \sin^2(\theta_{ti})) \frac{\partial^2 u_p}{\partial x_3^2} \\
& + \frac{(q_4 - q_4 \cos(4\theta_{ti}))}{8q_3^4} \frac{\partial^4 F_1}{\partial x_h^4} \\
& + \frac{q_4 \sin(4\theta_{ti})}{2q_3^4} \frac{\partial^4 F_1}{\partial x_h^3 \partial x_3} \\
& + \frac{q_4 + 3q_4 \cos(4\theta_{ti})}{4q_3^4} \frac{\partial^4 F_1}{\partial x_h^2 \partial x_3^2} \\
& - \frac{q_4 \sin(4\theta_{ti})}{2q_3^4} \frac{\partial^4 F_1}{\partial x_h \partial x_3^3} \\
& + \frac{q_4 - q_4 \cos(4\theta_{ti})}{8q_3^4} \frac{\partial^4 F_1}{\partial x_3^4} \\
& + \sum_{i=1}^9 Q_i \frac{\partial^8 F_2}{\partial x_h^{9-i} \partial x_3^{i-1}}.
\end{aligned} \tag{3.22}$$

Equation 3.22 is the acoustic wave equation for TTI media we propose in this work. In the numerical solution, the spatial derivatives can be discretized by the finite-difference method. The 9-point centered finite-difference scheme is used for each spatial direction, regardless of the derivative order. We use a 3-point centered finite-difference scheme for time discretization. The evaluations for F_1 and F_2 require one FFT and two iFFTs in space. Thus, to solve equation 3.19, one needs the hybrid finite-difference/pseudospectral method.

To extend equation 3.22 to 3D, there are 3 steps: (1) We substitute $n_{p,h}^2 = n_{p,1}^2 + n_{p,2}^2$ into equation 3.18, where 1 and 2 are the x_1 - and x_2 - directions in 3D. As such, the 2D kinematics for a VTI medium of our proposed method is extended to 3D. (2) Another rotation transform is needed to derive the acoustic TTI kinematics,

$$\begin{bmatrix} n_{p,1} \\ n_{p,2} \\ n_{p,3} \end{bmatrix} \rightarrow \begin{bmatrix} \cos(\theta_{ti}) & 0 & \sin(\theta_{ti}) \\ 0 & 1 & 0 \\ -\sin(\theta_{ti}) & 0 & \cos(\theta_{ti}) \end{bmatrix} \begin{bmatrix} \cos(\phi_{ti}) & -\sin(\phi_{ti}) & 0 \\ \sin(\phi_{ti}) & \cos(\phi_{ti}) & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} n_{p,1} \\ n_{p,2} \\ n_{p,3} \end{bmatrix}, \quad (3.23)$$

where ϕ_{ti} is the azimuth angle of the symmetry axis relative to positive x_2 -axis, counterclockwise. (3) With the 3D acoustic TTI kinematics, the 3D acoustic TTI wave equation can be obtained using the same procedure as the 2D case. However, an explicit 3D acoustic TTI equation is not provided due to the inherent complexity involved in its expression.

At each time step, the finite-difference scheme expands the computational domain 4 grids in all of the axis directions and the pseudospectral method expands the computational grids infinitely in space. So, the bottom line of the stability is dependent on the finite-difference method. According to Kelly, Ward, Treitel, and Alford (1976) and Alkhalifah (2000), the Courant-Friedrichs-Lewy (CFL) condition of our hybrid finite-difference/pseudospectral is $\frac{\sqrt{2} \max(v)\Delta t}{4 \min(\Delta x_h, \Delta x_3)} < 1$ for 2D and $\frac{\sqrt{3} \max(v)\Delta t}{4 \min(\Delta x_1, \Delta x_2, \Delta x_3)} < 1$ for 3D, where Δt is the time grid spacing and $\Delta x_i, i \in \{h, 1, 2, 3\}$ is the spatial grid spacing.

3.2.4 The true 2D acoustic TTI solution

For benchmarking, we also need to compare our proposed method and some existing methods together with the exact solution of the acoustic TTI wavefield. We start from equation 3.5 with sign + for P-wave mode. Plugging equation 3.20 into equation 3.5, we obtain the exact solution of the acoustic TTI kinematics. Similar to the process to obtain our proposed TTI acoustic wave equation, the acoustic TTI wave equation for the exact solution is derived. The exact solution of $u_p(t, x)$ is given by,

$$\frac{\partial^2 u_p}{\partial t^2} = \frac{1}{(2\pi)^{N_{\text{dim}}/2}} \int_{\Omega_k} H_p(x, k) \mathbb{F}(u_p)(k) \exp(\mathbf{i}k \cdot x) dk, \quad (3.24)$$

where

$$\begin{aligned}
 H_p &= \frac{\tilde{b} - \sqrt{\tilde{b}^2 - 4\tilde{c}}}{2}, \\
 \tilde{b} &= - (A_{11} + A_{44})(k_h \cos(\theta_{ti}) - k_3 \sin(\theta_{ti}))^2 \\
 &\quad + (A_{33} + A_{44})(k_h \sin(\theta_{ti}) + k_3 \cos(\theta_{ti}))^2, \\
 \tilde{c} &= A_{11}A_{44}(k_h \cos(\theta_{ti}) - k_3 \sin(\theta_{ti}))^4 + \left(\frac{A_{11}A_{44} + 2\eta A_{11}A_{33}}{1 + 2\eta} + A_{33}A_{44} \right) \\
 &\quad (k_h \cos(\theta_{ti}) - k_3 \sin(\theta_{ti}))^2 (k_h \sin(\theta_{ti}) + k_3 \cos(\theta_{ti}))^2 \\
 &\quad + A_{33}A_{44}(k_h \sin(\theta_{ti}) + k_3 \cos(\theta_{ti}))^4,
 \end{aligned} \tag{3.25}$$

In an arbitrary medium, an iFFT in space is required for each grid point of x . However, normally, there are some grid points with the same material properties, and only one iFFT in space is needed for the region with identical material properties. Equation 3.24 refers to the exact acoustic TTI wave equation. H_p is the wavefield extrapolator and can be used for accuracy analysis (Details given in Numerical Examples).

3.3 Numerical examples

3.3.1 A VTI kinematics benchmark

We initiate our study by examining the P-wave kinematics as defined in equation 3.18. To provide comprehensive comparisons, we reference several other works (Chu et al., 2011; Pestana et al., 2012; Stovas et al., 2020; Y. Zhang et al., 2011). In particular, we adopt equation 12 from Stovas et al. (2020), equation 20 from Chu et al. (2011), equation 14 from Y. Zhang et al. (2011), and equation 22 from Pestana et al. (2012) for calculating P-wave phase velocities, aligning our methodology with these established references. Equation 3.18 offers our solution for the P-wave phase velocity in VTI media.

Greenhorn shale (Jones & Wang, 1981) is used to benchmark the accuracy of the kinematics of our proposed solution and some other solutions (Figure 3.1). The relative error of the estimated P-wave phase velocity is computed by $E_r = \frac{|v_{\text{est}} - v_{\text{exa}}|}{|v_{\text{exa}}|}$, where v_{est} is the estimated solution and v_{exa} is the exact solution. It is shown that our proposed solution has an accuracy similar to that proposed by Stovas et al. (2020), both of which perform better than the solution by Pestana et al. (2012), Chu et al. (2011), and Y. Zhang et al. (2011). We further test the proposed phase velocity solution for 1000 randomly generated samples. The relative errors along all directions are stacked over those samples (Figure 3.2). It is found that our proposed method retains high accuracy among the other compared methods.

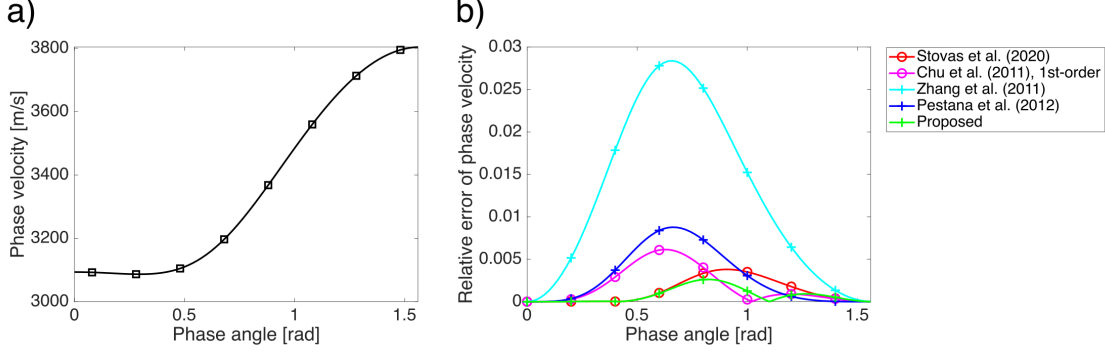


Figure 3.1: The P-wave kinematics for Greenhorn shale. $A_{11} = 1.447 \times 10^7$ Pa m³/kg, $A_{33} = 9.57 \times 10^6$ Pa m³/kg, $A_{13} = 4.51 \times 10^6$ Pa m³/kg, $A_{44} = 2.28 \times 10^6$ Pa m³/kg, thus, $\eta = 0.3409$. (a) Exact P-wave phase velocity. (b) The relative error of the estimated P-wave phase velocity E_r .

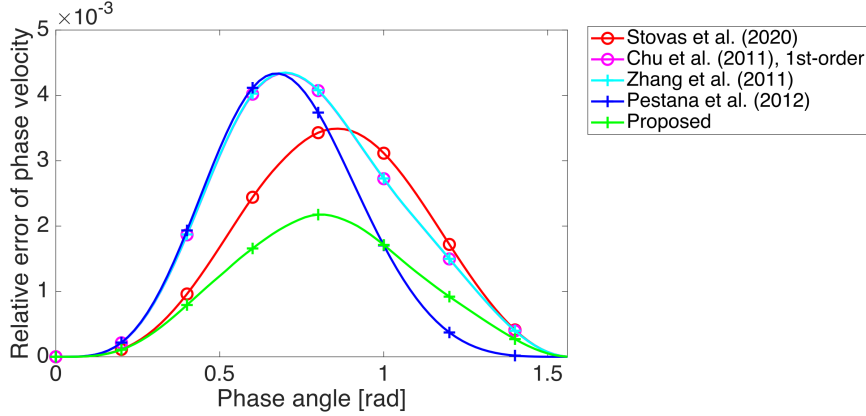


Figure 3.2: The average stacked relative error of the estimated P-wave phase velocity for randomly generated 1000 VTI samples using a uniform distribution. The stacking is performed for each phase angle. $A_{11} \sim U_{\text{dis}}(1.1 \times 10^7 \text{ Pa m}^3/\text{kg}, 1.7 \times 10^7 \text{ Pa m}^3/\text{kg})$, $A_{33} \sim U_{\text{dis}}(7 \times 10^6 \text{ Pa m}^3/\text{kg}, 1.25 \times 10^7 \text{ Pa m}^3/\text{kg})$, $A_{13} \sim U_{\text{dis}}(1.5 \times 10^6 \text{ Pa m}^3/\text{kg}, 3 \times 10^6 \text{ Pa m}^3/\text{kg})$, $A_{44} \sim U_{\text{dis}}(3.5 \times 10^6 \text{ Pa m}^3/\text{kg}, 4.5 \times 10^6 \text{ Pa m}^3/\text{kg})$. U_{dis} is for uniform distribution. Thus $\eta \in [-0.2475, 0.7100]$ in our experiment.

3.3.2 A wavefield extrapolator comparison to the low-rank approach

The example in this subsection illustrates a wavefield extrapolator comparison between our approach and the low-rank method. We have selected a velocity model with varying properties along the x_3 -axis (see Figure 3.3). Solving for the true acoustic wavefield in a general heterogeneous TTI medium is infeasible, so we focus on computing the wavefield extrapolator H_p at specific locations (Figure 3.4), as defined by equation 3.24.

In our experiment, we set the rank to 4 for both x and k (Fomel et al., 2013). The rank was adjusted until the low-rank approximation demonstrated performance comparable to our proposed method: the maximum absolute value of the relative error for both methods falls within a similar magnitude. Figure 3.4 shows the accuracy of the wavefield extrapolator by our proposed method and the low-rank method. The relative error is computed using $\frac{H_{p,app} - H_{p,true}}{|H_{p,true}|}$, where $H_{p,app}$ is an approximation of the wavefield extrapolator H_p and $H_{p,true}$ is the true solution. Our proposed method evidently outperforms the low-rank approach at shallower depths but becomes less competitive at greater depths. For solving the corresponding wavefields by the two methods, 4 iFFTs are necessary in the low-rank approximations, and only 2 iFFTs are needed in our proposed method.

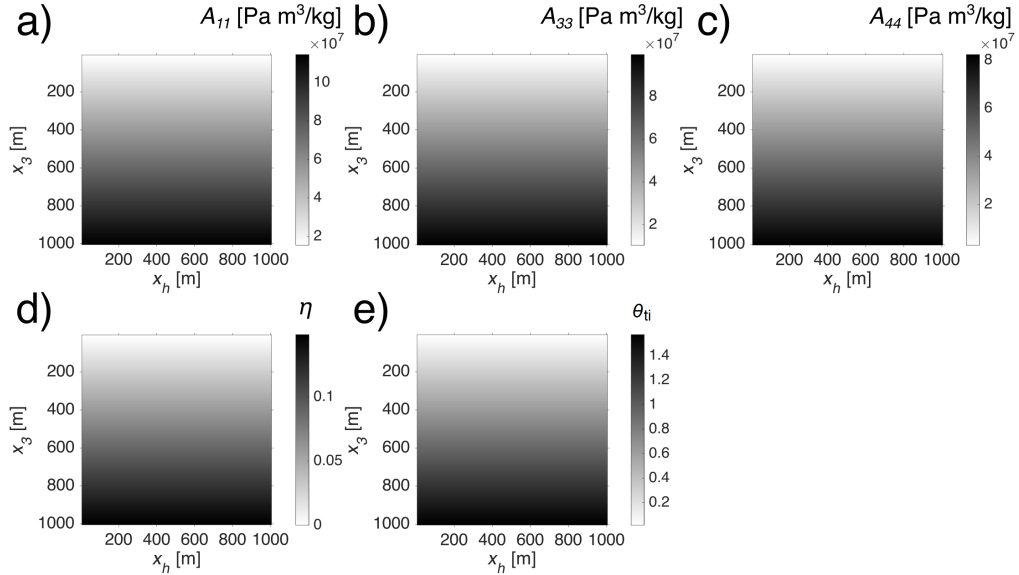
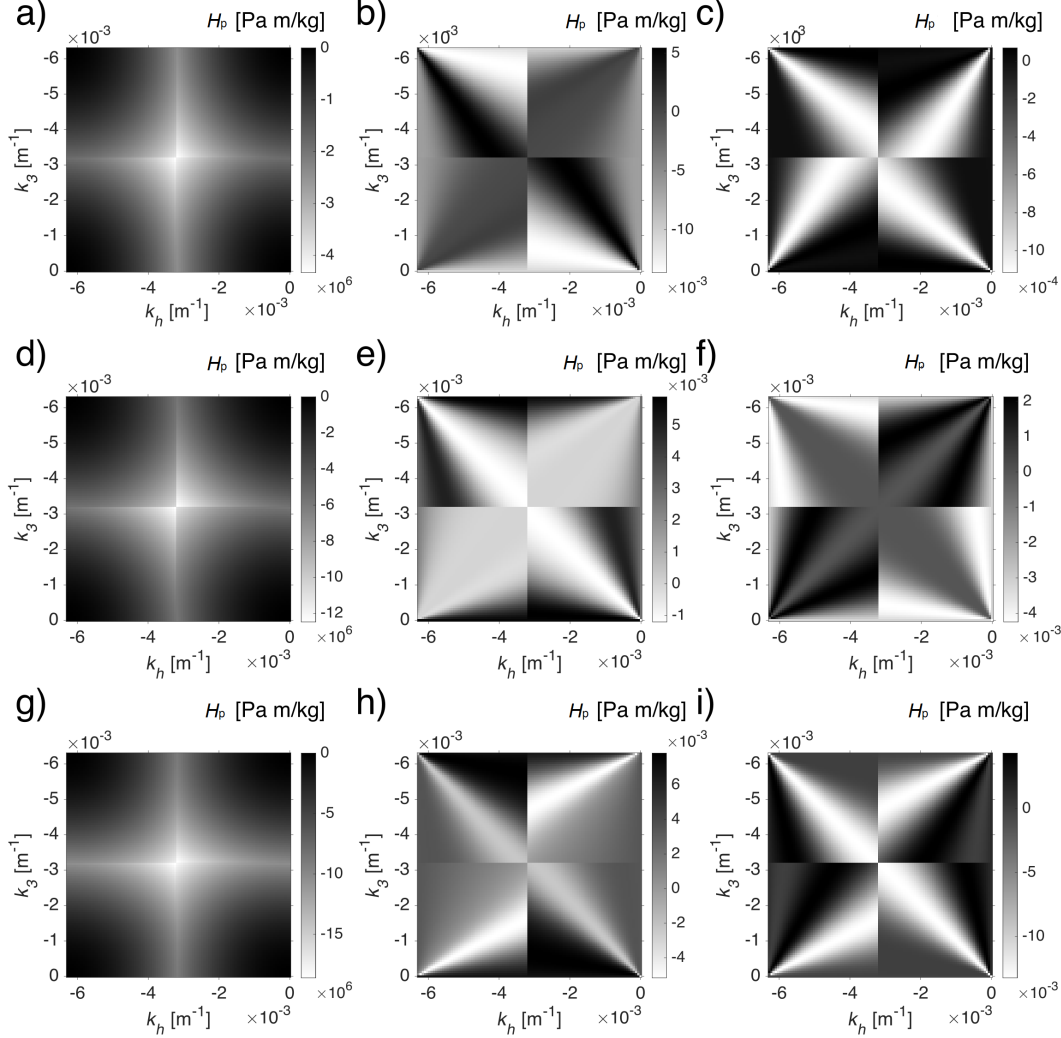


Figure 3.3: The model utilized in the numerical experiment for comparing the wavefield extrapolator of the proposed method with the low-rank approach. The model is only a function of x_3 .



3.3.3 A wavefield test for the Greenhorn shale

We assessed the accuracy of wavefield simulations in a homogeneous medium comprising Greenhorn shale, comparing our method to other existing approaches (Alkhalifah, 2000; L. Zhang & Ohen, 2022). In Figure 3.5b, it is evident that all methods provide accurate approximations of the acoustic wavefield, except for the case in Figure 3.5 from Alkhalifah (2000), which exhibits S-wave artifacts. Though S-wave artifacts are typically minor in scenarios where the source is located in a weakly anelliptic medium, eliminating them remains vital for stable acoustic wavefield simulations in TTI media with varying tilt angles θ_{ti} . Our proposed method, as depicted in Figure 3.6, showcases the highest accuracy for this particular profile.

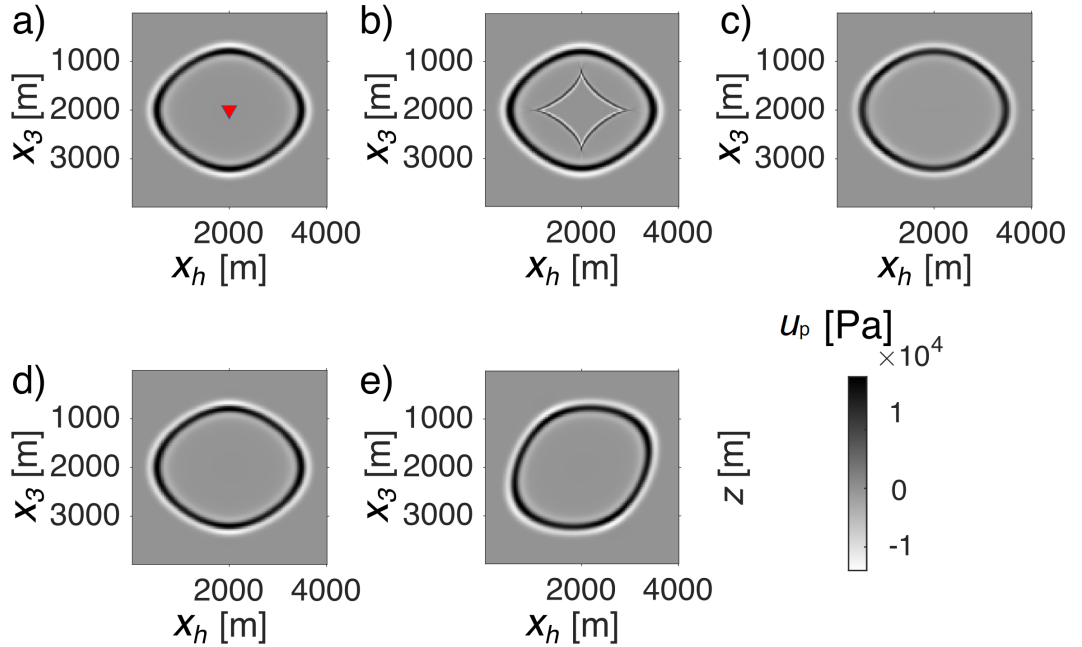


Figure 3.5: The acoustic wavefield for the Greenhorn shale at 0.5 s. The red down-pointing triangle depicted in (a) serves as the source and will retain the same marker throughout. The source signature is Ricker wavelet, 10 Hz. (a) Exact solution. (b) VTI wavefield by Alkhalifah (2000). (c) VTI wavefield by L. Zhang and Ohen (2022). (d) Proposed VTI solution. (e) Proposed TI solution with $\theta_{ti} = \frac{\pi}{6}$.

3.3.4 A wedge model

To demonstrate the stability of our proposed TTI acoustic wave equations under varying θ_{ti} and negative η , we generated wavefields for a wedge-shaped medium and compared our method to L. Zhang and Ohen (2022). The acoustic TTI wavefield

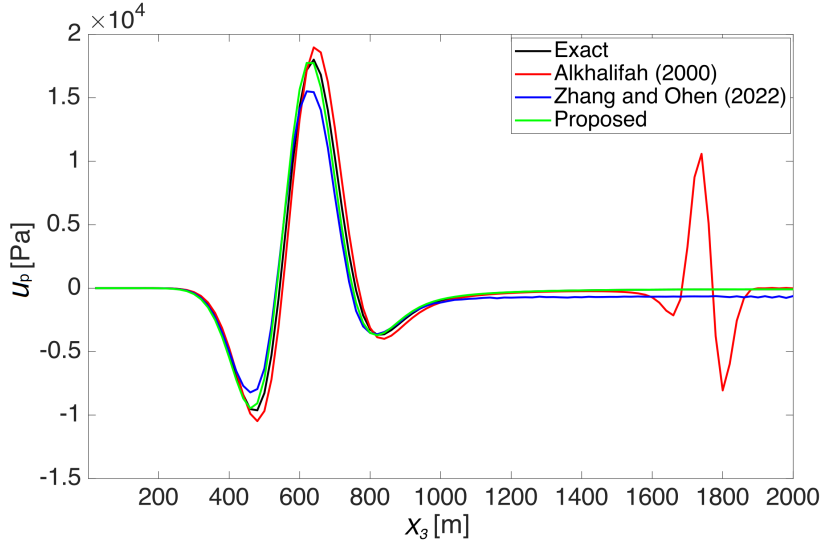


Figure 3.6: The profile at $x_h = 1600$ m for Figure 3.5 a-d.

derived from Alkhalifah (2000) is not shown due to instability. The wedge medium is depicted in Figure 3.7 and contains negative η (see Figure 3.7a).

Figure 3.8 shows that both our method and the approach by L. Zhang and Ohen (2022) maintain stable solutions on this wedge model, even in the presence of negative η and rapid changes in θ_{ti} . The shot gathers in Figure 3.9 also demonstrate that both of the two approximations are reliable while our proposed method is more accurate than L. Zhang and Ohen (2022) according to Figure 3.9d.

We also benchmarked computational time for this model, using a spatial discretization of 200×200 and 5000 time steps. On a workstation equipped with an AMD EPYC 7302P CPU and 128 GB of RAM, L. Zhang and Ohen (2022) took 91 seconds per shot, while our proposed method required 103 seconds. The method by L. Zhang and Ohen (2022) proves slightly more time-efficient.

3.3.5 An interference test

As previously mentioned, the Poynting vector provides an approximation of the dominant phase directions for the acoustic wavefield. Methods, such as those discussed in L. Zhang and Ohen (2022) and Xu and Zhou (2014), may exhibit reduced accuracy in the presence of strong interference. In this test, we demonstrate that our method is not affected by this issue.

We created a scenario where three sources were simultaneously activated (referred to as u_p for this case). The wavefield was calculated for each source and then stacked together as $u_{p,sta}$. When interference does not affect the solution, $u_{p,sta}$ equals u_p . We conducted this experiment using the model shown in Figure 3.10, where the

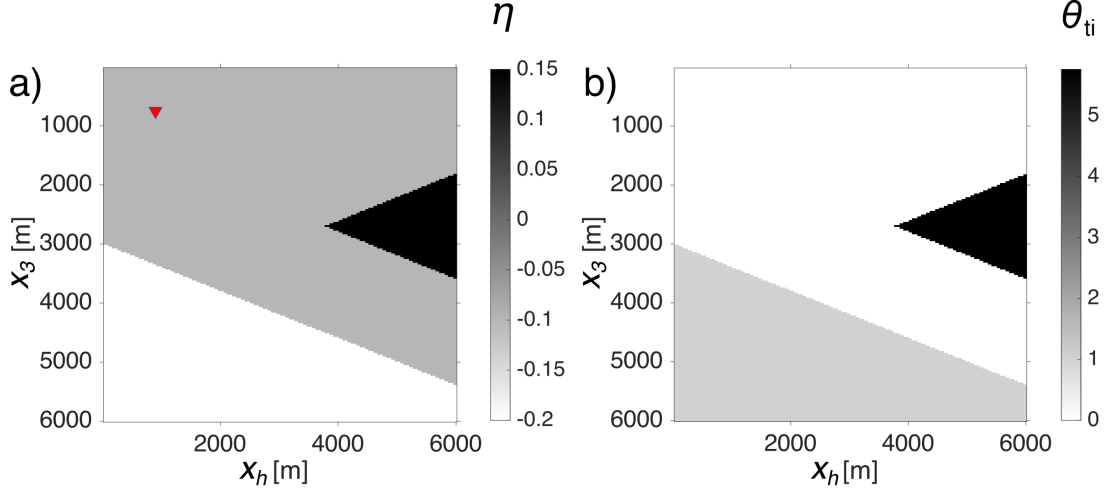


Figure 3.7: The wedge model. (a) η . (b) θ_{ti} . The other parameters are set homogeneous along the space. $A_{11} = 1.4 \times 10^7$ Pa m³/kg, $A_{33} = 1.0 \times 10^7$ Pa m³/kg, $A_{44} = 2.0 \times 10^6$ Pa m³/kg. 30 receivers are placed evenly at $x_h \in [780, 5130]$ m and $x_3 = 750$ m.

three sources can induce wavefield interference.

We compared our proposed method with one of the Poynting methods, L. Zhang and Ohen (2022). Figure 3.11 displays the shot gather, showcasing both u_p and $u_{p,sta}$. Generally, both methods yield similar wavefield solutions. However, a slight interference issue is observed in Figure 3.11f for L. Zhang and Ohen (2022), whereas our proposed method demonstrates identical seismograms for u_p and $u_{p,sta}$.

3.3.6 An application to 2D RTM

To show the reliability of the proposed equation, we use a modified TTI version of the Hess model (Figure 3.12). The imaging condition is calculated using

$$I_{R,3}(x) = \frac{1}{N_{source}} \sum_{i=1}^{N_{source}} \frac{\int_{\Omega_t} u_{p,u}^i(x, t) u_{p,d}^i(x, t) dt}{\int_{\Omega_t} u_{p,d}^i(x, t)^2 dt}, \quad (3.26)$$

where $I_{R,3}$ is the reflectivity image; Ω_t is the full space of t ; i is the source index; N_{source} is the source numbers; $u_{p,u}$ is the upgoing wavefield and $u_{p,d}$ is the downgoing wavefield. The observed data is computed using our proposed acoustic TTI equation from the model in Figure 3.12. The migration model is the same. For TTI RTM, both $u_{p,d}$ and $u_{p,u}$ are calculated using the model in Figure 3.12, but $\theta_{ti} = 0$ for VTI RTM and $\eta = 0$ for elliptic RTM.

Figure 3.13b displays the migration results. To evaluate the significance of the tilt angle and anellipticity, we introduced several imperfections in the migration

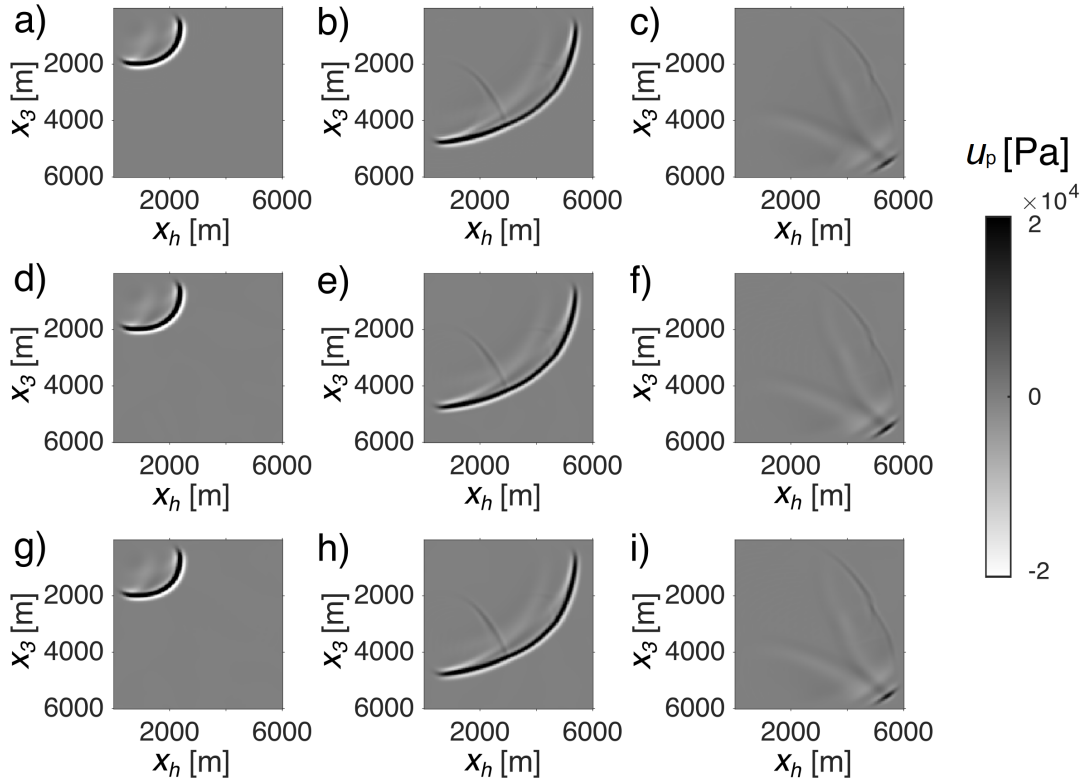


Figure 3.8: The acoustic wavefield for the model in Figure 3.7. The source signature is a Ricker wavelet, 10 Hz. Acoustic approximations (a)(d)(g) are at 0.5 s, (b)(e)(h) at 1 s and (c)(f)(i) at 1.8 s. (a) - (c) are computed by the method of L. Zhang and Ohen (2022) and (d) - (f) are from our proposed method. (g) - (i) are exact solutions.

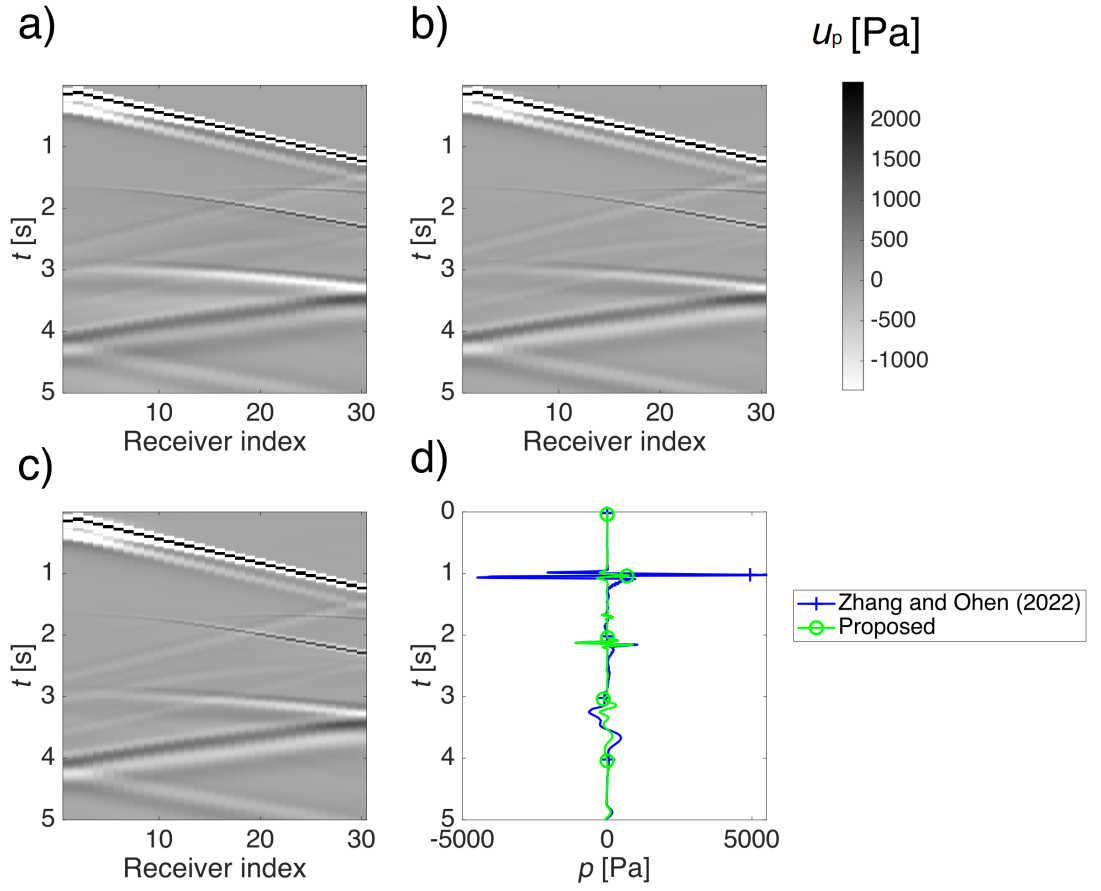


Figure 3.9: The shot gathers for the wavefield in Figure 3.8 and the setup in Figure 3.7, (a) for L. Zhang and Ohen (2022), (b) for the proposed method, (c) for the exact solution, and (d) differences between the approximations and the true solution for the receiver index 25.

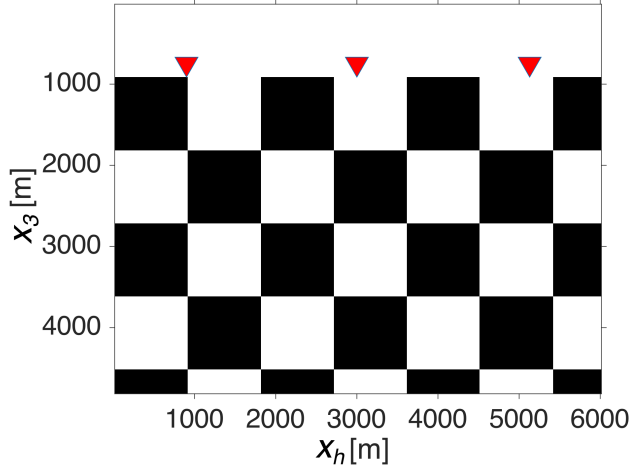


Figure 3.10: The checkerboard model. For white region, $A_{11} = 1.4 \times 10^7$ Pa m³/kg, $A_{33} = 1.0 \times 10^7$ Pa m³/kg, $A_{44} = 7 \times 10^6$ Pa m³/kg, $\eta = 0.2$, $\theta_{ti} = \frac{\pi}{3}$. For the black region, the values of the above parameters are doubled.

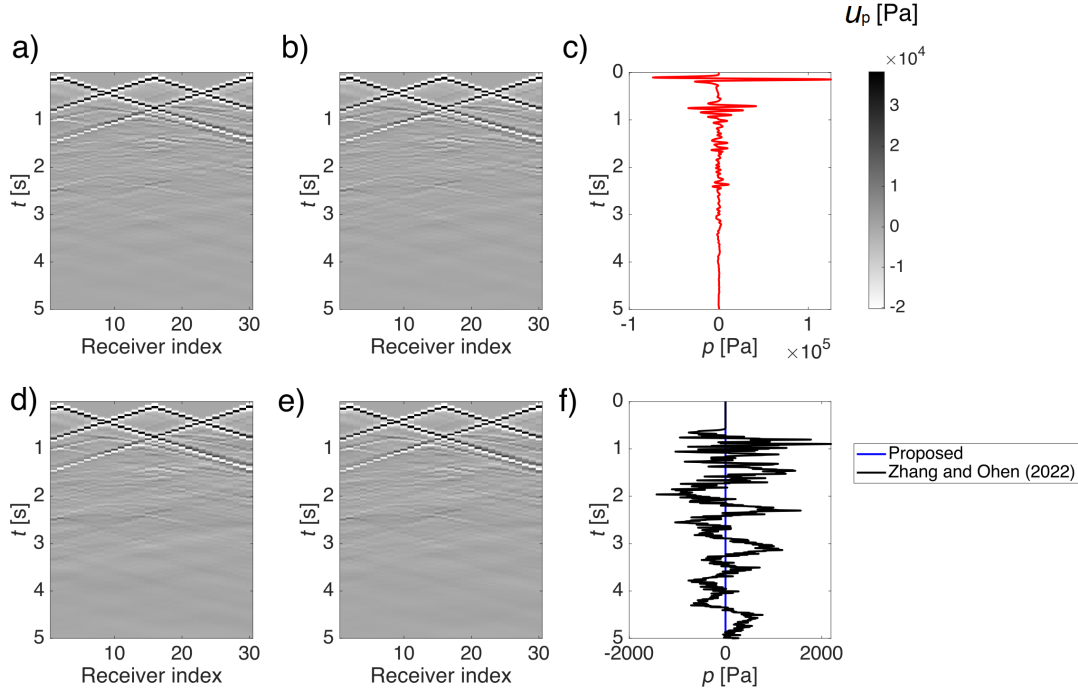


Figure 3.11: The shot gathers for Figure 3.10. Ricker wavelet, 10 Hz. u_p is the solution when 3 sources are shot simultaneously, and $u_{p,sta}$ is the stacked solution for each shot wavefield, (a) u_p by our proposed method, (b) $u_{p,sta}$ by our proposed method, (c) True seismograms for receiver index 15, (d) u_p by L. Zhang and Ohen (2022), (e) $u_{p,sta}$ by L. Zhang and Ohen (2022), and (f) $u_{p,sta} - u_p$ for L. Zhang and Ohen (2022) and the proposed method.

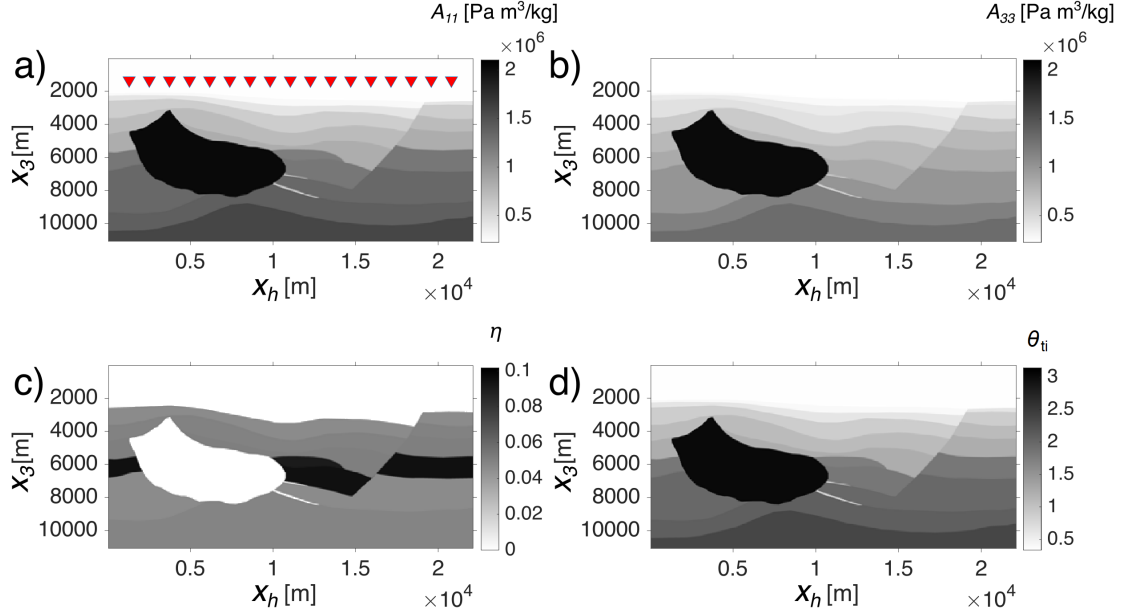


Figure 3.12: The modified Hess model. (a) A_{11} . (b) A_{33} . (c) η . (d) An additional tilt angle θ_{ti} of the symmetry axis is added to the original Hess model. 201 receivers are placed evenly at $x_h \in [1281, 25681]$ m and $x_3 = 1342$ m.

processes in Figure 3.13c and d. For reference, Figure 3.13a represents the vertical differentiation of the model A_{33} , revealing the true interfaces.

Notably, the TI migration in Figure 3.13b produces the clearest image. However, due to inaccurate wavefield simulations, Figure 3.13c and 3.13d exhibit more migration noise and a loss of reflector accuracy.

3.3.7 A 3D example

This subsection gives the wavefield simulation and RTM for the modified 3D EAGE/SEG Salt model. Our proposed 2D acoustic TTI wave equation is extended to 3D using the procedure given in Subsection An acoustic TTI medium. The EAGE/SEG Salt model is modified. As the original 3D EAGE/SEG Salt model is isotropic, we add to its complexity by introducing some anisotropic TTI parameters (Figure 3.14). We obtain our acoustic TTI model by $A_{11} = v^2$, $A_{33} = 1.4A_{11}$, $\eta = 0.2 \frac{A_{11}}{\max(A_{11})}$, $\theta_{ti} = \frac{\pi}{2} \frac{A_{11}}{\max(A_{11})}$, $\phi_{ti} = 2\pi \frac{A_{11}}{\max(A_{11})}$, where v is the original isotropic velocity of 3D EAGE/SEG model. Figure 3.15 is an example of the wavefield and Figure 3.16 is the RTM image. For RTM, the observed data is computed from the model in Figure 3.14. The migration model stays the same.

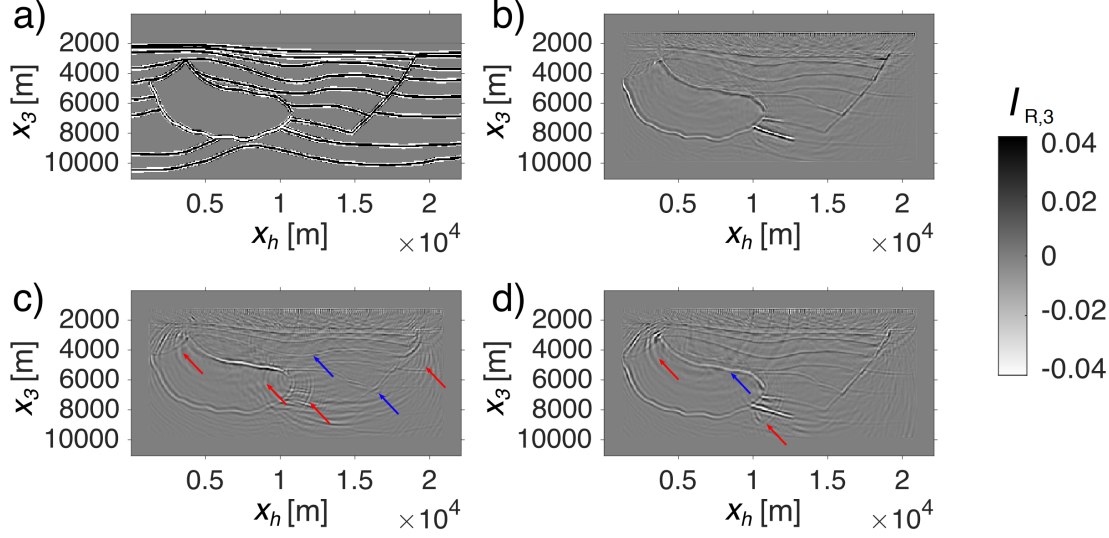


Figure 3.13: The reflectivity images $I_{R,3}$. Source: Ricker wavelet, 5 Hz. 17 sources are placed evenly at $x_h \in [1281, 20862] m$, $x_3 = 1342m$, and each shot gather is generated using a single source. (a) is the image derived from the vertical differentiation of A_{33} . (b)(c) and (d) are for TTI, VTI, and elliptic migration images, respectively. In (c) and (d), certain artifacts (indicated by red arrows) are generated, and some reflectors are inaccurately constructed (shown in blue) due to inappropriate physics.

3.4 Discussions

We propose a new acoustic TTI wave equation. We start from the acoustic approximation by Stovas et al. (2020) for VTI media. The Christoffel equation of the new acoustic wave equation only contains the P-wave propagating mode, so it is free from S-wave artifacts. However, the acoustic wave equation by Stovas et al. (2020) has a coupled (x, k) anellipticity term, so it is not cheap to solve with the hybrid finite-difference/pseudospectral method. A treatment of the anellipticity term using Taylor series can decouple the space and wavenumber dependencies analytically, then the hybrid finite-difference/pseudospectral method can be applied. For a 2D case, only one FFT in space, two iFFTs in space, and additional spatial differentiations using the finite-difference method are required at each time step. This method is not dependent on media heterogeneity and is not prone to wavefield interference.

In the kinematics analysis and the wavefield extrapolator comparison, we show that our proposed method has a competitive accuracy. Followed by the stability test of a wedge model and the interference test, our proposed method is stable under fast change of θ_{ti} , negative η , and is not affected by the interference issues. Those factors are key for an accurate and stable acoustic TTI wavefield simulation. We compare the computational cost of our proposed method with one of the cheap Poynting vector methods. Our proposed method is not much more expensive than the efficient Poynting vector method while retaining more accuracy in certain circumstances.

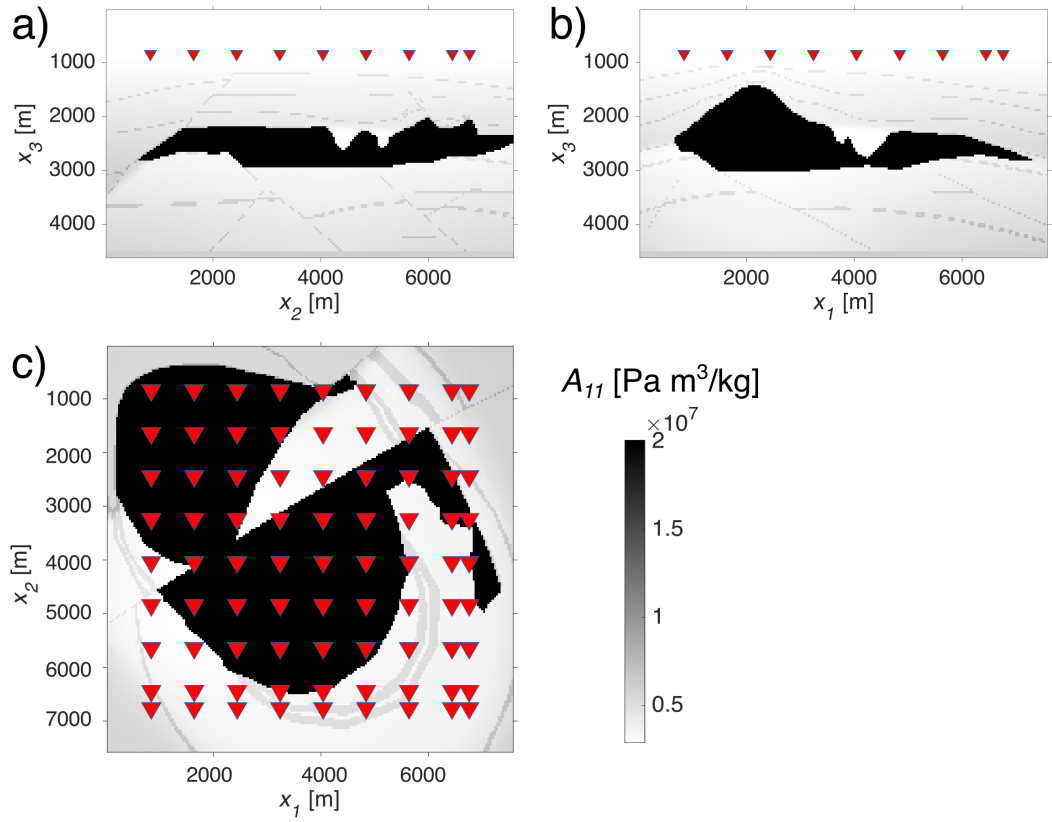


Figure 3.14: The modified 3D EAGE/SEG Salt model. (a) Slice at $x_1 = 4400$ m. (b) Slice at $x_2 = 4400$ m. (c) Slice at $x_3 = 2800$ m. The down-pointing triangles are the imaging source locations projected on those planes. 421875 receivers are evenly distributed in this range, $x_1 \in [840, 6760]$ m, $x_2 \in [840, 6760]$ m, $x_3 = 840$ m.

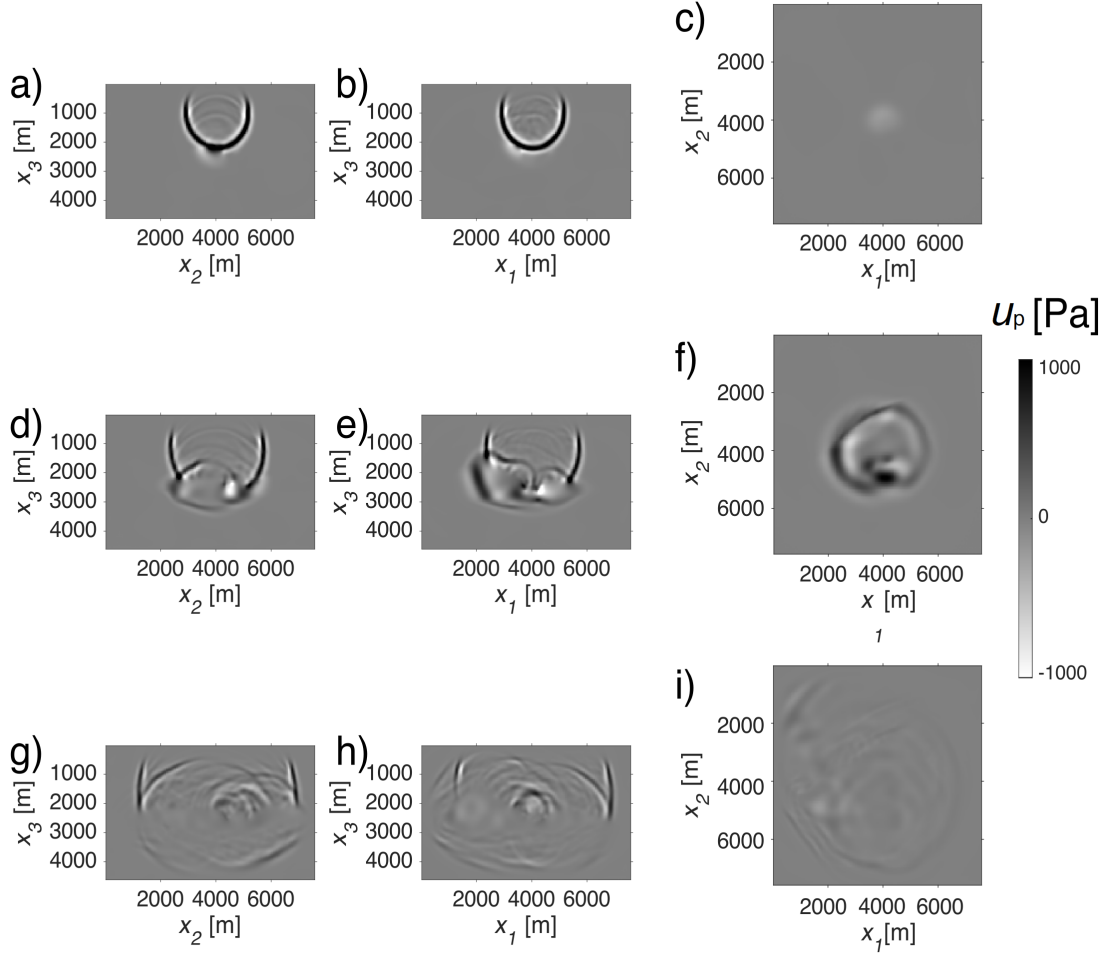


Figure 3.15: The solution of the wavefield by the proposed method. The source is placed at $[4440 \ 4440 \ 840]^T$ m. Source: Ricker wavelet, 7 Hz. The three rows correspond to $t = 0.3, 0.4, 0.6$ s. The three columns correspond to the same slices in Figure 3.14 for (a), (b), and (c).

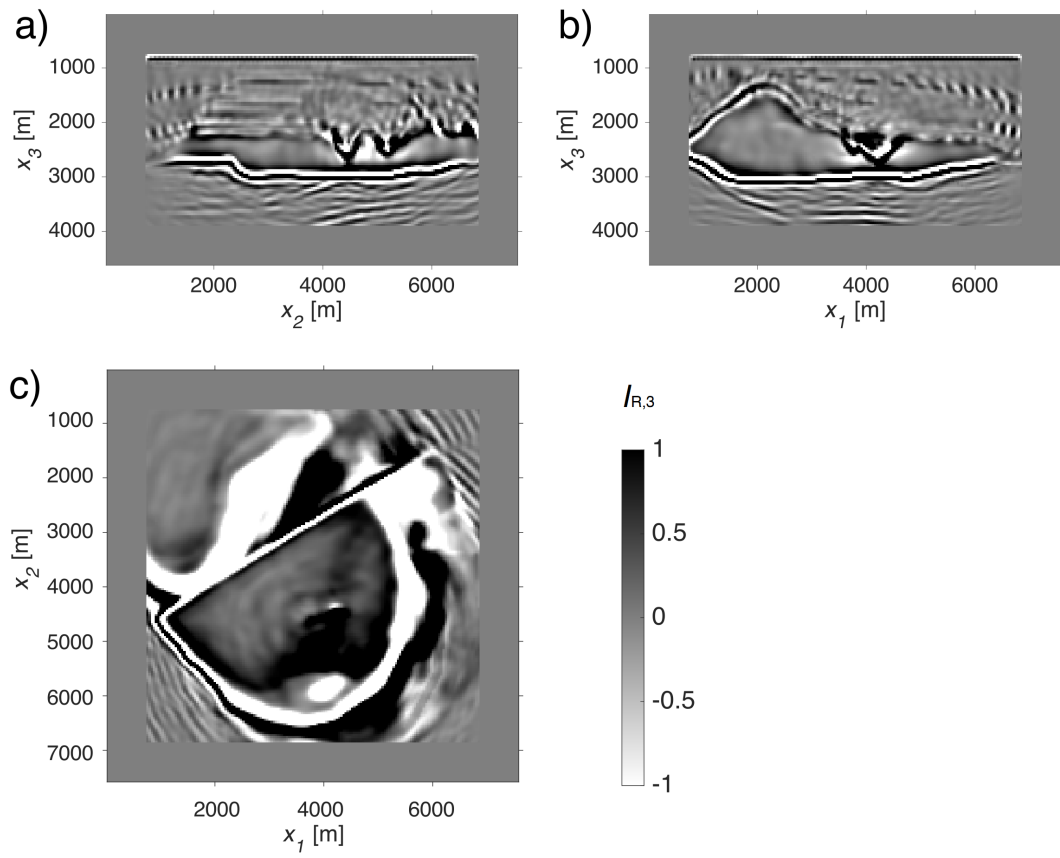


Figure 3.16: The RTM image for the slice in Figure 3.14, same slices as in Figure 3.14 for (a), (b), and (c).

The RTM example shows the necessity to include anellipticity η and tilt angle of the symmetry axis θ_{ti} in seismic imaging to avoid losing information. However, including more physical properties in the method application is more computationally expensive. For instance, including the anellipticity term requires special treatment to separate the space and wavenumbers. The tilt angle of the symmetry axis complicates the spatial derivatives.

3.5 Conclusions

This work presents a new acoustic wave equation for TTI media. Its solution does not suffer from S-wave artifacts. The equation is obtained by applying a Taylor series expansion to the anellipticity term of the existing pure P-wave equation with a rotation to the wavenumber vector. Numerically, the acoustic TTI wave equation can be solved by the hybrid finite-difference/pseudospectral method. This method is not expensive and maintains high accuracy. The wavefield solution always appears stable and devoid of artifacts, even for challenging media with fast-changing tilt angles. The RTM for the modified 2D Hess model demonstrates the necessity of considering the ellipticity and tilt angle for the TI media. The example for the modified 3D EAGE/SEG Salt model demonstrates the method is applicable to 3D. As a result, our proposed acoustic TTI wave equation holds considerable promise for advancing seismic imaging and analysis in scenarios where Earth's subsurface exhibits the utmost complexity and variability.

Chapter 4

Iterative total variation regularization for first-break travel time inversion

This work was presented at the following conference: Zhang, Y., Bachmayr, M., & De Siena, L. (2023). Total variation regularization for first-break travel time inversion. 84th EAGE Annual Conference & Exhibition, 2023(1), 1-5.

In addition, the work has been submitted as a journal article and is in revision: Zhang, Y., Bachmayr, M., & De Siena, L. (submitted in 2024). Iterative total variation regularization for first-break travel time inversion. *Geophysical Prospecting*.

Y.Z. derived the equations for the adjoint problem, developed the code, conducted the numerical experiments, drafted the initial manuscript, and performed the revisions. M.B. conceived the research idea, derived the equations for total variation, reviewed the manuscript, and supervised the research. L.D.S. verified the equations and contributed to manuscript revisions.

Abstract

Total Variation (TV) regularization is widely utilized in geophysical inverse problems for reconstructing blocky subsurface parameters. However, due to the non-linearity of its first-order optimality condition, convergence in such inverse problems can be slow. A common strategy in geophysics for addressing TV-related inverse problems is the Modified Total Variation (MTV) approach, which involves an alternating minimization for L^2 -regularized and TV-denoising subproblems that require careful tuning. In this study, we introduce a more efficient and accurate approach to solving TV-related inverse problems: Bregman iterative TV regularization with Levenberg–Marquardt-type linearization, which can be implemented using the

primal-dual Newton’s method. By linearizing the forward operator, the only non-linearity in each step comes from the regularization function. We illustrate this with the eikonal equation inverse problem as an example, utilizing the adjoint method to formulate the Fréchet derivatives, thereby circumventing the need for cumbersome ray tracing. Through numerical experiments, we benchmark our method against MTV for boundary reconstruction. We also compare the approach with smoothness penalized approaches, L-BFGS applied to the least squares functional and H^1 regularization. Our method demonstrates less-effort parameter tuning, reduced artifacts and faster convergence rates compared to MTV and the other methods. This workflow is adaptable to other physical problems by employing suitable linearization strategies.

4.1 Introduction

Seismic travel-time inversion has been widely applied to reconstruct subsurface velocities. The technique has become standard in geophysics in combination with Tikhonov regularization in H^1 (where the L^2 -norm of the gradient serves as the regularization functional, see, e.g. Rawlinson and Sambridge (2003) and Ajo-Franklin, Minsley, and Daley (2007)). However, this regularization blurs the reconstruction of medium discontinuities, as its optimality condition contains a second-order elliptic differential operator, which has a smoothing effect.

Total Variation (TV) regularization is an established technique for reconstructing sharp boundaries (Osher, Burger, Goldfarb, Xu, & Yin, 2005; Rudin, Osher, & Fatemi, 1992; Vogel & Oman, 1996; X. Zhang, Burger, Bresson, & Osher, 2010). Formally, it corresponds to an L^1 -norm penalty of the model gradient, thus enforcing sparsity of the gradient of the inverted model; in other words, the reconstructed model consists of piecewise constant regions bounded by sharp edges. Such reconstructions are preferred when there is prior knowledge that the model exhibits strong discontinuities, such as the interface between a salt dome and its overburden.

The TV regularization originated from image restoration, targeting image denoising and deblurring/deconvolution. The TV-based image denoising was first introduced in Rudin et al. (1992) with an objective function consisting of an image misfit and a non-linear TV regularization term - the Rudin-Osher-Fatemi (ROF) model. The time marching method (Rudin et al., 1992) converges slowly due to stability constraints. Vogel and Oman (1996) proposed a fixed point (FP) iteration approach; however, this method is still only linearly convergent. Finally, Chan, Golub, and Mulet (1996) introduced a linearization of the non-linear TV term based on a dual variable, where the primal-dual problem is solved using Newton’s method. This approach leads to reliable superlinear convergence.

An alternative solution to the TV regularized image denoising problem is the split-type method. Goldstein and Osher (2009) linearized the spatial differentiation term in the TV-norm of an image deconvolution problem, splitting the L^2 - and L^1 -norms.

This ensures the L^2 -norm problem is differentiable while the L^1 -norm is an ROF problem that can be solved by the shrinkage formula. Another penalization approach is to introduce deblurred and denoised images in the image deconvolution problem (Y. Huang et al., 2008). Image deblurring is involved in the L^2 -norm problem, while image denoising is contained in the $L^2 - L^1$ problem and solved with the Chambolle’s projection algorithm (Chambolle, 2004; Y. Huang et al., 2008). For each subproblem in those methods, efficient algorithms are available; however, several tuning parameters controlling the penalization terms must be carefully adjusted to achieve optimal results.

In geophysics, TV regularization has gained prominence due to the need to reconstruct sharp interfaces of subsurface objects. Geophysical inverse and image deblurring problems have different forward operators. A geophysical forward operator is usually a nonlinear operator mapping some physical parameters to the surface observations via a Partial Differential Equation (PDE). Most TV-related studies in geophysics first replace the identity matrix in the image denoising of Y. Huang et al. (2008) by a geophysical forward operator. Then the minimization problem can be split into two subproblems: one only involves L^2 -norm and the other is an image denoising problem. The L^2 -norm problem can be easily solved by standard optimization methods. The denoising steps can be done by any method for TV minimization. Gholami (2015) and Gholami and Zabihi Naeini (2019) applied the framework to acoustic impedance and Dix inversions. Other studies used the method for solving waveform inversion (Lin & Huang, 2015a), least-square reverse time migration (Lin & Huang, 2015b) and travel time inversions (Jiang & Zhang, 2017, 2018). In geophysics, this method is named Modified Total Variation (MTV).

Because the MTV approach introduces several tuning parameters, achieving optimal results in its geophysics inversion framework is often not straightforward. In this work, we explore an alternative approach based on the primal-dual linearization in image denoising proposed by Chan et al. (1996). Despite the advantage of requiring fewer parameters to tune, few studies have explored its application to geophysical problems. Bachmayr and Burger (2009) proposed a Levenberg-Marquardt type linearization, together with TV regularization based on the Bregman distance, for the diffusion equation problem.

Our method is based on the following observation: directly using TV regularization with a nonlinear forward problem results in a challenging non-convex optimization problem at each step. We apply a Levenberg-Marquardt method to the forward operator in the L^2 -norm data misfit. At each iteration, this method combines such a Levenberg-Marquardt linearization at the previous iterate, which ensures that the data misfit is a convex problem with a regularization by the total variation Bregman distance (Bachmayr & Burger, 2009).

Different from MTV, where the L^2 -norm problem and the TV problem are weakly coupled, our method still maintains the close coupling between the forward problem and TV-regularization. In the present context, this iterative scheme has the benefit of substantial computational simplifications. In particular, because the forward

operator in each step is linearized, the Primal-Dual Newton’s (PDN) method (Andersen, Christiansen, Conn, & Overton, 2000; Chan et al., 1996) can be used to solve the first-order optimality condition. This method is among the most efficient for TV minimization. In the following context, our method is abbreviated as TV-LM.

TV-LM also differs from MTV in terms of computational structure. MTV alternatively solves two sub-problems. Firstly, it addresses a model update problem involving data misfit and constraints with the TV-denoised image. In this context, Fréchet derivatives are not necessary if a gradient-based method is employed (Lin & Huang, 2015a), even if they must be computed if using some Newton’s methods (Jiang & Zhang, 2017). In this case, the update of the denoised image follows the algorithm described by Goldstein and Osher (2009), which involves several inner iterations to refine the denoised image. TV-LM solves the aforementioned linearized first-order optimality condition using Newton’s method, which always requires constructing the Fréchet derivatives. This construction is achieved by solving the adjoint equation. At each outer iteration, a non-linear problem is solved to update the model, and as a result, each outer iteration involves multiple inner iterations.

In this work, we apply our algorithm TV-LM to the eikonal equation inverse problem. We examine a standard Levenberg–Marquardt-type regularization of the model gradient – in the following referred to as H^1 regularization for simplicity – and the corresponding iterative TV regularization using three toy models, comprising the Marmousi model (Billette & Brandsberg-Dahl, 2005) and the 2004 BP model (Billette & Brandsberg-Dahl, 2005). To assess performance of our method, we also include L-BFGS (Leung & Qian, 2006) and MTV for comparisons. A noise test is also performed for 2004 BP model. The L-BFGS method uses the L^2 -norm data misfit as the objective function. In the inversion process, a Gauss smoothing filter is applied to the descent direction to accelerate the convergence.

Our results show that TV-LM not only provides superior reconstruction of sharp interfaces compared to MTV but also offers greater computational efficiency in our numerical tests. In contrast, methods such as the L-BFGS method based on L^2 -norm misfit functional and H^1 regularization are not suitable for such mission.

In Section 4.2, we first present the general inversion framework based on TV-LM and then apply it to the eikonal equation problem for 2D travel time inversion. In Section 4.3, we discuss other approaches to solving geophysics inverse problems, including the L-BFGS method based on L^2 -norm misfit functional, H^1 regularization, and MTV. Section 4.4 showcases numerical tests to demonstrate the different behaviors of the aforementioned inversion methods. Finally, Section 4.5 provides discussions and concluding remarks.

4.2 Method

4.2.1 Total variation regularization of inverse problems

We consider the problem of travel time inversion on a bounded domain $\Omega \subset \mathbb{R}^2$ representing the medium, where $x = [x_1 \ x_3]^T$ denotes the Cartesian coordinates. The methods studied in this work are based on total variation penalty terms, which have been proposed for image denoising problems by Rudin et al. (1992) and have been applied to general linear inverse problems by Acar and Vogel (1994) and to nonlinear inverse problems by Chavent and Kunisch (1997), Chan and Tai (2003) and Ascher, Haber, and Huang (2006).

A cost function E targeting seismic velocity $v(x)$ with total variation regularization is given by

$$E = \min_v h(v) + \alpha \text{TV}(v) , \quad (4.1)$$

h is the data misfit defined by

$$h(v) = \frac{1}{2} \int_{\Omega_r} |t(v) - t_{\text{obs}}|^2 dx ; \quad (4.2)$$

$\Omega_r \subset \Omega$ is the region where receivers are placed; t is the calculated solution for the given seismic velocity, in this application, travel time; t_{obs} is the observed data; α is the scaling parameter; and the isotropic total variation $\text{TV}(v)$ defined by

$$\text{TV}(v) = \sup \left\{ \int_{\Omega} v \nabla \cdot w \, dx : w \in C^1(\Omega, \mathbb{R}^2) \text{ with } \text{supp } w \subset \Omega, \sup_{\Omega} |w| \leq 1 \right\} , \quad (4.3)$$

is used as a regularization term. Note that in the special case of v that is weakly differentiable, we have

$$\text{TV}(v) = \int_{\Omega} |\nabla v| \, dx . \quad (4.4)$$

However, functions of finite total variation can also have jump discontinuities, and regularization methods using this penalty term can thus reconstruct sharp edges. $\text{TV}(v)$ is, in general, not differentiable in the classical sense, which complicates the numerical treatment of first-order optimality conditions for equation 4.1. Methods exploiting the convexity of tv based on dual formulations, as developed in Carter (2002) and Chambolle (2004), are applicable mainly to denoising problems with the identity as the forward operator, but not to our setting. Here, we instead replace $\text{tv}(v)$ by a standard smoothed approximation,

$$\text{TV}_{\epsilon_{\text{tv}}}(v) \approx \int_{\Omega} \sqrt{|\nabla v|^2 + \epsilon_{\text{tv}}^2} \, dx , \quad (4.5)$$

where ϵ_{tv} is a small positive number. As pointed out in Vogel and Oman (1996), larger values of ϵ_{tv} tend to round off sharp edges, and a larger α leads to approximations with fewer piecewise constant regions.

4.2.2 Bregman distance

Osher et al. (2005) have proposed a generally applicable construction of total variation-based iterative regularization methods based on the corresponding Bregman distance (Bregman, 1967). The Bregman distance associated to $\text{TV}_{\epsilon_{\text{tv}}}(v)$ is given by

$$D_{\text{TV}_{\epsilon_{\text{tv}}}}(v_{i+1}, v_i) = \text{TV}_{\epsilon_{\text{tv}}}(v_{i+1}) - \text{TV}_{\epsilon_{\text{tv}}}(v_i) - \langle \text{TV}'_{\epsilon_{\text{tv}}}(v_i), v_{i+1} - v_i \rangle, \quad (4.6)$$

where $\langle \cdot, \cdot \rangle$ is the dot product; v_i and v_{i+1} are the velocities at two consecutive iterations.

Now, instead of minimizing $h(v) + \alpha \text{TV}_{\epsilon_{\text{tv}}}(v)$, we solve the inversion with TV-based Bregman distance as the regularization: for a given v_i ,

$$v_{i+1} = \underset{v}{\operatorname{argmin}} h(v) + \alpha D_{\text{TV}_{\epsilon_{\text{tv}}}}(v, v_i). \quad (4.7)$$

In image denoising, this type of iteration has been observed to be superior to conventional TV regularization in that it leads to reconstructions with better contrast (Goldstein & Osher, 2009). In our setting of nonlinear inverse problems, it has been studied by Bachmayr and Burger (2009), who also considered combinations with iterative linearizations of Levenberg–Marquardt or Landweber type.

4.2.3 Levenberg–Marquardt-type linearization

We combine the iterative scheme in equation 4.7 with a linearization of t with respect to v to ensure the convexity of the minimization problem at each iteration. Linearizing the nonlinear mapping t inside the quadratic data misfit functional yields

$$h_1(v) = \frac{1}{2} \int_{\Omega_r} |t_i + t'(v_i)(v - v_i) - t_{\text{obs}}|^2 dx. \quad (4.8)$$

where i is the iteration index for the outer iteration; $t'(v_i)$ is the derivative of the travel time with respect to the velocity at iteration i .

The iterative Levenberg–Marquardt regularization in $H^1(\Omega)$ is defined by the minimization problem

$$v_{i+1} = \underset{v}{\operatorname{argmin}} h_1(v) + \alpha \int_{\Omega} |\nabla(v - v_i)|^2 dx. \quad (4.9)$$

Following Bachmayr and Burger (2009) to obtain an analogous scheme for TV regularization, we insert equations 4.8 and (4.6) into equation 4.7 and replace h in equation 4.7 with h_1 to formulate the new optimization problem,

$$v_{i+1} = \underset{v}{\operatorname{argmin}} \frac{1}{2} \int_{\Omega_r} |t_i + t'(v_i)(v - v_i) - t_{\text{obs}}|^2 dx + \alpha (\text{TV}_{\epsilon_{\text{tv}}}(v) - \langle \xi_i, v \rangle), \quad (4.10)$$

with

$$\xi = \text{TV}'_{\epsilon_{\text{tv}}}(v) = -\nabla \cdot \frac{\nabla v}{\sqrt{|\nabla v|^2 + \epsilon_{\text{tv}}^2}} . \quad (4.11)$$

Solving for the first-order optimality condition of equation 4.10 yields v_{i+1} . Combining with equation 4.11, we thus need to solve

$$t'(v_i)^*(t_i + t'(v_i)(v - v_i) - t_{\text{obs}}) + \alpha(\text{tv}'_{\epsilon_{\text{tv}}}(v) - \xi_i) = 0 , \quad (4.12)$$

for $v_{i+1} = v$, and then update

$$\xi_{i+1} = \xi_i - \frac{1}{\alpha} t'(v_i)^*(t_i + t'(v_i)(v_{i+1} - v_i) - t_{\text{obs}}) , \quad (4.13)$$

where $t'(v_i)^*$ is the adjoint operator of $t'(v_i)$ with respect to the L^2 -inner product. In equation 4.12, $t'(v_i)$ is only evaluated for t at receivers.

4.2.4 A primal-dual Newton's method

As $h_l(v)$ is linear with respect to v , the only remaining nonlinearity arises from the term $D_{\text{TV}_{\epsilon_{\text{tv}}}}(v, v_i)$; in other words, at each iteration i , equation 4.12 is nonlinear with respect to v due to the presence of $\text{TV}'_{\epsilon_{\text{tv}}}(v)$. Thus, iterative methods are also necessary to approximate the solution of equation 4.12. We denote the iterates of this inner iteration with index j , those of the outer iteration defined by equation 4.10 with index i . For given v_i , to obtain v_{i+1} as in equation 4.12 we iteratively solve

$$-\alpha \nabla \cdot \frac{\nabla u_v}{\sqrt{|\nabla u_v|^2 + \epsilon_{\text{tv}}^2}} + t'(v_i)^*(t'(v_i)u_v - f_i) - \alpha \xi_i = 0 , \quad (4.14)$$

for u_v , where

$$f_i = t'(v_i)v_i + t_{\text{obs}} - t_i , \quad (4.15)$$

and where we can use v_i as the starting value of the inner iteration. The computed approximate solution of equation 4.14 is then assigned to v_{i+1} .

There are several approaches to solving the system in equation 4.14, for instance, the Fixed Point (FP) or lagged diffusivity method (Vogel & Oman, 1996) where for a given approximation $(u_v)_j$ of the solution u_v , the next approximation $(u_v)_{j+1}$ is obtained as the solution of

$$-\alpha \nabla \cdot \frac{\nabla (u_v)_{j+1}}{\sqrt{|\nabla (u_v)_j|^2 + \epsilon_{\text{tv}}^2}} + t'(v_i)^*(t'(v_i)(u_v)_{j+1} - f_i) - \alpha \xi_i = 0 . \quad (4.16)$$

The resulting iterative method is robust but only linearly convergent.

Chan et al. (1996) proposed a new linearization method for $\text{TV}'_{\epsilon_{\text{TV}}}(v)$ that improves performance, the Primal-Dual Newton's method (PDN). This method is based on introducing the auxiliary (dual) variable

$$w = \frac{\nabla u_v}{\sqrt{|\nabla u_v|^2 + \epsilon_{\text{TV}}^2}}, \quad (4.17)$$

with $|w| \leq 1$. With the linearization in equation 4.17, we obtain a system of PDEs for u_v and w . For example, assuming a 2D case in $x = [x_1 \ x_3]^T$ coordinates, we can rewrite (4.14) as

$$\begin{aligned} -\alpha \nabla \cdot w + t'(v_i)^*(t'(v_i)u_v - f_i) - \alpha \xi_i &= 0, \\ w_1 \sqrt{|\nabla u_v|^2 + \epsilon_{\text{TV}}^2} - \frac{\partial u_v}{\partial x_1} &= 0, \\ w_3 \sqrt{|\nabla u_v|^2 + \epsilon_{\text{TV}}^2} - \frac{\partial u_v}{\partial x_3} &= 0, \end{aligned} \quad (4.18)$$

where $w = [w_1 \ w_3]^T$.

The system can be solved efficiently using Newton's method. At each iteration, updating w requires a line search to preserve the constraint $|w| \leq 1$; as noted in Vogel (2002), in practice, no line search is observed to be necessary for the update of u_v . Note that in contrast to the effective global convergence of this primal-dual Newton's method, the direct application of Newton's method to equation 4.14 generally leads to an extremely small convergence radius.

4.2.5 Eikonal equation

Under the assumption of infinite frequency (ray theory), the first arrival of a single wave mode in isotropic media can be described by the eikonal equation,

$$|\nabla t|^2 = \frac{1}{v^2}, \quad (4.19)$$

where the boundary condition is $t(x_s) = 0$, with x_s the source location for one event. To complete the inversion framework, we still need two steps: the numerical solution of the eikonal equation for the computation of t and the linearization of the eikonal equation for $t'(v)$.

The eikonal equation can be solved in the framework of Hamiltonian systems by the method of characteristics (Biondi, Fomel, & Alkhalifah, 2001; ČErvený, 1990; Rueger, 1993; U. b. Waheed, Alkhalifah, & Wang, 2015) or using finite-difference approximations, such as fast marching (Gill & Murray, 1978; Rawlinson & Sambridge, 2018) and fast sweeping methods (U. B. Waheed, Yarman, & Flagg, 2014; Zhao, 2004). The numerical Hamiltonian framework yields a consistent and accurate discretization scheme. However, there are three challenges: It has to mitigate the non-differentiability at intersections of characteristics; a large system of nonlinear

equations has to be solved after discretization, which can be slow for large problems; and it may suffer from the shadow zone problem (Julian & Gubbins, 1977). An interpolation is required if receivers are located in the gap of the rays. Here, we thus follow a finite-difference approach, which does not suffer from the aforementioned issues (Zhao, 2004) and is usually suitable for large-scale eikonal equation problems. The fast sweeping method has the complexity of $O(N_x)$ for an N_x -grid problem, while the fast marching method has $O(N_x \log N_x)$. Thus, the fast-sweeping method used in this work has an advantage when it comes to large models.

In inversion techniques, one needs to either construct the Fréchet derivatives for Newton’s method or compute the gradient of the data misfit with respect to the model parameters for gradient-based methods. It is common to use the adjoint eikonal equation to evaluate the gradient of the data misfit with respect to the model parameters, denoted as $h'(v)$ (see, e.g., Tong (2021), Leung and Qian (2006), Taillandier et al. (2009)). In our framework, we specifically require the Fréchet derivatives of the travel time function $t'(v)$.

In geophysics, the computation of the Fréchet derivatives of the travel time function $t'(v)$ has traditionally relied on ray tracing methods (Rawlinson, Kool, & Sambridge, 2018; Rawlinson & Sambridge, 2003), even though the forward problem can be solved using the fast marching or fast sweeping methods. However, there are two main issues with ray tracing: (1) it suffers from the shadow zone problem; and (2) using the travel time gradient for ray tracing can be inaccurate, potentially leading to inaccurate ray paths.

In light of these challenges, we compute the Fréchet derivatives using the adjoint travel time, which can be efficiently solved using the fast sweeping method (J.-W. Huang & Bellefleur, 2012; Leung & Qian, 2006; Taillandier et al., 2009). Appendix B provides a detailed illustration of how to compute the adjoint eikonal equation and how to use the adjoint travel time to construct the Fréchet derivatives.

4.2.6 Numerical implementation

Suppose the velocity v is defined on an equispaced grid in two space dimensions with $N_{\text{space},1} \times N_{\text{space},3}$ grids, $\{(ih_{\text{space},1}, jh_{\text{space},3}) : i = 1, 2, \dots, N_{\text{space},1}, j = 1, 2, \dots, N_{\text{space},3}\}$. Here $h_{\text{space},1}$ and $h_{\text{space},3}$ are the grid spacings in x_1 - and x_3 -directions, respectively. The value of v at grid (i, j) is $\mathbf{v}_{i,j}$. Through the eikonal equation, one can obtain the travel time $\mathbf{t}_{i,j}$ at each node of the velocity $\mathbf{v}_{i,j}$. Let $\mathbf{v}$ and $\mathbf{t}$ denote the vectors (grid functions) obtained from the discretization of v and t .

Equation A9 has to be evaluated for each source-receiver pair. Say we have N_{sr} source-receiver pairs, there is a sequence of δt_e , $e = 1, 2, \dots, N_{\text{sr}}$. This means there are N_{sr} corresponding linear equations as in equation A9. We can rewrite these

equations in matrix form,

$$\delta \mathbf{t} = \mathbf{G} \delta \mathbf{v} , \quad (4.20)$$

where $\mathbf{G}$ is the discretization of $t'(v)$. Equation 4.20 is the discrete version of equation A9.

To implement the travel time inversion, we combine the Bregman iteration given by equations 4.12 and (4.13) as an outer iteration with the primal-dual Newton's method obtained from the discretization of equation 4.18. We denote by $\mathbf{u}$, $\mathbf{w}_1$ and $\mathbf{w}_3$ the coefficient vectors for u_v , w_1 and w_3 discretized analogously to v above. We also define the discretized differential operators $\mathbf{D}_1$ and $\mathbf{D}_3$ by

$$\begin{aligned} (\mathbf{D}_1 \mathbf{v})_{i,j} &= \frac{\mathbf{v}_{i,j} - \mathbf{v}_{i-1,j}}{h_{\text{space},1}} , \\ (\mathbf{D}_3 \mathbf{v})_{i,j} &= \frac{\mathbf{v}_{i,j} - \mathbf{v}_{i,j-1}}{h_{\text{space},3}} , \end{aligned} \quad (4.21)$$

where $\mathbf{D}_1$ and $\mathbf{D}_3$ are both combined with Neumann boundary conditions.

For the discrete Newton updates $\Delta \mathbf{u}_v$, $\Delta \mathbf{w}_1$ and $\Delta \mathbf{w}_3$, following Vogel (2002), we obtain

$$\Delta \mathbf{u}_v = (\mathbf{G}^T \mathbf{G} + \alpha \mathbf{A}_1)^{-1} \mathbf{b}_1 , \quad (4.22)$$

where $\mathbb{I}$ is the identity matrix, and

$$\begin{aligned} \mathbf{A}_1 &= \mathbf{D}_1^T \text{diag} \left(\text{diag}(\mathbf{a}_2)^{-1} \left[1 + \text{diag}(\mathbf{a}_2)^{-1} \text{diag} \left(\mathbf{D}_1(\mathbf{u}_v)_j \right) (\mathbf{w}_1)_j \right] \right) \mathbf{D}_1 \\ &\quad + \mathbf{D}_1^T \text{diag} \left(\text{diag}(\mathbf{a}_2)^{-2} \text{diag} \left(\mathbf{D}_1(\mathbf{u}_v)_j \right) (\mathbf{w}_1)_j \right) \mathbf{D}_3 \\ &\quad + \mathbf{D}_3^T \text{diag} \left(\text{diag}(\mathbf{a}_2)^{-1} \left[1 + \text{diag}(\mathbf{a}_2)^{-1} \text{diag} \left(\mathbf{D}_3(\mathbf{u}_v)_j \right) (\mathbf{w}_3)_j \right] \right) \mathbf{D}_3 \\ &\quad + \mathbf{D}_3^T \text{diag} \left(\text{diag}(\mathbf{a}_2)^{-2} \text{diag} \left(\mathbf{D}_3(\mathbf{u}_v)_j \right) (\mathbf{w}_1)_j \right) \mathbf{D}_3 , \\ \mathbf{b}_1 &= \mathbf{G}^T \left(\mathbf{f}_i - \mathbf{G}(\mathbf{u}_v)_j \right) \\ &\quad + \alpha \xi_i + \alpha \mathbf{D}_1 \text{diag}(\mathbf{a}_2)^{-1} \mathbf{D}_1(\mathbf{u}_v)_j + \alpha \mathbf{D}_3 \text{diag}(\mathbf{a}_2)^{-1} \mathbf{D}_3(\mathbf{u}_v)_j , \end{aligned} \quad (4.23)$$

where $\mathbf{a}_2$ is the discretization of $\sqrt{\left(\frac{\partial(u_v)_j}{\partial x_1^2}\right)^2 + \left(\frac{\partial(u_v)_j}{\partial x_3}\right)^2 + \epsilon_{\text{tv}}^2}$ and $\mathbf{u}_v$ is the discretization of u_v .

The updates for the dual variables are given by,

$$\begin{aligned}
\Delta \mathbf{w}_1 &= -(\mathbf{w}_1)_j \\
&\quad + \text{diag}(\mathbf{a}_2)^{-1} \left[\mathbf{D}_1(\mathbf{u}_v)_{j+1} + \text{diag} \left(1 + \text{diag}(\mathbf{a}_2)^{-1} \text{diag}(\mathbf{D}_1(\mathbf{u}_v)_{j+1}) (\mathbf{w}_1)_j \right) \right] \\
&\quad \mathbf{D}_1 \Delta \mathbf{u}_v \\
&\quad + \text{diag}(\mathbf{a}_2)^{-1} \text{diag}(\mathbf{D}_3(\mathbf{u}_v)_{j+1}) \text{diag}((\mathbf{w}_1)_j) \mathbf{D}_3 \Delta \mathbf{u} , \\
\Delta \mathbf{w}_3 &= -(\mathbf{w}_3)_j \\
&\quad + \text{diag}(\mathbf{a}_2)^{-1} \left[\mathbf{D}_3(\mathbf{u}_v)_{j+1} + \text{diag} \left(1 + \text{diag}(\mathbf{a}_2)^{-1} \text{diag}(\mathbf{D}_3(\mathbf{u}_v)_{j+1}) (\mathbf{w}_3)_j \right) \right] \\
&\quad \mathbf{D}_3 \Delta \mathbf{u}_v \\
&\quad + \text{diag}(\mathbf{a}_2)^{-1} \text{diag}(\mathbf{D}_1(\mathbf{u}_v)_{j+1}) \text{diag}((\mathbf{w}_3)_j) \mathbf{D}_1 \Delta \mathbf{u} ,
\end{aligned} \tag{4.24}$$

where $\text{diag}(\mathbf{v})$ represents the square diagonal matrix with the elements of the vector $\mathbf{v}$ on the main diagonal; $\mathbf{f}$, $\mathbf{a}_2$ and $\boldsymbol{\xi}$ is the vector from the discretization of f , a_2 and ξ , respectively; and $\boldsymbol{\xi}_1 = \mathbf{0}$.

To avoid the storage of $\mathbf{G}^T \mathbf{G}$, we can solve an equivalent to equation 4.22,

$$\begin{bmatrix} \Delta \mathbf{u}_v \\ \Delta \phi_d \end{bmatrix} = \begin{bmatrix} \alpha \mathbf{A}_1 & \mathbf{G}^T \\ \mathbf{G} & -\mathbb{I} \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{b}_1 \\ \mathbf{0} \end{bmatrix} , \tag{4.25}$$

where $\Delta \phi_d$ is the dual variable corresponding the primal $\Delta \mathbf{u}_v$. For the inner iterations, it is not necessary to store $\Delta \phi_d$, and the update for (u, w) system is,

$$\begin{aligned}
(\mathbf{u}_v)_{j+1} &= (\mathbf{u}_v)_j + \Delta \mathbf{u}_v , \\
(\mathbf{w}_1)_{j+1} &= (\mathbf{w}_1)_j + \alpha_w \Delta \mathbf{w}_1 , \\
(\mathbf{w}_3)_{j+1} &= (\mathbf{w}_3)_j + \alpha_w \Delta \mathbf{w}_3 ,
\end{aligned} \tag{4.26}$$

where α_w is the step length for the update of w which ensures $\|w\|_\infty \leq 1$.

During the outer iterations, in principle, α is set large at the beginning and decreases gradually with the outer iterations until the data misfit h reaches the noise level. Following Morozov's discrepancy principle as analyzed for this method by Bachmayr and Burger (2009), in the case of noisy data, the iterative regularization is stopped when the data misfit h reaches the level of data noise.

We now summarize the procedure for performing the Levenberg–Marquardt-type iterative TV regularization. Algorithm 4.1 gives the outer iteration procedure. Algorithm 4.2 provide the inner iterations by PDN. For TV-LM, we select $n_{\text{inner}} = 10$ and $\varepsilon_{\text{inner}} = 10^{-4}$.

Algorithm 4.1: Outer iteration of TV-LM to update v .

$n_{\text{outer}} :=$ maximum outer iterations.

$\epsilon_{\text{outer}} :=$ stopping criterion for the data misfit.

$\mathbf{v}_1 :=$ initial guess for the velocity.

for $i = 1 : n_{\text{outer}}$ **do**

 Compute travel time by equation 4.19 and data misfit h by equation 4.2.

 Compute $\mathbf{G}$ using equations A7 - (A9) and (4.20).

if $h \leq \epsilon_{\text{outer}}$ **then**

 | Break.

end

 Perform inner iterations for the update of $\mathbf{u}_v$, according to Algorithm 4.2.

 Assign the last iterate $\mathbf{u}_v$ in the inner iterations to $\mathbf{v}_{i+1}$. Compute ξ_{i+1} using equation 4.13.

end

Return $\mathbf{v}_{i+1}$

Algorithm 4.2: Inner iteration of TV-LM to update $\mathbf{u}_v$ by PDN.

$n_{\text{inner}} :=$ maximum inner iterations.

$\epsilon_{\text{inner}} :=$ stopping criterion for $\mathbf{u}_v$.

$(\mathbf{u}_v)_1 := \mathbf{v}_i$.

$(\mathbf{w}_1)_1 :=$ initial guess, generally a zero vector.

$(\mathbf{w}_3)_1 :=$ initial guess, generally a zero vector.

for $j = 1 : n_{\text{inner}}$ **do**

 Compute $\Delta \mathbf{u}_v$ by equation 4.25.

 Compute $\Delta \mathbf{w}_1$ and $\Delta \mathbf{w}_3$ by equation 4.24.

 Update $(\mathbf{u}_v)_{j+1}$, $(\mathbf{w}_1)_{j+1}$ and $(\mathbf{w}_3)_{j+1}$ by equation 4.26.

 Perform a line search,

$$\alpha_w = \max_{\tau_w} \{0 \leq \tau_w \leq 1 \mid \| ((\mathbf{w}_1)_j + \tau_w \Delta \mathbf{w}_1)^2 + ((\mathbf{w}_3)_j + \tau_w \Delta \mathbf{w}_3)^2 \|_{\infty} \leq 1\} \quad (4.27)$$

 Update $(\mathbf{u}_v)_{j+1}$, $(\mathbf{w}_1)_{j+1}$ and $(\mathbf{w}_3)_{j+1}$ as in equation 4.26.

if $\|(\mathbf{u}_v)_{j+1} - (\mathbf{u}_v)_j\|_2^2 \leq \epsilon_{\text{inner}} \|(\mathbf{u}_v)_j\|_2^2$ **then**

 | Break.

end

end

Return $(\mathbf{u})_{j+1}$.

4.3 Other methods

To benchmark the accuracy and efficiency of our algorithm, we compare it with two other algorithms: L-BFGS and MTV.

The L-BFGS method uses the adjoint-state method to compute the gradient of updating velocity $h'(v)$, stacked for all shot-gathers. According to Leung and Qian (2006); Taillandier et al. (2009), the numerical solution of $h'(v)$ contains irregularities due to spikes. Simply using $h'(v)$ will force one to pick very small α_{lbfgs} . So, the Gauss image filter is imposed for $h'(v)$ as a regularity to achieve a faster convergence. The velocity update for the L-BFGS method is given by

$$\mathbf{v}_{i+1} = \mathbf{v}_i - \alpha_{\text{lbfgs}} \mathbf{H}_{\text{lbfgs}}^{-1}(\mathbf{G}_{\mathbf{h},\mathbf{v}})_{\text{Gauss}(\sigma_{\text{gauss}})} , \quad (4.28)$$

where $\mathbf{G}_{\mathbf{h},\mathbf{v}}$ is the discrete $h'(v)$; α_{lbfgs} is the step length obtained in the line search; $\mathbf{H}_{\text{lbfgs}}^{-1}$ is the approximated inverse of Hessian obtained by L-BFGS; $(\cdot)_{\text{gauss}(\sigma_{\text{gauss}})}$ is a Gauss image filter with a standard deviation σ_{gauss} .

MTV is a popular TV method in geophysics (Jiang & Zhang, 2017, 2018; Lin & Huang, 2015a, 2015b). The MTV problem is generally decomposed into an L^2 problem and an $L^1 - L^2$ coupled image denoising problem. The objective function of MTV is

$$E_{\text{mtv}}(v, \eta_{\text{mtv}}) = \min_{v, \eta_{\text{mtv}}} h(v) + \lambda_{\text{tv},1} \int_{\Omega} (v - \eta_{\text{mtv}})^2 dx + \lambda_{\text{tv},2} \text{TV}(\eta_{\text{mtv}}) , \quad (4.29)$$

where $\lambda_{\text{tv},1}$ and $\lambda_{\text{tv},2}$ are balancing parameters. This problem can be split into two subproblems,

$$\begin{aligned} \text{Subproblem 1: } v &= \underset{v}{\operatorname{argmin}} h(v) + \lambda_{\text{tv},1} \int_{\Omega} (v - \eta_{\text{mtv}})^2 dx , \\ \text{Subproblem 2: } \eta_{\text{mtv}} &= \underset{\eta_{\text{mtv}}}{\operatorname{argmin}} \lambda_{\text{tv},1} \int_{\Omega} (v - \eta_{\text{mtv}})^2 dx + \lambda_{\text{tv},2} \text{TV}(\eta_{\text{mtv}}) . \end{aligned} \quad (4.30)$$

The two subproblems can be solved iteratively, where $\lambda_{\text{tv},1}$ and $\lambda_{\text{tv},2}$ have to be tuned carefully. Subproblem 1 only contains an L^2 -norm problem, which can be solved using Newton's method. Subproblem 2 is an image-denoising problem - its solution is provided, for instance, the split-Bregman approach by Goldstein and Osher (2009); Y. Wang, Yang, Yin, and Zhang (2008).

4.4 Numerical tests

In this section, we show first the perturbation test using the adjoint-state technique. Then, we perform inversions with L-BFGS, H^1 and TV regularizations using three models, a synthetic toy model with sharp interfaces, the Marmousi model (Brougois et al., 1990) and 2004 BP (Billette & Brandsberg-Dahl, 2005). Inversion by H^1

regularization in this section means Levenberg–Marquardt-type regularization of the model gradient, computed by equation 4.9. TV-LM is solved by algorithms 4.1 and 4.2. We first compute the first-break arrival time with the toy model and then try to reconstruct the velocity using these “true data”.

4.4.1 Perturbation test

First, we test the perturbation in equation A9. The background model is given in Figure 4.1. We choose the source and the 10th receiver and evaluate equations A7 and (A8) for λ_t (Figure 4.2). Now, if δv is known, δt at a certain location can be computed using equation A9. We perturb the velocity in Figure 4.1 by 0.1 times the background, computing the exact travel time difference and the difference by the adjoint-state equation (Figure 4.3). Approximation error arises due to the solution of the linearized eikonal equation, with this error progressively diminishing as the data misfit approaches zero. The travel time perturbation approximated by the adjoint-state equation follows the trend of the exact solution. The travel time perturbation approximated by the adjoint-state equation is also more accurate with a near offset.

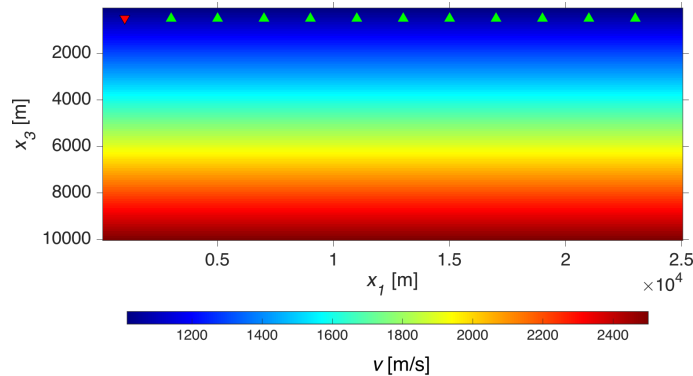


Figure 4.1: Background model for the perturbation test. Down pointing triangles – sources. Up pointing triangles – receivers. The same marker convention is consistently applied throughout the work.

4.4.2 A toy model

In this part, we consider a synthetic model (Figure 4.4a), where the sources and receivers are placed at the top. The true travel time is computed first for the model in Figure 4.4a, and the inversion is performed with a starting model given in Figure 4.4b, by L-BFGS, H^1 regularization, MTV and TV-LM. We want to compare the inversion performance using different methods and the effects of the inversion parameters.

To analyze the inversion performance, we present three metrics: the data misfit h ,

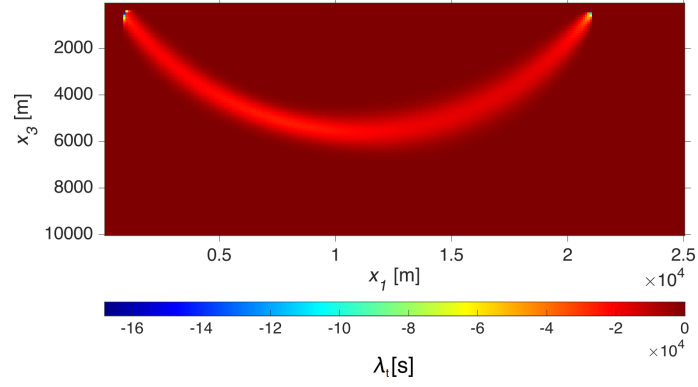


Figure 4.2: Adjoint travel time λ_t for the source and 10th receiver in the model in Figure 4.1.

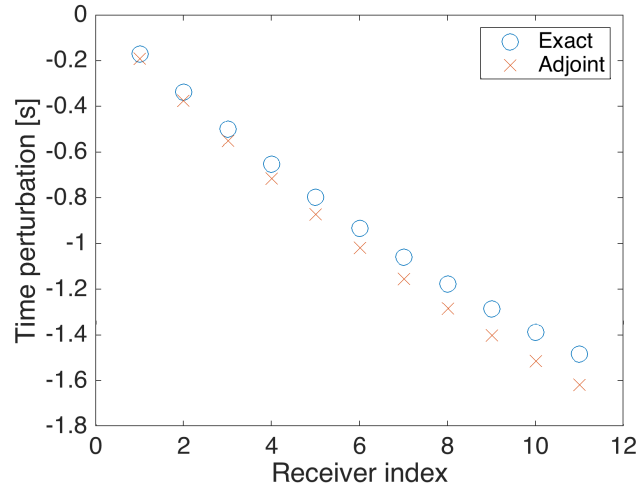


Figure 4.3: Exact δt and approximated δt using the adjoint-state equation method. The velocity perturbation is $\delta v = 0.1v$ in Figure 4.1.

the Bregman distance $D_{\text{TV}_{\epsilon_{\text{TV}}}}(v_t, v_i)$, and the L^2 -norm model misfit $\frac{1}{2} \int_{\Omega} (v_i - v_t)^2 dx$. Here, v_i denotes the inverted velocity at a certain outer iteration, and v_t represents the ground truth. These metrics are plotted as functions of elapsed time. Although the L^2 -norm model misfit is commonly used in geophysics to measure differences between models, we also include the Bregman distance to study the quality of the inversion because the Bregman distance contains TV functional and it focuses on sharpness features.

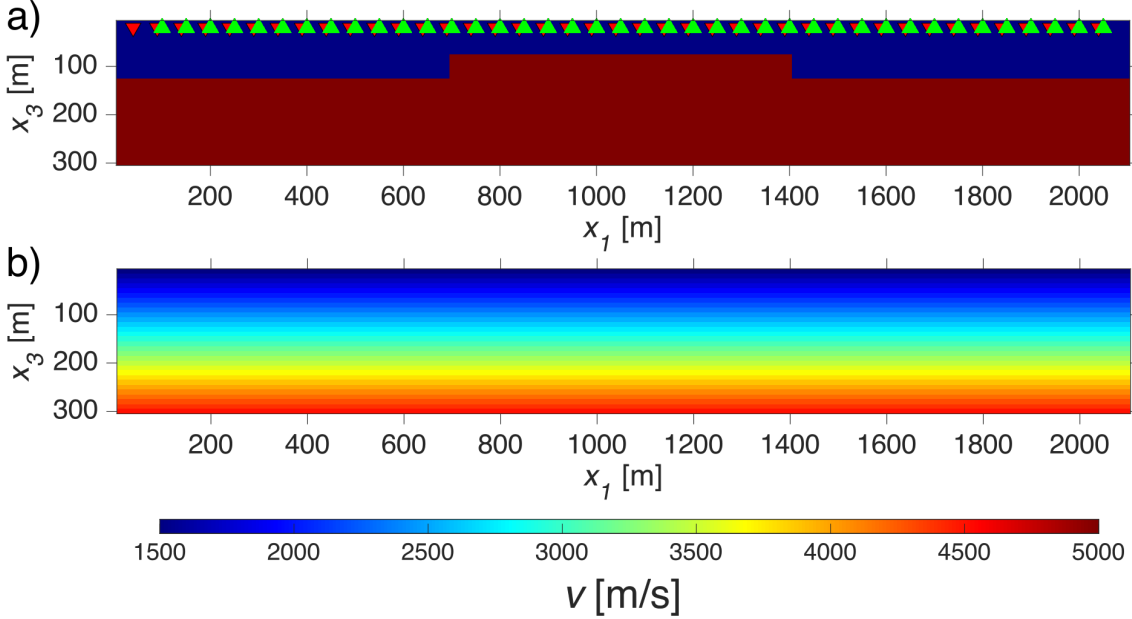


Figure 4.4: (a) The true toy model. (b) The starting model for inversion. The model is discretized into 210 (x_1) by 30 (x_3) grids.

Table 4.1 shows the tuning parameters used for the inversion of the toy model. For L-BFGS and MTV, we use the trial and error test to select the parameters that lead to the lowest data misfit within the certain iterations in our experiments. For TV-LM, we vary the regularization parameters to test their effect on model reconstruction. We take $\alpha = 10^{-5}$ and $\epsilon_{\text{TV}} = 10^{-1}$ as the reference and change them alternatively. For $\alpha = 5 \times 10^{-5}$, we need more iterations because the data misfit decay is slower than the reference. For MTV, more iterations are assigned because its convergence is slower than TV-LM with reference parameters.

Figure 4.5 gives the result of the inversion by all the methods mentioned in Table 4.1. Figure 4.6 is a profile taking at $x = 100$ m. Figure 4.5a shows L-BFGS constructs a smooth result. Figure 4.5b is also smooth but sharper than Figure 4.5a. The other methods in Figures 4.5c-f can reconstruct sharp interfaces. In Figure 4.6, the difference is clearer. MTV reconstructs a smoother model than TV-LM with reference parameters. The result obtained using larger α is similar to TV-LM with reference parameters, as long as the iterations are sufficient. A larger ϵ_{TV} slightly smooths out the interface.

In Figure 4.7, L-BFGS has the worst reconstruction accuracy, even if its computation

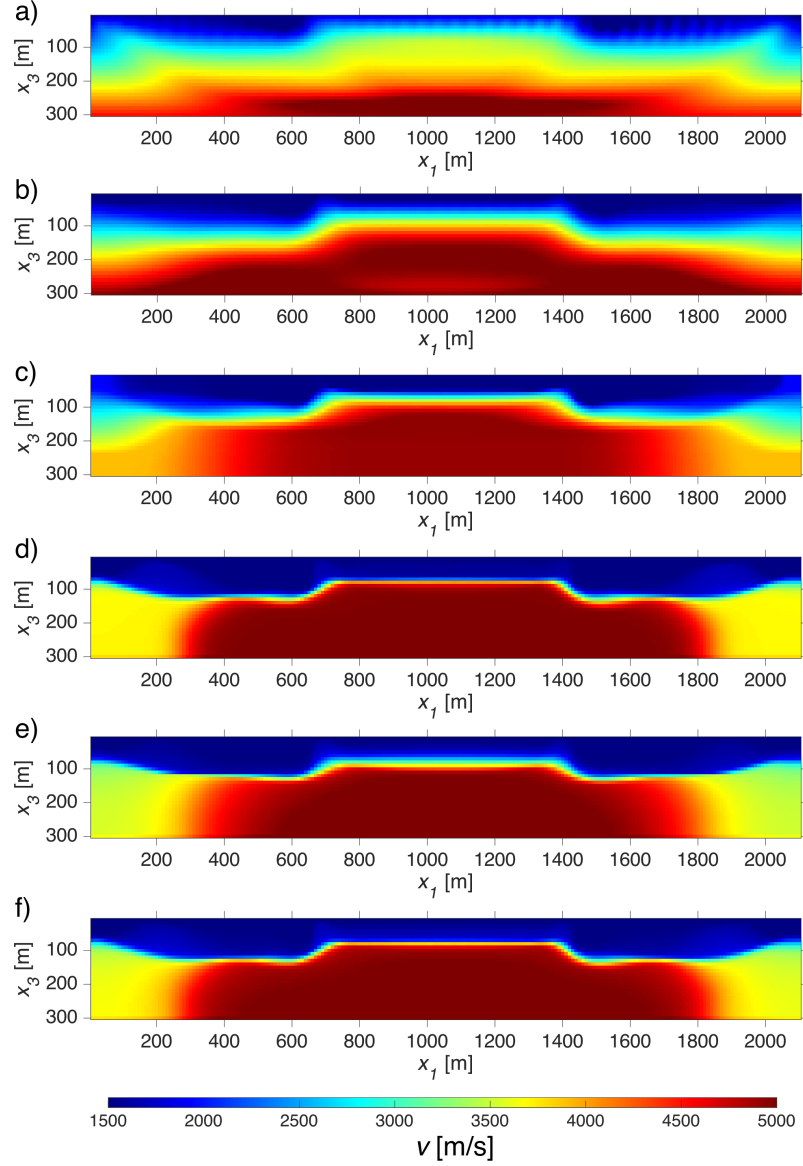


Figure 4.5: Inversions for the model in Figure 4.4 by (a) L-BFGS, (b) H^1 -regularization, (c) MTV, (d) TV-LM with reference parameters, (e) TV-LM with large ϵ_{tv} , (f) TV-LM with large α .

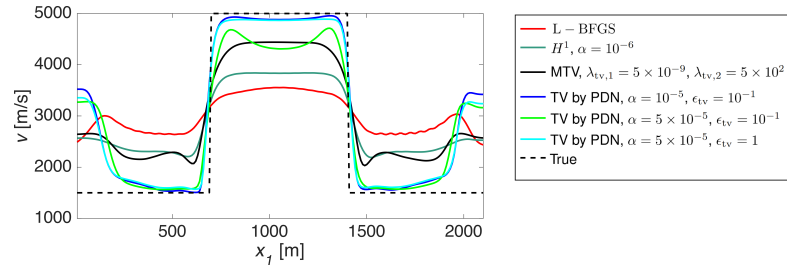


Figure 4.6: Profile at $x_3 = 100$ m for Figure 4.5.

Table 4.1: Tuning parameters in the inversion for the toy model.

Method	Tuning parameters	Iterations
L-BFGS	σ_{gauss} is set 3,2,1	Each σ_{gauss} is assigned with three iteration
H^1	$\alpha = 10^{-6}$	9
MTV	$\lambda_{\text{tv},1} = 10^{-9}$, $\lambda_{\text{tv},2} = 10^{-7}$	30
TV-LM with reference parameters	$\alpha = 10^{-5}$, $\epsilon_{\text{tv}} = 10^{-1}$	9 outer iterations
TV-LM with large α	$\alpha = 5 \times 10^{-5}$, $\epsilon_{\text{tv}} = 10^{-1}$	30 outer iterations
TV-LM with large ϵ_{tv}	$\alpha = 10^{-5}$, $\epsilon_{\text{tv}} = 10$	9 outer iterations

is cheap. H^1 regularization behaves poorly, as a smooth prior is assumed. MTV is slowest in both data and model misfit decays, though it significantly shrinks the L^2 -norm model misfit and the Bregman distance. Its boundary reconstruction is worse than TV-LM with reference parameters. TV-LM with reference parameters is efficient and results in the lowest L^2 -norm model misfit and Bregman distance. With a larger α , TV-LM converges more slowly. Both the data and model misfits are slightly larger for TV-LM with large ϵ_{tv} , compared to TV-LM with reference parameters, as the latter tends to construct smoother interface.

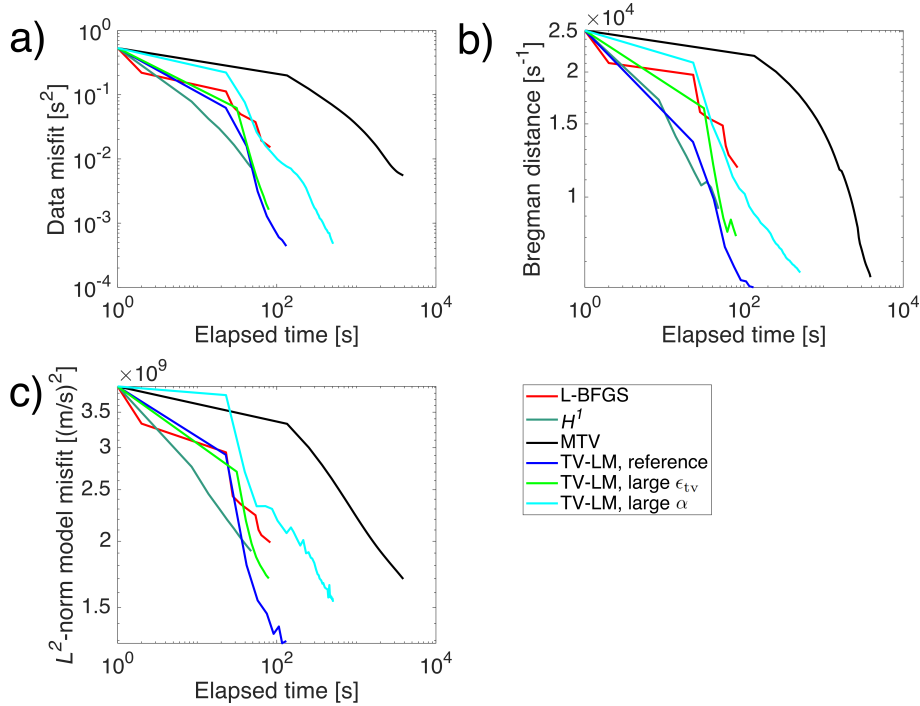


Figure 4.7: Inversion performance for the toy model. (a) Data misfit. (b) Bregman distance, $D_{\text{TV}_{\epsilon_{\text{tv}}}}(v_t, v_i)$. $\epsilon_{\text{tv}} = 10^{-2}$ for this and subsequent evaluation of Bregman distance in the inversion performance analysis. (c) L^2 -norm model misfit.

Table 4.2: Tuning parameters in the inversion for the Marmousi model.

Method	Tuning parameters	Iterations
L-BFGS	σ_{gauss} is set 3,2,1.	Each σ_{gauss} is assigned with 10 iterations.
H^1	α is linearly ranging from 10^{-4} to 10^{-6} .	10
MTV	$\lambda_{\text{tv},1} = 10^{-8}$. $\lambda_{\text{tv},2} = 10^{-9}$.	10
TV-LM	α is linearly ranging from 10^{-4} to 10^{-7} . $\epsilon_{\text{tv}} = 1$.	10 outer iterations

4.4.3 Marmousi model

We examine the algorithm using a more realistic model: the Marmousi model (Figure 4.8a). The starting model of our inversion is given in Figure 4.8b. We compare L-BFGS, MTV, H^1 regularization and TV-LM.

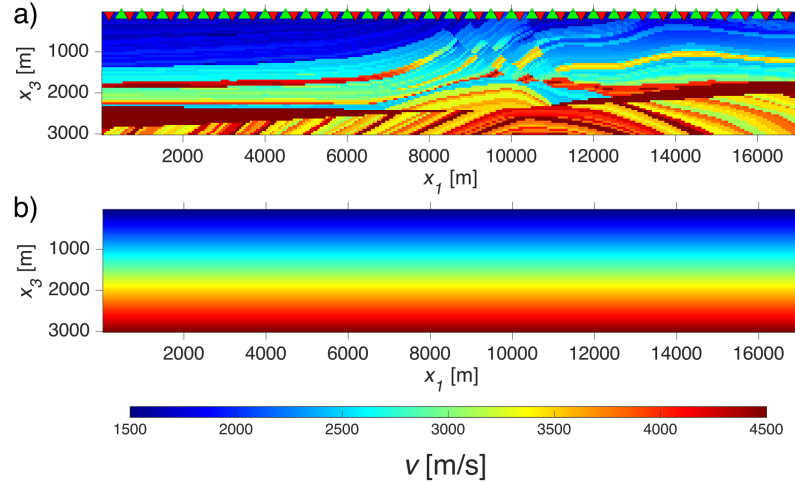


Figure 4.8: (a) Marmousi model. (b) Starting model for the inversion. The model is discretized into 340 by 60 grids. The maximum offset for the source-receiver pair is set at 10000 m.

We need to choose the appropriate parameters using the continuation method, so finding the maximum and minimum regularization parameters is key. For H^1 and TV-LM, we choose first large regularization parameters, which produce relatively homogeneous models for the first several outer iterations. We then test the lower bound, producing source and receiver imprints. Once the two bounds are set, we use the continuation to decrease the regularization parameters gradually. For MTV, the tuning parameters are set based on trial and error. The parameters are given in Table 4.2.

Figure 4.9 displays the inversion results. All four methods provide robust velocity estimations, but TV-LM produces the sharpest velocity interface. As shown in Figure 4.10, all methods effectively reduce the data misfit. L-BFGS exhibits fluctu-

ations in the Bregman distance, while TV-LM achieves the lowest levels of both the Bregman distance and the L^2 -norm model misfit. Although MTV is also robust, it demonstrates lower efficiency and worse boundary reconstruction than TV-LM.

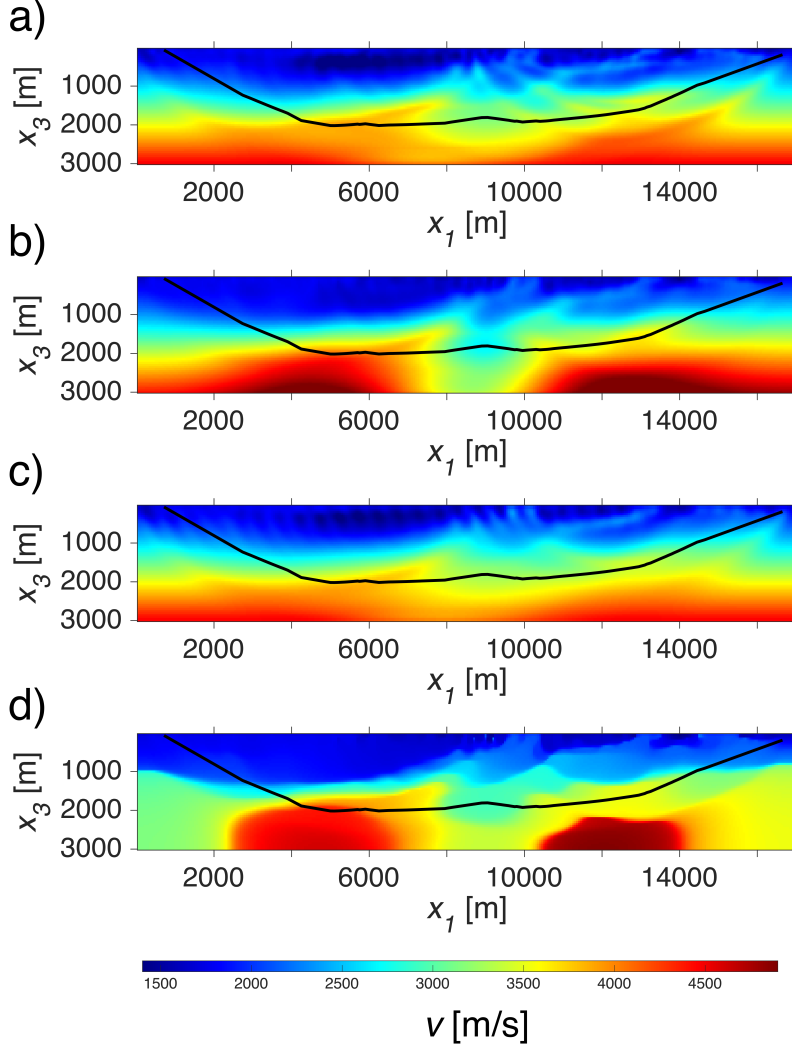


Figure 4.9: Inversions for the Marmousi model by (a) L-BFGS, (b) H^1 , (c) MTV and (d) TV-LM. We use the adjoint travel time to identify the region where the inversion is trusted, as in Figure 4.2. The area above the black line is the trusted region. The model misfit and the Bregman distance are only evaluated for the trusted region. This applies to the following experiments in this work.

4.4.4 2004 BP model

Finally, we test on the 2004 BP model in two scenarios: one with clean data and the other with noisy data. Figure 4.11 shows the true model, the initial model, and the configurations. The inversion parameters for these tests are detailed in Table 4.3.

The noise-free inversion results are presented in Figure 4.12. All four methods successfully reconstruct a low-velocity layer bounded by a high-velocity bottom.

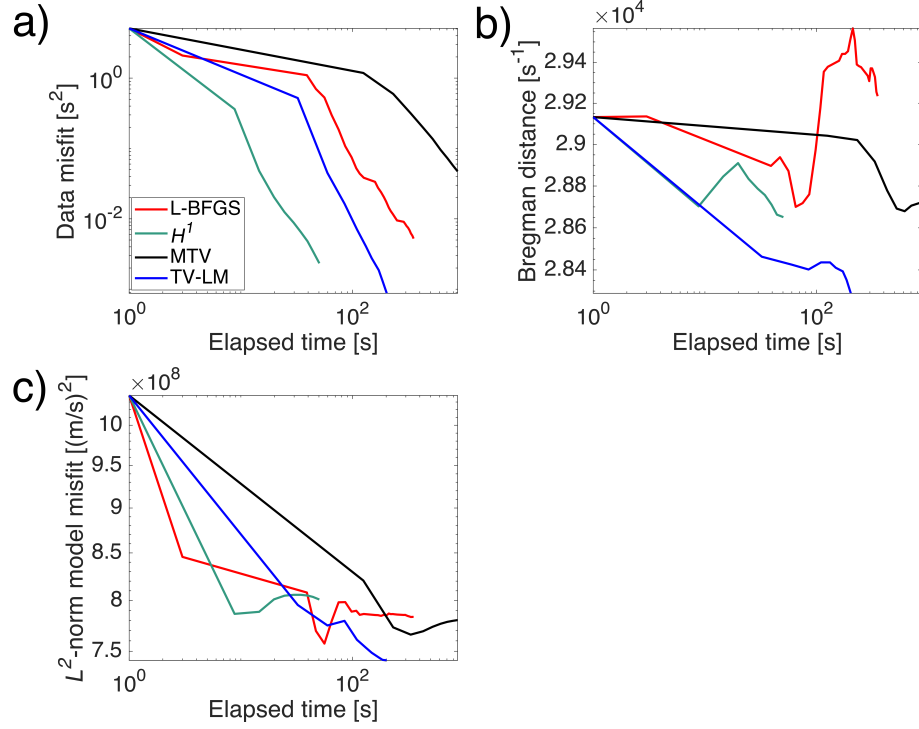


Figure 4.10: Inversion performance for the Marmousi model.

However, the results from L-BFGS and H^1 exhibit blurring at the boundary interface (Figures 4.12a,b). The boundary reconstructed by Figure 4.12c is similar to Figure 4.12b. The result obtained using TV-LM can recover the sharpest boundary (Figure 4.12d).

Figure 4.13 demonstrates that TV-LM is the most effective method among those tested. MTV also performs well in terms of both the Bregman distance and the L^2 -norm, but it requires the most time. L-BFGS and H^1 exhibit similar performance.

For the noisy case, we generate noisy observed data $\tilde{d}_{\text{obs}}$ such that $\tilde{d}_{\text{obs}} = d_{\text{obs}}N_{\text{dis}}(1, 0.02)$, which means a random number that satisfies normal distribution with mean of 1 and standard deviation of 0.02 for this shot gather. The inversion parameters are given

Table 4.3: Tuning parameters in the noise-free inversion for the 2004 BP model.

Method	Tuning parameters	Iterations
L-BFGS	σ_{gauss} is set 6,4,2,1.	Each σ_{gauss} is assigned with 20 iterations.
H^1	α is linearly ranging from -1 to 3.2×10^{-4} .	40
MTV	$\lambda_{\text{tv},1} = 10^{-7}$. $\lambda_{\text{tv},2} = 10^{-6}$.	20
TV-LM	α is linearly ranging from 3×10^{-1} to 3.2×10^{-4} . $\epsilon_{\text{tv}} = 10^{-1}$.	20 outer iterations

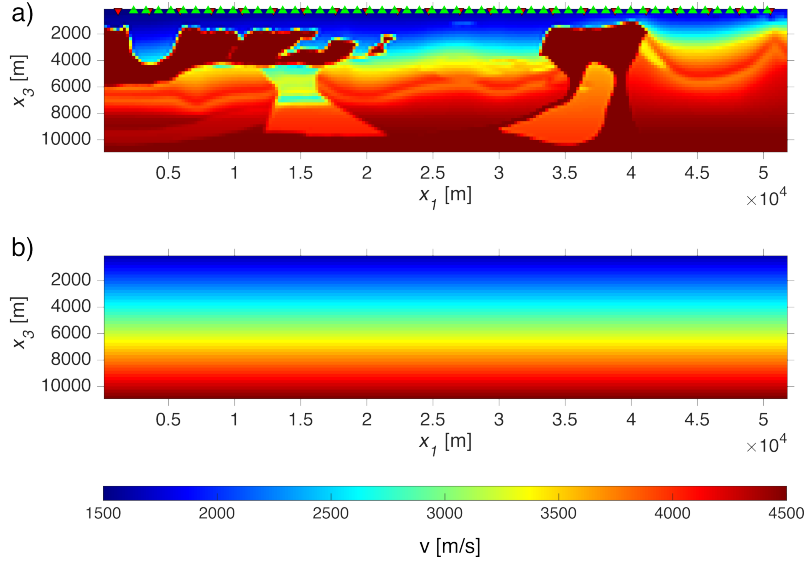


Figure 4.11: (a) 2004 BP model. (b) The initial model for the inversion. The model is discretized into 220 by 46 grids. The maximum offset for the source-receiver pair is set at 43200 m.

in Table 4.4. Figures 4.14 and 4.15 show the inversion results. The inversion results show that all methods perform worse with noisy data compared to the noise-free case. However, TV-LM still constructs the best velocity model.

4.5 Conclusions

We have proposed a travel time inversion algorithm based on Bregman iterative total variation regularization with Levenberg–Marquardt-type linearization. The main goal of this study is to find a feasible and relatively cheap solution to reconstruct

Table 4.4: Tuning parameters in the noisy inversion for the 2004 BP model. The iteration number shown in this table is the maximum possible iteration. The progress is stopped if the method stagnates or the data misfit reaches the noise level.

Method	Tuning parameters	Iterations
L-BFGS	σ_{gauss} is set 6,4,3.	Each σ_{gauss} is assigned with 10 iterations.
H^1	α is linearly ranging from 1 to 3.2×10^{-3} .	40.
MTV	$\lambda_{\text{tv},1}$ is set 10^{-7} . $\lambda_{\text{tv},2} = 10^{-6}$.	20.
TV-LM	α is linearly ranging from 3×10^{-1} to 3.2×10^{-4} . $\epsilon_{\text{tv}} = 10^{-1}$.	20 outer iterations.

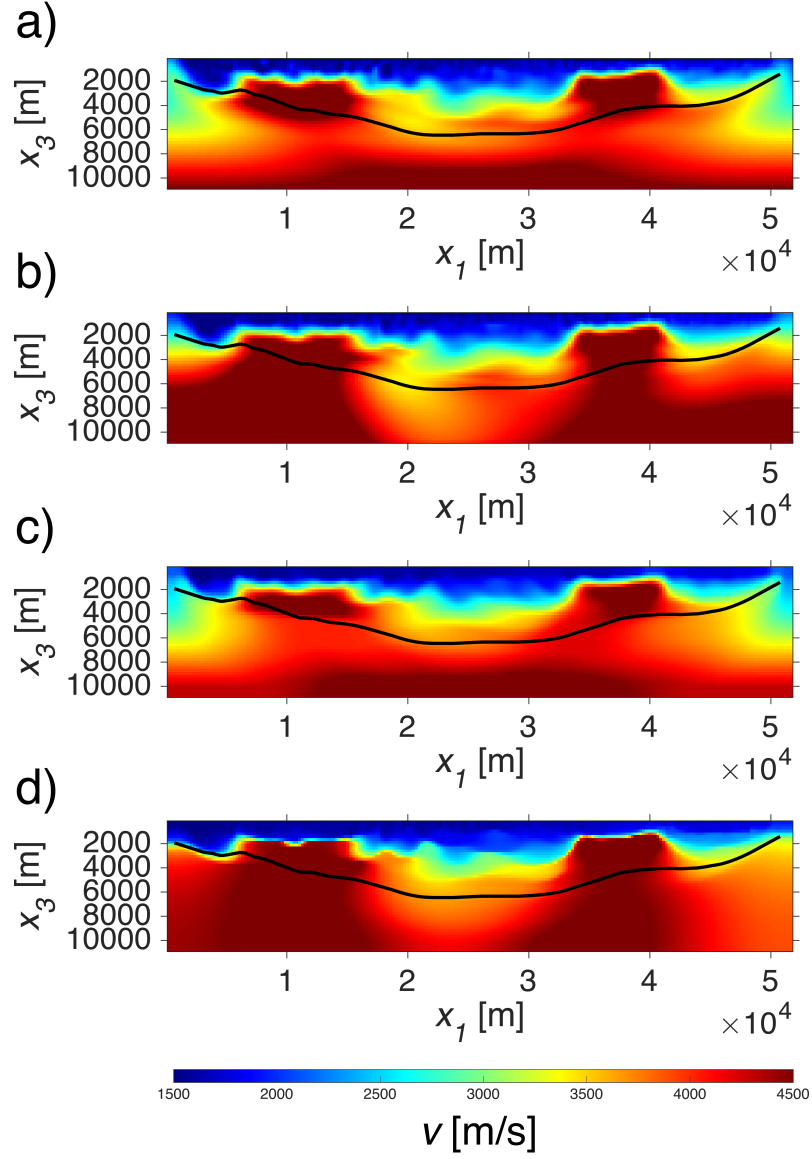


Figure 4.12: Noise-free inversions for the 2004 BP model by (a) L-BFGS, (b) H^1 , (c) MTV and (d) TV-LM.

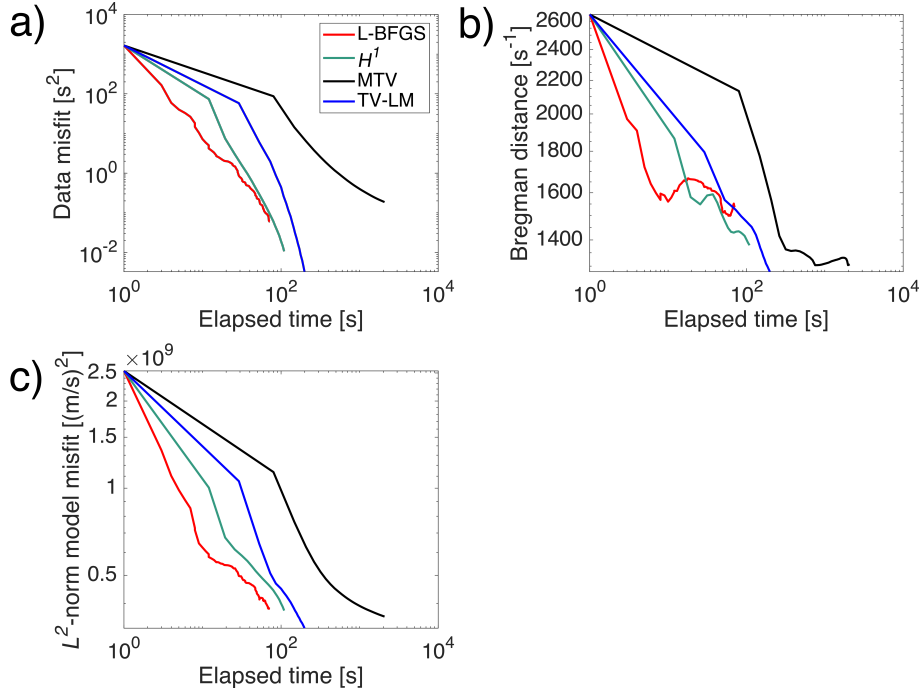


Figure 4.13: Inversion performance for the 2004 BP model under noise-free conditions.

subsurface velocity with a prior comprising a sharp interface. Although TV regularization is usually more expensive than the smooth reconstruction technique (H^1 regularization), with our approach we can use efficient computational methods for handling the TV minimization problems. In particular, the problem in each outer iteration is convex as the forward operator is linearized in the data misfit term. To solve the TV optimality condition we can apply the rapidly and globally convergent primal-dual Newton's method. In contrast, MTV solves the L^2 problem followed by an L^1 problem at each iteration. While these subproblems are straightforward to solve, their weak coupling makes parameter tuning challenging and can lead to slow convergence and worse reconstructions than TV-LM.

In all of our experiments, TV-LM shows the best construction of the original model (among lowest Bregman distance and L^2 -norm model misfit). The L-BFGS method is efficient, but it yields smooth reconstructions and does not always guarantee a lower Bregman distance with iteration. The H^1 method is reliable, though it tends to reconstruct a smooth interface, and usually, one needs to use the continuation method to determine α to yield a sharper interface. In all of our examples, MTV converges slowest. In practice, MTV has two parameters to tune $\lambda_{tv,1}$ and $\lambda_{tv,2}$, making the inversion more sensitive to input parameters. Though TV-LM also contains two tuning parameters, α and ϵ_{tv} , one can fix ϵ_{tv} empirically as it only affects the sharpness of the result. For α , the upper and lower bounds can be sampled, and the continuation is applied in the outer iterations. Both MTV and TV-LM are suitable for reconstructing sharp boundaries. However, TV-LM is always more efficient and reliable for the sharp interface reconstruction.

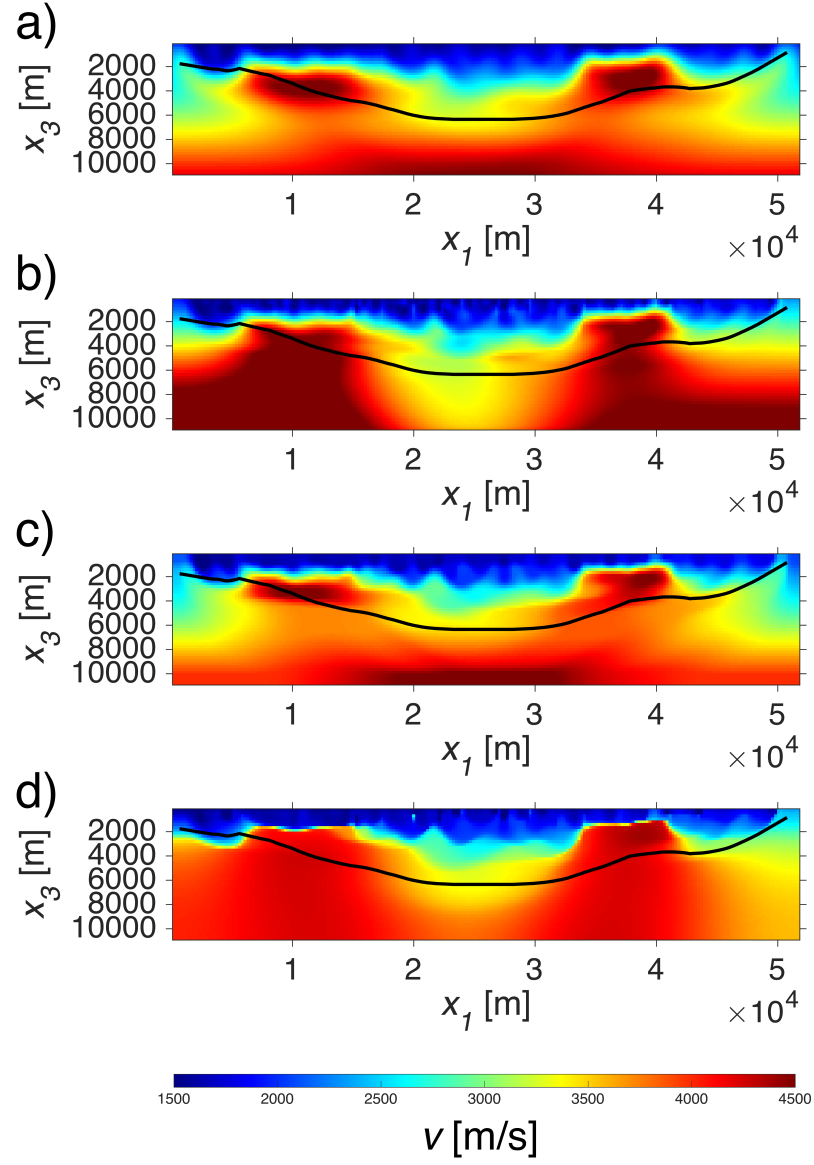


Figure 4.14: Inversions of noisy data for the 2004 BP model by (a) L-BFGS, (b) H^1 , (c) MTV and (d) TV-LM.

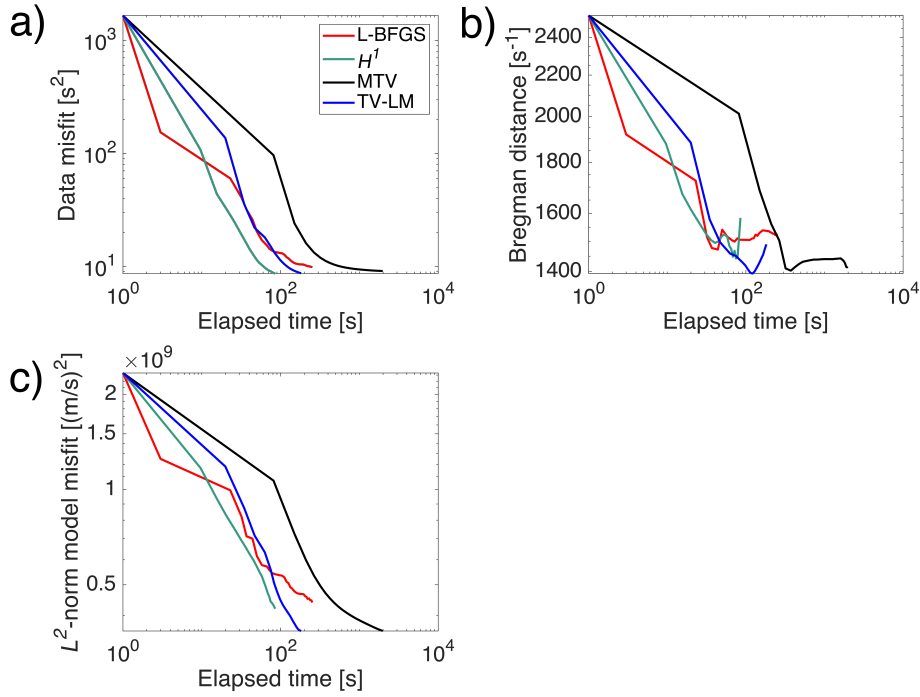


Figure 4.15: Inversion performance for the 2004 BP model under noisy conditions.

In this study, we show only one PDE example – the eikonal equation. However, the inversion framework can be applied to any physics with a computationally feasible linearization. The result show a clear potential of the method when targeting salt intrusion, volcanics, or any medium presenting sharp vertical contrasts. This is true even in the presence of high ambient noise, making it a suitable technique for marine environments.

Chapter 5

Simultaneous velocity and reflectivity estimation using extended source waveform inversion

The work presented in this chapter has been submitted as a journal article: Zhang, Y., and De Siena, L. (2025). Simultaneous velocity and reflectivity estimation using extended source waveform inversion. *GEOPHYSICS*.

Y.Z. formulated the research idea, derived the equations, developed the code, performed validation, and drafted the manuscript. L.D.S. reviewed the manuscript and contributed to its revision and improvement.

Abstract

Joint Migration Inversion (JMI) aims to estimate background velocity and reflectivity simultaneously. However, like conventional Full Waveform Inversion (FWI), local minima are often encountered due to the nonlinearity between model parameters and the observed wavefield. This study introduces a new algorithm, Extended Source Joint Migration Inversion (ES-JMI), which addresses these challenges through two strategies: (1) using a velocity-impedance parameterized wave equation as the forward model, and (2) implementing Extended Source waveform Inversion (ESI) to mitigate local minima issues. The velocity-impedance parameterization considers reflectivity in the forward problem, enabling both velocity and reflectivity updates in the inversion process. We demonstrate, from the plane wave perspective, that the velocity gradient tends to build low-wavenumber components. In contrast, the impedance gradient enhances high-wavenumber components, explaining the role of velocity as the background model and impedance as the basis for reflectivity. Our

ES-JMI employs a source defocusing annihilator-weighted quadratic penalty formulation of the constrained optimization problem, solved using a reduced approach. In ES-JMI, the data augmented wavefield is assumed to be dependent on model parameters and is approximated efficiently using the scalar fitting method. Then, the data augmented wavefield is used to compute the subsurface model update. Numerical tests on the Marmousi and 2004 BP models, comparing conventional FWI with least-square migration, reduced-space JMI, Wavefield Reconstruction JMI (WR-JMI), and ES-JMI, demonstrate that ES-JMI delivers more reliable results in challenging settings.

5.1 Introduction

A key objective of seismic exploration is to estimate both the velocity and reflectivity of the subsurface (i.e., Joint Migration Inversion or JMI). Full Waveform Inversion (FWI) is a powerful method for generating high-resolution velocity models (Plessix, 2006; Robein, 2010; Virieux & Operto, 2009). Reflectivity models are typically derived through migration techniques based on these velocity models. Least-squares reverse time migration (LS-RTM), which relies on the wave equation, is particularly appealing as it utilizes the full wavefield to construct the reflectivity model, including multiples (Bai & Yilmaz, 2021; Liu & Peter, 2018; Mu et al., 2020; Zand, Ghasemzadeh, Gholami, & Malcolm, 2023).

Building on the strengths of both FWI and LS-RTM, D. J. Verschuur, Staal, and Berkhout (2016) proposed a joint migration inversion (JMI) algorithm that utilizes the full wavefield. This algorithm employs local propagation and scattering operators to simulate the wavefield rather than parameterizing the model based on the wave equation. However, a significant limitation of this method is its inability to handle horizontal wave propagation effectively (E. Verschuur, Qu, Kim, & Sun, 2020). Additionally, because this approach does not rely on a Partial Differential Equation (PDE) to describe the wavefield, extending the search space to mitigate local minima is not straightforward. In contrast, W. Zhou, Brossier, Operto, and Virieux (2015) developed a wave-equation-based method and proposed a JMI algorithm that formulates a cost function decomposed into contributions from diving and reflected waves. This approach separates the low-wavenumber and high-wavenumber components to update the background velocity and reflectivity, respectively. However, the need for alternating updates of velocity and reflectivity increases the computational burden, and the algorithm requires preprocessing of the observed data to distinguish between the diving and reflection components.

Reverse Time Migration (RTM) images are often contaminated by low-wavenumber artifacts when the background velocity model contains sharp interfaces (Fletcher, Fowler, Kitchenside, & Albertin, 2006; Mulder & Plessix, 2004). Early research indicated that using an impedance kernel could enhance the high-wavenumber components in the reflectivity image (Berkhout, 2014). Douma, Yingst, Vasconcelos, and

Tromp (2010) demonstrated that, in the velocity-impedance parameterized acoustic wave equation, employing the impedance kernel is equivalent to applying a Laplacian filter to the image, which acts as a high-pass filter to amplify the high-wavenumber content. Ramos-Martinez et al. (2016) proposed using the velocity kernel for long-wavelength updates in waveform inversion. Compared to the conventional velocity update, this long-wavelength velocity update reduces the high-wavenumber reflectivity imprint, resulting in better illumination of deeper targets. Y. Yang et al. (2020) suggested using a velocity-reflectivity parameterized acoustic wave equation for forward modeling. Based on this parameterization, Y. Yang et al. (2020) explored the simultaneous estimation of velocity and reflectivity, treating reflectivity as an auxiliary variable that can be derived from impedance. This inversion scheme is equivalent to performing FWI for the smooth or background velocity and Least-Squares Reverse Time Migration (LS-RTM) for the reflectivity.

The formulation of the velocity-impedance (or velocity-reflectivity) parameterized wave equation provides an integrated approach to predicting the full wavefield, relying on two key parameters: velocity and reflectivity. This PDE can be used as a constraint for seismic data fitting, which is then converted into a Lagrangian formulation (Haber & Ascher, 2000). Typically, this approach involves finding the stationary point by applying the Newton method under the Karush-Kuhn-Tucker (KKT) conditions. However, this approach requires the storage of all wavefields, making it impractical to solve large-scale industry problems. To address this, practitioners often employ a reduced system where the physics (i.e., the forward problem equation) is fully enforced (Plessix, 2006; Virieux & Operto, 2009). While this simplifies the computational burden, it introduces a highly non-linear relationship between the model parameters and the cost function (e.g., van Leeuwen & Herrmann, 2013). Due to the oscillatory nature of the wave equation, the reduced method is prone to cycle-skipping issues and can easily become trapped in a poor local minimum (Jiménez Tejero, Dagnino, Sallarès, & Ranero, 2015; Pan & Innanen, 2016; Yao, Silva, Warner, Wu, & Yang, 2019). This problem is exacerbated when low-frequency data are missing or when a good initial model is unavailable.

Efforts from various perspectives have been made to mitigate the cycle-skipping issue in waveform inversion. The primary strategy is to develop a cost function that is smoother and has a larger attraction basin, thereby reducing the inversion’s dependence on the initial model or the availability of low-frequency data. For example, more robust metrics for measuring data misfit have been proposed, such as adaptive waveform inversion (Warner & Guasch, 2016), dynamic time warping (Ma & Hale, 2013; M. Wang et al., 2016), and optimal transport methods (Engquist, Froese, & Yang, 2016; Ramos-Martínez, Qiu, Kirkebo, Valenciano, & Yang, 2018).

Another approach involves extending the search space, making the dependency on the model parameters more convex. Techniques in this category include Contrast Source Inversion (CSI) (Abubakar, Hu, Habashy, & van den Berg, 2009; Abubakar, Hu, van den Berg, & Habashy, 2008), Wavefield Reconstruction Inversion (WRI) (Aravkin & van Leeuwen, 2012; van Leeuwen & Herrmann, 2016), and volume Source Extended Inversion (ESI) (G. Huang et al., 2018b). These model extension meth-

ods share a common goal: to expand the search space by relaxing the physical constraints, thereby rendering the cost function more convex.

Due to the ease of implementation of the frequency-domain wave equation, many early studies focused on solving problems in the frequency domain. The wavefield reconstruction inversion was introduced by van Leeuwen and Herrmann (2013), which utilizes a penalty method. This approach optimizes both the wavefield and the model parameters simultaneously. By optimizing for the wavefield, the data augmented wave equation is obtained, and their optimization process involves alternating updates between the data augmented wavefield and the model parameters. Additionally, Aghamiry et al. (2019b) proposed an augmented Lagrangian method, solved using the alternating direction method of multipliers (ADMM). ADMM is less sensitive to the size of the physics penalty but requires the storage of the wavefield and dual variables for each shot.

Intensive efforts have been made to address the challenges of implementing time-domain WRI or ESI. The primary bottleneck lies in computing the time-domain data augmented wavefield. Exact solutions for this wavefield are generally computationally expensive (Aghamiry et al., 2020a; Guo, Operto, Gholami, & Aghamiry, 2024; Rizzuti et al., 2019) because they require inner iterations to update the wavefield. C. Wang et al. (2016) introduced an approximation to the data augmented wavefield by considering a large value for the physics penalty. Building on this, Gholami et al. (2022) proposed a scalar fitting method for the data space diagonal Hessian, which is equivalent to performing one inner iteration of the steepest descent to solve the extended data residual, offering better accuracy than the approach by C. Wang et al. (2016). Furthermore, more advanced methods for approximating the data space Hessian were developed by Guo et al. (2024).

We refer to the vanilla penalty method as WRI (e.g., van Leeuwen & Herrmann, 2013, 2016), and the weighted source extension using an annihilator as (ESI) (e.g., Fu & Symes, 2017b; G. Huang et al., 2018b). A detailed comparison between WRI and ESI is provided in P. Yang and Zhou (2024).

In this work, we propose a new algorithm called Extended Source Joint Migration Inversion (ES-JMI) in the time domain. This algorithm utilizes the velocity-impedance parameterized acoustic wave equation as the forward problem, integrating both background velocity and reflectivity into the wavefield’s dependencies. This formulation offers two key advantages: (1) the simultaneous updating of velocity and impedance (reflectivity) and (2) a straightforward implementation of source extension to address the local minimum issue. Similar to the data augmented wavefield approach used in WRI (Gholami et al., 2022), we approximate the ESI data augmented wavefield through the scalar fitting. The optimization is conducted within a reduced objective framework (Aravkin & van Leeuwen, 2012; van Leeuwen & Herrmann, 2016), which eliminates the need to store all wavefields simultaneously for each source.

We performed numerical tests on the Marmousi and 2004 BP models, comparing

conventional FWI followed by Least-Squares Reverse Time Migration (LS-RTM), conventional Joint Migration Inversion (JMI), JMI via Wavefield Reconstruction (or WR-JMI), and our proposed ES-JMI. In these tests, low-frequency data was filtered out, and a crude initial iterate was provided. The results demonstrate competitive background velocity estimation and effective reflectivity reconstruction with ES-JMI.

The outline of the work is as follows. In Section 5.2, we describe our forward problem: the acoustic wave equation parameterized in velocity and impedance. Section 5.3 describes the all-at-once method, which is followed by the reduced-space method in Section 5.4. Section 5.5 details the inversion framework, from the computation of the data augmented wavefield to the update of the model parameters. In Section 5.6, we prove that the velocity gradient acts as a low-pass filter while the impedance gradient functions as a high-pass filter. The velocity component is thus favored for background velocity, and the impedance part is utilized to construct reflectivity. Section 5.7 presents numerical tests for the Marmousi model and 2004 BP across different methods. Finally, some open problems are discussed in Section 5.8, and Section 5.9 offers conclusions.

5.2 Forward problem

The acoustic wave equation parameterized in acoustic wave speed v and impedance Z can be cast as (Luo & Wu, 2018)

$$\frac{1}{v(x)Z(x)} \frac{\partial^2 u(x, t)}{\partial t^2} - \nabla \cdot \left(\frac{v(x)}{Z(x)} \nabla u(x, t) \right) = \delta(x - x_s) s(t) , \quad (5.1)$$

where $x \in \Omega \subseteq \mathbb{R}^{N_{\text{dim}}}$ is the N_{dim} -dimensional spatial coordinate; t is the time and $0 \leq t \leq T$; u is the acoustic pressure; $\delta(\cdot)$ is Dirac delta; x_s is the source location; s is the source wavelet. In this work, the wave equations with Perfectly Matched Layers (PML) are solved by the finite-difference method.

The boundary condition for u is given by,

$$u(x, 0) = 0, x \in \Omega , \quad (5.2)$$

$$\frac{\partial u(x, 0)}{\partial t} = 0, x \in \Omega , \quad (5.3)$$

$$u(x, t) = 0, x \in \partial\Omega \text{ and } 0 \leq t \leq T , \quad (5.4)$$

$$\nabla u(x, t) = 0, x \in \partial\Omega \text{ and } 0 \leq t \leq T . \quad (5.5)$$

Equation 5.1 is a linear PDE, thus can be written as

$$\mathbf{A}(\mathbf{v}, \mathbf{Z})\mathbf{u} = \mathbf{b}_{\text{obs}} , \quad (5.6)$$

where $\mathbf{A} \in \mathbb{R}^{(N_x N_t) \times (N_x N_t)}$ is the discrete operator of $\frac{1}{vZ} \frac{\partial^2}{\partial t^2} - \nabla \cdot \left(\frac{v}{Z} \nabla \right)$ and $\mathbf{b}_{\text{obs}} \in \mathbb{R}^{(N_x N_t) \times 1}$ is the discrete $\delta(x - x_s)s(t)$; $\mathbf{u}^{(N_x N_t) \times 1}$ is the column vector for discrete

u ; $\mathbf{v}^{N_x \times 1}$ and $\mathbf{Z}^{N_x \times 1}$ is column vectors consisted of discrete v and Z , respectively. We denote N_x as the number of grid points in space and N_t as the number of time steps.

According to the definition (Y. Yang et al., 2020), the seismic vector reflectivity I_R can be connected to the impedance by

$$I_R = [I_{R,1} \ I_{R,3}]^T = \frac{1}{2Z} \nabla Z . \quad (5.7)$$

5.3 All-at-once method

Our goal is to solve the following optimization problem to estimate both v and Z ,

$$\min_{\mathbf{v}, \mathbf{Z}, \mathbf{u}} E(\mathbf{v}, \mathbf{Z}, \mathbf{u}), \ E = \frac{1}{2} \|\mathbf{R}\mathbf{u} - \mathbf{d}_{\text{obs}}\|_2^2 \text{ s.t. } \mathbf{A}\mathbf{u} = \mathbf{b}_{\text{obs}} , \quad (5.8)$$

where $\mathbf{R} \in \mathbb{R}^{(N_r N_t) \times (N_x N_t)}$ and N_r is the number of receivers. With the reconstruction of Z , one can compute reflectivity I_R directly. Thus, this minimization solves both velocity and reflectivity; This is the reason why the algorithm is named joint migration inversion (JMI).

The associated Lagrangian of equation 5.8 can be written as,

$$\min_{\mathbf{v}, \mathbf{Z}, \mathbf{u}} E_{\text{lag}}(\mathbf{v}, \mathbf{Z}, \mathbf{u}), \ E_{\text{lag}} = \frac{1}{2} \|\mathbf{R}\mathbf{u} - \mathbf{d}_{\text{obs}}\|_2^2 - \mathbf{u}_{\mathbf{a}}^T (\mathbf{A}\mathbf{u} - \mathbf{b}_{\text{obs}}) , \quad (5.9)$$

where $(\cdot)^T$ denotes the transpose; $\mathbf{u}_{\mathbf{a}} \in \mathbb{R}^{(N_x N_t) \times (N_x N_t)}$ is the column vector consisting of discrete Lagrangian multipliers or the adjoint wavefield.

The all-at-once method solves equation 5.9 using the KKT condition, which is the first-order optimality condition for $\mathbf{v}$, $\mathbf{Z}$, $\mathbf{u}$ and $\mathbf{u}_{\mathbf{a}}$,

$$\nabla_{\mathbf{v}} E_{\text{lag}} = \mathbf{0} , \quad (5.10)$$

$$\nabla_{\mathbf{Z}} E_{\text{lag}} = \mathbf{0} , \quad (5.11)$$

$$\nabla_{\mathbf{u}} E_{\text{lag}} = \mathbf{0} , \quad (5.12)$$

$$\nabla_{\mathbf{u}_{\mathbf{a}}} E_{\text{lag}} = \mathbf{0} . \quad (5.13)$$

The system in equations 5.10 - 5.13 can be solved by the iterative Newton-type method. The detailed solution for equation 5.12 is given in equation A17 in Appendix C. The first-order optimality condition for u_a is,

$$\mathbf{A}\mathbf{u} - \mathbf{b}_{\text{obs}} = \mathbf{0} , \quad (5.14)$$

which is just the forward problem stated in equation 5.6.

Equations 5.10 and 5.11 correspond to the gradient calculations for $\mathbf{v}$ and $\mathbf{Z}$, and they are given by equation A24 in Appendix D.

The all-at-once method linearizes equations 5.10 - 5.13 for $\mathbf{v}$, $\mathbf{Z}$, $\mathbf{u}$ and $\mathbf{u}_a$ and solves for the update for all of those variables simultaneously at each iteration (Haber & Ascher, 2000). However, this is usually computationally intractable in the time domain, as it requires storing all the wavefields for different shots so that the dimensionality of the system becomes prohibitively large.

5.4 Reduced objective

Practitioners in geophysics usually seek a simplified version of the full KKT condition: they successively solve equation 5.12 for u , 5.13 for u_a and A24 for the update of v and Z .

As such, the physics equation 5.12 is enforced at each iteration, so that the optimization problem of equation 5.8 becomes,

$$\min_{\mathbf{v}, \mathbf{Z}} E_{\text{red}}(\mathbf{v}, \mathbf{Z}), \quad E_{\text{red}} = \frac{1}{2} \|\mathbf{R}\mathbf{A}(\mathbf{v}, \mathbf{Z})^{-1} \mathbf{b}_{\text{obs}} - \mathbf{d}_{\text{obs}}\|_2^2. \quad (5.15)$$

The relationship between E_{red} and the model parameters, v and Z becomes highly non-convex due to the oscillating nature of the operator $\mathbf{A}(\mathbf{v}, \mathbf{Z})^{-1}$ (van Leeuwen & Herrmann, 2013).

Equation 5.15 simultaneously estimates both velocity and impedance. The estimated impedance can then be used to compute reflectivity, which is referred to as joint migration inversion (JMI) in this work, as in the formulation of Y. Yang et al. (2020).

5.5 Source extension

To mitigate the local minimum issues, one can use the form of the physics penalty method instead of the reduced-space approach with enforced physics as in equation 5.15. The physics penalty optimization problem reads,

$$\min_{\mathbf{v}, \mathbf{Z}, \mathbf{u}} E_{\text{es}, \lambda}(\mathbf{v}, \mathbf{Z}, \mathbf{u}), \quad E_{\text{es}, \lambda} = \frac{1}{2} \|\mathbf{R}\mathbf{u} - \mathbf{d}_{\text{obs}}\|_2^2 + \frac{\lambda^2}{2} \|\mathbf{C}_{\mathbf{A}} (\mathbf{A}\mathbf{u} - \mathbf{b}_{\text{obs}})\|_2^2, \quad (5.16)$$

where $\mathbf{C}_{\mathbf{A}}$ is the so-called annihilator operator. In this work, we name equation 5.16 ES-JMI.

Some choices exist for $\mathbf{C}_{\mathbf{A}}$ (Symes, 2008). In this work, we use the distance-dependent annihilator (G. Huang et al., 2018a, 2018b; P. Yang & Zhou, 2024),

$$C_{\mathbf{A}}(x, x_s) = |x - x_s| + \varepsilon_{\mathbf{A}}, \quad (5.17)$$

where ε_A is a small positive number to avoid zero division in the optimization process; $C_A(x, x_s)$ is the continuous form annihilator. We choose ε_A to be 0.1 times the mean value of $|x - x_s|$. If $\mathbf{C}_A$ is chosen as an identity matrix, then equation 5.16 is reduced to wavefield reconstruction inversion based JMI (WR-JMI),

$$\min_{\mathbf{v}, \mathbf{Z}, \mathbf{u}} E_{\text{wr}, \lambda}(\mathbf{v}, \mathbf{Z}, \mathbf{u}), \quad E_{\text{wr}, \lambda} = \frac{1}{2} \|\mathbf{R}\mathbf{u} - \mathbf{d}_{\text{obs}}\|_2^2 + \frac{\lambda^2}{2} \|\mathbf{A}\mathbf{u} - \mathbf{b}_{\text{obs}}\|_2^2. \quad (5.18)$$

Thus, WR-JMI can be viewed as a special case of ES-JMI.

The problem in equation 5.16 can be solved by alternately updating the source extension $\mathbf{u}$ and the model parameters $\mathbf{v}$ and $\mathbf{Z}$. This scheme is similar to that proposed by van Leeuwen and Herrmann (2013). However, the implementation in the time domain requires further development: we consider a reduced objective function based on equation 5.16,

$$\min_{\mathbf{v}, \mathbf{Z}} E_{\text{res}, \lambda}(\mathbf{v}, \mathbf{Z}), \quad E_{\text{res}, \lambda} = E_{\text{es}, \lambda}(\mathbf{v}, \mathbf{Z}, \tilde{\mathbf{u}}_\lambda) = \min_{\mathbf{u}} E_{\text{es}, \lambda}(\mathbf{v}, \mathbf{Z}, \mathbf{u}), \quad (5.19)$$

where

$$\tilde{\mathbf{u}}_\lambda(\mathbf{v}, \mathbf{Z}) = \underset{\mathbf{u}}{\operatorname{argmin}} E_{\text{es}, \lambda}(\mathbf{v}, \mathbf{Z}, \mathbf{u}). \quad (5.20)$$

The inversion now involves two tasks: solving the problem in equation 5.20 for the data augmented wavefield and then using equation 5.19 for the model update. In the following subsections, we explain how to address these two tasks in the time domain.

5.5.1 Data augmented wavefield

The objective in this subsection is to solve equation 5.20. Its first-order optimality condition reads,

$$(\mathbf{R}^T \mathbf{R} + \lambda^2 \mathbf{A}^T \mathbf{C}_A^T \mathbf{C}_A \mathbf{A}) \mathbf{u} = \mathbf{R}^T \mathbf{d}_{\text{obs}} + \lambda^2 \mathbf{A}^T \mathbf{C}_A^T \mathbf{C}_A \mathbf{b}_{\text{obs}}. \quad (5.21)$$

The solution to the normal equation 5.21 provides $\tilde{\mathbf{u}}_\lambda$ in equation 5.20. Solving equation 5.21 requires defining new variables,

$$\mathbf{A}\mathbf{u} = \mathbf{b}_e, \quad (5.22)$$

$$\mathbf{b}_e - \mathbf{b}_{\text{obs}} = \frac{1}{\lambda^2} \Delta \mathbf{b}, \quad (5.23)$$

$$\Delta \mathbf{d} = \mathbf{d}_{\text{obs}} - \mathbf{R}\mathbf{A}^{-1} \mathbf{b}_{\text{obs}}, \quad (5.24)$$

where $\Delta \mathbf{b}$ is named source extension, which measures the physics error. $\Delta \mathbf{d}$ is the reduced space data difference.

Equations 5.22, 5.23 and 5.24 allow for elimination of $\mathbf{u}$, $\mathbf{b}_{\text{obs}}$ and $\mathbf{d}_{\text{obs}}$, so that equation 5.21 can be recast in the best form to solve for $\Delta \mathbf{b}$,

$$\left(\frac{1}{\lambda^2} (\mathbf{R}\mathbf{A}^{-1})^T \mathbf{R}\mathbf{A}^{-1} + \mathbf{C}_A^T \mathbf{C}_A \right) \Delta \mathbf{b} = (\mathbf{R}\mathbf{A}^{-1})^T \Delta \mathbf{d}. \quad (5.25)$$

Equation 5.25 is similar to a deblurred problem where $\Delta \mathbf{b}$ is the deblurred image and $(\mathbf{R}\mathbf{A}^{-1})^T \Delta \mathbf{d}$ is the noisy image. The derivation from equation 5.21 to equation 5.25 can be found in Appendix E. After the solution for $\Delta \mathbf{b}$ in equation 5.25 is obtained ($\Delta \tilde{\mathbf{b}}_\lambda$), equations 5.22 and 5.23 can be used to obtain the solution of the data augmented wavefield $\tilde{\mathbf{u}}_\lambda$ from equation 5.20.

The question is now how to solve equation 5.25. Let us introduce a variable,

$$\mathbf{S} = \mathbf{R}\mathbf{A}^{-1} . \quad (5.26)$$

Equation 5.25 can thus be simplified as,

$$\begin{aligned} \Delta \tilde{\mathbf{b}}_\lambda &= \left(\frac{1}{\lambda^2} \mathbf{S}^T \mathbf{S} + \mathbf{C}_A^T \mathbf{C}_A \right)^{-1} \mathbf{S}^T \Delta \mathbf{d} \\ &= (\mathbf{C}_A^T \mathbf{C}_A)^{-1} \mathbf{S}^T \left(\frac{1}{\lambda^2} \mathbf{S} (\mathbf{C}_A^T \mathbf{C}_A)^{-1} \mathbf{S}^T + \mathbb{I} \right)^{-1} \Delta \mathbf{d} . \end{aligned} \quad (5.27)$$

where $\mathbb{I}$ is the identity matrix.

Equation 5.27 converts the source side Hessian,

$$\mathbf{H}_S = \frac{1}{\lambda^2} \mathbf{S}^T \mathbf{S} + \mathbf{C}_A^T \mathbf{C}_A . \quad (5.28)$$

to the receiver side Hessian,

$$\mathbf{H}_R = \frac{1}{\lambda^2} \mathbf{S} (\mathbf{C}_A^T \mathbf{C}_A)^{-1} \mathbf{S}^T + \mathbb{I} . \quad (5.29)$$

For a detailed derivation of equation 5.27, see Appendix F. Compared to equation 45 by P. Yang and Zhou (2024), which solves the same problem, our approach contains fewer wave equation evaluations and is less computationally expensive.

Equation 5.27 becomes,

$$\Delta \tilde{\mathbf{b}}_\lambda = (\mathbf{C}_A^T \mathbf{C}_A)^{-1} \mathbf{S}^T \mathbf{H}_R^{-1} \Delta \mathbf{d} . \quad (5.30)$$

We employ one iteration of a steepest-descent algorithm to solve equation 5.30 by Gholami et al. (2022). We set

$$\mathbf{F} = \mathbf{H}_R^{-1} . \quad (5.31)$$

Then,

$$\begin{aligned} \Delta \mathbf{d} &= \mathbf{F} \mathbf{H}_R \Delta \mathbf{d} \\ &= \mathbf{F} \left(\frac{1}{\lambda^2} \mathbf{S} (\mathbf{C}_A^T \mathbf{C}_A)^{-1} \mathbf{S}^T + \mathbb{I} \right) \Delta \mathbf{d} , \\ &= \mathbf{F} \Delta \mathbf{d}_b \end{aligned} \quad (5.32)$$

where

$$\Delta \mathbf{d}_b = \left(\frac{1}{\lambda^2} \mathbf{S} (\mathbf{C}_A^T \mathbf{C}_A)^{-1} \mathbf{S}^T \Delta \mathbf{d} + \Delta \mathbf{d} \right). \quad (5.33)$$

If we assume

$$\mathbf{F} = \alpha_{\text{hess}} \mathbb{I} \quad (5.34)$$

the solution α_{hess} can be computed by solving the following problem,

$$\alpha_{\text{hess}} = \underset{\alpha_{\text{hess}}}{\operatorname{argmin}} \|\alpha_{\text{hess}} \Delta \mathbf{d}_b - \Delta \mathbf{d}\|_2^2. \quad (5.35)$$

The optimality condition of equation 5.35 takes the form,

$$\alpha_{\text{hess}} = \frac{\Delta \mathbf{d}_b^T \Delta \mathbf{d}}{\Delta \mathbf{d}_b^T \Delta \mathbf{d}_b}. \quad (5.36)$$

This condition is evaluated for every shot gather.

According to Guo et al. (2024), λ can be estimated using,

$$\alpha_{\text{fraction}} = \frac{\left\| \frac{1}{\lambda^2} \mathbf{S} (\mathbf{C}_A^T \mathbf{C}_A)^{-1} \mathbf{S}^T \Delta \mathbf{d} \right\|_2^2}{\|\Delta \mathbf{d}\|_2^2}, \quad (5.37)$$

where α_{fraction} is selected around 10^{-2} in our study. Having the solution $\Delta \tilde{\mathbf{b}}_\lambda$, one can insert it into equations 5.23 and 5.22 to obtain the reduced space data-augmented wavefield $\tilde{\mathbf{u}}_\lambda$.

5.5.2 Model parameters update

With the solution of $\tilde{\mathbf{u}}_\lambda$ inserted into equation 5.19, we can update the model parameters. From equation 5.19, the reduced space model update can be recast as,

$$\begin{aligned} \nabla_{\mathbf{v}} E_{\text{res},\lambda}(\mathbf{v}, \mathbf{Z}) &= \nabla_{\mathbf{v}} E_{\text{es},\lambda}(\mathbf{v}, \mathbf{Z}, \tilde{\mathbf{u}}_\lambda(\mathbf{v}, \mathbf{Z})) + \nabla_{\mathbf{u}} E_{\text{es},\lambda}(\mathbf{v}, \mathbf{Z}, \tilde{\mathbf{u}}_\lambda(\mathbf{v}, \mathbf{Z})) \nabla_{\mathbf{v}} \mathbf{u}(\mathbf{v}, \mathbf{Z}) \\ &= \nabla_{\mathbf{v}} E_{\text{es},\lambda}(\mathbf{v}, \mathbf{Z}, \tilde{\mathbf{u}}_\lambda(\mathbf{v}, \mathbf{Z})), \end{aligned} \quad (5.38)$$

$$\begin{aligned} \nabla_{\mathbf{Z}} E_{\text{res},\lambda}(\mathbf{v}, \mathbf{Z}) &= \nabla_{\mathbf{Z}} E_{\text{es},\lambda}(\mathbf{v}, \mathbf{Z}, \tilde{\mathbf{u}}_\lambda(\mathbf{v}, \mathbf{Z})) + \nabla_{\mathbf{u}} E_{\text{es},\lambda}(\mathbf{v}, \mathbf{Z}, \tilde{\mathbf{u}}_\lambda(\mathbf{v}, \mathbf{Z})) \nabla_{\mathbf{Z}} \mathbf{u}(\mathbf{v}, \mathbf{Z}) \\ &= \nabla_{\mathbf{Z}} E_{\text{es},\lambda}(\mathbf{v}, \mathbf{Z}, \tilde{\mathbf{u}}_\lambda(\mathbf{v}, \mathbf{Z})). \end{aligned} \quad (5.39)$$

Note that $\nabla_{\mathbf{u}} E_{\text{es},\lambda}(\mathbf{v}, \mathbf{Z}, \tilde{\mathbf{u}}_\lambda(\mathbf{v}, \mathbf{Z})) = 0$ because of equation 5.20, thus,

$$\begin{aligned} \nabla_{\mathbf{v}} E_{\text{res},\lambda}(\mathbf{v}, \mathbf{Z}) &= \nabla_{\mathbf{v}} E_{\text{es},\lambda}(\mathbf{v}, \mathbf{Z}, \Delta \tilde{\mathbf{b}}_\lambda(\mathbf{v}, \mathbf{Z})) \\ &= \lambda^2 \left(\mathbf{C}_A \frac{\partial \mathbf{A} \tilde{\mathbf{u}}_\lambda}{\partial \mathbf{v}} \right)^T \mathbf{C}_A \Delta \tilde{\mathbf{b}}_\lambda, \end{aligned} \quad (5.40)$$

$$\begin{aligned} \nabla_{\mathbf{Z}} E_{\text{res},\lambda}(\mathbf{v}, \mathbf{Z}) &= \nabla_{\mathbf{Z}} E_{\text{es},\lambda}(\mathbf{v}, \mathbf{Z}, \Delta \tilde{\mathbf{b}}_\lambda(\mathbf{v}, \mathbf{Z})) \\ &= \lambda^2 \left(\mathbf{C}_A \frac{\partial \mathbf{A} \tilde{\mathbf{u}}_\lambda}{\partial \mathbf{Z}} \right)^T \mathbf{C}_A \Delta \tilde{\mathbf{b}}_\lambda. \end{aligned} \quad (5.41)$$

Similar to Appendix C, one can compute the updates of v and Z , respectively,

$$\frac{\partial E_{\text{res},\lambda}}{\partial v(x)} = \lambda^2 C_A(x, x_s)^2 \int_0^T \frac{1}{v(x)^2 Z(x)} c_3(x, t) - \frac{1}{Z(x)} c_4(x, t) dt, \quad (5.42)$$

$$\frac{\partial E_{\text{res},\lambda}}{\partial Z(x)} = \lambda^2 C_A(x, x_s)^2 \int_0^T \frac{1}{v(x) Z(x)^2} c_3(x, t) + \frac{v(x)}{Z(x)^2} c_4(x, t) dt, \quad (5.43)$$

and

$$c_3(x, t) = \frac{\partial^2 \tilde{u}_\lambda(x, t)}{\partial t^2} \Delta \tilde{b}_\lambda(x, t), \quad (5.44)$$

$$c_4(x, t) = \nabla \tilde{u}_\lambda(x, t) \cdot \nabla \left(\Delta \tilde{b}_\lambda(x, t) \right), \quad (5.45)$$

where $\tilde{u}_\lambda(x, t)$ and $\Delta \tilde{b}_\lambda(x, t)$ is the continuous form of $\tilde{\mathbf{u}}_\lambda$, $\Delta \tilde{\mathbf{b}}_\lambda$, respectively.

The boundary condition of $\tilde{u}_\lambda$ is given by,

$$\begin{aligned} \tilde{u}_\lambda(x, 0) &= 0, x \in \Omega, \\ \frac{\partial \tilde{u}_\lambda(x, 0)}{\partial t} &= 0, x \in \Omega, \\ \tilde{u}_\lambda(x, t) &= 0, x \in \partial\Omega \text{ and } 0 \leq t \leq T, \\ \nabla \tilde{u}_\lambda(x, t) &= 0, x \in \partial\Omega \text{ and } 0 \leq t \leq T, \end{aligned} \quad (5.46)$$

and the boundary condition of $\Delta \tilde{b}_\lambda$ is given by,

$$\begin{aligned} \Delta \tilde{b}_\lambda(x, T) &= 0, x \in \Omega, \\ \frac{\partial \Delta \tilde{b}_\lambda(x, T)}{\partial t} &= 0, x \in \Omega, \\ \Delta \tilde{b}_\lambda(x, t) &= 0, x \in \partial\Omega \text{ and } 0 \leq t \leq T, \\ \nabla \left(\Delta \tilde{b}_\lambda(x, t) \right) &= 0, x \in \partial\Omega \text{ and } 0 \leq t \leq T. \end{aligned} \quad (5.47)$$

A sparse approximation of Hessian (Aravkin & van Leeuwen, 2012; van Leeuwen & Herrmann, 2016) for velocity can be computed by,

$$\begin{aligned} \mathbf{H}_{\mathbf{v}} &= \nabla_{\mathbf{v}}^2 E_{\text{res}}(\mathbf{v}, \mathbf{Z}) = \nabla_{\mathbf{v}}^2 E_{\text{es}}(\mathbf{v}, \mathbf{Z}, \tilde{\mathbf{u}}_\lambda((\mathbf{v}, \mathbf{Z}))) + \nabla_{\mathbf{v}, \mathbf{u}}^2 E_{\text{es}}(\mathbf{v}, \mathbf{Z}, \tilde{\mathbf{u}}_\lambda((\mathbf{v}, \mathbf{Z}))) \nabla_{\mathbf{v}} \tilde{u}_\lambda(\mathbf{v}, \mathbf{Z}) \\ &\approx \nabla_{\mathbf{v}}^2 E_{\text{es}}(\mathbf{v}, \mathbf{Z}, \tilde{\mathbf{u}}_\lambda((\mathbf{v}, \mathbf{Z}))) \\ &= \lambda^2 \frac{\partial \mathbf{G}_{\mathbf{v}}^T}{\partial \mathbf{m}} \mathbf{C}_{\mathbf{A}}^T \mathbf{C}_{\mathbf{A}} (\mathbf{A} \mathbf{u} - \mathbf{b}_{\text{obs}}) + \lambda^2 \mathbf{G}_{\mathbf{v}}^T \mathbf{C}_{\mathbf{A}}^T \mathbf{C}_{\mathbf{A}} \mathbf{G}_{\mathbf{v}} \\ &\approx \lambda^2 \left(\frac{\partial \mathbf{A} \tilde{\mathbf{u}}_\lambda}{\partial \mathbf{v}} \right)^T \mathbf{C}_{\mathbf{A}}^T \mathbf{C}_{\mathbf{A}} \left(\frac{\partial \mathbf{A} \tilde{\mathbf{u}}_\lambda}{\partial \mathbf{v}} \right). \end{aligned} \quad (5.48)$$

Similarly, the sparse approximation of Hessian for impedance can be computed by,

$$\mathbf{H}_{\mathbf{Z}} = \nabla_{\mathbf{Z}}^2 E_{\text{res}}(\mathbf{v}, \mathbf{Z}) \approx \lambda^2 \left(\frac{\partial \mathbf{A} \tilde{\mathbf{u}}_\lambda}{\partial \mathbf{Z}} \right)^T \mathbf{C}_{\mathbf{A}}^T \mathbf{C}_{\mathbf{A}} \left(\frac{\partial \mathbf{A} \tilde{\mathbf{u}}_\lambda}{\partial \mathbf{Z}} \right). \quad (5.49)$$

However, as both $\frac{\partial \mathbf{A}\tilde{\mathbf{u}}_\lambda}{\partial \mathbf{v}}$ and $\frac{\partial \mathbf{A}\tilde{\mathbf{u}}_\lambda}{\partial \mathbf{Z}}$ contains non-diagonal elements, evaluations of equations 5.48 and 5.49 are expensive.

In practice, the diagonal of the Hessian is approximated by the source and receiver wavefield illumination. Du, Fletcher, Mobley, and Yarman (2012) and Plessix and Mulder (2004) used similar approaches. The diagonal entry of the Hessians $\mathbf{H}_\mathbf{v}$ and $\mathbf{H}_\mathbf{Z}$ is given by

$$(\mathbf{H}_\mathbf{v})_{i,i} = (\mathbf{H}_\mathbf{Z})_{i,i} = \left(\int_0^T \tilde{u}_\lambda(\mathbf{x}_i, t)^2 \tilde{\Delta} b_\lambda(\mathbf{x}_i, t)^2 dt \right)^{\alpha_{\text{illu}}} . \quad (5.50)$$

where $i = 1, 2, 3, \dots, N_x$; $0 \leq \alpha_{\text{illu}} \leq 1$ is chosen optimally to minimize the data misfit. The magnitude of the Hessian approximated by equation 5.50 has a scaling effect compared to the original Hessian. We ignore the scaling factor as this will be adjusted in the line search. Through trial-and-error experiments, we find that $\alpha_{\text{illu}} = 0.5$ is suitable for our problem.

5.5.3 Algorithm

Cycle skipping often results in the overestimation or underestimation of parameters, leading to instability during the inversion process. It is essential to apply a bound constraint to help stabilize the inversion. We introduce a barrier function,

$$\begin{aligned} E_{\text{regres},\lambda}(\mathbf{v}, \mathbf{Z}) = & E_{\text{res},\lambda}(\mathbf{v}, \mathbf{Z}) \\ & + \sum (-\alpha_{\text{bv}} \ln(\mathbf{v} - v_l) - \alpha_{\text{bv}} \ln(v_u - \mathbf{v}) \\ & - \alpha_{\text{bZ}} \ln(\mathbf{Z} - Z_l) - \alpha_{\text{bZ}} \ln(Z_u - \mathbf{Z})) . \end{aligned} \quad (5.51)$$

In this expression, the parameters are defined as follows: α_{bv} and α_{bZ} represent the bound constraint regularization sizes for v and Z , respectively. The terms v_l and v_u denote the lower and upper bounds for v , while Z_l and Z_u denote the lower and upper bounds for Z , respectively. Thus, the corresponding gradient for model parameters can be computed by

$$\nabla_{\mathbf{v}} E_{\text{regres},\lambda} = \nabla_{\mathbf{v}} E_{\text{res},\lambda} - \alpha_{\text{bv}}(\mathbf{v} - v_l)^{-1} - \alpha_{\text{bv}}(v_u - \mathbf{v})^{-1} , \quad (5.52)$$

$$\nabla_{\mathbf{Z}} E_{\text{regres},\lambda} = \nabla_{\mathbf{Z}} E_{\text{res},\lambda} - \alpha_{\text{bZ}}(\mathbf{Z} - Z_l)^{-1} - \alpha_{\text{bZ}}(Z_u - \mathbf{Z})^{-1} . \quad (5.53)$$

Our Hessian of $E_{\text{regres},\lambda}$ ignores the regularization,

$$\nabla_{\mathbf{v}}^2 E_{\text{regres},\lambda} = \nabla_{\mathbf{v}}^2 E_{\text{res},\lambda} \approx \mathbf{H}_\mathbf{v} , \quad (5.54)$$

$$\nabla_{\mathbf{Z}}^2 E_{\text{regres},\lambda} = \nabla_{\mathbf{Z}}^2 E_{\text{res},\lambda} \approx \mathbf{H}_\mathbf{Z} . \quad (5.55)$$

Because the magnitude of impedance Z is much bigger than the magnitude of the velocity v , it is necessary to parameterize impedance Z such that both v and Z fall into the similar magnitude. In our case, we parameterize Z in the unit of $[2000 \text{ kg m}^{-2} \text{ s}^{-1}]$ during the inversion. The output of Z is rescaled back to the SI unit in the result.

We refer to the problem in equation 5.51 as extended-source joint migration inversion (ES-JMI), which leads to Algorithm 5.1. When the annihilator is chosen as $\mathbf{C}_A = \mathbb{I}$, the method reduces to joint migration inversion by wavefield reconstruction inversion (WR-JMI). Additionally, if $\lambda = \infty$, ES-JMI simplifies to the reduced-space method for JMI, which we refer to as JMI in this work. Those three methods will be benchmarked in the numerical tests. Using conventional FWI followed by estimating only impedance in JMI is equivalent to using FWI and LS-RTM, we refer this to conventional FWI in our numerical tests.

Algorithm 5.1: ES-JMI algorithm with possible bound constraint.

$n_{\max} :=$ maximum iterations.

$\varepsilon_{\text{esjmi}} :=$ stopping criterion.

$\mathbf{v} :=$ initial guess for the velocity.

$\mathbf{Z} :=$ initial guess for the impedance.

for $i = 1 : n_{\max}$ **do**

Compute $\Delta \mathbf{d}$ by equation 5.24.

Compute $\Delta \mathbf{d}_b$ by equation 5.33.

Estimate α_{hess} by equation 5.36, compute $\mathbf{F}$ in equation 5.34 and $\mathbf{H}_R^{-1}$ in equation 5.31.

if $i == 1$ **then**

Estimate λ using equation 5.37. This λ will be fixed when $i > 1$.

end

Compute $\Delta \tilde{\mathbf{b}}_\lambda$ by equation 5.30.

Compute $E_{\text{regres},\lambda}(\mathbf{v}, \mathbf{Z})$ by equation 5.51.

if $E_{\text{regres},\lambda} \leq \varepsilon_{\text{esjmi}}$ **then**

Break

end

Update model parameters by

$$\begin{bmatrix} \mathbf{v} \\ \mathbf{Z} \end{bmatrix} = \begin{bmatrix} \mathbf{v} \\ \mathbf{Z} \end{bmatrix} - \alpha_{\text{step}} \begin{bmatrix} \mathbf{H}_v & 0 \\ 0 & \mathbf{H}_Z \end{bmatrix}^{-1} \begin{bmatrix} \nabla_v E_{\text{regres},\lambda} \\ \nabla_Z E_{\text{regres},\lambda} \end{bmatrix} \quad (5.56)$$

where α_{step} is the step length found by Wolfe condition.

end

Return $\mathbf{v}$ and $\mathbf{Z}$.

5.6 Connections to band-pass filter

In this section, we analyze the nature of the band-pass filter of the gradient for v and Z in equations 5.42 and 5.43. It is known that if an acoustic wave equation is parameterized in (v, Z) , then in the inversion process, the v -update is dominated by low-wavenumber components, while high-wavenumber components mainly contribute the Z -update. With boundary conditions in equations 5.46 and 5.47, we

know that

$$\int_0^T \frac{\partial^2 \tilde{u}_\lambda(x, t)}{\partial t^2} \Delta \tilde{b}_\lambda(x, t) dt = \int_0^T \tilde{u}_\lambda(x, t) \frac{\partial^2 \Delta \tilde{b}_\lambda(x, t)}{\partial t^2} dt . \quad (5.57)$$

For the velocity gradient, we start from equation 5.42 and eliminate the time derivatives using the wave equation defined by equation 5.1,

$$\begin{aligned} \frac{\partial E_{\text{res}}}{\partial v(x)} &= C_A(x, x_s)^2 \frac{1}{Z(x)} \int_0^T \left(\frac{1}{2} \frac{1}{v(x)^2} \frac{\partial^2 \tilde{u}_\lambda(x, t)}{\partial t^2} \Delta \tilde{b}_\lambda(x, t) \right. \\ &\quad + \frac{1}{2} \frac{1}{v(x)^2} \tilde{u}_\lambda(x, t) \frac{\partial^2 \Delta \tilde{b}_\lambda(x, t)}{\partial t^2} \\ &\quad \left. - \nabla \tilde{u}_\lambda(x, t) \cdot \nabla \tilde{b}_\lambda(x, t) \right) dt \\ &= C_A(x, x_s)^2 \frac{1}{2Z(x)} \int_0^T \left(\frac{Z(x)}{v(x)} \nabla \frac{v(x)}{Z(x)} \cdot \nabla \tilde{u}_\lambda(x, t) \Delta \tilde{b}_\lambda(x, t) \right. \\ &\quad + \nabla^2 \tilde{u}_\lambda(x, t) \Delta \tilde{b}_\lambda(x, t) \\ &\quad + \frac{Z(x)}{v(x)} \nabla \frac{Z(x)}{v(x)} \cdot \nabla \Delta \tilde{b}_\lambda(x, t) \tilde{u}_\lambda(x, t) \\ &\quad + \nabla^2 \Delta \tilde{b}_\lambda(x, t) \tilde{u}_\lambda(x, t) \\ &\quad \left. - 2 \nabla \tilde{u}_\lambda(x, t) \cdot \Delta \tilde{b}_\lambda(x, t) \right) dt . \end{aligned} \quad (5.58)$$

Now, we consider plane wave solutions with a single angular frequency ω to $\tilde{u}_\lambda(x, t)$ and $\tilde{b}_\lambda(x, t)$ (Shearer, 2009),

$$\begin{aligned} \tilde{u}_\lambda(x, t) &= A_u(\omega) \exp(-i\omega t + ik(x) \cdot x) , \\ \Delta \tilde{b}_\lambda(x, t) &= A_b(\omega) \exp(-i\omega t + ik_b(x) \cdot x) . \end{aligned} \quad (5.59)$$

where $i = \sqrt{-1}$.

Inserting equation 5.59 into equation 5.58 yields

$$\begin{aligned} \frac{\partial E_{\text{res}}}{\partial v} &= C_A^2 \frac{1}{2Z} \int_0^T \left(\frac{Z}{v} \nabla \frac{v}{Z} \cdot (ik) + (ik)^2 + \frac{Z}{v} \nabla \frac{v}{Z} \cdot (ik_b) + (ik_b)^2 - 2i^2 k \cdot k_b \right) \\ &\quad A_u(\omega) A_b(\omega) \exp(-2i\omega t + i(k + k_b) \cdot x) dt . \end{aligned} \quad (5.60)$$

Equation 5.60 can be identified as a wave, and the real part has a physical meaning,

$$\begin{aligned} \text{Re}\left(\frac{\partial E_{\text{res}}}{\partial v}\right) &= C_A^2 \frac{1}{v} \int_0^T \left(\left(\frac{1}{2} \nabla \frac{v}{Z} \cdot (k + k_b) \right) A_u(\omega) A_b(\omega) \sin(-2\omega t + (k + k_b) \cdot x) \right. \\ &\quad \left. + \frac{v}{2Z} (k - k_b)^2 A_u(\omega) A_b(\omega) \cos(-2\omega t + (k + k_b) \cdot x) \right) dt \end{aligned} \quad (5.61)$$

Table 5.1: Wave properties of the second term in the gradient computation.

Gradient	Amplitude scaling	Wave vector	Wavelength
v	$(k - k_b)^2$	$k + k_b$	$2\pi\sqrt{(k + k_b)^{-2}}$
Z	$(k + k_b)^2$	$k + k_b$	$2\pi\sqrt{(k + k_b)^{-2}}$

The first term of equation 5.61 corresponds to the contribution due to the density change. The second term is a wave, and we name it pseudo-wave for the second term. Its amplitude scales to $(k - k_b)^2$, and this is dependent on the wave vector is $k + k_b$.

We can also derive a similar expression for Z ,

$$\begin{aligned} \text{Re}\left(\frac{\partial E_{\text{res}}}{\partial Z}\right) = & C_A^2 \frac{1}{Z} \int_0^T \left(\left(\frac{1}{2} \nabla \frac{v}{Z} \cdot (k + k_b) \right) A_u(\omega) A_b(\omega) \sin(-2\omega t + (k + k_b) \cdot x) \right. \\ & \left. + \frac{v}{2Z} (k + k_b)^2 A_u(\omega) A_b(\omega) \cos(-2\omega t + (k + k_b) \cdot x) \right) dt. \end{aligned} \quad (5.62)$$

$\text{Re}(\frac{\partial E_{\text{res}}}{\partial Z})$ has identical first term in the integral. And the second term is slightly different. The amplitude of the pseudo wave in the second term scales to $(k + k_b)^2$. Table 5.1 gives an overview of the wave properties of the pseudo wave.

We can assume $|k| = |k_b|$ as they are the wave vectors at the same location and in the same medium. The only difference between k and k_b is the direction. Let's define the angle between the two vectors as θ_k . An input wavelength of $4\pi^2 (k + k_b)^{-2} = 4\pi^2 (k^2 + 2k^2 \cos(\theta_k))^{-2}$ will be scaled by $(k - k_b)^2 = 2k^2 - 2k^2 \cos(\theta_k)$ for the amplitude of v , and will be scaled by $(k + k_b)^2 = 2k^2 + 2k^2 \cos(\theta_k)$ for the amplitude of Z . As $\cos(\theta_k)$ is monotonic in the range of $[0, \pi]$, for larger wavelengths, the scaling for v is greater, while for Z it is smaller. v thus tends to select larger wavelengths, and Z tends to select smaller wavelengths. An illustration is given in Figure 5.1.

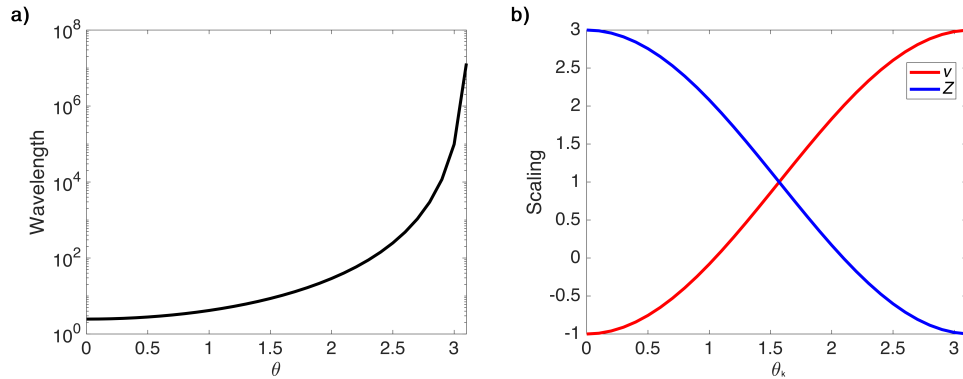


Figure 5.1: Band-pass filter for the second term in the gradient computation.

5.7 Numerical tests

In this section, we show two example models, Marmousi (Irons, 1988) and 2004 BP (Billette & Brandsberg-Dahl, 2005), to test our method against alternative methodologies. Wavefield simulations are performed using the accuracy $\mathcal{O}(\Delta t^2, \Delta x^6)$. For each experiment, 15 layers of Perfectly Matched Layers (PML) are assigned.

5.7.1 Marmousi

The ground truth of the Marmousi model is shown in Figure 5.2. A total of 38 sources and 124 receivers are evenly placed at $x_3 = 281.25$ m. The starting model is displayed in Figure 5.3. Two wavelets are used, as illustrated in Figure 5.4. Wavelet 1 is a Ricker wavelet with a peak frequency of 5 Hz, with components below 2.5 Hz filtered out. Wavelet 2 is also a Ricker wavelet with a peak frequency of 5 Hz but with components below 5 Hz filtered out. The low-frequency components are removed from the data to represent a more realistic scenario. The model is discretized into a grid of 401 by 140 cells in the x_1 - and x_3 - directions, respectively, with a grid spacing of 18.75 m. The time domain is discretized into 1600 samples with a time interval of 0.002 s.

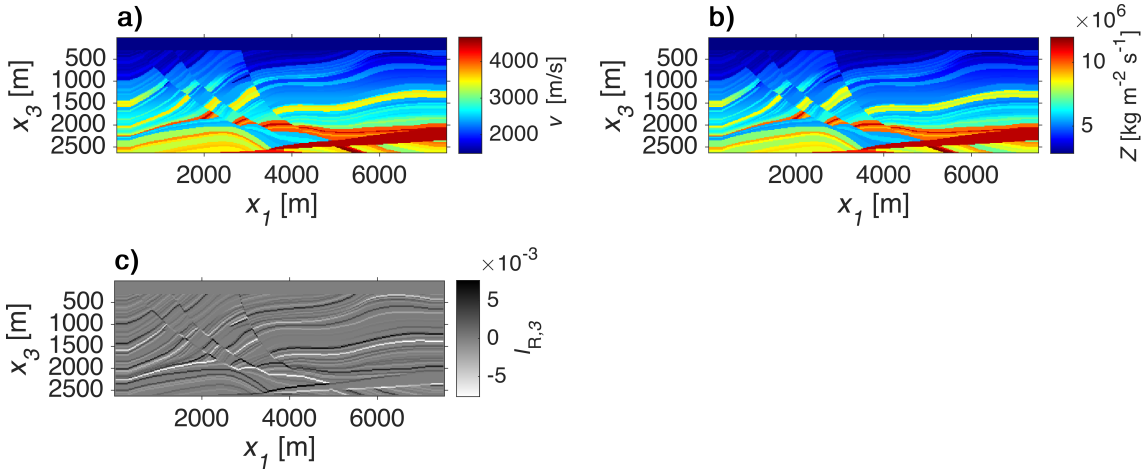


Figure 5.2: Velocity (a), impedance (b), and vertical reflectivity (c) ground truth for the Marmousi.

Four methods are tested using the Marmousi model. Method 1 runs 50 iterations of conventional FWI to estimate velocity at Step 1. Then, 10 iterations are performed to estimate Z by equation 5.15 at Step 2: this is similar to a least-square migration. This sequence is performed for each wavelet. We named this method conventional FWI. Method 2 solves the inversion using Algorithm 5.1, but with $\lambda = \infty$: this corresponds to the reduced space method in equation 5.15. We named this experiment JMI. For Method 3, we use Algorithm 5.1 with $\mathbf{C}_A = \mathbb{I}$, which corresponds to the WR-JMI. For method 4, Algorithm 5.1 is utilized with $\mathbf{C}_A$ defined in 5.17, i.e., it is the ES-JMI.

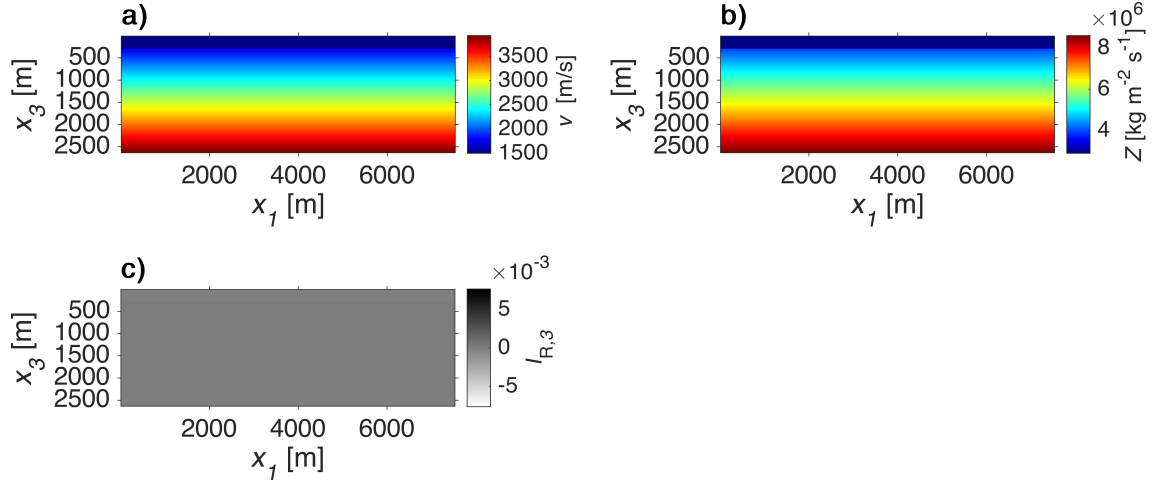


Figure 5.3: Initial iterate for the inversion of velocity (a), impedance (b) and vertical reflectivity (c) of the Marmousi.

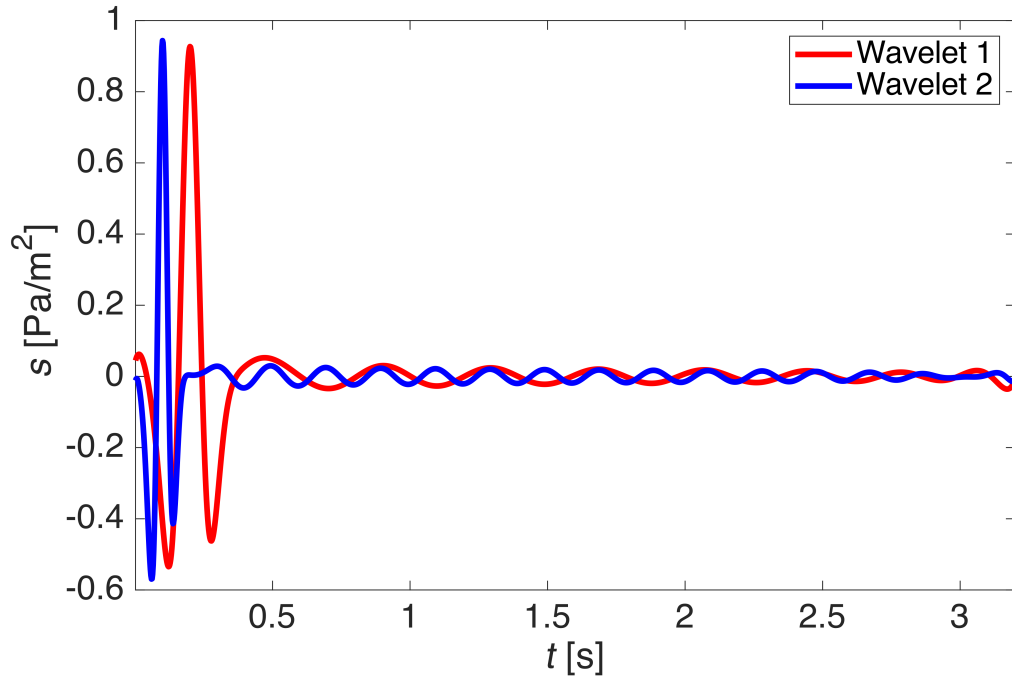


Figure 5.4: Wavelets utilized in the inversion of Marmousi.

For Methods 2-4, 3 steps are performed for each wavelet. Step 1 only updates v for 10 iterations. Step 2 updates both v and Z for 30 iterations. Step 3 only updates Z for 10 iterations. λ for Methods 3 and 4 is estimated using equation 5.37 at the first iteration for each wavelet. According to our experiment, no bounded constraint is necessary for the Marmousi model.

Figure 5.5 shows the update direction for Wavelet 1 at the first iteration. Compared to other methods, the update for v contains both low- and high-wavenumber components (Figure 5.5a). The velocity update for JMI, WR-JMI, and ES-JMI is smoother. The update direction for Z in Figure 5.5h resolves deeper features than Figure 5.5d,f.

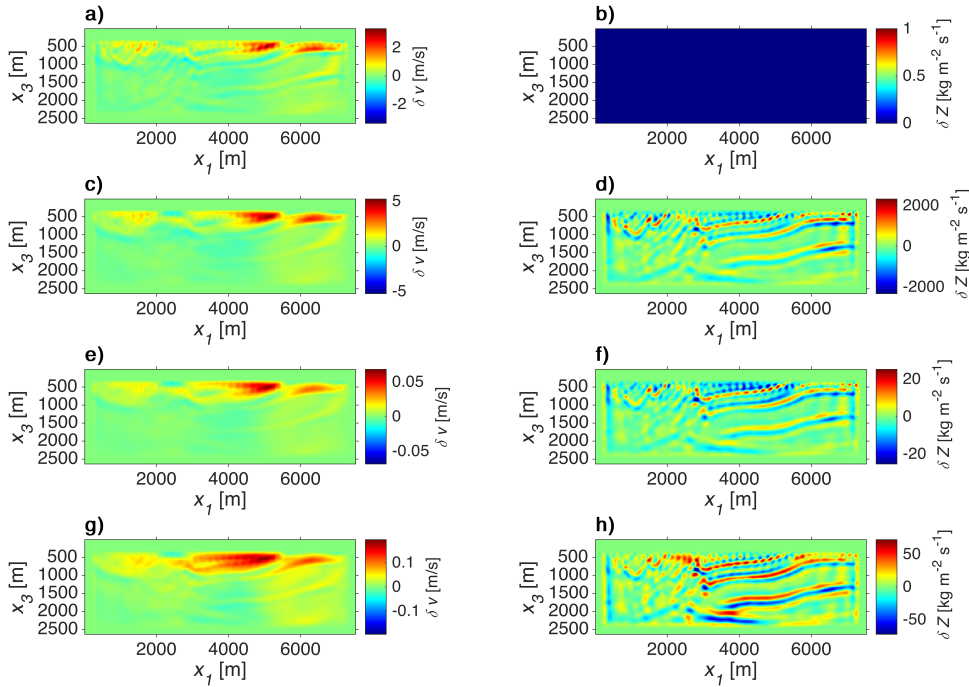


Figure 5.5: Update direction for the first iteration of Wavelet 1. The update direction for v is shown in the first column, while the update direction for Z is displayed in the second column. The four rows correspond to conventional FWI, JMI, WR-JMI, and ES-JMI, respectively. The conventional FWI updates only v in the velocity build-up so that the update direction for Z is missing.

Figure 5.6 is the last iterate for each method. Conventional FWI fails to recover a reliable velocity model, as it reduces the data misfit with iterations but ultimately results in a higher model misfit compared to the initial estimate. Significant artifacts in the deeper regions are observed (Figure 5.6a) due to cycle skipping. As the reflectivity is dependent on the velocity model, Figure 5.6c is not accurate. JMI improves on conventional FWI because of the velocity build-up (Figure 5.6d). However, it still contains visible artifacts in the velocity and reflectivity models. WR-JMI (Figure 5.6g) obtains a reliable velocity model but lacks deep velocity and reflectivity

information. ES-JMI is able to reconstruct accurate velocity and reflectivity models at all depths (Figure 5.6k,l).

Figure 5.7 shows the convergence history. The reduced space data misfit is computed by $\frac{1}{2} \|\mathbf{R}\mathbf{u} - \mathbf{d}_{\text{obs}}\|_2^2$ where $\mathbf{u}$ is the solution of the wave equation. For WR-JMI and ES-JMI, $\frac{1}{2} \|\mathbf{R}\mathbf{u}_\lambda - \mathbf{d}_{\text{obs}}\|_2^2$ where $\mathbf{u}_\lambda$ is the data augmented wavefield. The model misfit for v is computed by $\frac{1}{2} \|\mathbf{v}_i - \mathbf{v}_t\|_2^2$ where $\mathbf{v}_i$ is the velocity at iteration i and $\mathbf{v}_t$ is the true velocity. The physics misfit is computed by $\frac{\lambda^2}{2} \|\mathbf{A}\mathbf{u} - \mathbf{b}_{\text{obs}}\|_2^2 = \frac{1}{2\lambda^2} \|\Delta\tilde{\mathbf{b}}_\lambda\|_2^2$.

Although conventional FWI and JMI can reduce the data misfit (Figure 5.7a,c), their performance in reducing the model misfit is poor, indicating that they are trapped in a local minimum. In contrast, both WR-JMI and ES-JMI achieve a significantly lower model misfit for v (Figure 5.7a,c), with ES-JMI performing better. Additionally, the reconstructed data misfit and the physics misfit decrease with iterations for both WR-JMI and ES-JMI (Figure 5.7b,d).

Figure 5.8 displays the synthetic seismograms. For conventional FWI and JMI, the data residuals remain high, whereas ES-JMI shows the lowest data residual.

5.7.2 2004 BP

We also tested our algorithm on the more challenging 2004 BP model, which is characterized by high contrasts. Figure 5.9 shows the ground truth, while Figure 5.10 presents the starting model. 25 sources and 80 receivers are evenly placed at $x_3 = 937.5$ m. Figure 5.11 illustrates the wavelets used in our experiment. Wavelet 1 is a Ricker wavelet with a 2 Hz peak frequency, from which components below 1 Hz have been removed. Wavelet 2 is a Ricker wavelet with a 4 Hz peak frequency, with components below 2 Hz removed. The model is discretized into a grid of 271 by 135 cells in the x_1 - and x_3 -directions, respectively, with a grid spacing of 62.5 m. The time domain is discretized into 3000 samples with a time interval of 0.005 s.

Figure 5.12 displays the update direction for the first iterate. All of them have similar update directions for v . For Z , ES-JMI contains smoother components at the top (Figure 5.12h).

For this model, a bounded constraint is necessary. We choose the following parameters to ensure stable inversion: $\alpha_{bv} = 1$, $v_l = 1000$ [m/s], $v_u = 7000$ [m/s], $\alpha_{bZ} = 1$, $Z_l = 1000$ [kg m⁻² s⁻¹], and $Z_u = 8000$ [kg m⁻² s⁻¹]. Despite these constraints, conventional FWI, JMI, and WR-JMI fail to produce reliable results. Therefore, we only present the results from ES-JMI alongside those from WR-JMI for comparison. For both WR-JMI and ES-JMI, three steps are performed for each wavelet: Step 1 updates v for a maximum of 10 iterations, Step 2 updates both v and Z for a maximum of 50 iterations, and Step 3 updates only Z for a maximum of 10 iterations. Some iterations may stagnate, leading to a transition to the next step.

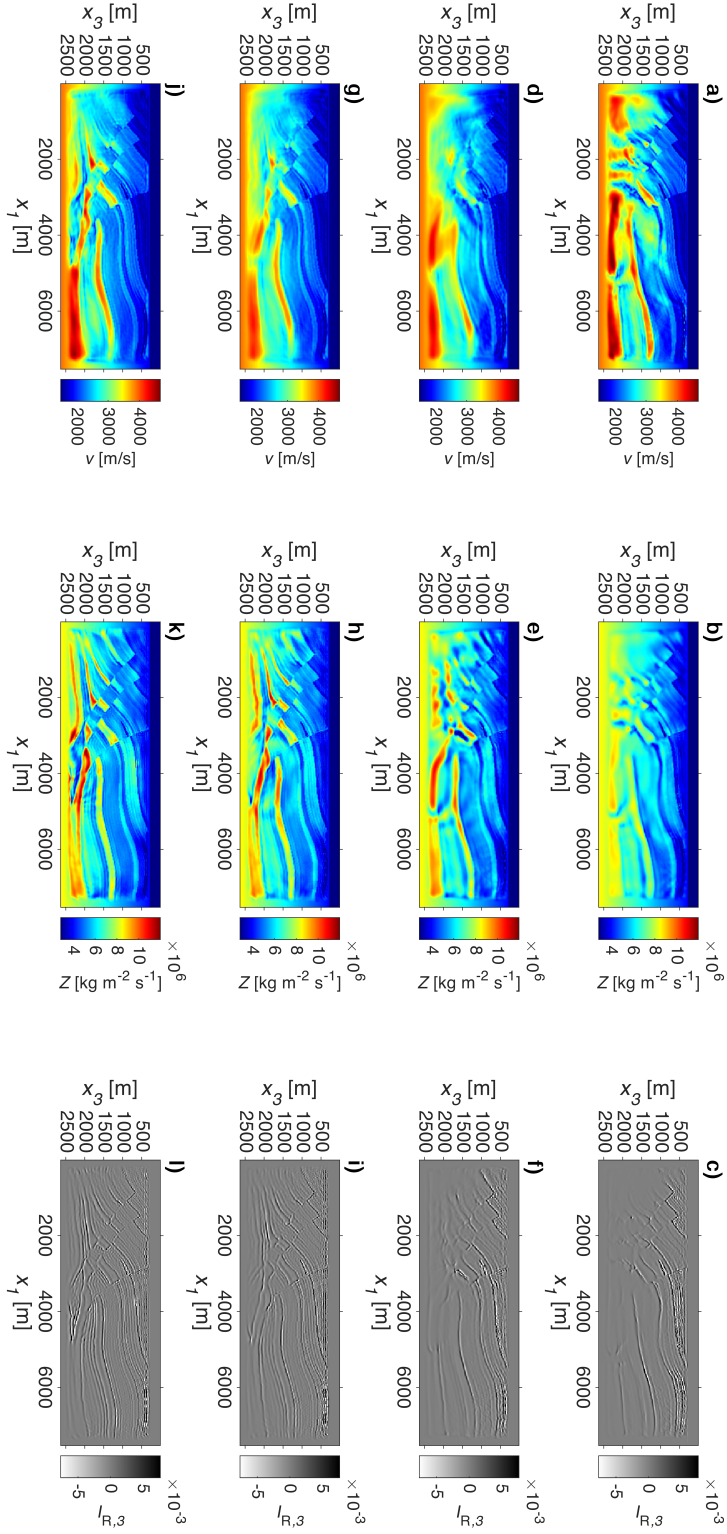


Figure 5.6: Inversion results for Marmousi. The first column represents v , the second column represents Z , and the third column shows the vertical reflectivity $I_{R,3}$. The four rows correspond to conventional FWI, JMI, WR-JMI, and ES-JMI, respectively. Z in conventional FWI is derived at Step 2 of the conventional FWI process.

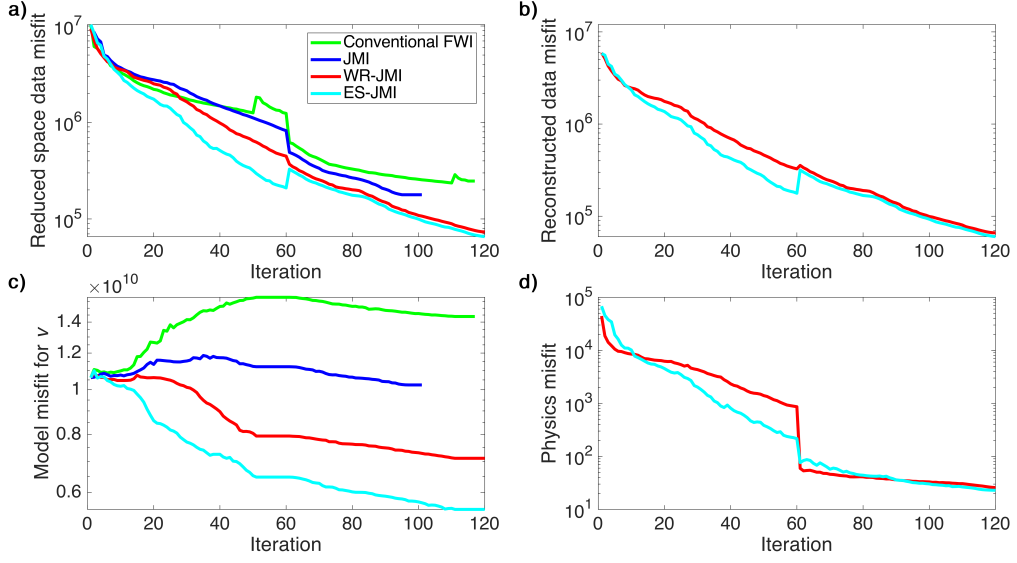


Figure 5.7: Convergence history of the different methods for the Marmousi inversions.

Figure 5.13 shows the inversion results for the 2004 BP model. WR-JMI converges to a local minimum during the inversion, so it is unable to reduce the data misfit appropriately. In contrast, ES-JMI provides a good estimate for both velocity and vertical reflectivity. However, the impedance inversion for ES-JMI remains suboptimal, possibly due to parameter crosstalk between v and Z . The superiority of ES-JMI is evident in the comparisons between convergence histories (Figure 5.14) and seismograms (Figure 5.15).

5.8 Discussions

While the initial results are promising, ES-JMI can still be improved. Firstly, the data augmented wavefield solution in equation 5.21 is not solved exactly but approximated. As a consequence, equations 5.38 and 5.39 do not strictly hold. Frequency-domain data augmented wavefields can be solved exactly and efficiently, but the time-domain problem remains computationally challenging and is usually solved by approximations. Though approximations for the data augmented wavefield are typically used to solve the time domain problem, the error analysis is worth further investigation.

Secondly, our study estimates the size of the physics penalty, λ , during the data deblurring step at the first iteration and then fixes this value for subsequent iterations. In practice, a continuation strategy for λ would be more suitable for ensuring that the data misfit decreases further; however, this approach typically requires additional iterations to achieve satisfactory results. Therefore, developing an effective

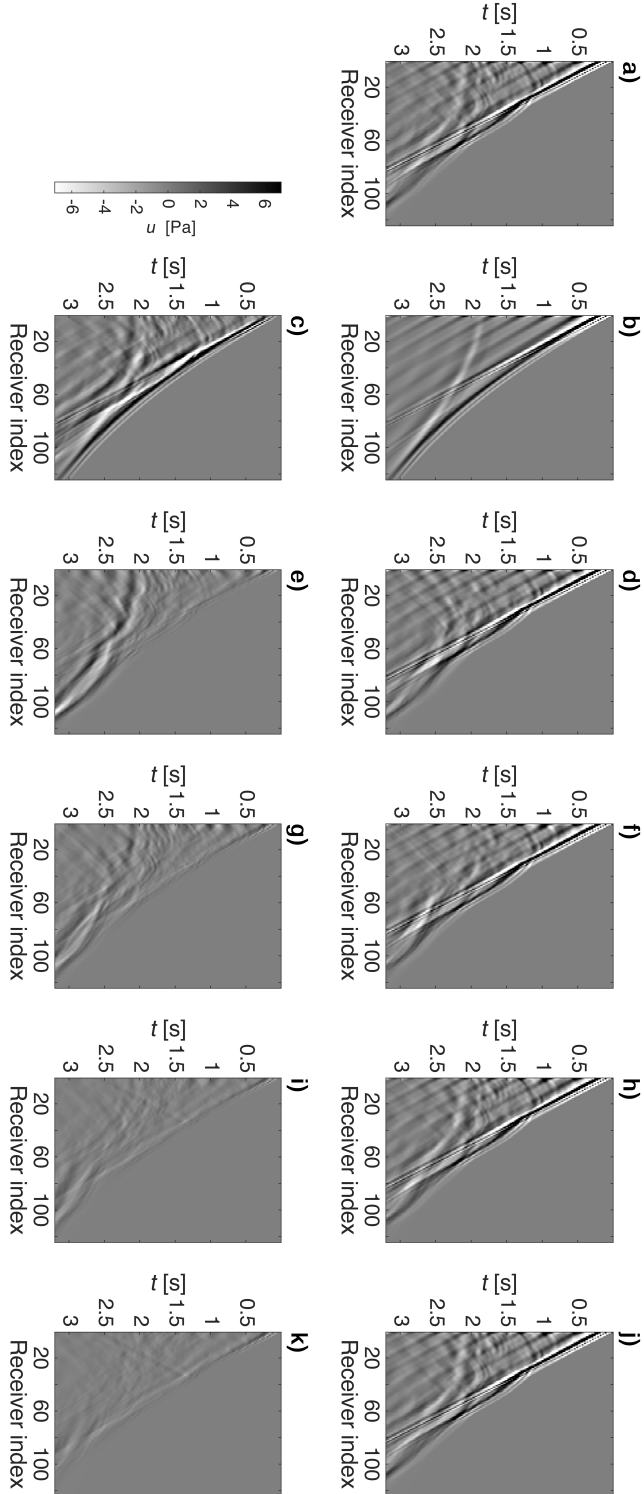


Figure 5.8: Seismograms for the first source from the left. The seismograms are stacked for Wavelets 1 and 2. (a) Observed data. (b) Data from initial iterate. (d)(f)(h)(j) is data by conventional FWI, JMI, WR-JMI, and ES-JMI, respectively. The second row is the subtraction of the first row by the observed data in (a).

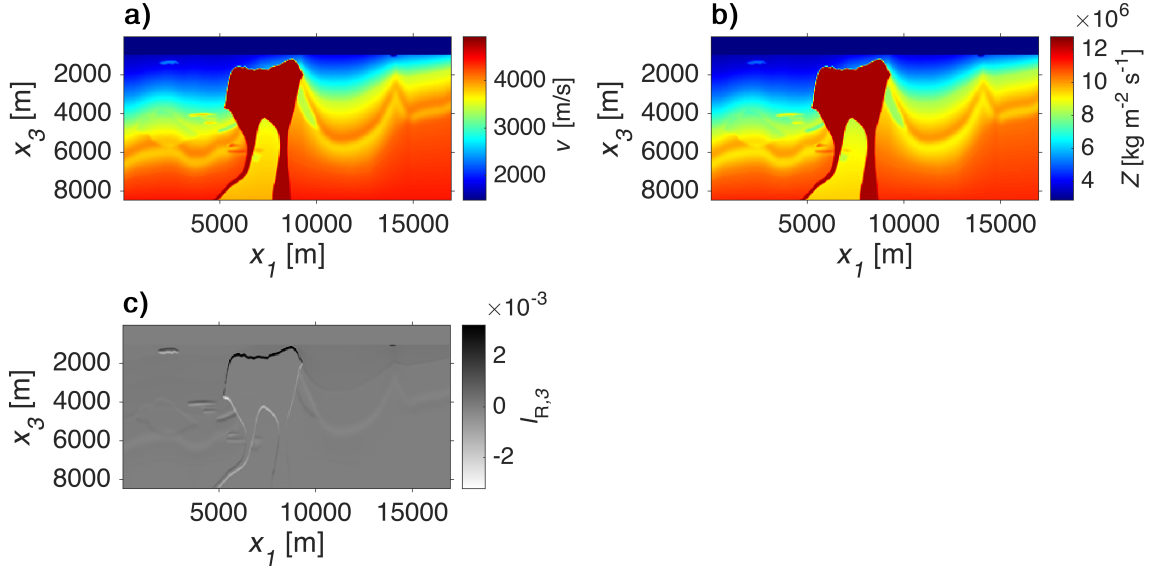


Figure 5.9: Velocity (a), impedance (b) and vertical reflectivity (c) ground truth for 2004 BP.

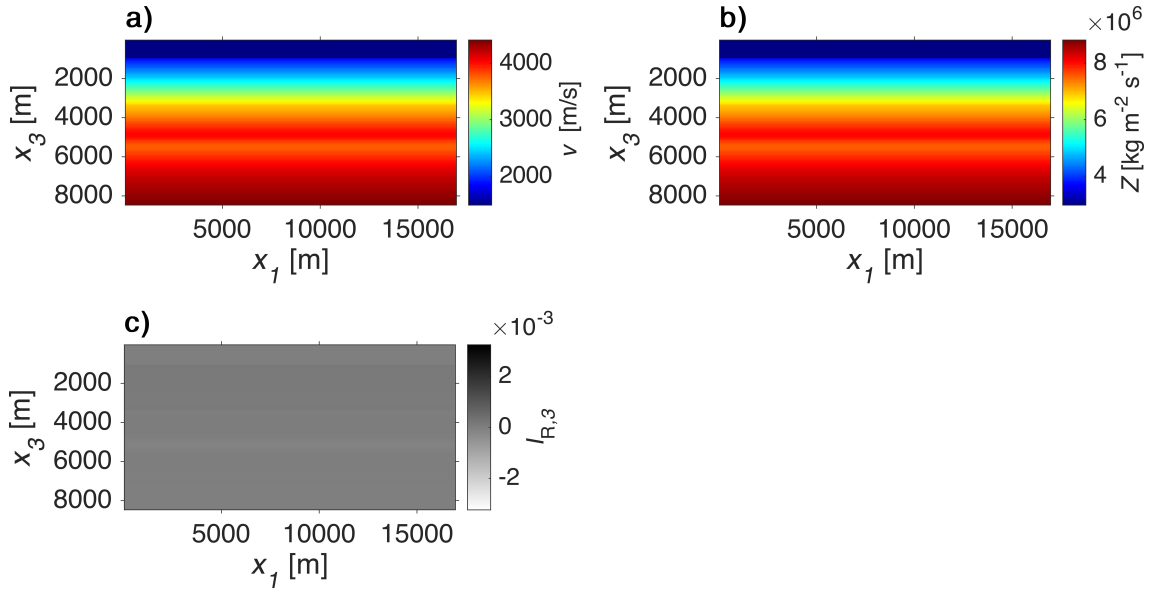


Figure 5.10: Initial iterate for the inversion of velocity (a), impedance (b), and vertical reflectivity (c) for 2004 BP.

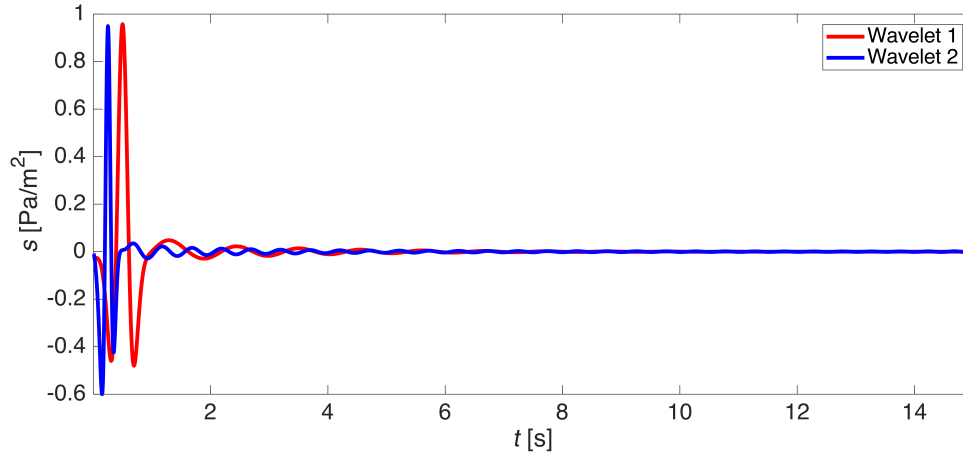


Figure 5.11: Wavelets utilized in the inversion of 2004 BP.

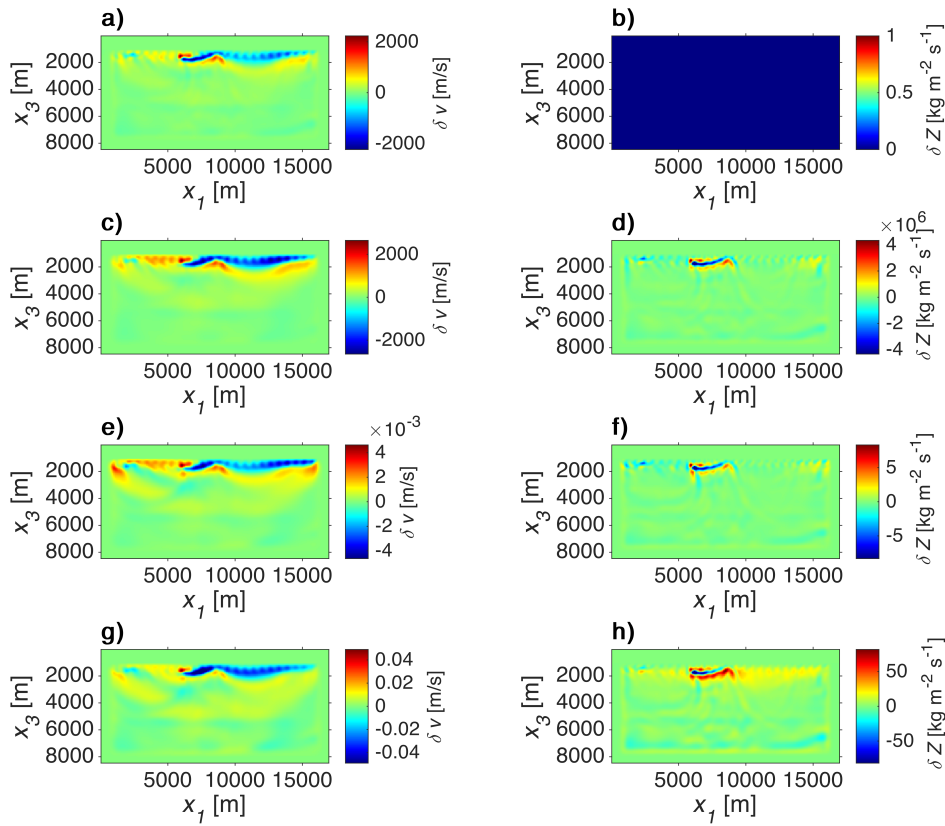


Figure 5.12: Same as Figure 5.5 but for 2004 BP model.

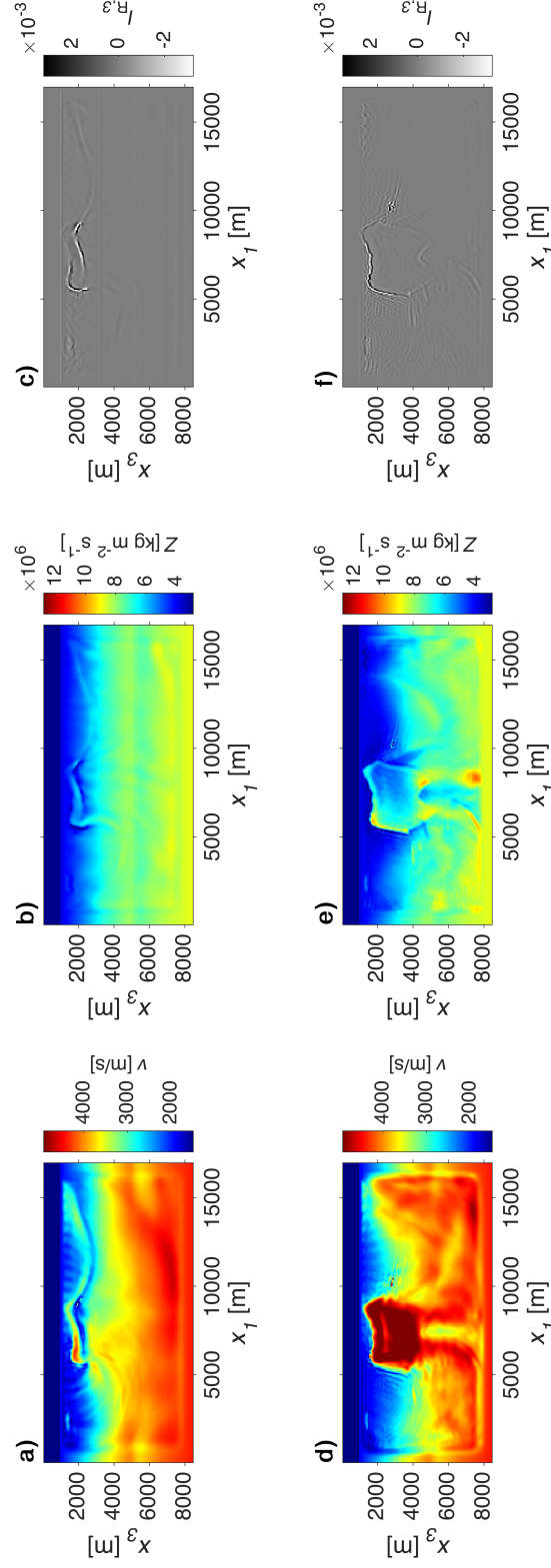


Figure 5.13: Same column arrangements as in Figure 5.6 but for 2004 BP model. First row is for WR-JMI and the second is for ES-JMI.

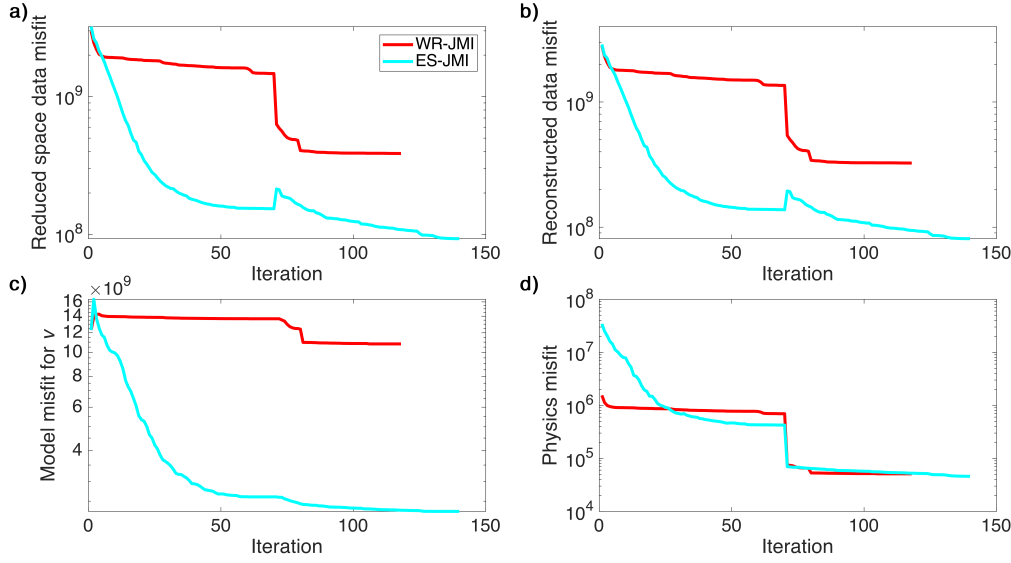


Figure 5.14: Convergence histories of the different methods for the 2004 BP inversion.

strategy for designing the continuation of λ is essential.

5.9 Conclusions

This work proposes a new algorithm that uses reduced-space extended source waveform inversion to address the joint migration inversion problem, which we refer to as ES-JMI. The forward problem is formulated using a velocity-impedance parameterized wave equation. This PDE combines the conventional velocity and reflectivity estimates into a single equation, streamlining the process.

ES-JMI solves the weighted physics penalty formulation, where the physics is relaxed during the inversion process to reduce sensitivity to cycle skipping, which often occurs with poor initial models or in the absence of low-frequency data. In this approach, the data augmented wavefield is assumed to be directly dependent on the model parameters, eliminating the need to update the wavefield based on the previous iteration and thereby avoiding simultaneous storage of the wavefield.

However, solving the data augmented wavefield requires an exact solution to the data domain Hessian, which traditionally involves iterative solvers, with each iteration requiring two evaluations of the wave equation. To address this issue, we use scalar fitting to approximate the scaled identity of the data domain Hessian, allowing us to bypass this computational bottleneck. These strategies enable our algorithm to be efficiently implemented for large-scale problems.

We demonstrate that the velocity update primarily captures the low-wavenumber

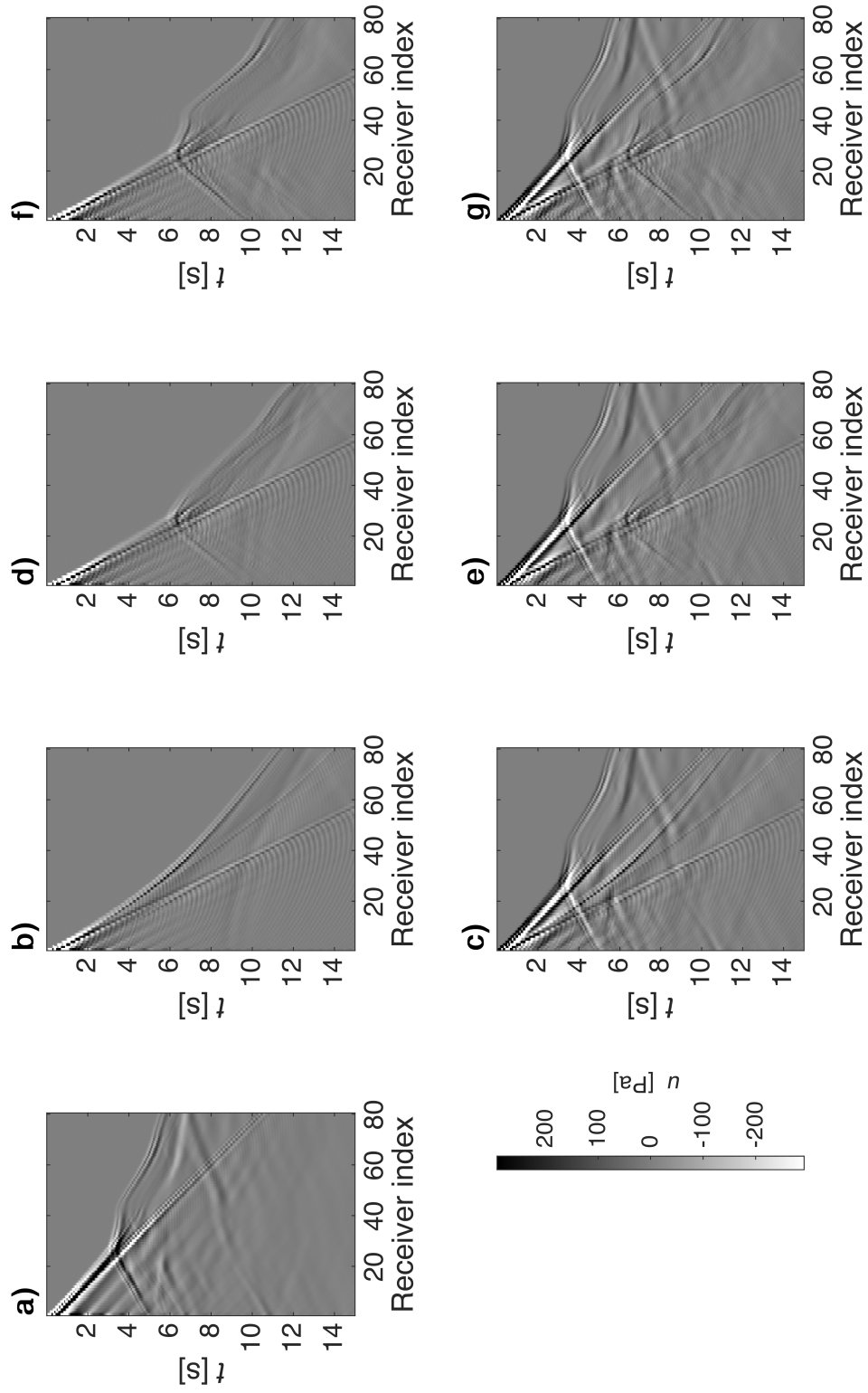


Figure 5.15: Seismograms for the first source from the left. The seismograms are stacked for Wavelets 1 and 2. (a) Observed data. (b) Data from initial iterate. (d)(f) are the data from the final iterate by conventional WR-JMI and ES-JMI. The second row is the subtraction of the first row by the observed data in (a).

components while the impedance update absorbs the high-wavenumber components when considering a plane-wave solution. This explains why the velocity iterates can be interpreted as the background velocity model, while the impedance is well-suited for computing reflectivity.

Our numerical results demonstrate that the proposed ES-JMI is less prone to local minimum issues, even when a poor initial iterate is provided, and the low-frequency components are absent from the data.

Chapter 6

Source independent extended source waveform inversion based on augmented Lagrangian

The work presented in this chapter is independent and has not been submitted for publication as a journal article. The author would like to acknowledge valuable discussions with Prof. Markus Bachmayr regarding the all-at-once approach and total variation regularization techniques. The author also extends gratitude to Prof. Luca De Siena for his valuable suggestions on improving the manuscript.

Abstract

Conventional Full Waveform Inversion (FWI) depends on low-frequency data and a well-chosen initial model to ensure global convergence. Typically, the source signature used in FWI must be predetermined from the data, and the resolution of conventional FWI is highly sensitive to the accuracy of this source estimation. In this work, we propose a source-independent Iteratively Refined Extended Source waveform Inversion with Source Estimation (SE-IR-ESI). The approach incorporates an additional step to estimate the source wavelet within the inversion process. To eliminate the need for laborious tuning to balance data misfit and physics misfit, we adopt the augmented Lagrangian (AL) framework and solve it using the Peaceman-Rachford Splitting Method (PRSM) with relaxation factors. Furthermore, we enhance SE-IR-ESI with Total Variation (TV) regularization and bounded constraints for the model update. TV regularization is implemented via Levenberg-Marquardt linearization combined with Bregman iteration, and the resulting system is solved using the primal-dual Newton's method. We validate our method using a subset of the Marmousi model with a starting frequency of 3 Hz and a subset of 2004 BP model with a starting frequency of 2,5 Hz. Results show that SE-IR-ESI, even without regularization, achieves reliable reconstruction. Incorporating TV regularization

and bounded constraints further reduces the model misfit. In contrast, conventional reduced-space FWI suffers from cycle skipping under the same conditions.

6.1 Introduction

A key objective in seismic exploration is to estimate the subsurface velocity field, which is critical for understanding the Earth’s structure and composition. Full Waveform Inversion (FWI) has become a widely adopted technique in geophysics to achieve this goal (Choi, Min, & Shin, 2008; Liu, Beller, Lei, Peter, & Tromp, 2000; Tarantola, 1984b). FWI seeks to minimize the mismatch between simulated seismic data—generated using a physics based numerical engine – and observed field data. This process is formulated as a partial differential equation (PDE) constrained optimization problem, requiring the solution of a series of forward and inverse problems governed by PDEs (Haber & Ascher, 2000; Tarantola & Valette, 1982).

FWI has undergone significant evolution over the years. Haber and Ascher (2000) described the use of the Karush-Kuhn-Tucker (KKT) conditions to address the PDE-constrained optimization problem, introducing the so-called “all-at-once” method. This approach solves the forward wavefields and their Lagrange multipliers for all seismic shots simultaneously. However, the method is computationally prohibitive and impractical for large-scale industry applications. To overcome this limitation, geophysicists have sought more efficient alternatives. One of the most widely adopted strategies is the “reduced-space” method (Kaderli, McChesney, & Minkoff, 2018; Liu et al., 2000; Pratt, 1999; Tarantola, 1984b; Tarantola & Valette, 1982). In this approach:

- The forward wavefield is computed by solving the forward PDE.
- The Lagrange multipliers (or adjoint wavefields) are obtained by back-propagating the residuals from the receivers.
- The model parameters are updated using both the forward and adjoint wavefields.

While the reduced-space approach significantly reduces computational demands, it introduces a highly nonlinear dependency of the cost function on the model parameters. This stems from the oscillatory nature of the wave equation, making the method susceptible to “cycle skipping” – a phenomenon that leads to suboptimal convergence (Aravkin & van Leeuwen, 2012; van Leeuwen & Herrmann, 2013). To mitigate this issue, conventional FWI (i.e., the reduced-space method) typically requires both an accurate initial model and low-frequency data.

Another issue with conventional FWI is its reliance on accurate source estimation. Pratt (1999) demonstrated that the relationship between the source wavefield and

the source signature is linear, suggesting that the source signature can be directly estimated once the source wavefield is simulated. However, the source signature is dependent on the initial model. Lee and Kim (2003) and B. Zhou and Greenhalgh (2003) proposed normalizing each seismogram in a shot gather by the sum of all seismograms, effectively removing the dependency on the source signature in the normalized data. However, the normalized residuals do not satisfy the wave equation and therefore cannot be back-projected.

In this work, we aim to propose an algorithm that addresses both cycle skipping and source signature issues.

A significant number of studies have been conducted to mitigate cycle skipping in FWI. Broadly, these methods can be categorized into two approaches: (1) proposing improved metrics to measure data misfit and (2) enlarging the search space to make the cost function more linear and convex with respect to the model parameters. The first category includes methods such as adaptive waveform inversion (Warner & Guasch, 2016), dynamic time warping (Ma & Hale, 2013; M. Wang et al., 2016), and optimal transport techniques (Engquist et al., 2016; Ramos-Martínez et al., 2018). In this work, we adopt the second approach, enlarging the search space to make the inversion process more robust against cycle skipping. Initially, van Leeuwen and Herrmann (2013) introduced the penalty method to solve the inversion problem, calling this approach Wavefield Reconstruction Inversion (WRI). In their formulation, the problem is addressed using alternating optimization between the data augmented wavefield and the model parameters. At each iteration, the model update is represented as a quadratic function and solved using Newton’s method. However, this approach requires a continuation strategy to dynamically adjust the penalty size, ensuring that the physics constraint is increasingly enforced over iterations, which promotes global convergence. Later, van Leeuwen and Herrmann (2016) addressed the same problem as van Leeuwen and Herrmann (2013) but employed a reduced-space approach. This method yields a fully ranked Hessian with a natural sparse approximation. However, the reduced objective maintains a nonlinear relationship with the parameters. Building on these efforts, Aghamiry et al. (2019b) reformulated the WRI problem using the augmented Lagrangian method, leading to the development of Iteratively Refined WRI (IR-WRI). A notable advantage of IR-WRI over WRI is that the right-hand side of the WRI system is also updated during iterations. This refinement more effectively reduces both data misfit and physics errors. Furthermore, IR-WRI can be implemented with a fixed and moderate penalty size, as demonstrated by Aghamiry et al. (2019a).

The source independency can also be incorporated into WRI-type methods. Fang, Wang, and Herrmann (2018) proposed an algorithm for source signature estimation during WRI. Their approach solves the data augmented wavefield and the source signature simultaneously within a least-squares problem. This method does not increase the computational cost compared to conventional WRI. However, the problem is solved using the quasi-Newton (L-BFGS) method, which makes parameter estimation nonlinear. Another development is extended source waveform inversion (ESI). Fu and Symes (2017b) and G. Huang et al. (2018a) proposed using an an-

nihilator to measure the extent of nonphysicality. This approach shares a similar idea with WRI (van Leeuwen & Herrmann, 2013), with the primary difference being that the penalty term is weighted such that the source term can be omitted. In their work, the problem is solved using the variable projection method to improve computational efficiency. The L-BFGS method is employed for optimization, but this method still results in nonlinear parameter estimation. Additionally, it does not use a Newton update, leading to slower convergence.

In this work, we propose a new algorithm that incorporates Extended Source waveform Inversion (ESI) with the augmented Lagrangian method. We solve a modified ESI problem by combining the approaches of G. Huang et al. (2018a) and Fang et al. (2018), such that the inverse problem becomes independent of the source signature while also enabling the estimation of the source signature. The modified ESI problem is reformulated within an augmented Lagrangian framework, allowing the penalty size to remain fixed throughout the iterations. The augmented Lagrangian is solved using the Peaceman-Rachford splitting method with relaxation factors (Aghamiry et al., 2019b; He, Liu, Wang, & Yuan, 2014). We name our algorithm SE-IR-ESI. To further enhance immunity to cycle skipping, we incorporate Total Variation (TV) regularization and bounded constraints. This algorithm is named Iteratively Refined Extended Source waveform Inversion with Source Estimation and Total Variation (TV-SE-IR-ESI). Unlike the split-Bregman method described in Aghamiry et al. (2019a), we implement a Levenberg-Marquardt type TV regularization for the model parameter updates (Bachmayr & Burger, 2009), where Bregman iterations are applied. The model update system is then solved using the primal-dual Newton’s method (Chan et al., 1996). Unlike the split-Bregman method, where model updates and sparsity promotion are decoupled, our approach maintains a tight coupling between these two components. The bounded constraints are implemented using the projection method (Esser, Guasch, Leeuwen, Aravkin, & Herrmann, 2016). We tested our algorithm on the Marmousi model using data above 3 Hz and 2004 BP model with data above 2.5 Hz. Our results demonstrate that the proposed algorithm successfully reconstructs the model even when low-frequency data is absent and the initial model is coarse. The source signature can be estimated during the inversion. Additionally, incorporating moderate TV regularization effectively reduces the model misfit to a lower level for high contrast models such as 2004 BP.

6.2 Conventional FWI

The frequency domain acoustic wave equation can be cast as

$$\frac{(-i\omega)^2}{v(x)^2}u(x) - \nabla^2 u(x) = \delta(x - x_s)s(\omega), \quad (6.1)$$

where $i = \sqrt{-1}$; $x \in \Omega \subseteq \mathbb{R}^{N_{\text{dim}}}$ is the N_{dim} -dimensional spatial coordinate; u is the acoustic pressure; $\delta(\cdot)$ is Dirac delta; x_s is the source location; s is the source

wavelet. In this work, the wave equations with Perfectly Matched Layers (PML) are solved by the finite-difference method.

One can rewrite equation 6.1 into a discretized form,

$$\mathbf{A}(\mathbf{v})\mathbf{u} = \mathbf{b}_{\text{obs}} , \quad (6.2)$$

where $\mathbf{A}(\mathbf{v}) = \text{diag}((-i\omega)^2./\mathbf{v}.^2) - \nabla^2$; $\mathbf{u}$ and $\mathbf{b}_{\text{obs}}$ is the discretized version of $u(x)$ and $\delta(x - x_s)s(\omega)$, perspectively.

In full waveform inversion, we try to solve the following minimization problem,

$$\min_{\mathbf{v}, \mathbf{u}} \frac{1}{2} \|\mathbf{R}\mathbf{u} - \mathbf{d}_{\text{obs}}\|_2^2 \quad \text{s.t.} \quad \mathbf{A}(\mathbf{v})\mathbf{u} - \mathbf{b}_{\text{obs}} = \mathbf{0} , \quad (6.3)$$

According to Haber and Ascher (2000), the problem in equation 6.3 can be recast in Lagrangian,

$$\min_{\mathbf{v}, \mathbf{u}, \mathbf{u}_a} E_1, \quad E_1 = \frac{1}{2} \|\mathbf{R}\mathbf{u} - \mathbf{d}_{\text{obs}}\|_2^2 - \mathbf{u}_a^* (\mathbf{A}(\mathbf{v})\mathbf{u} - \mathbf{b}_{\text{obs}}) , \quad (6.4)$$

where $\mathbf{u}_a$ is the Lagrangian multiplier or adjoint wavefield. Solving the KKT condition in equation 6.4 gives the algorithm for simutenously update $\mathbf{v}, \mathbf{u}, \mathbf{u}_a$. This is named “all-at-once” approach. This method needs to solve the wavefields for all shots simutenously. And it is not widely used in geophysics industry because the dimentionality of the KKT condition is prohibitaly high.

In practice, practioners solve the problem in the reduced-space (Tarantola, 1984a),

$$E_{1,\text{red}}(\mathbf{v}) = \frac{1}{2} \|\mathbf{R}\mathbf{A}(\mathbf{v})^{-1}\mathbf{b}_{\text{obs}} - \mathbf{d}_{\text{obs}}\|_2^2 , \quad (6.5)$$

For each shot, the forward problem can be evaluated separately. So the problem becomes cheap to compute. This is the conventional FWI. Because the physics is enforced, the oscinating nature of $\mathbf{A}(\mathbf{v})^{-1}$ makes $E_{1,\text{red}}$ dependent on $\mathbf{v}$ in a highly nonlinear manner.

6.3 Source extension

We start from the penalty method by van Leeuwen and Herrmann (2013),

$$\min_{\mathbf{v}, \mathbf{u}} E_2, \quad E_2 = \frac{1}{2} \|\mathbf{R}\mathbf{u} - \mathbf{d}_{\text{obs}}\|_2^2 + \frac{\lambda^2}{2} \|\mathbf{A}(\mathbf{v})\mathbf{u} - \mathbf{b}_{\text{obs}}\|_2^2 , \quad (6.6)$$

where λ is the weight for the penalty. The penalty method is less prone to local minimum when λ is chosen small. The problem in equation 6.6 can be either solved by the reduced space method (van Leeuwen & Herrmann, 2016) or alternating optimization method (van Leeuwen & Herrmann, 2013).

To make the problem independent, G. Huang et al. (2018a) proposed the extended source waveform inversion,

$$\min_{\mathbf{v}, \mathbf{u}} E_3, \quad E_3 = \frac{1}{2} \|\mathbf{R}\mathbf{u} - \mathbf{d}_{\text{obs}}\|_2^2 + \frac{\lambda^2}{2} \|\mathbf{C}_\mathbf{A} (\mathbf{A}(\mathbf{v})\mathbf{u} - \mathbf{b}_{\text{obs}})\|_2^2 \quad (6.7)$$

$$\approx \frac{1}{2} \|\mathbf{R}\mathbf{u} - \mathbf{d}_{\text{obs}}\|_2^2 + \frac{\lambda^2}{2} \|\mathbf{C}_\mathbf{A} \mathbf{A}(\mathbf{v})\mathbf{u}\|_2^2, \quad (6.8)$$

where

$$\mathbf{C}_\mathbf{A} = \text{diag}(|\mathbf{x} - x_s| + \varepsilon_\mathbf{A}^2), \quad (6.9)$$

and $\varepsilon_\mathbf{A}$ is a small positive number to avoid singularity; $\mathbf{C}_\mathbf{A}$ is the so-called annihilator operator.

To further make ESI capable of estimating the source signature, we modify the problem in equation 6.8,

$$\min_{s, \mathbf{v}, \mathbf{u}} E_4, \quad E_4 = \frac{1}{2} \|\mathbf{R}\mathbf{u} - \mathbf{d}_{\text{obs}}\|_2^2 + \frac{\lambda^2}{2} \|\mathbf{C}_\mathbf{A} (\mathbf{A}(\mathbf{v})\mathbf{u} - \mathbf{b}_{\text{obs}}(s))\|_2^2, \quad (6.10)$$

where

$$\mathbf{b}_{\text{obs}}(s) = s\mathbf{B}_{\text{obs}}, \quad (6.11)$$

and $\mathbf{B}_{\text{obs}}$ is the discretization of $\delta(x - x_s)$.

Similarly, equation 6.10 can be also solved by the reduced space method (Fu & Symes, 2017a, 2017b; G. Huang et al., 2018a), which still leaves the parameters estimation nonlinear.

6.4 Augmented Lagrangian and TV regularization

In this work, we solve the Augmented Lagrangian (AL) of equation 6.10 by alternating direction method of multipliers (ADMM).

To make the following derivation more convenient, we rewrite equation 6.10 in a more compact form,

$$\min_{s, \mathbf{v}, \mathbf{u}} E_4, \quad E_4 = \frac{1}{2} \|\mathbf{R}\mathbf{u} - \mathbf{d}_{\text{obs}}\|_2^2 + \frac{\lambda^2}{2} \left\| \mathbf{C}_\mathbf{A} [\mathbf{A}(\mathbf{v}) \quad -\mathbf{B}_{\text{obs}}] \begin{bmatrix} \mathbf{u} \\ s \end{bmatrix} \right\|_2^2 \quad (6.12)$$

$$= \frac{1}{2} \left\| \begin{bmatrix} \mathbf{R} & 0 \\ \lambda \mathbf{C}_\mathbf{A} \mathbf{A}(\mathbf{v}) & -\lambda \mathbf{C}_\mathbf{A} \mathbf{B}_{\text{obs}} \end{bmatrix} \begin{bmatrix} \mathbf{u} \\ s \end{bmatrix} - \begin{bmatrix} \mathbf{d}_{\text{obs}} \\ \mathbf{0} \end{bmatrix} \right\|_2^2 \quad (6.13)$$

$$= \frac{1}{2} \|\mathbf{F}(\mathbf{v})\mathbf{U}(s, \mathbf{u}) - \mathbf{e}_\mathbf{a}\|_2^2, \quad (6.14)$$

where

$$\mathbf{F}(\mathbf{v}) = \begin{bmatrix} \mathbf{R} & 0 \\ \lambda \mathbf{C}_A \mathbf{A}(\mathbf{v}) & -\lambda \mathbf{C}_A \mathbf{B}_{\text{obs}} \end{bmatrix}, \quad (6.15)$$

$$\mathbf{U}(s, \mathbf{u}) = \begin{bmatrix} \mathbf{u} \\ s \end{bmatrix}, \quad (6.16)$$

$$\mathbf{e}_a = \begin{bmatrix} \mathbf{d}_{\text{obs}} \\ \mathbf{0} \end{bmatrix}. \quad (6.17)$$

$$(6.18)$$

Now a Lagrangian multiplier is introduced for equation 6.14,

$$\min_{\mathbf{v}, \mathbf{U}, \mathbf{U}_a} E_5, \quad E_5 = \frac{1}{2} \|\mathbf{F}(\mathbf{v})\mathbf{U}(s, \mathbf{u}) - \mathbf{e}_a\|_2^2 - \mathbf{U}_a^* (\mathbf{F}(\mathbf{v})\mathbf{U}(s, \mathbf{u}) - \mathbf{e}_a) \quad (6.19)$$

$$= \frac{1}{2} \|\mathbf{F}(\mathbf{v})\mathbf{U}(s, \mathbf{u}) - \mathbf{e}_a\|_2^2 - \mathbf{U}_a^* (\mathbf{F}(\mathbf{v})\mathbf{U}(s, \mathbf{u}) - \mathbf{e}_a) \quad (6.20)$$

$$+ \frac{1}{2} \|\mathbf{U}_a\|_2^2 - \frac{1}{2} \|\mathbf{U}_a\|_2^2 \quad (6.21)$$

$$= \frac{1}{2} \|\mathbf{F}(\mathbf{v})\mathbf{U}(s, \mathbf{u}) - \mathbf{e}_a - \mathbf{U}_a\|_2^2 - \frac{1}{2} \|\mathbf{U}_a\|_2^2 \quad (6.22)$$

Problem in equation 6.22 is named Iteratively Refined Extended Source waveform Inversion (IR-ESI). The seemly complicated optimization problem can be simplified to alternating update scheme. According to Nocedal and Wright (2006), the method of multipliers can be applied to equation 6.22,

$$(\mathbf{v}_{i+1}, \mathbf{U}_{i+1}) = \underset{\mathbf{v}, \mathbf{U}}{\text{argmin}} E_5(\mathbf{v}, \mathbf{U}, \mathbf{U}_{a,i}), \quad (6.23)$$

$$\mathbf{U}_{a,i+1} = \underset{\mathbf{U}_a}{\text{argmin}} E_5(\mathbf{v}_{i+1}, \mathbf{U}_{i+1}, \mathbf{U}_a), \quad (6.24)$$

$$= \mathbf{U}_{a,i} + \mathbf{e}_a - \mathbf{F}(\mathbf{v}_{i+1})\mathbf{U}_{i+1}. \quad (6.25)$$

Let us introduce some auxiliary variables,

$$\mathbf{U}_{a,i} = \begin{bmatrix} \mathbf{d}_i \\ \lambda \mathbf{C}_A \mathbf{b}_i \end{bmatrix}, \quad (6.26)$$

then the problem in equation 6.23 becomes,

$$(s_{i+1}, \mathbf{u}_{i+1}) = \underset{s, \mathbf{u}}{\text{argmin}} \frac{1}{2} \|\mathbf{F}(\mathbf{v}_i)\mathbf{U}(s, \mathbf{u}) - \mathbf{e}_a - \mathbf{U}_{a,i}\|_2^2 \quad (6.27)$$

$$= \frac{1}{2} \left\| \begin{bmatrix} \mathbf{R} & 0 \\ \lambda \mathbf{C}_A \mathbf{A}(\mathbf{v}_i) & -\lambda \mathbf{C}_A \mathbf{B} \end{bmatrix} \begin{bmatrix} \mathbf{u} \\ s \end{bmatrix} - \begin{bmatrix} \mathbf{d} \\ \mathbf{0} \end{bmatrix} - \begin{bmatrix} \mathbf{d}_i \\ \lambda \mathbf{C}_A \mathbf{b}_i \end{bmatrix} \right\|_2^2. \quad (6.28)$$

Now we solve the optimization problem combining equations 6.23 and 6.25,

$$(s_{i+1}, \mathbf{u}_{i+1}) = \underset{s, \mathbf{u}}{\operatorname{argmin}} \frac{1}{2} \|\mathbf{R}\mathbf{u} - \mathbf{d}_{\text{obs}} - \mathbf{d}_i\| + \frac{\lambda^2}{2} \|\mathbf{C}_A (\mathbf{A}(\mathbf{v}_i)\mathbf{u} - \mathbf{b}_{\text{obs}}(s) - \mathbf{b}_i)\|_2^2, \quad (6.29)$$

$$\mathbf{v}_{i+1} = \underset{\mathbf{v}}{\operatorname{argmin}} \frac{1}{2} \|\mathbf{C}_A (\mathbf{A}(\mathbf{v})\mathbf{u}_{i+1} - \mathbf{b}_{\text{obs}}(s_{i+1}) - \mathbf{b}_i)\|_2^2, \quad (6.30)$$

$$\mathbf{d}_{i+1} = \mathbf{d}_i + \mathbf{d}_{\text{obs}} - \mathbf{R}\mathbf{u}_{i+1}, \quad (6.31)$$

$$\mathbf{b}_{i+1} = \mathbf{b}_i + \mathbf{b}_{\text{obs}}(s_{i+1}) - \mathbf{A}(\mathbf{v}_{i+1})\mathbf{u}_{i+1}. \quad (6.32)$$

Compared to Aghamiry et al. (2019b, their equation 8), our algorithm estimates source wavelet in equation 6.29. And an additional annihilator $\mathbf{C}_A$ appears in equations 6.29 and 6.30.

In equations 6.29 - 6.32, the dual variables are only updated once after the primal variables are computed. A more efficient approach is the Peaceman-Rachford splitting method with relaxation factors (Aghamiry et al., 2019b; He et al., 2014),

$$(s_{i+1}, \mathbf{u}_{i+1}) = \underset{s, \mathbf{u}}{\operatorname{argmin}} \frac{1}{2} \|\mathbf{R}\mathbf{u} - \mathbf{d}_{\text{obs}} - \mathbf{d}_i\| + \frac{\lambda^2}{2} \|\mathbf{C}_A (\mathbf{A}(\mathbf{v}_i)\mathbf{u} - \mathbf{b}_{\text{obs}}(s) - \mathbf{b}_i)\|_2^2, \quad (6.33)$$

$$\mathbf{d}_{i+1} = \mathbf{d}_i + \mathbf{d}_{\text{obs}} - \mathbf{R}\mathbf{u}_{i+1}, \quad (6.34)$$

$$\mathbf{b}_{i+\frac{1}{2}} = \mathbf{b}_i + \alpha_{\text{prsm}} (\mathbf{b}_{\text{obs}}(s_{i+1}) - \mathbf{A}(\mathbf{v}_{i+1})\mathbf{u}_{i+1}), \quad (6.35)$$

$$\mathbf{v}_{i+1} = \underset{\mathbf{v}}{\operatorname{argmin}} \frac{1}{2} \left\| \mathbf{C}_A \left(\mathbf{A}(\mathbf{v})\mathbf{u}_{i+1} - \mathbf{b}_{\text{obs}}(s_{i+1}) - \mathbf{b}_{i+\frac{1}{2}} \right) \right\|_2^2, \quad (6.36)$$

$$\mathbf{b}_{i+1} = \mathbf{b}_{i+\frac{1}{2}} + \alpha_{\text{prsm}} (\mathbf{b}_{\text{obs}}(s_{i+1}) - \mathbf{A}(\mathbf{v}_{i+1})\mathbf{u}_{i+1}), \quad (6.37)$$

where $\alpha_{\text{prsm}} \in (0, 1)$ is the step length. In our experiments, $\alpha_{\text{prsm}} = 0.5$.

6.4.1 Optimization over s and $\mathbf{u}$

To solve equation 6.29, we simply take the first-order optimality condition for both s and $\mathbf{u}$,

$$\begin{bmatrix} \mathbf{u}_{i+1} \\ s_{i+1} \end{bmatrix} = \begin{bmatrix} \mathbf{R} & \mathbf{0}_R \\ \lambda \mathbf{C}_A \mathbf{A}(\mathbf{v}_i) & -\lambda \mathbf{C}_A \mathbf{B}_{\text{obs}} \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{d}_{\text{obs}} + \mathbf{d}_i \\ \lambda \mathbf{C}_A \mathbf{b}_i \end{bmatrix}, \quad (6.38)$$

where $\mathbf{0}_R$ is a zero matrix whose size is the same with $\mathbf{R}$. Equation 6.38 is the data augmented wave equation.

6.4.2 Optimization over $\mathbf{v}$

Another key step is to solve equation 6.36 to update the model parameters $\mathbf{v}$. To make the problem tractable to solve, a linearization is performed for the term

$$\mathbf{A}(\mathbf{v})\mathbf{u}_{i+1},$$

$$\mathbf{v}_{i+1} = \underset{\mathbf{v}}{\operatorname{argmin}} \frac{1}{2} \left\| \mathbf{C}_{\mathbf{A}} \left(\mathbf{A}(\mathbf{v}_i)\mathbf{u}_{i+1} + \mathbf{G}(\mathbf{v}_i, \mathbf{u}_{i+1})(\mathbf{v} - \mathbf{v}_i) - \mathbf{b}_{\text{obs}}(s_{i+1}) - \mathbf{b}_{i+\frac{1}{2}} \right) \right\|_2^2, \quad (6.39)$$

$$(6.40)$$

where

$$\mathbf{G}(\mathbf{v}_i, \mathbf{u}) = \left. \frac{\partial \mathbf{A}(\mathbf{v})\mathbf{u}}{\partial \mathbf{v}} \right|_{\mathbf{v}=\mathbf{v}_i} = -2(-i\omega)^2 \operatorname{diag}(\mathbf{v}^{-3})\mathbf{u}. \quad (6.41)$$

Now equation 6.39 becomes linear with respect to $\mathbf{v}$ at each iteration and one can take the first-order optimality for updating $\mathbf{v}$.

In our study, we introduce a total variation regularization for equation 6.39. In our study, we employ the Levenberg-Marquardt type TV,

$$\begin{aligned} \mathbf{v}_{i+1} = \underset{\mathbf{v}}{\operatorname{argmin}} \frac{1}{2} \left\| \mathbf{C}_{\mathbf{A}} \left(\mathbf{A}(\mathbf{v}_i)\mathbf{u}_{i+1} + \mathbf{G}(\mathbf{v}_i, \mathbf{u}_{i+1})(\mathbf{v} - \mathbf{v}_i) - \mathbf{b}_{\text{obs}}(s_{i+1}) - \mathbf{b}_{i+\frac{1}{2}} \right) \right\|_2^2 \\ + \alpha D_{\text{TV}_{\epsilon_{\text{tv}}}}(\mathbf{v}, \mathbf{v}_i), \end{aligned} \quad (6.42)$$

where α is the size of the regularization;

$$D_{\text{TV}_{\epsilon_{\text{tv}}}}(\mathbf{v}_{i+1}, \mathbf{v}_i) = \text{TV}_{\epsilon_{\text{tv}}}(\mathbf{v}_{i+1}) - \text{TV}_{\epsilon_{\text{tv}}}(\mathbf{v}_i) + \langle \text{TV}'_{\epsilon_{\text{tv}}}(\mathbf{v}_i), \mathbf{v}_{i+1} - \mathbf{v}_i \rangle, \quad (6.43)$$

and

$$\text{TV}_{\epsilon_{\text{tv}}}(\mathbf{v}) \approx \sum \sqrt{\left(\frac{\partial v(\mathbf{x})}{\partial x_1} \right)^2 + \left(\frac{\partial v(\mathbf{x})}{\partial x_3} \right)^2 + \epsilon_{\text{tv}}^2}. \quad (6.44)$$

where $\sum$ denotes the sum for all grid points; ϵ_{tv} is a small positive number to avoid singularity.

Insert equation 6.43 into equation 6.42, we have

$$\begin{aligned} \mathbf{v}_{i+1} = \underset{\mathbf{v}}{\operatorname{argmin}} \frac{1}{2} \left\| \mathbf{C}_{\mathbf{A}} \left(\mathbf{A}(\mathbf{v}_i)\mathbf{u}_{i+1} + \mathbf{G}(\mathbf{v}_i, \mathbf{u}_{i+1})(\mathbf{v} - \mathbf{v}_i) - \mathbf{b}_{\text{obs}}(s_{i+1}) - \mathbf{b}_{i+\frac{1}{2}} \right) \right\|_2^2 \\ + \alpha (\text{TV}_{\epsilon_{\text{tv}}}(\mathbf{v}) - \langle \text{TV}'(\mathbf{v}_i), \mathbf{v} \rangle). \end{aligned} \quad (6.45)$$

We introduce an auxiliary variable,

$$\boldsymbol{\xi} = \text{TV}'(\mathbf{v}) = -\nabla \cdot \frac{\nabla v(\mathbf{x})}{\sqrt{|\nabla v(\mathbf{x})|^2 + \epsilon_{\text{tv}}^2}}. \quad (6.46)$$

Equation 6.45 can be now recast as the Bregman iteration,

$$\begin{aligned} \mathbf{v}_{i+1} = \underset{\mathbf{v}}{\operatorname{argmin}} \frac{1}{2} \left\| \mathbf{C}_A \left(\mathbf{A}(\mathbf{v}_i) \mathbf{u}_{i+1} + \mathbf{G}(\mathbf{v}_i, \mathbf{u}_{i+1})(\mathbf{v} - \mathbf{v}_i) - \mathbf{b}_{\text{obs}}(s_{i+1}) - \mathbf{b}_{i+\frac{1}{2}} \right) \right\|_2^2 \\ + \alpha (\text{TV}_{\epsilon_{\text{tv}}}(\mathbf{v}) - \langle \boldsymbol{\xi}_i, \mathbf{v} \rangle) , \end{aligned} \quad (6.47)$$

$$\begin{aligned} \boldsymbol{\xi}_{i+1} = \boldsymbol{\xi}_i - \frac{1}{\alpha} (\mathbf{C}_A \mathbf{G}(\mathbf{v}_i, \mathbf{u}_{i+1}))^* \mathbf{C}_A \\ \left(\mathbf{A}(\mathbf{v}_i) \mathbf{u}_{i+1} + \mathbf{G}(\mathbf{v}_i, \mathbf{u}_{i+1})(\mathbf{v}_{i+1} - \mathbf{v}_i) - \mathbf{b}_{\text{obs}}(s_{i+1}) - \mathbf{b}_{i+\frac{1}{2}} \right) . \end{aligned} \quad (6.48)$$

Because the first-order optimality condition of equation 6.47 is nonlinear for $\mathbf{v}$, we need inner iterations to solve it. With the change of variable from v to u_v for the inner iterations, we have

$$\begin{aligned} -\alpha \nabla \cdot \frac{\nabla u_v(\mathbf{x})}{\sqrt{|\nabla u_v(\mathbf{x})|^2 + \epsilon_{\text{tv}}^2}} \\ + (\mathbf{C}_A \mathbf{G}(\mathbf{u}_{\mathbf{v}_i}, \mathbf{u}_{i+1}))^* \mathbf{C}_A \left(\mathbf{A}(\mathbf{v}_i) \mathbf{u}_{i+1} + \mathbf{G}(\mathbf{v}_i, \mathbf{u}_{i+1})(\mathbf{u}_{\mathbf{v}} - \mathbf{v}_i) - \mathbf{b}_{\text{obs}}(s_{i+1}) - \mathbf{b}_{i+\frac{1}{2}} \right) = 0 , \end{aligned} \quad (6.49)$$

where $\mathbf{u}_{\mathbf{v}} = u_v(\mathbf{x})$.

The primal-dual Newton's method is employed to solve equation 6.49. It is noticed that the only nonliterary in equation 6.49 is coming from the term $\frac{\nabla u_v(\mathbf{x})}{\sqrt{|\nabla u_v(\mathbf{x})|^2 + \epsilon_{\text{tv}}^2}}$.

Thus, a linearization is applied to the term,

$$w(\mathbf{x}) = \frac{\nabla u_v(\mathbf{x})}{\sqrt{|\nabla u_v(\mathbf{x})|^2 + \epsilon_{\text{tv}}^2}} , \quad (6.50)$$

where $w(\mathbf{x}) = [w_1(\mathbf{x}) \ w_3(\mathbf{x})]^*$ in 2D and $|w(\mathbf{x})| \leq 1$.

With the linearization in equation 6.50, we obtain a new linear system of PDEs for $\mathbf{u}_{\mathbf{v}}$ and $w(\mathbf{x})$,

$$\begin{aligned} -\alpha \nabla \cdot w(\mathbf{x}) + (\mathbf{C}_A \mathbf{G}(\mathbf{u}_{\mathbf{v}_i}, \mathbf{u}_{i+1}))^* \\ \mathbf{C}_A \left(\mathbf{A}(\mathbf{v}_i) \mathbf{u}_{i+1} + \mathbf{G}(\mathbf{v}_i, \mathbf{u}_{i+1})(\mathbf{u}_{\mathbf{v}} - \mathbf{v}_i) - \mathbf{b}_{\text{obs}}(s_{i+1}) - \mathbf{b}_{i+\frac{1}{2}} \right) = 0 , \end{aligned} \quad (6.51)$$

$$w_1(\mathbf{x}) \sqrt{|\nabla u_v(\mathbf{x})|^2 + \epsilon_{\text{tv}}^2} - \frac{\partial u_v(\mathbf{x})}{\partial x_1} = 0 , \quad (6.52)$$

$$w_3(\mathbf{x}) \sqrt{|\nabla u_v(\mathbf{x})|^2 + \epsilon_{\text{tv}}^2} - \frac{\partial u_v(\mathbf{x})}{\partial x_3} = 0 . \quad (6.53)$$

The system can be solved efficiently using Newton's method. At each iteration, updating w requires a line search to preserve the constraint $|w(\mathbf{x})| \leq 1$.

Suppose $\mathbf{v} = v(\mathbf{x})$ and $v(\mathbf{x}_{e,l}) = \mathbf{v}_{e,l}$ is defined on an equispaced grid in two space dimensions with $N_{\text{space},1} \times N_{\text{space},3}$ grids, $\{(eh_{\text{space},1}, lh_{\text{space},3}) : e = 1, 2, \dots, N_{\text{space},1}, l =$

$1, 2, \dots, N_{\text{space},3}\}$. Here $h_{\text{space},1}$ and $h_{\text{space},3}$ are the grid spacings in x_1 - and x_3 -directions, respectively. To solve the system in equations 6.51 - 6.51, we firstly define the discretized differential operators $\mathbf{D}_1$ and $\mathbf{D}_3$,

$$(\mathbf{D}_1 \mathbf{v})_{e,l} = \frac{\mathbf{v}_{e,l} - \mathbf{v}_{e-1,l}}{h_{\text{space},1}}, \quad (6.54)$$

$$(\mathbf{D}_3 \mathbf{v})_{e,l} = \frac{\mathbf{v}_{e,l} - \mathbf{v}_{e,l-1}}{h_{\text{space},3}}. \quad (6.55)$$

For the discrete Newton updates $\Delta \mathbf{u}$, $\Delta \mathbf{w}_1$ and $\Delta \mathbf{w}_3$, following Vogel (2002), we obtain,

$$\Delta \mathbf{u}_{\mathbf{v}} = ((\mathbf{C}_A \mathbf{G}(\mathbf{v}_i, \mathbf{u}_{i+1}))^* \mathbf{C}_A \mathbf{G}(\mathbf{v}_i, \mathbf{u}_{i+1}) + \alpha \mathbf{A}_1)^{-1} \mathbf{b}_1, \quad (6.56)$$

where

$$\begin{aligned} \mathbf{A}_1 = & \mathbf{D}_1^* \text{diag} \left(\text{diag}(\mathbf{a}_2)^{-1} \left[1 + \text{diag}(\mathbf{a}_2)^{-1} \text{diag} \left(\mathbf{D}_1 (\mathbf{u}_{\mathbf{v}})_j \right) (\mathbf{w}_1)_j \right] \right) \mathbf{D}_1 \\ & + \mathbf{D}_1^* \text{diag} \left(\text{diag}(\mathbf{a}_2)^{-2} \text{diag} \left(\mathbf{D}_1 (\mathbf{u}_{\mathbf{v}})_j \right) (\mathbf{w}_1)_j \right) \mathbf{D}_3 \\ & + \mathbf{D}_3^* \text{diag} \left(\text{diag}(\mathbf{a}_2)^{-1} \left[1 + \text{diag}(\mathbf{a}_2)^{-1} \text{diag} \left(\mathbf{D}_3 (\mathbf{u}_{\mathbf{v}})_j \right) (\mathbf{w}_3)_j \right] \right) \mathbf{D}_3 \\ & + \mathbf{D}_3^* \text{diag} \left(\text{diag}(\mathbf{a}_2)^{-2} \text{diag} \left(\mathbf{D}_3 (\mathbf{u}_{\mathbf{v}})_j \right) (\mathbf{w}_1)_j \right) \mathbf{D}_3, \quad (6.57) \\ \mathbf{b}_1 = & - (\mathbf{C}_A \mathbf{G}(\mathbf{v}_i, \mathbf{u}_{i+1}))^* \mathbf{C}_A \\ & \left(\mathbf{A}(\mathbf{v}_i) \mathbf{u}_{i+1} - \mathbf{G}(\mathbf{v}_i, \mathbf{u}_{i+1})((\mathbf{u}_{\mathbf{v}})_j - \mathbf{v}_i) - \mathbf{b}_{\text{obs}} - \mathbf{b}_i \right) \\ & \alpha \mathbf{D}_1 \text{diag}(\mathbf{a}_2)^{-1} \mathbf{D}_1 (\mathbf{u}_{\mathbf{v}})_j + \alpha \mathbf{D}_3 \text{diag}(\mathbf{a}_2)^{-1} \mathbf{D}_3 (\mathbf{u}_{\mathbf{v}})_j, \end{aligned}$$

where $\mathbf{a}_2$ is the discretization of $\sqrt{\left(\frac{\partial(u_v)_j}{\partial x_1}\right)^2 + \left(\frac{\partial(u_v)_j}{\partial x_3}\right)^2 + \epsilon_{\text{tv}}^2}$ and $(u_v)_j$ is continuous form u_v at inner iteration j .

If we impose a bounded constraint for $\mathbf{v}$, we use the projection for the inner iterate $\mathbf{u}_{\mathbf{v}}$,

$$\Delta \tilde{\mathbf{u}}_{\mathbf{v}} = ((\mathbf{C}_A \mathbf{G}(\mathbf{v}_i, \mathbf{u}_{i+1}))^* \mathbf{C}_A \mathbf{G}(\mathbf{v}_i, \mathbf{u}_{i+1}) + \alpha \mathbf{A}_1)^{-1} \mathbf{b}_1, \quad (6.58)$$

$$\Delta \mathbf{u}_{\mathbf{v}} = \max(\mathbf{v}_1 - (\mathbf{u}_{\mathbf{v}})_j, \min(\mathbf{v}_{\mathbf{u}} - (\mathbf{u}_{\mathbf{v}})_j, -\Delta \tilde{\mathbf{u}}_{\mathbf{v}})), \quad (6.59)$$

where $\mathbf{v}_1$ and $\mathbf{v}_{\mathbf{u}}$ is the lower and upper bounds of the velocity, respectively.

And the dual updates can be computed by,

$$\begin{aligned}\Delta \mathbf{w}_1 = & -(\mathbf{w}_1)_j \\ & + \text{diag}(\mathbf{a}_2)^{-1} \left(\mathbf{D}_1(\mathbf{u}_v)_{j+1} + \text{diag} \left(1 + \text{diag}(\mathbf{a}_2)^{-1} \text{diag} \left(\mathbf{D}_1(\mathbf{u}_v)_{j+1} \right) (\mathbf{w}_1)_j \right) \right) \\ & \mathbf{D}_1 \Delta \mathbf{u}_v \\ & + \text{diag}(\mathbf{a}_2)^{-1} \text{diag} \left(\mathbf{D}_3(\mathbf{u}_v)_{j+1} \right) \text{diag} \left((\mathbf{w}_1)_j \right) \mathbf{D}_3 \Delta \mathbf{u}_v ,\end{aligned}\tag{6.60}$$

$$\begin{aligned}\Delta \mathbf{w}_3 = & -(\mathbf{w}_3)_j \\ & + \text{diag}(\mathbf{a}_2)^{-1} \left(\mathbf{D}_3(\mathbf{u}_v)_{j+1} + \text{diag} \left(1 + \text{diag}(\mathbf{a}_2)^{-1} \text{diag} \left(\mathbf{D}_3(\mathbf{u}_v)_{j+1} \right) (\mathbf{w}_3)_j \right) \right) \\ & \mathbf{D}_3 \Delta \mathbf{u}_v \\ & + \text{diag}(\mathbf{a}_2)^{-1} \text{diag} \left(\mathbf{D}_1(\mathbf{u}_v)_{j+1} \right) \text{diag} \left((\mathbf{w}_3)_j \right) \mathbf{D}_1 \Delta \mathbf{u}_v ,\end{aligned}\tag{6.61}$$

where $\mathbf{w}_1 = w_1(\mathbf{x})$ and $\mathbf{w}_3 = w_3(\mathbf{x})$.

And the update for the inner iteration is given by,

$$(\mathbf{u}_v)_{j+1} = (\mathbf{u}_v)_j + \Delta \mathbf{u}_v ,\tag{6.62}$$

$$(\mathbf{w}_1)_{j+1} = (\mathbf{w}_1)_j + \alpha_w \Delta \mathbf{w}_1 ,\tag{6.63}$$

$$(\mathbf{w}_3)_{j+1} = (\mathbf{w}_3)_j + \alpha_w \Delta \mathbf{w}_3 ,\tag{6.64}$$

where α_w is the step length for the update of w which ensures $\|w\|_\infty \leq 1$,

$$\alpha_w = \max_{\tau_w} \{0 \leq \tau_w \leq 1 \mid \|(\mathbf{w}_1)_j + \tau_w \Delta \mathbf{w}_1\|^2 + \|(\mathbf{w}_3)_j + \tau_w \Delta \mathbf{w}_3\|^2\|_\infty \leq 1\} .\tag{6.65}$$

An algorithm for SE-IR-ESI with TV solved by the primal-dual Newton's method and bounded constraint is given in algorithm 6.1.

Algorithm 6.2 is for the inner iterations. In practice, we found that 1 inner iteration is enough because SE-IR-ESI with TV does not estimate the velocity accurately in the early iterations so that it does not make much sense to iterate so much for the TV denoising.

Algorithm 6.1 for SE-IR-ESI with TV is not sensitive to the choice of λ , compared to the conventional WRI (Aghamiry et al., 2019b). λ can be fixed during the iteration. According to van Leeuwen and Herrmann (2016), we get an estimate of $\|\mathbf{R}\mathbf{A}^{-1}\|_2^2$.

6.5 Numerical tests

6.5.1 Marmousi

We test our algorithm on the a subset of the Marmousi model. Figure 6.1 gives the ground truth and the initial iterate for the inversion. We place 75 sources and 373 receivers evenly at $x_1 \in [300, 7237.5]$ m and $x_3 = 281.25$ m.

Algorithm 6.1: Outer iteration for TV-SE-IR-ESI with TV solved by the primal-dual Newton's method and bounded constraint.

$n_{\text{outer}} :=$ maximum outer iterations.
for $i = 1 : n_{\text{outer}}$ **do**
 Compute $\mathbf{u}_{i+1}$ and s_{i+1} by equation 6.38.
 Compute $\mathbf{d}_{i+1}$ by equation 6.34.
 Compute $\mathbf{b}_{i+\frac{1}{2}}$ by equation 6.35.
 if $i=1$ **then**
 Employ inner iteration in algorithm 6.2 to update $\mathbf{u}_{\mathbf{v}}$, $\mathbf{w}_1$ and $\mathbf{w}_3$, with initial iterates being zero vectors. The final inner iterate $\mathbf{u}_{\mathbf{v}}$ is assigned to $\mathbf{v}_{i+1}$.
 else
 Employ inner iteration in algorithm 6.2 to update $\mathbf{u}_{\mathbf{v}}$, $\mathbf{w}_1$ and $\mathbf{w}_3$, with initial iterates being the final iterates from last round of inner iterations. The final inner iterate $\mathbf{u}_{\mathbf{v}}$ is assigned to $\mathbf{v}_{i+1}$.
 end
 Compute ξ_{i+1} by equation 6.48.
 Compute $\mathbf{b}_{i+1}$ by equation 6.37.
 if *Data misfit reaches the threshold* **then**
 Break.
 end
end
Return $\mathbf{v}_{i+1}$ and $\mathbf{s}_{i+1}$.

Algorithm 6.2: Inner iteration for TV-SE-IR-ESI solved by the primal-dual Newton's method and bounded constraint.

$n_{\text{inner}} :=$ maximum inner iterations.
for $i = 1 : n_{\text{inner}}$ **do**
 Compute $\Delta \mathbf{u}_{\mathbf{v}}$ by equations 6.58 and 6.59.
 Compute $\Delta \mathbf{w}_1$ and $\Delta \mathbf{w}_3$ by equations 6.60 and 6.61
 Compute $(\mathbf{u}_{\mathbf{v}})_{j+1}$ by equation 6.62.
 Compute $(\mathbf{w}_1)_{j+1}$ and $(\mathbf{w}_3)_{j+1}$ by equations 6.63 - 6.65.
 if *The relative change of $\mathbf{u}_{\mathbf{v}}$ reaches the threshold* **then**
 Break.
 end
end
Return $(\mathbf{u}_{\mathbf{v}})_{j+1}$, $(\mathbf{w}_1)_{j+1}$ and $(\mathbf{w}_3)_{j+1}$.

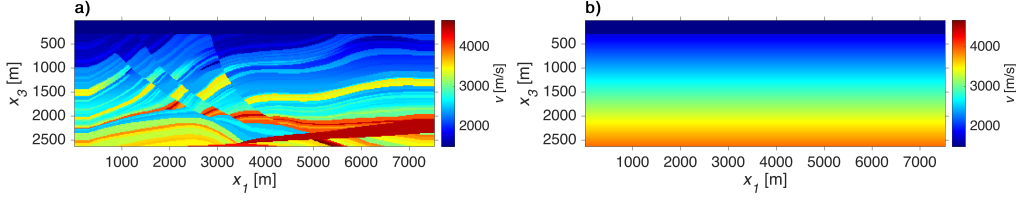


Figure 6.1: Marmousi model. (a) Ground truth. (b) Initial iterate.

The source wavelet is a Ricker wavelet with dominant frequency of 4.7 Hz. The frequency point selected for the inversion is given in Figure 6.2. Note that the data less than 3 Hz is not included in the inversion.

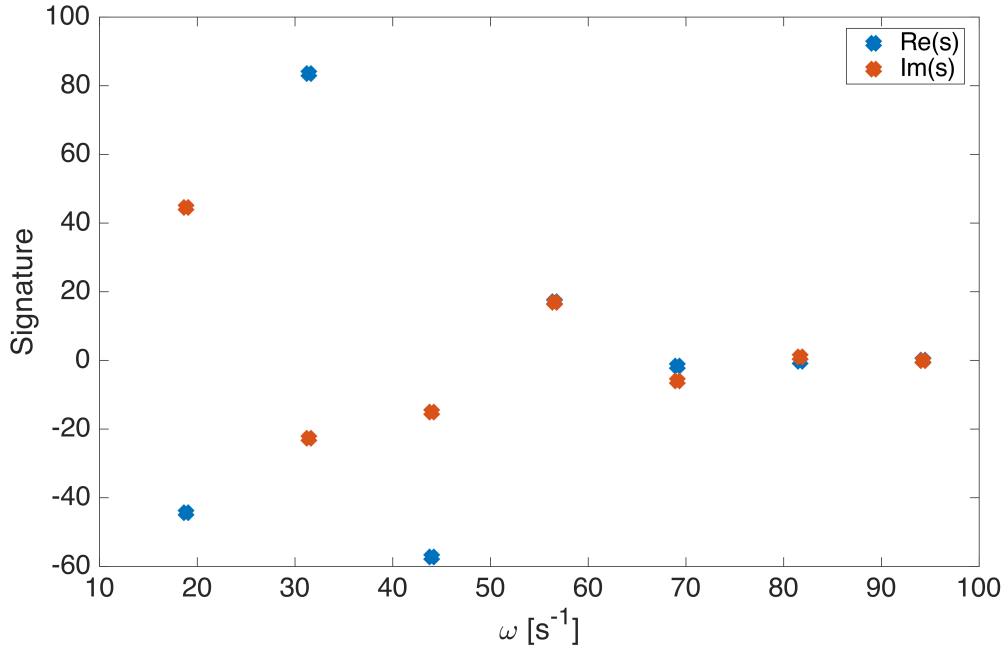


Figure 6.2: Source wavelet for the Marmousi model experiment.

In our experiments, we use conventional FWI (solved in the reduced space), SE-IR-ESI with and without TV regularizations. The estimation according to van Leeuwen and Herrmann (2016) suggests $\lambda = 1$. For all experiments for SE-IR-ESI, a bounded constraint, whose lower and upper bounds are the same with the ground truth, is employed. For SE-IR-ESI with TV regularizations, we tested three values for α , 10^{-8} , 10^{-7} and 10^{-6} . $\epsilon_{tv} = 10^{-2}$ for the size of the regularization. The frequency continuation strategy is used. For each frequency, a maximum of 30 iterations are performed. The inversion is ceased when the relative change of the total misfit is less than 0.01.

Figure 6.3 presents the inversion results. Figure 6.3a, obtained using conventional FWI, suffers from cycle skipping and fails to recover the main features of the Mar-

mousi model. In contrast, Figures 6.3b–e produce more reliable models with similar overall structures. The piece-wise constants are reduced with larger α .

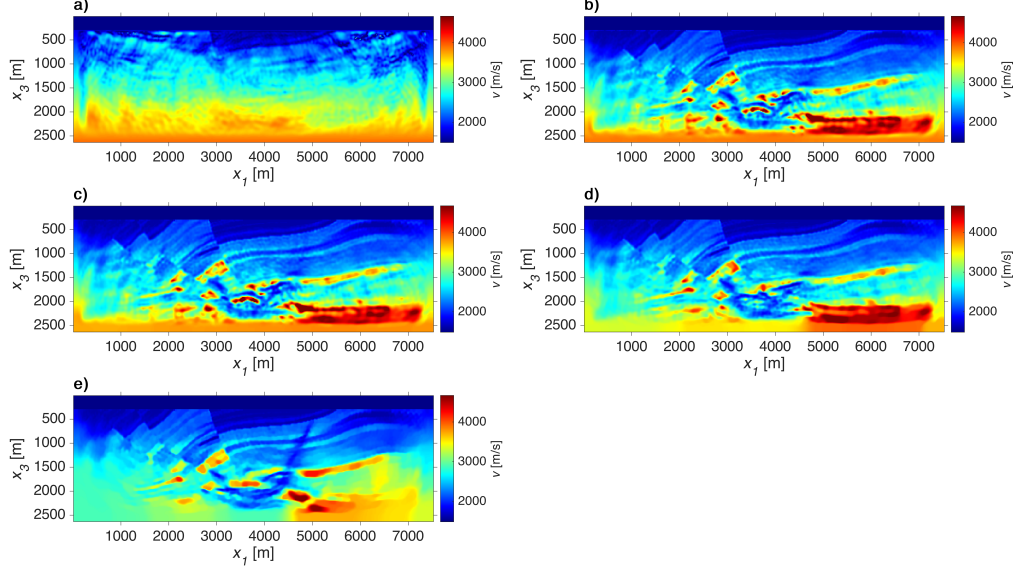


Figure 6.3: Inversion results for the Marmousi model. (a) is obtained by the conventional FWI. (b), (c), (d) and (e) is computed using SE-IR-ESI without regularizations, with TV reuglarization $\alpha = 10^{-8}$, with TV reuglarization $\alpha = 10^{-7}$, and with TV reuglarization $\alpha = 10^{-6}$. All TV regularizations are assigned with $\epsilon_{tv} = 10^{-2}$. The bounded constraints of (b), (c), (d) and (e) are set the minimum and maximum velocities of the ground truth.

Figure 6.4 illustrates the convergence history of the experiments. Figure 6.4a shows the data misfit for conventional FWI. The computation of the data misfit E_d , physics misfit E_p , and model misfit E_m is given in equations 6.66–6.68,

$$E_d(\mathbf{u}) = \frac{1}{2} \|\mathbf{R}\mathbf{u} - \mathbf{d}\|_2^2, \quad (6.66)$$

$$E_p(\mathbf{v}) = \frac{\lambda^2}{2} \|\mathbf{C}_A(\mathbf{A}(\mathbf{v}) - \mathbf{b}_{\text{obs}})\|_2^2, \quad (6.67)$$

$$E_m(\mathbf{v}) = \frac{1}{2} \|\mathbf{v} - \mathbf{v}_t\|_2^2. \quad (6.68)$$

Although the data misfit decreases with iterations for conventional FWI (Figure 6.4a), the model misfit in Figure 6.4d does not exhibit a clear downward trend. Figure 6.4b indicates that data misfits of SE-IR-ESI are generally higher when TV regularization is applied compared to cases without it. Meanwhile, Figure 6.4c shows that the physics misfit from SE-IR-ESI remains similar regardless of the presence of TV regularization. In Figure 6.4d, the model misfits for SE-IR-ESI without TV regularization and with $\alpha = 10^{-8}$ or $\alpha = 10^{-7}$ are comparable. However, when $\alpha = 10^{-6}$, the model misfit convergence slows significantly due to excessive TV penalization. In practice, an optimal α can be determined by running multiple

inversions with different values and selecting the one that yields the lowest data misfit, as defined in equation 6.66.

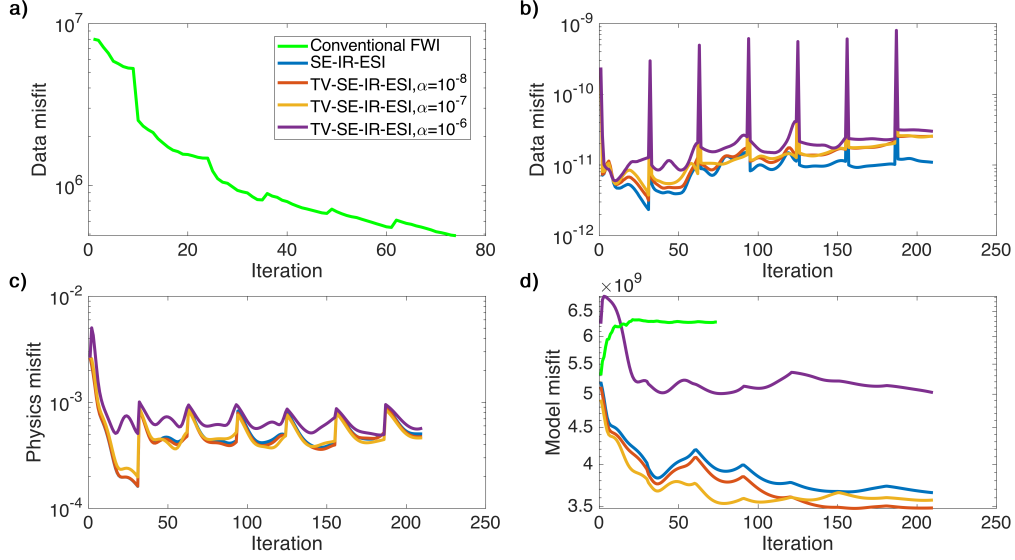


Figure 6.4: Convergence history for the Marmousi model inversion.

Figure 6.5 presents the wavefields computed from different models. In Figure 6.5c, the wavefield obtained using the conventional FWI model deviates the most from the ground truth shown in Figure 6.5a. In contrast, Figures 6.5d–g closely resemble the ground truth. Additionally, Figure 6.6 further confirms that SE-IR-ESI is less susceptible to cycle skipping compared to conventional FWI.

Table 6.1 gives the wavelet estimate from our SE-IR-ESI algorithm. It shows that the wavelets obtained for all frequencies honors the ground truth.

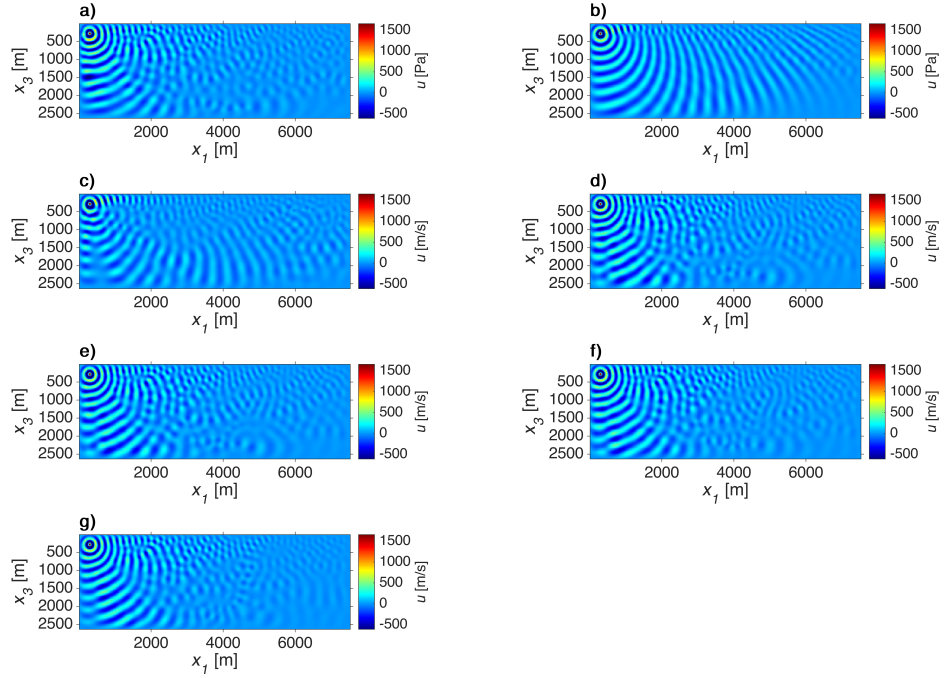


Figure 6.5: Acoustic wavefield simulated for the left-most source, 9 Hz source. (a) Marmousi ground truth, (b) initial iterate, (c) velocity recovered by conventional FWI, (d) SE-IR-ESI without regularization, (e) SE-IR-ESI with TV regularization $\alpha = 10^{-8}$, (f) SE-IR-ESI with TV regularization $\alpha = 10^{-7}$, and (g) SE-IR-ESI with TV regularization $\alpha = 10^{-6}$.

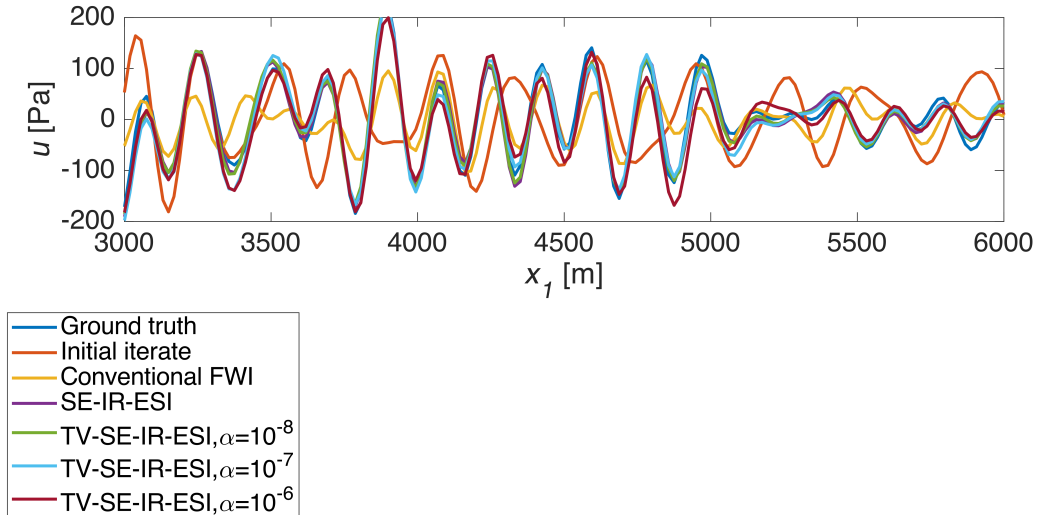


Figure 6.6: Wavefield sampled at receiver depth for Figure 6.5.

Table 6.1: Wavelet estimation for the Marmousi model.

Experiment	Frequency1	Frequency2	Frequency3	Frequency4	Frequency5	Frequency6	Frequency7
Ground truth	-44.3+44.5i	83.5-22.7i	-57.3-15.1i	17.1+16.9i	-1.65-5.98i	-0.259+1.03i	0.0897-0.0801i
SE-IR-ESI	-44.5+44.5i	82.9-22.7i	-57.3-14.7i	17.0+17.1i	-1.77-6.01i	-0.263+1.04i	0.0891-0.0818i
TV-SE-IR-ESI, $\alpha = 10^{-8}$	-44.5+44.5i	82.9-22.7i	-57.3-14.8i	17.1+17.1i	-1.79-6.02i	-0.260+1.04i	0.0890-0.0815i
TV-SE-IR-ESI, $\alpha = 10^{-7}$	-44.7+44.4i	83.4-22.6i	-57.4-15.0i	17.2+17.3i	-1.83-6.06i	-0.266+1.05i	0.0891-0.0825i
TV-SE-IR-ESI, $\alpha = 10^{-6}$	-42.7+45.5i	83.6-21.7i	-58.6-15.6i	17.8+17.7i	-1.75-6.25i	-0.267+1.08i	0.0910-0.0793i

6.5.2 2004 BP

The algorithms are also tested for the 2004 BP model. Figure 6.7 presents the ground truth and the initial iterate for the inversion. The setup includes 241 sources and 243 receivers, uniformly distributed along the line $x_1 \in [300, 7237.5]$ m at a depth of $x_3 = 281.25$ m. The source wavelet is provided in Figure 6.8.

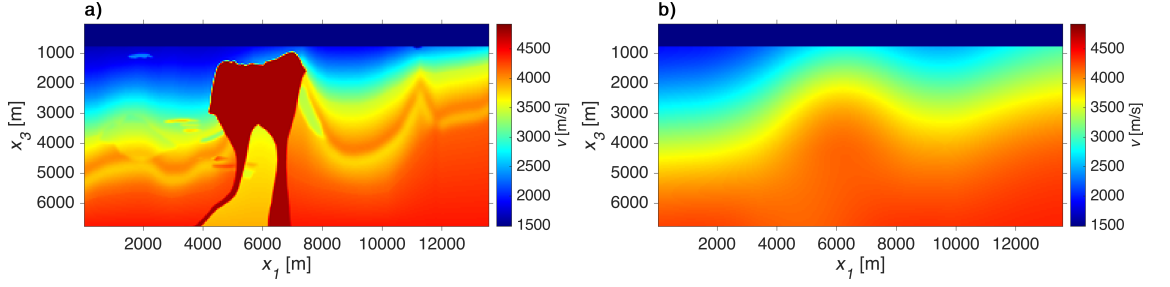


Figure 6.7: 2004 BP model. (a) Ground truth. (b) Initial iterate.

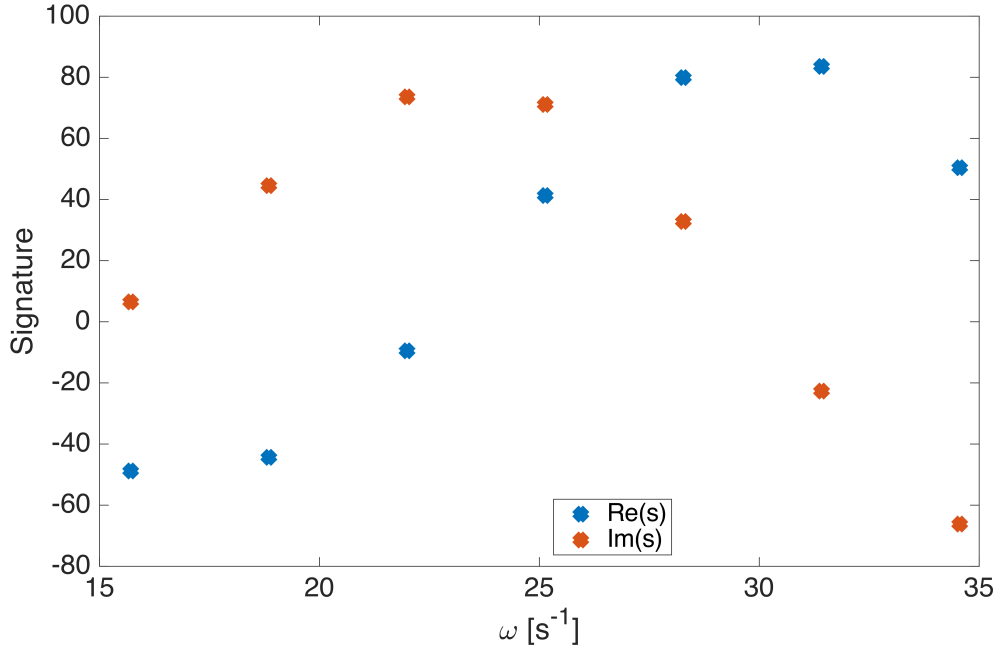


Figure 6.8: Source wavelet for the 2004 BP model experiment.

We use the same algorithms as the Marmousi model. For SE-IR-ESI, the estimation according to van Leeuwen and Herrmann (2016) suggests $\lambda = 1$. For all experiments for SE-IR-ESI, a bounded constraint, whose lower and upper bounds are the same with the ground truth, is employed. For SE-IR-ESI with TV regularizations, three values are used for the size of the TV regularization, α , 10^{-7} , 10^{-6} and 10^{-5} . $\epsilon_{tv} = 10^{-2}$. The frequency continuation strategy is used. For each frequency, a maximum of 50 iterations are performed. The inversion is ceased when the relative change of the total misfit is less than 0.01.

Figure 6.9 presents the final iterates for different inversion experiments. The conventional FWI result in Figure 6.9a completely deviates in the wrong update direction, yielding an unreliable outcome. Notably, the main feature of the salt structure is not recovered at all. Figure 6.9b shows the result of SE-IR-ESI without TV regularization. Figure 6.9c exhibits a similar result to Figure 6.9b, as the magnitude of the TV regularization is small. Among all the inversions, Figure 6.9d achieves the best result, with TV regularization also enhancing the recovery of deeper structures. In contrast, Figure 6.9e fails to recover deep features entirely, as the large TV penalty restricts the number of piece-wise constants.

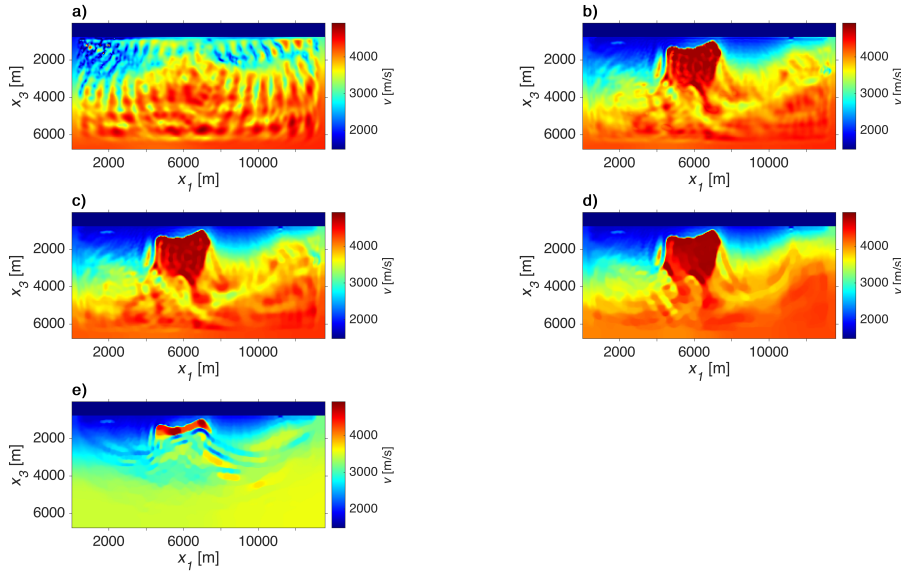


Figure 6.9: Inversion results for the 2004 BP model. (a) is obtained by the conventional FWI. (b), (c), (d) and (e) is computed using SE-IR-ESI without regularizations, with TV regularization $\alpha = 10^{-7}$, with TV regularization $\alpha = 10^{-6}$, and with TV regularization $\alpha = 10^{-5}$. All TV regularizations are assigned with $\epsilon_{tv} = 10^{-2}$. The bounded constraints of (b), (c), (d) and (e) are set the minimum and maximum velocities of the ground truth.

Figure 6.10 illustrates the convergence history for the inversion of the 2004 BP model. Similar to the Marmousi model experiment, conventional FWI reduces the data misfit while increasing the model misfit. Among the SE-IR-ESI experiments, the one with $\alpha = 10^{-6}$ achieves the lowest model misfit and also results in the lowest final physics and data misfits. However, when $\alpha = 10^{-5}$ is applied in TV-SE-IR-ESI, the strong TV regularization causes slow convergence and traps the inversion in a local minimum far from the ground truth.

Figure 6.11 displays the wavefield simulations generated from the final iterates of different inversion methods. Figure 6.11a shows the wavefield simulated from the ground truth, where a longer wavelength is observed in the high-velocity region of the salt. In contrast, Figure 6.11c, which corresponds to the final iterate from conventional FWI, fails to reconstruct the long wavelength component within the

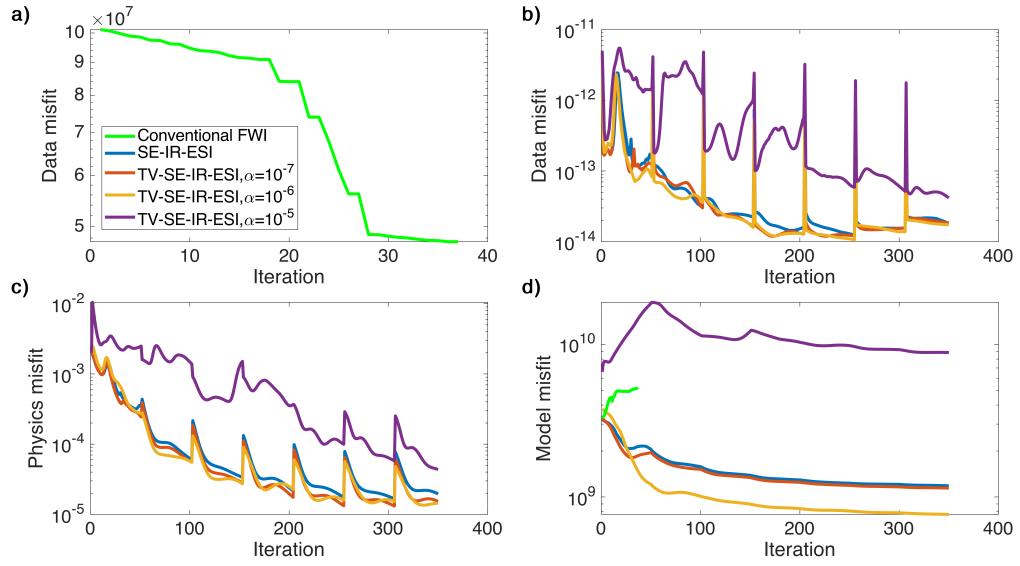


Figure 6.10: Convergence history for the 2004 BP model inversion.

salt. Figures 6.11d, e, and f successfully reconstruct the long wavelength wavefield. However, Figure 6.11g also fails to capture these features as the wrong velocity model. Figure 6.12 demonstrates that the wavefield at the surface, generated by the SE-IR-ESI models, aligns well with the observations. This confirms that SE-IR-ESI is less prone to cycle-skipping issues. In contrast, conventional FWI is highly dependent on the initial iterate.

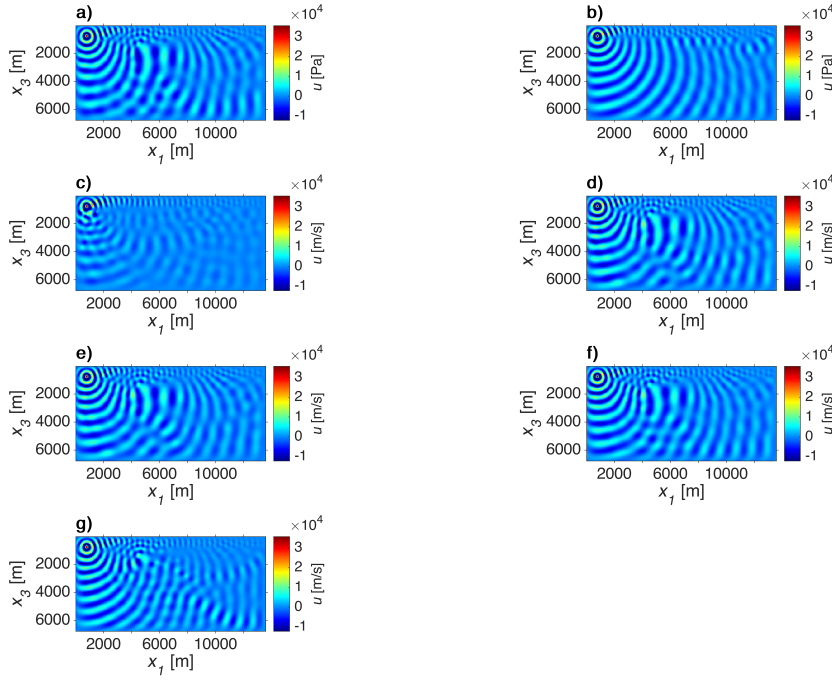


Figure 6.11: Acoustic wavefield simulated for the left-most source, 4 Hz source. (a) 2004 BP ground truth, (b) initial iterate, (c) velocity recovered by conventional FWI, (d) SE-IR-ESI without regularization, (e) SE-IR-ESI with TV regularization $\alpha = 10^{-7}$, (f) SE-IR-ESI with TV regularization $\alpha = 10^{-6}$, and (g) SE-IR-ESI with TV regularization $\alpha = 10^{-5}$.

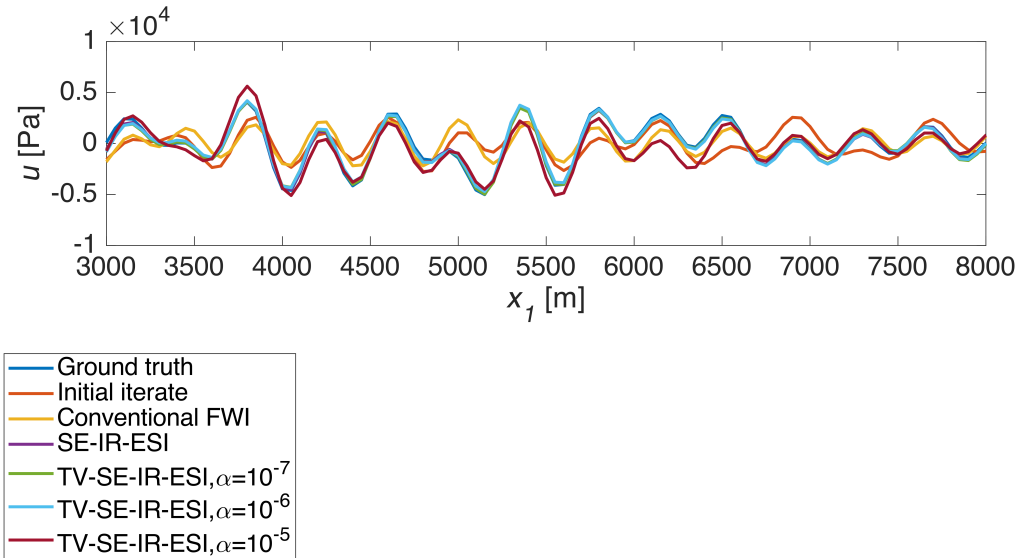


Figure 6.12: Wavefield sampled at receiver depth for Figure 6.11.

Table 6.2: Wavelet estimation for the 2004 BP model.

Experiment	Frequency1	Frequency2	Frequency3	Frequency4	Frequency5	Frequency6	Frequency7
Ground truth	-48.8 + 6.51i	-44.4 + 44.5i	-9.50 + 73.6i	41.3 + 71.1i	79.9 + 32.8i	83.5 - 22.7i	50.4 - 66.2i
SE-IR-ESI	-49.2 + 6.52i	-44.6 + 45.2i	-9.28 + 74.3i	42.2 + 71.4i	80.8 + 32.4i	83.8 - 23.7i	49.8 - 67.1i
TV-SE-IR-ESI, $\alpha = 10^{-7}$	-48.8 + 6.44i	-44.5 + 44.9i	-9.39 + 73.8i	41.7 + 71.2i	80.2 + 32.5i	83.6 - 23.1i	50.2 - 66.4i
TV-SE-IR-ESI, $\alpha = 10^{-6}$	-48.8 + 6.25i	-44.3 + 44.9i	-9.38 + 73.7i	41.7 + 71.0i	80.2 + 32.5i	83.6 - 23.0i	50.2 - 66.5i
TV-SE-IR-ESI, $\alpha = 10^{-5}$	-47.6 + 7.36i	-44.6 + 45.5i	-10.4 + 70.0i	42.6 + 72.0i	80.2 + 33.1i	83.9 - 23.1i	50.6 - 66.5i

6.6 Conclusions

We have developed a new algorithm named SE-IR-ESI, addressing key limitations of conventional approaches. First, our algorithm eliminates the need for prior knowledge of the source signature. Second, it is robust against cycle skipping, allowing for the use of a crude initial model and achieving successful inversion without low-frequency data. This data-driven approach minimizes the need for human intervention, ensuring reliable convergence.

Our TV regularization is implemented using Levenberg-Marquardt linearization combined with Bregman distance, with inner iterations efficiently solved by the primal-dual Newton’s method. Notably, our experiments indicate that only a single inner iteration is required. Unlike split-type methods such as modified total variation (Lin & Huang, 2015a) and split-Bregman (Aghamiry et al., 2019a), our algorithm maintains strong coupling between the model update and the TV denoising problem, enhancing its stability and performance. This allows the algorithm to be applied to real world high contrast media, such as salt structures and volcanoes.

Additionally, our algorithm is general and can be applied to other linear PDE-constrained optimization problems. The only required adaptation is in computing the matrix $\mathbf{G}$ in equation 6.41, which can either be evaluated analytically, as demonstrated in our work, or obtained through automatic differentiation. Since $\mathbf{G}$ is typically sparse (or at least no denser than $\mathbf{A}$), the algorithm remains computationally efficient for other linear PDE applications.

Chapter 7

Conclusions

Identification of subsurface parameters has been an evolving technique in geophysics. Early geophysical parameter estimation required substantial human interaction, but this process was often subjective and prone to inaccuracies. For instance, in the velocity analysis of seismic processing, practitioners had to pick an appropriate normal moveout (NMO) velocity to flatten the common midpoint (CMP) gather (Yilmaz & Doherty, 2001), a task heavily dependent on the experience of the practitioner.

With the development of high-performance computing (HPC), more parameter identification work can now be done through data-driven methods. The first requirement for using these methods is the establishment of a robust forward model, typically a partial differential equation (PDE) derived from theoretical foundations. For example, the acoustic and elastic wave equations are derived from the equation of motion and Hooke's law (Carcione, 2007). The second step involves solving a physics-constrained optimization (inverse problem) to find the optimal values for the parameters. From an HPC perspective, the efficiency of codes for both the forward and inverse problems can also be significantly improved.

In this thesis, the author has explored several possibilities to improve subsurface parameters identification, including both forward models and inversions.

For seismic anisotropy, the author presented a new acoustic TTI wave equation. A key aspect of this work is that the proposed wave equation is obtained by analytical approximation from the dispersion relations (using Taylor series) and this method is not dependent on the model heterogeneity (such as low-rank approximations). Compared to the previous methods, the proposed wave equation demonstrates competitive accuracy in both kinematics and wavefield extrapolators. Applications to RTM demonstrate the necessity of including anisotropy in seismic imaging.

The second contribution of this thesis is the application of TV-LM to geophysical inverse problems. TV-LM is introduced to the eikonal equation problem and benchmarked against several conventional methods, such as L-BFGS applied to the least-squares functional, H^1 regularization, and MTV. Unlike conventional MTV,

where total variation (TV) is implemented using the split-type method, the proposed TV-LM applies Levenberg-Marquardt linearization to the forward problem in the cost function. This makes the TV problem and the model update more closely coupled. The associated first-order optimality condition is efficiently solved using the primal-dual Newton's method. Numerical tests show that TV-LM converges to better minimizers compared to the previous methods and is more computationally efficient. The formulation presented in this work is general and can be applied to other problems where the physics can be linearized. The author also extends TV-LM to the ESI problem, which will be discussed later.

A third major contribution of this thesis is the exploration of extending the search space for FWI. Conventional FWI is solved in the reduced objective for the KKT condition, where the physics is enforced at each iteration. Relaxing the physics constraint enlarges the search space, making the inversion less vulnerable to cycle skipping. Wavefield reconstruction inversion (WRI) achieves this by enlarging the search space. Extended source inversion (ESI) generalizes WRI by using an annihilator as the weighting function for the physics misfit. The author explores the time-domain joint migration inversion (JMI) problem by enlarging the search space. Specifically, JMI is formulated using a velocity-impedance parameterized wave equation. It is shown that velocity captures the low wavenumber component, while impedance absorbs the high wavenumber component. Thus, velocity serves as the background velocity, and impedance is suitable for computing reflectivity. This forward problem is integrated into the source extension for inversion (ES-JMI). As ES-JMI is implemented in the time domain, the reduced objective is employed for efficiency. Benchmark tests demonstrate that ES-JMI is less vulnerable to cycle skipping compared to conventional FWI followed by least-squares migration, JMI with reduced objective for the Lagrangian problem (the conventional JMI in this thesis), and JMI with wavefield reconstruction inversion (WR-JMI).

Finally, the author combines TV-LM and Extended Source Inversion (ESI) to enhance immunity to cycle skipping in Full Waveform Inversion (FWI). The ESI is modified to estimate the source wavelet directly, eliminating the need for a predefined source wavelet input. In this work, the forward problem is solved in the frequency domain. The ESI is reformulated using augmented Lagrangian (AL), which distinguishes the proposed method from conventional FWI and reduced space ESI by generating a sparse Hessian. The new algorithm is named SE-IR-ESI. For the acoustic wave equation, the Hessian is diagonal, making it computationally efficient. By incorporating TV-LM, the model construction promotes sparsity in the model gradient, enabling the reconstruction of models with sharp interfaces.

To conclude, this thesis presents several novel approaches for enhancing subsurface imaging through improved wave equation modeling and advanced inversion techniques. By addressing key challenges such as seismic anisotropy, regularization, and cycle skipping, the proposed methods offer promising improvements in both computational efficiency and inversion accuracy. The introduction of a new acoustic TTI wave equation, the development of the TV-LM regularization method for the eikonal equation, the implementation of ES-JMI and TV-SE-IR-ESI contribute to

more reliable and robust seismic imaging. These advancements not only enable more precise subsurface parameter estimation but also pave the way for future research in geophysical inverse problems. The methodologies presented in this work are poised to make a substantial impact on both theoretical developments and practical applications in geophysics, medical imaging, and nondestructive testing.

Chapter 8

Outlook

While the results of the thesis are promising, some aspects of the proposed methods warrant further investigation.

8.1 Comparisons for different TV methods

The comparison between TV-SE-IR-ESI, where TV-LM is employed, and the split-Bregman method (e.g., Aghamiry et al., 2019a) has not been addressed in this thesis. Within the geophysics community, no direct comparison currently exists between TV-LM and split-type methods, such as MTV and split-Bregman, in the context of waveform inversion problems.

In Chapter 4, the author compares TV-LM and MTV for solving the eikonal equation, highlighting that TV-LM offers advantages such as requiring less tedious tuning and achieving more efficient convergence to a better minimizer.

However, conventional FWI approaches prohibit the use of TV-LM, as the Hessian is dense and lacks sparse approximations. Additionally, the Hessian grows with the number of source-receiver pairs, which becomes a significant issue for large-scale problems. The work in Chapter 6 addresses this by enabling the implementation of TV-LM, as its Hessian is sparse, can be stored in a running sum, and is independent of source-receiver pairings. Thus, a promising direction is to promote TV-LM to a larger community. But this requires a study to show the advantage of TV-LM over MTV and split-Bregman for the waveform inversions.

In Chapter 6, which focuses on extended source waveform inversion, the author conducts preliminary experiments with TV-LM, MTV, and split-Bregman. Due to time constraints, these preliminary results were not included in Chapter 6. The results indicate that TV-LM requires the fewest tuning parameters, while MTV converges more slowly than the other two methods. The split-Bregman method,

though similar to MTV in terms of tuning parameters, is more efficient than TV-LM but generally results in a larger model misfit compared to TV-LM. Indeed, further comparative studies on these methods could be conducted in the future.

Given that current computational resources now enable the use of TV-LM, it holds promise for broader adoption, provided it is rigorously benchmarked against MTV and split-Bregman methods. With this, it has greater potential for recovering high-contrast subsurface media.

8.2 Applications to real data

A holy grail in geophysics is to achieve fully automated imaging. However, this remains a challenging task due to issues with data quality and complex subsurface conditions. The work presented in Chapters 5 and 6 moves us a step closer to automation. Nevertheless, there is still considerable progress to be made.

For instance, the work in Chapter 6 makes the waveform inversion problem independent of the source signature. However, it does not account for surface waves or elastic cases. To implement a truly “full” waveform inversion, all wave types must be included in the forward model, eliminating the need for processes such as top mute on the data.

To achieve this, the author needs to modify the forward problem by incorporating more realistic physics. However, more complex physics introduces cross-talk among parameters, as they often share similar radiation patterns (e.g., Keating & Innanen, 2020). Several approaches have been proposed to mitigate this cross-talk. For example, Diallo and Appel (2000) suggested using additional information, such as vertical seismic profiles and crosswell data, to constrain the inversion. Luo and Wu (2018) proposed separating the sensitivity in the angle domain, while Pan and Innanen (2015) identified parameter trade-offs in the Hessian and introduced a multi-parameter Hessian approximation to suppress these trade-offs in elastic FWI.

That said, it is straightforward to extend the work in Chapters 5 and 6 to more realistic scenarios, such as the elastic case with surface waves, and to integrate strategies to suppress parameter cross-talk. This would further reduce human intervention in the inversion process and improve subsurface imaging quality.

8.3 Medical imaging

Medical imaging and geophysical imaging share a common objective: non-destructive imaging that relies on limited data to produce high-fidelity images. There exists a unified framework between seismic and medical imaging, as both are based on wave physics and can be described using the same objective function. Yin et al. (2020)

proposed a unified framework for photoacoustic imaging, seismic imaging and photoextended imaging.

In addition, some researchers also propose to use acoustic/elastic waves for medical imaging (Lin, Yue, & Zhang, 2020; Riyazifar, Aghamiry, & Gholami, 2024). Conventional imaging methods, such as CT and MRI, have limitations, including reduced sensitivity to early-stage tumors, the use of ionizing radiation in CT, and the high maintenance costs associated with MRI. In contrast, FWI is not affected by these limitations and has shown improved sensitivity to early-stage tumors (X. Li et al., 2025).

The imaging techniques developed in this thesis extend beyond geophysics and hold potential in various fields where imaging plays a critical role. The methods presented in Chapter 3 could be applied to investigate human body anisotropy, while those discussed in Chapter 5 may help study the reflectivity of internal human structures. Additionally, the approach outlined in Chapter 6 offers direct applicability to medical imaging, eliminating the need for source signature estimation. Combined with TV-LM regularization, this technique enables the construction of higher-resolution images, unlocking new possibilities for non-invasive diagnostic methods.

References

- Abubakar, A., Hu, W., Habashy, T. M., & van den Berg, P. M. (2009). Application of the finite-difference contrast-source inversion algorithm to seismic full-waveform data [Journal Article]. *Geophysics*, 74(6), WCC47-WCC58. doi: 10.1190/1.3250203
- Abubakar, A., Hu, W., van den Berg, P. M., & Habashy, T. M. (2008). A finite-difference contrast source inversion method [Journal Article]. *Inverse Problems*, 24(6). doi: 10.1088/0266-5611/24/6/065004
- Acar, R., & Vogel, C. R. (1994). Analysis of bounded variation penalty methods for ill-posed problems [Journal Article]. *Inverse Problems*, 10(6), 1217. Retrieved from <https://dx.doi.org/10.1088/0266-5611/10/6/003> doi: 10.1088/0266-5611/10/6/003
- Aghamiry, H. S., Gholami, A., & 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 [Journal Article]. *Geophysical Journal International*. doi: 10.1093/gji/ggz189
- Aghamiry, H. S., Gholami, A., & Operto, S. (2019b). Improving full-waveform inversion by wavefield reconstruction with the alternating direction method of multipliers [Journal Article]. *GEOPHYSICS*.
- Aghamiry, H. S., Gholami, A., & Operto, S. (2020a). Accurate and efficient data-assimilated wavefield reconstruction in the time domain [Journal Article]. *Geophysics*, 85(2), A7-A12. doi: 10.1190/geo2019-0535.1
- Aghamiry, H. S., Gholami, A., & Operto, S. (2020b). Multi-parameter wavefield reconstruction inversion for wavespeed and attenuation with bound constraints and total variation regularization [Journal Article]. *Geophysics*, 1-54. doi: 10.1190/geo2019-0596.1
- Ajo-Franklin, J. B., Minsley, B. J., & Daley, T. M. (2007). Applying compactness constraints to differential traveltime tomography [Journal Article]. *GEOPHYSICS*, 72(4), R67-R75. Retrieved from <https://doi.org/10.1190/1.2742496> (doi: 10.1190/1.2742496) doi: 10.1190/1.2742496
- Aki, K., & Lee, W. H. K. (1976). Determination of three-dimensional velocity anomalies under a seismic array using first p arrival times from local earthquakes: 1. a homogeneous initial model [Journal Article]. *Journal of Geophysical Research (1896-1977)*, 81(23), 4381-4399. Retrieved from <https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/JB081i023p04381> doi: <https://doi.org/10.1029/JB081i023p04381>

REFERENCES

- Aki, K., & Richards, P. (2002). *Quantitative seismology* [Book]. University Science Books.
- Alkhalifah, T. (1998a). Acoustic approximations for processing in transversely isotropic media [Journal Article]. *Geophysics*, 63(2).
- Alkhalifah, T. (1998b). Acoustic approximations for processing in transversely isotropic media [Journal Article]. *GEOPHYSICS*, 63(2), 623-631. Retrieved from <https://library.seg.org/doi/abs/10.1190/1.1444361> doi: 10.1190/1.1444361
- Alkhalifah, T. (2000). An acoustic wave equation for anisotropic media [Journal Article]. *Geophysics*, 65, 1239-1250.
- Alkhalifah, T. (2002). Traveltime computation with the linearized eikonal equation for anisotropic media [Journal Article]. *Geophysical Prospecting*, 50, 373-382.
- Andersen, K. D., Christiansen, E., Conn, A. R., & Overton, M. L. (2000). An efficient primal-dual interior-point method for minimizing a sum of euclidean norms [Journal Article]. *SIAM Journal on Scientific Computing*, 22(1), 243-262.
- Aravkin, A. Y., & van Leeuwen, T. (2012). Estimating nuisance parameters in inverse problems [Journal Article]. *Inverse Problems*, 28(11). doi: 10.1088/0266-5611/28/11/115016
- Ascher, U. M., Haber, E., & Huang, H. (2006). On effective methods for implicit piecewise smooth surface recovery [Journal Article]. *SIAM Journal on Scientific Computing*, 28(1), 339-358. Retrieved from <https://doi.org/10.1137/040617261> (doi: 10.1137/040617261) doi: 10.1137/040617261
- Bachmayr, M., & Burger, M. (2009). Iterative total variation schemes for nonlinear inverse problems [Journal Article]. *Inverse Problems*, 25(10).
- Bai, J., & Yilmaz, O. (2021). Image-domain least-squares reverse-time migration using point spread functions [Journal Article]. *82nd EAGE Annual Conference & Exhibition, 2021*, 1-5.
- Berkhout, A. J. (2014). Review paper: An outlook on the future of seismic imaging, part iii: Joint migration inversion [Journal Article]. *Geophysical Prospecting*, 62(5), 950-971. doi: 10.1111/1365-2478.12158
- Billette, F., & Brandsberg-Dahl, S. (2005). The 2004 bp velocity benchmark [Journal Article]. *67th EAGE Conference and Exhibition*.
- Biondi, B., Fomel, S., & Alkhalifah, T. (2001). Focusing eikonal equation and global tomography [Journal Article]. *Stanford Exploration Project*.
- Bregman, L. (1967). The relaxation method of finding the common point of convex sets and its application to the solution of problems in convex programming [Journal Article]. *USSR Computational Mathematics and Mathematical Physics*, 7(3), 200-217.
- Brougois, A., Bourget, M., Lailly, P., Poulet, M., Ricarte, P., & Versteeg, R. (1990). Marmousi, model and data [Journal Article]. Retrieved from <https://www.earthdoc.org/content/papers/10.3997/2214-4609.201411190> doi: <https://doi.org/10.3997/2214-4609.201411190>
- Calandra, H., Baina, R., Hanitzsch, C., & Rousseau, J. H. L. (2001). Improving 3d kirchoff prestack depth migration: Why not use regularization and multi-pathing? [Book Section]. In *Seg technical program expanded abstracts*

- 2001 (p. 1049-1052). Retrieved from <https://library.seg.org/doi/abs/10.1190/1.1816261> doi: 10.1190/1.1816261
- Calvert, R. (2005). *Insights and methods for 4d reservoir monitoring and characterization* [Book]. Society of Exploration Geophysicists. Retrieved from <https://doi.org/10.1190/1.9781560801696> doi: 10.1190/1.9781560801696
- Carcione, J. M. (1990). Wave propagation in anisotropic linear viscoelastic media: theory and simulated wave fields [Journal Article]. *Geophysics*.
- Carcione, J. M. (2007). *Wave fields in real media: Wave propagation in anisotropic, anelastic, porous and electromagnetic media* [Book]. Elsevier.
- Carter, J. (2002). *Dual methods for total variation-based image restoration* (Thesis).
- Caspari, E., Pevzner, R., Gurevich, B., Dance, T., Ennis-King, J., Cinar, Y., & Lebedev, M. (2015). Feasibility of co2 plume detection using 4d seismic: Co2erc otway project case study — part 1: Rock-physics modeling [Journal Article]. *GEOPHYSICS*, 80(4), B95-B104. Retrieved from <https://library.seg.org/doi/abs/10.1190/geo2014-0459.1> doi: 10.1190/geo2014-0459.1
- Chambolle, A. (2004). An algorithm for total variation minimization and applications [Journal Article]. *J. Math. Imaging Vision*, 20, 89-97.
- Chan, T. F., Golub, G. H., & Mulet, P. (1996). A nonlinear primal-dual method for total variation-based image restoration [Journal Article]. *SIAM Journal on Scientific Computing*, 20(6).
- Chan, T. F., & Tai, X.-C. (2003). Identification of discontinuous coefficients in elliptic problems using total variation regularization [Journal Article]. *SIAM Journal on Scientific Computing*, 25(3), 881-904. Retrieved from <https://epubs.siam.org/doi/abs/10.1137/S1064827599326020> doi: 10.1137/s1064827599326020
- Chavent, G., & Kunisch, K. (1997). Regularization of linear least squares problems by total bounded variation [Journal Article]. *ESAIM: COCV*, 2, 359-376. Retrieved from <https://doi.org/10.1051/cocv:1997113>
- Cheng, J., Alkhalifah, T., Wu, Z., Zou, P., & Wang, C. (2016). Simulating propagation of decoupled elastic waves using low-rank approximate mixed-domain integral operators for anisotropic media [Journal Article]. *Geophysics*, 81(2), T63-T77.
- Cheng, J., & Fomel, S. (2014). Fast algorithms for elastic-wave-mode separation and vector decomposition using low-rank approximation for anisotropic media [Journal Article]. *Geophysics*, 79(4), C97-C110.
- Choi, Y., Min, D. J., & Shin, C. (2008). Frequency-domain elastic full waveform inversion using the new pseudo-hessian matrix: Experience of elastic marmousi-2 synthetic data [Journal Article]. *Bulletin of the Seismological Society of America*, 98(5), 2402-2415. doi: 10.1785/0120070179
- Chu, C., Macy, B. K., & Anno, P. D. (2011). Approximation of pure acoustic seismic wave propagation in tti media [Journal Article]. *GEOPHYSICS*, 76(5), WB97-WB107.
- De Siena, L., Pezzo, E. D., & Bianco, F. (2011). A scattering image of campi flegrei from the autocorrelation functions of velocity tomograms [Journal Article]. *Geophysical Journal International*, 184(3), 1304-1310. doi: 10.1111/j.1365-246X.2010.04911.x

REFERENCES

- De Siena, L., Thomas, C., & Aster, R. (2014). Multi-scale reasonable attenuation tomography analysis (murat): An imaging algorithm designed for volcanic regions [Journal Article]. *Journal of Volcanology and Geothermal Research*, 277, 22-35. doi: 10.1016/j.jvolgeores.2014.03.009
- Diallo, M. S., & Appel, E. (2000). Acoustic wave propagation in saturated porous media: reformulation of the biot/squirt flow theory [Journal Article]. *Journal of Applied Geophysics*, 44(4), 313-325.
- Douma, H., Yingst, D., Vasconcelos, I., & Tromp, J. (2010). On the connection between artifact filtering in reverse-time migration and adjoint tomography [Journal Article]. *Geophysics*, 75(6), S219-S223. doi: 10.1190/1.3505124
- Du, X., Fletcher, R., Mobley, E., & Yarman, E. (2012). *Source and receiver illumination compensation for reverse-time migration* [Conference Paper]. doi: 10.3997/2214-4609.20148793
- Engquist, B., Froese, B. D., & Yang, Y. (2016). Optimal transport for seismic full waveform inversion [Journal Article]. *Communications in Mathematical Sciences*, 14(8), 2309-2330. doi: 10.4310/CMS.2016.v14.n8.a9
- Engquist, B., & Ying, L. (2009). A fast directional algorithm for high frequency acoustic scattering in two dimensions [Journal Article]. *Commun. Math. Sci.*, 7(2), 327-345.
- ČErvený, V. (1990). *Seismic ray theory* [Book]. Boston, MA: Springer US. Retrieved from https://doi.org/10.1007/0-387-30752-4_134 doi: 10.1007/0-387-30752-4_134
- Esser, E., Guasch, L., Leeuwen, T. v., Aravkin, A. Y., & Herrmann, F. J. (2016). Total-variation regularization strategies in full-waveform inversion [Journal Article]. *arXiv*.
- Fang, Z., Wang, R., & Herrmann, F. J. (2018). Source estimation for wavefield-reconstruction inversion [Journal Article]. *GEOPHYSICS*, 83(4), R345-R359. Retrieved from <https://doi.org/10.1190/geo2017-0700.1> (doi: 10.1190/geo2017-0700.1) doi: 10.1190/geo2017-0700.1
- Fletcher, R. P., Fowler, P. J., Kitchenside, P., & Albertin, U. (2006). Suppressing unwanted internal reflections in prestack reverse-time migration [Journal Article]. *GEOPHYSICS*, 71(6), E79-E82. Retrieved from <https://doi.org/10.1190/1.2356319> (doi: 10.1190/1.2356319) doi: 10.1190/1.2356319
- Fomel, S. (1997). Traveltime computation with the linearized eikonal equation [Journal Article]. *SEP report*, 94, 123-131.
- Fomel, S., Ying, L., & Song, X. (2013). Seismic wave extrapolation using lowrank symbol approximation [Journal Article]. *Geophysical Prospecting*, 61(3), 526-536. doi: 10.1111/j.1365-2478.2012.01064.x
- Fu, L., & Symes, W. W. (2017a). An adaptive multiscale algorithm for efficient extended waveform inversion [Journal Article]. *Geophysics*, 82(3), R183-R197. doi: 10.1190/geo2016-0426.1
- Fu, L., & Symes, W. W. (2017b). A discrepancy-based penalty method for extended waveform inversion [Journal Article]. *Geophysics*, 82(5), R287-R298. doi: 10.1190/geo2016-0326.1
- Gholami, A. (2015). Nonlinear multichannel impedance inversion by total-variation regularization [Journal Article]. *Geophysics*, 80(5), R217-R224. doi: 10.1190/

- geo2015-0004.1
- Gholami, A. (2024). An extended gauss-newton method for full-waveform inversion [Journal Article]. *GEOPHYSICS*, 89(3), R261-R274. Retrieved from <https://library.seg.org/doi/abs/10.1190/geo2022-0673.1> doi: 10.1190/geo2022-0673.1
- Gholami, A., Aghamiry, H., & Operto, S. (2021). *Data reconstruction inversion: An augmented lagrangian based full-waveform inversion in the time domain* [Conference Paper]. doi: 10.3997/2214-4609.202112940
- Gholami, A., Aghamiry, H. S., & Operto, S. (2022). Extended full waveform inversion in the time domain by the augmented lagrangian method [Journal Article]. *Geophysics*, 87(1), R63–R77.
- Gholami, A., & Zabihi Naeini, E. (2019). 3d dix inversion using bound-constrained total variation regularization [Journal Article]. *Geophysics*, 84(3), R311-R320. doi: 10.1190/geo2018-0194.1
- Gill, P. E., & Murray, W. (1978). Algorithms for the solution of the nonlinear least-squares problem [Journal Article]. *SIAM Journal on Numerical Analysis*, 15(5), 977-992.
- Goldstein, T., & Osher, S. (2009). The split bregman method for l1-regularized problems [Journal Article]. *SIAM Journal on Imaging Sciences*, 2(2), 323-343.
- Guasch, L., Calderon Agudo, O., Tang, M. X., Nachev, P., & Warner, M. (2020). Full-waveform inversion imaging of the human brain [Journal Article]. *NPJ Digit Med*, 3, 28. Retrieved from <https://www.ncbi.nlm.nih.gov/pubmed/32195363> (Guasch, Lluís Calderon Agudo, Oscar Tang, Meng-Xing Nachev, Parashkev Warner, Michael eng England 2020/03/21 NPJ Digit Med. 2020 Mar 6;3:28. doi: 10.1038/s41746-020-0240-8. eCollection 2020.) doi: 10.1038/s41746-020-0240-8
- Guo, G., Operto, S., Gholami, A., & Aghamiry, H. S. (2024). Time-domain extended-source full-waveform inversion: Algorithm and practical workflow [Journal Article]. *Geophysics*, 89(2), R73-R94. doi: 10.1190/geo2023-0055.1
- Haber, E., & Ascher, U. (2000). On optimization techniques for solving nonlinear inverse problems [Journal Article]. *Inverse Problems*, 16. doi: 10.1088/0266-5611/16/5/309
- He, B., Liu, H., Wang, Z., & Yuan, X. (2014). A strictly contractive peaceman-rachford splitting method for convex programming [Journal Article]. *SIAM Journal on Optimization*, 24, 1011–1040.
- Henderson, H. V., & Searle, S. R. (1980). On deriving the inverse of a sum of matrices [Journal Article].
- Huang, G., Nammour, R., & Symes, W. W. (2018a). Source-independent extended waveform inversion based on space-time source extension: Frequency-domain implementation [Journal Article]. *Geophysics*, 83(5), R449-R461. doi: 10.1190/geo2017-0333.1
- Huang, G., Nammour, R., & Symes, W. W. (2018b). Volume source-based extended waveform inversion [Journal Article]. *Geophysics*, 83(5), R369-R387. doi: 10.1190/geo2017-0330.1
- Huang, J.-W., & Bellefleur, G. (2012). Joint transmission and reflection traveltime

REFERENCES

- tomography using the fast sweeping method and the adjoint-state technique [Journal Article]. *Geophysical Journal International*, 188(2), 570-582. doi: 10.1111/j.1365-246X.2011.05273.x
- Huang, Y., Ng, M. K., & Wen, Y.-W. (2008). A fast total variation minimization method for image restoration [Journal Article]. *Multiscale Modeling and Simulation*, 7(2), 774-795.
- Irons, T. (1988). *Marmousi model* [Generic]. Institut Français du Pétrole.
- Jain, K. C., & deFigueiredo, R. J. P. (1982). *Concepts and techniques in oil and gas exploration* [Book]. Society of Exploration Geophysicists. Retrieved from <https://doi.org/10.1190/1.9781560802778> doi: 10.1190/1.9781560802778
- Jiang, W., & Zhang, J. (2017). First-arrival traveltime tomography with modified total-variation regularization [Journal Article]. *Geophysical Prospecting*, 65(5), 1138-1154.
- Jiang, W., & Zhang, J. (2018). 3d first-arrival traveltime tomography with modified total variation regularization [Journal Article]. *Journal of Geophysics and Engineering*, 15(1), 207-223.
- Jiménez Tejero, C. E., Dagnino, D., Sallarès, V., & Ranero, C. R. (2015). Comparative study of objective functions to overcome noise and bandwidth limitations in full waveform inversion [Journal Article]. *Geophysical Journal International*, 203(1), 632-645. doi: 10.1093/gji/ggv288
- Jones, L. E. A., & Wang, H. F. (1981). Ultrasonic velocities in cretaceous shales from the williston basin [Journal Article]. *Geophysics*, 46(3), 288-297.
- Julian, B., & Gubbins, D. (1977). Three-dimensional seismic ray tracing [Journal Article]. *GEOPHYSICS*, 43(1), 95-113.
- Kaderli, J., McChesney, M. D., & Minkoff, S. E. (2018). A self-adjoint velocity-stress full-waveform inversion approach to microseismic source estimation [Journal Article]. *Geophysics*, 83(5), 413-427.
- Keating, S., & Innanen, K. A. (2020). Parameter crosstalk and leakage between spatially separated unknowns in viscoelastic full-waveform inversion [Journal Article]. *GEOPHYSICS*, 85(4), R397-R408. Retrieved from <https://doi.org/10.1190/geo2019-0370.1> (doi: 10.1190/geo2019-0370.1) doi: 10.1190/geo2019-0370.1
- Kelly, K. R., Ward, R. W., Treitel, S., & Alford, R. M. (1976). Synthetic seismograms: A finite -difference approach [Journal Article]. *GEOPHYSICS*, 41(1), 2-27. Retrieved from <https://library.seg.org/doi/abs/10.1190/1.1440605> doi: 10.1190/1.1440605
- Lee, K. H., & Kim, H. J. (2003). Source-independent full-waveform inversion of seismic data [Journal Article]. *Geophysics*, 68, 2010-2015.
- Leung, S., & Qian, J. (2006). An adjoint state method for three-dimensional transmission traveltime tomography using first-arrivals [Journal Article]. *COMM. MATH. SCI.*, 4, 249-266.
- Li, B., & Stovas, A. (2021a). Decoupled approximation and separate extrapolation of p- and sv-waves in transversely isotropic media [Journal Article]. *Geophysics*, 86(4), C133-C142. doi: 10.1190/geo2020-0232.1
- Li, B., & Stovas, A. (2021b). Decoupling approximation of p- and s-wave phase

- velocities in orthorhombic media [Journal Article]. *Geophysics*, 87(2), T169-T182. doi: 10.1190/geo2021-0394.1
- Li, Q., Zhou, H., Zhang, Q., Chen, H., & Sheng, S. (2015). Efficient reverse time migration based on fractional laplacian viscoacoustic wave equation [Journal Article]. *Geophysical Journal International*, 204(1), 488-504. doi: 10.1093/gji/ggv456
- Li, X., Cueto, C., Blanco, J. C., Landless, J., Toulemonde, M. E. G., Bates, O. A., ... Agudo, O. C. (2025). Early-stage stomach cancer detection using ultrasound full-waveform inversion: a numerical study [Journal Article]. *SPIE Medical Imaging 2025*.
- Li, Z.-C., Lin, Y.-Z., Zhang, K., Li, Y.-Y., & Yu, Z.-N. (2018). Time-domain wave-field reconstruction inversion [Journal Article]. *Applied Geophysics*, 14(4), 523-528. doi: 10.1007/s11770-017-0629-6
- Lin, Y., & Huang, L. (2015a). Acoustic- and elastic-waveform inversion using a modified total-variation regularization scheme [Journal Article]. *Geophysical Journal International*, 200(1), 489-502. doi: 10.1093/gji/ggu393
- Lin, Y., & Huang, L. (2015b). *Least-squares reverse-time migration with modified total-variation regularization* [Conference Paper]. doi: 10.1190/segam2015-5914034.1
- Lin, Y., Yue, W., & Zhang, Z. (2020). Full-waveform inversion with applications to subsurface characterization and breast cancer detection [Book Section]. In *Seg technical program expanded abstracts 2020* (p. 3497-3501). Society of Exploration Geophysicists. Retrieved from <https://doi.org/10.1190/segam2020-3424462.1> (doi:10.1190/segam2020-3424462.1) doi: doi:10.1190/segam2020-3424462.110.1190/segam2020-3424462.1
- Liu, Q., Beller, S., Lei, W., Peter, D., & Tromp, J. (2000). Preconditioned bfgs-based uncertainty quantification in elastic full waveform inversion [Journal Article]. *Geophys. J. Int.*, 142.
- Liu, Q., & Peter, D. (2018). One-step data-domain least-squares reverse time migration [Journal Article]. *Geophysics*, 83(4), R361-R368. doi: 10.1190/geo2017-0622.1
- Luo, J., & Wu, R.-S. (2018). Velocity and density reconstruction based on scattering angle separation [Journal Article]. *Pure and Applied Geophysics*, 175(12), 4371-4387. doi: 10.1007/s00024-018-1916-8
- Ma, Y., & Hale, D. (2013). Wave-equation reflection traveltime inversion with dynamic warping and full-waveform inversion [Journal Article]. *GEOPHYSICS*, 78(6), R223-R233. Retrieved from <https://doi.org/10.1190/geo2013-0004.1> (doi: 10.1190/geo2013-0004.1) doi: 10.1190/geo2013-0004.1
- Métivier, L., Brossier, R., Virieux, J., & Operto, S. (2013). Full waveform inversion and the truncated newton method [Journal Article]. *SIAM Journal on Scientific Computing*, 35(2), B401-B437.
- Mu, X., Huang, J., Yang, J., Guo, X., & Guo, Y. (2020). Least-squares reverse time migration in tti media using a pure qp-wave equation [Journal Article]. *Geophysics*, 85(4), S199-S216. doi: 10.1190/geo2019-0320.1
- Mulder, W. A., & Plessix, R. E. (2004). A comparison between one-way and two-way wave-equation migration [Journal Article]. *GEOPHYSICS*, 69(6),

REFERENCES

- 1491-1504. Retrieved from <https://doi.org/10.1190/1.1836822> (doi: 10.1190/1.1836822) doi: 10.1190/1.1836822
- Nocedal, J., & Wright, S. J. (2006). *Numerical optimization* [Book]. New York, NY, USA: Springer.
- Osher, S., Burger, M., Goldfarb, D., Xu, J., & Yin, W. (2005). An iterative regularization method for total variation-based image restoration [Journal Article]. *Multiscale Modeling and Simulation*, 4(2), 460-489. doi: 10.1137/040605412
- Pan, W., & Innanen, K. (2015). Suppressing cross-talk for elastic fwi with multi-parameter approximate hessian and its parameter-type approximation [Journal Article]. *CREWES Research Report*.
- Pan, W., & Innanen, K. (2016). Elastic full-waveform inversion: density effects, cycle-skipping, and inter-parameter mapping [Journal Article]. *CREWES Research Report*, 28.
- Parra, J. O. (1997). The transversely isotropic poroelastic wave equation including the biot and the squirt mechanisms: Theory and application [Journal Article]. *Geophysics*, 62, 309-318.
- Pedersen, r., Ursin, B., & Helgesen, H.-K. (2010). One-way wave-equation migration of compressional and converted waves in a vti medium [Journal Article]. *GEOPHYSICS*, 75(6), S237-S248. Retrieved from <https://library.seg.org/doi/abs/10.1190/1.3509466> doi: 10.1190/1.3509466
- Pestana, R. C., Ursin, B., & Stoffa, P. L. (2012). Rapid expansion and pseudo spectral implementation for reverse time migration in vti media [Journal Article]. *Journal of Geophysics and Engineering*, 9(3), 291-301. doi: 10.1088/1742-2132/9/3/291
- Pevzner, R., Caspari, E., Gurevich, B., Dance, T., & Cinar, Y. (2015). Feasibility of co2 plume detection using 4d seismic: Co2crc otway project case study — part 2: Detectability analysis [Journal Article]. *GEOPHYSICS*, 80(4), B105-B114. Retrieved from <https://library.seg.org/doi/abs/10.1190/geo2014-0460.1> doi: 10.1190/geo2014-0460.1
- Plessix, R. E. (2006). A review of the adjoint-state method for computing the gradient of a functional with geophysical applications [Journal Article]. *Geophysical Journal International*, 167(2), 495-503. doi: 10.1111/j.1365-246X.2006.02978.x
- Plessix, R. E., & Mulder, W. A. (2004). Frequency-domain finite-difference amplitude-preserving migration [Journal Article]. *Geophysical Journal International*, 157(3), 975-987. doi: 10.1111/j.1365-246X.2004.02282.x
- Poplavko, Y. M. (2019). Chapter 2 - mechanical properties of solids [Book Section]. In Y. M. Poplavko (Ed.), *Electronic materials* (p. 71-93). Elsevier. Retrieved from <http://www.sciencedirect.com/science/article/pii/B9780128157800000025> doi: <https://doi.org/10.1016/B978-0-12-815780-0.00002-5>
- Pratt, R. G. (1999). Seismic waveform inversion in the frequency domain, part 1: Theory and verification in a physical scale model [Journal Article]. *GEOPHYSICS*, 64(3), 888-901. Retrieved from <https://library.seg.org/doi/abs/10.1190/1.1444597> doi: 10.1190/1.1444597
- Ramos-Martinez, J., Crawley, S., Zou, K., Valenciano, A., Qiu, L., Chemingui, N.,

- & Long, A. (2016). A robust gradient for long wavelength fwi updates [Journal Article]. *ASEG Extended Abstracts*.
- Ramos-Martínez, J., Qiu, L., Kirkebo, J., Valenciano, A. A., & Yang, Y. (2018). *Long-wavelength fwi updates beyond cycle skipping* [Conference Paper]. doi: 10.1190/segam2018-2998433.1
- Rawlinson, N., Kool, M. d., & Sambridge, M. (2018). Seismic wavefront tracking in 3d heterogeneous media: applications with multiple data classes [Journal Article]. *Exploration Geophysics*, 37(4), 322-330. doi: 10.1071/eg06322
- Rawlinson, N., & Sambridge, M. (2003). Seismic traveltime tomography of the crust and lithosphere [Journal Article]. *Advances in geophysics*, 46, 81-199.
- Rawlinson, N., & Sambridge, M. (2018). The fast marching method: An effective tool for tomographic imaging and tracking multiple phases in complex layered media [Journal Article]. *Exploration Geophysics*, 36(4), 341-350. doi: 10.1071/eg05341
- Riyazifar, Y., Aghamiry, H. S., & Gholami, A. (2024, 26-28 June 2024). *Breast cancer imaging by source-extended full-waveform inversion* [Conference Paper]. doi: 10.1109/MeMeA60663.2024.10596772
- Rizzuti, G., Louboutin, M., Wang, R., Daskalakis, E., & Herrmann, F. (2019). *A dual formulation for time-domain wavefield reconstruction inversion* [Conference Paper]. doi: 10.1190/segam2019-3216760.1
- Robein, E. (2010). *Seismic imaging: A review of the techniques, their principles, merits and limitations* [Book]. EAGE Publications. Retrieved from <https://books.google.no/books?id=QbDySAAACAAJ>
- Robertsson, J. O. A., Blanch, J. O., & Symes, W. W. (1994). Viscoelastic finite-difference modeling [Journal Article]. *Geophysics*, 59(9), 1444-1456. doi: 10.1190/1.1443701
- Rossato, D., Gutierrez, F. D., Guarneri, G. A., Passarin, T. A. R., Pires, G. P., & Pipa, D. R. (2023). Full waveform inversion for ndt using ultrasonic linear arrays [Journal Article]. *Research and Review Journal of Nondestructive Testing*, 1(1). doi: 10.58286/28117
- Rudin, L. I., Osher, S., & Fatemi, E. (1992). Nonlinear total variation based noise removal algorithms [Journal Article]. *Physica D: Nonlinear Phenomena*, 60(1-4), 259-268.
- Rueger, A. (1993). *Dynamic ray tracing and its application in triangulated media* (Thesis).
- Schaeken, T., Hoogerbrugge, L., & Verschuur, E. (2022). A non-reflecting wave equation through directional wave-field suppression and its finite difference implementation [Journal Article]. *Sci Rep*, 12(1), 407. Retrieved from <https://www.ncbi.nlm.nih.gov/pubmed/35013407> (Schaeken, Teun Hoogerbrugge, Leo Verschuur, Eric eng England 2022/01/12 Sci Rep. 2022 Jan 10;12(1):407. doi: 10.1038/s41598-021-04064-3.) doi: 10.1038/s41598-021-04064-3
- Shearer, P. (2009). *Introduction to seismology* [Book]. Cambridge University Press.
- Sibley, E. C. P., Madgett, A. S., Lawrence, J. M., Elsdon, T. S., Marnane, M. J., Fernandes, P. G., & Demer, D. (2023). Quantifying the ability of imaging sonar to identify fish species at a subtropical artificial reef [Journal Article]. *ICES Journal of Marine Science*. doi: 10.1093/icesjms/fsad156

REFERENCES

- Song, X., & Alkhalifah, T. (2013). Modeling of pseudoacoustic p-waves in orthorhombic media with a low-rank approximation [Journal Article]. *Geophysics*, 78(4), 33-40.
- Stovas, A. (2017). Kinematic parameters of pure- and converted-mode waves in elastic orthorhombic media [Journal Article]. *Geophysical Prospecting*, 65(2), 426-452. doi: 10.1111/1365-2478.12420
- Stovas, A. (2020). On parameterization in monoclinic media with a horizontal symmetry plane [Journal Article]. *Geophysics*, 1-58. doi: 10.1190/geo2020-0301.1
- Stovas, A., Alkhalifah, T., & Waheed, U. b. (2020). Pure p- and s-wave equations in transversely isotropic media [Journal Article]. *Geophysical Prospecting*, 68(9), 2762-2769. doi: 10.1111/1365-2478.13026
- Symes, W. W. (2008). Migration velocity analysis and waveform inversion [Journal Article]. *Geophysical Prospecting*, 56(6), 765-790. doi: 10.1111/j.1365-2478.2008.00698.x
- Taillandier, C., Noble, M., Chauris, H., & Calandra, H. (2009). First-arrival traveltimes tomography based on the adjoint-state method [Journal Article]. *Geophysics*, 74(6), 1-10.
- Tarantola, A. (1984a). Inversion of seismic reflection data in the acoustic approximation [Journal Article]. *Geophysics*, 49(8), 1259-1266. doi: 10.1190/1.1441754
- Tarantola, A. (1984b). Linearized inversion of seismic reflection data [Journal Article]. *Geophysical prospecting*.
- Tarantola, A., & Valette, B. (1982). Generalized nonlinear inverse problems solved using the least squares criterion [Journal Article]. *Reviews of Geophysics*, 20(2), 219-232. Retrieved from <https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/RG020i002p00219> doi: <https://doi.org/10.1029/RG020i002p00219>
- Thomsen, L. (1986). Weak elastic anisotropy [Journal Article]. *Geophysics*, 51(10), 1954-1966.
- Thomsen, L. (2002). *Understanding seismic anisotropy in exploration and exploitation* (2nd ed.) [Book]. Society of Exploration Geophysicists.
- Tong, P. (2021). Adjoint-state traveltimes tomography: Eikonal equation-based methods and application to the anza area in southern california [Journal Article]. *Journal of Geophysical Research: Solid Earth*, 126(5). doi: 10.1029/2021jb021818
- Tromp, J., Tape, C., & Liu, Q. (2004). Seismic tomography, adjoint methods, time reversal and banana-doughnut kernels [Journal Article]. *Geophysical Journal International*, 160(1), 195-216. doi: 10.1111/j.1365-246X.2004.02453.x
- Tsvankin, I. (2012). *Seismic signatures and analysis of reflection data in anisotropic media* [Book]. United States of America: Society of Exploration Geophysicists.
- Tsvankin, I., Gaiser, J., Grechka, V., van der Baan, M., & Thomsen, L. (2010). Seismic anisotropy in exploration and reservoir characterization: An overview [Journal Article]. *GEOPHYSICS*, 75(5), 75A15-75A29. Retrieved from <https://doi.org/10.1190/1.3481775> (doi: 10.1190/1.3481775) doi: 10.1190/1.3481775
- Tsvankin, I., & Thomsen, L. (1994). Nonhyperbolic reflection moveout in anisotropic

- media [Journal Article]. *GEOPHYSICS*, 59(8), 1290-1304. Retrieved from <https://library.seg.org/doi/abs/10.1190/1.1443686> doi: 10.1190/1.1443686
- van Leeuwen, T., & Herrmann, F. J. (2013). Mitigating local minima in full-waveform inversion by expanding the search space [Journal Article]. *Geophysical Journal International*, 195(1), 661-667. doi: 10.1093/gji/ggt258
- van Leeuwen, T., & Herrmann, F. J. (2016). A penalty method for pde-constrained optimization in inverse problems [Journal Article]. *Inverse Problems*, 32(1). doi: 10.1088/0266-5611/32/1/015007
- Verschuur, D. J., Staal, X. R., & Berkhout, A. J. (2016). Joint migration inversion: Simultaneous determination of velocity fields and depth images using all orders of scattering [Journal Article]. *The Leading Edge*, 35(12), 1037-1046. doi: 10.1190/tle35121037.1
- Verschuur, E., Qu, S., Kim, Y. S., & Sun, Y. (2020). Joint migration inversion: features and challenges [Journal Article]. *Journal of Geophysics and Engineering*, 17(3), 525-538. doi: 10.1093/jge/gxaa012
- Virieux, J., & Operto, S. (2009). An overview of full-waveform inversion in exploration geophysics [Journal Article]. *Geophysics*, 74(6), WCC1-WCC26. doi: 10.1190/1.3238367
- Vogel, C. R. (2002). *Computational methods for inverse problems* [Book].
- Vogel, C. R., & Oman, M. (1996). Iterative methods for total variation denoising [Journal Article]. *SIAM Journal on Scientific Computing*, 17(1), 227-238.
- Waheed, U. b., Alkhalifah, T., & Wang, H. (2015). Efficient traveltime solutions of the acoustic ti eikonal equation [Journal Article]. *Journal of Computational Physics*, 282, 62-76. Retrieved from <https://www.sciencedirect.com/science/article/pii/S002199911400758X> doi: <https://doi.org/10.1016/j.jcp.2014.11.006>
- Waheed, U. B., Yarman, C. E., & Flagg, G. (2014). An iterative, fast-sweeping-based eikonal solver for 3d tilted anisotropic media [Journal Article]. *GEOPHYSICS*, 80(3), 49-58.
- Wang, C., Yingst, D., Farmer, P., & Leveille, J. (2016). *Full-waveform inversion with the reconstructed wavefield method* [Conference Paper]. doi: 10.1190/segam2016-13870317.1
- Wang, M., Xie, Y., Xu, W. Q., Loh, F. C., Xin, K., Chuah, B. L., ... Wolfarth, S. (2016). Dynamic-warping full-waveform inversion to overcome cycle skipping [Journal Article]. *SEG Technical Program Expanded Abstracts 2016*.
- Wang, Y., Yang, J., Yin, W., & Zhang, Y. (2008). A new alternating minimization algorithm for total variation image reconstruction [Journal Article]. *SIAM Journal on Imaging Sciences*, 1(3), 248-272. Retrieved from <https://epubs.siam.org/doi/abs/10.1137/080724265> doi: 10.1137/080724265
- Warner, M., & Guasch, L. (2016). Adaptive waveform inversion: Theory [Journal Article]. *GEOPHYSICS*, 81(6), R429-R445. Retrieved from <https://doi.org/10.1190/geo2015-0387.1> (doi: 10.1190/geo2015-0387.1) doi: 10.1190/geo2015-0387.1
- Wei, Z., Mei, J., Wu, Z., Zhang, Z., Huang, R., & Wang, P. (2023). Pushing seismic resolution to the limit with fwi imaging [Journal Article]. *The Leading Edge*,

- 42(1), 24-32. doi: 10.1190/tle42010024.1
- Whitmore, N., Ramos-Martinez, J., Yang, Y., & Valenciano, A. (2020). Full wave-field modeling with vector reflectivity [Journal Article]. *82nd EAGE Annual Conference & Exhibition, 2020*.
- Wu, Z., Alkhalifah, T., & Zhang, Z. (2019). A partial-low-rank method for solving acoustic wave equation [Journal Article]. *Journal of Computational Physics*, 385, 1-12. doi: 10.1016/j.jcp.2019.01.054
- Xu, S., Stovas, A., Alkhalifah, T., & Mikada, H. (2020). New acoustic approximation for transversely isotropic media with a vertical symmetry axis [Journal Article]. *Geophysics*, 85(1), C1-C12. doi: 10.1190/geo2019-0100.1
- Xu, S., & Zhou, H. (2014). Accurate simulations of pure quasi-p-waves in complex anisotropic media [Journal Article]. *Geophysics*, 79(6), T341-T348.
- Yang, D., & Zhang, Z. (2002). Poroelastic wave equation including the biot/squirt mechanism and the solid/fluid coupling anisotropy [Journal Article]. *Wave Motion*, 35, 223-245.
- Yang, J., & Zhu, H. (2018). Viscoacoustic reverse time migration using a time-domain complex-valued wave equation [Journal Article]. *Geophysics*, 1-92.
- Yang, P., & Zhou, W. (2024). Algorithmic analysis towards time-domain extended source waveform inversion [Journal Article]. *Pure and Applied Geophysics*. doi: 10.1007/s00024-024-03556-3
- Yang, Y., Ramos-Martinez, J., Whitmore, D., Huang, G., & Chemingui, N. (2021a). *Simultaneous inversion of velocity and reflectivity* [Conference Paper]. doi: 10.1190/segam2021-3594405.1
- Yang, Y., Ramos-Martinez, J., Whitmore, D., Huang, G., & Chemingui, N. (2021b). Simultaneous velocity and reflectivity inversion: Fwi + lsrtm [Journal Article]. *82nd EAGE Annual Conference & Exhibition, 2021*, 1-5.
- Yang, Y., Ramos-Martinez, J., Whitmore, D., Valenciano, A., & Chemingui, N. (2020). Full waveform inversion using wave equation reflectivity modeling [Journal Article]. *EAGE 2020 Annual Conference & Exhibition, 2020*.
- Yao, G., Silva, N. V. d., Warner, M., Wu, D., & Yang, C. (2019). Tackling cycle skipping in full-waveform inversion with intermediate data [Journal Article]. *GEOPHYSICS*, 84(3), R411-R427.
- Yilmaz, z., & Doherty, S. M. (2001). *Seismic data analysis* (Vol. 1) [Book]. Science.
- Yin, Z., Orozco, R., Witte, P., Louboutin, M., Rizzuti, G., & Herrmann, F. J. (2020). Extended source imaging ? a unifying framework for seismic and medical imaging [Book Section]. In *Seg technical program expanded abstracts 2020* (p. 3502-3506). Society of Exploration Geophysicists. Retrieved from <https://doi.org/10.1190/segam2020-3426999.1> (doi:10.1190/segam2020-3426999.1) doi: doi:10.1190/segam2020-3426999.110.1190/segam2020-3426999.1
- Yoon, K., Guo, M., Cai, J., & Wang, B. (2011). 3d rtm angle gathers from source wave propagation direction and dip of reflector [Book Section]. In *Seg technical program expanded abstracts* (p. 3136-3140). Retrieved from <https://library.seg.org/doi/abs/10.1190/1.3627847> doi: 10.1190/1.3627847
- Zand, T., Ghasemzadeh, H., Gholami, A., & Malcolm, A. (2023). Least-squares reverse time migration with shifted total variation regularization [Journal Article]. *GEOPHYSICS*, 88(2), 59-69.

- Zhan, G., Pestana, R. C., & Stoffa, P. L. (2013). An efficient hybrid pseudospectral/finite-difference scheme for solving the tti pure p-wave equation [Journal Article]. *Journal of Geophysics and Engineering*, 10(2). doi: 10.1088/1742-2132/10/2/025004
- Zhang, L., & Ohen, E. (2022). A stable tti acoustic wave equation [Journal Article]. *Second International Meeting for Applied Geoscience and Energy*, 2527-2531. doi: 10.1190/image2022-3747266.1
- Zhang, X., Burger, M., Bresson, X., & Osher, S. (2010). Bregmanized nonlocal regularization for deconvolution and sparse reconstruction [Journal Article]. *SIAM Journal on Imaging Sciences*, 3(3), 253-276. doi: 10.1137/090746379
- Zhang, Y., Zhang, H., & Zhang, G. (2011). A stable tti reverse time migration and its implementation [Journal Article]. *GEOPHYSICS*, 76(3), WA3-WA11.
- Zhao, H. (2004). A fast sweeping method for eikonal equations [Journal Article]. *Mathematics of computation*, 74(250), 603-627.
- Zhou, B., & Greenhalgh, S. A. (2003). Crosshole seismic inversion with normalized full-waveform amplitude data [Journal Article]. *Geophysics*, 68, 1320-1330.
- Zhou, H., Zhang, G., & Bloor, R. (2006). An anisotropic acoustic wave equation for modeling and migration in 2d tti media [Journal Article]. *SEG Technical Program Expanded Abstracts 2006*, 194-198.
- Zhou, W., Brossier, R., Operto, S., & Virieux, J. (2015). Full waveform inversion of diving and reflected waves for velocity model building with impedance inversion based on scale separation [Journal Article]. *Geophysical Journal International*, 202(3), 1535-1554. doi: 10.1093/gji/ggv228
- Zhou, Z.-Z., Howard, M., & Mifflin, C. (2011). Use of rtm full 3d subsurface angle gathers for subsalt velocity update and image optimization: Case study at shenzi field [Journal Article]. *GEOPHYSICS*, 76(5), WB27-WB39. Retrieved from <https://library.seg.org/doi/abs/10.1190/geo2011-0065.1> doi: 10.1190/geo2011-0065.1

REFERENCES

List of figures

1.1	A schematic of the seismic experiment. To determine the properties of the subsurface medium, a seismic source is activated at the surface, generating waves that propagate through the subsurface. These waves are recorded by receivers at the surface, providing data that can be used to compute the subsurface parameters.	1
1.2	A synthetic subsurface acoustic model v defined in a 2D space $[x_1 \ x_3]^T$ created by Irons (1988) – Marmousi model.	2
1.3	Simulations based on the model in Figure 1.2. The source is placed at $[300, 281.25]^T$ m. 292 receivers are evenly placed at $x_1 \in [300, 16669]$ m, $x_3 = 281.25$ m. (a),(b) and (c) are the acoustic wavefields at different time steps, simulated by the acoustic wave equation. (d) is the travel time obtained by the eikonal equation.	3
1.4	Observed data sampled for the numerical experiment at Figure 1.3, (a) for the wavefield and (b) for the travel time.	4
1.5	A schematic illustrating cycle skipping in Full Waveform Inversion (FWI) (Virieux & Operto, 2009). Cycle skipping is a primary cause of non-uniqueness in FWI. The upper dashed line represents the simulated data, which deviates by more than half of the period, $T/2$, from the recorded data (shown by the middle solid line). In this case, FWI mismatches cycle n to cycle $n - 1$. The bottom dashed line represents another simulated data set, which deviates by less than $T/2$ from the observed data. Here, FWI successfully matches the two signals. Cycle skipping is affected by the initial model and frequency band.	5
3.1	The P-wave kinematics for Greenhorn shale. $A_{11} = 1.447 \times 10^7$ Pa m ³ /kg, $A_{33} = 9.57 \times 10^6$ Pa m ³ /kg, $A_{13} = 4.51 \times 10^6$ Pa m ³ /kg, $A_{44} = 2.28 \times 10^6$ Pa m ³ /kg, thus, $\eta = 0.3409$. (a) Exact P-wave phase velocity. (b) The relative error of the estimated P-wave phase velocity E_r	36

- 3.2 The average stacked relative error of the estimated P-wave phase velocity for randomly generated 1000 VTI samples using a uniform distribution. The stacking is performed for each phase angle. $A_{11} \sim U_{\text{dis}}(1.1 \times 10^7 \text{ Pa m}^3/\text{kg}, 1.7 \times 10^7 \text{ Pa m}^3/\text{kg})$, $A_{33} \sim U_{\text{dis}}(7 \times 10^6 \text{ Pa m}^3/\text{kg}, 1.25 \times 10^7 \text{ Pa m}^3/\text{kg})$, $A_{13} \sim U_{\text{dis}}(1.5 \times 10^6 \text{ Pa m}^3/\text{kg}, 3 \times 10^6 \text{ Pa m}^3/\text{kg})$, $A_{44} \sim U_{\text{dis}}(3.5 \times 10^6 \text{ Pa m}^3/\text{kg}, 4.5 \times 10^6 \text{ Pa m}^3/\text{kg})$. U_{dis} is for uniform distribution. Thus $\eta \in [-0.2475, 0.7100]$ in our experiment. . . . 36
- 3.3 The model utilized in the numerical experiment for comparing the wavefield extrapolator of the proposed method with the low-rank approach. The model is only a function of x_3 37
- 3.4 The acoustic wavefield extrapolators H_p for the model given in Figure 3.3. (a) - (c) are for the location $x_h = 30 \text{ m}$, $x_3 = 100 \text{ m}$. (d) - (f) are for the location $x_h = 30 \text{ m}$, $x_3 = 500 \text{ m}$. (g) - (i) are for the location $x_h = 30 \text{ m}$, $x_3 = 900 \text{ m}$. (a), (d), and (g) are the true solutions. (b), (e), and (h) are the relative errors of the low-rank approach, solved with 4 ranks for x and k , respectively. (c), (f), and (i) are the relative errors of our proposed solutions. 38
- 3.5 The acoustic wavefield for the Greenhorn shale at 0.5 s. The red down-pointing triangle depicted in (a) serves as the source and will retain the same marker throughout. The source signature is Ricker wavelet, 10 Hz. (a) Exact solution. (b) VTI wavefield by Alkhalifah (2000). (c) VTI wavefield by L. Zhang and Ohen (2022). (d) Proposed VTI solution. (e) Proposed TI solution with $\theta_{\text{ti}} = \frac{\pi}{6}$ 39
- 3.6 The profile at $x_h = 1600 \text{ m}$ for Figure 3.5 a-d. 40
- 3.7 The wedge model. (a) η . (b) θ_{ti} . The other parameters are set homogeneous along the space. $A_{11} = 1.4 \times 10^7 \text{ Pa m}^3/\text{kg}$, $A_{33} = 1.0 \times 10^7 \text{ Pa m}^3/\text{kg}$, $A_{44} = 2.0 \times 10^6 \text{ Pa m}^3/\text{kg}$. 30 receivers are placed evenly at $x_h \in [780, 5130] \text{ m}$ and $x_3 = 750 \text{ m}$ 41
- 3.8 The acoustic wavefield for the model in Figure 3.7. The source signature is a Ricker wavelet, 10 Hz. Acoustic approximations (a)(d)(g) are at 0.5 s, (b)(e)(h) at 1 s and (c)(f)(i) at 1.8 s. (a) - (c) are computed by the method of L. Zhang and Ohen (2022) and (d) - (f) are from our proposed method. (g) - (i) are exact solutions. 42
- 3.9 The shot gathers for the wavefield in Figure 3.8 and the setup in Figure 3.7, (a) for L. Zhang and Ohen (2022), (b) for the proposed method, (c) for the exact solution, and (d) differences between the approximations and the true solution for the receiver index 25. 43
- 3.10 The checkerboard model. For white region, $A_{11} = 1.4 \times 10^7 \text{ Pa m}^3/\text{kg}$, $A_{33} = 1.0 \times 10^7 \text{ Pa m}^3/\text{kg}$, $A_{44} = 7 \times 10^6 \text{ Pa m}^3/\text{kg}$, $\eta = 0.2$, $\theta_{\text{ti}} = \frac{\pi}{3}$. For the black region, the values of the above parameters are doubled. 44

3.11	The shot gathers for Figure 3.10. Ricker wavelet, 10 Hz. u_p is the solution when 3 sources are shot simultaneously, and $u_{p,sta}$ is the stacked solution for each shot wavefield, (a) u_p by our proposed method, (b) $u_{p,sta}$ by our proposed method, (c) True seismograms for receiver index 15, (d) u_p by L. Zhang and Ohen (2022), (e) $u_{p,sta}$ by L. Zhang and Ohen (2022), and (f) $u_{p,sta} - u_p$ for L. Zhang and Ohen (2022) and the proposed method.	44
3.12	The modified Hess model. (a) A_{11} . (b) A_{33} . (c) η . (d) An additional tilt angle θ_{ti} of the symmetry axis is added to the original Hess model. 201 receivers are placed evenly at $x_h \in [1281, 25681]$ m and $x_3 = 1342$ m.	45
3.13	The reflectivity images $I_{R,3}$. Source: Ricker wavelet, 5 Hz. 17 sources are placed evenly at $x_h \in [1281, 20862]$ m, $x_3 = 1342$ m, and each shot gather is generated using a single source. (a) is the image derived from the vertical differentiation of A_{33} . (b)(c) and (d) are for TTI, VTI, and elliptic migration images, respectively. In (c) and (d), certain artifacts (indicated by red arrows) are generated, and some reflectors are inaccurately constructed (shown in blue) due to inappropriate physics.	46
3.14	The modified 3D EAGE/SEG Salt model. (a) Slice at $x_1 = 4400$ m. (b) Slice at $x_2 = 4400$ m. (c) Slice at $x_3 = 2800$ m. The down-pointing triangles are the imaging source locations projected on those planes. 421875 receivers are evenly distributed in this range, $x_1 \in [840, 6760]$ m, $x_2 \in [840, 6760]$ m, $x_3 = 840$ m.	47
3.15	The solution of the wavefield by the proposed method. The source is placed at $[4440 \ 4440 \ 840]^T$ m. Source: Ricker wavelet, 7 Hz. The three rows correspond to $t = 0.3, 0.4, 0.6$ s. The three columns correspond to the same slices in Figure 3.14 for (a), (b), and (c).	48
3.16	The RTM image for the slice in Figure 3.14, same slices as in Figure 3.14 for (a), (b), and (c).	49
4.1	Background model for the perturbation test. Down pointing triangles – sources. Up pointing triangles – receivers. The same marker convention is consistently applied throughout the work.	64
4.2	Adjoint travel time λ_t for the source and 10th receiver in the model in Figure 4.1.	65
4.3	Exact δt and approximated δt using the adjoint-state equation method. The velocity perturbation is $\delta v = 0.1v$ in Figure 4.1.	65
4.4	(a) The true toy model. (b) The starting model for inversion. The model is discretized into 210 (x_1) by 30 (x_3) grids.	66

LIST OF FIGURES

4.5	Inversions for the model in Figure 4.4 by (a) L-BFGS, (b) H^1 -regularization, (c) MTV, (d) TV-LM with reference parameters, (e) TV-LM with large ϵ_{tv} , (f) TV-LM with large α	67
4.6	Profile at $x_3 = 100$ m for Figure 4.5.	67
4.7	Inversion performance for the toy model. (a) Data misfit. (b) Bregman distance, $D_{TV_{\epsilon_{tv}}}(v_t, v_i)$. $\epsilon_{tv} = 10^{-2}$ for this and subsequent evaluation of Bregman distance in the inversion performance analysis. (c) L^2 -norm model misfit.	68
4.8	(a) Marmousi imodel. (b) Starting model for the inversion. The model is discretized into 340 by 60 grids. The maximum offset for the source-receiver pair is set at 10000 m.	69
4.9	Inversions for the Marmousi model by (a) L-BFGS, (b) H^1 , (c) MTV and (d) TV-LM. We use the adjoint travel time to identify the region where the inversion is trusted, as in Figure 4.2. The area above the black line is the trusted region. The model misfit and the Bregman distance are only evaluated for the trusted region. This applies to the following experiments in this work.	70
4.10	Inversion performance for the Marmousi model.	71
4.11	(a) 2004 BP model. (b) The initial model for the inversion. The model is discretized into 220 by 46 grids. The maximum offset for the source-receiver pair is set at 43200 m.	72
4.12	Noise-free inversions for the 2004 BP model by (a) L-BFGS, (b) H^1 , (c) MTV and (d) TV-LM.	73
4.13	Inversion performance for the 2004 BP model under noise-free conditions.	74
4.14	Inversions of noisy data for the 2004 BP model by (a) L-BFGS, (b) H^1 , (c) MTV and (d) TV-LM.	75
4.15	Inversion performance for the 2004 BP model under noisy conditions.	76
5.1	Band-pass filter for the second term in the gradient computation.	91
5.2	Velocity (a), impedance (b), and vertical reflectivity (c) ground truth for the Marmousi.	92
5.3	Initial iterate for the inversion of velocity (a), impedance (b) and vertical reflectivity (c) of the Marmousi.	93
5.4	Wavelets utilized in the inversion of Marmousi.	93

5.5	Update direction for the first iteration of Wavelet 1. The update direction for v is shown in the first column, while the update direction for Z is displayed in the second column. The four rows correspond to conventional FWI, JMI, WR-JMI, and ES-JMI, respectively. The conventional FWI updates only v in the velocity build-up so that the update direction for Z is missing.	94
5.6	Inversion results for Marmousi. The first column represents v , the second column represents Z , and the third column shows the vertical reflectivity $I_{R,3}$. The four rows correspond to conventional FWI, JMI, WR-JMI, and ES-JMI, respectively. Z in conventional FWI is derived at Step 2 of the conventional FWI process.	96
5.7	Convergence history of the different methods for the Marmousi inversions.	97
5.8	Seismograms for the first source from the left. The seismograms are stacked for Wavelets 1 and 2. (a) Observed data. (b) Data from initial iterate. (d)(f)(h)(j) is data by conventional FWI, JMI, WR-JMI, and ES-JMI, respectively. The second row is the subtraction of the first row by the observed data in (a).	98
5.9	Velocity (a), impedance (b) and vertical reflectivity (c) ground truth for 2004 BP.	99
5.10	Initial iterate for the inversion of velocity (a), impedance (b), and vertical reflectivity (c) for 2004 BP.	99
5.11	Wavelets utilized in the inversion of 2004 BP.	100
5.12	Same as Figure 5.5 but for 2004 BP model.	100
5.13	Same column arrangements as in Figure 5.6 but for 2004 BP model. First row is for WR-JMI and the second is for ES-JMI.	101
5.14	Convergence histories of the different methods for the 2004 BP inversion.	102
5.15	Seismograms for the first source from the left. The seismograms are stacked for Wavelets 1 and 2. (a) Observed data. (b) Data from initial iterate. (d)(f) are the data from the final iterate by conventional WR-JMI and ES-JMI. The second row is the subtraction of the first row by the observed data in (a).	103
6.1	Marmousi model. (a) Ground truth. (b) Initial iterate.	118
6.2	Source wavelet for the Marmousi model experiment.	118

6.3	Inversion results for the Marmousi model. (a) is obtained by the conventional FWI. (b), (c), (d) and (e) is computed using SE-IR-ESI without regularizations, with TV reuglarization $\alpha = 10^{-8}$, with TV reuglarization $\alpha = 10^{-7}$, and with TV reuglarization $\alpha = 10^{-6}$. All TV regularizations are assigned with $\epsilon_{tv} = 10^{-2}$. The bounded constraints of (b), (c), (d) and (e) are set the minimum and maximum velocities of the ground truth.	119
6.4	Convergence history for the Marmousi model inversion.	120
6.5	Acoustic wavefield simulated for the left-most source, 9 Hz source. (a) Marmousi ground truth, (b) initial iterate, (c) velocity recovered by conventional FWI, (d) SE-IR-ESI without regularization, (e) SE-IR-ESI with TV reuglarization $\alpha = 10^{-8}$, (f) SE-IR-ESI with TV reuglarization $\alpha = 10^{-7}$, and (g) SE-IR-ESI with TV reuglarization $\alpha = 10^{-6}$	121
6.6	Wavefield sampled at receiver depth for Figure 6.5.	121
6.7	2004 BP model. (a) Ground truth. (b) Initial iterate.	123
6.8	Source wavelet for the 2004 BP model experiment.	123
6.9	Inversion results for the 2004 BP model. (a) is obtained by the conventional FWI. (b), (c), (d) and (e) is computed using SE-IR-ESI without regularizations, with TV reuglarization $\alpha = 10^{-7}$, with TV reuglarization $\alpha = 10^{-6}$, and with TV reuglarization $\alpha = 10^{-5}$. All TV regularizations are assigned with $\epsilon_{tv} = 10^{-2}$. The bounded constraints of (b), (c), (d) and (e) are set the minimum and maximum velocities of the ground truth.	124
6.10	Convergence history for the 2004 BP model inversion.	125
6.11	Acoustic wavefield simulated for the left-most source, 4 Hz source. (a) 2004 BP ground truth, (b) initial iterate, (c) velocity recovered by conventional FWI, (d) SE-IR-ESI without regularization, (e) SE-IR-ESI with TV reuglarization $\alpha = 10^{-7}$, (f) SE-IR-ESI with TV reuglarization $\alpha = 10^{-6}$, and (g) SE-IR-ESI with TV reuglarization $\alpha = 10^{-5}$	126
6.12	Wavefield sampled at receiver depth for Figure 6.11.	126

List of tables

4.1	Tuning parameters in the inversion for the toy model.	68
4.2	Tuning parameters in the inversion for the Marmousi model.	69
4.3	Tuning parameters in the noise-free inversion for the 2004 BP model.	71
4.4	Tuning parameters in the noisy inversion for the 2004 BP model. The iteration number shown in this table is the maximum possible iteration. The progress is stopped if the method stagnates or the data misfit reaches the noise level.	72
5.1	Wave properties of the second term in the gradient computation.	91
6.1	Wavelet estimation for the Marmousi model.	122
6.2	Wavelet estimation for the 2004 BP model.	127
A1	Variable correspondence for the problem $(\mathbf{A} + \mathbf{UBV})^{-1}$	170

List of acronyms

AL Augmented Lagrangian	10
ES-JMI Joint Migration Inversion by Extended Source waveform inversion	9
ESI Extended Source waveform Inversion	7
FWI Full Waveform Inversion	9
IR-ESI Iteratively Refined Extended Source waveform Inversion	111
LS-RTM Least-Squares Reverse Time Migration	10
JMI Joint Migration Inversion	9
KKT Karush-Kuhn-Tucker	6
MTV Modified Total Variation	8
PML Perfectly Matched Layers	92
qP Quasi P	15
qSH Quasi SH	15
qSV Quasi SV	15

SE-IR-ESI Iteratively Refined Extended Source waveform Inversion with Source Estimation	105
TTI Tilted Transversely Isotropic	8
TV Total Variation	8
TV-LM TV with Levenberg-Marquardt-type linearization	9
TV-SE-IR-ESI Iteratively Refined Extended Source waveform Inversion with Source Estimation and Total Variation	108
VTI Vertically Transversely Isotropic	14
WRI Wavefield Reconstruction Inversion	6

List of symbols

t	Time
x	Space coordinate
N_{dim}	Dimension of the space
u_{d}	Displacement vector of elastic waves
v	Phase velocity
v_{p}	P-wave phase velocity
v_{s}	S-wave phase velocity
c	4th-order Elastic stiffness tensor
C	Elastic stiffness matrix
u	Wavefield
u_{a}	Adjoint wavefield
θ_{ti}	Counterclockwise rotation angle of the TTI symmetry axis of the TI media relative to the positive x_3 -axis
ϕ_{ti}	Azimuth angle of the TTI symmetry axis relative to positive x_2 -axis, counterclockwise
A	Left-hand-side operator of the wave equation
b_{obs}	Right-hand-side of the wave equation (source term)
s	Source signature
$\text{TV}(\cdot)$	TV norm
$\text{TV}_{\epsilon_{\text{tv}}}(\cdot)$	Smoothed TV norm
ϵ_{tv}	A small positive number for the smoothed TV norm to avoid singularity
Z	Impedance
I_{R}	Reflectivity vector

Throughout the thesis, the author uses boldface notation to denote the discrete counterparts.

Appendix

A Coefficients for acoustic TTI wave equation

Coefficients $Q_i, i \in \{1, 2, \dots, 9\}$ in equations 3.21 and 3.22 are given by,

$$\begin{aligned}
Q_1 &= \frac{128}{q_3^2} (11 q_3 q_4 - 3 q_2 q_4 - 5 q_1 q_4 - \sigma_{q37} + \sigma_{34} + \sigma_{q33} + \sigma_{q31} \\
&\quad + \sigma_{q30} - \sigma_{q26} + \sigma_{q36} - \sigma_{q28} - \sigma_{q35} + \sigma_{q32}) , \\
Q_2 &= \frac{128}{q_3^2} (\sigma_{q17} - \sigma_3 - \sigma_{q16} - \sigma_2 - \sigma_{q10} + \sigma_1 - \sigma_{q15} + \sigma_9 + \sigma_{q14} - \sigma_{q13}) , \\
Q_3 &= \frac{128}{q_3^2} (44 q_3 q_4 - 12 q_2 q_4 - 20 q_1 q_4 - \sigma_{q29} - \sigma_{q25} + \sigma_{q27} - \sigma_{q24} - \sigma_{q19} \\
&\quad + \sigma_{q20} - \sigma_{q23} + \sigma_{q18} + \sigma_{q22} - \sigma_{q21}) , \\
Q_4 &= \frac{128}{q_3^2} (\sigma_{q12} - \sigma_3 - \sigma_{q11} - \sigma_2 + \sigma_8 + \sigma_1 + \sigma_7 - \sigma_6 - \sigma_5 + \sigma_4) , \\
Q_5 &= \frac{128}{q_3^2} (66 q_3 q_4 - 18 q_2 q_4 - 30 q_1 q_4 - 40 q_1 q_4 \cos(4\theta_{ti}) \\
&\quad - 40 q_2 q_4 \cos(4\theta_{ti}) + 120 q_3 q_4 \cos(4\theta_{ti}) \\
&\quad + 70 q_1 q_4 \cos(8\theta_{ti}) - 70 q_2 q_4 \cos(8\theta_{ti}) + 70 q_3 q_4 \cos(8\theta_{ti})) , \\
Q_6 &= \frac{128}{q_3^2} (\sigma_{q12} + \sigma_3 - \sigma_{q11} + \sigma_2 + \sigma_8 - \sigma_1 - \sigma_7 - \sigma_6 + \sigma_5 - \sigma_4) , \\
Q_7 &= \frac{128}{q_3^2} (44 q_3 q_4 - 12 q_2 q_4 - 20 q_1 q_4 + \sigma_{q29} - \sigma_{q25} - \sigma_{q27} - \sigma_{q24} \\
&\quad + \sigma_{q19} + \sigma_{q20} - \sigma_{q23} - \sigma_{q18} + \sigma_{q22} - \sigma_{q21}) , \\
Q_8 &= \frac{128}{q_3^2} (\sigma_{q17} + \sigma_3 - \sigma_{q16} + \sigma_2 - \sigma_{q10} - \sigma_1 + \sigma_{q15} + \sigma_9 - \sigma_{q14} + \sigma_{q13}) , \\
Q_9 &= \frac{128}{q_3^2} (11 q_3 q_4 - 3 q_2 q_4 - 5 q_1 q_4 + \sigma_{q37} + \sigma_{q34} - \sigma_{q33} + \sigma_{q31} - \sigma_{q30} \\
&\quad - \sigma_{q26} + \sigma_{q36} + \sigma_{q28} - \sigma_{q35} + \sigma_{q32}) ,
\end{aligned} \tag{A1}$$

where $\sigma_{qi}, i \in \{1, 2, \dots, 37\}$ are given by,

$$\begin{aligned}
\sigma_{q1} &= 48 q_3 q_4 \sin(4 \theta_{ti}), \sigma_{q2} = 16 q_2 q_4 \sin(4 \theta_{ti}), \\
\sigma_{q3} &= 16 q_1 q_4 \sin(4 \theta_{ti}), \sigma_{q4} = 56 q_3 q_4 \sin(8 \theta_{ti}), \\
\sigma_{q5} &= 56 q_2 q_4 \sin(8 \theta_{ti}), \sigma_{q6} = 56 q_3 q_4 \sin(6 \theta_{ti}), \\
\sigma_{q7} &= 56 q_1 q_4 \sin(8 \theta_{ti}), \sigma_{q8} = 56 q_1 q_4 \sin(6 \theta_{ti}), \\
\sigma_{q9} &= 24 q_3 q_4 \sin(6 \theta_{ti}), \sigma_{q10} = 24 q_1 q_4 \sin(6 \theta_{ti}), \\
\sigma_{q11} &= 24 q_3 q_4 \sin(2 \theta_{ti}), \sigma_{q12} = 24 q_1 q_4 \sin(2 \theta_{ti}), \\
\sigma_{q13} &= 8 q_3 q_4 \sin(8 \theta_{ti}), \sigma_{q14} = 8 q_2 q_4 \sin(8 \theta_{ti}), \\
\sigma_{q15} &= 8 q_1 q_4 \sin(8 \theta_{ti}), \sigma_{q16} = 8 q_3 q_4 \sin(2 \theta_{ti}), \\
\sigma_{q17} &= 8 q_1 q_4 \sin(2 \theta_{ti}), \sigma_{q18} = 56 q_3 q_4 \cos(6 \theta_{ti}), \\
\sigma_{q19} &= 56 q_1 q_4 \cos(6 \theta_{ti}), \sigma_{q20} = 48 q_3 q_4 \cos(4 \theta_{ti}), \\
\sigma_{q21} &= 28 q_3 q_4 \cos(8 \theta_{ti}), \sigma_{q22} = 28 q_2 q_4 \cos(8 \theta_{ti}), \\
\sigma_{q23} &= 28 q_1 q_4 \cos(8 \theta_{ti}), \sigma_{q24} = 16 q_2 q_4 \cos(4 \theta_{ti}), \\
\sigma_{q25} &= 16 q_1 q_4 \cos(4 \theta_{ti}), \sigma_{q26} = 12 q_3 q_4 \cos(4 \theta_{ti}), \\
\sigma_{q27} &= 8 q_3 q_4 \cos(2 \theta_{ti}), \sigma_{q28} = 4 q_3 q_4 \cos(6 \theta_{ti}), \\
\sigma_{q29} &= 8 q_1 q_4 \cos(2 \theta_{ti}), \sigma_{q30} = 4 q_1 q_4 \cos(6 \theta_{ti}), \\
\sigma_{q31} &= 4 q_2 q_4 \cos(4 \theta_{ti}), \sigma_{q32} = q_3 q_4 \cos(8 \theta_{ti}), \\
\sigma_{q33} &= 4 q_3 q_4 \cos(2 \theta_{ti}), \sigma_{q34} = 4 q_1 q_4 \cos(4 \theta_{ti}), \\
\sigma_{q35} &= q_2 q_4 \cos(8 \theta_{ti}), \sigma_{q36} = q_1 q_4 \cos(8 \theta_{ti}), \\
\sigma_{q37} &= 4 q_1 q_4 \cos(2 \theta_{ti}).
\end{aligned} \tag{A2}$$

B An adjoint approach for the linearized eikonal equation

To find the effect of the model perturbation on the travel time, we perturb the eikonal equation 4.19 by setting $v = v + \delta v$ and $t = t + \delta t$,

$$\nabla(t + \delta t) \cdot \nabla(t + \delta t) = \frac{1}{(v + \delta v)^2}. \tag{A3}$$

Expanding the left-hand side and applying the Taylor series up to first-order accuracy, one obtains

$$|\nabla t|^2 + 2\nabla t \cdot \nabla \delta t + |\nabla \delta t|^2 \approx \frac{1}{v^2} - 2\frac{\delta v}{v^3}. \tag{A4}$$

We assume that $|\delta t| \ll |t|$ so that $|\nabla \delta t|^2$ can be ignored in equation (A4). Subtracting this from equation (4.19), the equation becomes,

$$\nabla t \cdot \nabla \delta t = -\frac{\delta v}{v^3}. \tag{A5}$$

We consider a perturbation $\delta t(x_r)$ where $x_r \notin \partial\Omega$ is a receiver location. Then we integrate equation (A5) over Ω and apply integrate by parts,

$$\begin{aligned}\delta t(r) &= \delta t(r) + \int_{\Omega} \lambda_t (\nabla t \cdot \nabla \delta t + \frac{\delta v}{v^3}) dx \\ &= \delta t(r) + \int_{\partial\Omega} \delta t \lambda_t \nabla t \cdot n dx + \int_{\Omega} (-\delta t \nabla \cdot (\lambda_t \nabla t)) dx + \int_{\Omega} \frac{\lambda_t \delta v}{v^3} dx ,\end{aligned}\tag{A6}$$

We assume vanishing boundary condition for λ_t , i.e.,

$$\lambda_t(x) = 0, x \in \partial\Omega ,\tag{A7}$$

Now we can choose

$$\nabla \cdot (\lambda_t \nabla t) = -\delta(x - x_r), x \in \Omega ,\tag{A8}$$

where $\delta(\cdot)$ is Dirac delta, and

$$\delta t(x_r) = \int_{\Omega} \frac{\lambda_t \delta v}{v^3} dx .\tag{A9}$$

The variable λ_t is the adjoint-state variable, and equations (A7) and (A8) are the adjoint-state equations. To evaluate $t'(v)$, one needs to compute λ_t using equations (A7) and (A8) first and then insert it to equation (A9) for $t'(v)$.

C Adjoint wave equation

To solve equation 5.12, we rewrite equation 5.9 in the continuous form and apply integrate-by-parts for both time and space derivatives,

$$\begin{aligned}
 E_{\text{lag}}(v, Z, u) &= \frac{1}{2} \int_0^T \int_{\Omega_r} (u(x, t) - u_{\text{obs}}(x, t))^2 dx dt \\
 &\quad - \int_0^T \int_{\Omega} u_a(x, t) \left(\frac{1}{v(x)Z(x)} \frac{\partial^2 u(x, t)}{\partial t^2} - \nabla \cdot \left(\frac{v(x)}{Z(x)} \nabla u(x, t) \right) \right) dx dt \\
 &= \frac{1}{2} \int_0^T \int_{\Omega_r} (u - u_{\text{obs}})^2 dx dt \\
 &\quad - \left(\int_{\Omega} \frac{1}{vZ} \frac{\partial u}{\partial t} u_a \Big|_0^T dx - \int_{\Omega} \frac{1}{vZ} \frac{\partial u_a}{\partial t} u \Big|_0^T dx + \int_0^T \int_{\Omega} \frac{1}{vZ} u \frac{\partial^2 u_a}{\partial t^2} dx dt \right. \\
 &\quad - \int_0^T \int_{\partial\Omega} u_a \left(\frac{v}{Z} \nabla u \right) \cdot n_b dx \\
 &\quad + \int_0^T \int_{\partial\Omega} u \left(\frac{v}{Z} \nabla u_a \right) \cdot n_b dx dt \\
 &\quad \left. - \int_0^T \int_{\Omega} u \nabla \cdot \left(\frac{v}{Z} \nabla u_a \right) dx dt \right), \tag{A10}
 \end{aligned}$$

where n_b is the outward unit normal vector for the boundary of Ω ; u_{obs} is the observed data; Ω_r is the receiver location.

We can further assume

$$u_a(x, T) = 0, x \in \Omega, \tag{A11}$$

$$\frac{\partial u_a(x, T)}{\partial t} = 0, x \in \Omega, \tag{A12}$$

$$u_a(x, t) = 0, x \in \partial\Omega \text{ and } 0 \leq t \leq T, \tag{A13}$$

$$\nabla u_a(x, t) = 0, x \in \partial\Omega \text{ and } 0 \leq t \leq T. \tag{A14}$$

With boundary conditions of in equations 5.5 and A11 - A14, equation A10 can be simplified

$$\begin{aligned}
 E_{\text{lag}}(v, Z, u, u_a) &= \frac{1}{2} \int_0^T \int_{\Omega_r} (u - u_{\text{obs}})^2 dx dt \\
 &\quad - \int_0^T \int_{\Omega} \left(\frac{1}{vZ} u \frac{\partial^2 u_a}{\partial t^2} - u \nabla \cdot \left(\frac{v}{Z} \nabla u_a \right) \right) dx dt, \tag{A15}
 \end{aligned}$$

Take the variation of E_{lag} with respect to u ,

$$\begin{aligned} \delta E_{\text{lag}} = & \int_0^T \int_{\Omega_r} (u - u_{\text{obs}}) \delta u \, dx \, dt \\ & - \int_0^T \int_{\Omega} \left(\int_{\Omega} \frac{1}{vZ} \frac{\partial^2 u_a}{\partial t^2} \delta u - \nabla \cdot \left(\frac{v}{Z} \nabla u_a \right) \delta u \right) dx \, dt , \end{aligned} \quad (\text{A16})$$

According to equation 5.12, we can formulate the adjoint wave equation,

$$\begin{aligned} \frac{1}{v(x)Z(x)} \frac{\partial^2 u_a(x, t)}{\partial t^2} - \nabla \cdot \left(\frac{v(x)}{Z(x)} \cdot u_a(x, t) \right) &= u(x, t) - u_{\text{obs}}(x, t), \quad x \in \Omega_r , \\ \frac{1}{v(x)Z(x)} \frac{\partial^2 u_a(x, t)}{\partial t^2} - \nabla \cdot \left(\frac{v(x)}{Z(x)} \cdot u_a(x, t) \right) &= 0, \quad x \notin \Omega_r . \end{aligned} \quad (\text{A17})$$

According to equations A11 and A12, u_a has the final condition in time. So, equation A17 is run backward in time.

D Velocity and impedance gradients for the Lagrangian

For equation 5.8, we consider a small perturbation δv that leads to a perturbation of δE_{lag} ,

$$E_{\text{lag}}(\mathbf{v} + \delta \mathbf{v}, \mathbf{Z}, \mathbf{u}, \mathbf{u}_a) = \frac{1}{2} \|\mathbf{R}\mathbf{u} - \mathbf{d}_{\text{obs}}\|_2^2 - \mathbf{u}_a^T (\mathbf{A}(\mathbf{v} + \delta \mathbf{v}, \mathbf{Z})\mathbf{u} - \mathbf{b}) . \quad (\text{A18})$$

Taking the subtraction of equation A18 from equation 5.8,

$$\delta E_{\text{lag}} = -\mathbf{u}_a^T ((\mathbf{A}(\mathbf{v} + \delta \mathbf{v}, \mathbf{Z}) - (\mathbf{A}(\mathbf{v}, \mathbf{Z}))\mathbf{u} - \mathbf{b}) . \quad (\text{A19})$$

In a continuous form, equation A19 can recast as,

$$\begin{aligned} \delta E_{\text{lag}} = & \int_0^T \int_{\Omega} \frac{\delta v(x)}{v(x)^2 Z(x)} \frac{\partial^2 u(x, t)}{\partial t^2} u_a(x, t) \\ & + \nabla \frac{\delta v(x)}{Z(x)} \cdot \nabla u(x, t) u_a(x, t) \\ & + \frac{\delta v(x)}{Z(x)} \nabla^2 u(x, t) u_a(x, t) dx \, dt . \end{aligned} \quad (\text{A20})$$

For the second term, we apply integrate-by-parts for space derivatives and assume

$u(x, t) = 0$ and $u_a(x, t) = 0$ for $x \in \partial\Omega$,

$$\begin{aligned}
 \int_{\Omega} \nabla \frac{\delta v(x)}{Z(x)} \cdot \nabla u(x, t) u_a(x, t) dx &= \int_{\Omega} (u(x, t) \nabla u(x, t)) \cdot \nabla dx \frac{\delta v(x)}{\delta Z(x)} dx \\
 &= \int_{\Omega} -\frac{\delta v(x)}{Z(x)} \nabla \cdot (u_a(x, t) \nabla u(x, t)) dx \\
 &= \int_{\Omega} -\frac{\delta v(x)}{Z(x)} \nabla u(x, t) \cdot \nabla u_a(x, t) \\
 &\quad - \frac{\delta v(x)}{Z(x)} \nabla^2 u(x, t) u_a(x, t) dx .
 \end{aligned} \tag{A21}$$

Thus, equation A18 now becomes,

$$\delta E_{\text{lag}} = \int_0^T \int_{\Omega} \frac{\delta v(x)}{v(x)^2 Z(x)} \frac{\partial^2 u(x, t)}{\partial t^2} u_a(x, t) - \frac{\delta v(x)}{Z(x)} \nabla u(x, t) \cdot \nabla u_a(x, t) dx dt . \tag{A22}$$

Equation A22 describes how a small perturbation $\delta v(x)$ affects δE_{lag} . Similarly, we can derive δE_{lag} caused by $\delta Z(x)$,

$$\delta E_{\text{lag}} = \int_{-\infty}^{\infty} \int_{\Omega} \frac{\delta Z(x)}{v(x) Z(x)^2} \frac{\partial^2 u(x, t)}{\partial t^2} u_a(x, t) + \frac{\delta v(x) v(x)}{Z(x)^2} \nabla u(x, t) \cdot \nabla u_a(x, t) dx dt \tag{A23}$$

Reformulating equations A22 and A23 in the gradient form yields,

$$\begin{aligned}
 \frac{\partial E_{\text{lag}}}{\partial v(x)} &= \int_0^T \frac{1}{v(x)^2 Z(x)} c_1(x, t) - \frac{1}{Z(x)} c_2(x, t) dt , \\
 \frac{\partial E_{\text{lag}}}{\partial Z(x)} &= \int_0^T \frac{1}{v(x) Z(x)^2} c_1(x, t) + \frac{v(x)}{Z(x)^2} c_2(x, t) dt ,
 \end{aligned} \tag{A24}$$

where

$$c_1(x, t) = \frac{\partial^2 u(x, t)}{\partial t^2} u_a(x, t) , \tag{A25}$$

$$c_2(x, t) = \nabla u(x, t) \cdot \nabla u_a(x, t) . \tag{A26}$$

E From data augmented wavefield to source extension

Let us left multiply $(\mathbf{A}^{-1})^T$ on both sides of equation 5.21,

$$((\mathbf{A}^{-1})^T \mathbf{R}^T \mathbf{R} + \lambda^2 \mathbf{C}_A^T \mathbf{C}_A \mathbf{A}) \mathbf{u} = (\mathbf{A}^{-1})^T \mathbf{R}^T \mathbf{d}_{\text{obs}} + \lambda^2 \mathbf{C}_A^T \mathbf{C}_A \mathbf{b} . \tag{A27}$$

Equation A27 is equivalent to

$$\begin{aligned} & (\mathbf{R}\mathbf{A}^{-1})^T \mathbf{R}\mathbf{A}^{-1} \mathbf{A}\mathbf{u} + \lambda^2 \mathbf{C}_A^T \mathbf{C}_A \mathbf{A}\mathbf{u} \\ &= (\mathbf{R}\mathbf{A}^{-1})^T \mathbf{d}_{\text{obs}} + \lambda^2 \mathbf{C}_A^T \mathbf{C}_A \mathbf{b} + (\mathbf{R}\mathbf{A}^{-1})^T \mathbf{R}\mathbf{A}^{-1} \mathbf{b} - (\mathbf{R}\mathbf{A}^{-1})^T \mathbf{R}\mathbf{A}^{-1} \mathbf{b} . \end{aligned} \quad (\text{A28})$$

Combining some terms in equation A28 yields,

$$(\mathbf{R}\mathbf{A}^{-1})^T \mathbf{R}\mathbf{A}^{-1} (\mathbf{A}\mathbf{u} - \mathbf{b}) + \lambda^2 \mathbf{C}_A^T \mathbf{C}_A (\mathbf{A}\mathbf{u} - \mathbf{b}) = (\mathbf{R}\mathbf{A}^{-1})^T (\mathbf{d}_{\text{obs}} - \mathbf{R}\mathbf{A}^{-1} \mathbf{b}) . \quad (\text{A29})$$

Using the settings in equations 5.22 - 5.24, A29 can be simplified to

$$\left(\frac{1}{\lambda^2} (\mathbf{R}\mathbf{A}^{-1})^T \mathbf{R}\mathbf{A}^{-1} + \mathbf{C}_A^T \mathbf{C}_A \right) \Delta \mathbf{b} = (\mathbf{R}\mathbf{A}^{-1})^T \Delta \mathbf{d} , \quad (\text{A30})$$

which is just equation 5.25.

F Receiver side Hessian

In this section, we show how to obtain the derivation in equation 5.27. To achieve it, we need to utilize some inference of $(\mathbf{A} + \mathbf{UBV})^{-1}$ by Henderson and Searle (1980),

$$\begin{aligned} & \left(\frac{1}{\lambda^2} \mathbf{S}^T \mathbf{S} + \mathbf{C}_A^T \mathbf{C}_A \right)^{-1} \mathbf{S}^T \\ &= \left((\mathbf{C}_A^T \mathbf{C}_A)^{-1} - (\mathbf{C}_A^T \mathbf{C}_A)^{-1} \mathbf{S}^T \left(\mathbb{I} + \frac{1}{\lambda^2} \mathbf{S} (\mathbf{C}_A^T \mathbf{C}_A)^{-1} \mathbf{S}^T \right)^{-1} \frac{1}{\lambda^2} \mathbf{S} (\mathbf{C}_A^T \mathbf{C}_A)^{-1} \right) \mathbf{S}^T \end{aligned} \quad (\text{A31})$$

$$= (\mathbf{C}_A^T \mathbf{C}_A)^{-1} \mathbf{S}^T - (\mathbf{C}_A^T \mathbf{C}_A)^{-1} \mathbf{S}^T \left(\mathbb{I} + \frac{1}{\lambda^2} \mathbf{S} (\mathbf{C}_A^T \mathbf{C}_A)^{-1} \mathbf{S}^T \right)^{-1} \frac{1}{\lambda^2} \mathbf{S} (\mathbf{C}_A^T \mathbf{C}_A)^{-1} \mathbf{S}^T \quad (\text{A32})$$

$$= (\mathbf{C}_A^T \mathbf{C}_A)^{-1} \mathbf{S}^T \left(\mathbb{I} - \left(\mathbb{I} + \frac{1}{\lambda^2} \mathbf{S} (\mathbf{C}_A^T \mathbf{C}_A)^{-1} \mathbf{S}^T \right)^{-1} \frac{1}{\lambda^2} \mathbf{S} (\mathbf{C}_A^T \mathbf{C}_A)^{-1} \mathbf{S}^T \right) \quad (\text{A33})$$

$$= (\mathbf{C}_A^T \mathbf{C}_A)^{-1} \mathbf{S}^T \left(\mathbb{I} + \frac{1}{\lambda^2} \mathbf{S} (\mathbf{C}_A^T \mathbf{C}_A)^{-1} \mathbf{S}^T \right)^{-1} . \quad (\text{A34})$$

Equation A32 and A34 employ equation 23 in Henderson and Searle (1980) with the variable correspondence in Table A1.

Table A1: Variable correspondence for the problem $(\mathbf{A} + \mathbf{UBV})^{-1}$.

Equation number in our study	Variable in our study	Variable in Henderson and Searle (1980)	Equation number in Henderson and Searle (1980)
$\frac{1}{\lambda^2}\mathbb{I}$	A32	$\mathbf{B}$	23
$\mathbf{S}^T$	A32	$\mathbf{U}$	23
$\mathbf{S}$	A32	$\mathbf{V}$	23
$(\mathbf{C}_A^T \mathbf{C}_A^T)^{-1}$	A32	$\mathbf{A}$	23
$\mathbb{I}$	A34	$\mathbf{A}^{-1}$	23
$\frac{1}{\lambda^2}\mathbf{S}$	A34	$\mathbf{U}$	23
$(\mathbf{C}_A^T \mathbf{C}_A^T)^{-1}$	A34	$\mathbf{B}$	23
$\mathbf{S}^T$	A34	$\mathbf{V}$	23

CV

Yi Zhang

March 14, 2025

Email: zhang9@outlook.de

Github: <https://github.com/deconvolution/>

Born: 1993, Hunan, China

Research Interests

Wave equation based modelling and inversion; Regularization; Anisotropic Wave Equation; Neural Network surrogate model

Education

2023	Exchange at University of Lausanne to learn surrogate model
2020-2025	Ph.D. in Geophysics, Max Planck Graduate Center with Johannes Gutenberg University Mainz, Germany
2018-2019	Erasmus Exchange, Petroleum Geophysics, Norwegian University of Science and Technology (NTNU), Norway
2017-2020	M.Sc. in Applied Geosciences, RWTH Aachen, Germany
2012-2016	B.Eng. in Petroleum Geosciences, China University of Petroleum Beijing, China