One-Way Reflection Waveform Inversion:

Migration and Tomography

ONE-WAY REFLECTION WAVEFORM INVERSION:

MIGRATION AND TOMOGRAPHY

Dissertation

for the purpose of obtaining the degree of doctor
at Delft University of Technology
by the authority of the Rector Magnificus, prof. dr. ir. T.H.J.J. van der Hagen,
chair of the Board for Doctorates
to be defended publicly on
Thursday 23 January 2025 at 12:30 o'clock

by

Siamak ABOLHASSANI

Master of Science in Geophysics,
University of Tehran, Iran,
born in Tehran, Iran.

This dissertation has been approved by the promotors.

Composition of the doctoral committee:

Rector Magnificus,	chairperson
Dr. ir. D.J. Verschuur,	Delft University of Technology, promotor
Prof. dr. S. Stallinga,	Delft University of Technology, promotor

Independent members:

Prof. dr. ir. E.C. Slob	Delft University of Technology
Prof. dr. ir. F.C. Vossepoel	Delft University of Technology
Prof. dr. H. Chauris	Mines ParisTech, France
Dr. G. Lambaré	Viridien, France
Dr. N. Amini	Shearwater, UK

The work in this dissertation was conducted at the Imaging Physics (ImPhys) department, Faculty of Applied Sciences, Delft University of Technology, and was financially supported by the Delphi consortium.

Keywords: Reflected Seismic Data, Full Waveform Inversion, Reflection Waveform Inversion, Joint Migration Inversion, Migration, Tomography, Velocity Model Building, Resolution, Gradient Preconditioning, Seismic Multiple Scattering, One-Way Propagation

Printed by: Gildeprint

ISBN 978-94-6366-999-3

An electronic version of this dissertation is available at
http://repository.tudelft.nl/.

To my wife Sanaz
and my parents Parvin and Farhad.

CONTENTS

SUMMARY

The introduction and adoption of seismic full waveform inversion (FWI) revolutionized Earth's subsurface imaging practices. FWI uses all the information in the seismic data (amplitudes and phases) to reconstruct a detailed Earth's subsurface model. However, it does come with limitations. Beyond the reach of refractions and diving waves, FWI cannot effectively reconstruct subsurface layers. This led to the development of reflection waveform inversion (RWI), which exclusively uses the pair of transmission-after-reflection wavepaths to sample deeper compared to FWI. RWI reconstructs the background velocity model of the subsurface by alternating between a migration loop and a tomography loop.

Despite the theoretical appeal, RWI has its own share of limitations. This dissertation investigates the barriers limiting the optimal performance of reflection waveform inversion in the context of one-way RWI (ORWI), a variation of reflection waveform inversion that adopts one-way wavefield propagators to forward model seismic reflection data. After exploring the barriers, the dissertation offers solutions to improve the reliability, accuracy, and convergence of ORWI.

The dissertation acknowledges several barriers that limit the optimal performance of conventional/standard ORWI. First, ORWI relies on accurate subsurface images. However, limited-resolution images with unpreserved amplitudes, resulting from the migration loop, lead to suboptimal background velocity updates. This issue also extends to the tomography loop, where limited-resolution tomographic wavepaths impede optimal background velocity updates. Second, ORWI overlooks the immediate impact that updating the velocity model has on the reflectivity model, as the reflectors' positions in depth remain fixed while the background velocity is updated. This oversight leads to inconsistent reflectivity and velocity models in ORWI, introducing full-wave inconsistencies in the short-offset residual waveforms for tomography. Third, similar to other seismic waveform inversion techniques, ORWI suffers from the detrimental effect of including cycle-skipped data from long offsets.

To mitigate the barriers, the dissertation proposes a range of solutions. Initially, the dissertation introduces a computationally efficient high-resolution migration algorithm called preconditioned least-squares wave-equation migration (PLS-WEM) through depth-dependent gradient preconditioning. PLS-WEM reconstructs high-resolution, amplitude-preserved seismic images in fewer iterations.

Following that, by incorporating PLS-WEM into standard ORWI, the dissertation enhances ORWI, achieving improved reflectivity imaging and thereby reconstructing tomograms that are more representative of the true subsurface layers. The dissertation also proposes the following data solutions: (a) Muting short-offset residual waveforms in the tomography data to reduce the adverse imprint of inconsistencies between the reflectivity and velocity models on the tomographic gradient of ORWI. (b) Building on (a), extending the migration offset to the maximum effective migration offset (MEMO) to enhance both the signal-to-noise ratio and the illumination of the reflectivity model.

(c) Introducing a data selection algorithm to minimize the impact of cycle-skipped long-offset data.

The dissertation then presents high-resolution ORWI (HR-ORWI) technology, which leverages depth-dependent gradient preconditioning in both migration and tomography loops to reconstruct optimal tomograms in fewer cycles.

The dissertation next evaluates three approaches to depth-dependent gradient preconditioning: conventional, source-interference-free, and source-interference-inclusive. Numerical results show the superiority of the source-interference-inclusive approach, offering enhanced resolution, reduced computational demands, and resilience to source interference.

Lastly, the dissertation develops a mathematical framework that integrates early-arrival waveform inversion with ORWI through the subspace gradient method, combining the strengths of transmission and transmission-after-reflection wavepaths to enhance tomogram reconstruction.

In conclusion, this dissertation offers a comprehensive set of solutions to overcome the limitations of ORWI, facilitating its broader adoption and application in seismic exploration and velocity model building.

SAMENVATTING

De introductie en invoering van seismische full waveform inversie (FWI) heeft de beeldvorming van de ondergrond van de aarde revolutionair veranderd. FWI maakt gebruik van alle informatie in seismische data (amplitude en fase) om een gedetailleerd model van de ondergrond van de aarde te reconstrueren. Desondanks zijn er beperkingen. Diepe ondergrondlagen, buiten het bereik van refracties en zogenoemde brekende golven, kan FWI niet effectief reconstrueren. Dit heeft geleid tot de ontwikkeling van reflection waveform inversie (RWI), die uitsluitend gebruikmaakt van transmissie-na-reflectie-golfpaden om dieper te bemonsteren met de beschikbare dataset in vergelijking met traditionele FWI. Door af te wisselen tussen een migratielus en een tomografielus reconstrueert RWI het achtergrond snelheidsmodel van de ondergrond.

Ondanks de theoretische aantrekkingskracht kent RWI ook zijn beperkingen. Dit proefschrift onderzoekt de barrières die de optimale prestatie van reflection waveform inversie belemmeren, in de context van one-way RWI (ORWI), een variant van reflection waveform inversie die gebruikmaakt van een-weg golfveldpropagatoren om seismische reflectiedata voorwaarts te modelleren. Na het identificeren van de barrières presenteert dit proefschrift oplossingen die bedoeld zijn om de prestaties van ORWI te verbeteren in termen van betrouwbaarheid, nauwkeurigheid en convergentie.

Het proefschrift erkent verschillende barrières die de prestaties van standaard ORWI beperken. Ten eerste is ORWI afhankelijk van nauwkeurige afbeeldingen van de reflectoren in de ondergrond. Echter, laag-resolutiebeelden met niet-bewaarde amplitudes, resulterend uit de migratielus, leiden tot suboptimale bijwerkingen van het achtergrond snelheidsmodel. Dit probleem strekt zich ook uit tot de tomografielus, waar laag-resolutie tomografische golfpaden optimale bijwerkingen van het achtergrond snelheidsmodel in de weg staan. Ten tweede negeert ORWI het directe effect dat het bijwerken van het snelheidsmodel heeft op het reflectiviteitsmodel, omdat de posities van de reflectoren in de diepte gelijk blijven terwijl het achtergrond snelheidsmodel wordt bijgewerkt. Dit leidt tot inconsistente reflectiviteit- en snelheidsmodellen in ORWI, wat inconsistenties introduceert in de kort-offset residuele golfvormen voor tomografie. Ten derde lijdt ORWI, net als andere seismische waveform inversietechnieken, onder de nadelige effecten van het in ogenschouw nemen van cycle-skipped data van lange offsets.

Om deze barrières aan te pakken, stelt het proefschrift een reeks oplossingen voor. Allereerst introduceert het een computationeel efficiënt hoge-resolutie migratiealgoritme genaamd preconditioned least-squares wave-equation migration (PLS-WEM) door middel van diepte-afhankelijke gradient preconditionering. Dit algoritme genereert hoge-resolutie, amplitude-bewaarde seismische beelden in minder iteraties.

Vervolgens bouwt het voort op standaard ORWI door PLS-WEM te integreren om de reflectiviteitsbeelden te verbeteren en tomogrammen te reconstrueren die representatiever zijn voor de werkelijke ondergrondlagen. Het proefschrift stelt ook het volgende voor: (a) Het dempen van korte-offset residuele golfvormen in de tomografiedata om de nadelige

invloed van inconsistenties tussen reflectiviteits- en snelheidsmodellen op de tomografische gradient van ORWI te verminderen. (b) Het uitbreiden van de migratie-offset naar de maximale effectieve migratie-offset (MEMO) om zowel de signaal-ruisverhouding als de illuminatie van het reflectiviteitsmodel te verbeteren. (c) Een data-selectie-algoritme om de impact van cycle-skipped lange-offset data te minimaliseren.

Vervolgens introduceert het proefschrift high-resolution ORWI (HR-ORWI). Deze techniek, voortbouwend op de diepte-afhankelijke preconditionering in zowel de migratie- als tomografielussen, bereikt optimale tomogrammen in minder cycli.

Daarnaast vergelijkt het proefschrift drie verschillende diepte-afhankelijke preconditioneringen (updatevergelijkingen) voor least-squares wave-equation migration: conventioneel, zonder broninterferentie, en met broninterferentie. Deze vergelijking benadrukt de superioriteit van de laatste aan de hand van numerieke voorbeelden, waarbij verbeterde resolutie, verminderde computationele intensiteit en weerstand tegen broninterferentie worden aangetoond.

Tot slot biedt het proefschrift een wiskundige basis die vroege-aankomst waveform inversie combineert met ORWI via de subspace gradient-methode. Dit benut de sterke punten van zowel transmissie- als transmissie-na-reflectie-golfpaden voor verbeterde tomogramgeneratie.

Dit proefschrift presenteert kortom een uitgebreide reeks oplossingen om de beperkingen van ORWI aan te pakken, waarmee het de weg vrijmaakt voor breder gebruik en optimale prestaties in seismische exploratie en snelheidsmodelbouw.

1

INTRODUCTION

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

– ISAAC NEWTON

1.1. EXPLORATION SEISMOLOGY

Exploration seismology is an applied branch of seismology that primarily focuses on understanding Earth's subsurface structures and physical properties through the study of elastic wave propagation inside the Earth. The study starts with generating seismic waves with explosives, vibrators, or air guns placed on the Earth's surface, at depth, on the seafloor, or towed behind survey vessels in a marine environment. These waves travel into the Earth and interact with different geological layers, making direct, refracted, and reflected seismic waves. Seismic sensors, known as geophones or hydrophones, planted on the Earth's surface, at depth, on the seafloor, or towed behind survey vessels in a marine environment, record the returning seismic waves. Exploration seismologists then process the returning signals recorded by the sensors to create a seismic profile. A seismic profile represents the time it takes for the seismic waves to travel (back and forth) into the Earth as well as the amplitudes of the returned signals. A seismic profile could be indeed interpreted as a time picture of the studied subsurface.

Exploration seismology has long served as the mainstay of oil and gas exploration. This role is indeed attributed to its capabilities in providing precise subsurface imaging (from shallow to deep), identifying potential deposits of oil and gas resources, and optimizing the utilization of oil and gas reservoirs (monitoring applications).

Over the years, and particularly in recent times with the global shift towards cleaner energy alternatives, exploration seismology has actively broadened its applications beyond the traditional use linked to oil and gas. It would seem that exploration seismology continues to remain a central player in unlocking an essential component of the world's

green-energy future: geothermal energy power, through assessing and mapping subsurface heat reservoirs. Moreover, amidst the urgent challenge of climate change, exploration seismology evidently serves as a core element in navigating carbon capture and storage (CCS) technologies, aiding in the identification/monitoring of geologically suitable formations for securely sequestering CO_2. Additionally, in the pursuit of critical minerals for the "green" era, such as, among others, Copper, Lithium, Cobalt, Graphite, Aluminum, Nickel, and Manganese, exploration seismology facilitates the sustainable supply chain of those minerals necessary for the production of battery storage, solar panels, wind turbines, and electric vehicle components (renewable energy technologies). Furthermore, exploration seismology has taken a leading role in guiding and optimizing wind farm installation by characterizing potential sites and assessing the risk of geological hazards—earthquakes or unstable subsurface conditions—which could impact the long-term stability of wind turbines.

Advanced imaging algorithms, including **seismic migration** and **tomography**, may be employed to further process the seismic profiles. The objective is to accurately position the recorded seismic waves in time in their genuine subsurface locations in depth, allowing for a clear understanding of the geological attributes (structures and properties) beneath the Earth's surface.

1.2. SCALE SEPARATION: MIGRATION + TOMOGRAPHY

While seismic migration in the presence of a fixed velocity model explains the amplitudes of observed seismic waves and retrieves a high-wavenumber subsurface model, seismic tomography explains the moveouts existing in observed seismic waves and retrieves a low-wavenumber subsurface model (Figure 1.1). Although, in 1985, the model scale separation, the separation between the tomography and migration worlds, is acknowledged by Claerbout (1985), Tarantola (1984) ambitiously tries to seek a broadband-wavenumber subsurface model by introducing full waveform inversion (FWI) thorough taking advantage of all geometrical and scattering properties of observed waves (Figure 1.1), without using any sort of model scale separation. More precisely, Tarantola's FWI, least-squares FWI, is a data-fitting problem by defining an l2-norm objective function, comparing the Euclidean distance between all properties of the predicted and observed seismic waves, from now on called seismic waveforms or data, and involves an iterative closed-loop process (Figure 1.2) via forward and adjoint operators, based on the wave equation, until the data residual almost vanishes.

However, without any sort of model scale separation, Tarantola's FWI faces a serious challenge called cycle-skipping (Virieux and Operto, 2009). The cycle-skipping challenge is defined as having an estimated model that is not kinematically mature/accurate enough to predict data with less than half a cycle time shift (Δt) with respect to the observed data at each iteration. To avoid cycle-skipping, there exist two basic but significant remedies. The first remedy is known as the frequency continuation strategy, firstly introduced by Kolb et al. (1986) and Bunks et al. (1995). They showed that the likelihood of cycle skipping and being trapped in a local minimum situation is lower with low-frequency information compared to high-frequency information. This is because low frequencies exhibit a more linear relation with the low-wavenumber components of the model (Sirgue and Pratt, 2004). As is shown in Figure 1.3 schematically, a more convex objective function is

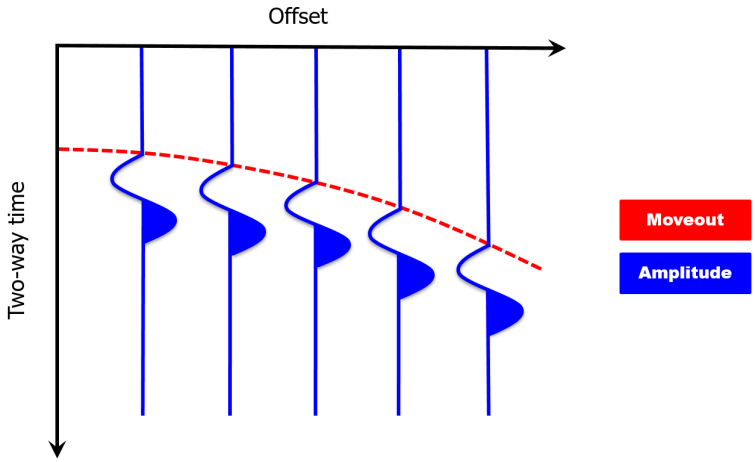


Figure 1.1: Tarantola's FWI takes advantage of both amplitudes and moveouts.

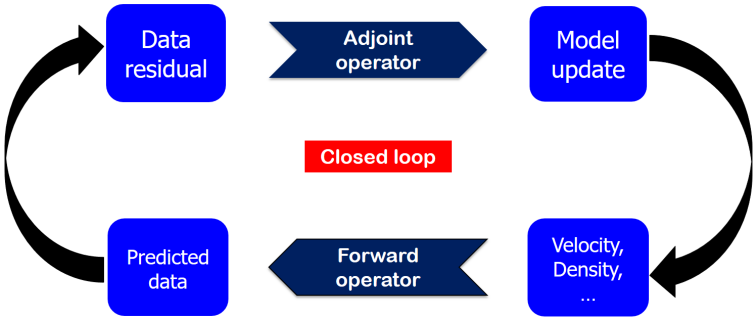


Figure 1.2: FWI workflow as a closed loop.

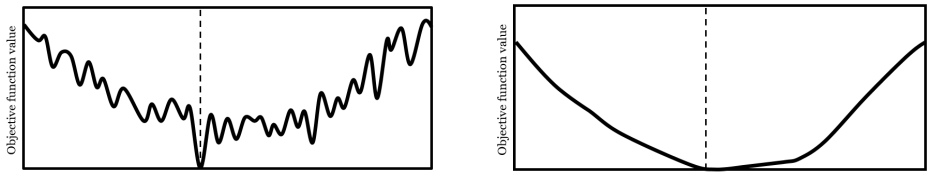


Figure 1.3: The risk of cycle skipping and being trapped in local minima with low-frequency information is less than high-frequency information. (a) A non-convex objective function is achieved using high-frequency information. (b) A more convex objective function is achieved using low-frequency information.

achieved using low-frequency information, increasing the chance of reaching the global minimum. However, the low-frequency information recorded by the data acquisition systems of that time was not low enough in order to follow the frequency continuation

strategy in practice.

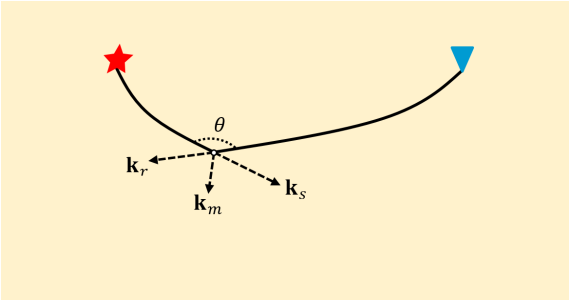


Figure 1.4: Illustration of the diffraction tomography principle. The source-, $\mathbf{k}_s$, and receiver-side plane wave vectors, $\mathbf{k}_r$, make the local illumination angle, θ, and model wavevector, $\mathbf{k}_m$, at a point diffractor.

The advent of modern seismic acquisition systems, including broadband sources and wide-aperture systems, is acknowledged as the second remedy to cycle-skipping. After the advent of broadband sources (about 3 Hz) and wide-aperture seismic acquisition systems (about 8 km), the flame of hope was rekindled that the FWI gradient could be able to retrieve a broadband wavenumber subsurface model as the modern acquisition systems can acquire long offsets and low frequencies, which are necessary to make the low local model wavenumbers. This is illustrated by the generalized diffraction tomography principle, introduced by Devaney (1982), where the local model wavevector ($\mathbf{k}_m$), sampled by the single-scattering-based FWI gradient through the zero-lag cross-correlation of the source-side and receiver-side local plane waves at each point diffractor within the medium (Figure 1.4), is given by

$$\mathbf{k}_m = \mathbf{k}_s + \mathbf{k}_r = 2k_0 \cos\left(\frac{\theta}{2}\right)\mathbf{n}, \tag{1.1}$$

where θ is called the local scattering/illumination angle, made by the source-side, $\mathbf{k}_s$, and receiver-side plane wave vectors, $\mathbf{k}_r$, at a point diffractor, k_0 is the nominal wavenumber, and $\mathbf{n}$ is a unit vector in the direction of $\mathbf{k}_m$. Note that Equation 1.1 is called Devaney's equation from now on. Sirgue and Pratt (2004) reconstruct the low-wavenumber components of the subsurface model with the simultaneous use of low frequencies and long offsets.

1.3. DIFFERENT WAVEPATHS IN THE FWI GRADIENT

Based on Devaney's equation, for a 360° and broadband data, the corresponding local model-wavenumber spectrum is completely white. However, band-limited seismic data and surface acquisitions cause a series of wavenumbers, that are typically present in the model-wavenumber spectrum, not to be sampled. As a result, the model-wavenumber spectrum will be inherently colored due to the gap in the spectrum. Before exploring the information gap, in order to understand better the rest of the discussion, different wavepaths constructing the FWI gradient are introduced. We know that the FWI gradient is the superposition of zero-lag cross-correlations between the source and receiver wavefields traveling in two different regimes, the same travel directions, either

upwards or downwards, which is called propagative regime, or opposite travel directions, which is called scattering regime, see Table 1.1. For a homogeneous model with one reflector, meaningful cross-correlations between the direct and adjoint energies have been represented in Figure 1.5a. Additionally, corresponding to each type of cross-

Table 1.1: Different regimes in the FWI gradient and their corresponding illumination angles.

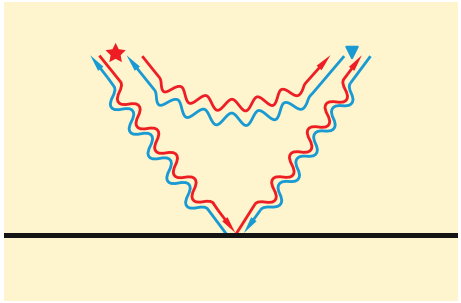
regime	illumination angle
propagative	$\Theta \approx 180°$
scattering	$\Theta < 180°$

correlation, a type of wavepath is defined, where the model parameter update is done along it. Accordingly, three different wavepaths are introduced (Figure 1.5b): (a) the transmission wavepath, marked by number 1, which is propagative-based so contributes to updating low-propagative wavenumbers, (b) the pair of transmission-after-reflection wavepaths, marked by number 2, which are propagative-based so contribute in updating low-propagative wavenumbers, and (c) the migration wavepath, marked by number 3, which is scattering-based so responsible for updating high reflective wavenumbers. Note that, in Figure 1.5b, the intensity of the blue color serves to illustrate the amplitude of the wavepaths schematically. As is seen, the pair of wavepaths marked by number 2 is shown by a less intense blue color, and this is due to having much weaker amplitudes than the others because of being scaled by the reflection coefficient of the reflector in the medium. This means that the FWI gradient is dominated by the transmission and migration wavepaths (Figure 1.6). From now on, based on Devaney's equation, the transmission (1) and migration wavepaths (3) are called low- and high-wavenumber wavepaths, respectively.

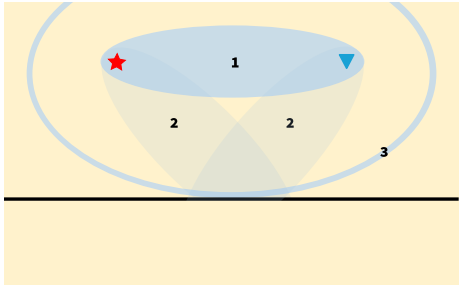
To confirm the theory presented in this part, two reflection-acquisition acoustic FWI cases are represented. While a circular velocity anomaly lies within the reach of the low-wavenumber wavepath in case (a) (Figure 1.7a), a similar anomaly lies beyond the reach of the low-wavenumber wavepath in case (b) (Figure 1.7b). The corresponding inverted models, both started from a 3000-meter-per-second homogeneous model, are shown in Figures 1.7c and 1.7d. While both kinematics and dynamics of the true velocity model are properly extracted in case (a) by the FWI gradient, there is no trace of tomography in case (b), and all is done is migration, so only the reflectors are retrieved in its corresponding inverted model.

1.4. SAMPLED LOCAL MODEL WAVENUMBERS BY FWI

Generally, sampling of high reflective-wavenumbers is nonlinearly dependent on the sampling of low-propagative wavenumbers in FWI. As a result, if low-propagative wavenumbers are not sampled accurately enough, high-reflective wavenumbers are not sampled properly. This means that low-propagative wavenumber sampling is a critical thing. For a single frequency and a homogeneous background velocity model, the local model wavenumber spectrum corresponding to each wavepath existing in the FWI gradient can



(a)



(b)

Figure 1.5: Different cross-correlations and corresponding wavepaths constructing the FWI gradient. (a) Meaningful zero-lag cross-correlations of the source-side and receiver-side local plane waves at each point diffractor, constructing the FWI gradient. (b) The wavepaths constructing the FWI gradient: the transmission wavepath, marked by number 1, the pair of transmission-after-reflection wavepaths, marked by number 2, and the migration wavepath, marked by number 3.

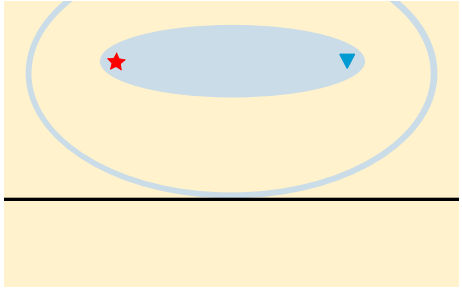
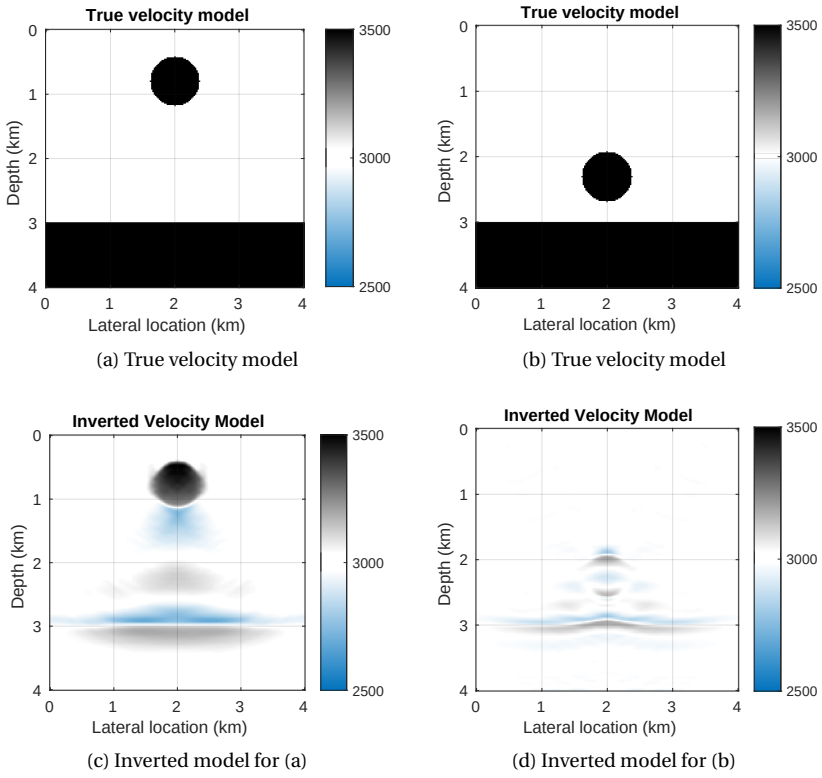


Figure 1.6: The FWI gradient is dominated by only the transmission and migration wavepaths.

be shown analytically (Zhou et al., 2018). Following the approach of Zhou et al. (2018), based on Devaney's equation, if one examines the local model wavenumbers sampled by the low-wavenumber wavepath in the FWI gradient (refer to the first wavenumber experiment in Appendix 1.A), it is evident that:

- The low-wavenumber wavepath originally samples vertical wavevectors with relative magnitudes smaller than 0.5, making the low part of the spectrum and con-



(a) True velocity model

(b) True velocity model

(c) Inverted model for (a)

(d) Inverted model for (b)

Figure 1.7: Two reflection-acquisition acoustic FWI cases are represented. (a) and (b) True velocity models. (c) and (d) Corresponding inverted models. While both kinematics and dynamics of the true velocity model are properly extracted in case (a) by the FWI gradient, there is no trace of tomography in case (b), and all that is done is the migration of the reflectors.

tributing to illuminating horizontal layers (see Figure 1.A.2).

- At deeper depths within the medium, the illumination power of the low-wavenumber wavepath naturally becomes weaker as we can not extend the maximum offset to infinity to record all the refractions coming back from the deep Earth, that is, the larger the maximum-offset-to-depth ratio, the wider the vertical low local model wavenumber spectrum bandwidth (see Figure 1.A.2).

- There are unsampled regions in the spectrum (see the two blank semicircles in Figure 1.A.2) due to the low sensitivity of the FWI gradient to the pair of transmission-after-reflection wavepaths. Indeed, the unsampled regions correspond to the low wavevectors but mostly the horizontal ones, contributing to illuminating near-vertical or dipping layers.

If one explores the impacts of large and small maximum-offset-to-depth ratios on the local model wavenumber spectrum sampled by the low-wavenumber wavepath (refer to

the second wavenumber experiment in Appendix 1.A), the following insights are gained:

- Large maximum-offset-to-depth ratios significantly contribute to extracting a wide vertical low local model wavenumber bandwidth (see Figure 1.A.5).
- Small maximum-offset-to-depth ratios lead to missing very low vertical wavevectors. Recognizing that, for a natural transition from a low-resolution velocity model to a high-resolution velocity model, the inclusion of these missing wavevectors becomes essential (see Figure 1.A.5).
- Wavevectors with pure horizontal wavenumbers are sampled when both the source and receiver are positioned at infinity (see Figure 1.A.5).
- Examining the interrelation between the sampling patterns of the low-wavenumber and high-wavenumber wavepaths within the FWI gradient reveals that they do not overlap but form a complementary relation (see Figure 1.A.5).

1.5. SAMPLED LOCAL MODEL WAVENUMBERS BY REFLECTION TOMOGRAPHY

Although the regions unsampled by the low-wavenumber wavepath of the FWI gradient could be recovered by starting the frequency continuation strategy from ultra-low frequencies (Operto et al., 2015), such a recovery is not feasible, at least for deep targets, due to the typical small maximum-offset-to-depth ratios and band-limited seismic data. Reflection tomography could be considered a suitable alternative to recover the unsampled regions.

Reflection tomography is a cyclic/iterative technique designed to enhance seismic migration by refining an initial background velocity model. Of all the reflection tomography variations, those integrating wave-equation forward modeling have gained prominence, for their superior handling of wave propagation in complex media. One may choose to conduct reflection waveform tomography in either the data domain or the image domain. The succeeding chapters will thoroughly review this versatility, but this study is exclusively centered on the data-domain approach, i.e., reflection waveform inversion (e.g., Xu et al., 2012).

Reflection waveform inversion (RWI) originated to sample deep subsurface targets using pure reflection data. Operating as a data-domain reflection waveform tomography technique, RWI involves mapping a temporary stacked image (migration) in each cycle to refine the subsurface background velocity model based on that image. This cyclic process indeed contributes to a refined representation of the subsurface image by cyclically updating the background velocity model (Figure 1.8).

If one compares the sampling pattern of the pair of transmission-after-reflection wavepaths, which is active in the reflection tomography gradient, with the other wavepaths active in the FWI gradient (refer to the third wavenumber experiment in Appendix 1.A), the following insights are gained:

- The wavenumber spectrums that are resolved do not overlap and are complementary to each other (see Figure 1.A.7).

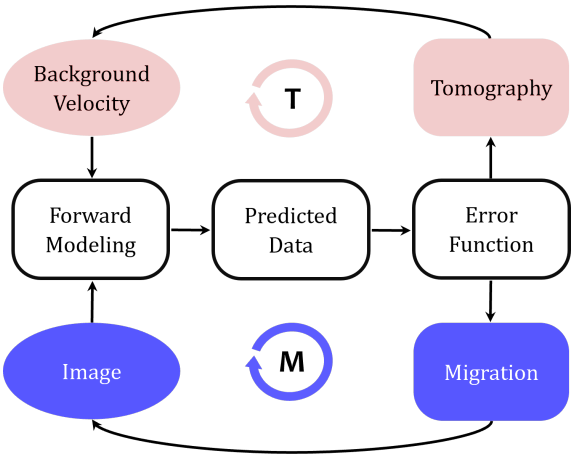


Figure 1.8: RWI cycle in which background velocity (reflection tomography) and image reconstructions (migration) alternate.

- While the pair of transmission-after-reflection wavepaths (reflection tomography gradient) sample low horizontal wavevectors, the transmission wavepath (the tomographic component of the FWI gradient) sample vertical low wavevectors (see Figure 1.A.7).

- By reducing the maximum offset by half, a gap appears in both sampled spectrums. Indeed, by decreasing the maximum offset, two things happen: at an identical depth level, (1) the tomographic component of the FWI gradient would be unable to sample similar low vertical wavevectors as before, and (2) the reflection tomography gradient would sample fewer vertical wavevectors than before, but it preserves its sampling power of low horizontal wavevectors, which is beneficial when, for any reason, acquiring the long-offset is not feasible (see Figure 1.A.7).

Throughout this dissertation, our research is supported by a reflection waveform inversion technology facilitated by a one-way forward modeling scheme known as primary wavefield modeling (PWMod) (Berkhout, 2014), which we abbreviate as one-way reflection waveform inversion (ORWI). Adequate parameterization in PWMod (angle-independent reflectivity and background velocity) leads to a natural scale separation in ORWI. This separation allows for the independent calculation of migration and tomography gradients within each ORWI cycle, thereby freeing tomograms from high-reflective model wavenumbers. Based on a zero initial reflectivity model, a smooth initial background velocity model, and pure reflection data, ORWI begins with the migration loop (see the "**M**" loop, in blue, in the cycle displayed in Figure 1.8). It proceeds to alternate between the reflection tomography loop (see the "**T**" loop, in red, in the cycle displayed in Figure 1.8), and the migration loop until reaching convergence with accurate solutions for the background velocity and image.

1

1.6. MOTIVATION OF THE DISSERTATION

Despite its conceptual appeal, RWI could face a few limitations due to multiple factors, which are as follows:

1. Non-optimal background velocity updates in RWI may partly linked to low-resolution seismic images with unpreserved amplitudes (migration loop), and the same reasoning can also be used to attribute the poor background velocity updates of RWI to low-resolution tomographic wavepaths (tomography loop).

2. The alternating approach of conventional RWI imposes a forced decoupling between reflectivity and velocity model updates—disregarding simultaneous model updates. This leads to inconsistent reflectivity and velocity models in RWI, introducing full-wave inconsistencies in the short-offset residual waveforms used for tomography.

3. All waveform inversion techniques, including RWI, are prone to being negatively impacted by the inclusion of cycle-skipped long-offset data.

By addressing these limitations in the context of ORWI, this dissertation aims to improve the accuracy, robustness and convergence of ORWI, paving the way for its wider adoption and impact in seismic imaging and velocity model building.

1.7. DISSERTATION STRUCTURE

To guide you through the remainder of this dissertation, I will briefly overview the content of each of the subsequent chapters below:

As low-resolution images with unpreserved amplitudes can degrade the background velocity updates retrieved by ORWI, **Chapter 2** proposes a computationally efficient preconditioned least-squares wave-equation migration (PLS-WEM) algorithm for estimating high-resolution, amplitude-preserved seismic images. By leveraging a one-way wavefield extrapolation technique in the frequency-depth domain, PLS-WEM construct the approximate Hessian operator for gradient preconditioning recursively depth by depth. PLS-WEM decomposes and reduces the massive approximate Hessian operator for the entire domain into multiple smaller operators, each corresponding to a specific depth level. As a result, with PLS-WEM, each time a depth-dependent approximate Hessian and its reciprocal is calculated, only a fraction of the total model parameters is involved, which saves computational costs.

Chapter 3 supercharges standard ORWI. Not only does it integrate PLS-WEM into standard ORWI for high-resolution, amplitude-preserved reflectivity imaging, but it also tackles the other ORWI limitations using new proposals. It recommends muting inconsistent short-offset residual waveforms for tomography to handle the imprint of inconsistent reflectivity and velocity models. Building upon the muting of short-offset residuals for tomography, it suggests extending the migration offset beyond short offsets to enhance reflectors' illumination and signal-to-noise ratio. This chapter also introduces a data-selection algorithm to exclude the negative impact of cycle-skipped long-offset data from tomography. These proposals unite to tackle the challenges of conventional ORWI, aiming for tomograms of the highest fidelity.

Chapter 4 proposes a high-resolution ORWI (HR-ORWI) algorithm to achieve optimal tomograms in fewer cycles through a cost-friendly gradient preconditioning in both imaging and tomography loops. By leveraging the depth-dependent preconditioning introduced in Chapter 2, HR-ORWI attempts to not only improve the seismic image resolution and accuracy but also develop the required mathematical groundwork to integrate the depth-dependent preconditioning concept for background velocity reconstruction. This two-fold effort enables the generation of high-resolution and optimal tomograms through HR-ORWI.

Investigating depth-dependent preconditioning for least-squares wave-equation migration, **Chapter 5** compares three update equations: conventional, free of source interference, and including source interference. It demonstrates the superiority of the update equation including source interference through numerical examples, showing superior resolution, reduced computational intensity, and resilience to source interference.

While the transmission wavepath mostly samples horizontal layers tomographically, the transmission-after-reflection wavepaths mostly sample dipping layers tomographically. To benefit from both, **Chapter 6** proposes combining early-arrival waveform inversion (EWI) and ORWI with the aid of the subspace gradient method. Note that this chapter is only establishing a framework and does not yet present conclusive results.

Finally, **Chapter 7** brings this dissertation to a close. Here, we recap the key findings, draw conclusions, and offer suggestions for future research that could build upon and expand our work, propelling the field further.

1.8. CONTRIBUTIONS

This dissertation significantly advances seismic subsurface imaging and velocity model building through the following contributions:

1. **Investigating ORWI limitations:** This research investigates the limitations of one-way reflection waveform inversion (ORWI), which is crucial for understanding the challenges in seismic velocity model reconstruction.

2. **Introduction of PLS-WEM:** Preconditioned least-squares wave-equation migration (PLS-WEM) is proposed as a high-resolution, amplitude-preserved imaging algorithm, which is then incorporated into standard ORWI to enhance its performance.

3. **Innovative data solutions:** This includes muting short-offset residual waveforms to address inconsistencies in reflectivity and velocity models, extending the migration offset for better reflectivity model illumination and signal-to-noise ratio, and introducing a data-selection algorithm to exclude cycle-skipped long-offset data—all aimed at enhancing ORWI performance.

4. **Introduction of HR-ORWI:** High-resolution ORWI (HR-ORWI) is introduced, leveraging depth-dependent gradient preconditioning in both imaging and tomography loops to reconstruct optimal tomograms in fewer cycles.

5. **Investigating various approaches for depth-dependent preconditioning:** Three different approaches to depth-dependent gradient preconditioning for least-squares

wave-equation migration are investigated, offering insights into preserving amplitudes and enhancing resolution while reducing computational demands.

6. **A novel mathematical framework for integrating EWI and ORWI:** A novel mathematical framework is developed, synergizing early-arrival waveform inversion with one-way reflection waveform inversion using the subspace gradient method to reconstruct broadband tomograms.

Overall, these contributions advance the field of seismic reflection tomography, providing enhanced insights into the Earth's subsurface model.

REFERENCES

Berkhout, A. J. (2014). "Review paper: An outlook on the future seismic imaging, part I: forward and reverse modelling". In: *Geophysical Prospecting* 62.5, pp. 911–930.

Bunks, C. et al. (1995). "Multiscale seismic waveform inversion Geophysics". In: *Geophysics* 60.5, pp. 1457–1473.

Claerbout, J. (1985). *Imaging the Earth's Interior*. Blackwell Scientific Publication.

Devaney, A. J. (1982). "A filtered backprojection algorithm for diffraction tomography". In: *Ultrasonic imaging* 4, pp. 336–350.

Kolb, P., F. Collino, and P. Lailly (1986). "Pre-stack inversion of a 1-D medium". In: *Proceedings of the IEEE* 74.3, pp. 498–508.

Operto, S. et al. (2015). "Efficient 3-D frequency-domain mono-parameter full-waveform inversion of ocean-bottom cable data: application to Valhall in the visco-acoustic vertical transverse isotropic approximation". In: *Geophysical Journal International* 202.2, pp. 1362–1391.

Sirgue, L. and R. G. Pratt (2004). "Efficient waveform inversion and imaging: A strategy for selecting temporal frequencies". In: *Geophysics* 69.1, pp. 231–248.

Tarantola, A. (1984). "Inversion of seismic reflection data in the acoustic approximation". In: *Geophysics* 49.8, pp. 1259–1266.

Virieux, J. and S. Operto (2009). "An overview of full-waveform inversion in exploration geophysics". In: *Geophysics* 74.6, WCC1–WCC26.

Xu, S. et al. (2012). "Inversion on reflected seismic wave". In: SEG Annual International Meeting, Expanded Abstracts.

Zhou, W. et al. (2018). "Velocity model building by waveform inversion of early arrivals and reflections: A 2D case study with gas-cloud effects". In: *Geophysics* 83.2, R141–R157.

APPENDIX

1.A. WAVENUMBER ANALYSES

This appendix explores the sampling patterns of wavepaths within the FWI and RWI gradients in the $k_x - k_z$ plane, relying on Devaney's equation. Through a few experiments, we aim to understand how the wavepaths distribute and sample information in this space.

1.A.1. FIRST WAVENUMBER EXPERIMENT

Following Zhou et al. (2018), relying on Devaney's equation, the first wavenumber experiment is designed to analyze and illustrate the sampled wavenumbers by the low-wavenumber wavepath in the Full Waveform Inversion (FWI) gradient. For three sets of parameters listed in Table 1.A.1, different in trial depth level, three local model wavenumber spectrums corresponding to the low-wavenumber wavepath are displayed in Figure 1.A.2. Note that, each point in the local model wavenumber spectrums displayed in this report denotes the tail of a wavevector. The resultant spectrums are displayed in Figure 1.A.2.

Table 1.A.1: First wavenumber experiment: three sets of parameters used for illustrating the illuminated wavenumbers sampled by the low-wavenumber wavepath.

case	frequency (Hz)	background velocity (m/s)	lateral distribution of sources (km)	maximum offset (km)	trial depth (km)
1	10	2000	[-10, 10]	±10	1
2	10	2000	[-10, 10]	±10	3
3	10	2000	[-10, 10]	±10	6

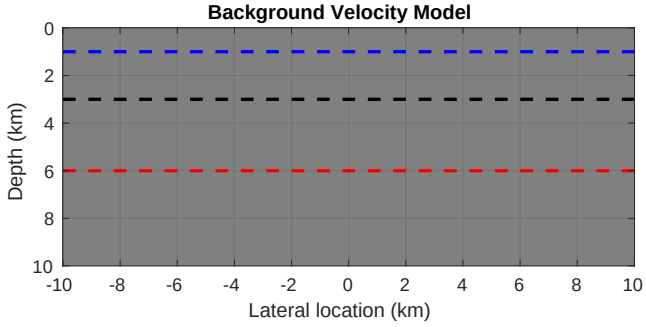


Figure 1.A.1: The trial depth levels are marked blue (1 km deep), black (3 km deep), and red (5 km deep) in a homogenous background velocity model (2000 m/s).

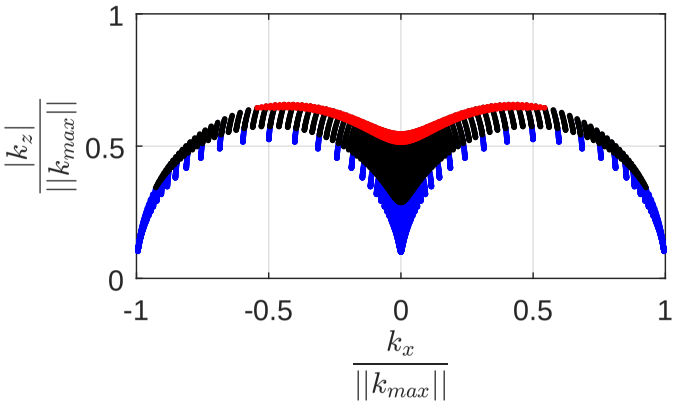


Figure 1.A.2: For three sets of parameters listed in Table 1.A.1, different in trial depth level, three local model wavenumber spectrums corresponding to the low-wavenumber wavepath are displayed.

1.A.2. SECOND WAVENUMBER EXPERIMENT

To see the effects of large and small maximum-offset-to-depth ratios on the local model wavenumber spectrum sampled by the low-wavenumber wavepath in the FWI gradient, the second wavenumber experiment is designed. For two sets of parameters listed in Table 1.A.2, different in maximum offset, two local model wavenumber spectrums corresponding to the low-wavenumber wavepath are displayed in Figure 1.A.4.

Table 1.A.2: Second wavenumber experiment: two sets of parameters used for illustrating the illuminated wavenumbers sampled by the low-wavenumber wavepath.

case	frequency (Hz)	background velocity (m/s)	lateral distribution of sources (km)	maximum offset (km)	trial depth (km)
1	10	2000	[-10, 10]	$\pm\infty$	3
2	10	2000	[-10, 10]	± 10	3

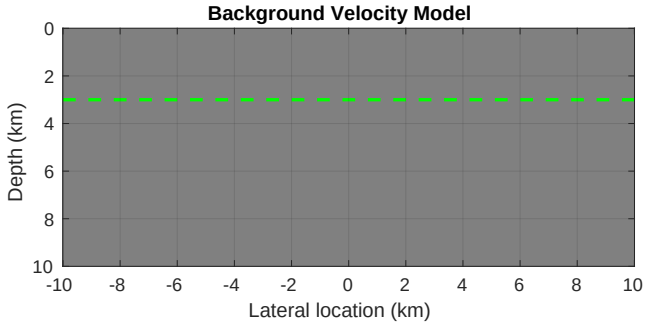
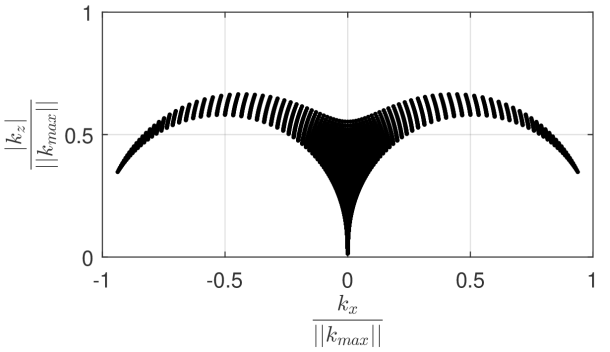
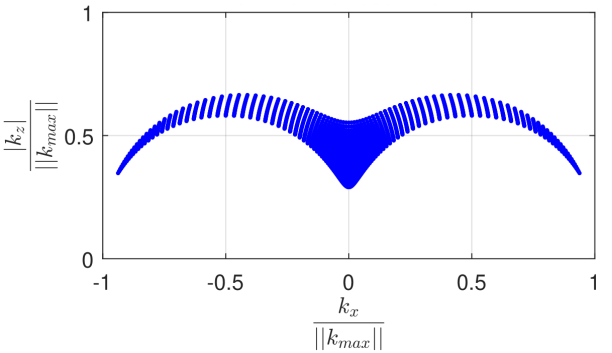


Figure 1.A.3: The trial depth level is marked green (3 km deep) in a homogenous background velocity model (2000 m/s).

For the same set of parameters, but with both the maximum offset and the lateral distribution of sources extending from $-\infty$ to $+\infty$ km, Figure 1.A.5 depicts the sampled model wavenumber spectrum by the low- and high-wavenumber wavepaths within the FWI gradient.



(a) Maximum offset = $\pm\infty$



(b) Maximum offset = ± 10

Figure 1.A.4: For two sets of parameters listed in Table 1.A.2, different in maximum offset, two local model wavenumber spectrums corresponding to the low-wavenumber wavepath are displayed. (a) The local model wavenumber spectrum is sampled by the low-wavenumber wavepath with an infinite maximum offset. (b) The local model wavenumber spectrum is sampled by the low-wavenumber wavepath with a maximum offset of ± 10 km.

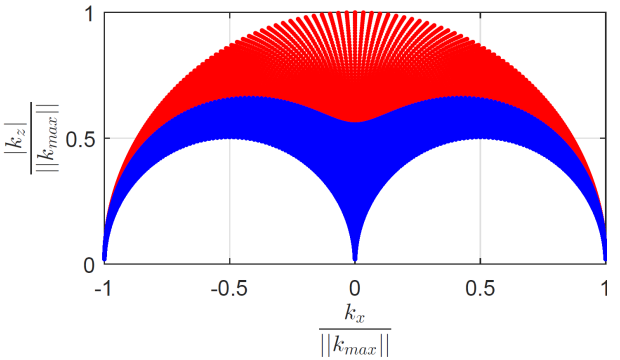


Figure 1.A.5: For the set of parameters listed in Table 1.A.2, when both the maximum offset and the lateral distribution of sources extend from $-\infty$ to $+\infty$ km, the local model wavenumber spectrums corresponding to the low- and high-wavenumber wavepaths are displayed.

1

1.A.3. THIRD WAVENUMBER EXPERIMENT

To evaluate the sampling pattern of the pair of transmission-after-reflection wavepaths, which is active in the reflection tomography gradient, in connection to the wavepaths in the FWI gradient, the third wavenumber experiment is designed. For two sets of parameters listed in Table 1.A.3, different in maximum offset, two local model wavenumber spectrums corresponding to all introduced wavepaths are displayed in Figure 1.A.7.

Table 1.A.3: Third wavenumber experiment: two sets of parameters used for illustrating the illuminated wavenumbers sampled by all introduced wavepaths.

case	frequency (Hz)	background velocity (m/s)	lateral distribution of sources (km)	maximum offset (km)	trial depth (km)	reflector depth (km)
1	10	2000	0-20	±10	1	1.3
2	10	2000	0-20	±5	1	1.3

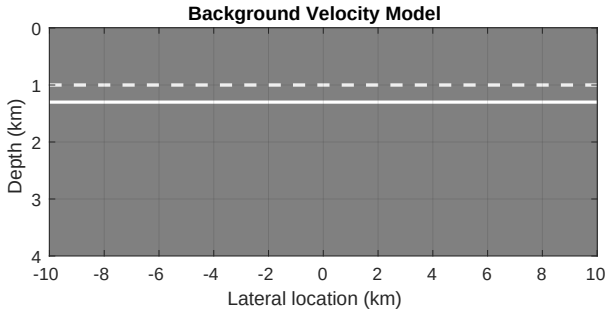
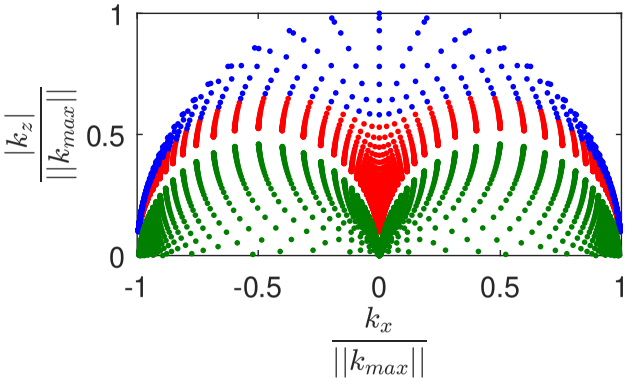
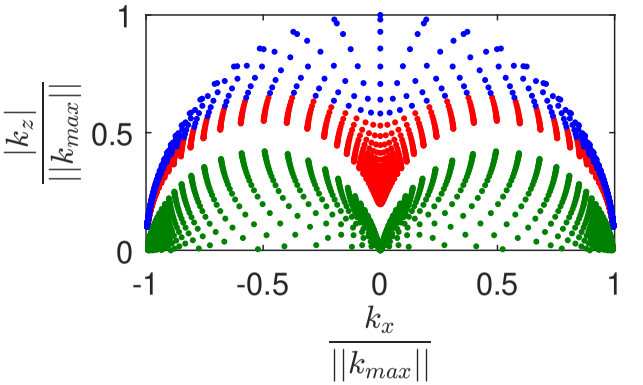


Figure 1.A.6: The trial depth level (1 km deep) and reflector (1.3 km deep) are shown by a dashed white line and a solid white line, respectively, in a homogenous background velocity model (2000 m/s).



(a) Maximum offset = ±10 km



(b) Maximum offset = ±5 km

Figure 1.A.7: For two sets of parameters listed in Table 1.A.3, different in maximum offset, two local model wavenumber spectrums corresponding to all introduced wavepaths (blue: migration wavepath; red: transmission wavepath; green: transmission-after-reflection wavepath) are displayed. (a) The local model wavenumber spectrum is sampled by all introduced wavepaths with a maximum offset of ±10 km. (b) The local model wavenumber spectrum sampled by all introduced wavepaths with the maximum offset of ±5 km.

2

EFFICIENT PRECONDITIONED LEAST-SQUARES WAVE-EQUATION MIGRATION

"Success is often the result of taking a misstep in the right direction."

– AL BERNSTEIN

Since the appearance of wave-equation migration, many have tried to improve the resolution and effectiveness of this technology. Least-squares wave-equation migration is one of those attempts that tries to fill the gap between the migration assumptions and reality in an iterative manner. However, these iterations do not come cheap. A proven solution to limit the number of least-squares iterations is to correct the gradient direction within each iteration via the action of a preconditioner that approximates the inverse Hessian. However, the Hessian computation, or even the Hessian approximation computation, in large-scale seismic imaging problems involves an expensive computational bottleneck, making it unfeasible. Therefore, we propose an efficient computation of the Hessian approximation operator, in the context of one-way wave-equation migration (WEM) in the space-frequency domain. We build the Hessian approximation operator depth by depth, considerably reducing the operator size each time it is calculated. We prove the validity of our proposed method with two numerical examples. We then extend our proposal to the framework of full-wavefield migration, which is based on WEM principles but includes interbed multiples. Finally, this efficient preconditioned least-squares full-wavefield migration is successfully applied to a dataset with strong interbed multiple scattering.

This chapter is based on the following paper: **Abolhassani, S.**, & Verschuur, D. J. (2024). Efficient preconditioned least-squares wave-equation migration. *Geophysics*, 89(3), S275-S288.

2.1. INTRODUCTION

Seismic migration, also called seismic imaging, has been an imperative tool in characterizing the Earth's subsurface geological structures in the search for subsurface resources over the last half-century. Numerous research efforts have been attempted to contribute to the theoretical developments of modern seismic migration, being one- or two-way wave-equation migration, notable among which are Claerbout (1971) and Claerbout and Doherty (1972) for introducing finite-difference migration, Stolt (1978) and Gazdag (1978) for inventing and developing migration in the wavenumber-frequency domain, Berkhout and Wulfften (1979) for introducing migration as a spatial deconvolution in the space-frequency domain, and also Baysal et al. (1983), Whitmore (1983), and McMechan (1983) for pioneering the use of reverse-time migration (RTM).

Most wave-equation migration algorithms share the same imaging mechanism. They first propagate the source wavefield forward into the medium while also propagating the receiver wavefield backward into the medium. They then construct the subsurface image by applying an imaging condition (Jones, 2014) on the forward and backward propagated wavefields at every trial image point. Note that wave-equation migration can also be interpreted and implemented as a generalized diffraction stack migration (Schuster, 2002; Zhan et al., 2014).

Despite the similarities, wave-equation migration algorithms differ in a couple of aspects, among which are the dimension—either time or depth—along which they propagate wavefields, and the scalar Helmholtz equation—either one-way or two-way—they solve to propagate wavefields. Generally speaking, the most popular wave-equation migration algorithms can be split into two main kinds: one-way wave-equation migration (WEM) and RTM (Etgen et al., 2009; Jones, 2014). While RTM propagates wavefields along the time axis through the numerical solution of the two-way Helmholtz equation via direct methods such as finite-difference, spectral-element, and finite-element modeling, WEM propagates wavefields along the depth axis through the numerical solution of the one-way Helmholtz equation. WEM algorithms, to avoid the well-known computational overhead of direct solutions to the Helmholtz equation, are mostly built on the solution of an approximation to the square root of the Helmholtz operator, which is cheap, specifically for 3D cases (Mulder and Plessix, 2004). Although RTM outperforms WEM in addressing large-propagation angles (greater than $\pm 90^\circ$), so delivering superior images in complex subsurface regions, it is a more computationally-intensive and memory-demanding operation. Hence, WEM is still a frequently used migration technique in the industry for high-frequency imaging or large-scale datasets (Mehta et al., 2017), if steep reflecting structures are not expected. Additionally, WEM principles may be used in the so-called full-wavefield modeling (Berkhout, 2014b) to generate multiple scattering in order to apply full-wavefield migration (Berkhout, 2014a). From another perspective, however, it may be argued that interpreting such comparative statements between WEM and RTM as the outright dismissal of one over the other is a misconception. In fact, the preference for one over the other could be significantly affected by various factors, particularly the geological complexity of the medium, the current stage of the project—it is in the early stage or full production mode—and, more importantly, efficient project resource allocation.

WEM and RTM algorithms both have proven effective on their own provided that

a couple of fundamental migration assumptions are satisfied: (a) their input dataset is regularly sampled, (b) their input dataset is free of multiples if the first-order Born approximation is followed, (c) there is no amplitude problem in their input dataset due to the source energy variation/dissipation, (d) there is an accurate migration velocity model, (e) there is a broadband source function, and (f) there is an accurate migration operator. Otherwise, their output would not be an ideal representation of the subsurface reflectivity model, as they suffer from migration artifacts (Jones, 2018).

To resolve better images and close the gap between the migration assumptions and reality, the iterative least-squares migration (LSM) concept was introduced and leveled up any standard migration algorithm into a local minimization problem (e.g., Cole and Karrenbach, 1992; Nemeth et al., 1999; Chavent and Plessix, 1999; Duquet and Marfurt, 1999). While the number of iterations increases, any LSM technique can effectively suppress the part of migration artifacts that appear by irregular acquisition geometry, a band-limited source function, and geometrical spreading (Huang et al., 2014). This iterative process, however, comes with high computational costs, so keeping the number of iterations to a minimum is crucial.

To reduce the number of least-squares iterations, one remedy is to precondition the gradient vector in each iteration with an approximation to the reciprocal of the Hessian matrix-operator (Pratt et al., 1998). As a result, the model perturbation vector $\Delta\mathbf{m}$ reads,

$$\mathbf{H}^{\mathrm{a}}\Delta\mathbf{m} = -\mathbf{g}, \tag{2.1}$$

in which $\mathbf{g}$ and $\mathbf{H}^{\mathrm{a}}$ denote the gradient and Hessian approximation, respectively, and are given by,

$$\mathbf{g} = \mathbf{J}^{\dagger}\Delta\mathbf{d}, \tag{2.2}$$

$$\mathbf{H}^{\mathrm{a}} = \mathbf{J}^{\dagger}\mathbf{J}, \tag{2.3}$$

where $\Delta\mathbf{d}$ is the data error vector, $\dagger$ denotes adjoint, and $\mathbf{J}$ is the forward (Born) modeling operator, or Jacobian matrix-operator, each column of which represents the wavefield scattered by a small perturbation of the model parameter m_i (i denotes the location of the model parameter here) while all the other model parameters are kept fixed. It mathematically follows that while diagonal Hessian elements carry the scattered wavefield auto-correlations, off-diagonal Hessian elements carry the scattered wavefield cross-correlations of neighboring model parameters in the medium (Operto et al., 2013). Although it is expected that the scattered wavefields exhibit only auto-correlations (i.e., only on-diagonal coefficients in $\mathbf{H}^{\mathrm{a}}$ have values), this is not the case. Indeed, due to the limited bandwidth of seismic sources and the proximity of model parameters in space, they are also partially cross-correlated (both on-diagonal and off-diagonal coefficients in $\mathbf{H}^{\mathrm{a}}$ have values). It follows that the Hessian approximation matrix has a diagonally-dominant structure rather than a pure diagonal one, as noted by Pratt et al. (1998).

As is clear from equation 2.1, each element of the gradient vector is a weighted sum of all the model perturbations. This clarifies how blurred will be a migrated image using a forward modeling operator with nonunitary columns (i.e. $\mathbf{H}^{\mathrm{a}}$ is not an identity matrix) and, at the same time, explains why LSM after the action of the inverse Hessian approximation on the gradient converges faster (Aoki and Schuster, 2009). However, for large-scale seismic imaging problems, even such an approximation to the Hessian

operator has a considerable computational burden—if it is rebuilt/updated in each iteration—as it requires in each iteration to construct, invert, and multiply a matrix of size: number of model parameters × number of model parameters.

To make the computation of the Hessian operator feasible, some have replaced it with a diagonal approximation (e.g., Beydoun and Mendes, 1989; Chavent and Plessix, 1999; Shin et al., 2001), and others have approximated the entire set of coefficients with a manageable computational burden. Unlike the reciprocal of the diagonal coefficients, which only corrects the migration image for the amplitude-related artifacts, the reciprocal of the entire set of coefficients can correct the migration image for both bandwidth- and amplitude-related artifacts. Therefore, to gain a better image resolution, there have been many proposals to approximate the entire Hessian matrix, either in the data domain or image domain, among which the most remarkable are as follows. Hu et al. (2001) approximate the inverse Hessian matrix by a migration deconvolution filter constructed based on a single reference acoustic velocity profile $v(z)$, leaving laterally-invariant deconvolution filters. For a layered medium with lateral velocity variations, Yu et al. (2006) construct a set of different deconvolution filters based on multiple acoustic velocity profiles to apply on different parts of the migrated image. Guitton (2004) approximates the Hessian inverse matrix by designing a matching convolution filter for each point in the image space to locally match a reference image to its corresponding migrated image in a least-squares sense. For a target-oriented imaging problem, Valenciano et al. (2006) compute a sparse-structure Hessian matrix explicitly through the cross-correlation of the source and receiver Green's functions within the target area. Lecomte (2008) and Fletcher et al. (2016) calculate the Hessian approximation matrix using point spread functions (PSFs), in which each PSF measures the scattered wavefield for every point in the image space. Metivier et al. (2014) estimate the product of the Hessian approximation matrix and any vector that lies within the subsurface model space through the second-order adjoint-state method and then solve equation 2.1 using conjugate-gradient iterations. Assis and Schleicher (2021) use the same technique as Metivier et al. (2014) but in the context of one-way reflection waveform inversion. Lu et al. (2018), in the context of viscoacoustic anisotropic WEM, compute $\mathbf{J}$ and $\mathbf{J}^{\dagger}$ implicitly and then solve equation 2.1 using an iterative sparse solver of the LSQR family. In their recent study, Yang et al. (2021) approximate the inverse Hessian by comparing the S-transform spectra of two images of the subsurface: one obtained through Born migration, and the other obtained by remodeling the Born image and then remigrating it.

Over the past years, there have also been significant studies on transforming the migration operator from an adjoint to a pseudo-inverse under the high-frequency assumption, commonly known as a "true-amplitude migration operator", mostly taking into account the geometrical spreading loss. In such context, Zhang et al. (2007) achieve a true-amplitude WEM by introducing the high-frequency approximation of true-amplitude downgoing (forward) and upgoing (backward) wavefields at every trial image point. Kiyashchenko et al. (2007) propose a true-amplitude correlation-based imaging condition applicable with a finite-difference solution of the one-way wave equation, mathematically equivalent to the high-frequency approximation of the LSM solution (Bleistein et al., 2001). There exist alternative true-amplitude migration schemes that explicitly take into account the transmission effect as well (e.g., Deng and McMechan, 2007). ten Kroode

(2012) derives a pseudo-inverse to the subsurface extended scale-separated Kirchhoff modeling operator, and Hou and Symes (2015) obtain a pseudo-inverse to the subsurface extended Born modeling operator. In the same spirit as Hou and Symes, Chauris and Cocher (2017) construct a pseudo-inverse to the subsurface extended Born modeling operator but after linearizing the phase of the Born operator, promoting shorter offsets, smaller dips, and vertical rays in practice. A pseudo-inverse operator, more importantly, can be advantageous in a least-squares framework too as it can be applied to the data residual to speed up the convergence rate of the least-squares inversion scheme that relies merely on the steepest descent update direction (e.g., Hou and Symes, 2016).

This paper aims to present a cost-friendly Hessian approximation operator in the context of space-frequency domain least-squares one-way wave-equation migration (LS-WEM) relying on angle-independent reflection/transmission coefficients (Berkhout, 1982; Berkhout, 2014a). To this purpose, we build the Hessian approximation operator depth by depth, significantly reducing the operator size each time it is calculated (Abolhassani and Verschuur, 2022). We also rebuild/update the Hessian approximation operator within each iteration, taking into account the updated transmission effects. This paper is organized as follows. First, we present our preconditioned LS-WEM theory (forward and inverse problems), in which the reciprocal of our proposed depth-based Hessian approximation operator is used to precondition the gradient direction. This is followed by two numerical examples for two synthetic datasets, including only primary reflections, one from a velocity model with a lens-shaped inclusion and the other from the SEG/EAGE overthrust velocity model. In addition, we evaluate the effectiveness of our proposed approach within the context of full-wavefield migration (Berkhout, 2014a) with a numerical example for a synthetic dataset including strong interbed multiples. We finish with a final discussion and conclusions.

2.2. PRECONDITIONED LEAST-SQUARES ONE-WAY WAVE-EQUATION MIGRATION (PLS-WEM)

The theoretical and algorithmic aspects of PLS-WEM are presented here.

2.2.1. FORWARD PROBLEM

The Kirchhoff integral for homogeneous fluids describes how to model the pressure wavefield at an arbitrary point inside a closed surface S when the pressure wavefield and the normal component of the particle velocity everywhere on S are known. The Kirchhoff integral can be simplified by choosing a plane surface for S and absorbing boundary conditions for one-way Green's functions (see, e.g., Berkhout and Wapenaar, 1989). The resultant integral is then represented as the Rayleigh II integral, which is the basis here for one-way forward and adjoint wavefield extrapolations in a 2D inhomogeneous acoustic medium.

With an extrapolation step bounded by origin and destination depth levels, the 2D forward extrapolated wavefield, in which the wavefields existing at all lateral positions located at the origin depth level (z_ℓ) are extrapolated in the +z direction towards a lateral position located at the destination depth level ($z_{\ell+1}$) (Figure 2.1), is defined as (known as

the Rayleigh II integral),

$$p^+(x_j,z_{\ell+1},\omega_f)=\frac{1}{2\pi}\int_{-\infty}^{+\infty}\tilde{p}^+\left(k_x,z_\ell,\omega_f\right)e^{-i\left(\sqrt{\left(\frac{\omega_f}{v_j}\right)^2-k_x^2}\right)|\Delta z|}e^{-ik_xx_j}\,dk_x, \tag{2.4}$$

where v_j is the acoustic velocity at position $(x_j,z_{\ell+1})$, ω_f denotes a given angular frequency component, k_x is the horizontal wavenumber, Δz denotes the laterally constant extrapolation step (the vertical distance between the origin and destination depth levels) and is assumed small enough to be considered homogeneous vertically, and $p^+(x_j,z_{\ell+1},\omega_f)$ and $\tilde{p}^+\left(k_x,z_\ell,\omega_f\right)$ are the monochromatic downgoing acoustic wavefields that read the following forward and inverse Fourier conventions,

$$\tilde{p}(k_{x_j})=\int_{-\infty}^{+\infty}p(x)e^{ik_{x_j}x}dx, \tag{2.5}$$

$$p(x_j)=\frac{1}{2\pi}\int_{-\infty}^{+\infty}\tilde{p}\left(k_x\right)\mathrm{e}^{-ik_xx_j}\,dk_x. \tag{2.6}$$

Inserting equation 2.5 into 2.4 gives a convolution integral along the x-axis,

$$\begin{aligned}p^+(x_j,z_{\ell+1},\omega_f)&=\frac{1}{2\pi}\int_{-\infty}^{+\infty}p^+(x,z_\ell,\omega_f)\overbrace{\left(\int_{-\infty}^{+\infty}e^{+i\left(\sqrt{\left(\frac{\omega_f}{v_j}\right)^2-k_x^2}\right)|\Delta z|}e^{-ik_x(x_j-x)}\,dk_x\right)}^{w^+\text{: extrapolation kernel}}dx\\&=\frac{1}{2\pi}\int_{-\infty}^{+\infty}p^+\left(x,z_\ell,\omega_f\right)w^+_{z_{\ell+1};z_\ell}\left(x_j-x,\omega_f\right)dx,\end{aligned} \tag{2.7}$$

in which the downward wavefield extrapolation is marked by w^+, and x_j denotes the convolution lag. Rewriting equation 2.7 into the vector-matrix form gives the following equations (Berkhout, 1982),

$$\mathbf{p}^+(z_{\ell+1},\omega_f)=\mathbf{W}^+_{z_{\ell+1};z_\ell}\mathbf{p}^+(z_\ell,\omega_f), \tag{2.8}$$

$$\mathbf{p}^-(z_\ell,\omega_f)=\mathbf{W}^-_{z_\ell;z_{\ell+1}}\mathbf{p}^-(z_{\ell+1},\omega_f), \tag{2.9}$$

where $\mathbf{p}^-(z_\ell,\omega_f)$ is the monochromatic upgoing acoustic wavefield at z_ℓ, and $\mathbf{p}^+(z_\ell,\omega_f)$ is the monochromatic downgoing acoustic wavefield at z_ℓ, $\mathbf{p}^+(z_{\ell+1},\omega_f)$ is the monochromatic downgoing acoustic wavefield at $z_{\ell+1}$, $\mathbf{p}^-(z_{\ell+1},\omega_f)$ is the monochromatic upgoing acoustic wavefield at $z_{\ell+1}$, $\mathbf{W}^+_{z_{\ell+1};z_\ell}$ is the downward $(z_\ell\rightarrow z_{\ell+1})$ wavefield extrapolation matrix-operator, and $\mathbf{W}^-_{z_\ell;z_{\ell+1}}$ is the upward $(z_{\ell+1}\rightarrow z_\ell)$ wavefield extrapolation matrix-operator.

With an extrapolation step without lateral velocity variations ($v_j=v$ everywhere at the destination depth level), $\mathbf{W}$ becomes a Toeplitz matrix-operator, i.e., each row vector-operator of $\mathbf{W}$ comes with a finite length of n_x and contains the same elements as others but is moved by one element to the right compared to its preceding row. However, with an extrapolation step including lateral velocity variations (a different v_j for every x_j at the destination depth level), $\mathbf{W}$ turns into a space-variant convolution matrix-operator, i.e., each row vector-operator of $\mathbf{W}$ is defined based on a locally-averaged velocity so comes with a short spatial extent of L such that $L\ll n_x$ (Berkhout, 1982; Thorbecke et al., 2004).

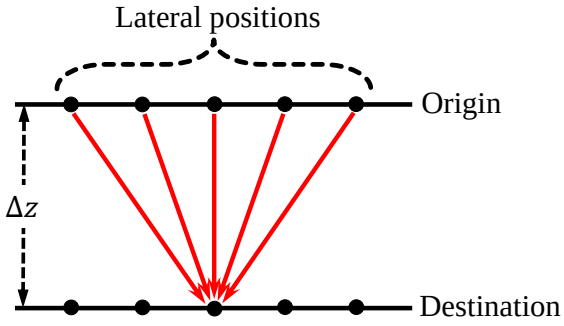


Figure 2.1: Extrapolation from all existing lateral positions located at the origin depth level towards a lateral position located at the destination depth level. Note that Δz denotes the extrapolation step.

To model monochromatic angle-independent primary seismic data at the Earth's surface (traditional surface seismic data) for a given shot location, the following equations, extrapolating the seismic source signature downward from z_0 (the Earth's surface) to z_N and then upward from z_N to z_0, are solved in the space-frequency domain,

$$\mathbf{p}^{-}_{\text{mod}}\left(z_0,\omega_f\right)=\sum_{m=N}^{1}\mathbf{U}^{-}_{z_0;z_m}\left(\mathbf{r}^{\cup}\left(z_m\right)\circ\mathbf{p}^{+}_{\text{mod}}\left(z_m,\omega_f\right)\right), \tag{2.10}$$

$$\mathbf{p}^{+}_{\text{mod}}\left(z_m,\omega_f\right)=\mathbf{U}^{+}_{z_m;z_0}\mathbf{s}^{+}\left(z_0,\omega_f\right), \tag{2.11}$$

$$\mathbf{U}^{+}_{z_m;z_0}=\left[\prod_{n=m-1}^{1}\mathbf{W}^{+}_{z_{n+1};z_n}\mathbf{T}^{+}\left(z_n\right)\right]\mathbf{W}^{+}_{z_1;z_0}, \tag{2.12}$$

$$\mathbf{U}^{-}_{z_0;z_m}=\left[\prod_{n=1}^{m-1}\mathbf{W}^{-}_{z_{n-1};z_n}\mathbf{T}^{-}\left(z_n\right)\right]\mathbf{W}^{-}_{z_{m-1};z_m}, \tag{2.13}$$

in which $\mathbf{p}^{-}_{\text{mod}}\left(z_0,\omega_f\right)$ denotes the monochromatic upgoing wavefield received at the depth level z_0, $\mathbf{p}^{+}_{\text{mod}}\left(z_m,\omega_f\right)$ indicates the monochromatic downgoing wavefield received at the depth level z_m, $\mathbf{r}^{\cup}\left(z_m\right)$ represents the angle-independent upward reflectivity vector-operator at z_m, $\mathbf{T}^{+}\left(z_n\right)$ is the downward transmission diagonal-operator and reads $\mathbf{T}^{+}=\mathbf{I}+\text{diag}\left(\mathbf{r}^{\cup}\right)$, $\mathbf{T}^{-}$ is the upward transmission diagonal-operator and reads $\mathbf{T}^{-}=\mathbf{I}-\text{diag}\left(\mathbf{r}^{\cup}\right)$, $\mathbf{U}^{+}_{z_m;z_0}$ is called the total downward extrapolator and contains all the downward extrapolation operators ($\mathbf{W}^{+}$) coupled with the downward transmission operators required to reach from z_0 to z_m, $\mathbf{U}^{-}_{z_0;z_m}$ is called the total upward extrapolator and contains all the upward extrapolation operators ($\mathbf{W}^{-}$) coupled with the upward transmission operators required to reach from z_m to z_0, $\mathbf{s}^{+}\left(z_0,\omega_f\right)$ indicates the monochromatic downgoing physical source at the Earth's surface, N is the total number of depth levels, and finally the symbol $\circ$ means the Hadamard product (Figure 2.2). Henceforward, we refer to the above-described primary wavefield modeling approach as PWMod.

As clear from equations 2.10 to 2.13, PWMod calculates the Earth's primary two-way response via a recursive summation (equation 2.10) in depth by including reflection and transmission effects. The recursive summation, indeed, includes the multiplication of the total upward extrapolator and the upward scattered wavefield at each depth level.

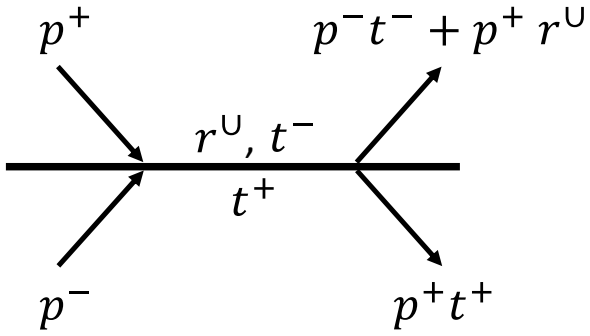


Figure 2.2: Wavefield updating in PWMod with angle-independent reflections at a given depth level, where p^+ is the downgoing wavefield, p^- is the upgoing wavefield, $r^\cup$ represents the upward reflectivity, t^+ and t^- denote the downward and upward transmission.

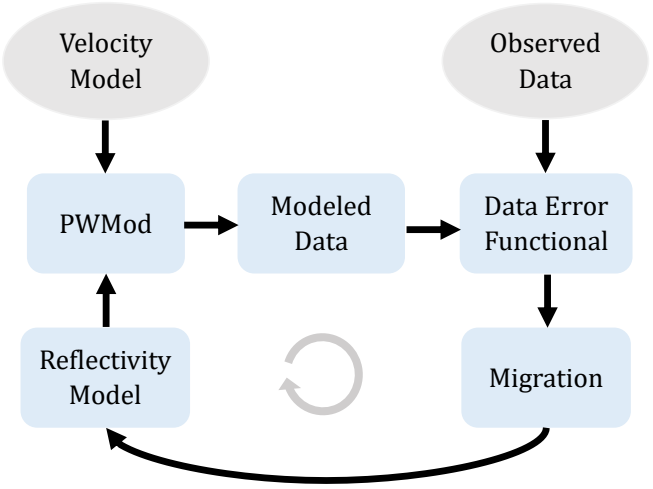


Figure 2.3: LS-WEM cycle. LS-WEM starts with an accurate migration velocity model (fixed in the cycle) and a zero-reflectivity model (variable in the cycle). Then, using PWMod, the modeled data is generated and contrasted with the multiple-free observed data using a data error functional. The data error is then back-projected into the reflectivity model to build/update the reflectors through a scaled steepest descent algorithm. The cycle repeats itself until the data error almost disappears.

2.2.2. INVERSE PROBLEM

LS-WEM is a seismic depth migration technology based on a data-fitting process that seeks the Earth's reflectivity model. LS-WEM iteratively minimizes a data error functional in a least-squares sense (Schuster, 2017). To minimize the data error functional, LS-WEM follows a migration-demigration cycle, as depicted in Figure 2.3. The LS-WEM cycle starts with an accurate migration velocity model fixed in the cycle and a zero-reflectivity model variable in the cycle. The modeled data is then generated via PWMod, and the data error is calculated and back-projected into the reflectivity model to build/update the reflectors via a steepest descent algorithm. The cycle iterates until the data error almost vanishes.

To introduce preconditioning in the LS-WEM algorithm, we seek the least-squares solution of the linear system (Lines and Treitel, 1984; Tarantola, 1984),

$$\Delta \mathbf{d}^{-} = \mathbf{J}\Delta \mathbf{r}^{\cup}, \tag{2.14}$$

where $\Delta \mathbf{d}^{-}$ represents the error between the observed and modeled data at z_0, $\Delta \mathbf{r}^{\cup}$ denotes the total upward reflectivity perturbation, and $\mathbf{J}$ is given by,

$$\mathbf{J} = \frac{\partial \mathbf{p}^{-}_{\text{mod}}(z_0)}{\partial \mathbf{r}^{\cup}}. \tag{2.15}$$

Therefore, to solve equation 2.14, we minimize the following error functional (Shin et al., 2001; Jang et al., 2009; Oh and Min, 2013), as a quadratic function of $\Delta \mathbf{r}^{\cup}(\omega_f)$,

$$\mathscr{C} = \frac{1}{2}\sum_{f=1}^{N_f}\sum_{s=1}^{N_s}\left\|\Delta \mathbf{d}^{-}_{s}(\omega_f) - \mathbf{J}_s(\omega_f)\,\Delta \mathbf{r}^{\cup}(\omega_f)\right\|_2^2, \tag{2.16}$$

in which

$$\Delta \mathbf{d}^{-}_{s}(\omega_f) = \mathbf{p}^{-}_{\text{obs},s}(z_0,\omega_f) - \mathbf{p}^{-}_{\text{mod},s}(z_0,\omega_f,\mathbf{r}^{\cup}), \tag{2.17}$$

and $\mathbf{p}^{-}_{\text{obs},s}(z_0,\omega_f)$ is the monochromatic multiple-free observed data recorded at the Earth's surface for shot s, $\mathbf{p}^{-}_{\text{mod},s}(z_0,\omega_f,\mathbf{r}^{\cup})$ represents the monochromatic modeled primary data at the Earth's surface for shot s, $\Delta \mathbf{d}^{-}_{s}(\omega_f)$ is the monochromatic residual data for shot s, ω_f represents an angular frequency component, N_s is the total number of shot locations, N_f indicates the total number of frequency components, and $\mathbf{r}^{\cup}$ is ordered as,

$$\mathbf{r}^{\cup} = \begin{bmatrix} \mathbf{r}^{\cup}(z_0) \\ \mathbf{r}^{\cup}(z_1) \\ \vdots \\ \mathbf{r}^{\cup}(z_N) \end{bmatrix}. \tag{2.18}$$

Minimizing equation 2.16 with respect to $\Delta \mathbf{r}^{\cup}(\omega_f)$ gives the descent direction,

$$\Delta \mathbf{r}^{\cup} = -\sum_{f=1}^{N_f}\Delta \mathbf{r}^{\cup}(\omega_f), \tag{2.19}$$

in which

$$\Delta\mathbf{r}^{\cup}(\omega_f) = \Re\left\{\overbrace{\left(\sum_{s=1}^{N_s}\mathbf{J}_s^{\dagger}(\omega_f)\mathbf{J}_s(\omega_f)\right)}^{\mathbf{H}_s^{\mathrm{a}}(\omega_f)}\right\}^{-1}\Re\left\{\overbrace{\sum_{s=1}^{N_s}\mathbf{J}_s^{\dagger}(\omega_f)\Delta\mathbf{d}_s^{-}(\omega_f)}^{\mathbf{g}_s(\omega_f)}\right\}, \tag{2.20}$$

and, to link the linearized and non-linear inverse problems, $\mathbf{g}$ and $\mathbf{H}^{\mathrm{a}}$ are referred to as the gradient and Hessian approximation (also known as the Gauss-Newton Hessian approximation). The introduced frequency-dependent preconditioning operation here exhibits certain parallels with the deconvolution imaging condition introduced by Valenciano and Biondi (2003) that accounts for source deconvolution.

Using equation 2.20, the reflectivity model can be updated iteratively via,

$$\mathbf{r}_{k+1}^{\cup} = \mathbf{r}_k^{\cup} + \alpha_k\Delta\mathbf{r}_k^{\cup}, \tag{2.21}$$

where α denotes the minimization step length, and k denotes the current iteration. Since with PWMod, the acoustic wavefield is accessible depth by depth, the gradient vector in equation 2.20 can be rewritten as,

$$\mathbf{g}_s(\omega_f) = \begin{bmatrix}\mathbf{g}_s(z_0,\omega_f)\\ \vdots \\ \mathbf{g}_s(z_N,\omega_f)\end{bmatrix} = \Big[\overbrace{\frac{\partial\mathbf{p}_{\mathrm{mod},s}^{-}(z_0,\omega_f)}{\partial\mathbf{r}^{\cup}(z_0)}\cdots\frac{\partial\mathbf{p}_{\mathrm{mod},s}^{-}(z_0,\omega_f)}{\partial\mathbf{r}^{\cup}(z_N)}}^{\text{total Jacobian}}\Big]^{\dagger}\Delta\mathbf{d}_s^{-}(\omega_f), \tag{2.22}$$

in which each element of $\mathbf{g}$ is a vector representing the gradient associated with the model parameters located at a given depth level, and each column of the total Jacobian denotes the partial derivative of the upgoing modeled wavefield at the receiver locations with respect to the model parameters located at a given depth level. What is important to note here is that equation 2.22 allows us to construct not only the gradient vector depth by depth but also the Jacobian matrix, paving the way for constructing the Hessian approximation matrix with a reduced number of elements at each depth level.

Each column of the total Jacobian matrix, after calculating the corresponding partial derivatives, is simplified to,

$$\begin{aligned}\frac{\partial\mathbf{p}_{\mathrm{mod},s}^{-}(z_0,\omega_f)}{\partial\mathbf{r}^{\cup}(z_m)} = \Bigg[&\mathbf{U}_{z_0;z_m}^{-}\begin{pmatrix}p_{1\,\mathrm{mod},s}^{+}(z_m,\omega_f)\\ 0\\ \vdots\\ 0\end{pmatrix}\mathbf{U}_{z_0;z_m}^{-}\begin{pmatrix}0\\ p_{2\,\mathrm{mod},s}^{+}(z_m,\omega_f)\\ \vdots\\ 0\end{pmatrix}\\ \cdots\ &\mathbf{U}_{z_0;z_m}^{-}\begin{pmatrix}0\\ 0\\ \vdots\\ p_{n_x\,\mathrm{mod},s}^{+}(z_m,\omega_f)\end{pmatrix}\Bigg],\end{aligned} \tag{2.23}$$

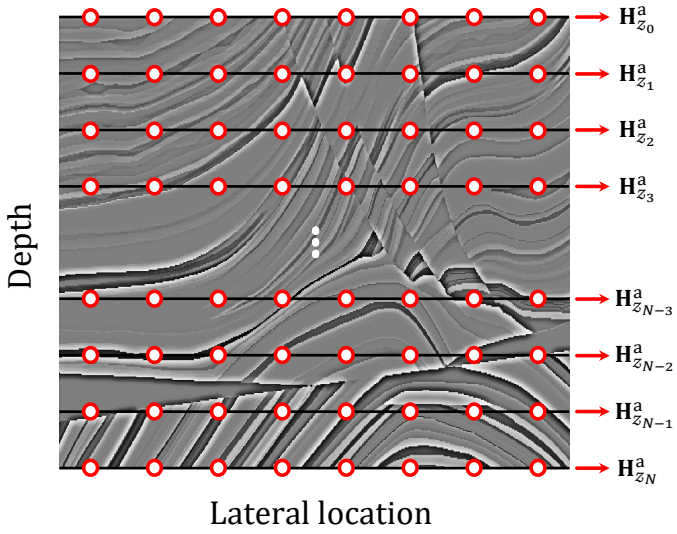


Figure 2.4: Decomposing the massive Hessian approximation operator into several small operators, each related to a depth level.

where n_x denotes the number of grid points at z_m (inversion and modeling grids are similar), and $p^+_{j\ \mathrm{mod},s}(z_m,\omega_f)$ is a complex number representing the downgoing modeled wavefield at the jth grid point of z_m. As the vectors multiplied by $\mathbf{U}^-_{z_0;z_m}$ in equation 2.23 each have only one active element, the multiplications can be simplified to,

$$\frac{\partial \mathbf{p}^-_{\mathrm{mod},s}(z_m,\omega_f)}{\partial \mathbf{r}^{\cup}(z_m)} = \Big[p^+_{1\ \mathrm{mod},s}(z_m,\omega_f)\,\mathbf{U}^-_{*,1\,z_0;z_m} \quad p^+_{2\ \mathrm{mod},s}(z_m,\omega_f)\,\mathbf{U}^-_{*,2\,z_0;z_m} \quad \dots \quad p^+_{n_x\,\mathrm{mod},s}(z_m,\omega_f)\,\mathbf{U}^-_{*,n_x\,z_0;z_m}\Big], \tag{2.24}$$

in which $\mathbf{U}^-_{*,j\,z_0;z_m}$ denotes the jth column of $\mathbf{U}^-_{z_0;z_m}$.

According to equation 2.3, the Hessian approximation matrix for each pair of source and frequency is expressed as,

$$\mathbf{H}^{\mathrm{a}}_s(\omega_f) = \Big[\frac{\partial \mathbf{p}^-_{\mathrm{mod},s}(z_0,\omega_f)}{\partial \mathbf{r}^{\cup}}\Big]^{\dagger}\Big[\frac{\partial \mathbf{p}^-_{\mathrm{mod},s}(z_0,\omega_f)}{\partial \mathbf{r}^{\cup}}\Big], \tag{2.25}$$

and as seen, it shows a square and symmetric structure with the massive dimension of $n_m \times n_m$, where n_m represents the total number of model parameters in the whole medium. As an alternative to constructing one massive, computationally unfeasible Hessian approximation operator all at once, where it is required to perform the cross-correlation between the partial derivative wavefields associated with model parameters located at all depth levels (the whole medium), we split up the massive Hessian approximation operator into several smaller operators, each calculated for the model parameters located at a different depth level (n_x model parameters are located at each depth level).

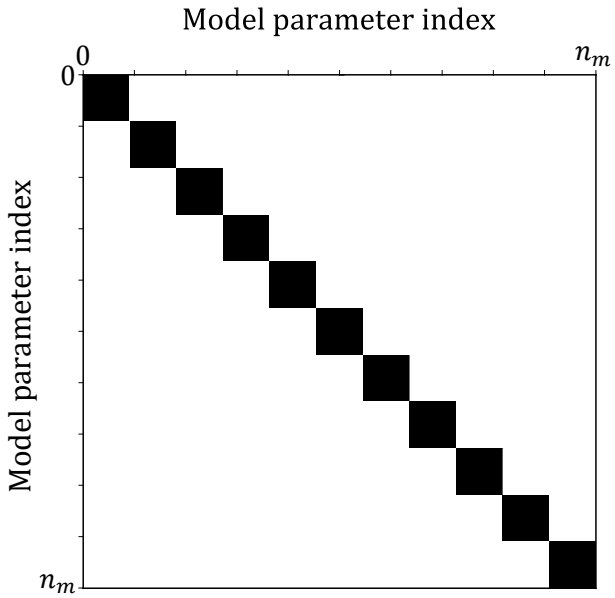


Figure 2.5: Proposed small Hessian approximation operators assembled into one big operator. Each block is a matrix of $n_x \times n_x$ related to a depth level, where $n_x = \frac{n_m}{N}$ and $N = 11$.

To do so, it suffices to write,

$$\mathbf{H}_s^{\mathrm{a}}(z_m,\omega_f) = \Big[\frac{\partial \mathbf{p}^-_{\mathrm{mod},s}(z_0,\omega_f)}{\partial \mathbf{r}^{\cup}(z_m)}\Big]^{\dagger}\Big[\frac{\partial \mathbf{p}^-_{\mathrm{mod},s}(z_0,\omega_f)}{\partial \mathbf{r}^{\cup}(z_m)}\Big], \tag{2.26}$$

where the reciprocal of $\mathbf{H}_s^{\mathrm{a}}(z_m,\omega_f)$ compensates for geometrical spreading and spatial correlations of neighboring model parameters, while also conducting source deconvolution.

Therefore, to precondition every depth-level gradient, we calculate an individual tiny Hessian operator with the dimension of $n_x \times n_x$, where $n_x = \frac{n_m}{N}$ (Figure 2.4), reducing the number of model parameters by a factor of N whenever the Hessian operator is calculated. This is feasible since equation 2.24 enables us to construct $\frac{\partial \mathbf{p}^-_{\mathrm{mod},s}(z_0,\omega_f)}{\partial \mathbf{r}^{\cup}}$ for each depth level in the medium. This obviously allows us to decompose the Hessian approximation operator into several small operators, each of which only carries the correlation of partial derivative wavefields associated with a single depth level. These minimal operators are computationally cheap to build, store, and invert in each iteration, enabling us to precondition the gradient vector efficiently. Figure 2.5 represents the structure of such an approximate Hessian.

2.3. NUMERICAL EXAMPLES OF PLS-WEM

To investigate the effect of the suggested approximate Hessian while solving equation 2.19 in each iteration, we contrast PLS-WEM with a scaled version of LS-WEM. In this scaled version, the LS-WEM gradient vector is scaled by the diagonal components of an inverse approximate Hessian (see Plessix and Mulder, 2004, equation 27), where the Hessian operator itself is given by,

$$\mathbf{H} = \frac{\partial^2 \mathscr{C}'}{\partial \mathbf{r}^{\cup 2}} = \frac{\partial^2}{\partial \mathbf{r}^{\cup 2}} \left(\frac{1}{2} \sum_{f=1}^{N_f} \sum_{s=1}^{N_s} \left\| \Delta \mathbf{d}_s^- \left(\omega_f \right) \right\|_2^2 \right), \tag{2.27}$$

and the approximate Hessian operator is derived by neglecting the terms in the derivative above that depend on the residual data, as discussed in Cova and Innanen (2013).

Both PLS-WEM and LS-WEM methods are tested here on two synthetic datasets associated with a lens-shaped inclusion model and the SEG/EAGE overthrust model.

2.3.1. LENS-SHAPED INCLUSION MODEL

In this example, our true model is a homogeneous model with a lens-shaped anomaly in the middle (Figure 2.6). The true model is represented by 201 × 173 grid points in the horizontal and vertical directions. The horizontal and vertical grid sizes are 15 and 7 m, respectively. 41 shot points with 75 m spacing are laid out on top of the model, and 201 receivers with 15 m spacing are used per shot to record the reflection data (fixed-spread acquisition). The source function used in this example is a Ricker wavelet with a dominant frequency of 10 Hz. Our forward modeling tool for generating the observed dataset (only primaries) is PWMod, applied to the true model, the so-called inverse crime. The recording time is 1.4 s, and the largest offset used in the migration process is 3 km. Employing an accurate migration velocity model with sharp reflectors, we apply

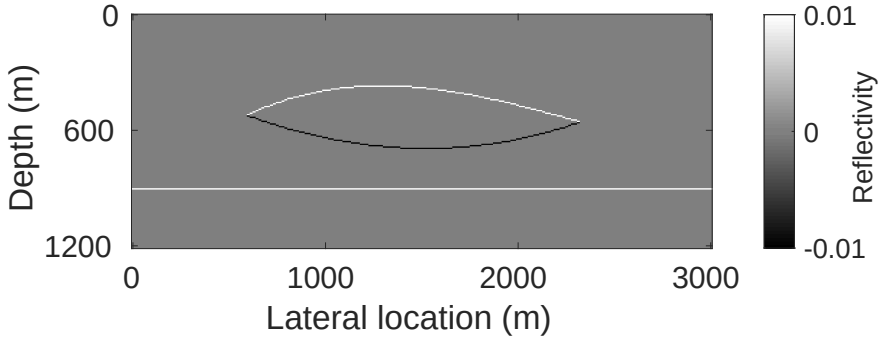
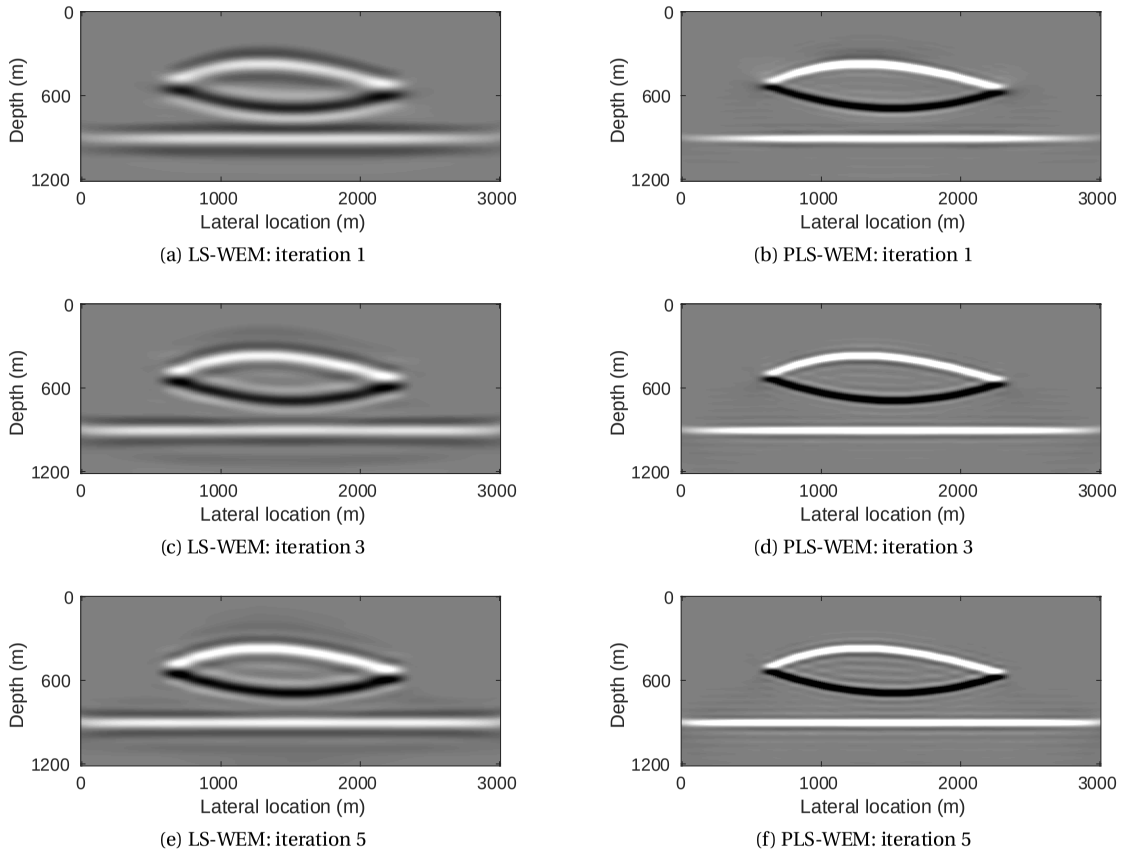


Figure 2.6: True lens-shaped inclusion reflectivity model.

both LS-WEM and PLS-WEM in such a setup to evaluate the action of the reciprocal of the suggested depth-based Hessian approximation operator on the gradient. Figure 2.7 depicts the results. Figures 2.7a, 2.7c, and 2.7e display the LS-WEM results after 1, 3, and 5 iterations, respectively. Figures 2.7b, 2.7d, and 2.7f show the PLS-WEM results after 1, 3, and 5 iterations, respectively.



(a) LS-WEM: iteration 1

(b) PLS-WEM: iteration 1

(c) LS-WEM: iteration 3

(d) PLS-WEM: iteration 3

(e) LS-WEM: iteration 5

(f) PLS-WEM: iteration 5

Figure 2.7: Estimated reflectivity models using the LS-WEM and PLS-WEM methods after 1, 3, and 5 iterations, associated with the true lens-shaped inclusion model shown in Figure 2.6. (a), (c), and (e) Estimated reflectivity models by LS-WEM after 1, 3, and 5 iterations, respectively. (b), (d), and (f) Estimated reflectivity models by PLS-WEM after 1, 3, and 5 iterations, respectively.

The results show that the reciprocal of the suggested depth-based Hessian approximation operator significantly impacts the gradient vector from very early iterations. Image deconvolution and preserved amplitudes are clearly the most noticeable impacts of employing the suggested preconditioner. Therefore, the PLS-WEM images demonstrate superior focusing compared to the LS-WEM images. As seen in Figures 2.7e and 2.7f, even after 5 least-squares iterations, the LS-WEM image quality does not reach the image quality of PLS-WEM after its first iteration.

To compare the convergence characteristics of LS-WEM and PLS-WEM, their normalized data error values in each iteration in the log scale are plotted in Figure 2.12a. As evident, the PLS-WEM method shows a faster convergence ratio and also reaches smaller data errors, which means its estimated image fits the observed data better.

2.3.2. SEG/EAGE OVERTHRUST MODEL

In the next example, our true model is one vertical slice of the 3D SEG/EAGE overthrust model (Figure 2.8) (Aminzadeh et al., 1994). The selected model contains 501 × 151 grid points in the horizontal and vertical directions, with a grid interval of 24 m. 101 shot points with an interval of 120 m are used on top of the model, and 501 receivers per shot with an interval of 24 m collect the reflection data (fixed-spread acquisition). PWMod is used to generate the observed reflection data (only primaries) using a Ricker wavelet with a dominant frequency of 10 Hz as the source function, the so-called inverse crime. The trace length is 2.6 s, and the largest offset used in the migration process is 4 km. With the aid of an accurate migration velocity model including sharp reflectors, both LS-WEM and PLS-WEM are tested in such a setup to examine the action of the reciprocal of the suggested depth-based Hessian approximation operator on the gradient. The results are shown in Figure 2.9. Figures 2.9a, 2.9c, and 2.9e display the LS-WEM results after 1, 3, and 5 iterations, respectively. Figures 2.9b, 2.9d, and 2.9f show the PLS-WEM results after 1, 3, and 5 iterations, respectively.

The results confirm the effectiveness of the action of the suggested preconditioner on the gradient. From the first iteration, balanced-amplitude reflectors and image deconvolution are again the most visible outcomes of applying the depth-based Hessian approximation inverse on the gradient vector. While LS-WEM in its early iterations leaves us with an image affected by the limited bandwidth of the seismic data, PLS-WEM mitigates such an unfavorable effect and estimates a high-resolution image from early iterations. As demonstrated in Figures 2.9e and 2.9f, even after 5 iterations, unfocused

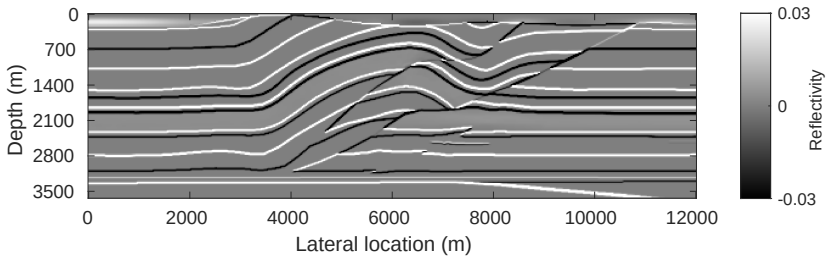
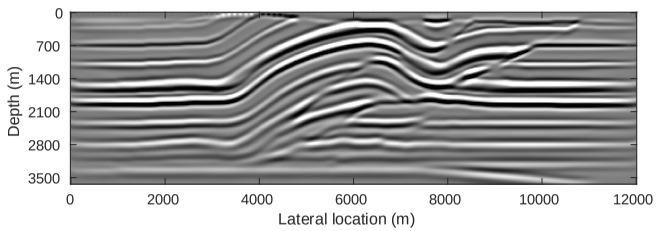
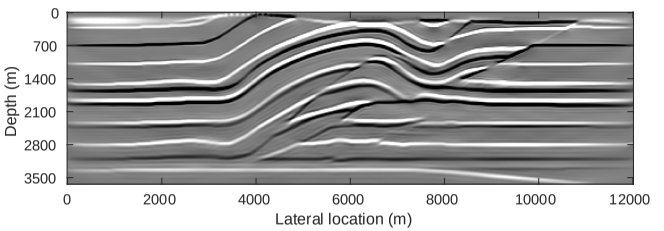


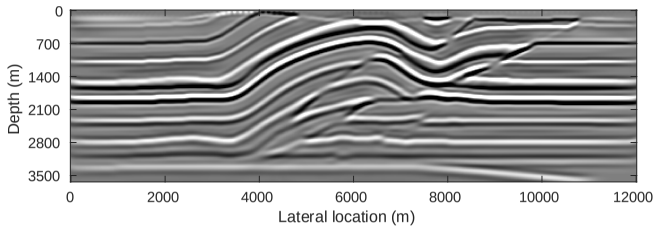
Figure 2.8: True SEG/EAGE overthrust reflectivity model.



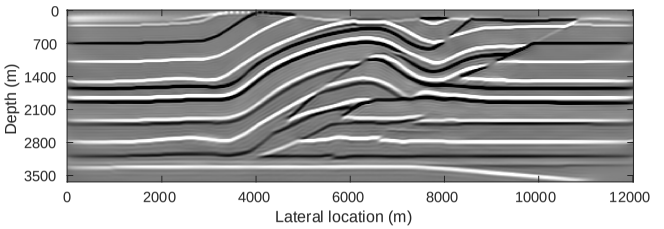
(a) LS-WEM: iteration 1



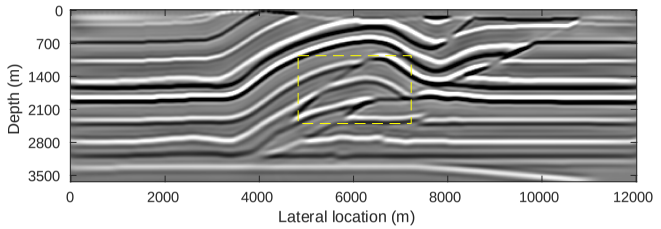
(b) PLS-WEM: iteration 1



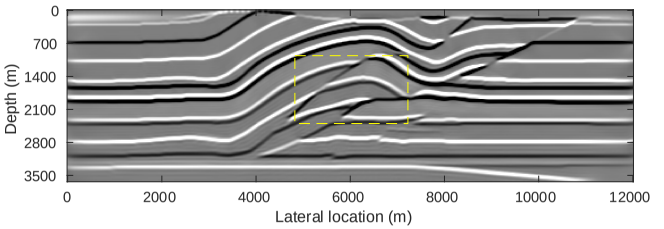
(c) LS-WEM: iteration 3



(d) PLS-WEM: iteration 3

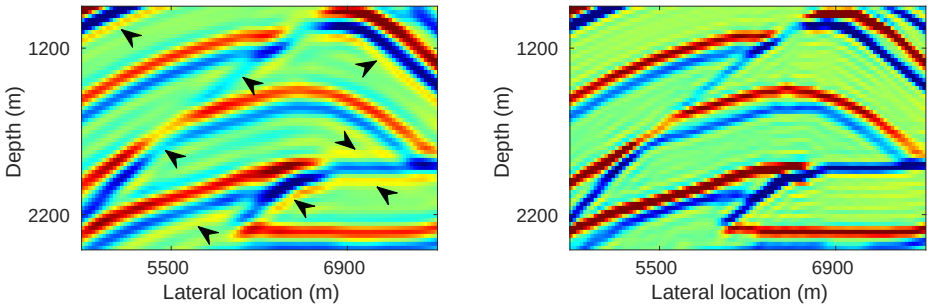


(e) LS-WEM: iteration 5



(f) PLS-WEM: iteration 5

Figure 2.9: Estimated reflectivity models using the LS-WEM and PLS-WEM methods after 1, 3, and 5 iterations, associated with the true SEG/EAGE overthrust reflectivity model shown in Figure 2.8. (a), (c), and (e) Estimated reflectivity models by LS-WEM after 1, 3, and 5 iterations, respectively. (b), (d), and (f) Estimated reflectivity models by PLS-WEM after 1, 3, and 5 iterations, respectively.



(a) LS-WEM: iteration 5

(b) PLS-WEM: iteration 5

Figure 2.10: Magnified sections of the estimated SEG/EAGE overthrust reflectivity models by LS-WEM and PLS-WEM after 5 iterations, shown in Figures 2.9e and 2.9f. (a) Magnified section of the estimated reflectivity model by LS-WEM. (b) Magnified section of the estimated model by PLS-WEM.

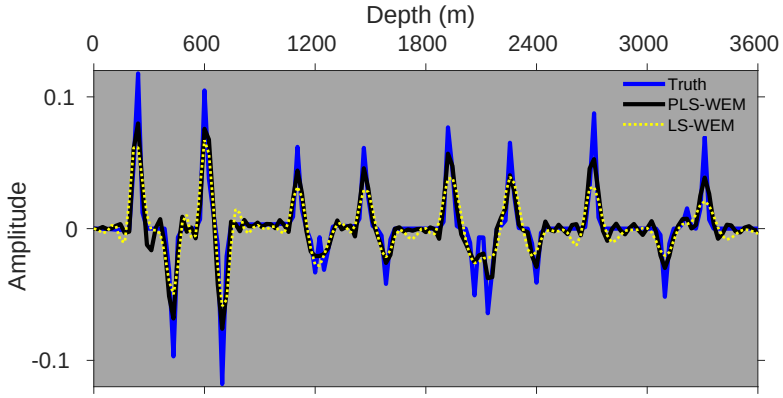
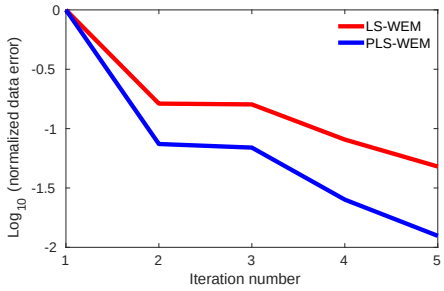


Figure 2.11: Estimated 1D reflectivity profiles using the LS-FWM and PLS-FWM methods after 5 iterations, associated with the SEG/EAGE overthrust model. The profiles are taken from the estimated reflectivity models shown in Figures 2.9e and 2.9f and represent the lateral location of 6 km.

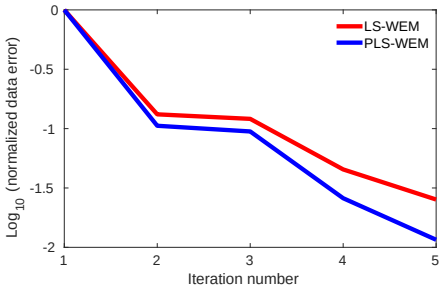
reflection energies are visible in the LS-WEM image, but the PLS-WEM image appears to be focused clearly.

In Figure 2.10, a specific region of the estimated images after 5 iterations, specified by dashed yellow rectangles in Figures 2.9e and 2.9f, is presented to obtain magnified sections for deeper investigation. While in Figure 2.10b (PLS-WEM) faults and curvatures are accurately imaged together with focused, stronger, accurate reflectors, in Figure 2.10a (LS-WEM) faults and curvatures are mapped inaccurately together with unfocused, weaker reflectors. Figure 2.11 also compares the 1D vertical reflectivity profiles estimated by LS-WEM and PLS-WEM with the true profile at the lateral location of 6 km.

The convergence properties of LS-WEM and PLS-WEM are compared in Figure 2.12b by plotting their data error values for each iteration. As expected, the PLS-WEM method shows a higher convergence ratio and achieves smaller data errors, which means that its estimated image is more in line with the observed data. Figure 2.13 represents the initial

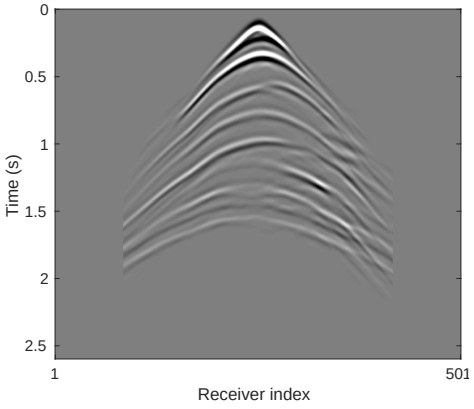


(a) Lens-shaped inclusion model

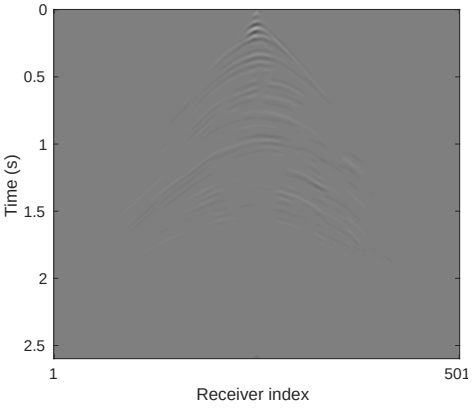


(b) SEG/EAGE overthrust model

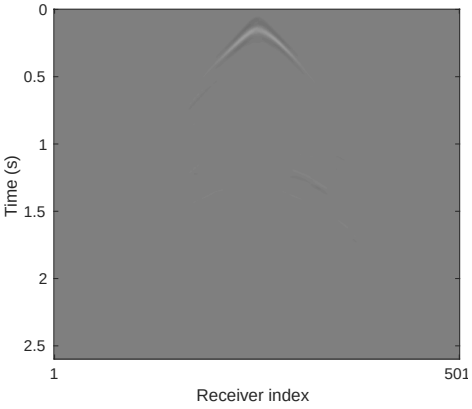
Figure 2.12: Convergence curves. (a) Lens-shaped inclusion model. (b) The SEG/EAGE overthrust model.



(a) Initial



(b) LS-WEM: iteration 5



(c) PLS-WEM: iteration 5

Figure 2.13: Residual data associated with the SEG/EAGE overthrust model. (a) Initial residual data. (b) Final residual data with LS-WEM (iteration 5). (c) Final residual data with PLS-WEM (iteration 5).

and final residual data, confirming the convergence properties.

2.4. APPLICATION TO LEAST-SQUARES FULL-WAVEFIELD MIGRATION

In the preceding examples, we ignored multiple scattering while modeling the seismic data. As a result, to avoid matching data events with different scattering orders, known as cross-talk noise or multiple imprints, the observed data events caused by multiple scattering were required to be detected and eliminated prior to the migration process using a multiple elimination algorithm (Berkhout and Verschuur, 1997; Weglein et al., 1997; Weglein et al., 2003; Berkhout and Verschuur, 2005; Groenestijn and Verschuur, 2009; Ypma and Verschuur, 2013; Slob et al., 2014; Siahkoohi et al., 2019; Zhang and Slob, 2020; Thorbecke et al., 2021). While eliminating multiple scattering offers benefits, it can also be a challenging task or exceed budgetary limits in some seismic applications. Therefore, from this perspective, it would be ideal if a migration algorithm could effectively manage multiple scattering together with first-order scattering waves.

Some studies have shown that when a migration algorithm effectively manages multiple scattering, it can enhance illumination, especially in areas where primary imaging fails to provide adequate illumination due to missing/masked-by-multiple primary reflections in the observed data (Berkhout and Verschuur, 2016; Davydenko and Verschuur, 2017; Lu et al., 2018; Slob et al., 2021). Moreover, when reflectors are better focused from early iterations, through an effective gradient preconditioning, multiple scattering modeling will be more accurate, which should result in a better reduction of cross-talk noise on the final stacked image. This is indeed expected as the outcome of a clearer distinction between the primary and multiple reflections during data matching, especially when they overlap.

Taking into account interbed multiples in modeling, the following example first examines the successful application (enhanced resolution and stronger reflectivities) and subsequently investigates the expected improvement (better cross-talk noise reduction), if any, of the suggested depth-based preconditioner within the context of the least-squares full-wavefield migration (LS-FWM) algorithm, which was first introduced by Berkhout (2014a).

LS-FWM is a non-linear least-squares one-way wave-equation migration technique that uses the full (primary and multiples) two-way Earth's response by incorporating multiple scattering (via iterative modeling) into the forward modeling theory already presented in the forward problem section. LS-FWM and LS-WEM share the same inversion cycle (Figure 2.3), with the difference that LS-FWM uses full-wavefield modeling (FWMod) (Berkhout, 1982; Berkhout, 2014b) rather than PWMod. The LS-FWM gradient vector, similar to the LS-WEM gradient vector, in each iteration is scaled by the diagonal components of the inverse Hessian approximation derived by the linear part of equation 2.27.

Angle-independent FWMod models angle-independent primary reflections at the Earth's surface by extrapolating the seismic source signature first downward from z_0 to z_N and then upward from z_N to z_0 by accounting for reflection and transmission effects, which is called one roundtrip/iteration. Subsequent FWMod roundtrips contribute to

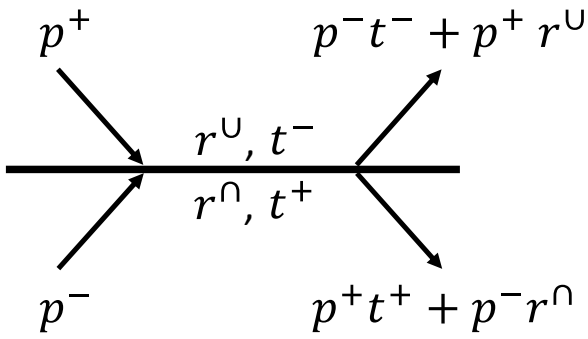


Figure 2.14: Wavefield updating in FWMod with angle-independent reflections at a given depth level, where p^+ is the downgoing wavefield, p^- is the upgoing wavefield, $r^\cup$ represents the upward reflectivity, $r^\cap$ shows the downward reflectivity, t^+ and t^- denote the downward and upward transmission.

new orders of multiple scattering while still explaining reflection and transmission effects. Compared to PWMod, FWMod takes into account the multiple scattering by replacing equation 2.11 with,

$$\mathbf{p}^+_{\text{mod}}(z_m,\omega_f) = \sum_{n=0}^{m-1} \mathbf{U}^+_{z_m;z_n}\Big(\mathbf{s}^+(z_n,\omega_f) + \mathbf{r}^\cap(z_m)\circ\mathbf{p}^-_{\text{mod}}(z_m,\omega_f)\Big), \tag{2.28}$$

where $\mathbf{r}^\cap(z_m)$ represents the angle-independent downward reflectivity vector-operator at z_m, and $\mathbf{r}^\cap = -\mathbf{r}^\cup$ (acoustic medium assumption) (Figure 2.14).

As is clear, FWMod in its first roundtrip calculates the Earth's two-way response (primary reflections) via a recursive summation in depth by including reflection and transmission effects. FWMod has control over multiple scattering, and each roundtrip (other than the first roundtrip) adds an order of scattering to the primaries. As a result, with FWMod, modeling primary reflections, surface-related multiples, interbed multiples, and total reflection wavefields are easily accessible.

PLS-FWM minimizes the same error functional as PLS-WEM, i.e., equation 2.16, with the difference that $\mathbf{p}^-_{\text{obs}}$ and $\mathbf{p}^-_{\text{mod}}$ in the equation are now the full (primaries and multiples) observed and modeled data at the Earth's surface. Other details of the method can be found in Berkhout (2014a). As a result, for deriving the suggested preconditioned model update in the PLS-FWM framework, nothing mathematically changes from what is already derived for PLS-WEM, as FWMod can be seen as an iterative PWMod, where the upgoing/downgoing wavefields stored at different depth levels in former roundtrips are used as the input wavefields for the subsequent roundtrips. Therefore, the same theory is here employed to build the preconditioned model update for PLS-FWM in each iteration.

2.5. NUMERICAL EXAMPLES OF PLS-WEM

To evaluate the application and impact of the suggested depth-based preconditioner within the framework of least-squares full-wavefield migration, we use a multi-reflector wedge model placed between two horizontal reflectors, in which each reflector is only due to the density contrast, as we utilize a homogeneous velocity of 2000 m/s to build the

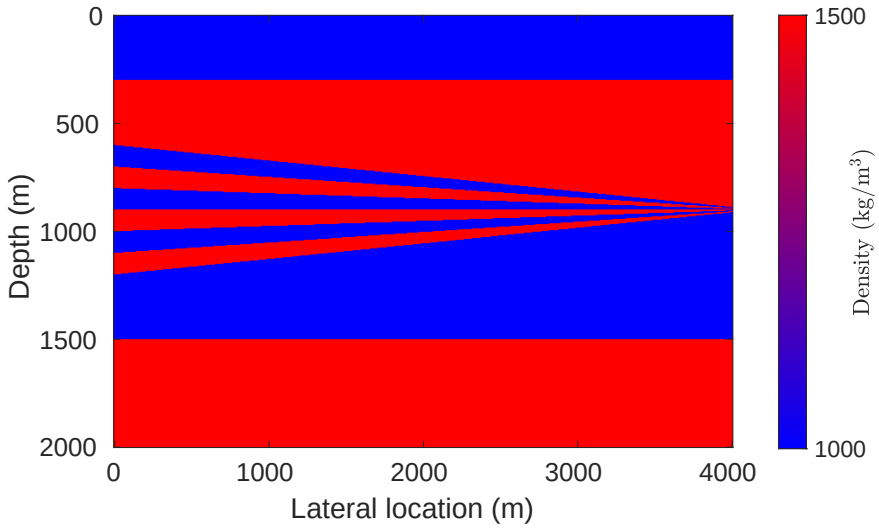
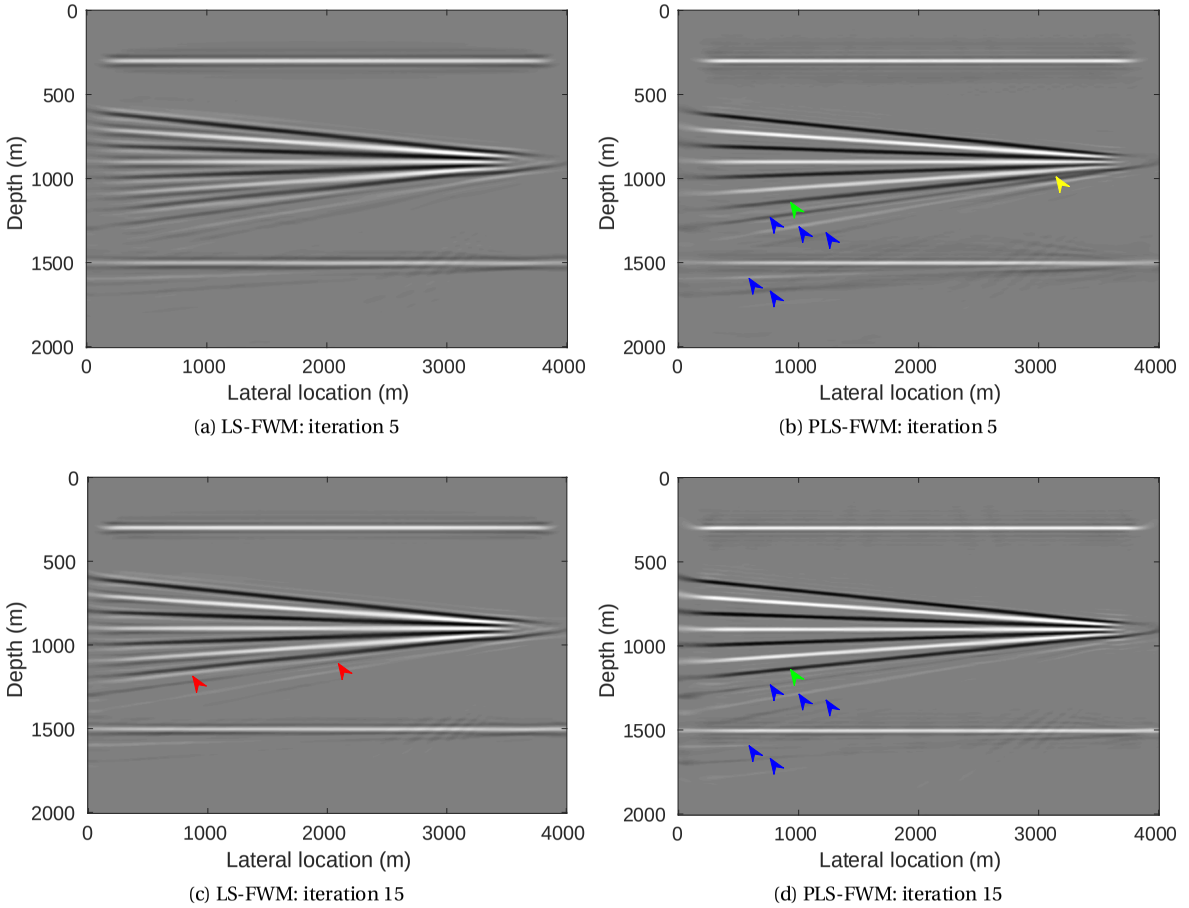


Figure 2.15: Variations in the true resonant wedge density model in kg/m^3. The change in density between the layers is 50%, and the dipping layers all have identical thicknesses.

reflectivity model. This extreme example yields alternating local reflection coefficients of $r = 1/5$ and $r = -1/5$ in the model, causing medium-strength interbed multiples in the data compared to the reflection coefficients used in Slob et al. (2021). Figure 2.15 shows the true resonant wedge density model in kg/m^3. To add the resonant feature to the model, the dipping layers in the wedge part all come in the same thickness. Besides, while each layer on the left side has a thickness of 100 m, the thickness approaches zero on the right. An acoustic finite-difference modeling scheme with a grid size of 1 m generates the observed reflection data. As a source wavelet, a 15 Hz Ricker wavelet is used. 31 shots with a 120 m shot spacing are deployed on top of the model such that the first shot is placed at 200 m and the last shot at 3800 m. Each observed shot gather includes 401 traces with 10 m trace spacing (fixed-spread acquisition), and the length of the traces is 2.92 s. With the same source-receiver setup and the true homogeneous velocity model, FWMod with a horizontal grid size of 10 m and a vertical grid size of 4 m generates the modeled reflection data in every iteration. The maximum frequency used in the migration is 40 Hz, and the largest offset used in the migration is 4 km. Note that in Slob et al. (2021) it is already shown that such a model can be viewed as an extreme scenario when interbed multiples mask primaries. With this experimental setup, LS-FWM and PLS-FWM are employed here to assess the extent of achievement of the reciprocal of the suggested depth-based Hessian operator in preconditioning the gradient direction within the context of full-wavefield migration. Figure 2.16 displays the results.

Figures 2.16a and 2.16b show the LS-FWM and PLS-FWM output images after 5 iterations, respectively. For the purpose of analysis, the output images can be laterally split into three main regions: the first 1000 m of the lateral distance of the image (left-edge region), from 1000 to 3000 m of the lateral distance of the image (middle region), and the last 1000 m of the lateral distance of the image (right-edge region). For the left-



(a) LS-FWM: iteration 5

(b) PLS-FWM: iteration 5

(c) LS-FWM: iteration 15

(d) PLS-FWM: iteration 15

Figure 2.16: Estimated reflectivity models using the LS-FWM and PLS-FWM methods after 5 and 15 iterations, associated with the true resonant wedge density model shown in Figure 2.15. (a) and (c) Estimated reflectivity models by LS-WEM after 5 and 15 iterations, respectively. (b) and (d) Estimated reflectivity models by PLS-WEM after 5 and 15 iterations, respectively.

edge region as deep as the dipping layers, LS-FWM delivers an alternating sequence of positive and negative amplitudes. Consequently, it is challenging to determine which ones represent the real dipping reflectors. In contrast, by removing the source signature and balancing the reflector amplitudes, PLS-FWM delivers a clean view of the real dipping reflectors there. For the middle region as deep as the dipping layers, we notice that PLS-FWM compared to LS-FWM, effectively performs better and maps the reflectors with less blurriness. For the right-edge region as deep as the dipping layers, where the resolution limit is reached, we notice that for recovering the fifth dipping reflector (yellow arrow), while PLS-FWM does not stably function, LS-FWM functions slightly more stable.

We can also see that after 5 iterations, both images exhibit a common characteristic, that is, the cross-talk noise or equally ghost reflectors generated in the image owing to the presence of interbed multiples in the data. As a result, the last real dipping reflector is not retrieved as strongly as the others, as denoted by the green arrow in Figure 2.16b, and also several ghost reflectors are built, as indicated by the blue arrows in Figure 2.16b.

Figures 2.16c and 2.16d display the LS-FWM and PLS-FWM results after 15 iterations, respectively. Comparing the images reveals that the ghost reflectors existing in the fifth-iteration images are now attenuated after 15 iterations as shown by the blue arrows in Figure 2.16d. In addition, we observe that the weak real dipping reflector in the fifth-iteration images is now recovered much stronger after addressing both primaries and interbed multiples (see also Slob et al. (2021)), as denoted by the green arrows in Figure 2.16d, but still not as strong as the other real dipping reflectors. We also notice that the ghost reflectors are marginally better reduced by PLS-WEM, as denoted by the red arrows in Figure 2.16c; see also the discussion section.

Figure 2.17 compares the 1D reflectivity profiles corresponding to the LS-FWM and PLS-FWM images in iteration 15, depicted in Figures 2.16c and 2.16d, at the lateral location of 2150 m. It clearly proves the accuracy of the PLS-FWM image in comparison to the LS-FWM image in terms of the resolved resolution, reflectivity magnitudes, and reduction of multiple cross-talk (red arrows). Therefore, we can state that PLS-FWM outperforms LS-FWM in terms of image quality.

2.6. DISCUSSION

Despite the limitations outlined in the introduction, the industry has continued to embrace the practice of mapping primary energies back into the image space using LS-WEM due to its notable advantage of lower computational cost when compared to LS-RTM. Although LS-WEM has proven itself quite effective in recovering decent images, we observed that our newly developed PLS-WEM algorithm outperforms LS-WEM. It shows faster convergence and superior resolution. Moreover, it is efficient in computation when compared to an alternative solution that computes a huge Hessian approximation operator for the entire medium. While feasible, it still remains time-wise expensive compared to LS-WEM, with the current implementation and contemporary computing resources. However, based on our internal investigations, PLS-WEM with source subsampling in the computation of Hessian approximation operators (Matharu and Sacchi, 2019)—one source out of three—can generate similar high-resolution images, comparable to this study, in half the iterations of LS-WEM, although its wall time for each iteration is observed to be double to triple.

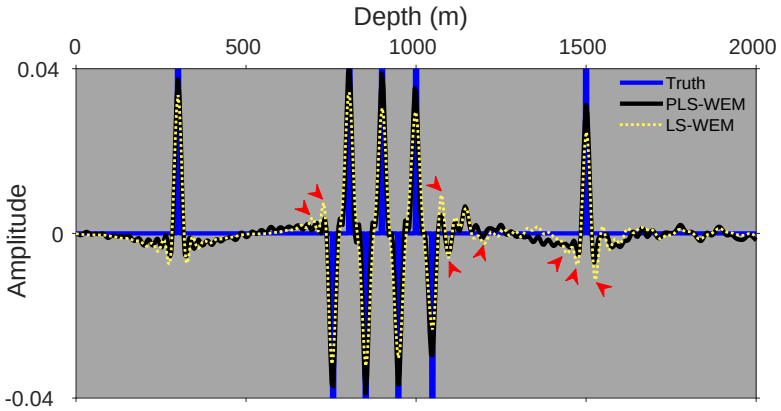


Figure 2.17: Estimated 1D reflectivity profiles using the LS-FWM and PLS-FWM methods after 15 iterations, associated with the resonant wedge model. The profiles are taken from the estimated reflectivity models shown in Figures 2.16c and 2.16d and represent the lateral location of 2150 m.

It is important to acknowledge that both LS-WEM and PLS-WEM algorithms presented in this paper may underperform due to flaws in the multiple elimination stage, poor wavelet estimation, the angle-independent reflection coefficient assumption, steep reflectors, and an inaccurate migration velocity model.

In the case of the full-wavefield migration example, while we expected a better resolution and stronger reflection coefficients from PLS-FWM compared to LS-FWM, we also hoped for more reduction of cross-talk noise. According to the resonant wedge example results, even though PLS-FWM achieves a greater resolution and stronger reflectivities, the cross-talk noise is not better reduced. As previously stated, better-focused reflectors are expected to end in more accurate modeling of multiple scattering and hence better attenuation of multiple imprints. However, this doesn't seem to work as well as we thought it would, although it gave slight improvements in reducing the cross-talk noise, see the red arrows on Figures 2.16c and 2.17. This capacity may be partially underused due to the double-edged nature of preconditioning; the preconditioner does its function on both real and ghost reflectors in each iteration, giving high-resolution ghost and real reflectors for modeling, leading to a conflict between real and fake reflections in the corresponding data minimization problem, still ending up in a local minimum. From another perspective, the underperformance could also be attributed to the highly nonlinear nature of the resonant wedge experiment. Lastly, the underperformance might be linked to the FWMod convergence issue with spatially inhomogeneous media (McMaken, 1986). To ascertain the main cause, further investigation must be conducted.

Typically, the convergence of LSM exhibits a faster rate in the initial iterations and a slower rate in the subsequent iterations. In Figures 2.12*a* and 2.12*b*, it can be observed that the data error between the second and third iterations remains almost identical, indicating a minimal change in the image update. The observed phenomenon can be attributed to the nonlinear dynamics introduced into the modeling process by the implicit accounting for the transmission effects in the modeled data, as shown by the factor: $\left(1+r^{\cup}\right)\left(1-r^{\cup}\right)=1-\left(r^{\cup}\right)^2$. In the first iteration of LSM, there is no information

about the reflection coefficients. This lack of information may lead to an overcorrection of the transmission coefficients with the nonlinear factor, resulting in modeled data that is not amplitude-consistent with the observed data. In order to correct this, the next iteration of LSM performs data amplitude balancing, resulting in a relatively similar data error when compared to the previous iteration. With each subsequent iteration, this issue gradually irons itself out, bringing the modeled data into close alignment with the observed data.

Superior imaging can also help build high-resolution and strong tomographic updates in reflection waveform inversion (RWI). RWI is described as a method for constructing a low-wavenumber velocity model via confining the full waveform inversion sensitivity kernel along the transmission-after-reflection wavepaths. RWI sequentially solves a multi-parameter primary-reflection-data-driven inverse problem with velocity and reflectivity as the parameters. In RWI, reflectors are first mapped using a migration technology, and then the tomographic update is built based on the mapped reflectors. As a result, in such a flow, a higher-resolution image with stronger amplitudes may lead to a more consistent and stronger tomographic update, something that our proposed PLS-WEM can facilitate.

2.7. CONCLUSION

We mathematically showed how the Hessian approximation operator and its reciprocal as the gradient preconditioner could be efficiently built for LS-WEM in a depth marching regime. We showed that our proposed cost-friendly PLS-WEM algorithm, thanks to PWMod, comes with a minimal computational effort in each iteration in comparison to its alternatives that compute a massive Hessian approximation for the entire medium. Using two numerical examples, we verified how the proposed preconditioner effectively cut the migration artifacts generated by the band-limited nature of seismic data and preserves the reflectivity amplitudes. We also confirmed that the improvement in migration images results in a faster convergence ratio and a better data fit for PLS-WEM. We also used the proposed preconditioner in the context of full-wavefield migration. With an extreme scenario generating strong interbed multiples, we showed that PLS-FWM outperforms LS-FWM in the same two ways that PLS-WEM outperforms LS-WEM. That is, PLS-FWM removes the source signature from the image quickly and recovers stronger and more accurate reflectivities than LS-FWM. We also observed that PLS-FWM only provided limited improvement in minimizing the cross-talk noise, contrary to our expectations.

REFERENCES

Abolhassani, S. and E. Verschuur (2022). “Fast Gauss-Newton full-wavefield migration”. In: SEG/AAPG Second International Meeting for Applied Geoscience & Energy, Expanded Abstracts., pp. 2709–2713.

Aminzadeh, F. et al. (1994). “SEG/EAGE 3-D modeling project: 2nd update”. In: *The Leading Edge* 13.9, pp. 949–952.

Aoki, N. and G. T. Schuster (2009). “Fast least-squares migration with a deblurring filter”. In: *Geophysics* 74 (6), WCA83–WCA93.

Assis, C. A. and J. Schleicher (2021). "Introduction of the Hessian in joint migration inversion and improved recovery of structural information using image-based regularization". In: *Geophysics* 86 (6), R777–R793.

Baysal, E., D. D. Kosloff, and J. W. C. Sherwood (1983). "Reverse time migration". In: *Geophysics* 48.11, pp. 1514–1524.

Berkhout, A. J. (1982). *Seismic migration, imaging of acoustic energy by wave field extrapolation, A: theoretical aspects*. Elsevier.

— (2014a). "Review paper: An outlook on the future of seismic imaging, part II: Full-wavefield migration". In: *Geophysical Prospecting* 62.5, pp. 931–949.

— (2014b). "Review paper: An outlook on the future seismic imaging, part I: forward and reverse modelling". In: *Geophysical Prospecting* 62.5, pp. 911–930.

Berkhout, A. J. and D. J. Verschuur (1997). "Estimation of multiple scattering by iterative inversion, part I: Theoretical considerations". In: *Geophysics* 62.5, pp. 1586–1595.

— (2005). "Removal of internal multiples with the common-focus-point (CFP) approach: Part 1 — Explanation of the theory". In: *Geophysics* 70 (6), pp. V45–V60.

— (2016). "Enriched seismic imaging by using multiple scattering". In: *The Leading Edge* 35.2, pp. 128–133.

Berkhout, A. J. and C. P. A. Wapenaar (1989). "One-way versions of the Kirchhoff integral". In: *Geophysics* 54.4, pp. 460–467.

Berkhout, A. J. and P. V. Wulfften (1979). "Migration in terms of Spatial Deconvolution". In: *Geophysical Prospecting* 27.1, pp. 261–291.

Beydoun, W. B. and M. Mendes (1989). "Elastic ray-Born l_2-migration/inversion". In: *Geophysical Journal International* 97.1, pp. 151–160.

Bleistein, N., J. Cohen, and J. Stockwell (2001). *Mathematics of Multidimensional Seismic Imaging and Inversion*. Springer.

Chauris, H. and E. Cocher (2017). "From migration to inversion velocity analysis". In: *Geophysics* 82 (3), S207–S223.

Chavent, G. and R. E. Plessix (1999). "An optimal true-amplitude least-squares prestack depth-migration operator". In: *Geophysics* 64.2, pp. 508–515.

Claerbout, J. F. (1971). "Towards a unified theory of reflector mapping". In: *Geophysics* 36.3, pp. 467–481.

Claerbout, J. F. and S. Doherty (1972). "Downward continuation of moveout-corrected seismogram". In: *Geophysics* 37.5, pp. 741–768.

Cole, S. and M. Karrenbach (1992). "Least-squares Kirchhoff migration". In: *Stanford Exploration Project (SEP) Report*.

Cova, R. and K. A. Innanen (2013). "Approximate- vs. full-Hessian in FWI: 1D analytical and numerical experiments". In: *CREWES Research Report*.

Davydenko, M. and D. J. Verschuur (2017). "Full-wavefield migration: Using surface and internal multiples in imaging". In: *Geophysical Prospecting* 65.1, pp. 7–21.

Deng, F. and G. A. McMechan (2007). "True-amplitude prestack depth migration". In: *Geophysics* 72 (3), S155–S166.

Duquet, B. and K. J. Marfurt (1999). "Filtering coherent noise during prestack depth migration". In: *Geophysics* 64.4, pp. 1054–1066.

Etgen, J., S. H. Gray, and Y. Zhang (2009). "An overview of depth imaging in exploration geophysics". In: *Geophysics* 74 (6), WCA5–WCA17.

Fletcher, R. P. et al. (2016). “Least-squares migration — Data domain versus image domain using point spread functions”. In: *The Leading Edge* 35.2, pp. 157–162.
Gazdag, J. (1978). “Wave equation migration with the phase-shift method”. In: *Geophysics* 43.7, pp. 1342–1351.
Groenestijn, G. J. van and D. J. Verschuur (2009). “Estimating primaries by sparse inversion and application to near-offset data reconstruction”. In: *Geophysics* 74 (3), A23–A28.
Guitton, A. (2004). “Amplitude and kinematic corrections of migrated images for nonunitary imaging operators”. In: *Geophysics* 69.4, pp. 1017–1024.
Hou, J. and W. Symes (2016). “Approximate Gauss-Newton iteration for full-waveform inversion”. In: SEG Annual International Meeting, Expanded Abstracts., pp. 1163–1168.
Hou, J. and W. W. Symes (2015). “An approximate inverse to the extended Born modeling operator”. In: *Geophysics* 80 (6), R331–R349.
Hu, J., G. T. Schuster, and P. Valasek (2001). “Poststack migration deconvolution”. In: *Geophysics* 66.3, pp. 939–952.
Huang, Y. et al. (2014). “Making the most out of least-squares migration”. In: *The Leading Edge* 33.9, pp. 954–960.
Jang, U., D.-J. Min, and C. Shin (2009). “Comparison of scaling methods for waveform inversion”. In: *Geophysical Prospecting* 57.1, pp. 49–59.
Jones, I. F. (2014). “Tutorial: Migration imaging conditions”. In: *First Break* 32.12, pp. 45–55.
— (2018). *Velocities, imaging, and waveform inversion*. EAGE.
Kiyashchenko, D. et al. (2007). “A modified imaging principle for true-amplitude wave-equation migration”. In: *Geophysical Journal International* 168.3, pp. 1093–1104.
Lecomte, I. (2008). “Resolution and illumination analyses in PSDM: A ray-based approach”. In: *The Leading Edge* 27.5, pp. 650–663.
Lines, L. and S. Treitel (1984). “A review of least-squares inversion and its application to geophysical problems”. In: *Geophysical Prospecting* 32.2, pp. 159–186.
Lu, S. et al. (2018). “Least-squares full-wavefield migration”. In: *The Leading Edge* 37.1, pp. 46–51.
Matharu, G. and M. Sacchi (2019). “A subsampled truncated-Newton method for multiparameter full-waveform inversion”. In: *Geophysics* 84.3, R333–R340.
McMaken, H. (1986). “On the convergence of the Bremmer series for the Helmholtz equation in 2-D”. In: *Wave Motion* 8.3, pp. 277–283.
McMechan, G. A. (1983). “Migration by extrapolation of time-dependent boundary values”. In: *Geophysical Prospecting* 31.3, pp. 413–420.
Mehta, K. et al. (2017). “One-Way wave equation migration at scale on GPUs using directive based programming”. In: IEEE International Parallel and Distributed Processing Symposium, pp. 224–233.
Metivier, L. et al. (2014). “Full waveform inversion and the truncated Newton method: Quantitative imaging of complex subsurface structures”. In: *Geophysical Prospecting* 62.6, pp. 1353–1375.
Mulder, W. A. and R. E. Plessix (2004). “A comparison between one-way and two-way wave-equation migration”. In: *Geophysics* 69.6, pp. 1491–1504.

Nemeth, T., C. Wu, and G. T. Schuster (1999). "Least-squares migration of incomplete reflection data". In: *Geophysics* 64.1, pp. 208–221.

Oh, J.-W. and D.-J. Min (2013). "Weighting technique using backpropagated wavefields incited by deconvolved residuals for frequency-domain elastic full waveform inversion". In: *Geophysical Journal International* 194.1, pp. 322–347.

Operto, S. et al. (2013). "A guided tour of multiparameter full-waveform inversion with multicomponent data: From theory to practice". In: *The Leading Edge* 32.9, pp. 1040–1054.

Plessix, R. E. and W. A. Mulder (2004). "Frequency-domain finite-difference amplitude-preserving migration". In: *Geophysical Journal International* 157.3, pp. 975–987.

Pratt, R. G., C. Shin, and G. J. Hicks (1998). "Gauss–Newton and full Newton methods in frequency-space seismic waveform inversion". In: *Geophysical Journal International* 133.2, pp. 341–362.

Schuster, G. T. (2017). *Seismic inversion*. SEG.

Schuster, G. T. (2002). "Reverse-time migration= generalized diffraction stack migration". In: SEG Annual International Meeting, Expanded Abstracts., pp. 1280–1283.

Shin, C., C. S. Jang, and D.-J. Min (2001). "Improved amplitude preservation for prestack depth migration by inverse scattering theory". In: *Geophysical Prospecting* 49.5, pp. 592–606.

Siahkoohi, A., D. J. Verschuur, and F. J. Herrmann (2019). "Surface-related multiple eliminations with deep learning". In: SEG Annual International Meeting, Expanded Abstracts., pp. 4629–4634.

Slob, E. et al. (2014). "Seismic reflector imaging using internal multiples with Marchenko-type equations". In: *Geophysics* 79 (2), S63–S76.

Slob, E., L. Zhang, and E. Verschuur (2021). "Marchenko multiple elimination and full-wavefield migration in a resonant pinch-out model". In: *Geophysics* 86 (5), WC1–WC9.

Stolt, R. H. (1978). "Migration by Fourier transform". In: *Geophysics* 43.1, pp. 23–48.

Tarantola, A. (1984). "Linearized inversion of seismic reflection data". In: *Geophysical Prospecting* 32.6, pp. 998–1015.

ten Kroode, F. (2012). "A wave-equation-based Kirchhoff operator". In: *Inverse Problems* 28.11, p. 115013.

Thorbecke, J. et al. (2021). "Implementation of the Marchenko multiple elimination algorithm". In: *Geophysics* 86 (2), F9–F23.

Thorbecke, J. W., K. Wapenaar, and G. Swinnen (2004). "Design of one-way wavefield extrapolation operators, using smooth functions in WLSQ optimization". In: *Geophysics* 69.4, pp. 1037–1045.

Valenciano, A. A., B. Biondi, and A. Guitton (2006). "Target-oriented wave-equation inversion". In: *Geophysics* 71 (4), A35–A38.

Valenciano, A. A. and B. Biondi (2003). "2-D deconvolution imaging condition for shot-profile migration". In: SEG Annual International Meeting, Expanded Abstracts., pp. 1–5.

Weglein, A. B. et al. (1997). "An inverse-scattering series method for attenuating multiples in seismic reflection data". In: *Geophysics* 62.6, pp. 1975–1989.

Weglein, A. B. et al. (2003). “Inverse scattering series and seismic exploration”. In: *Inverse Problems* 19 (6), R27–R83.
Whitmore, N. D. (1983). “Iterative depth migration by backward time propagation”. In: SEG Annual International Meeting, Expanded Abstracts., pp. 382–385.
Yang, J. et al. (2021). “Approximating the Gauss-Newton Hessian using a space-wavenumber filter and its applications in least-squares seismic imaging”. In: *IEEE Transactions on Geoscience and Remote Sensing* 60, pp. 1–13.
Ypma, F. H. C. and D. J. Verschuur (2013). “Estimating primaries by sparse inversion, a generalized approach”. In: *Geophysocal Prospecting* 61.1, pp. 94–108.
Yu, J. et al. (2006). “Prestack migration deconvolution”. In: *Geophysics* 71 (2), S53–S62.
Zhan, G. et al. (2014). “Generalized diffraction-stack migration and filtering of coherent noise”. In: *Geophysical Prospecting* 62.3, pp. 427–442.
Zhang, L. and E. Slob (2020). “A field data example of Marchenko multiple elimination”. In: *Geophysics* 85 (2), S65–S70.
Zhang, Y. et al. (2007). “True-amplitude, angle-domain, common-image gathers from one-way wave-equation migrations”. In: *Geophysics* 72 (1), S49–S58.

3

IMPROVED ONE-WAY REFLECTION WAVEFORM INVERSION

"Life can only be understood backwards; but it must be lived forwards."

– SOREN KIERKEGAARD

Conventional reflection waveform inversion solves a two-parameter seismic inverse problem alternately for subsurface reflectivity and acoustic background velocity as the model parameters. It seeks to reconstruct a low-wavenumber velocity model of the subsurface from pure reflection data cyclically, through alternating migration and tomography loops, such that the remodeled data fits the observed data. Low-resolution seismic images with unpreserved amplitudes, cycle skipping in long-offset data, and full-wave inconsistency in the short-offset data due to inconsistent reflectivity and velocity models are perceived as the main reasons for poor tomographic updates and slow convergence in conventional reflection waveform inversion. In the context of one-way reflection waveform inversion, this paper addresses the listed limitations through four main components. First, it augments one-way reflection waveform inversion with a computationally affordable preconditioned least-squares wave-equation migration algorithm, to ensure high-resolution reflectors with preserved amplitudes. Second, the paper verifies how well the full-wave consistency condition in the short-offset data is satisfied in one-way reflection waveform inversion and suggests muting inconsistent short-offset residual waveforms in the tomography loop to attenuate their adverse imprint. Third, with short-offset residual waveforms muted, the paper suggests extending the migration offset beyond short offsets to improve both the illumination and the signal-to-noise ratio of the reflectors. Fourth, the paper presents a data-selection algorithm to exclude the damaging effect of the cycle-skipped long-offset data in the tomography loop. The effectiveness of the proposed one-way reflection waveform inversion algorithm is finally validated through three numerical examples, demonstrating its capability to recover high-fidelity tomograms.

This chapter is based on the following manuscript: **Abolhassani, S.**, Verschuur, D. J. (2024) Improved One-Way Reflection Waveform Inversion and Strategies for Optimal Offset Selection. Submitted to *Geophysical Prospecting*, GP-2024-4391.R3

3.1. INTRODUCTION

Since the early 1980s, multiple efforts have been made by geoscientists to recover high-resolution subsurface models (e.g., velocity, density) by explaining observed seismic waveforms. In a big step forward, Tarantola (1984) introduced full waveform inversion (FWI) in the acoustic approximation. Conventional FWI solves a minimization problem by defining an L2-norm misfit function that measures the difference between the modeled and observed seismic waveforms to reconstruct high-resolution subsurface models. While conventional FWI has shown great success in reconstructing high-fidelity shallow subsurface models described by the field data (e.g., Vigh et al., 2011; Liu et al., 2012; Plessix et al., 2013), it has shown incapable of retrieving deep subsurface targets, where refracted and diving waves do not reach due to the limited offset range in the acquired data (e.g., Irabor and Warner, 2016; Vigh et al., 2016; Jones, 2019; Brittan and Jones, 2019). Given the technical barriers of acquiring ultra-long-offset seismic data, the expenses of modern seismic data acquisition equipment for acquiring long offsets, physical obstacles in acquisition areas, and short-offset legacy data, conventional FWI cannot always provide the optimum outcome.

While observed refracted and diving waves suffer from limited penetration depth for a given offset, the reflected waves can penetrate deeper. To benefit from pure reflection data and illuminate deep targets, Mora (1989) elaborated on the significance of scale separation, separating the retrieval of high-reflective local model wavenumbers (migration process, also called imaging) from the retrieval of low-propagative local model wavenumbers (tomography process). After Mora (1989), multiple variations of migration-based velocity analysis (MVA) relying on model scale separation have been developed. Notably, among these variants, those incorporating wave-equation forward modeling (Born approximation) have emerged as significant due to their enhanced handling of wave propagation in complex geological environments. Specifically, wave-equation MVA (WEMVA) algorithms, as introduced by Sava and Biondi (2004), typically aim to optimize an image-domain error function computed from common image gathers (CIGs). This optimization is achieved using either the classical semblance or a differential semblance (DS) function, which evaluates the flatness or coherency of the CIGs (Chauris and Noble, 2001; Shen et al., 2003). WEMVA can also be expressed within the extended-image framework as described by Shen (2005) and further explored by Lameloise et al. (2015). Shen and Symes (2008) demonstrated the use of the DS function in conjunction with subsurface-offset CIGs. Later, Symes (2008) developed the theory of a joint framework for combining WEMVA and FWI, known as tomographic full-waveform inversion (TFWI). In the following years, Almomin and Biondi (2012), Biondi and Almomin (2012), and Sun and Symes (2013) practically showed that TFWI works, although expensively due to the extended-domain imaging condition employed and the number of the Green's functions involved.

In the meanwhile and with the same thinking as WEMVA, several reflection waveform inversion (RWI) techniques also originated by combining wave-equation migration and FWI into a single framework (e.g., Xu et al., 2012a; Zhou et al., 2012). RWI, like WEMVA, relies on the idea of model scale separation; however, such a separation between the migration and tomography sensitivity kernels in waveform inversion has always been challenging. So far, there are three main alternatives to address the challenge: (a)

scattering-angle filtering (e.g., Alkhalifah, 2014; Xie, 2015; Kazei et al., 2016; Wu and Alkhalifah, 2017; Yao et al., 2018; Yao et al., 2019a), (b) wavefield decomposition (e.g., Liu et al., 2011; Wang et al., 2013; Tang et al., 2013; Irabor and Warner, 2016; Chi et al., 2017; Lian et al., 2018), and (c) Born modeling (e.g., Xu et al., 2012a; Xu et al., 2012b; Zhou et al., 2012; Sun et al., 2016; Vigh et al., 2016; Wang et al., 2018). To improve the RWI gradient, some studies examined merging early-arrival waveform inversion and RWI (Zhou et al., 2015; Wu and Alkhalifah, 2015).

Berkhout (2012) and Berkhout (2014b) introduced an algorithm for reflection waveform inversion—based on wavefield decomposition into upgoing and downgoing waves—and named it joint migration inversion (JMI). Wavefield decomposition in JMI relies on a one-way acoustic wavefield extrapolation scheme known as full wavefield modeling (FWMod), which includes both primaries and controlled-order multiples (Berkhout, 2014a). Angle-independent FWMod is set up on two classes of parameters: acoustic background velocity and angle-independent reflectivity. Due to the adequate parametrization in FWMod, the tomographic and imaging sensitivity kernels in JMI appear independent. Recently, Lu et al. (2018), Dong et al. (2018), Romahn et al. (2021), Hassine et al. (2022), and Soubaras et al. (2023) have also utilized the one-way acoustic wave equation for the purpose of wave-equation migration and reflection waveform inversion. Constraining FWMod to single-scattered waveforms—primary wavefield modeling, in short PWMod—we use the same notion as JMI for reflection waveform inversion here and refer to it as one-way reflection waveform inversion (ORWI). As an inverse problem, ORWI alternately solves a two-parameter minimization problem with a misfit function that measures the reflection data (primaries) error in the least-squares sense. Both classes of parameters in ORWI—background velocity and angle-independent reflectivity—are reconstructed in a cyclic process, including least-squares migration and reflection tomography loops (Algorithm 3.1). Based on a zero initial reflectivity model, a smooth initial background velocity model, and full-offset reflection data in both migration and tomography loops, the flow begins with the migration loop and proceeds to alternate between the loops until reaching convergence with accurate solutions for the background velocity and reflectivity.

Algorithm 3.1: Cyclic workflow of ORWI

```
1 Providing a smooth initial background velocity and zero initial
    reflectivity model when cycle number k = 1.
2 while convergence not reached do
3 |   for i ← 1 to m do
4 |   |   UPDATE the reflectivity model using one-way wave-equation
  |   |     while keeping the background velocity model fixed (migration).
5 |   for j ← 1 to n do
6 |   |   UPDATE the background velocity model using one-way
  |   |     wave-equation while keeping the reflectivity model fixed
  |   |     (reflection tomography).
7 |   k = k + 1   // The reflectivity model (image) can be
  |     reset to zero at this point.
```

Building true amplitude and high-resolution reflectors can considerably help any RWI technique recover strong and in-phase tomographic wavepaths in each cycle (Gomes and Yang, 2018; Chen et al., 2020). However, iterative least-squares migration (LSM) techniques, e.g., least-squares reverse time migration (LS-RTM) and least-squares one-way wave-equation migration (LS-WEM), are often expensive to employ in RWI because LSM requires iterations to be converged optimally. Although the number of iterations can be cut down by preconditioning the gradient direction with the reciprocal of the Hessian information, the Hessian computation turns prohibitively costly when dealing with large-scale seismic problems (Lines and Treitel, 1984). For cost reduction, Beydoun and Mendes (1989), Chavent and Plessix (1999), Shin et al. (2001), and Plessix and Mulder (2004) approximated the diagonal coefficients of the Hessian matrix rather than full coefficients, and others have tried to approximate the complete coefficients at an affordable cost in either data or image domain. Pratt et al. (1998) calculated the Gauss-Newton approximation of the Hessian matrix (also known as the linear or approximate Hessian). Hu et al. (2001), Guitton (2004), and Yu et al. (2006) approximated the inverse Hessian by constructing deconvolution filters in the image domain. For a target-oriented imaging problem, explicit computation of a sparse Hessian matrix via cross-correlation of the source and receiver Green's functions is feasible (Valenciano et al., 2006). Choi et al. (2008) replaced the Hessian with a scaled pseudo-Hessian. Point spread functions have also been used to reduce the relevant cost via computing local Hessian matrices in the image domain (Lecomte, 2008; Fletcher et al., 2016). The truncated-Newton method is also used to implicitly approximate the application of the inverse Hessian on the gradient vector (Metivier et al., 2013; Assis and Schleicher, 2021; Assis et al., 2024). Recently, researchers have also proposed explicit inverse formulas for the one/two-way modeling operators in the high-frequency limit, mostly compensating for geometrical spreading. In this context, Zhang et al. (2007) and Kiyashchenko et al. (2007) proposed true-amplitude wave-equation migration schemes with one-way forward modeling operators. ten Kroode (2012) derived an explicit inversion expression for the Kirchhoff modeling operator, while ten Kroode (2014) and Hou and Symes (2015) derived similar inversion expressions for the Born modeling operator, both in the subsurface-extended domain. In a similar vein as Hou and Symes (2015), Chauris and Cocher (2017) approximated a pseudo-inverse operator for the subsurface-extended Born modeling operator; Chauris and Cocher (2018) presented a comparison of the two methods. A pseudo-inverse operator can also be used as a preconditioner in a gradient-descent inversion scheme to speed up convergence, as demonstrated by Hou and Symes (2016) and Qin and Lambaré (2016). In this paper, to build the approximate Hessian information with affordable computational cost, we use the efficient preconditioned least-squares wave-equation migration (PLS-WEM) algorithm recently introduced by Abolhassani and Verschuur (2024). PLS-WEM constructs the approximate Hessian operator recursively depth by depth, thanks to PWMod. It decomposes and reduces the massive approximate Hessian operator for the entire domain into sub-operators relevant to each depth level. PLS-WEM can recover reflectors to be fit for our purpose in one to five iterations at best.

It is well understood that two-way traveltime varies with both reflector's depth and background velocity, a fundamental problem known as ambiguity in the depth-velocity determination problems. From a traveltime perspective, RWI also aims to resolve this

ambiguity and estimate the subsurface model by measuring how much of the overall traveltime error in the reflection data is due to errors in depth versus errors in background velocity. As a waveform inversion technique, RWI achieves this relying on full-wave consistency between short-offset modeled and observed data (Verschuur et al., 2016). However, since the alternating approach of conventional RWI ignores the fact that a change in the velocity model immediately changes the reflectivity model—by fixing the reflectors' positions in depth while updating the background velocity—the full-wave consistency condition in the short-offset data does not perfectly hold. This issue, referred to hereafter as the imprint of inconsistent reflectivity and velocity models, is also known as reflectivity-velocity coupling, phase ambiguity in the short-offset data, and reflectivity-velocity conflict (Baina and Valensi, 2018; Audebert and Cocher, 2020; Valensi and Baina, 2021; Provenzano et al., 2023). To address the imprint of inconsistent reflectivity and velocity in RWI, several approaches have been suggested so far. Many have minimized the imprint of inconsistent reflectivity and velocity by using migration with short-offset data or even near-zero offset data (e.g., Zhou et al., 2015; Guo and Alkhalifah, 2017; Wang et al., 2018; Li et al., 2019; Yao et al., 2020). Another solution is pseudo-time domain RWI, where reflectors' positions are calculated in vertical time and are less dependent on background velocity (e.g., Qu et al., 2020; Provenzano et al., 2023). Recently, Chen et al. (2020) used a rolling-offset strategy to deal with the imprint, Valensi and Baina (2021) proposed to ensure full-wave consistency in the zero-offset data by accounting for the reflectivity-velocity consistency in the RWI tomographic gradient, and Liang et al. (2022) proposed replacing the conventional image in RWI with an image extracted from stacked, flattened common image gathers. This paper, first investigates how effectively the full-wave consistency condition holds for the short-offset reflcetion data within the context of ORWI, and then suggests to mute the short-offset inconsistent waveforms in the residual data domain to minimize the imprint of inconsistent reflectivity and velocity on the ORWI tomographic gradient. Relying on the mitigation of imprints of inconsistent reflectivity and velocity models, this paper also proposes extending the migration offset beyond near-zero or short offsets to avoid limitations in the reflectors' illumination and amplitudes. The extended migration offset is defined as the offset at which the misfit function no longer significantly decreases. This offset is referred to as MEMO, standing for maximum effective migration offset.

Cycle skipping (Virieux and Operto, 2009) in long-offset data due to a poor initial background velocity model is perceived as another hurdle in RWI. Building a kinematic-accurate initial background velocity model, multi-scaling (Bunks et al., 1995; Pratt et al., 1998), and adopting alternative misfit functions rather than traditional waveform-based misfit functions have been the most common solutions. Notable among the alternative misfit functions are traveltime-based misfit functions (Luo and Schuster, 1991; Hoop and Der Hilst, 2005; van Leeuwen and Mulder, 2010; Ma and Hale, 2013; Luo et al., 2016), unwrapped-phase misfit functions (Choi and Alkhalifah, 2013; Choi and Alkhalifah, 2014; Choi et al., 2015), envelope-based misfit functions (Bozdağ et al., 2011; Wu et al., 2014; Chi et al., 2014; Chen et al., 2018), optimal-transport-based misfit functions (Métivier et al., 2016; Yang and Engquist, 2018), and the normalized integration method (Chauris et al., 2012; Donno et al., 2013). In this paper, we first illustrate the detrimental contribution of cycle-skipped long-offset data on the tomographic gradient of ORWI. We then introduce,

inspired by van Leeuwen and Mulder (2010), a time-domain two-step data-selection algorithm excluding the contribution of the cycle-skipped long-offset residual data in each iteration of the tomography loop to obtain a high-fidelity tomographic update. The algorithm relies on global and local cross-correlations between the observed and modeled reflection data in the time domain.

This paper is ordered as follows. After the introduction, we first review the theoretical features of ORWI and reintroduce the PLS-WEM technology to replace the migration core of ORWI. Next, we examine how far the full-wave consistency condition is supported in the context of ORWI. Following this, we explain how we deal with the conflict of reflectivity and velocity in reflection waveform inversion. We then introduce the concept of MEMO and its significance. We also present a workflow to extract the uncycle-skipped part of the long-offset reflection data in each tomography loop. We evaluate our proposal through three numerical examples: two dipping layers placed in a linear background velocity model, the flat section of the Marmousi model, and the faulted section of the Marmousi model. We end with a discussion and conclusions.

3.2. STANDARD ONE-WAY REFLECTION WAVEFORM INVERSION (ORWI)

ORWI is an acoustic reflection waveform tomography tool working based on wavefield decomposition into upgoing and downgoing waves. Wavefield decomposition in ORWI relies on a one-way wavefield extrapolation scheme known as PWMod (Berkhout, 2014a), based on first-order scattering and also informed of transmission effects. PWMod perceives wave propagation as a one-way extrapolation from multiple points at z_ℓ to a point at $z_{\ell+1}$, where z_ℓ and $z_{\ell+1}$ denotes two subsequent virtual depth levels, respectively. The absolute value of the vertical distance between z_ℓ and $z_{\ell+1}$ is called the extrapolation step and is considered small enough to be assumed homogeneous vertically. The upward and downward data extrapolations are performed using extrapolation kernels, which are the analytical solutions to the acoustic wave equation in a homogeneous medium in the frequency-wavenumber domain. PWMod defines the 2D forward extrapolated wavefield from all lateral positions located at z_ℓ towards a lateral position located at $z_{\ell+1}$ in the +z direction as a frequency-space convolution integral along the x-axis (Rayleigh integral II in 2D)

$$p^+(x_j, z_{\ell+1}, \omega_f) = \int_{-\infty}^{+\infty} p^+\left(x, z_\ell, \omega_f\right) w^+_{z_{\ell+1};z_\ell}\left(x_j; x_j - x, \omega_f\right) dx, \tag{3.1}$$

in which $w^+_{z_{\ell+1};z_\ell}$ denotes the monochromatic downward wavefield extrapolation kernel from the virtual depth level z_ℓ to one lateral position at the virtual depth level $z_{\ell+1}$, the index j means the lateral position at $z_{\ell+1}$, x_j is the convolution lag, ω_f is a given angular frequency, $p^+(x, z_\ell, \omega_f)$ is the monochromatic downgoing wavefield located at z_ℓ, and $p^+(x_j, z_{\ell+1}, \omega_f)$ is the monochromatic downgoing wavefield at the lateral position x_j located at $z_{\ell+1}$. The wavefield extrapolation kernel in equation 3.1 reads

$$w^+_{z_{\ell+1};z_\ell}\left(x_j; x_o, \omega_f\right) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} e^{-i\sqrt{\left(\frac{\omega_f}{v_j}\right)^2 - k_x^2}|\delta z|} e^{-ik_x x_o} dk_x, \tag{3.2}$$

in which the integral represents the inverse Fourier transform of the downward extrapolation kernel—downward phase-shift extrapolation kernel—in the frequency-wavenumber domain in the x-direction, the index o means the lateral position at z_ℓ, $x_j; x_o$ means from x_o to x_j, δz represents the extrapolation step and equals the vertical distance between z_ℓ and $z_{\ell+1}$, v_j represents the velocity value at x_j between the virtual depth levels z_ℓ and $z_{\ell+1}$, and k_x is the horizontal wavenumber.

The downward extrapolation kernel in its matrix form is expressed as

$$\mathbf{W}^{+}_{z_{\ell+1};z_\ell} = \begin{bmatrix} \mathbf{w}^{+}_{z_{\ell+1};z_\ell}\left(x_1; x, \omega_f\right) & & & \\ & \ddots & & \\ & & \mathbf{w}^{+}_{z_{\ell+1};z_\ell}\left(x_j; x, \omega_f\right) & \\ & & & \ddots \\ & & & \mathbf{w}^{+}_{z_{\ell+1};z_\ell}\left(x_{n_x}; x, \omega_f\right) \end{bmatrix}, \tag{3.3}$$

where n_x shows the total number of model parameters located on each virtual depth level, $\mathbf{W}^{+}_{z_{\ell+1};z_\ell}$ is a square matrix of dimension $n_x \times n_x$, and $\mathbf{w}^{+}_{z_{\ell+1};z_\ell}\left(x_j; x, \omega_f\right)$ represents a row vector. For a virtual layer without lateral velocity variations, $\mathbf{W}^{+}_{z_{\ell+1};z_\ell}$ becomes a Toeplitz matrix (Berkhout, 1982, page 164), and for a laterally inhomogeneous virtual layer, defined by locally averaged velocities, $\mathbf{W}^{+}_{z_{\ell+1};z_\ell}$ becomes a space-variant convolutional matrix (Thorbecke et al., 2004). Figure 3.1a schematically displays one row of $\mathbf{W}^{+}_{z_{\ell+1};z_\ell}$. If more clarification is required, the extrapolation kernel is further described in Berkhout (2014a).

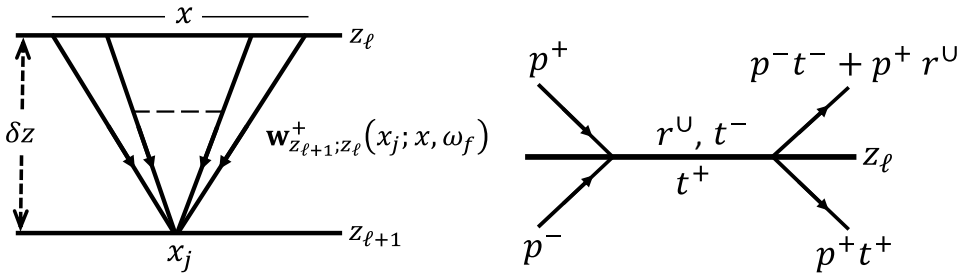


Figure 3.1: (a) Schematic representation of the jth row of the downward extrapolation matrix $\mathbf{W}^{+}_{z_{\ell+1};z_\ell}$ for a virtual layer bounded between two virtual depth levels. (b) Schematic representation of how the incident wavefields interact (reflection and transmission) at a particular virtual depth level in PWMod, where $r^\cup$ represents an upward reflection scalar, t^+ represents a downward transmission scalar, and t^- represents an upward transmission scalar.

PWMod reads the following equations—in the vector-matrix form—to model the angle-independent primary reflection data (traditional surface seismic data excluding direct waves, refractions, and diving waves) including reflection and transmission effects

(Figure 3.1b)

$$\mathbf{p}^{-}_{\mathrm{mod}}\left(z_0,\omega_f\right)=\sum_{m=N}^{1}\left[\left[\prod_{n=1}^{m-1}\mathbf{W}^{-}_{z_{n-1};z_n}\mathbf{T}^{-}\left(z_n\right)\right]\mathbf{W}^{-}_{z_{m-1};z_m}\right]\left(\mathbf{r}^{\cup}\left(z_m\right)\circ\mathbf{p}^{+}_{\mathrm{mod}}\left(z_m,\omega_f\right)\right), \tag{3.4}$$

$$\mathbf{p}^{+}_{\mathrm{mod}}\left(z_m,\omega_f\right)=\left[\left[\prod_{n=m-1}^{1}\mathbf{W}^{+}_{z_{n+1};z_n}\mathbf{T}^{+}\left(z_n\right)\right]\mathbf{W}^{+}_{z_1;z_0}\right]\mathbf{s}^{+}\left(z_0,\omega_f\right), \tag{3.5}$$

in which $\mathbf{p}^{-}_{\mathrm{mod}}\left(z_0,\omega_f\right)$ is the monochromatic upgoing wavefield modeled at z_0, $\mathbf{p}^{+}_{\mathrm{mod}}\left(z_m,\omega_f\right)$ is the monochromatic downgoing wavefield modeled at z_m. The matrix-matrix multiplication operations enclosed by the outer square brackets in equation 3.4 contain the upward extrapolation operator ($\mathbf{W}^{-}$) along with the upward transmission diagonal-matrix-operator ($\mathbf{T}^{-}\left(z_n\right)=\mathbf{I}-\mathrm{diag}(\mathbf{r}^{\cup}\left(z_n\right))$, encapsulating all the operators required for a wavefield to reach from z_m to z_0. The matrix-matrix multiplication operations enclosed by the outer square brackets in equation 3.5 contain the downward extrapolation operator ($\mathbf{W}^{+}$) along with the downward transmission diagonal-matrix-operator ($\mathbf{T}^{+}\left(z_n\right)=\mathbf{I}+\mathrm{diag}(\mathbf{r}^{\cup}\left(z_n\right))$, encapsulating all the operators required for a wavefield to reach from z_0 to z_m. Here, $\mathbf{r}^{\cup}\left(z_m\right)$ represents the angle-independent upward reflectivity vector-operator at z_m, $\mathbf{s}^{+}\left(z_0,\omega_f\right)$ shows the monochromatic downgoing physical source at the Earth's surface, N is the total number of virtual depth levels, and the symbol $\circ$ means element-wise product. Note that, diag(.) represents a mathematical function that receives a vector and gives a diagonal matrix as its output, and $\mathbf{I}$ means the identity matrix.

ORWI inverts for reflectivity and background velocity (two different classes of parameters) by minimizing the errors in primary reflection waveforms for each class of parameters alternately. Standard ORWI minimizes the following misfit function iteratively

$$\mathscr{C}(\mathbf{r}^{\cup},\mathbf{v})=\frac{1}{2}\sum_{f=1}^{N_f}\sum_{s=1}^{N_s}\|\mathbf{p}^{-}_{\mathrm{obs},s}\left(z_0,\omega_f\right)-\mathbf{p}^{-}_{\mathrm{mod},s}\left(z_0,\omega_f;\mathbf{r}^{\cup},\mathbf{v}\right)\|_2^2, \tag{3.6}$$

in which $\mathbf{p}^{-}_{\mathrm{obs},s}\left(z_0,\omega_f\right)$ represents the observed monochromatic multiple-free reflection waveforms for shot s recorded at the receiver locations, $\mathbf{p}^{-}_{\mathrm{mod},s}\left(z_0,\omega_f\right)$ represents the modeled monochromatic primary reflection waveforms for shot s recorded at the receiver locations, the vector difference $\mathbf{p}^{-}_{\mathrm{obs},s}\left(z_0,\omega_f\right)-\mathbf{p}^{-}_{\mathrm{mod},s}\left(z_0,\omega_f\right)$ is called the monochromatic residual data vector for shot s and represented later by $\delta\mathbf{d}^{-}_{s}\left(\omega_f\right)$, $\mathbf{v}$ represents the background velocity vector, $\mathbf{r}^{\cup}$ represents the angle-independent upward reflectivity vector, N_s is the total number of shots, and N_f is the total number of angular frequency components.

Taking the partial derivatives of the misfit function with respect to $\mathbf{r}^{\cup}$ and $\mathbf{v}$ gives the

total descent direction as

$$\mathbf{g}=\begin{bmatrix}\mathbf{g}^{r}(z_0)\\ \vdots\\ \mathbf{g}^{r}(z_N)\\ \mathbf{g}^{v}(z_0)\\ \vdots\\ \mathbf{g}^{v}(z_N)\end{bmatrix}=-\sum_{f=1}^{N_f}\sum_{s=1}^{N_s}\Re\left\{\overbrace{\left[\overbrace{\frac{\partial\mathbf{p}^{-}_{\text{mod},s}\left(z_0,\omega_f\right)}{\partial\mathbf{r}^{\cup}(z_0)}\cdots\frac{\partial\mathbf{p}^{-}_{\text{mod},s}\left(z_0,\omega_f\right)}{\partial\mathbf{r}^{\cup}(z_N)}\frac{\partial\mathbf{p}^{-}_{\text{mod},s}\left(z_0,\omega_f\right)}{\partial\mathbf{v}(z_0)}\cdots\frac{\partial\mathbf{p}^{-}_{\text{mod},s}\left(z_0,\omega_f\right)}{\partial\mathbf{v}(z_N)}}^{\mathbf{J}_s(\omega_f)}\right]^{\dagger}\delta\mathbf{d}^{-}_{s}(\omega_f)}^{\mathbf{g}_s(\omega_f)}\right\},\tag{3.7}$$

in which $\mathbf{g}$ denotes the total two-parameter gradient vector, $\mathbf{g}^{r}(z_m)$ shows the reflectivity gradient at a given depth level for migration, $\mathbf{g}^{v}(z_m)$ shows the background velocity gradient between z_m and z_{m+1} for tomography (for simplicity, shown in notation at z_m), $\mathbf{g}_s(\omega_f)$ is the total gradient vector for one shot and one frequency, $\dagger$ means the adjoint operator, $\mathbf{J}_s\left(\omega_f\right)$ is the total Jacobian matrix for one shot and one frequency, and each column of $\mathbf{J}_s\left(\omega_f\right)$ denotes the partial derivative of the upgoing modeled wavefield at the receiver locations with respect to the model parameters of one class, either reflectivity or background velocity, located at a given depth level.

The two-parameter nature of ORWI—as a minimization problem—is turned into a single-parameter by keeping a class of parameters constant while updating the other class. As a result, equation 3.7 is broken into two parts

$$\mathbf{g}^{r}=-\sum_{f=1}^{N_f}\sum_{s=1}^{N_s}\Re\left\{\left[\frac{\partial\mathbf{p}^{-}_{\text{mod},s}\left(z_0,\omega_f\right)}{\partial\mathbf{r}^{\cup}(z_0)}\cdots\frac{\partial\mathbf{p}^{-}_{\text{mod},s}\left(z_0,\omega_f\right)}{\partial\mathbf{r}^{\cup}(z_N)}\right]^{\dagger}\delta\mathbf{d}^{-}_{s}\left(\omega_f\right)\right\},\tag{3.8}$$

$$\mathbf{g}^{v}=-\sum_{f=1}^{N_f}\sum_{s=1}^{N_s}\Re\left\{\left[\frac{\partial\mathbf{p}^{-}_{\text{mod},s}\left(z_0,\omega_f\right)}{\partial\mathbf{v}(z_0)}\cdots\frac{\partial\mathbf{p}^{-}_{\text{mod},s}\left(z_0,\omega_f\right)}{\partial\mathbf{v}(z_N)}\right]^{\dagger}\delta\mathbf{d}^{-}_{s}\left(\omega_f\right)\right\},\tag{3.9}$$

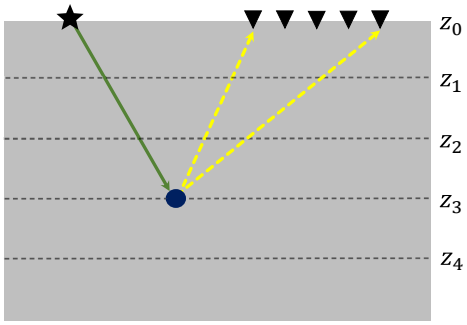
where $\mathbf{g}^{r}$ shows the total reflectivity gradient for migration and $\mathbf{g}^{v}$ shows the total velocity gradient for tomography. Indeed, with such a gradient separation, ORWI turns into an iterative cycle alternating between LS-WEM and least-squares one-way wave-equation reflection tomography (LS-WET) until the data residual vanishes (Algorithm 3.1).

Expanding one component of equation 3.8, which denotes the reflectivity gradient for a particular depth level, frequency component, and shot, yields (Figure 3.2)

$$\mathbf{g}^{r}_{s}(z_m,\omega_f)=-\Re\left\{\left(\mathbf{p}^{+}_{\text{mod,s}}\left(z_m,\omega_f\right)\right)^{*}\circ\overbrace{\left(\mathbf{L}^{-\dagger}_{z_0;z_m}\delta\mathbf{d}^{-}_{s}\left(\omega_f\right)\right)}^{\text{back-propagated wavefield}}\right\},\tag{3.10}$$

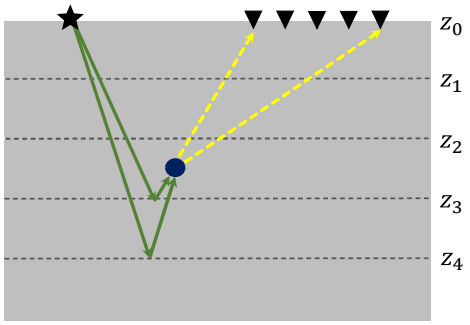
in which $\mathbf{L}^{-}_{z_0;z_m}$ is defined as

$$\mathbf{L}^{-}_{z_0;z_m}=\left[\prod_{n=1}^{m-1}\mathbf{W}^{-}_{z_{n-1};z_n}\mathbf{T}^{-}(z_n)\right]\mathbf{W}^{-}_{z_{m-1};z_m}.\tag{3.11}$$

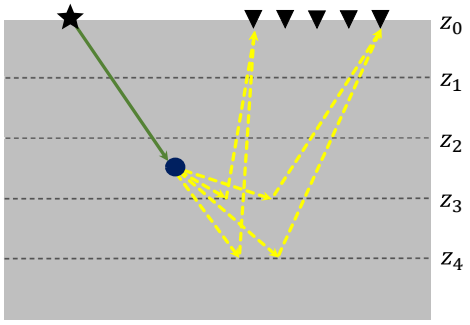


(a)

Figure 3.2: Schematic representation of one element of the reflectivity gradient for one shot, equation 3.10, in a medium including five virtual depth levels. Green signifies forward wavefields, while yellow denotes back-propagated wavefields.



(a) Receiver-side background velocity gradient



(b) Source-side background velocity gradient

Figure 3.3: Schematic representation of one element of the background velocity gradient for one shot, equation 3.12, in a medium including five virtual depth levels. (a) Receiver-side background velocity gradient for one shot: the term $\mathbf{g}_s^{v_A}(z_m,\omega_f)$ from equation 3.12. (b) Source-side background velocity gradient for one shot: the term $\mathbf{g}_s^{v_B}(z_m,\omega_f)$ from equation 3.12. Green signifies forward wavefields, while yellow denotes back-propagated wavefields.

Expanding one component of equation 3.9, which denotes the velocity gradient for a particular depth level, frequency component, and shot, yields (Figure 3.3)

$$\mathbf{g}_s^{v}(z_m,\omega_f)=\mathbf{g}_s^{v_A}(z_m,\omega_f)+\mathbf{g}_s^{v_B}(z_m,\omega_f)=-\Re\left\{\mathbf{G}^{-^*}_{z_m;z_{m+1}}\mathbf{q}^{-^*}_{\text{mod,s}}\left(z_{m+1},\omega_f\right)\circ\overbrace{\left(\left(\mathbf{L}^{-}_{z_0;z_m}\mathbf{T}^{-}\left(z_m\right)\right)^{\dagger}\delta\mathbf{d}_s^{-}\left(\omega_f\right)\right)}^{\text{back-propagated wavefield}}\right\}$$

$$-\Re\left\{\mathbf{G}^{+^*}_{z_{m+1};z_m}\mathbf{q}^{+^*}_{\text{mod,s}}\left(z_m,\omega_f\right)\circ\overbrace{\left(\left[\mathbf{L}^{-}_{z_0;z_{m+1}}\mathbf{R}^{\cup}\left(z_{m+1}\right)+\left[\sum_{m'=N}^{m+2}\mathbf{L}^{-}_{z_0;z_{m'}}\mathbf{R}^{\cup}\left(z_{m'}\right)\mathbf{L}^{+}_{z_{m'};z_{m+1}}\right]\mathbf{T}^{+}\left(z_{m+1}\right)\right]^{\dagger}\delta\mathbf{d}_s^{-}\left(\omega_f\right)\right)}^{\text{back-propagated wavefield}}\right\}, \tag{3.12}$$

and given the assumption that $\mathbf{W}^{+}_{z_{m+1};z_m}$ approximates $\mathbf{W}^{-}_{z_m;z_{m+1}}$, $\mathbf{G}$ is defined as

$$\mathbf{G}^{+}_{z_{m+1};z_m}=\mathbf{G}^{-}_{z_m;z_{m+1}}=\begin{bmatrix}\frac{\partial\mathbf{w}_{1,\star}}{\partial v_1}\\ \frac{\partial\mathbf{w}_{2,\star}}{\partial v_2}\\ \vdots\\ \frac{\partial\mathbf{w}_{n_x,\star}}{\partial v_{n_x}}\end{bmatrix}, \tag{3.13}$$

in which $\mathbf{w}_{j,\star}$ means the jth row of $\mathbf{W}^{-}_{z_m;z_{m+1}}$, and

$$\begin{aligned}\mathbf{q}^{-}_{\text{mod,s}}\left(z_{m+1},\omega_f\right)&=\mathbf{T}^{-}\left(z_{m+1}\right)\mathbf{p}^{-}_{\text{mod,s}}\left(z_{m+1},\omega_f\right)+\mathbf{R}^{\cup}\left(z_{m+1}\right)\mathbf{p}^{+}_{\text{mod,s}}\left(z_{m+1},\omega_f\right),\\ \mathbf{q}^{+}_{\text{mod,s}}\left(z_m,\omega_f\right)&=\mathbf{T}^{+}\left(z_m\right)\mathbf{p}^{+}_{\text{mod,s}}\left(z_m,\omega_f\right),\end{aligned} \tag{3.14}$$

in which $\mathbf{R}^{\cup}(z_{m+1})=\text{diag}\left(\mathbf{r}^{\cup}(z_{m+1})\right)$. It is important to recall that while equation 3.12 represents the background velocity gradient at z_m in notation, it updates the background velocity over the depth interval between z_m and z_{m+1} in practice.

To optimize the update direction, standard ORWI preconditions both gradient vectors as follows (equivalent to equation 27 in Plessix and Mulder (2004))

$$\delta\mathbf{r}^{\cup}(z_m)=\Re\left\{\overbrace{\sum_{f=1}^{N_f}\sum_{s=1}^{N_s}\left(\mathbf{p}^{+}_{\text{mod,s}}\left(z_m,\omega_f\right)\circ\mathbf{p}^{+^*}_{\text{mod,s}}\left(z_m,\omega_f\right)\right)\circ\left(\sum_{g=1}^{N_g}\left[\mathbf{L}^{-}_{z_0;z_m}\right]_{g,\star}\circ\left[\mathbf{L}^{-^*}_{z_0;z_m}\right]_{g,\star}\right)^{T}}^{\text{preconditioner (a column vector)}}\right\}^{-1}\circ\mathbf{g}^{r}(z_m), \tag{3.15}$$

and

$$\delta\mathbf{v}(z_m)=\Re\left\{\overbrace{\sum_{f=1}^{N_f}\sum_{s=1}^{N_s}\left(\mathbf{q}^{-}_{\mathrm{mod,s}}\left(z_{m+1},\omega_f\right)\circ\mathbf{q}^{-^*}_{\mathrm{mod,s}}\left(z_{m+1},\omega_f\right)\right)\circ\left(\sum_{g=1}^{N_g}\left[\Theta\right]_{g,\star}\circ\left[\Theta^*\right]_{g,\star}\right)^T}^{\text{preconditioner (a column vector)}}\right\}^{-1}\circ\mathbf{g}^{v_A}(z_m)$$
$$+\Re\left\{\overbrace{\sum_{f=1}^{N_f}\sum_{s=1}^{N_s}\left(\mathbf{q}^{+}_{\mathrm{mod,s}}\left(z_m,\omega_f\right)\circ\mathbf{q}^{+^*}_{\mathrm{mod,s}}\left(z_m,\omega_f\right)\right)\circ\left(\sum_{g=1}^{N_g}\left[\Psi\right]_{g,\star}\circ\left[\Psi^*\right]_{g,\star}\right)^T}^{\text{preconditioner (a column vector)}}\right\}^{-1}\circ\mathbf{g}^{v_B}(z_m), \tag{3.16}$$

with

$$\begin{aligned}\Theta&=\mathbf{L}^{-}_{z_0;z_m}\mathbf{T}^{-}(z_m)\,\mathbf{G}^{-}_{z_m;z_{m+1}},\\ \Psi&=\left[\mathbf{L}^{-}_{z_0;z_{m+1}}\mathbf{R}^{\cup}(z_{m+1})+\left[\sum_{m'=N}^{m+2}\mathbf{L}^{-}_{z_0;z_{m'}}\mathbf{R}^{\cup}(z_{m'})\mathbf{L}^{+}_{z_{m'};z_{m+1}}\right]\mathbf{T}^{+}(z_{m+1})\right]\mathbf{G}^{+}_{z_{m+1};z_m},\end{aligned} \tag{3.17}$$

where $\delta\mathbf{r}^{\cup}(z_m)$ represents the reflectivity update/perturbation vector at z_m, $\delta\mathbf{v}(z_m)$ is the background velocity update/perturbation vector between z_m and z_{m+1}.

Finally, standard ORWI updates the total angle-independent upward reflectivity and background velocity vectors as

$$\mathbf{r}^{\cup}_{k+1}=\mathbf{r}^{\cup}_{k}+\alpha^{r}_{k}\delta\mathbf{r}^{\cup}_{k}, \tag{3.18}$$

$$\mathbf{v}_{k+1}=\mathbf{v}_{k}+\alpha^{v}_{k}\delta\mathbf{v}_{k}, \tag{3.19}$$

in which k means the current cycle number, $\delta\mathbf{r}^{\cup}$ is the total reflectivity perturbation vector, $\delta\mathbf{v}$ is the total velocity perturbation vector, and α^{v} and α^{r} represent the minimization step lengths for estimating reflectivity and background velocity, respectively. For α^{r}_{k}, we use

$$\alpha^{r}_{k}=\frac{\delta\mathbf{d}^{\dagger}\left(\mathbf{r}^{\cup}_{k}+\delta\mathbf{r}^{\cup}_{k},\mathbf{v}_k\right)\delta\mathbf{d}\left(\mathbf{r}^{\cup}_{k},\mathbf{v}_k\right)}{\delta\mathbf{d}^{\dagger}\left(\mathbf{r}^{\cup}_{k}+\delta\mathbf{r}^{\cup}_{k},\mathbf{v}_k\right)\delta\mathbf{d}\left(\mathbf{r}^{\cup}_{k}+\delta\mathbf{r}^{\cup}_{k},\mathbf{v}_k\right)}, \tag{3.20}$$

and for α^{v}_{k}, we use

$$\alpha^{v}_{k}=\frac{\delta\mathbf{d}^{\dagger}\left(\mathbf{r}^{\cup}_{k+1},\mathbf{v}_k+\delta\mathbf{v}_k\right)\delta\mathbf{d}\left(\mathbf{r}^{\cup}_{k+1},\mathbf{v}_k\right)}{\delta\mathbf{d}^{\dagger}\left(\mathbf{r}^{\cup}_{k+1},\mathbf{v}_k+\delta\mathbf{v}_k\right)\delta\mathbf{d}\left(\mathbf{r}^{\cup}_{k+1},\mathbf{v}_k+\delta\mathbf{v}_k\right)}, \tag{3.21}$$

where $\delta\mathbf{d}$ is the superposition of all shots and the frequency components within the range. For more information on the gradient derivation, see Sun et al. (2019).

3.3. ENHANCED MIGRATION ALGORITHM IN ORWI

In this section, we introduce the first of our four proposed improvements to the ORWI methodology, with which we aim to enhance the migration algorithm embedded within

ORWI. LSM techniques have proven to be computationally expensive as their convergence requires multiple iterations. Gradient preconditioning helps reduce the number of LSM iterations. The preconditioned least-squares one-way wave-equation migration (PLS-WEM) introduced by Abolhassani and Verschuur (2024) is a depth least-squares migration technology that recovers high-resolution and accurate seismic images by reconstructing the approximate Hessian information recursively in depth; PLS-WEM accelerates the convergence of the migration process after employing the reciprocal of the approximate Hessian information for gradient preconditioning. PLS-WEM is relatively fast when compared to similar technologies because it recursively calculates the approximate Hessian operator depth by depth rather than calculating one huge approximate Hessian for the entire medium. This reduces the huge approximate Hessian operator for the entire domain into smaller, depth-dependent operators. Furthermore, as PLS-WEM uses PWMod for forward modeling, this facilitates to switch from LS-WEM to PLS-WEM in the standard ORWI cycle. This section provides an overview of the PLS-WEM theory.

To incorporate gradient preconditioning into equation 3.8, PLS-WEM minimizes the misfit function (Shin et al., 2001; Jang et al., 2009; Oh and Min, 2013)

$$\mathscr{C}\left(\delta\mathbf{r}^{\cup}\right)=\frac{1}{2}\sum_{f=1}^{N_f}\sum_{s=1}^{N_s}\left\|\delta\mathbf{d}_s^{-}\left(\omega_f\right)-\mathbf{J}_s^{r}\left(\omega_f\right)\delta\mathbf{r}^{\cup}\left(\omega_f\right)\right\|_2^2, \tag{3.22}$$

where $\mathbf{J}_s^{r}\left(\omega_f\right)=\dfrac{\partial\mathbf{p}^{-}_{\mathbf{mod},s}\left(z_0,\omega_f\right)}{\partial\mathbf{r}^{\cup}}$ and is called the total monochromatic reflectivity Jacobian for shot s.

The stationary point of equation 3.22 with respect to $\delta\mathbf{r}^{\cup}\left(\omega_f\right)$ is given by

$$\delta\mathbf{r}^{\cup}=-\sum_{f=1}^{N_f}\left(\Re\left\{\sum_{s=1}^{N_s}\mathbf{H}_s^{r,\mathrm{a}}(\omega_f)\right\}^{-1}\Re\left\{\sum_{s=1}^{N_s}\mathbf{g}_s^{r}(\omega_f)\right\}\right), \tag{3.23}$$

with $\mathbf{g}_s^{r}(\omega_f)$ and $\mathbf{H}_s^{r,\mathrm{a}}(\omega_f)$ defined as

$$\mathbf{g}_s^{r}(\omega_f)=\mathbf{J}_s^{r\dagger}(\omega_f)\delta\mathbf{d}_s^{-}\left(\omega_f\right), \tag{3.24}$$

$$\mathbf{H}_s^{r,\mathrm{a}}(\omega_f)=\mathbf{J}_s^{r\dagger}(\omega_f)\mathbf{J}_s^{r}(\omega_f), \tag{3.25}$$

in which $\mathbf{J}_s^{r}(\omega_f)$ denotes the monochromatic reflectivity Jacobian for shot s, $\mathbf{g}_s^{r}(\omega_f)$ represents the monochromatic reflectivity gradient for shot s (mathematically derived in equation 3.10 for each virtual depth level), and $\mathbf{H}_s^{r,\mathrm{a}}(\omega_f)$ refers to the monochromatic reflectivity approximate Hessian for shot s compensating for geometrical spreading, spatial correlations of neighboring model parameters on the same virtual depth levels (Pratt et al., 1998), and also conducting source deconvolution as it acts on the gradient vector frequency by frequency.

Thanks to PWMod, each column of the total monochromatic reflectivity Jacobian for shot s, introduced in equation 3.24, satisfies

$$\mathbf{J}_s^{r}(z_m,\omega_f)=\frac{\partial\mathbf{p}^{-}_{\mathrm{mod},s}(z_0,\omega_f)}{\partial\mathbf{r}^{\cup}(z_m)}. \tag{3.26}$$

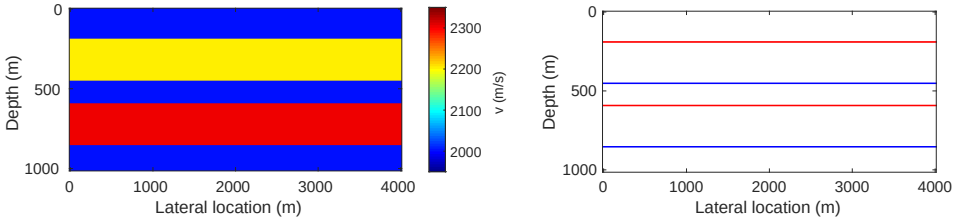
Abolhassani and Verschuur (2022) showed that $\mathbf{J}_s^r(z_m,\omega_f)$ can be expressed as follows

$$\mathbf{J}_s^r\left(z_m,\omega_f\right)=\left[\prod_{n=m-1}^{1}\mathbf{W}^-_{z_{n-1};z_n}\mathbf{T}^-\left(z_n\right)\right]\mathbf{W}^-_{z_{m-1};z_m}\left[\begin{pmatrix}p^+_{1\,\mathrm{mod},s}\left(z_m,\omega_f\right)\\0\\\vdots\\0\end{pmatrix}\begin{pmatrix}0\\p^+_{2\,\mathrm{mod},s}\left(z_m,\omega_f\right)\\\vdots\\0\end{pmatrix}\cdots\begin{pmatrix}0\\0\\\vdots\\p^+_{n_x\,\mathrm{mod},s}\left(z_m,\omega_f\right)\end{pmatrix}\right], \tag{3.27}$$

in which n_x denotes the number of gridpoints at z_m, $p^+_{j\,\mathrm{mod},s}\left(z_m,\omega_f\right)$ is the monochromatic downgoing modeled wavefield at the jth gridpoint of z_m, and $\mathbf{J}_s^r(z_m,\omega_f)$ is a matrix of dimension $n_x\times n_x$. Plugging equation 3.27 into equation 3.25 yields $\mathbf{H}_s^{r,\mathrm{a}}(z_m,\omega_f)$. Adopting equation 3.23 for each virtual depth level, the reflectivity update vector $\delta\mathbf{r}^{\cup}(z_m)$ is now constructed. Clearly, equation 3.23 requires a direct inversion of the approximate Hessian matrix for each frequency component. Finally, equation 3.18 is used to update the total angle-independent upward reflectivity vector.

PLS-WEM is computationally more feasible compared to alternative technologies that approximate Hessian coefficients at once, as it operates in a depth-marching regime. PLS-WEM achieves this by decomposing and reducing the entire approximate Hessian operator into multiple smaller sub-operators. Each of these sub-operators is relevant to a single virtual depth level and captures only the correlation of the partial derivative wavefields at that depth level. If assembled together for all virtual depth levels, these operators form a block-diagonal presentation of the approximate Hessian operator for the entire medium. With PLS-WEM, each time $\mathbf{H}_s^{r,\mathrm{a}}(z_m,\omega_f)$ and its reciprocal are calculated, indeed, only $1/N$ of the total number of model parameters is involved (Abolhassani and Verschuur, 2024).

To evaluate the performance of ORWI based on PLS-WEM compared to ORWI based on LS-WEM (standard ORWI), for a flat-layered medium shown in Figure 3.4, we compare the background velocity model update estimated by standard ORWI to that estimated by ORWI equipped with PLS-WEM, both after one cycle. To avoid amplitude fitting in tomography, we use relative amplitude scaling proposed by Son et al. (2013). With a zero initial reflectivity model, a homogeneous acoustic velocity model of 2000 m/s, a Ricker wavelet with a peak frequency of 10 Hz, surface acquisition, the reflection dataset within the range of 0-15 Hz, a maximum offset of 500 m for migration, and a maximum offset of 1000 m for tomography (no cycle-skipping), we run the comparison. After 1 cycle of ORWI, including 1x LS migration and 1x LS tomography, the resultant images and tomograms are shown in Figure 3.5. As expected, ORWI based on PLS-WEM delivers a superior tomographic update compared to ORWI based on LS-WEM. The superposition of tomographic wavepaths obtained based on PLS-WEM is stronger and more geometrically consistent with the true layers. This enhanced performance is attributed to the improved amplitude preservation and resolution power of PLS-WEM, as shown in Figure 3.5d.



(a) True background velocity model

(b) True reflectivity model

Figure 3.4: Flat-layered medium. (a) True background velocity model. (b) True reflectivity model.

3.4. OPTIMIZING THE TOMOGRAPHIC GRADIENT IN ORWI

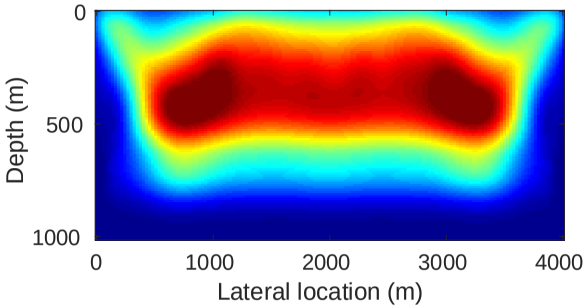
In this section, we first examine how well the full-wave consistency condition is met in ORWI. Next, we suggest three strategies for offset selection (in both migration and tomography loops) to optimize the tomographic updates in ORWI.

3.4.1. TOMOGRAPHIC UPDATE ANALYSIS

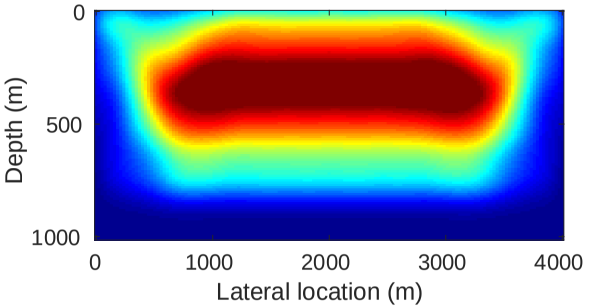
RWI solution primarily relies on a full-wave consistency condition between the anchor portions of the modeled and observed reflection data (Verschuur et al., 2016). To meet this condition, the migration loop must predict the reflectors' positions in depth so that the anchor portions of the observed and modeled waveforms (i.e., short-offset waveforms) remain consistent in terms of amplitude and traveltime, resulting in zero energy for tomography. This implies that, for example, for a simple model with a flat reflector and a homogeneous overburden, the velocity update direction remains unique whether the homogeneous initial velocity is greater or less than the true velocity: it either increases or decreases the initial velocity. As soon as the full-wave consistency condition does not hold, confusion arises in the velocity update direction, where the long and short-offset data act oppositely, which can result in getting trapped in a local minimum situation. Figure 3.6 shows an example of reflection data from a flat reflector medium when the full-wave consistency condition fails in the short-offset data (this data is derived after one cycle of reflection waveform inversion).

To verify whether or not the full-wave consistency in the short-offset data is well supported by ORWI, we here display the tomographic update of ORWI for a single-reflector medium after one cycle. The medium consists of a single reflector and a homogeneous overburden with an acoustic velocity of 3000 m/s (Figure 3.7a). For the initial background velocity model, a homogeneous model with an acoustic velocity of 2700 m/s is used (Figure 3.7b). For such a medium, the tomographic updates following one cycle of ORWI, including 1x LS migration and 1x LS tomography, in four different data scenarios are displayed in Figure 3.8. The data scenarios are listed in Table 3.1. Note that in the table, offsets with absolute values smaller than 1000 m are counted as uncycle-skipped.

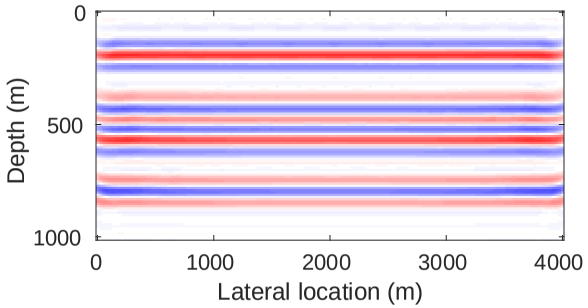
Scenario 1 evaluates the ORWI background velocity update when both data for imaging and tomography are cycle-skipped. As evident in Figure 3.8a, associated with Scenario 1, the background velocity update appears to be completely inefficient. While a positive



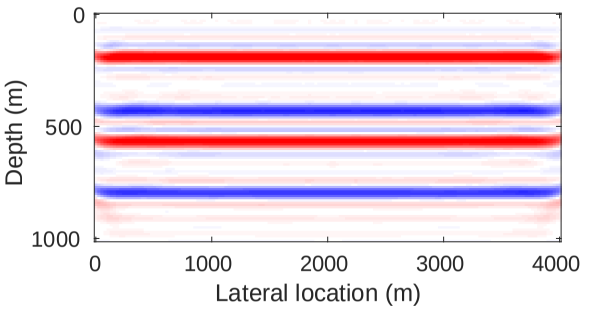
(a) Tomographic update via ORWI equipped with LS-WEM



(b) Tomographic update via ORWI equipped with PLS-WEM



(c) Image update with LS-WEM

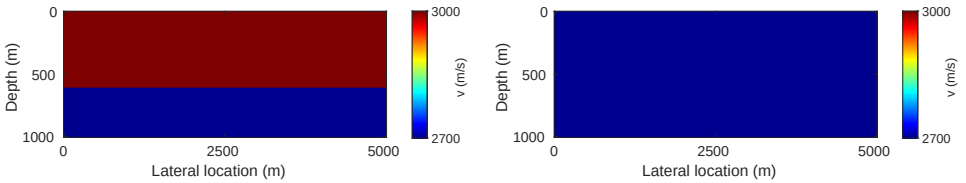


(d) Image update with PLS-WEM

Figure 3.5: Estimated background velocity models after 1 cycle of ORWI based on LS-WEM and PLS-WEM images, associated with the flat-layered medium. (a) and (c) Tomographic and imaging updates after 1 cycle of ORWI equipped with LS-WEM (standard ORWI). (b) and (d) Tomographic and imaging updates after 1 cycle of ORWI equipped with PLS-WEM image.

(a) $v_{initial} > v_{true}$ (b) $v_{initial} < v_{true}$

Figure 3.6: Modeled and observed waveforms for a flat reflector medium with a homogeneous overburden when the full-wave consistency condition is not met. (a) The homogeneous initial velocity is greater than the true velocity. (b) The homogeneous initial velocity is less than the true velocity. In both cases, blue arrows highlight the anchor portions of data where the full-wave inconsistency due to the conflict between updating the background velocity with the fixed reflector's depth appears.



(a) True background velocity model (b) Initial background velocity model

3

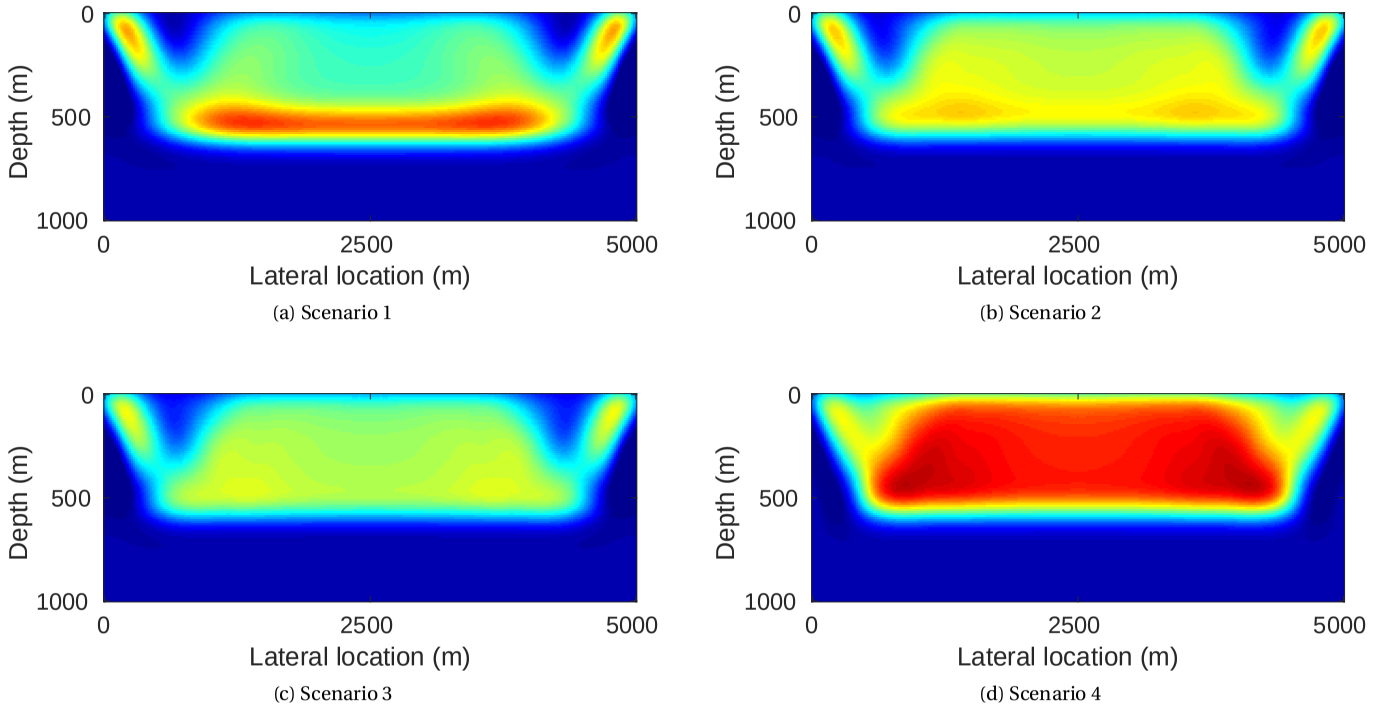
Figure 3.7: Single-reflector medium. (a) True background velocity model. (b) Initial background velocity model.

homogeneous velocity update everywhere above the reflector position is expected, an inhomogeneous mixture of positive and negative velocity updates is retrieved. To have a better understanding, the middle-shot subset of the tomographic update shown in Figure 3.8a is separately represented in Figure 3.9. According to Figure 3.9, the tomographic update can be described as the superposition of three batches of tomographic wavepaths updating against each other: uncycle-skipped mid-to-long-offset tomographic wavepaths, denoted by the gray arrow, increasing the background velocity above the reflector (correct update sign), cycle-skipped long-offset tomographic wavepaths, denoted by the blue arrow, decreasing the background velocity above the reflector (incorrect update sign), and short-offset tomographic wavepaths, denoted by the red arrow, decreasing the background velocity above the reflector (incorrect update sign). This complication in resolving the tomographic gradient indicates that the tomographic gradient of ORWI is degraded not only by cycle skipping in long-offset data due to longer propagation in the incorrect velocity but also by full-wave inconsistency in the short-offset data—the imprint of inconsistent reflectivity and velocity models. Note that narrowing the migration offset to short offsets, given the depth of investigation, and increasing the number of LS migration iterations can attenuate the imprint (Provenzano et al., 2023); this, however, comes at the price of increased computational cost. Moreover, there might be cases where the short-offset data is either missing or unusable due to technical reasons.

Table 3.1: Four different data scenarios for retrieving tomographic updates via ORWI.

	Data for imaging	**Data for tomography**	**Imprint of inconsistent reflectivity and velocity**
Scenario 1	cycle-skipped	cycle-skipped	present
Scenario 2	cycle-skipped	uncycle-skipped	present
Scenario 3	uncycle-skipped	uncycle-skipped	present
Scenario 4	uncycle-skipped	uncycle-skipped	muted

Scenario 2 evaluates the impact of excluding any contribution from the cycle-skipped long-offset data into the tomographic gradient to improve the tomographic gradient



(a) Scenario 1

(b) Scenario 2

(c) Scenario 3

(d) Scenario 4

Figure 3.8: Tomographic updates of ORWI associated with the single-reflector medium after one cycle in four different data scenarios introduced in Table 3.1. (a) Tomographic update estimated by ORWI when both data for imaging and tomography are cycle-skipped (Scenario 1). (b) Tomographic update estimated by ORWI after excluding any contribution of the cycle-skipped long-offset data into the tomographic gradient (Scenario 2). (c) Tomographic update estimated by ORWI after excluding any contribution of the cycle-skipped long-offset data into the imaging and tomographic gradients (Scenario 3). (d) Tomographic update estimated by ORWI after muting erroneous short-offset tomographic wavepaths (erroneous due to inconsistent reflectivity and velocity), coupled with cycle-skipping avoidance in both imaging and tomography (Scenario 4).

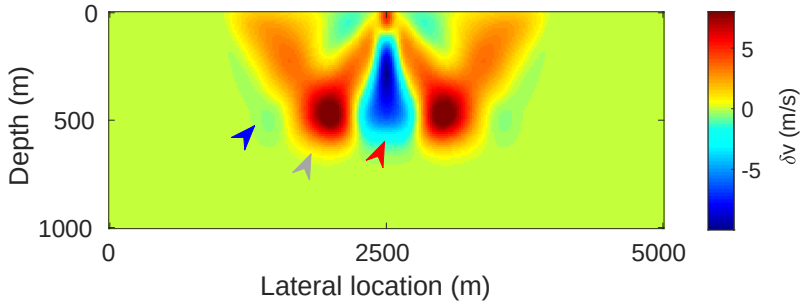


Figure 3.9: The middle-shot subset of the tomographic update shown in Figure 3.8a. The arrows represent different tomographic wavepath batches, as explained in the main body of the paper.

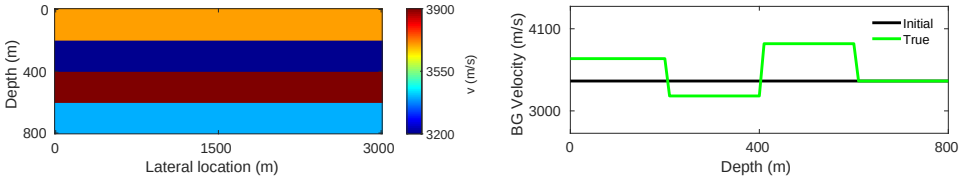
of ORWI. The background velocity update associated with Scenario 2 is represented in Figure 3.8b. Clearly, following the recovery of the canceled-out velocity updates after the exclusion of the cycle-skipped long-offset data, a fairly homogeneous background velocity update, compared to Scenario 1, is achieved. The improved update should not, however, obscure the fact that the full-wave inconsistency in the short-offset data still degrades the update and must be resolved.

Scenario 3 examines the impact of excluding any contribution from the cycle-skipped long-offset data into the imaging and tomographic gradients to improve the tomographic gradient of ORWI. The background velocity update associated with Scenario 3 is represented in Figure 3.8c. Compared to the previous scenario, the update is more homogeneous, indicating better image focusing due to cycle-skipping avoidance. Nevertheless, this does not guarantee an optimal tomographic update.

Scenario 4 examines the impact of muting erroneous short-offset tomographic wavepaths (erroneous due to inconsistent reflectivity and velocity) and excluding contributions from cycle-skipped long-offset data on the imaging and tomographic gradients, aiming to improve the tomographic gradient of ORWI. Figure 3.8d shows the associated background velocity update. As illustrated, Scenario 4 clearly enhances both the magnitude and homogeneity of the background velocity update.

3.4.2. TOMOGRAPHY OFFSET: SHORT AND MID-TO-LONG OFFSETS

Based on the analysis of Scenarios 1 to 4, to minimize the imprint of inconsistent reflectivity and velocity models on the tomographic gradient of ORWI, we propose muting the relevant erroneous tomographic wavepaths in the associated residual data gather, depending on the depth. As previously confirmed, these erroneous tomographic wavepaths are typically established at short offsets, where the full-wave consistency condition is not met. Figure 3.11a illustrates the conflict between reflectivity and velocity on the middle-shot (surface acquisition) tomographic gradient of ORWI after one cycle for a four-layer medium (Figure 3.10a), with the initial background velocity model being a homogeneous model of 3500 m/s (Figure 3.10b). Figure 3.11b shows the muted area in the middle-shot residual data gather, depending on the depth. As shown, the muted area increases “linearly” with depth and has a maximum width of 500 m.



(a) True background velocity model

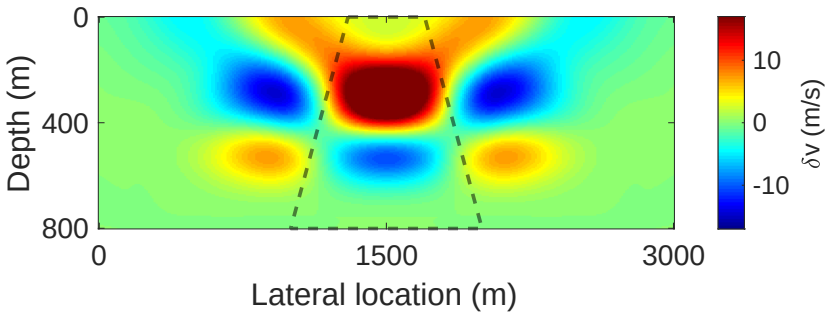
(b) 1D velocity profile

Figure 3.10: Four-layer medium. (a) True background velocity model. (b) 1D vertical background velocity profile, where BG stands for background.

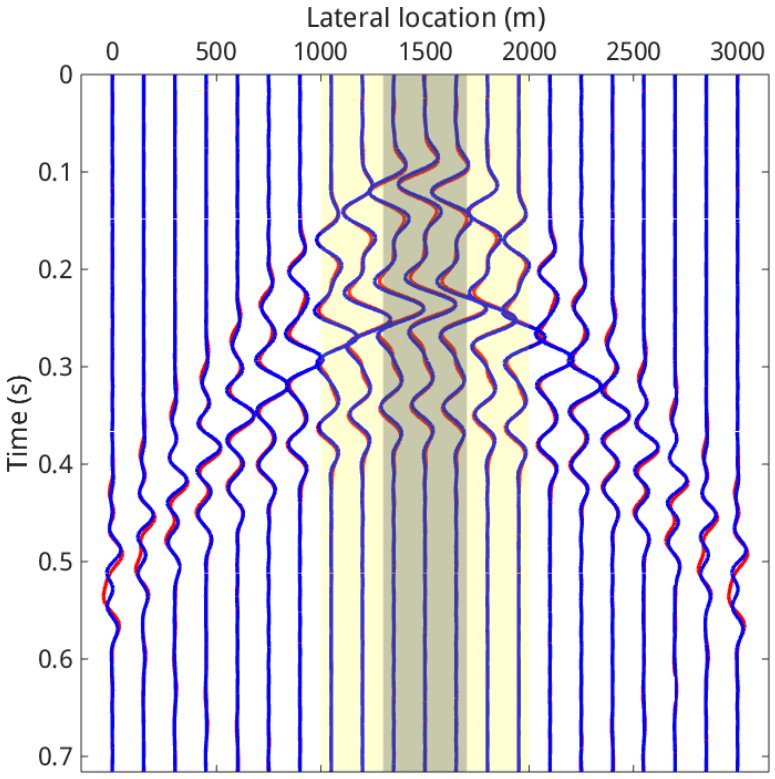
3.4.3. MIGRATION OFFSET

In the introduction, it was acknowledged that one approach to mitigate the imprint of inconsistent reflectivity and velocity in RWI is to use short-offset data for migration. Figure 3.12a shows the stacked image estimated by the migration mode of standard ORWI for the single-reflector medium (Figure 3.7) when the data for imaging is cycle-skipped (all migration offsets are included). As can be observed, due to the velocity error, not only do the reflector amplitudes highlighted in blue on the figure fail to dampen as one moves away from the predicted reflector position, but also a phase effect is evident in the recovered shape of the reflector. Such a partially out-of-focus image will definitely affect modeled waveforms in the tomography loop, degrading the full-wave consistency between the short-offset modeled and observed data for tomography. To lessen such error leakage from the migration loop to the tomography loop, a focused image with correct peak is required. The focused image may be obtained by either improving the background velocity model or muting the problematic isochrones in the imaging gradient. Since a better velocity model is what ORWI aims to achieve, the only way ahead would be to mute the problematic isochrones. Clearly, the problematic isochrones are the ones that are so affected by the velocity errors that they cannot stay in phase with the others, and this is expected to occur at long offsets where the data could be cycle-skipped. As a result, short-offset or even near-zero offset migration has become a widely accepted and effective approach in the RWI community. Figure 3.12b displays how muting the isochrones mapped with offsets of absolute values greater than 250 m results in a focused stacked image compared to the original stacked image shown in Figure 3.12a.

Stacked images from short- or near-zero offset data often suffer from a low signal-to-noise ratio (SNR) and poor illumination. While this can be resolved through LS iterations, it increases the computational cost. Extending the migration offset is another way to address the issue, although it compromises reflector focusing, as discussed (Figure 3.12). Despite this trade-off, extending the migration offset may be justified, as we already mute short-offset residual waveforms for tomography (to reduce the adverse impact of inconsistent reflectivity and velocity), where reflector unfocusing would cause the most significant damage. The most significant damage occurs at short offsets because there, the reflectivity inaccuracy dominates the velocity inaccuracy due to the shorter propagation time. Therefore, we suggest extending the migration offset. We extend the migration offset to the point where the misfit function no longer significantly decreases. At this offset, the amplitude fit does not get any better as the image is mapped by more and

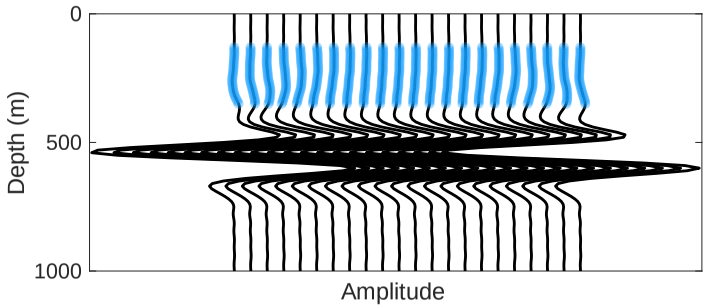


(a) Middle-shot tomographic gradient of ORWI

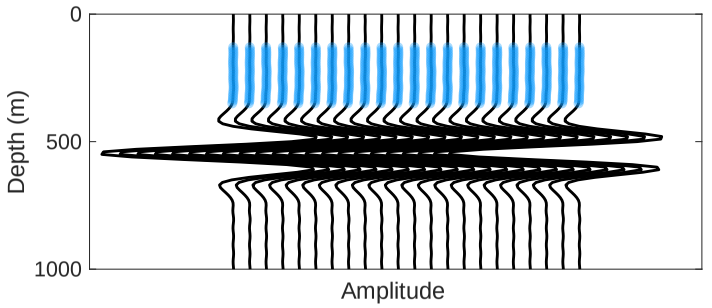


(b) Middle-shot data gather and muted traces based on depth

Figure 3.11: (a) The adverse imprint of inconsistent reflectivity and velocity is evident on the middle-shot tomographic gradient of ORWI after one cycle; the dashed trapezoid represents the updates with incorrect signs for each layer. (b) The gray and yellow areas in the middle-shot data gather corresponding to (a) indicate the sections of the reflection data where full-wave consistency is not satisfied (short offsets), at depths of 0 and 800 meters, respectively. We propose muting the short-offset residual waveforms in the tomography loop based on depth, with muting that increases linearly with depth, to minimize the corresponding erroneous tomographic updates.



(a)



(b)

Figure 3.12: Imaging updates of ORWI for the single-reflector medium (note that on both frames (a) and (b), each wiggle indicates a different lateral location; the lateral locations are ordered from 0 to 5000 m from left to right). (a) Image estimated by the migration mode of ORWI when the data for imaging is cycle-skipped (all migration offsets are included); the image is not focused enough. The lack of focus is evidenced by a phase effect observed in the recovered shape of the reflector and also by non-damped amplitudes preceding the reflector (highlighted in blue). (b) Image estimated by the migration mode of ORWI, where | offsets |> 250 m are muted; the image is focused enough now.

more out-of-phase data due to velocity errors. We refer to this as the maximum effective migration offset (MEMO). The MEMO may vary depending on the migration velocity, acquisition geometry, and frequency content, and averagely sits somewhere between the near-zero offsets and the maximum uncycle-skipped offset in the data. To set the MEMO, we follow the pseudo-code presented in Algorithm 3.2 as a pre-processing/QC step before running ORWI, which calculates the misfit function value associated with different ranges of offsets (e.g., intervals of 250 m) up to the maximum offset available in the acquisition. On the misfit-offset graph, we pick an offset right before the curve starts to become nearly horizontal on a logarithmic scale (see the discussion on the amplitude versus offset effect).

Algorithm 3.2: Migration offset analysis pseudo-code

```
for h ← 0 to max(offset) do
    image← 1x LS migration with h
    p⁻_mod(z_0) ← forward modeling with image and background velocity model
    full-offset residual data← p⁻_obs(z_0) − p⁻_mod(z_0)
    store the misfit function value associated with h
```

3.4.4. TOMOGRAPHY OFFSET: CYCLE-SKIPPED LONG OFFSETS

In every seismic waveform-fitting operation, cycle skipping occurs when the initial velocity model is not kinematically accurate enough in order to model the seismic waveforms with less than half a cycle time shift with respect to the observed waveforms in each iteration, leading to a local minimum situation in the corresponding minimization problem (Virieux and Operto, 2009). Earlier, we elaborated on the damaging effect of the cycle-skipped long-offset data on tomograms. This section presents a data-selection algorithm to minimize such a damaging effect. While employing an L2-norm waveform-based misfit function in ORWI, we develop an automated algorithm to recognize and exclude the cycle-skipped long-offset reflection data in each tomography loop based on the same idea suggested by van Leeuwen and Mulder (2010).

A given pair of modeled and observed seismic traces in the time domain is called cycle-skipped if their corresponding maximum correlation lag is greater than a reference lag, which is defined based on the dominant period of the observed trace. Based on this, we base our time-domain data-selection algorithm on the cross-correlation of the given traces within a sliding window (local cross-correlation), protected by a global cross-correlation of the given traces beforehand (Algorithm 3.3). The global cross-correlation serves as a kind of guard to keep the chain of the subsequent local cross-correlations away from facing the irregularities in the modeled waveform before forming maturely. Otherwise, the outputs of the local cross-correlations would not be accurate enough.

Algorithm 3.3: Uncycle-skipped data extraction pseudo-code

```
// Dobs and Dmod are 3D matrices representing the observed and modeled
   data (full-shot).
// nt× ng× ns is the size of Dobs.
// nt is the total number of time samples.
// ng and ns are the total number of receivers and the total number of
   shots.
// ℓw and ℓr are the correlation-window length and reference lag.
// ωc is the current maximum frequency in D.
// ωm is the peak frequency of the seismic wavelet.
// winc is the current correlation window.
function CycleSkippingCheck(Dobs, Dmod, ng, ns, nt, ℓw, ℓr, ωc, ωm)
    TIMETABLE ← a matrix of ones of size: ng × ns
    for ishot ← 1 to ns do
        φ ← determine the anchor offset index in the current shot gather
        ι ← φ
        κ ← φ − 1
        for m ← [1, 2] do
            flag ← true
            while (m = 1 and flag = true and ι ≤ ng) or (m = 2 and flag = true and κ ≤ 1) do
                dobs, nobs ← current observed trace and the associated first-break sample
                dmod ← current modeled trace
                if ωc > ωm then
                    dobs ← dobs[nobs : nt]
                    dmod ← dmod[nobs : nt]
                else
                    dobs ← dobs[1 : nt]
                    dmod ← dmod[1 : nt]
                calculate cross-correlation between dobs and dmod
                if max correlation lag < ℓr then
                    win ← [1 : ℓw : nt]
                    b ← length(win)
                    for τ ← 1 to b do
                        winc ← [1 : win(τ + 1)]
                        calculate cross-correlation between dobs and dmod in winc
                        if max correlation lag < ℓr then
                            if τ = b then
                                tmp ← nt
                            else
                                tmp ← nobs + winc(end)
                            if m = 1 then
                                TIMETABLE[ι, ishot] ← tmp
                            else
                                TIMETABLE[κ, ishot] ← tmp
                        else
                            flag ← false
                            break
                else
                    flag ← false
                if m = 1 then
                    ι = ι + 1
                else
                    κ = κ − 1
    return TIMETABLE
```

In Algorithm 3.3, for each source location, we run two 'for' loops independently (indicated by the counter "m"), one starting from the receiver index meeting the apex point of the shot gather and increasing, and one starting from the same index but decreasing. In each of these two loops, for each index, we perform global cross-correlation between the modeled and measured traces from the first-break time sample to the last time sample, to calculate the average time delay between the two signals. If the calculated time delay is less than a pre-defined reference time delay (reference lag), then we divide the signals

into windows and perform cross-correlation for each window, until a time delay greater than the reference lag is found. If the time delay is greater than the reference lag, we exit the 'for' loop and move on to the next source location. The time delays calculated during each iteration are stored in a time table. As a result, for a given shot gather, the reflection data aperture is not allowed to be extended from the apex point towards the positive/negative offsets unless it satisfies the relevant global cross-correlation as well as the local cross-correlations to the last window.

3.5. PROPOSED ORWI ALGORITHM

In an effort to resolve the limitations inherent to conventional reflection waveform inversion, as counted earlier and listed in the following, we propose our one-way reflection waveform inversion (ORWI) algorithm based on all of the analyses and discussions that we carried out so far. Our proposed ORWI algorithm incorporates the introduced cost-friendly preconditioned least-squares wave-equation migration (PLS-WEM) algorithm to retrieve high-resolution amplitude-preserved reflectors, thereby overcoming the challenge of low-resolution seismic images with unpreserved amplitudes in conventional reflection waveform inversion. Muting short-offset residual waveforms in the tomography loop is an additional aspect of our proposal that minimizes any full-wave inconsistency due to the fixed reflector's depth positions while updating the background velocity. Leveraging this data mute, we advocate extending the migration offset beyond short offsets by calculating the maximum effective migration offset (MEMO) to address limitations in reflection tomography linked to suboptimal SNR and illumination of reflectors. Furthermore, we propose a data-selection algorithm to exclude the negative impact of cycle-skipped long-offset data. The general features of our proposed ORWI algorithm is summarized in Table 3.2.

Table 3.2: General features of our proposed ORWI.

		Proposed ORWI
1	Each Cycle	LS migration and LS tomography
2	Migration Technology	PLS-WEM
3	Migration Offset	|offset| < MEMO based on Algorithm 3.2
4	Tomography Offset	mid to long
5	Cycle-Skipping in Tomography	Algorithm 3.3

It would now be insightful to compare our proposal (Table 3.2) to when it ignores any offset selection (i.e., it uses full offsets for migration and tomography) and when it uses the industry practice, (i.e., it uses short offsets for migration and full offsets for tomography). To make the comparison, we take the four-layer medium already displayed in Figure 3.10a, as the true model, and we choose a homogeneous 3500 m/s initial background velocity model (Figure 3.10b). We invert the 0-30 Hz reflection data (surface acquisition) all at once to estimate the background velocity model. Figure 3.13 displays the resulting graph

following Algorithm 3.2 for the MEMO calculation, considering the initial background velocity model and the full frequency band. Based on the graph, the MEMO is chosen to be 1000 m. After 15 cycles, each including 1x LS migration and 1x LS tomography, Figure 3.14 exhibits the resultant tomograms.

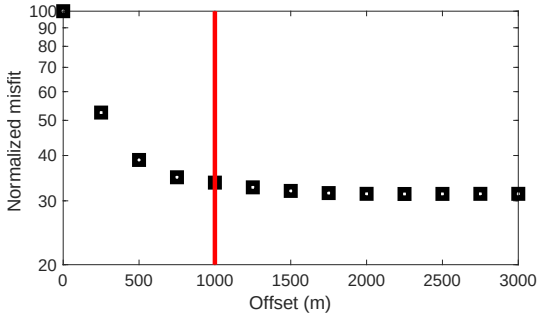


Figure 3.13: Migration offset analysis to extract the MEMO for the four-layer medium; the red line on the curve denotes the MEMO.

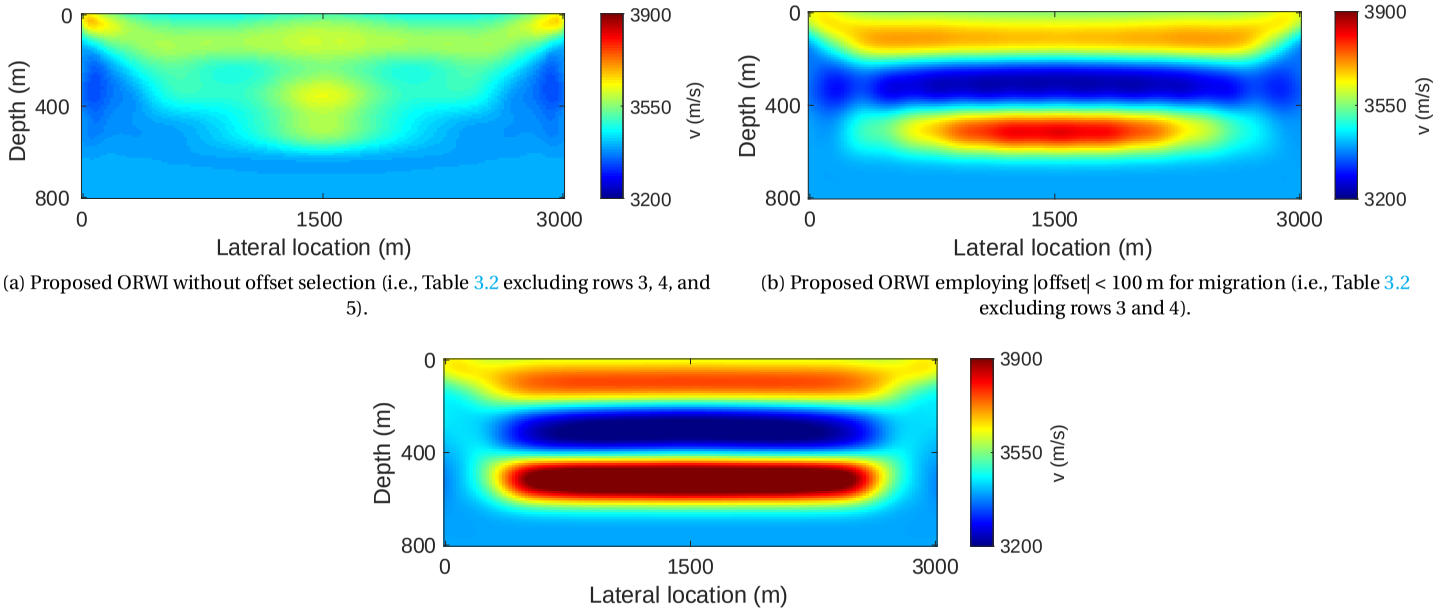
As observed, while our proposed ORWI accurately estimates the background velocity model of the four-layer medium from reflections within 15 cycles (Figure 3.14c), the two other approaches do not retrieve the layers as effectively as our proposal within the same number of cycles (Figures 3.14a and 3.14b). Although the tomogram estimated using the industry practice is a much better estimation than the one obtained without any offset selection, it remains less informative than the tomogram estimated by our proposed ORWI. Note that moving closer to a zero-offset data for migration and increasing the LS iterations for migration can improve the tomogram estimated using the industry practice, as earlier pointed out. This examination shows that the offset selection strategies in our proposed ORWI effectively refine the input data for migration and tomography, reaching the optimal solution in fewer cycles.

3.6. NUMERICAL EXAMPLES

In this section, we apply our proposed ORWI algorithm to three synthetic cases. In all cases, we refine velocity model updates in each cycle using a Gaussian smoothing operator with a standard deviation of 3 (gridpoint).

3.6.1. EXAMPLE 1: TWO-DIPPING-LAYER MODEL

First, we apply our proposed ORWI algorithm on a two-dipping-layer model, where two dipping layers are placed in a linear background velocity model (following Warner et al., 2018). Figures 3.15a and 3.15b show the true and initial background velocity models, respectively. The model is 5 km in width and 1.5 km in depth and is discretized by 251 $\times$ 301 gridpoints with a horizontal grid interval of 20 m and a vertical grid interval of 5 m. As shown, the main task here is to reconstruct the missing layers from the initial background velocity model. The observed dataset, containing only primary reflections, is modeled via PWMod and acquired using 51 surface shots every 100 m and 251 fixed receivers every 20 m. A Ricker wavelet with a peak frequency of 20 Hz is employed as the



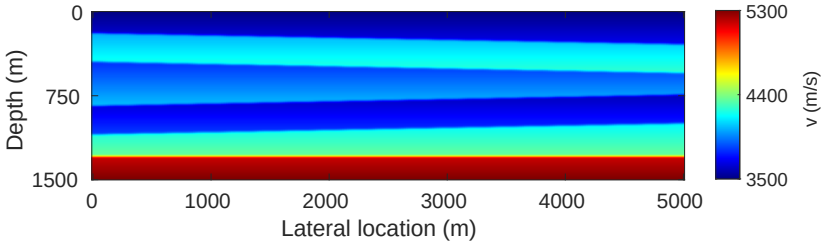
(a) Proposed ORWI without offset selection (i.e., Table 3.2 excluding rows 3, 4, and 5).

(b) Proposed ORWI employing |offset| < 100 m for migration (i.e., Table 3.2 excluding rows 3 and 4).

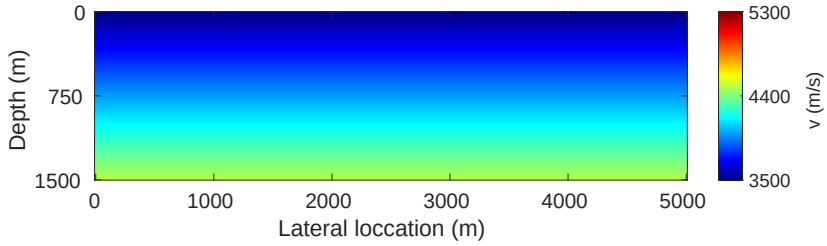
(c) Proposed ORWI (MEMO is set to 1000 m based on Figure 3.13).

Figure 3.14: Tomographic updates associated with the four-layer medium after 15 cycles. (a) Final background velocity model estimated after 15 cycles of the (a) proposed ORWI without any offset selection, (b) proposed ORWI using short-offset migration, and (c) proposed ORWI.

source function. The maximum available offset in the acquisition is limited to 4000 m. We record the reflection dataset for 1.8 s.



(a) True background velocity model



(b) Initial background velocity model

Figure 3.15: Two-dipping-layer model (Example 1). (a) True background velocity model. (b) Initial background velocity model.

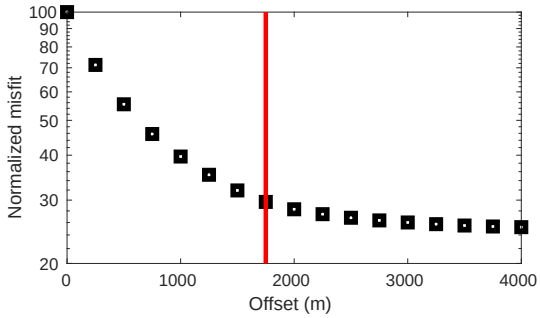


Figure 3.16: Migration offset analysis to extract MEMO (Example 1); the red line on the curve denotes the MEMO.

Figure 3.16 illustrates the migration offset analysis output graph, considering the initial background velocity model and the full frequency band, for this example. According to the misfit-offset graph, we choose the MEMO to be 1750 m (fixed). For tomography, we use offsets larger than 500 m, and any contribution from the cycle-skipped long-offset data into the tomographic gradient is excluded via Algorithm 3.3 in each cycle. As we progress and update the background velocity, additional offsets become relevant in tomography since they are not cycle-skipped anymore. This continues until all offsets are

included. Figure 3.17 shows the data fit prior to inversion, where cycle skipping is obvious in the dataset.

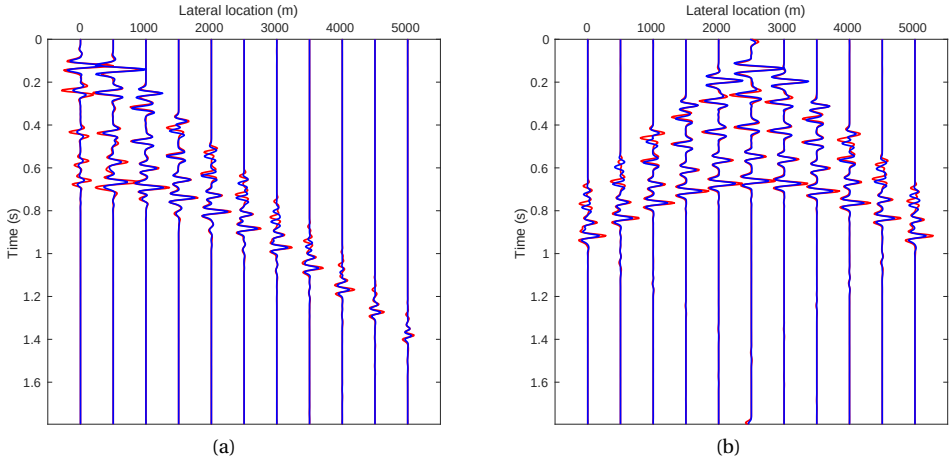


Figure 3.17: Data fit for the first and middle shots calculated in the true background velocity model, shown in red, and the initial background velocity model, shown in blue (Example 1). (a) First shot gather. (b) Middle shot gather.

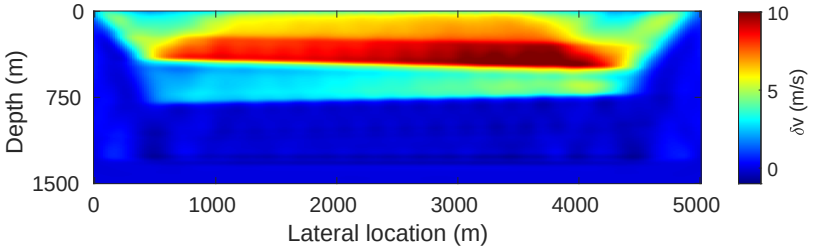


Figure 3.18: Estimated background velocity model perturbation after 1 cycle of the proposed ORWI (Example 1).

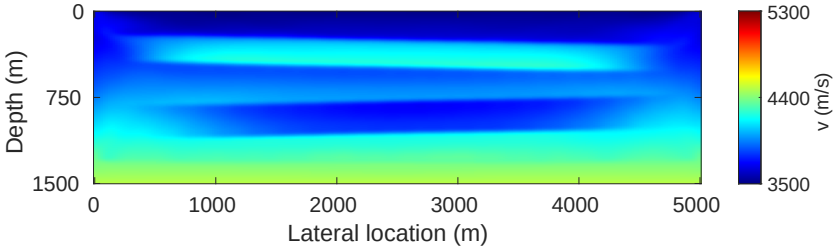
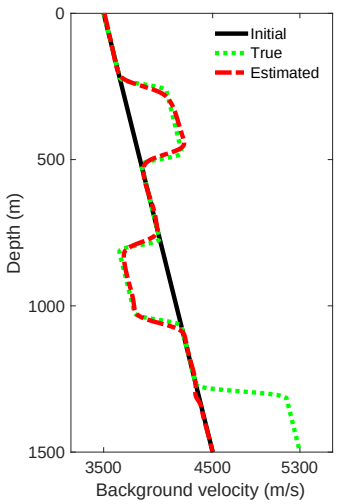
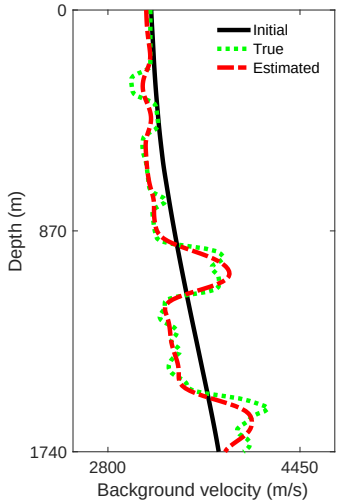


Figure 3.19: Final estimated background velocity model after 45 cycles of the proposed ORWI (Example 1).



(a) Two-dipping-layer model at the lateral location of 2500 m (Example 1)



(b) Flat section of the Marmousi model at the lateral location of 3000 m (Example 2)

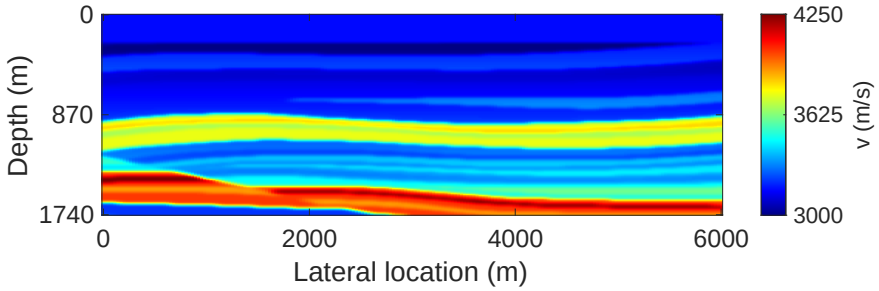
Figure 3.20: Vertical background velocity profiles. A vertical profile of the final estimated background velocity model already shown in (a) Figure 3.19 at the lateral location of 2500 m and (b) Figure 3.25 at the lateral location of 3000 m.

We invert the full-frequency band 0-60 Hz simultaneously (no multi-scaling strategy) in 45 cycles, each including 1x LS migration and 1x LS tomography. Figure 3.18 shows the estimated background velocity model perturbation after 1 cycle. As seen, the reconstructed background velocity perturbation accurately represents the true update direction from the very first cycle. Figure 3.19 shows the final background velocity model estimated after 45 cycles. The result shows a remarkable accuracy in recovering the missing dipping layers in the background velocity. This recovery comes with a high level of vertical resolution. Note that the finite lateral extent of the estimated layers is due to the finite extent of the acquisition. Also, note that we cannot update the deepest layer in the model, below 1300 m, using reflection tomography, as expected. One profile of the final estimated background velocity model is shown in Figure 3.20a. This profile once more illustrates the effective vertical illumination via our proposed ORWI algorithm.

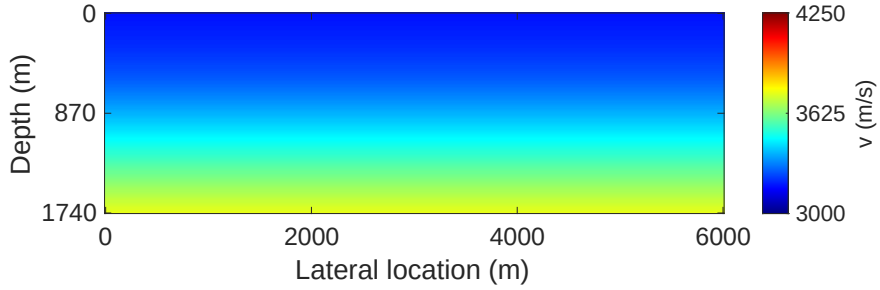
3.6.2. EXAMPLE 2: FLAT SECTION OF THE MARMOUSI MODEL

Next, we verify our proposed ORWI algorithm with a selected part of the Marmousi model that is almost flat. Figure 3.21a shows the true background velocity model, which is 6 km in width and 1,74 km in depth. The model is discretized by 251 × 175 gridpoints with a horizontal grid interval of 30 m and a vertical grid interval of 10 m. To construct the initial background velocity model, after smoothing out the true background velocity model using a 2D Gaussian filter with a standard deviation of 50 (gridpoint), we select a single vertical profile to construct the initial 1D model. Figure 3.21b shows the initial

background velocity model. 41 surface shots are used for acquisition, with a shot spacing of 150 m. There exist 251 fixed receivers distributed on the surface every 10 m. The recording time is 2.2 s. The maximum available offset is limited to 4000 m. Using a Ricker wavelet with a peak frequency of 15 Hz as the source function, the observed reflection dataset, containing only primaries, is generated by PWMod.



(a) True background velocity model



(b) Initial background velocity model

Figure 3.21: Flat section of the Marmousi model (Example 2). (a) True background velocity model. (b) Initial background velocity model.

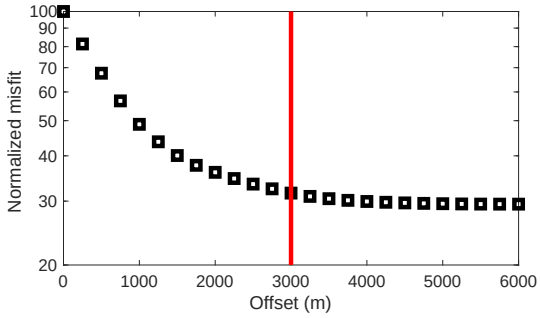


Figure 3.22: Migration offset analysis to extract MEMO (Example 2); the red line on the curve denotes the MEMO.

Figure 3.22 displays the migration offset analysis output graph, considering the initial

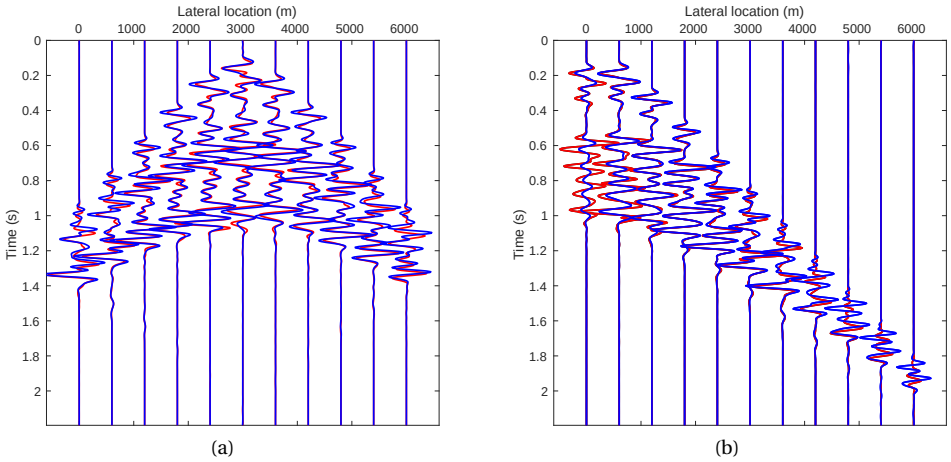


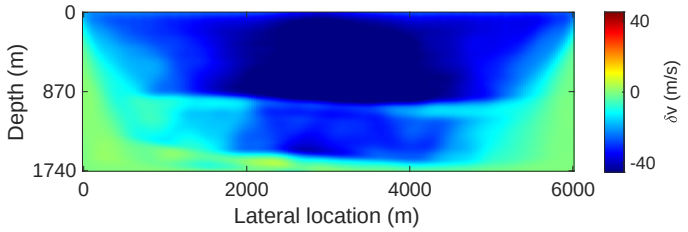
Figure 3.23: Data fit for the first and middle shots calculated in the true background velocity model, shown in red, and the initial background velocity model, shown in blue (Example 2). (a) First shot gather. (b) Middle shot gather.

background velocity model and the full frequency band. Based on the misfit-offset graph, the MEMO is fixed at 3000 m. Comparing the MEMO with the maximum available offset in the acquisition, we can infer that the cycle skipping is not that severe in this example. For tomography, offsets larger than 500 m are employed, and the cycle-skipped long-offset data contribution to the tomographic gradient is excluded in each cycle by Algorithm 3.3. As we advance in cycles and update the background velocity, more and more offsets contribute to tomography because they are no longer cycle-skipped. This process goes on until all offsets are addressed. Figure 3.23 demonstrates the data fit in the first- and middle-shot gathers prior to inversion, confirming the minimum cycle skipping in the dataset.

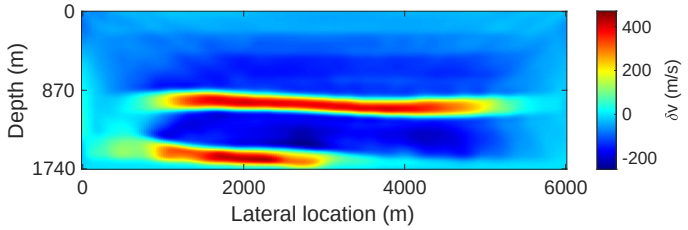
We invert the full-frequency band 2-45 Hz at once (no multi-scaling strategy). We perform 150 cycles of the proposed ORWI, each consisting of 1x LS migration and 1x LS tomography. Figures 3.24a to 3.24c display the cumulative background velocity model perturbation estimated after 1, 50, and 150 cycles. As seen, the recovery of the background velocity model begins with large-scale events and develops towards tiny events. One profile of the final estimated background velocity model is also displayed in Figure 3.20b, revealing a good vertical resolution. Figure 3.25 shows the final background velocity model estimated after 150 cycles of the proposed ORWI, confirming the effectiveness of the proposal.

3.6.3. EXAMPLE 3: FAULTED SECTION OF THE MARMOUSI MODEL

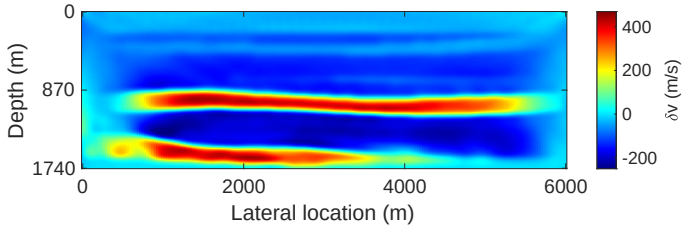
Finally, we validate our proposed ORWI algorithm using the faulted section of the Marmousi model (marine scenario). The true background velocity model, illustrated in Figure 3.26a, spans almost 7 km in width and 2.5 km in depth. This model is discretized into 296 × 104 gridpoints, with a horizontal grid interval of 25 m and a vertical grid interval of 25



(a) Estimated model perturbation after 1 cycle



(b) Estimated model perturbation after 50 cycles



(c) Estimated model perturbation after 150 cycles

Figure 3.24: Estimated background velocity model perturbations (Example 2). (a) Estimated background velocity model perturbation after 1 cycle. (b) Cumulative estimated background velocity model perturbation after 50 cycles. (c) Cumulative estimated background velocity model perturbation after 150 cycles.

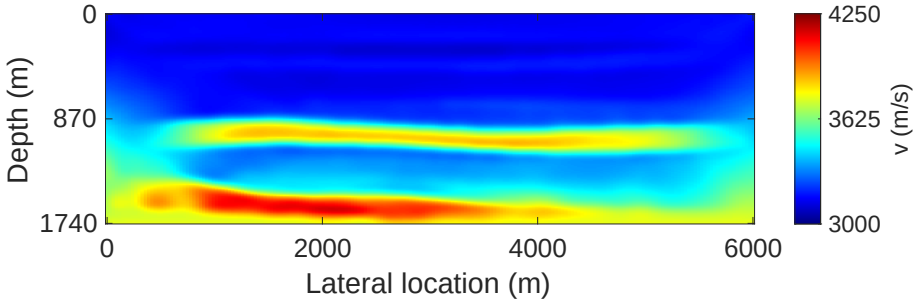
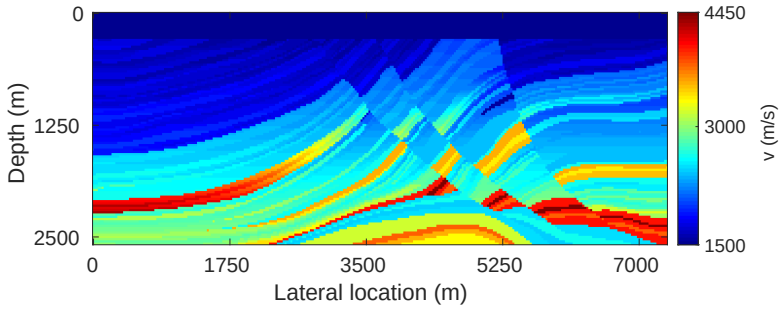


Figure 3.25: Final estimated background velocity model after 150 cycles of the proposed ORWI (Example 2).

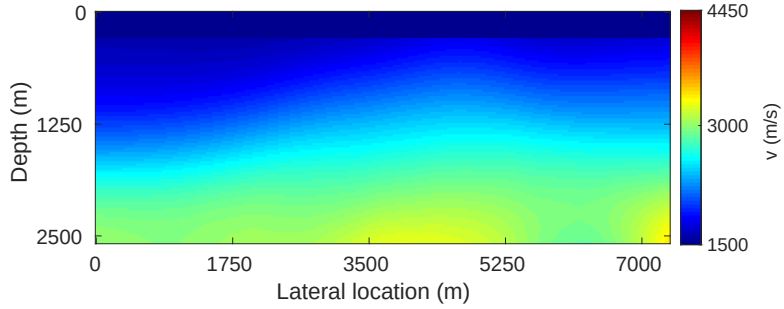
m. To create the initial background velocity model, we apply a 2D Gaussian filter with a standard deviation of 20 (gridpoint) to construct the true background velocity model. The

resulting initial background velocity model with a water layer on top is depicted in Figure 3.26b. For acquisition, 31 surface shots are used on the surface (the shots are spaced either 250 m apart or 225 m apart), and 296 fixed receivers are distributed on the surface at 25 m intervals. The recording time is 4.092 s, and the maximum available offset is limited to 4000 m. Using a Ricker wavelet with a peak frequency of 10 Hz as the source function, we generate the observed reflection dataset, only primaries, through PWMod. Figure 3.27 illustrates the output of migration offset analysis, considering the initial background velocity model and the full frequency band, through the misfit-offset graph. The MEMO is set at 2000 m based on the misfit-offset graph. In the tomography process, offsets exceeding 500 m are employed, and the contribution from the cycle-skipped long-offset to the tomographic gradient is excluded in each cycle through Algorithm 3.3. As we progress through cycles and update the background velocity, an increasing number of offsets contribute to tomography since they are no longer cycle-skipped. This process carries on until all offsets have been accounted for.

The inversion here employs a multi-scaling strategy, which starts from a low frequency of 2 Hz and gradually increases up to 21 Hz. Table 3.3 displays the frequency strategy used in the inversion, presenting details on the frequency ranges, number of total cycles, and number of iterations for both LS migration and LS tomography within each cycle. Note that the reflectivity model is reset back to zero after each frequency scale. Figures 3.28a



(a) True background velocity model



(b) Initial background velocity model

Figure 3.26: Faulted section of the Marmousi model (Example 3). (a) True background velocity model. (b) Initial background velocity model.

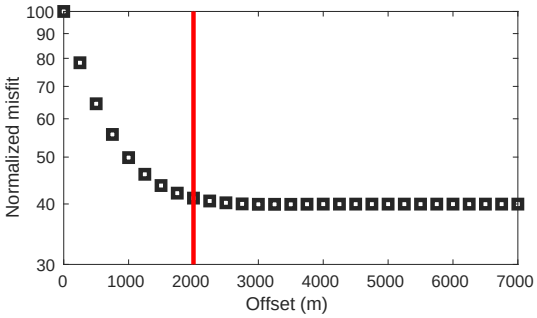


Figure 3.27: Migration offset analysis to extract MEMO (Example 3); the red line on the curve denotes the MEMO.

to 3.28c show the estimated background velocity models estimated after the frequency ranges of 2-7 Hz, 2-14 Hz, and 2-21 Hz, respectively, and Figures 3.29a to 3.29c present the stacked images mapped using the true, initial estimated, and final estimated background velocity models, respectively.

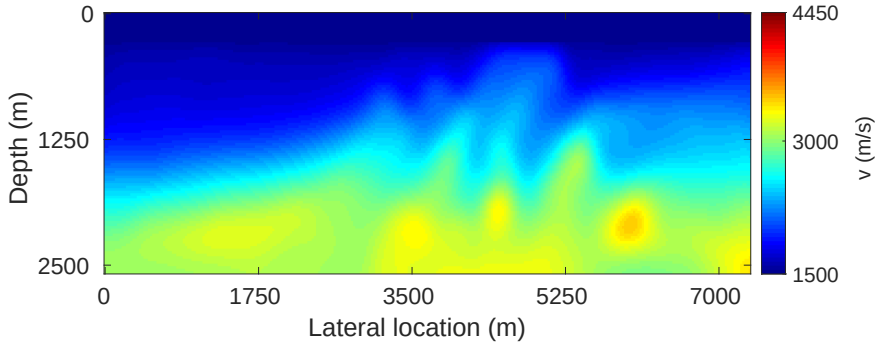
Table 3.3: Frequency strategy employed with Example 3.

	2-3 Hz	**2-5 Hz**	**2-7 Hz**	**2-14 Hz**	**2-21 Hz**
Total cycle	4	4	4	20	25
LS migration iteration in each cycle	5	5	5	3	2
LS tomography iteration in each cycle	5	5	5	3	2
Image reset to zero after each cycle	yes	yes	yes	yes	yes

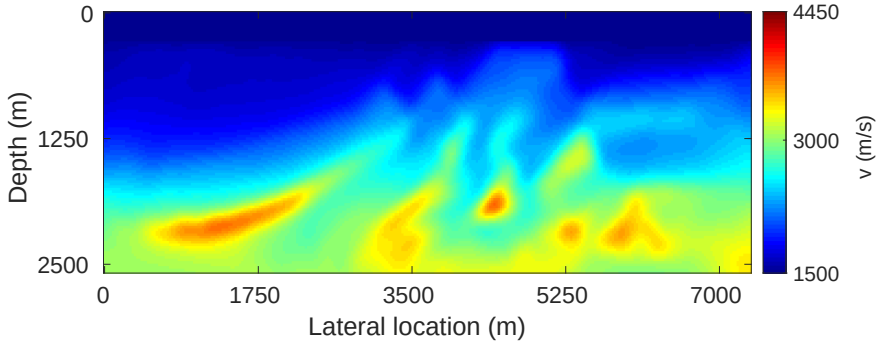
As can be seen in Figures 3.28 and 3.29, while not flawless, particularly in the deep right of the estimated model marked with a dashed ellipse in Figure 3.28c, the kinematic properties of the background velocity model during the inversion are improved significantly to incorporate the essential traveltime information for accurate depth migration. For a more precise evaluation of the overall accuracy in estimations, Figure 3.30 compares the image difference between the true and initial image, as well as the difference between the true and final estimated image.

3.7. DISCUSSION

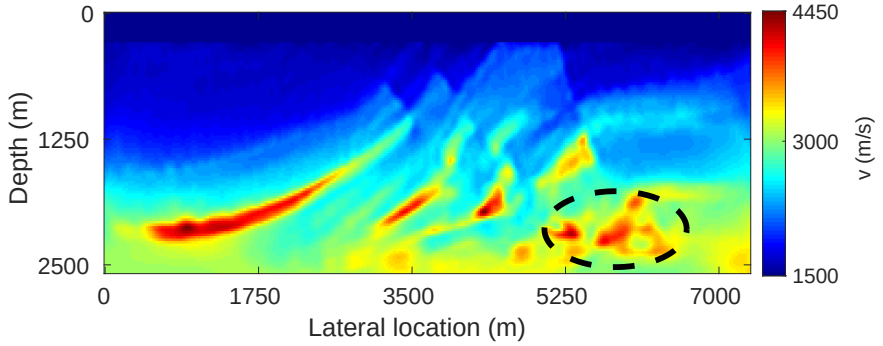
In the field of exploration seismology, FWI has been widely regarded as the ultimate solution for subsurface property reconstruction since its introduction. However, it was later discovered by the FWI community that the reflection-associated tomographic wavepaths in the FWI kernel were too weak to update deep low-wavenumber properties beyond the reach of refractions. In response to this limitation, RWI was developed to make full use of the reflection-associated tomographic wavepaths. Despite its potential, RWI has its



(a) 2-7 Hz

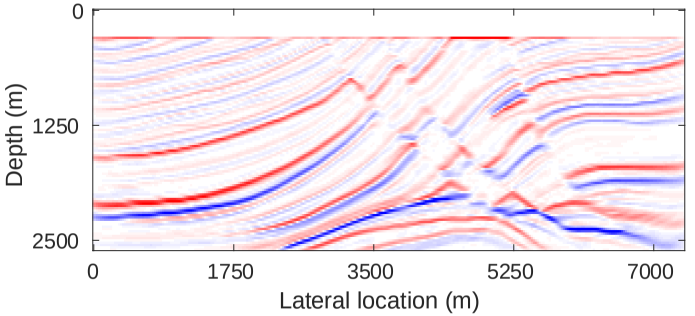


(b) 2-14 Hz

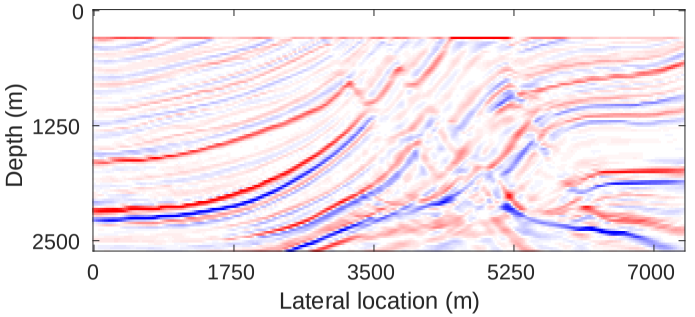


(c) 2-21 Hz

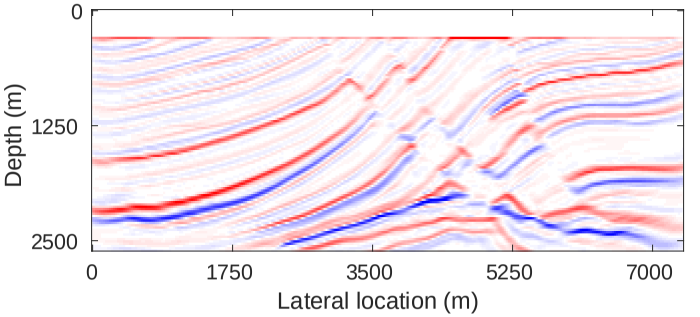
Figure 3.28: Estimated background velocity models using different frequency ranges (Example 3). (a) Estimated background velocity model with the frequency range of 2-7 Hz. (b) Estimated background velocity model with the frequency range of 2-14 Hz. (c) Estimated background velocity model with the frequency range of 2-21 Hz.



(a) True image mapped using the true background velocity model

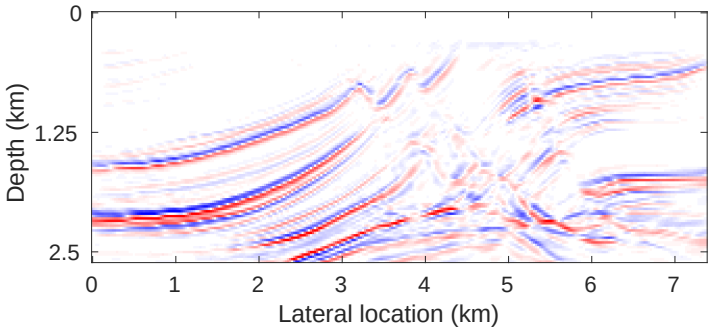


(b) Initial image mapped using the initial background velocity model

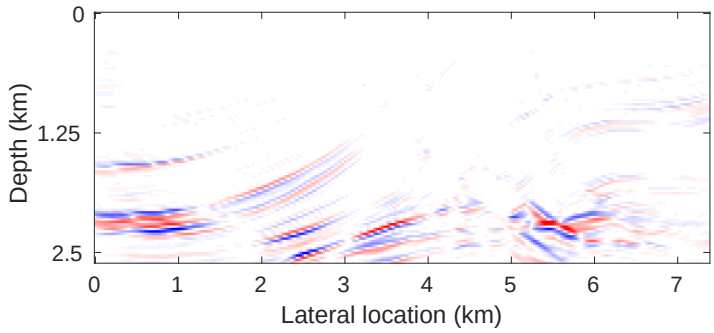


(c) Final estimated image mapped using the final estimated background velocity model

Figure 3.29: Estimated images using different background velocity models (Example 3). Estimated images using (a) the true background velocity model, (b) the initial background velocity model, and (c) the final estimated background velocity model.



(a) True image - Initial image



(b) True image - Final estimated image

Figure 3.30: Image difference. (a) Difference between the images mapped by the true and initial background velocity models. (b) Difference between the images mapped by the true and final estimated background velocity models.

own limitations, as counted earlier. To address these limitations, in the present study, we analyzed the current state of RWI in the context of one-way reflection waveform inversion (ORWI).

We acknowledge that our proposed solution for addressing inconsistent reflectivity and velocity models in ORWI is an intermediate approach and may not be perfect. It could be considered a "dirty solution" that requires further investigation and optimization. Other promising alternatives for addressing inconsistent reflectivity and velocity in RWI include pseudo-time domain RWI, as well as approaches proposed by Valensi and Baina (2021) and Liang et al. (2022).

In this research, we chose not to introduce any model-based regularization to our tomography misfit function. However, we believe that adding a model-based regularization, particularly a structure-oriented type of regularization guided by the reflectivity model (Masaya and Verschuur, 2018; Yao et al., 2019b; Provenzano et al., 2023), could positively influence the overall performance of the tomographic updates. We plan to explore this possibility in future investigations.

The multi-scaling strategy, in the context of reflection waveform inversion, involves inverting the reflection waveforms at multiple scales or frequencies, starting from low

frequencies towards higher frequencies. This is done to mitigate the effect of the cycle-skipping issue in the long-offset waveforms in the context of RWI. With Example 3, we showed that a multi-scaling strategy may be beneficial in ORWI wherever the full-band reflection dataset is too complex to be resolved tomographically with the reflection wavepaths all at once. The same is reported in Chen et al. (2020).

Another important aspect for discussion is the AVO effect. The ORWI process, given its angle-independent properties, cannot reproduce AVO on the data. Indeed, an angle-independent reflectivity model is insufficient for capturing AVO (Davydenko and Verschuur, 2017; Farshad and Chauris, 2020). Therefore, we propose to reduce AVO in the data before applying the ORWI methodology, as described in Qu and Verschuur (2021). The examples in this paper were created with the inverse crime approach; that is, we consider the optimum situation where AVO effects do not play a role, such that we can focus on the convergence aspects of ORWI.

There are additional challenges in real data applications. An accurate estimation of the source signal is essential. If there is a mismatch in the source signal, it will limit the resolution in imaging and, thereby, tomography. As the 2D field data still has propagated in a 3D world, amplitudes have to be corrected (3D to 2D transformation), and out-of-plane effects may not be well addressed by 2D ORWI. Naturally, when ORWI is applied in a full 3D sense to 3D field data, those amplitude and out-of-plane effects will no longer be present. Another challenge with field data is the potential presence of the anisotropy effect. While anisotropy effects can be included in ORWI, updating the anisotropy parameters is not trivial (Alshuhail and Verschuur, 2019). The Q-effect in the data can represent another challenge, and this can also be integrated into the ORWI process, as explained by Safari and Verschuur (2023).

Finally, the ORWI method can be extended to the 3D case, where we have wavefields measured along inline and crossline directions. The approximate Hessian matrix for the entire medium will be a 3D block-diagonal matrix, where each block exhibits non-zero elements on and around the main diagonals in three directions, forming a 3D band-diagonal matrix for each virtual data plane, which is still manageable to invert. It will, of course, incur higher computational costs, akin to the entire ORWI method, which is also more expensive in 3D.

3.8. CONCLUSION

This paper addressed reflection waveform inversion limitations to recover an improved background velocity model update. To do so, in the context of one-way reflection waveform inversion (ORWI), we replaced the least-squares wave-equation migration technology with an efficient preconditioned least-squares wave-equation migration technology to retrieve high-resolution reflectors with preserved amplitudes. To mitigate the adverse imprint of inconsistent reflectivity and velocity models (i.e., full-wave inconsistency in the short-offset data for tomography), we muted short-offset waveforms in the residual data gathers to suppress their associated erroneous tomographic wavepaths before they could establish in the tomographic gradient of ORWI. Thanks to the muting of short-offset residual waveforms, we adopted an extended migration offset (MEMO) rather than short offset to avoid limitations in reflection tomography linked to suboptimal SNR and illumination of the reflectivity model. We also proposed a data-selection algorithm to exclude

the damaging effect of the cycle-skipped long-offset data from the tomographic gradient of ORWI. The results of three numerical examples demonstrated the effectiveness of the proposed approach in recovering high-fidelity tomograms with good vertical resolution.

REFERENCES

Abolhassani, S. and E. Verschuur (2022). "Fast Gauss-Newton full-wavefield migration". In: SEG/AAPG Second International Meeting for Applied Geoscience & Energy, Expanded Abstracts.

Abolhassani, S. and D. J. Verschuur (2024). "Efficient preconditioned least-squares wave-equation migration". In: *Geophysics* 89.3, S275–S288.

Alkhalifah, T. (2014). "Scattering-angle based filtering of the waveform inversion gradients". In: *Geophysical Journal International* 200.1, pp. 363–373.

Almomin, A. and B. Biondi (2012). "Tomographic full waveform inversion: Practical and computationally feasible approach". In: SEG Annual International Meeting, Expanded Abstracts.

Alshuhail, A. A. and D. J. Verschuur (2019). "Robust estimation of vertical symmetry axis models via joint migration inversion: Including multiples in anisotropic parameter estimation". In: *Geophysics* 84.1, pp. C57–C74.

Assis, C. A. and J. Schleicher (2021). "Introduction of the Hessian in joint migration inversion and improved recovery of structural information using image-based regularization". In: *Geophysics* 86 (6), R777–R793.

Assis, C. A. et al. (2024). "Investigating Hessian-based inversion velocity analysis". In: *Geophysics* 89.2, R169–R186.

Audebert, F. and E. Cocher (2020). "Accounting for the depth-velocity coupling in reflection full waveform inversion (RFWI)". In: SEG Annual International Meeting, Expanded Abstracts.

Baina, R. and R. Valensi (2018). "Borrowing insight from travel-time reflection tomography to solve the depth reflectivity-velocity coupling issue of RFWI". In: 80th EAGE Annual Conference & Exhibition, Workshops.

Berkhout, A. J. (1982). *Seismic migration, imaging of acoustic energy by wave field extrapolation, A: theoretical aspects*. Elsevier.

— (2012). "Combining full wavefield migration and full waveform inversion, a glance into the future of seismic imaging". In: *Geophysics* 77.2, S43–S50.

— (2014a). "Review paper: An outlook on the future seismic imaging, part I: Forward and reverse modeling". In: *Geophysical Prospecting* 62.5, pp. 911–930.

— (2014b). "Review paper: An outlook on the future seismic imaging, part III: Joint Migration Inversion". In: *Geophysical Prospecting* 62.5, pp. 950–971.

Beydoun, W. B. and M. Mendes (1989). "Elastic ray-Born l_2-migration/inversion". In: *Geophysical Journal International* 97.1, pp. 151–160.

Biondi, B. and A. Almomin (2012). "Tomographic full waveform inversion (TFWI) by combining full waveform inversion with waveequation migration velocity analysis". In: SEG Annual International Meeting, Expanded Abstracts.

Bozdağ, E., J. Trampert, and J. Tromp (2011). "Misfit functions for full waveform inversion based on instantaneous phase and envelope measurements". In: *Geophysical Journal International* 185.2, pp. 845–870.

Brittan, J. and I. Jones (2019). "FWI evolution — From a monolith to a toolkit". In: *The Leading Edge* 38.3, pp. 179–184.

Bunks, C. et al. (1995). "Multiscale seismic waveform inversion Geophysics". In: *Geophysics* 60.5, pp. 1457–1473.

Chauris, H. and M. Noble (2001). "Two-dimensional velocity macro model estimation from seismic reflection data by local differential semblance optimization: applications to synthetic and real data sets". In: *Geophysical Journal International* 144.1, pp. 14–26.

Chauris, H. and E. Cocher (2017). "From migration to inversion velocity analysis". In: *Geophysics* 82 (3), S207–S223.

Chauris, H., D. Donno, and H. Calandra (2012). "Velocity estimation with the normalized integration method". In: 74th EAGE Conference & Exhibition incorporating EUROPEC, Extended Abstracts.

Chauris, H. and E. Cocher (2018). "Review of different expressions for the extended Born approximate inverse operator". In: 80th EAGE Annual Conference & Exhibition, Workshops.

Chavent, G. and R. E. Plessix (1999). "An optimal true-amplitude least-squares prestack depth-migration operator". In: *Geophysics* 64.2, pp. 508–515.

Chen, G.-X., R.-S. Wu, and S.-C. Chen (2018). "Reflection multi-scale envelope inversion". In: *Geophysical Prospecting* 66.7, pp. 1258–1271.

Chen, Y. et al. (2020). "Multiscale reflection phase inversion with migration deconvolution". In: *Geophysics* 85.1, R55–R73.

Chi, B., K. Gao, and L. Huang (2017). "Least-squares reverse time migration guided full-waveform inversion". In: SEG Annual International Meeting, Expanded Abstracts.

Chi, B., L. Dong, and Y. Liu (2014). "Full waveform inversion method using envelope objective function without low frequency data". In: *Journal of Applied Geophysics* 109, pp. 36–46.

Choi, Y. and T. Alkhalifah (2014). "Two-dimensional unwrapped phase inversion with damping and a Gaussian filter". In: 76th EAGE Conference & Exhibition, Extended Abstracts.

Choi, Y. and T. Alkhalifah (2013). "Frequency-domain waveform inversion using the phase derivative". In: *Geophysical Journal International* 195.3, pp. 1904–1916.

Choi, Y., T. Alkhalifah, and B. DeVault (2015). "Application of the unwrapped phase inversion to land data without source estimation". In: SEG Annual International Meeting, Expanded Abstracts.

Choi, Y., D.-J. Min, and C. Shin (2008). "Frequency-domain elastic full waveform inversion using the new pseudo-Hessian matrix: Experience of elastic Marmousi-2 synthetic data". In: *Bulletin of the Seismological Society of America* 98.5, pp. 2402–2415.

Davydenko, M. and D. J. Verschuur (2017). "Full-wavefield migration: Using surface and internal multiples in imaging". In: *Geophysical Prospecting* 65.1, pp. 7–21.

Dong, L. et al. (2018). "Correlation-based reflection waveform inversion by one-way wave equations". In: *Geophysical Prospecting* 66.8, pp. 1503–1520.

Donno, D., H. Chauris, and H. Calandra (2013). "Estimating the background velocity model with the normalized integration method". In: 75th EAGE Conference & Exhibition incorporating SPE EUROPEC, Extended Abstracts.

Farshad, M. and H. Chauris (2020). "From constant-to variable-density inverse extended Born modeling". In: *Geophysics* 85.4, S217–S232.
Fletcher, R. P. et al. (2016). "Least-squares migration — Data domain versus image domain using point spread functions". In: *The Leading Edge* 35.2, pp. 157–162.
Gomes, A. and Z. Yang (2018). "Improving reflection FWI reflectivity using LSRTM in the curvelet domain". In: SEG Annual International Meeting, Expanded Abstracts.
Guitton, A. (2004). "Amplitude and kinematic corrections of migrated images for nonunitary imaging operators". In: *Geophysics* 69.4, pp. 1017–1024.
Guo, Q. and T. Alkhalifah (2017). "Elastic reflection-based waveform inversion with a nonlinear approach". In: *Geophysics* 82.6, R309–R321.
Hassine, A. B. et al. (2022). "One-way waveform inversion (OWI)". In: 83rd EAGE Annual Conference & Exhibition, Extended Abstracts.
Hoop, M. V. de and R. D. van Der Hilst (2005). "On sensitivity kernels for wave-equation transmission tomography". In: *Geophysical Journal International* 160.2, pp. 621–633.
Hou, J. and W. Symes (2016). "Approximate Gauss-Newton iteration for full-waveform inversion". In: SEG Annual International Meeting, Expanded Abstracts., pp. 1163–1168.
Hou, J. and W. W. Symes (2015). "An approximate inverse to the extended Born modeling operator". In: *Geophysics* 80 (6), R331–R349.
Hu, J., G. T. Schuster, and P. Valasek (2001). "Poststack migration deconvolution". In: *Geophysics* 66.3, pp. 939–952.
Irabor, K. and M. Warner (2016). "Reflection FWI". In: SEG Annual International Meeting, Expanded Abstracts.
Jang, U., D.-J. Min, and C. Shin (2009). "Comparison of scaling methods for waveform inversion". In: *Geophysical Prospecting* 57.1, pp. 49–59.
Jones, I. F. (2019). "Tutorial: The mechanics of waveform inversion". In: *First Break* 37.5, pp. 31–43.
Kazei, V., E. Tessmer, and T. Alkhalifah (2016). "Scattering angle-based filtering via extension in velocity". In: SEG Annual International Meeting, Expanded Abstracts.
Kiyashchenko, D. et al. (2007). "A modified imaging principle for true-amplitude wave-equation migration". In: *Geophysical Journal International* 168.3, pp. 1093–1104.
Lameloise, C.-A., H. Chauris, and M. Noble (2015). "Improving the gradient of the image-domain objective function using quantitative migration for a more robust migration velocity analysis". In: *Geophysical Prospecting* 63.2, pp. 391–404.
Lecomte, I. (2008). "Resolution and illumination analyses in PSDM: A ray-based approach". In: *The Leading Edge* 27.5, pp. 650–663.
Li, Y. et al. (2019). "Elastic reflection waveform inversion with variable density". In: *Geophysics* 84.4, R553–R567.
Lian, S. et al. (2018). "Enhancing low-wavenumber components of full-waveform inversion using an improved wavefield decomposition method in the time-space domain". In: *Journal of Applied Geophysics* 157, pp. 10–22.
Liang, H., H. Zhang, and H. Liu (2022). "Analysis of the impact of demigration on traveltime-based reflection full-waveform inversion". In: Second International Meeting for Applied Geoscience & Energy, SEG/AAPG, Expanded Abstracts.

Lines, L. and S. Treitel (1984). "A review of least-squares inversion and its application to geophysical problems". In: *Geophysical Prospecting* 32.2, pp. 159–186.
Liu, F. et al. (2011). "An effective imaging condition for reverse-time migration using wavefield decomposition". In: *Geophysics* 76.6, S29–S39.
Liu, F. et al. (2012). "3D time-domain full waveform inversion of a Valhall OBC dataset". In: SEG Annual International Meeting, Expanded Abstracts.
Lu, S. et al. (2018). "Least-squares full-wavefield migration". In: *The Leading Edge* 37.1, pp. 46–51.
Luo, Y. and G. T. Schuster (1991). "Wave-equation traveltime inversion". In: *Geophysics* 56.5, pp. 645–653.
Luo, Y. et al. (2016). "Full-traveltime inversion". In: *Geophysics* 81.5, R261–R274.
Ma, Y. and D. Hale (2013). "Wave-equation reflection traveltime inversion with dynamic warping and full-waveform inversion". In: *Geophysics* 78.6, R223–R233.
Masaya, S. and D. Verschuur (2018). "Iterative reflectivity-constrained velocity estimation for seismic imaging". In: *Geophysical Journal International* 214.1, pp. 1–13.
Metivier, L. et al. (2013). "Full waveform inversion and the truncated Newton method". In: *SIAM Journal on Scientific Computing* 35.2, B401–B437.
Métivier, L. et al. (2016). "Measuring the misfit between seismograms using an optimal transport distance: application to full waveform inversion". In: *Geophysical Journal International* 205.1, pp. 345–377.
Mora, P. (1989). "Inversion = migration + tomography". In: *Geophysics* 54.12, pp. 1575–1586.
Oh, J.-W. and D.-J. Min (2013). "Weighting technique using backpropagated wavefields incited by deconvolved residuals for frequency-domain elastic full waveform inversion". In: *Geophysical Journal International* 194.1, pp. 322–347.
Plessix, R. E. and W. A. Mulder (2004). "Frequency-domain finite-difference amplitude-preserving migration". In: *Geophysical Journal International* 157.3, pp. 975–987.
Plessix, R. E. et al. (2013). "Multiparameter full waveform inversion: marine and land examples". In: *The Leading Edge* 32.9, pp. 1030–1038.
Pratt, R. G., C. Shin, and G. J. Hicks (1998). "Gauss–Newton and full Newton methods in frequency-space seismic waveform inversion". In: *Geophysical Journal International* 133.2, pp. 341–362.
Provenzano, G., R. Brossier, and L. Métivier (2023). "Robust and efficient waveform-based velocity-model building by optimal transport in the pseudotime domain: Methodology". In: *Geophysics* 88.2, U49–U70.
Qin, B. and G. Lambaré (2016). "Joint inversion of velocity and density in preserved-amplitude full-waveform inversion". In: SEG Annual International Meeting, Expanded Abstracts.
Qu, S. and D. Verschuur (2021). "An effective scheme of pseudo-time Joint migration inversion with an AVO mitigating workflow". In: 82th EAGE Annual Conference & Exhibition, Extended Abstracts.
Qu, S., Y. van den Brule, and D. J. Verschuur (2020). "A stable scheme of joint migration inversion in the pseudo-time domain". In: EAGE Annual Conference & Exhibition Online, Extended Abstracts.

Romahn, S., K. A. Innanen, and G. F. Margrave (2021). "Log-validated waveform inversion of reflection seismic data with wavelet phase and amplitude updating". In: *Geophysics* 86.4, R529–R546.
Safari, M. and J. Verschuur (2023). "Joint Migration Inversion including Q effects: towards Q estimation". In: 84th EAGE Annual Conference & Exhibition, Extended Abstracts.
Sava, P. and B. Biondi (2004). "Wave-equation migration velocity analysis. I. Theory". In: *Geophysical Prospecting* 52.6, pp. 593–606.
Shen, P. (2005). "Wave equation migration velocity analysis by differential semblance optimization". PhD thesis. Rice University.
Shen, P. and W. W. Symes (2008). "Automatic velocity analysis via shot profile migration". In: *Geophysics* 73.5, VE49–VE59.
Shen, P., W. W. Symes, and C. C. Stolk (2003). "Differential semblance velocity analysis by wave-equation migration". In: SEG Annual International Meeting, Expanded Abstracts.
Shin, C., C. S. Jang, and D.-J. Min (2001). "Improved amplitude preservation for prestack depth migration by inverse scattering theory". In: *Geophysical Prospecting* 49.5, pp. 592–606.
Son, M. et al. (2013). "Time domain full waveform inversion using a time-window and Huber function norm". In: *Journal of Seismic Exploration* 22.4, pp. 311–338.
Soubaras, R. et al. (2023). "Fast-track simultaneous velocity and reflectivity inversion using one-way modelling: a North Sea example". In: 84th EAGE Annual Conference & Exhibition.
Sun, D. and W. Symes (2013). "Waveform inversion via nonlinear differential semblance optimization". In: 75th EAGE Conference & Exhibition incorporating SPE EUROPEC, Workshops.
Sun, D. et al. (2016). "Reflection-based waveform inversion". In: SEG Annual International Meeting, Expanded Abstracts.
Sun, Verschuur, and Qu (2019). "Research note: Derivations of gradients in angle-independent joint migration inversion". In: *Geophysical Prospecting* 67.3, pp. 572–579.
Symes, W. W. (2008). "Migration velocity analysis and waveform inversion". In: *Geophysical Prospecting* 56.6, pp. 765–790.
Tang, Y. et al. (2013). "Tomographically enhanced full wavefield inversion". In: SEG Annual International Meeting, Expanded Abstracts.
Tarantola, A. (1984). "Inversion of seismic reflection data in the acoustic approximation". In: *Geophysics* 49.8, pp. 1259–1266.
ten Kroode, F. (2014). "A Lie group associated to seismic velocity estimation". In: Inverse Problems from Theory to Applications Conference, Institute of Physics Publishing, 142-146.
ten Kroode, F. (2012). "A wave-equation-based Kirchhoff operator". In: *Inverse Problems* 28.11, p. 115013.
Thorbecke, J. W., K. Wapenaar, and G. Swinnen (2004). "Design of one-way wavefield extrapolation operators, using smooth functions in wlsq optimization". In: *Geophysics* 69.4, pp. 1037–1045.
Valenciano, A. A., B. Biondi, and A. Guitton (2006). "Target-oriented wave-equation inversion". In: *Geophysics* 71 (4), A35–A38.

Valensi, R. and R. Baina (2021). “A time consistent waveform inversion (TWIN) method”. In: 82nd EAGE Annual Conference & Exhibition, Extended Abstracts.

van Leeuwen, T. and W. A. Mulder (2010). “A correlation-based misfit criterion for wave-equation traveltime tomography”. In: *Geophysical Journal International* 182.3, pp. 1383–1394.

Verschuur, D. J., X. R. Staal, and A. J. Berkhout (2016). “Joint migration inversion: Simultaneous determination of velocity fields and depth images using all orders of scattering”. In: *The Leading Edge* 35.12, pp. 1037–1046.

Vigh, D., J. Kapoor, and H. Li (2011). “Full-waveform inversion application in different geological settings”. In: SEG Annual International Meeting, Expanded Abstracts.

Vigh, D. et al. (2016). “Earth-model building from shallow to deep with full-waveform inversion”. In: *The Leading Edge* 35.12, pp. 1025–1030.

Virieux, J. and S. Operto (2009). “An overview of full-waveform inversion in exploration geophysics”. In: *Geophysics* 74.6, WCC1–WCC26.

Wang, F. et al. (2013). “Taking Advantage of Wave Field Decomposition in Full Waveform Inversion”. In: 75th EAGE Conference & Exhibition incorporating SPE EUROPEC, Extended Abstracts.

Wang, P. et al. (2018). “A demigration-based reflection full-waveform inversion workflow”. In: SEG Annual International Meeting, Expanded Abstracts.

Warner, M. et al. (2018). “High-resolution reflection FWI”. In: 80th EAGE Annual Conference & Exhibition, Extended Abstracts.

Wu, R. .-., J. Luo, and B. Wu (2014). “Seismic envelope inversion and modulation signal model”. In: *Geophysics* 79.6, WA13–WA24.

Wu, Z. and T. Alkhalifah (2015). “Simultaneous inversion of the background velocity and the perturbation in full-waveform inversion”. In: *Geophysics* 80.6, R317–R329.

— (2017). “Efficient scattering-angle enrichment for a nonlinear inversion of the background and perturbations components of a velocity model”. In: *Geophysical Journal International* 210.3, pp. 1981–1992.

Xie, X. (2015). “An angle-domain wavenumber filter for multi-scale full-waveform inversion”. In: SEG Annual International Meeting, Expanded Abstracts.

Xu, S. et al. (2012a). “Inversion on reflected seismic wave”. In: SEG Annual International Meeting, Expanded Abstracts.

Xu, S. et al. (2012b). “Full waveform inversion for reflected seismic data”. In: 74th EAGE Annual Conference & Exhibition.

Yang, Y. and B. Engquist (2018). “Analysis of optimal transport and related misfit functions in full-waveform inversion”. In: *Geophysics* 83.1, A7–A12.

Yao, G., D. Wu, and S. Wang (2020). “A review on reflection-waveform inversion”. In: *Petroleum Science* 17, pp. 334–351.

Yao, G. et al. (2018). “Separation of Migration and Tomography Modes of Full-Waveform Inversion in the Plane Wave Domain”. In: *Journal of Geophysical Research: Solid Earth* 123.2, pp. 1486–1501.

Yao, G. et al. (2019a). “Extraction of the tomography mode with nonstationary smoothing for full-waveform inversion”. In: *Geophysics* 84.4, R527–R537.

Yao, G., N. V. da Silva, and D. Wu (2019b). "Reflection-waveform inversion regularized with structure-oriented smoothing shaping". In: *Pure and Applied Geophysics* 176.12, pp. 5315–5335.

Yu, J. et al. (2006). "Prestack migration deconvolution". In: *Geophysics* 71 (2), S53–S62.

Zhang, Y. et al. (2007). "True-amplitude, angle-domain, common-image gathers from one-way wave-equation migrations". In: *Geophysics* 72 (1), S49–S58.

Zhou, H., L. Amundsen, and G. Zhang (2012). "Fundamental Issues in Full Waveform Inversion". In: SEG Annual International Meeting, Expanded Abstracts.

Zhou, W. et al. (2015). "Full waveform inversion of diving waves for velocity model building with impedance inversion based on scale separation". In: *Geophysical Journal International* 202.3, pp. 1535–1554.

4

HIGH-RESOLUTION ONE-WAY REFLECTION WAVEFORM INVERSION

"Correlation does not imply causation."

(STATISTICAL PRINCIPLE)

Reflection Waveform Inversion (RWI) is a technique that uses pure reflection data to estimate subsurface background velocity, relying on evolving seismic images. Conventional RWI operates in a cyclic workflow, with two key components in each cycle—migration and reflection tomography. Conventional RWI may result in suboptimal background velocity estimation, partly due to limited or unresolved resolution within each component in each cycle. While gradient preconditioning with the reciprocal of Hessian information helps resolve this issue in both components of RWI, it becomes impractical for a large number of model parameters. One-way reflection waveform inversion (ORWI) is a reflection waveform inversion technique in which the forward modeling scheme operates in one direction (downward and then upward) via virtual parallel depth levels within the medium. Leveraging the ORWI framework, we decompose and reduce the linear Hessian operator (also known as the approximate Hessian or Gauss-Newton Hessian) into multiple smaller sub-operators. In particular, the diagonal blocks of the mono-frequency approximate Hessian operators, each corresponding to a single depth level within the medium, are extracted and inverted to precondition the corresponding mono-frequency gradients in both the migration and reflection tomography components of ORWI. This depth-dependent gradient preconditioning transforms standard ORWI into a high-resolution, yet computationally feasible version aimed at addressing suboptimal velocity estimation, referred to as high-resolution ORWI (HR-ORWI). The effectiveness of the proposed approach is demonstrated through successful applications to synthetic data examples.

This chapter is based on the following paper: **Abolhassani, S.**, Hoogerbrugge, L., & Verschuur, D. J. (2025) One-Way Reflection Waveform Inversion with Depth-Dependent Gradient Preconditioning. *Geophysical Journal International*, 240(1), 652–672.

4.1. INTRODUCTION

Tarantola (1984) introduced the conventional full waveform inversion (FWI) in acoustic approximation, a well-known non-linear data-fitting technique to estimate high-resolution subsurface velocity models. Conventional FWI has become widely adopted for evaluating shallow subsurface targets. Indeed, incorporating refracted and diving waves, it offers a detailed subsurface model and has proven to be highly effective for this purpose (e.g., Vigh et al., 2011; Zou et al., 2014). However, when attempting to map the targets that lie beyond the reach of refracted and diving waves ("deep targets"), its performance may falter (e.g., Plessix et al., 2013; Brittan and Jones, 2019). This limitation arises because, while the sampling of high-reflective model wavenumbers is nonlinearly dependent on the sampling of low-propagative model wavenumbers in FWI (Audebert and Ortiz-Rubio, 2018), conventional FWI fails to actively sample low-propagative model wavenumbers within the deep targets because of weak transmission-after-reflection wavepaths within the FWI gradient, commonly known as "rabbit-ear" wavepaths. Scaled by reflection coefficients, the rabbit-ear wavepaths exhibit weaker amplitudes compared to other wavepaths within the FWI gradient, i.e., the transmission wavepath and migration isochrone (Fig. 4.1). This failure to sample low-propagative model wavenumbers significantly limits the capture of high-reflective model wavenumbers by conventional FWI, highlighting the significance of model scale separation for reflection tomography (Mora, 1989).

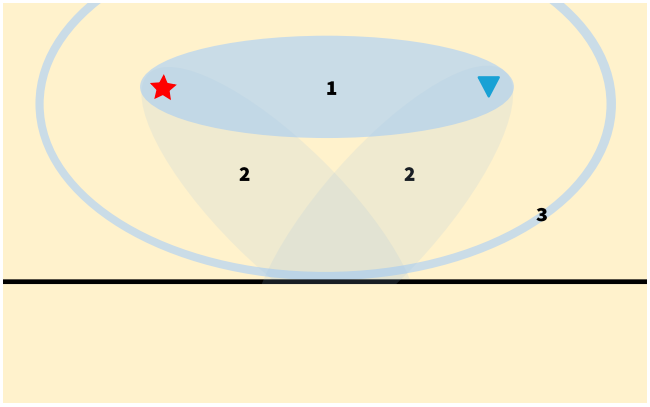


Figure 4.1: The wavepaths constructing the FWI gradient: the transmission wavepath, labeled as number 1, the pair of transmission-after-reflection wavepaths (rabbit-ear wavepaths), labeled as number 2, and the migration isochrone, labeled as number 3. As can be observed, the pair of transmission-after-reflection wavepaths appears in a lighter shade of blue, illustrating its much weaker amplitude compared to the other existing wavepaths. This occurs as a consequence of scaling by the reflection coefficient of the reflector within the medium. Hence, it can be concluded that the FWI gradient is dominated by the transmission wavepath and migration isochrone, while the pair of transmission-after-reflection wavepaths remains relatively inactive.

Over the years, several variants of migration-based velocity analysis (MVA), each built upon the principle of model scale separation based on Born modeling, have been developed. Among them, the ones incorporating wave-equation forward modeling have stood out due to the improved handling of wave propagation in complex media. Wave-equation MVA (WEMVA) algorithms (Sava and Biondi, 2004) typically optimize an image-domain

error function measured on the common image gathers (CIGs) either via a classical semblance or a differential semblance (DS) function, measuring the flatness (coherency property) in CIGs. While with the classical semblance function, the flatness measure of CIGs is given by the stack power of image amplitudes over surface offsets (Chauris and Noble, 2001), the flatness measure of CIGs with the DS function is given by squaring the L2 norm of the derivative of image amplitudes with respect to surface offsets (Shen et al., 2003). WEMVA may also be formulated in the subsurface offset domain, where, at the correct velocity, the image is expected to be focused at zero offset. In this domain, measuring the focusing of subsurface-offset CIGs mirrors the use of the classical semblance and DS function with surface CIGs (Shen, 2005; Shen and Symes, 2008; Lameloise et al., 2015). Alternatively, a scattering angle can serve as the extension parameter, with flatness measured on angle-domain CIGs (Biondi and Symes, 2004).

Expanding on the foundations of extended-domain WEMVA, Symes (2008) formulated a theory that laid the groundwork for integrating extended-domain WEMVA and FWI, leveraging their synergy. This theory was subsequently validated by Almomin and Biondi (2012). In recent initiatives to enhance MVA, Chauris and Cocher (2017) introduced the notion of inversion velocity analysis (IVA) by substituting standard migration with iterative migration to map cleaner CIGs, reducing migration artifacts and multiple crosstalk. Moreover, they adopted a pseudo-inverse of the Born modeling operator, rather than an adjoint, to reduce the migration iterations. In the same vein, Assis et al. (2022) introduced tomographic Hessian-based IVA.

Employing migrated images for velocity model building also triggered another trend of development known as reflection waveform inversion (RWI). RWI represents a wave-equation reflection tomography tool with a data-domain error function (Xu et al., 2012; Zhou et al., 2012; Berkhout, 2012; Wang et al., 2013). Conventional RWI in acoustic approximation alternately solves a two-parameter minimization problem for the background velocity and “reflectivity” of the subsurface. Although RWI and the migration-based traveltime waveform inversion (MBTT) (Clément et al., 2001) share some similarities, it is important to note that the migration step in MBTT operates on reflectivities in time rather than depth—images in time are largely invariant to velocity, unlike images in depth, but this assumes a 1D velocity model.

Unlike FWI, RWI is specifically tailored to sample deep targets tomographically harnessing exclusively reflection data through the pair of transmission-after-reflection wavepaths. RWI involves estimating the background velocity of the subsurface in a cyclic process as displayed in Fig. 4.2, relying on an evolving stacked image that fits near-offset observed and modeled “phase information" within each cycle. Therefore, it is expected that RWI naturally suppresses any ambiguities initiated by the phase mismatch in the near-offset reflection data Provenzano et al. (2023), particularly within the deep areas or the areas with challenging data acquisition conditions where FWI suffers severely. To improve the RWI tomogram, specifically within reach of refractions and diving waves, some studies have also investigated the impact of integrating early-arrival FWI (EWI) and RWI in order to steer the search direction toward recovering a full (horizontal and vertical) low local-model-wavenumber spectrum for the background velocity model (Zhou et al., 2015; Wu and Alkhalifah, 2015). For a full review of the other aspects of RWI in recent years, the reader can refer to Yao et al. (2020).

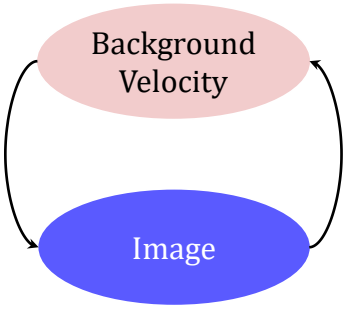


Figure 4.2: Conventional RWI cycle in which background velocity estimation and image reconstruction alternate.

Although RWI is built upon the notion of model scale separation, making this separation has been a struggle. To address the issue, several approaches have been put forth. Among others, we reference Alkhalifah (2014) for introducing the scattering-angle filtering solution, Wang et al. (2013) for adopting the wavefield decomposition solution, and Xu et al. (2012) for leveraging Born modeling as the solution. Additionally, Berkhout (2012) and Berkhout (2014) introduced the joint migration inversion (JMI) technique that makes use of a one-way forward modeling scheme called full-wavefield modeling (FWMod), parameterized based on reflectivity and acoustic background velocity, to fulfill the model scale separation by wavefield decomposition into upgoing and downgoing wavefields. FWMod models full (both primaries and multiples) downgoing and upgoing reflection wavefields within the JMI framework, in which zero-lag cross-correlations of same-direction source and receiver wavefields (either both upwards or both downwards) reconstruct the reflectivity model update (small-scale medium variations), and correlations of opposite-traveling source and receiver wavefields (one upgoing, one downgoing) estimate the background velocity model update (large-scale medium variations).

Adopting the concept of JMI for reflection waveform inversion and constraining FWMod to primary wavefield modeling (PWMod), assuming internal multiples are weak or resolved via pre-processing, this paper represents and investigates upon a one-way reflection waveform inversion technique, hereafter referred to as standard ORWI. The approporiate parameterization in PWMod leads to a natural scale separation—separation of migration isochrones and transmission-after-reflection wavepaths—in standard ORWI. This separation allows for the independent calculation of migration and tomography gradients within each ORWI cycle, thereby freeing tomograms from high-reflective model wavenumbers. Algorithm 4.1 shows the cyclic workflow of standard ORWI, which involves an alternating sequence between the image and background velocity reconstructions. Standard ORWI cyclically minimizes the error function

$$E = \frac{1}{2}\int d\omega \sum_{s} \Delta\mathbf{d}^{-}_{s,\omega}{}^{\dagger}\Delta\mathbf{d}^{-}_{s,\omega}, \tag{4.1}$$

where $\Delta\mathbf{d}^{-}$ is the residual data vector (upgoing) between the observed and modeled reflection data recorded at the Earth's surface, the subscript s indicates a source location, ω shows an angular frequency component, and the superscript $\dagger$ represents the complex

conjugate transpose operation.

Suboptimal background velocity estimation in RWI can be attributed, among others, to the use of seismic images with limited or unresolved resolution as well as suboptimal preservation of true amplitudes in each cycle (e.g., Hou and Symes, 2015; Chauris and Cocher, 2017; Gomes and Yang, 2017). With the same line of reasoning, one can also link suboptimal background velocity estimation in RWI to the use of tomograms with limited or unresolved resolution. Pratt et al. (1998) showed that the Newton optimization method in waveform inversion estimates accurate velocity models with higher resolution than the

Algorithm 4.1: Cyclic workflow of ORWI

Provide a smooth initial background velocity model and set the initial reflectivity model to zero when cycle number $k = 1$.
while convergence not reached **do**
  for $i \leftarrow 1$ to m **do**
    UPDATE the reflectivity model using one-way wave-equation while keeping the background velocity model fixed (migration).
  for $j \leftarrow 1$ to n **do**
    UPDATE the background velocity model using one-way wave-equation while keeping the reflectivity model fixed (reflection tomography).
  $k = k + 1$ ▷ The reflectivity model can be reset to zero at this point.

gradient decent method since it uses the reciprocal of second-order derivatives of the error function (inverse Hessian) to precondition the gradient. However, the calculation demand of the Hessian matrix-operator for large-scale problems in seismology renders the Newton method impossibly expensive in computing terms. Several attempts, either in the data or image domain, have been made thus far to approximate an effective yet cost-effective preconditioner: migration deconvolution filters (Hu et al., 2001; Yu et al., 2006), matching filters (Guitton, 2004; Aoki and Schuster, 2009; Guitton, 2017; Guo and Wang, 2020; Yang et al., 2021), point spread functions (Lecomte, 2008; Fletcher et al., 2016; Yang et al., 2022), target-oriented solutions (Valenciano et al., 2006; Tang, 2009), asymptotic pseudo inverses to the Born modeling operator (ten Kroode, 2014; Hou and Symes, 2015; Chauris and Cocher, 2017), linear Hessian (also known as the Gauss-Newton Hessian or approximate Hessian) (Pratt et al., 1998), pseudo Hessian (Choi et al., 2008; Jun et al., 2015), diagonal approximate Hessian (Chavent and Plessix, 1999; Shin et al., 2001; Plessix and Mulder, 2004), and approximating the application of the Hessian inverse operator on the search direction (Brossier et al., 2009; Asnaashari et al., 2013; Métivier et al., 2013; Assis and Schleicher, 2021). Within the context of iterative least-squares one-way wave-equation migration (LS-WEM), while Lu et al. (2018) obtained an approximate Hessian by implicitly formulating the modeling operator and its adjoint, Abolhassani and Verschuur (2022) introduced a computationally affordable, depth-dependent approximate Hessian, referred to henceforth as preconditioned LS-WEM (PLS-WEM).

This paper aims to increase the resolution power of standard ORWI tomograms, expanding upon the findings of Abolhassani and Verschuur (2022) on depth-dependent

preconditioning for image reconstruction. While employing depth-dependent preconditioners for high-resolution image reconstruction with standard ORWI, we concurrently develop the required mathematical groundwork to integrate the depth-dependent preconditioning concept for background velocity estimation. This two-fold effort facilitates more accurate velocity estimation with ORWI. We refer to it as high-resolution ORWI (HR-ORWI).

In the following sections, we will first review the theory of HR-ORWI. Then, we will present a brief overview of the depth-dependent gradient preconditioning concept for migration (PLS-WEM). Following that, we will mathematically demonstrate its relevance for tomography within the context of ORWI. Next, we will provide three numerical examples, where the first two compare standard ORWI and HR-ORWI, and the third validates that our cost-friendly high-resolution updates in both migration and tomography loops of HR-ORWI result in a reliable background velocity estimation for the Marmousi2 model. Lastly, we end with discussions and conclusions.

4.2. HR-ORWI THEORY

The theoretical aspects of high-resolution one-way reflection waveform inversion (HR-ORWI) are presented here. We rely on Berkhout (1981) and Berkhout (1982) for the forward problem theory.

4.2.1. FORWARD PROBLEM

The two-way acoustic wave equation for a homogeneous medium in the space-frequency domain can be written as

$$\frac{\partial^2 u}{\partial z^2} = -\frac{\omega^2}{c^2} u - \frac{\partial^2 u}{\partial x^2} - \frac{\partial^2 u}{\partial y^2}, \tag{4.2}$$

in which u represents the monochromatic pressure wavefield, and c is the velocity.

Replacing the differentiations in eq. 4.2 with spatial convolutions (Berkhout, 1982, p. 347) gives

$$\frac{\partial^2 u}{\partial z^2} = -\left(\frac{\omega^2}{c^2}\delta(x)\delta(y) + d_2(x) + d_2(y)\right) * u, \tag{4.3}$$

where $*$ denotes the spatial convolution, δ means the Dirac delta function, and the operators $d_2(x)$ and $d_2(y)$ represent spatial differential operators of order 2 with respect to x and y, respectively. When the initial condition $u(z_0)$ is available, eq. 4.3 reads

$$\left.\frac{\partial^2 u(x,y,z)}{\partial z^2}\right|_{z=z_0} = h_2(x,y) * u(x,y,z_0). \tag{4.4}$$

After splitting the total pressure wavefield into up- and downgoing wavefields ($u = u^+ + u^-$), eq. 4.4, following its transformation into the wavenumber-frequency domain within the $k_x - k_y$ plane, can be broken down into

$$\begin{aligned} \left.\frac{\partial \tilde{u}^+}{\partial z}\right|_{z=z_0} &= -i\tilde{h}_1^+ \tilde{u}^+(z_0), \\ \left.\frac{\partial \tilde{u}^-}{\partial z}\right|_{z=z_0} &= +i\tilde{h}_1^- \tilde{u}^-(z_0), \end{aligned} \tag{4.5}$$

where z is pointing downward, $\tilde{u}^+$ represents the monochromatic downgoing wavefield in the wavenumber domain, and $\tilde{u}^-$ is the monochromatic upgoing wavefield in the wavenumber domain. The partial derivatives $\frac{\partial \tilde{u}^+}{\partial z}$ and $\frac{\partial \tilde{u}^-}{\partial z}$ evaluated at z_0, are denoted as $\left.\frac{\partial \tilde{u}^+}{\partial z}\right|_{z_0}$ and $\left.\frac{\partial \tilde{u}^-}{\partial z}\right|_{z_0}$, respectively. Here, $i=\sqrt{-1}$, and $\tilde{h}_1^+ = \tilde{h}_1^- = \sqrt{\tilde{h}_2} = \sqrt{k^2 - k_x^2 - k_y^2}$, with $k = \frac{\omega}{c}$.

Inserting the derived partial derivative wavefields into the Taylor expansion for $\tilde{u}^{\pm}(z_0 \pm |\Delta z|)$ when $k^2 \geq k_x^2 + k_y^2$ yields

$$\begin{aligned} \tilde{u}^+(z_0 + |\Delta z|) &= \overbrace{e^{-i\sqrt{k^2-k_x^2-k_y^2}|\Delta z|}}^{\tilde{w}^+_{z_0+|\Delta z|;z_0}} \tilde{u}^+(z_0), \\ \tilde{u}^-(z_0 - |\Delta z|) &= \overbrace{e^{-i\sqrt{k^2-k_x^2-k_y^2}|\Delta z|}}^{\tilde{w}^-_{z_0-|\Delta z|;z_0}} \tilde{u}^-(z_0), \end{aligned} \tag{4.6}$$

where Δz denotes the extrapolation distance, $\tilde{w}^+_{z_0+|\Delta z|;z_0}$ is the downward propagator from z_0 to $z_0 + \Delta z$, and $\tilde{w}^-_{z_0-|\Delta z|;z_0}$ is the upward propagator from z_0 to $z_0 - \Delta z$. Reverting to the space-frequency domain, eq. 4.6 can be expressed as follows

$$u^{\pm}(x, y, z_0 \pm |\Delta z|) = w^{\pm}_{z_0 \pm |\Delta z|;z_0}(x, y, \omega) * u^{\pm}(x, y, z_0). \tag{4.7}$$

When dealing with extrapolation distances with lateral velocity changes, a matrix formulation is adopted to represent eq. 4.7 as

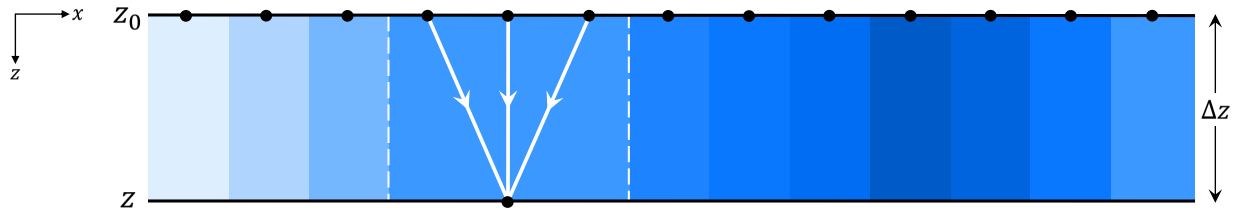
$$\mathbf{u}^+_{\omega}(z_0 + |\Delta z|) = \mathbf{W}^+_{z_0+|\Delta z|;z_0}\mathbf{u}^+_{\omega}(z_0), \tag{4.8}$$

$$\mathbf{u}^-_{\omega}(z_0 - |\Delta z|) = \mathbf{W}^-_{z_0-|\Delta z|;z_0}\mathbf{u}^-_{\omega}(z_0), \tag{4.9}$$

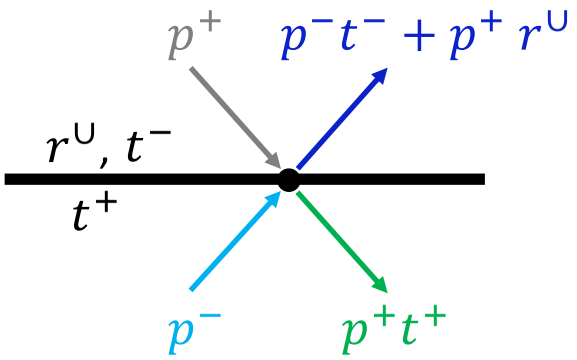
where $\mathbf{W}$ (extrapolation/propagation matrix) is a space-variant convolutional matrix rather than a Toeplitz matrix. Fig. 4.3 schematically illustrates one row of $\mathbf{W}^+_{z_0+|\Delta z|;z_0}$ propagating the wavefields downwards from z_0 to a lateral position at z using a local average velocity across an operator size of three—selected for illustrative purposes here—velocity cells. Note that the velocity cells are defined between the virtual depth levels. For more details on the space-variant extrapolation matrix and operator size, see Berkhout (1982) and Thorbecke et al. (2004). One should note that the space-variant extrapolation matrix used for this research remains an approximation valid under the condition of smooth lateral velocity changes. Interested readers are referred to Sun and Verschuur (2020) and Li et al. (2024) for insights on an accurate extrapolation matrix for strong lateral velocity changes.

By discretizing the subsurface into N virtual depth levels separated by Δz (Δz is supposed to be sufficiently small to be considered vertically homogeneous), the fundamental equation governing the modeling of monochromatic primary reflection data recorded at the Earth's surface (z_0) with one monochromatic physical source positioned at z_0 is given by

$$\mathbf{u}^-(z_0) = \sum_{m=1}^{N} \underline{\mathbf{W}}^-_{z_0;z_m}\mathbf{r}^{\cup}(z_m) \circ \underline{\mathbf{W}}^+_{z_m;z_0}\mathbf{s}^+(z_0), \tag{4.10}$$



(a)



(b)

Figure 4.3: (a) Schematic representation of downward wavefield extrapolation between two virtual depth levels, considering lateral velocity variation. This graphic represents one single row of $\mathbf{W}^+_{z_0+|\Delta z|;z_0}$, which propagates the wavefields to a lateral position at z using a local average velocity across an operator size of three velocity cells. Note that the background velocity cells are defined between the virtual depth levels. (b) Schematic representation of how the incident wavefield interacts (reflection and transmission) at a particular virtual depth level in PWMod, where $r^\cup$ means an upward reflectivity scalar, t^+ denotes a downward transmission scalar, and t^+ denotes an upward transmission scalar of a given grid point at a specific lateral location.

with

$$\underline{\mathbf{W}}^{+}_{z_m;z_0} = \mathbf{W}^{+}_{z_m;z_{m-1}}\mathbf{T}^{+}(z_{m-1})\cdots\mathbf{T}^{+}(z_1)\mathbf{W}^{+}_{z_1;z_0}, \tag{4.11}$$

$$\underline{\mathbf{W}}^{-}_{z_0,z_m} = \mathbf{W}^{-}_{z_0,z_1}\mathbf{T}^{-}(z_1)\cdots\mathbf{T}^{-}(z_{m-1})\mathbf{W}^{-}_{z_{m-1},z_m}, \tag{4.12}$$

where $\mathbf{u}^{-}(z_0)$ is the modeled monochromatic reflection data recorded at the Earth's surface (z_0), $\mathbf{s}^{+}(z_0)$ is the physical monochromatic downgoing source at z_0, $\mathbf{r}^{\cup}$ is the angle-independent upward reflectivity vector-operator at a layer boundary, $\mathbf{T}^{\pm}$ is the angle-independent downward/upward transmission matrix-operator at a layer boundary and explicitly defined by $\mathbf{T}^{\pm} = \mathbf{I} \pm \mathrm{diag}(\mathbf{r}^{\cup})$, and $\circ$ means the Hadamard product (Fig. 4.3b). Note that despite expressing the forward modeling equations in three dimensions (3D), for the remainder of this paper, we shall consider only the 2D case.

4.2.2. OPTIMIZATION PROBLEM

HR-ORWI minimizes a two-parameter error function by alternating updates of the two parameters: angle-independent upward reflectivity and background velocity (see Algorithm 4.1). The error function through a linearized forward problem approximation is defined as a quadratic function of frequency-dependent model perturbation vectors

$$E = \frac{1}{2}\int d\omega S(\delta\mathbf{m}_\omega), \tag{4.13}$$

with

$$S(\delta\mathbf{m}_\omega) = \sum_{s=1}^{N_s}\left(\Delta\mathbf{d}^{-}_{s,\omega} - \frac{\partial\mathbf{u}^{-}_{s,\omega}(z_0)}{\partial\mathbf{m}}\delta\mathbf{m}_\omega\right)^{\dagger}\left(\Delta\mathbf{d}^{-}_{s,\omega} - \frac{\partial\mathbf{u}^{-}_{s,\omega}(z_0)}{\partial\mathbf{m}}\delta\mathbf{m}_\omega\right), \tag{4.14}$$

and

$$\delta\mathbf{m}_\omega = \begin{pmatrix} \delta\mathbf{r}^{\cup}_{\omega} \\ \delta\mathbf{c}_\omega \end{pmatrix}, \tag{4.15}$$

where $\delta\mathbf{m}_\omega$ is the mono-frequency total model perturbation vector, $\delta\mathbf{r}^{\cup}_{\omega}$ represents the upward reflectivity perturbation vector (the total upward reflectivity perturbation vector $\delta\mathbf{r}^{\cup}$ is obtained by integrating over frequency), $\delta\mathbf{c}_\omega$ denotes the background velocity perturbation vector (the total background velocity perturbation vector $\delta\mathbf{c}$ is obtained by integrating over frequency), $\Delta\mathbf{d}^{-}_{s,\omega}$ represents the monochromatic residual data vector between the observed and modeled primary reflection data at z_0 for a given source location, $\frac{\partial\mathbf{u}^{-}_{s,\omega}(z_0)}{\partial\mathbf{m}}$ represents the monochromatic sensitivity matrix—the first-order partial derivative wavefield with respect to the model parameter change—for a given source location, s counts the source locations, and N_s shows the numbers of source locations.

Setting the derivative of eq. 4.13 with respect to $\delta\mathbf{m}_\omega$ to zero (gradient descent optimization method) gives

$$\frac{\partial E}{\partial\delta\mathbf{m}_\omega} = \frac{\partial S}{\partial\delta\mathbf{m}_\omega} = 0, \tag{4.16}$$

which leads to

$$\delta\mathbf{m}_\omega = \mathrm{Re}\left\{\overbrace{\sum_{s=1}^{N_s}\left[\frac{\partial\mathbf{u}^-_{s,\omega}(z_0)}{\partial\mathbf{m}}\right]^\dagger\left[\frac{\partial\mathbf{u}^-_{s,\omega}(z_0)}{\partial\mathbf{m}}\right]}^{\mathbf{H}^{\mathrm{a}}_\omega}\right\}^{-1}\mathrm{Re}\left\{\overbrace{\sum_{s=1}^{N_s}\left[\frac{\partial\mathbf{u}^-_{s,\omega}(z_0)}{\partial\mathbf{m}}\right]^\dagger\Delta\mathbf{d}^-_{s,\omega}}^{\mathbf{g}_\omega}\right\}, \tag{4.17}$$

with

$$\mathbf{H}^{\mathrm{a}}_\omega = \begin{bmatrix} \sum_{s=1}^{N_s}\left[\frac{\partial\mathbf{u}^-_{s,\omega}(z_0)}{\partial\mathbf{r}^\cup}\right]^\dagger\left[\frac{\partial\mathbf{u}^-_{s,\omega}(z_0)}{\partial\mathbf{r}^\cup}\right] & \sum_{s=1}^{N_s}\left[\frac{\partial\mathbf{u}^-_{s,\omega}(z_0)}{\partial\mathbf{r}^\cup}\right]^\dagger\left[\frac{\partial\mathbf{u}^-_{s,\omega}(z_0)}{\partial\mathbf{c}}\right] \\ \sum_{s=1}^{N_s}\left[\frac{\partial\mathbf{u}^-_{s,\omega}(z_0)}{\partial\mathbf{c}}\right]^\dagger\left[\frac{\partial\mathbf{u}^-_{s,\omega}(z_0)}{\partial\mathbf{r}^\cup}\right] & \sum_{s=1}^{N_s}\left[\frac{\partial\mathbf{u}^-_{s,\omega}(z_0)}{\partial\mathbf{c}}\right]^\dagger\left[\frac{\partial\mathbf{u}^-_{s,\omega}(z_0)}{\partial\mathbf{c}}\right] \end{bmatrix}, \tag{4.18}$$

and

$$\mathbf{g}_\omega = \begin{pmatrix} \mathbf{g}^r_\omega \\ \mathbf{g}^c_\omega \end{pmatrix}, \tag{4.19}$$

where $\mathbf{H}^{\mathrm{a}}_\omega$ is the mono-frequency approximate Hessian matrix for all shots (equivalent to the mono-frequency Gauss-Newton Hessian matrix), $\mathbf{g}_\omega$ shows the mono-frequency gradient vector, $\mathbf{g}^r_\omega$ represents the mono-frequency gradient vector for reflectivity, and $\mathbf{g}^c_\omega$ denotes the mono-frequency gradient vector for background velocity—note that with Gauss-Newton optimization in waveform inversion, the local convexity in inverse modeling, despite a nonlinear forward model, resembles dealing with a linear forward model in each iteration.

Following Jang et al. (2009) and Oh and Min (2013), the total model perturbation vector, after accounting for the negative sign indicating the descent direction, can be expressed as the linear combination of all the mono-frequency model perturbation vectors

$$\delta\mathbf{m} = -\int d\omega\delta\mathbf{m}_\omega, \tag{4.20}$$

where $\delta\mathbf{m}$ represents the total model perturbation vector.

In the n^{th} cycle, the total model perturbation vector satisfies

$$\mathbf{m}^{(n+1)} = \mathbf{m}^{(n)} + \alpha^{(n)}\delta\mathbf{m}^{(n)}, \tag{4.21}$$

where α denotes the optimization step length and reads

$$\alpha^{(n)} = \frac{\Delta\mathbf{d}^{-\dagger}\left(\mathbf{m}^{(n)}+\delta\mathbf{m}^{(n)}\right)\Delta\mathbf{d}^-\left(\mathbf{m}^{(n)}\right)}{\Delta\mathbf{d}^{-\dagger}\left(\mathbf{m}^{(n)}+\delta\mathbf{m}^{(n)}\right)\Delta\mathbf{d}^-\left(\mathbf{m}^{(n)}+\delta\mathbf{m}^{(n)}\right)}, \tag{4.22}$$

resulting from minimizing the residual data with respect to the optimization step length.

To reduce the computational load of solving the multi-parameter inverse problem, we first discard the off-diagonal blocks of the mono-frequency approximate Hessian,

which represent the cross-talk between parameter classes. This is achieved by setting the off-diagonal elements of $\mathbf{H}^{\mathrm{a}}_{\omega}$ to zero. As a result, eq. 4.18 transforms into

$$\mathbf{H}^{\mathrm{a}}_{\omega} = \begin{bmatrix} \mathbf{H}^{\mathrm{a},r}_{\omega} & \\ & \mathbf{H}^{\mathrm{a},c}_{\omega} \end{bmatrix}, \tag{4.23}$$

where $\mathbf{H}^{\mathrm{a},r}_{\omega}$ shows the mono-frequency approximate Hessian matrix to precondition $\mathbf{g}^{r}_{\omega}$, $\mathbf{H}^{\mathrm{a},c}_{\omega}$ shows the mono-frequency approximate Hessian matrix to precondition $\mathbf{g}^{c}_{\omega}$, and zeros are omitted for brevity. With this, eq. 4.21 turns into a system of equations with two equations with reduced and simpler terms. We then, instead of solving the two equations simultaneously, opt to solve the equations "independently" (i.e., minimal interdependency assumption between reflectivity and background velocity to further ease the computational load). This process involves an alternating scheme where we first solve the first equation to update $\mathbf{r}^{\cup}$ while keeping $\mathbf{c}$ fixed (migration), and then solve the second equation to update $\mathbf{c}$ while keeping $\mathbf{r}^{\cup}$ fixed (reflection tomography). This alternating scheme continues until convergence is achieved. This enables HR-ORWI to cyclically update both $\mathbf{r}^{\cup}$ and $\mathbf{c}$, with an alternating scheme offering partial mitigation of the neglected interdependency between $\mathbf{r}^{\cup}$ and $\mathbf{c}$ (see Algorithm 4.1).

In the following section, we present how to calculate $\mathbf{H}^{\mathrm{a},r}_{\omega}$ and $\mathbf{H}^{\mathrm{a},c}_{\omega}$ in a depth-marching regime. The gradient calculations are explained in detail in Sun et al. (2019).

4.2.3. DEPTH-DEPENDENT GRADIENT PRECONDITIONING

REFLECTIVITY

Since PWMod allows access to the acoustic wavefield at different virtual depth levels, $\mathbf{g}_{\omega}$ for the reflectivity class of parameters can be expressed as

$$\mathbf{g}^{r}_{\omega} = \begin{pmatrix} \mathbf{g}^{r}_{\omega}(z_0) \\ \vdots \\ \mathbf{g}^{r}_{\omega}(z_N) \end{pmatrix} = \begin{pmatrix} \sum_{s=1}^{N_s} \left[\frac{\partial \mathbf{u}^{-}_{s,\omega}(z_0)}{\partial \mathbf{r}^{\cup}(z_0)} \right]^{\dagger} \Delta \mathbf{d}^{-}_{s,\omega} \\ \vdots \\ \sum_{s=1}^{N_s} \left[\frac{\partial \mathbf{u}^{-}_{s,\omega}(z_0)}{\partial \mathbf{r}^{\cup}(z_N)} \right]^{\dagger} \Delta \mathbf{d}^{-}_{s,\omega} \end{pmatrix}, \tag{4.24}$$

and

$$\mathbf{H}^{\mathrm{a},r}_{\omega} = \begin{bmatrix} \mathbf{H}^{\mathrm{a},r}_{\omega}(z_0) & & \\ & \ddots & \\ & & \mathbf{H}^{\mathrm{a},r}_{\omega}(z_N) \end{bmatrix}, \tag{4.25}$$

where $\mathbf{H}^{\mathrm{a},r}_{\omega}$ shows a block-diagonal structure, and off-block-diagonal zeros are here omitted for brevity.

Within the LS-WEM context, Abolhassani and Verschuur (2022a) showed that the sensitivity matrix with respect to the upward reflectivity model at a given virtual depth

level (z_ℓ) for a pair of source and frequency component reads

$$\frac{\partial \mathbf{u}^-_{s,\omega}(z_0)}{\partial \mathbf{r}^\cup(z_\ell)} = \left[\underline{\mathbf{W}}^-_{z_0;z_\ell} \begin{pmatrix} u^+_{1\,s,\omega}(z_\ell) \\ 0 \\ \vdots \\ 0 \end{pmatrix} \underline{\mathbf{W}}^-_{z_0;z_\ell} \begin{pmatrix} 0 \\ u^+_{2\,s,\omega}(z_\ell) \\ \vdots \\ 0 \end{pmatrix} \right. \\ \left. \cdots \quad \underline{\mathbf{W}}^-_{z_0;z_\ell} \begin{pmatrix} 0 \\ 0 \\ \vdots \\ u^+_{n_x\,s,\omega}(z_\ell) \end{pmatrix} \right], \tag{4.26}$$

in which n_x denotes the number of gridpoints at z_m, and $u^+_{j\,s,\omega}(z_\ell)$ is a complex number representing the downgoing modeled wavefield at the jth grid point of z_ℓ.

Integrating eqs 4.24, 4.25, and 4.26 into eq. 4.17 and then into eq. 4.20 yield the total angle-independent upward reflectivity perturbation vector $\delta\mathbf{r}^\cup$. Finally, eq. 4.21 is used to update the total angle-independent upward reflectivity vector $\mathbf{r}^\cup$ in each cycle.

BACKGROUND VELOCITY

By similarity, we will here derive the sensitivity matrix for the background velocity model at a given virtual depth level (z_ℓ) for a pair of source and frequency component.

We rewrite the monochromatic primary reflection data recorded at the Earth's surface (eq. 4.10) into the superposition of two upgoing wavefields, one traveling from $[z_{\ell+1}, z_N]$ to z_0 and the other from $[z_1, z_\ell]$ to z_0,

$$\mathbf{u}^-_{s,\omega}(z_0) = \sum_{m=\ell+1}^{N} \underline{\mathbf{W}}^-_{z_0;z_m} \mathbf{r}^\cup(z_m) \circ \underline{\mathbf{W}}^+_{z_m;z_0} \mathbf{s}^+_\omega(z_0) + \sum_{m=1}^{\ell} \underline{\mathbf{W}}^-_{z_0;z_m} \mathbf{r}^\cup(z_m) \circ \underline{\mathbf{W}}^+_{z_m;z_0} \mathbf{s}^+_\omega(z_0). \tag{4.27}$$

The data sensitivity with respect to a single background velocity model parameter located at z_ℓ (i.e., a single column of the respective sensitivity matrix) is found by taking the derivative of eq. 4.27 with respect to $c_j(z_\ell)$

$$\frac{\partial \mathbf{u}^-_{s,\omega}(z_0)}{\partial c_j(z_\ell)} = \overbrace{\left\{ \sum_{m=\ell+1}^{N} \frac{\partial \underline{\mathbf{W}}^-_{z_0;z_m}}{\partial c_j(z_\ell)} \mathbf{r}^\cup(z_m) \circ \underline{\mathbf{W}}^+_{z_m;z_0} \mathbf{s}^+_\omega(z_0) \right\}}^{\mathbf{d}_a} + \overbrace{\left\{ \sum_{m=\ell+1}^{N} \underline{\mathbf{W}}^-_{z_0;z_m} \mathbf{r}^\cup(z_m) \circ \frac{\partial \underline{\mathbf{W}}^+_{z_m;z_0}}{\partial c_j(z_\ell)} \mathbf{s}^+_\omega(z_0) \right\}}^{\mathbf{d}_b}. \tag{4.28}$$

After reintroducing $\underline{\mathbf{W}}^-_{z_0;z_m}$ and $\underline{\mathbf{W}}^+_{z_m;z_0}$ as

$$\underline{\mathbf{W}}^-_{z_0;z_m} = \underline{\mathbf{W}}^-_{z_0;z_\ell} \mathbf{T}^-(z_\ell) \mathbf{W}^-_{z_\ell;z_{\ell+1}} \mathbf{T}^-(z_{\ell+1}) \underline{\mathbf{W}}^-_{z_{\ell+1};z_m}, \tag{4.29}$$

$$\underline{\mathbf{W}}^+_{z_m;z_0} = \underline{\mathbf{W}}^+_{z_m;z_{\ell+1}} \mathbf{T}^+(z_{\ell+1}) \mathbf{W}^+_{z_{\ell+1};z_\ell} \mathbf{T}^+(z_\ell) \underline{\mathbf{W}}^+_{z_\ell;z_0}, \tag{4.30}$$

and plugging them into eq. 4.28, we get

$$\begin{aligned}\mathbf{d}_a = &\underline{\mathbf{W}}^{-}_{z_0;z_\ell}\mathbf{T}^{-}(z_\ell)\frac{\partial \mathbf{W}^{-}_{z_\ell;z_{\ell+1}}}{\partial c_j(z_\ell)}\mathbf{T}^{-}(z_{\ell+1})\sum_{m=\ell+2}^{N}\underline{\mathbf{W}}^{-}_{z_{\ell+1};z_m}\mathbf{r}^{\cup}(z_m)\circ\underline{\mathbf{W}}^{+}_{z_m;z_0}\mathbf{s}^{+}_{\omega}(z_0)\\ &+\underline{\mathbf{W}}^{-}_{z_0;z_\ell}\mathbf{T}^{-}(z_\ell)\frac{\partial \mathbf{W}^{-}_{z_\ell;z_{\ell+1}}}{\partial c_j(z_\ell)}\mathbf{r}^{\cup}(z_{\ell+1})\circ\underline{\mathbf{W}}^{+}_{z_{\ell+1};z_0}\mathbf{s}^{+}_{\omega}(z_0),\end{aligned} \tag{4.31}$$

$$\begin{aligned}\mathbf{d}_b = &\sum_{m=\ell+2}^{N}\underline{\mathbf{W}}^{-}_{z_0;z_m}\mathbf{r}^{\cup}(z_m)\circ\mathbf{W}^{+}_{z_m;z_{\ell+1}}\mathbf{T}^{+}(z_{\ell+1})\frac{\partial \mathbf{W}^{+}_{z_{\ell+1};z_\ell}}{\partial c_j(z_\ell)}\mathbf{T}^{+}(z_\ell)\underline{\mathbf{W}}^{+}_{z_\ell;z_0}\mathbf{s}^{+}_{\omega}(z_0)\\ &+\underline{\mathbf{W}}^{-}_{z_0;z_{\ell+1}}\mathbf{r}^{\cup}(z_{\ell+1})\circ\frac{\partial \mathbf{W}^{+}_{z_{\ell+1};z_\ell}}{\partial c_j(z_\ell)}\mathbf{T}^{+}(z_\ell)\underline{\mathbf{W}}^{+}_{z_\ell;z_0}\mathbf{s}^{+}_{\omega}(z_0),\end{aligned} \tag{4.32}$$

in which given the assumption that $\mathbf{W}^{-}_{z_\ell;z_{\ell+1}}$ and $\mathbf{W}^{+}_{z_{\ell+1};z_\ell}$ are approximate to each other, the first partial derivatives of the extrapolation operator are defined as

$$\frac{\partial \mathbf{W}^{-}_{z_\ell;z_{\ell+1}}}{\partial c_j(z_\ell)} = \frac{\partial \mathbf{W}^{+}_{z_{\ell+1};z_\ell}}{\partial c_j(z_\ell)} = \begin{bmatrix}\mathbf{0}\\ \vdots\\ \mathbf{0}\\ \frac{\partial \mathbf{w}_{j,\star}}{\partial c_j}\\ \mathbf{0}\\ \vdots\\ \mathbf{0}\end{bmatrix}_{n_x\times n_x}, \tag{4.33}$$

and $\mathbf{w}_{j,\star}$ means the jth row of $\mathbf{W}^{-}_{z_\ell;z_{\ell+1}}$ or $\mathbf{W}^{+}_{z_{\ell+1};z_\ell}$, with each $\mathbf{0}$ (zero in boldface) denoting a row vector of dimension $1\times n_x$.

Fig. 4.4 provides a schematic representation of the physical interpretation of eqs 4.31 and 4.32 for a medium consisting of six virtual depth levels, including z_0. This illustration focuses on the scenario where c_j is positioned "at the virtual depth level z_2" in notation (between z_2 and z_3 in practice).

Now, after eq. 4.24, we can write,

$$\mathbf{g}^{c}_{\omega} = \begin{pmatrix}\mathbf{g}^{c}_{\omega}(z_0)\\ \vdots\\ \mathbf{g}^{c}_{\omega}(z_N)\end{pmatrix} = \begin{pmatrix}\sum_{s=1}^{N_s}\left[\frac{\partial \mathbf{u}^{-}_{s,\omega}(z_0)}{\partial \mathbf{c}(z_0)}\right]^{\dagger}\Delta\mathbf{d}^{-}_{s,\omega}\\ \vdots\\ \sum_{s=1}^{N_s}\left[\frac{\partial \mathbf{u}^{-}_{s,\omega}(z_0)}{\partial \mathbf{c}(z_N)}\right]^{\dagger}\Delta\mathbf{d}^{-}_{s,\omega}\end{pmatrix}, \tag{4.34}$$

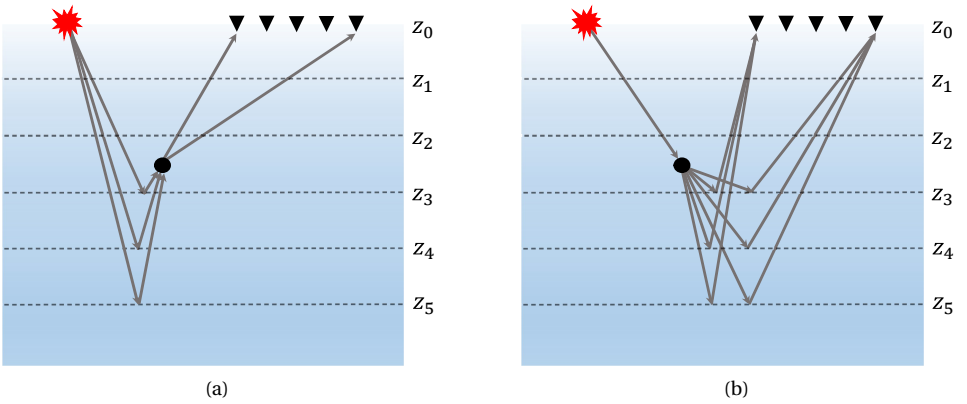


Figure 4.4: Schematic representation of $\mathbf{d}_a$ and $\mathbf{d}_b$ for one shot, eqs 4.31 and 4.32, in a medium with six virtual depth levels, including z_0. (a) Receiver-side background velocity data sensitivity for one shot (i.e., $\mathbf{d}_a$). (b) Source-side background velocity data sensitivity for one shot (i.e., $\mathbf{d}_b$).

and

$$\mathbf{H}_{\omega}^{\mathrm{a},c} = \begin{bmatrix} \mathbf{H}_{\omega}^{\mathrm{a},c}(z_0) & & \\ & \ddots & \\ & & \mathbf{H}_{\omega}^{\mathrm{a},c}(z_N) \end{bmatrix}, \tag{4.35}$$

where $\mathbf{H}_{\omega}^{\mathrm{a},c}$ again shows a block-diagonal structure, and off-block-diagonal zeros are here omitted for brevity.

Incorporating eqs 4.34, 4.35, and 4.28 into eq. 4.17 and subsequently into eq. 4.20 give the total background velocity perturbation vector $\delta\mathbf{c}$. Afterward, eq. 4.21 is employed to update the total background velocity vector $\mathbf{c}$ in each cycle.

According to equations 4.25 and 4.35, the depth-marching nature of HR-ORWI framework facilitates constructing mono-frequency, depth-dependent approximate Hessian operators. Each mono-frequency, depth-dependent approximate Hessian operator is calculated for model parameters located on a single virtual depth level (with n_x model parameters at each virtual depth level). This calculation involves both auto-correlation and cross-correlation between the partial derivative wavefields associated with a specific depth level, while disregarding any cross-correlation between the partial derivative wavefields of that depth level and those of other virtual depth levels. This choice, facilitated by the frequency-domain, depth-marching nature of one-way propagators, reduces the entire mono-frequency approximate Hessian, with a dimension of $n_m \times n_m$ ($n_m = N \times n_x$, representing the total number of model parameters within the entire medium), into multiple sub-operators, each tailored to a single virtual depth level (i.e., N operators with a reduced dimension of $n_x \times n_x$). Reducing the entire mono-frequency approximate Hessian into N smaller sub-operators lowers the memory and processing demands for constructing, storing, and inverting in large-scale problems, while still retaining essential information for preconditioning at each depth level. The inverse of $\mathbf{H}_{\omega}^{\mathrm{a,r}}(z_\ell)$ and $\mathbf{H}_{\omega}^{\mathrm{a,c}}(z_\ell)$

accounts for geometric spreading and the spatial correlations of neighboring model parameters at z_ℓ, while also performing frequency-wise source deconvolution. This approach approximates a block-diagonal representation of the mono-frequency Gauss-Newton Hessian (Pratt et al., 1998) if one assembles all mono-frequency, depth-dependent approximate Hessians into a single large matrix, representing the entire medium (Abolhassani and Verschuur, 2024).

4.2.4. HR-ORWI VERSUS STANDARD ORWI

Standard ORWI leverages the same forward problem theory as HR-ORWI and similarly updates reflectivity and background velocity by minimizing the errors in primary reflection waveforms for each class of parameters alternately, meaning that both standard ORWI and HR-ORWI follow the same inversion workflow, i.e., Algorithm 4.1.

Unlike HR-ORWI, standard ORWI minimizes eq. 4.13 as a quadratic function of $\delta\mathbf{m}$ rather than $\delta\mathbf{m}_\omega$. Setting the derivative of eq. 4.13 with respect to $\delta\mathbf{m}$ to zero (gradient descent optimization method) gives

$$\delta\mathbf{r}^{\cup} = -\,\mathrm{Re}\left\{\int d\omega\ \mathbf{H}_\omega^{\mathrm{a},r}(\mathbf{x},\mathbf{x})\right\}^{-1}\mathrm{Re}\left\{\int d\omega \sum_{s=1}^{N_s}\left[\frac{\partial\mathbf{u}^-_{s,\omega}(z_0)}{\partial\mathbf{r}^{\cup}}\right]^{\dagger}\Delta\mathbf{d}^-_{s,\omega}\right\}, \tag{4.36}$$

$$\delta\mathbf{c} = -\,\mathrm{Re}\left\{\int d\omega\ \mathbf{H}_\omega^{\mathrm{a},c}(\mathbf{x},\mathbf{x})\right\}^{-1}\mathrm{Re}\left\{\int d\omega \sum_{s=1}^{N_s}\left[\frac{\partial\mathbf{u}^-_{s,\omega}(z_0)}{\partial\mathbf{c}}\right]^{\dagger}\Delta\mathbf{d}^-_{s,\omega}\right\}, \tag{4.37}$$

where $\mathbf{H}_\omega^{\mathrm{a},r}(\mathbf{x},\mathbf{x})$ indicates that only the diagonal elements of the mono-frequency approximate Hessian possess values (see, Staal, 2015). This type of preconditioning only accounts for geometric spreading. Lastly, eq. 4.21 is employed in order to update the total angle-independent upward reflectivity vector $\mathbf{r}^{\cup}$ and the total background velocity vector $\mathbf{c}$ in each cycle.

We now use a two-layer model to evaluate the fidelity of the reflectivity and background velocity gradients in HR-ORWI versus Standard ORWI. The upper layer has a homogeneous background velocity of 3500 m/s, while the bottom layer has a homogeneous background velocity of 3000 m/s. The interface is at a depth of 800 m. 41 shots are triggered every 75 m on the surface. A 10 Hz Ricker wavelet is used as the source function with a fixed-spread acquisition geometry (the maximum available offset is 3000 m). We fix the migration offset—data offsets used for seismic migration—at 400 m (near-offset imaging) and use an initial background velocity model of 3400 m/s (i.e., ≈ 3% error). Fig. 4.5 display the stacked reflectivity perturbation ($\delta\mathbf{r}^{\cup}$) and one-source-receiver-pair background velocity perturbation, estimated after one cycle (1x LS migration and 1x LS tomography) of standard ORWI and HR-ORWI.

Comparing Figs. 4.5a and 4.5b (standard ORWI) with Figs. 4.5c and 4.5d (HR-ORWI) reveals that HR-ORWI provides significantly better resolution in both the image and tomogram than standard ORWI. In Figs. 4.5a and 4.5b (standard ORWI), the signature of the source function is visible in both the image and tomogram. The pair of transmission-after-reflection wavepaths in Fig. 4.5b (standard ORWI) are unfocused and band-limited, as manifested by the scattering of reflected energies over a larger area compared to Fig. 4.5d (HR-ORWI), which makes information extraction from the data more challenging. In contrast, Figs. 4.5c and 4.5d (HR-ORWI) show superior resolution in both the image

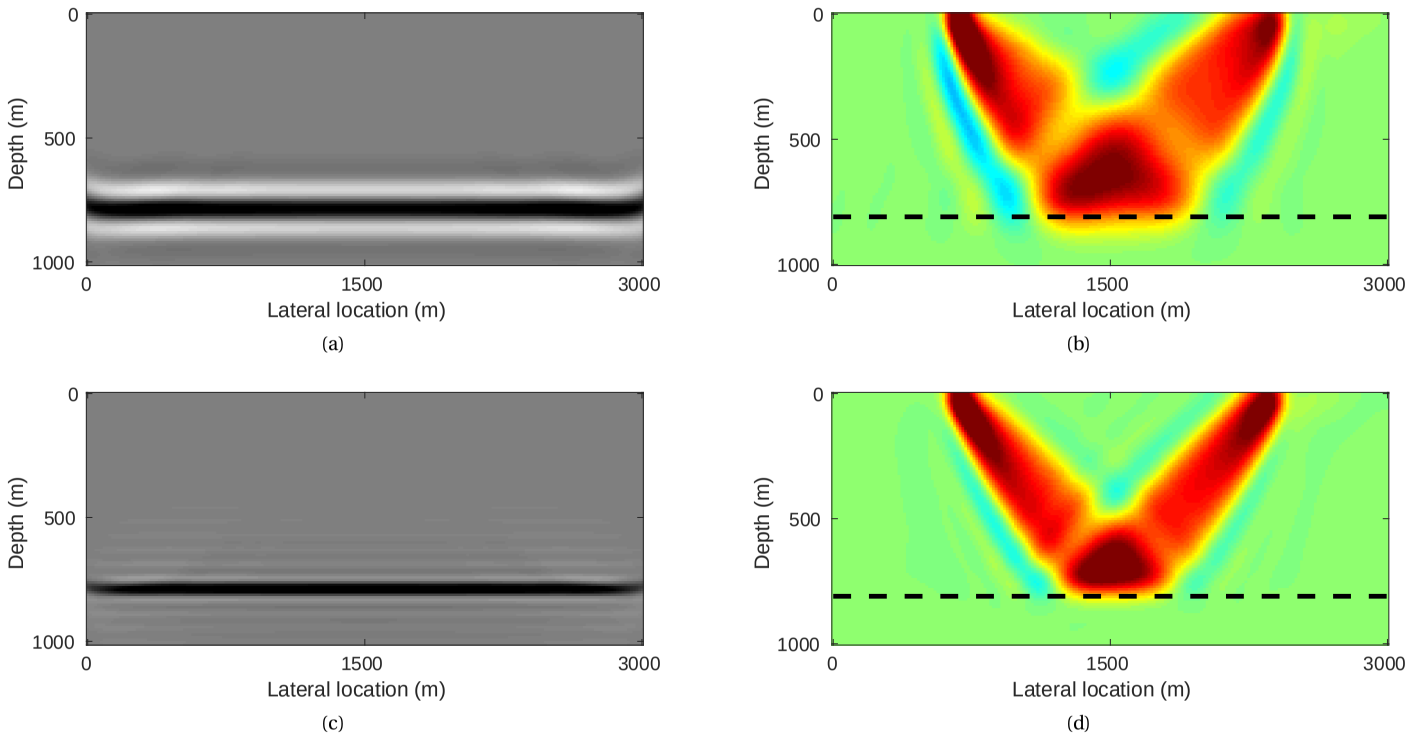
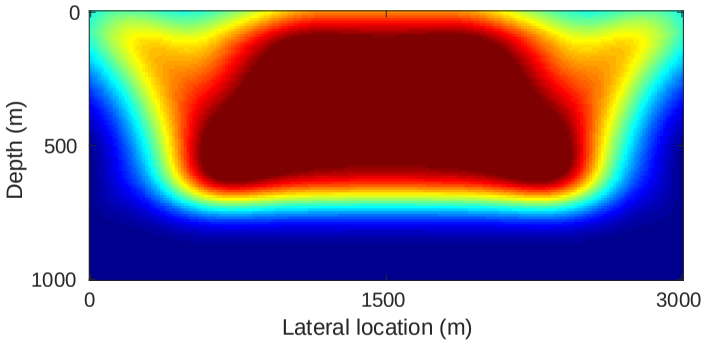
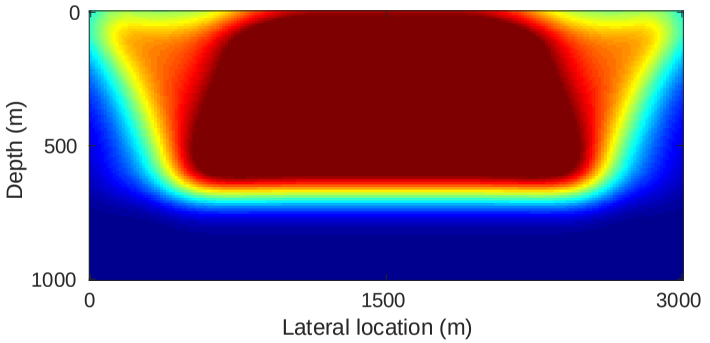


Figure 4.5: Stacked reflectivity perturbation and one-source-receiver-pair background velocity perturbation in a two-layer model after one cycle (1x LS migration and 1x LS tomography). (a, b) Standard ORWI: (a) Stacked reflectivity perturbation, (b) One-source-receiver-pair background velocity perturbation. (c, d) HR-ORWI: (c) Stacked reflectivity perturbation, (d) One-source-receiver-pair background velocity perturbation.



(a) Standard ORWI



(b) HR-ORWI

Figure 4.6: Background velocity perturbation corresponding to the two-layer model, estimated after one cycle (1x LS migration and 1x LS tomography) for (a) standard ORWI and (b) HR-ORWI.

and tomogram. This improvement is attributed to the source signature deconvolution effect embedded in the proposed depth-dependent approximate Hessian operators. With better-focused transmission-after-reflection wavepaths, HR-ORWI is better positioned to handle the complex interference of the wavepaths in challenging geological settings, supporting more reliable search directions than standard ORWI.

With acquisition configurations identical to those used in the experiment associated with Fig 4.5, we use the same two-layer model to compare the stacked background velocity perturbations estimated through HR-ORWI and standard ORWI. In the experiment, we use an initial background velocity model of 3000 m/s (i.e., ≈ 15% error). Additionally, we adopt the offset-selection strategy presented in Appendix 4.A. That is, we fix the migration offset at 800 m (maximum effective migration offset: MEMO), use the mid-to-far-offset residual data (|offsets| > 500 m) for tomography, and exclude the cycle-skipped far-offset data for tomography. Fig. 4.6 compares the respective background velocity perturbations ($\delta\mathbf{c}$) after one cycle (1x LS migration and 1x LS tomography) of HR-ORWI and standard ORWI. Upon comparison, it is observed that the superposition of tomographic wavepaths estimated through HR-ORWI exhibits a higher level of geometric consistency and homogeneity with the true layer. This observation highlights the accuracy of HR-ORWI in capturing subsurface features, resulting in a more faithful representation of the truth.

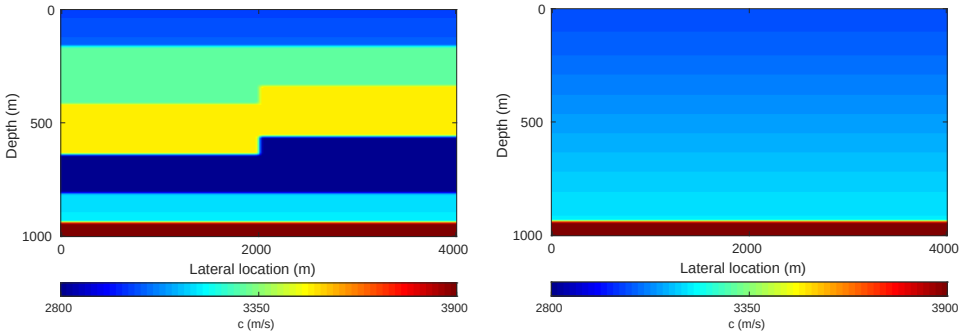
4.3. NUMERICAL EXAMPLES

First, we validate the effectiveness of HR-ORWI compared to standard ORWI using two synthetic examples. Second, we evaluate the performance of HR-ORWI with the synthetic data on the marine-environment Marmousi model (Martin et al., 2006). In all the examples, we adopt the offset-selection strategy presented in Appendix 4.A.

4.3.1. FAULT MODEL

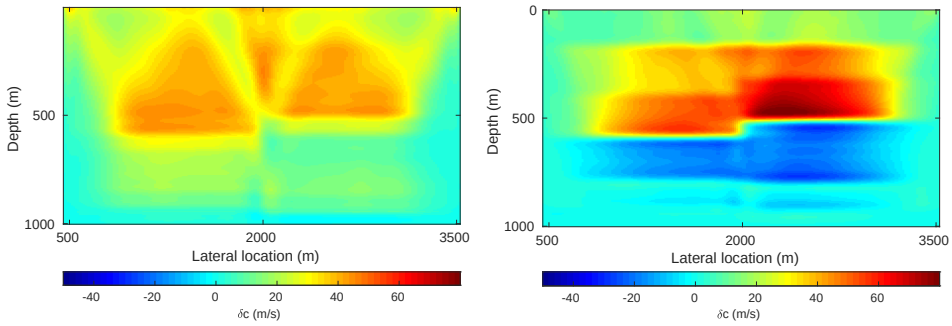
In the fault model example (Fig. 4.7a), the observed data is generated via PWMod (only primary reflections), employing a 10 Hz Ricker wavelet. The fault model is discretized with 201 gridpoints in the horizontal direction (20 m interval) and 201 gridpoints in the vertical direction (5 m interval). 41 shots with 100 m spacing are used, and each shot is recorded by 201 receivers with a fixed-spread acquisition geometry at the surface of the Earth. The maximum available offset in the data is limited to 3000 m. The data is recorded for a duration of 1.6 s. The initial background velocity model is a 1D linearly increasing gradient model (Fig. 4.7b), and the initial reflectivity model is zero. We fix the migration offset at 1000 m, use the mid-to-far-offset residual data (|offsets| > 500 m) for tomography, and exclude the cycle-skipped far-offset data for tomography in each cycle. The 0-30 Hz frequency band data is inverted at once, i.e., the inversion process does not involve a multi-scaling strategy (Bunks et al., 1995). Each cycle of ORWI includes 1x LS migration and 1x LS tomography. While Fig. 4.8 provides a comparison of the estimated background velocity perturbations (δc) via standard ORWI and HR-ORWI after one cycle, Fig. 4.9 compares the estimated tomograms using standard ORWI and HR-ORWI after 25 and 50 cycles.

Comparing Figs 4.8a and 4.8b reveals a clear optimization direction with HR-ORWI shortly after the first cycle, where the fault line and balanced-amplitude background



(a) True model (b) Initial model

Figure 4.7: Fault model example. (a) True background velocity model. (b) Initial background velocity model.

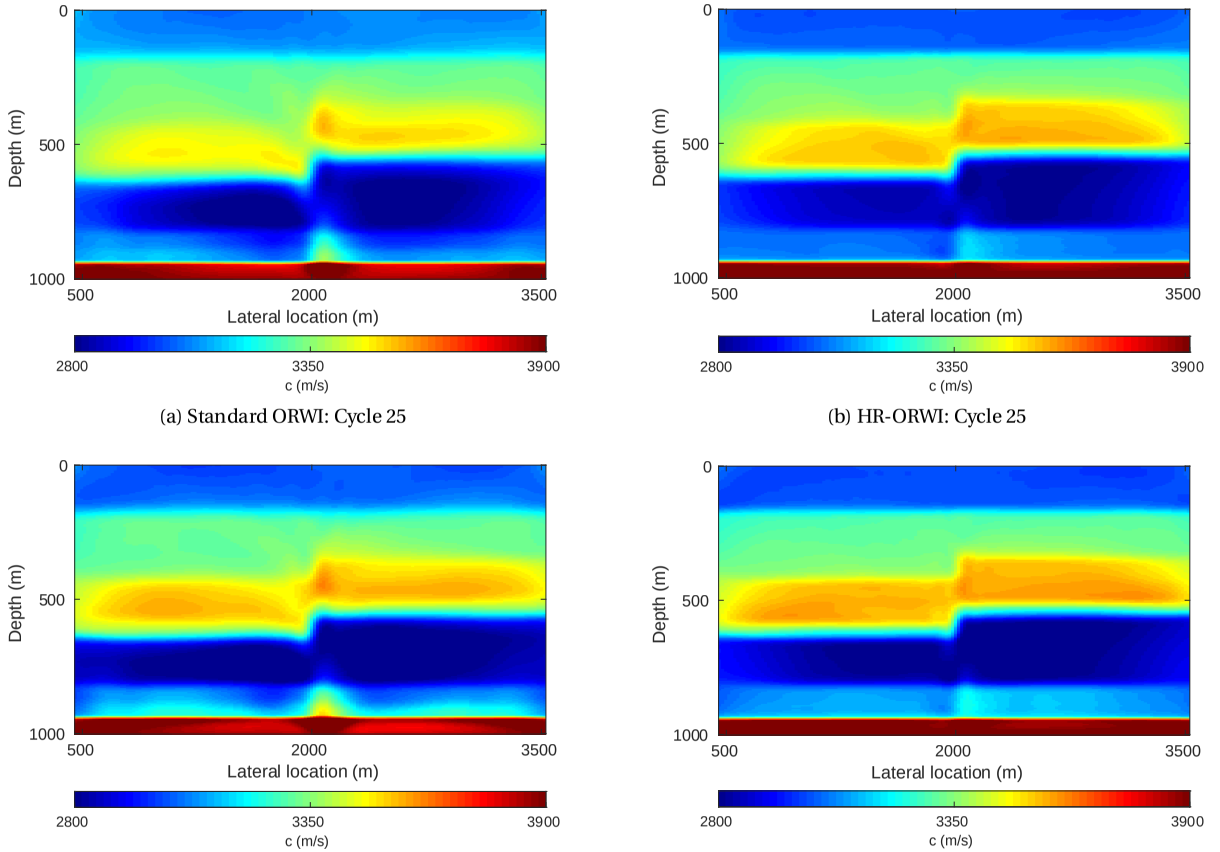


(a) Standard ORWI: Cycle 1 (b) HR-ORWI: Cycle 1

Figure 4.8: Background velocity perturbations associated with the fault model example estimated after one cycle (1x LS migration and 1x LS tomography) of ORWI. (a) Background velocity perturbation estimated via standard ORWI. (b) Background velocity perturbation estimated via HR-ORWI.

velocity perturbations are visible from shallow to deep. Emphasizing the reliability of HR-ORWI, this comparison also highlights the potential confusion that might be introduced with standard ORWI.

By comparing Figs 4.9a and 4.9c to Figs 4.9b and 4.9d, respectively, it becomes evident that the estimated tomogram via HR-ORWI after 25 cycles closely aligns with, or even excel, the estimated tomogram via standard ORWI after 50 cycles. Further comparison also confirms the existence of persistent and non-physical artifacts in tomograms estimated via standard ORWI from the first to the fiftieth cycle. The accuracy of the tomogram estimated via HR-ORWI after 50 cycles (Fig. 4.9d) is self-evident. Fig. 4.10 (convergence history) again validates that HR-ORWI exhibits faster convergence and is able to achieve a better data fit in fewer cycles compared to standard ORWI.



(a) Standard ORWI: Cycle 25

(b) HR-ORWI: Cycle 25

(c) Standard ORWI: Cycle 50

(d) HR-ORWI: Cycle 50

Figure 4.9: Estimated tomograms associated with the fault model example. (a) and (c) Estimated tomograms after 25 and 50 cycles (each cycle involves 1x LS migration and 1x LS tomography) of standard ORWI. (b) and (d) Estimated tomograms after 25 and 50 cycles (each cycle involves 1x LS migration and 1x LS tomography) of HR-ORWI.

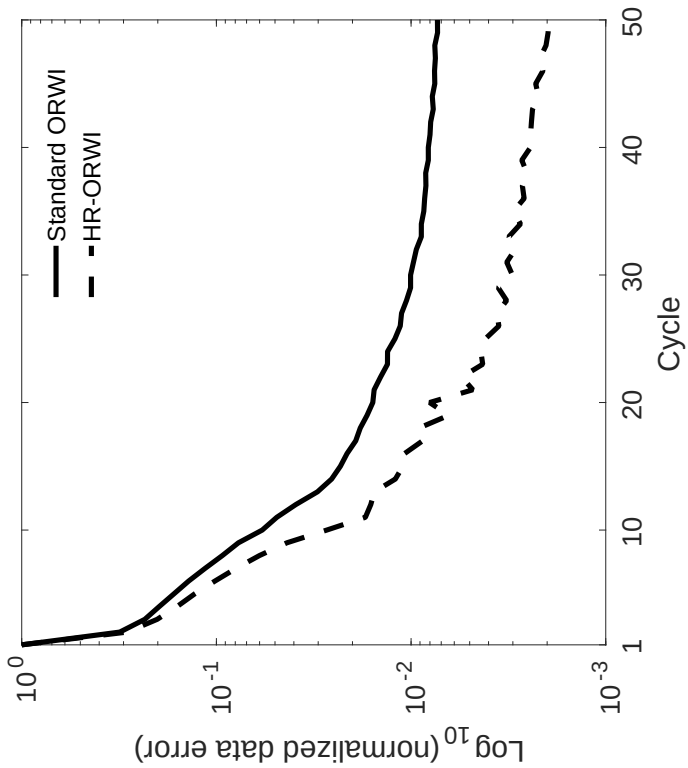
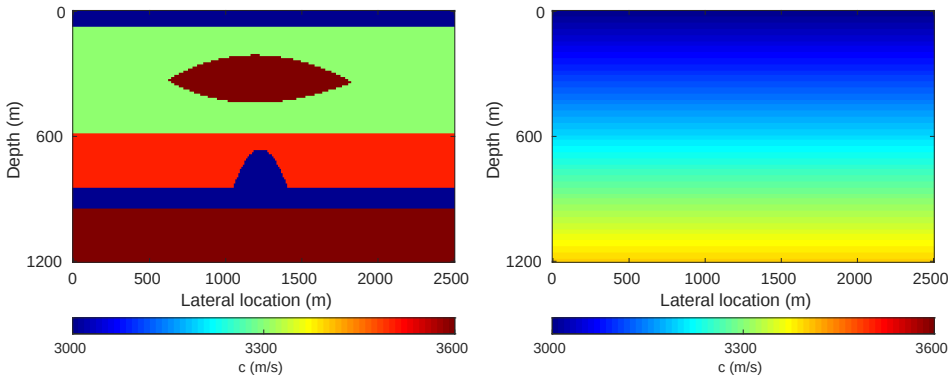


Figure 4.10: Convergence history (data error curve) associated with the fault model example.

4.3.2. RESERVOIR MODEL

In the reservoir model example, a buried low-velocity reservoir lies beneath a lens-shaped geological formation with high velocity (Fig. 4.11a), discretized into a grid of 251 points horizontally and 121 points vertically, both spaced at 10 m intervals. The observed dataset is acquired with a fixed-spread acquisition geometry, including 26 shots with 100 m spacing and 251 receivers, both distributed at the surface of the Earth. The dataset contains 1.4 s records, generated with a 15 Hz Ricker wavelet as the source function using PWMod, including only primary reflections. The maximum available offset in the dataset is limited to 2000 m. The initial background velocity model is a 1D linearly increasing gradient model, illustrated in Fig. 4.11b, and the initial reflectivity model is zero. We fix



(a) True model (b) Initial model

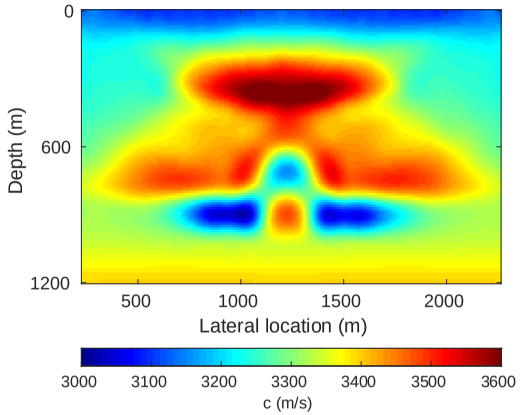
Figure 4.11: Reservoir model example. (a) True background velocity model. (b) Initial background velocity model.

the migration offset at 1000 m (MEMO), use the mid-to-far-offset residual data (|offsets| > 500 m) for tomography, and exclude the cycle-skipped far-offset data for tomography in each cycle. The 0-25 Hz frequency band data is inverted at once (no multi-scaling), and each cycle of ORWI includes 1x LS migration and 1x LS tomography. Fig. 4.12 compares the estimated tomograms using standard ORWI and HR-ORWI after 25 and 50 cycles.

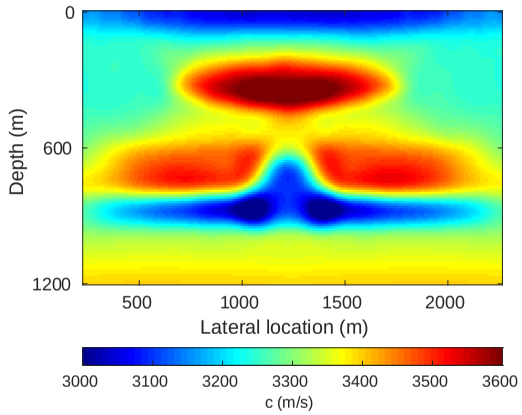
Comparing Fig. 4.12a with Fig. 4.12b reveals that standard ORWI, after 25 cycles, still struggles to distinguish between the lens body and the underlying flat layer, whereas HR-ORWI reliably resolves this. Furthermore, while standard ORWI completely misestimates the background velocity of the reservoir body, HR-ORWI shows an accurate retrieval.

Comparing Fig. 4.12c with Fig. 4.12b reveals that the tomogram estimated with HR-ORWI after 25 cycles not only aligns with the tomogram estimated with standard ORWI after 50 cycles but also surpasses in accuracy, especially at the reservoir level; this occurs while both show nearly identical data errors, as illustrated in Fig. 4.13, with no cycle-skipping observed in the fitted data. The accuracy of the tomogram estimated via HR-ORWI after 50 cycles (Fig. 4.12d) is self-evident.

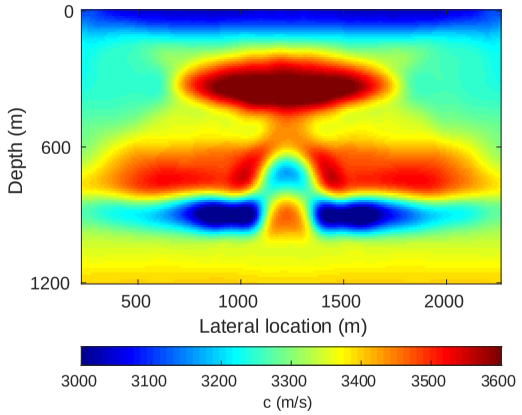
Fig. 4.13 (convergence history) again verifies that HR-ORWI comes with faster convergence and can reach a better data fit in fewer cycles compared to standard ORWI.



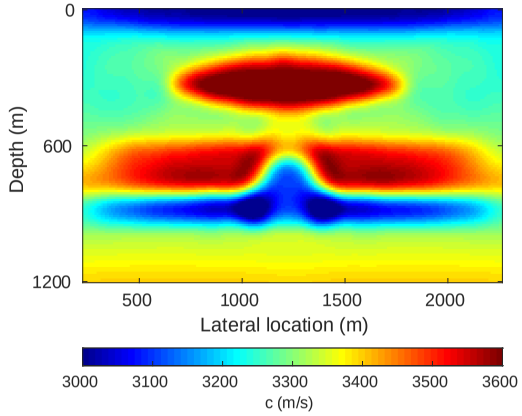
(a) Standard ORWI: Cycle 25



(b) HR-ORWI: Cycle 25



(c) Standard ORWI: Cycle 50



(d) HR-ORWI: Cycle 50

Figure 4.12: Estimated tomograms associated with the reservoir model example. (a) and (c) Estimated tomograms after 25 and 50 cycles (each cycle involves 1x LS migration and 1x LS tomography) of standard ORWI. (b) and (d) Estimated tomograms after 25 and 50 cycles (each cycle involves 1x LS migration and 1x LS tomography) of HR-ORWI.

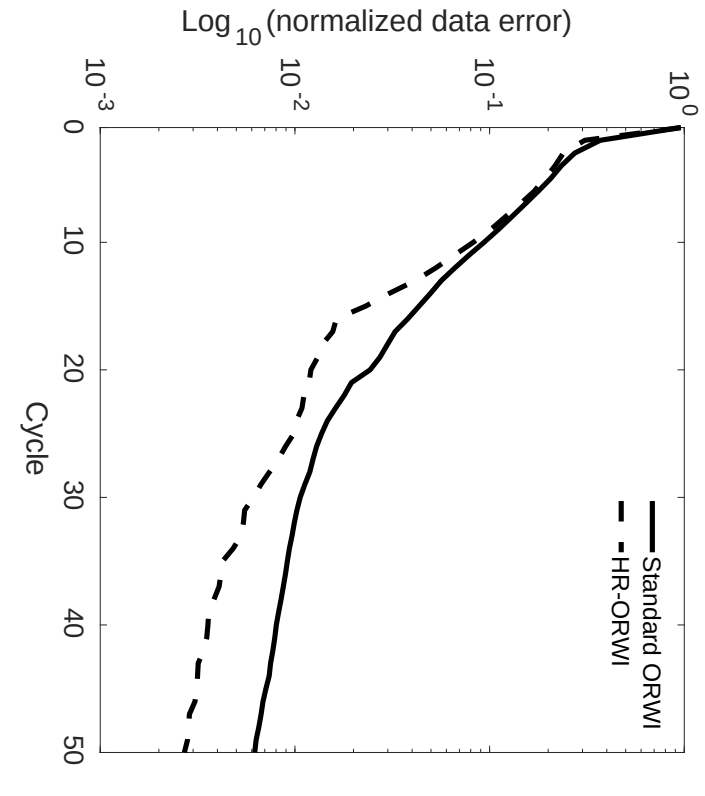
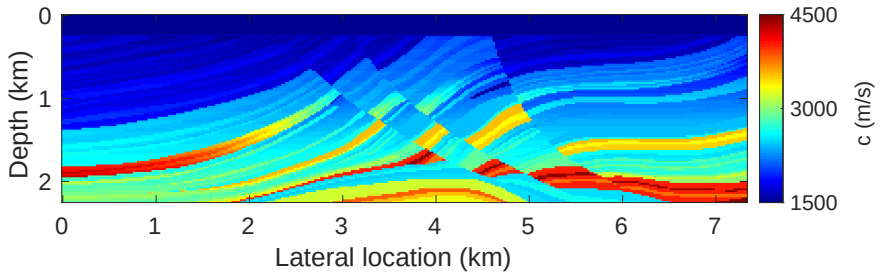


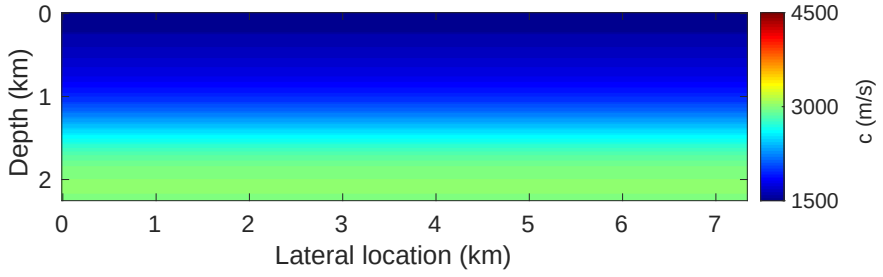
Figure 4.13: Convergence history (data error curve) associated with the reservoir model example.

4.3.3. MARMOUSI2 MODEL

The Marmousi2 model example (Fig. 4.14) contains 334 × 103 gridpoints with intervals of 22 m. The observed data is generated via PWMod (only primary reflections) with a 10 Hz Ricker wavelet as the source function. With a shooting interval of nearly 200 m, 41 shots are triggered at the surface of the Earth, and each shot is recorded by 334 receivers planted on the surface of the Earth (fixed-spread acquisition). The data is recorded for 4.096 s, and the maximum available offset is limited to 4 km. To build the initial background velocity model, the true background velocity model is smoothed out by applying a square Gaussian kernel with a standard deviation of 1100 m. Subsequently, a single vertical profile is chosen to construct the initial 1D model with a water layer on top, as shown in Fig. 4.14b. The initial reflectivity model is set to zero. We fix the migration



(a) True model



(b) Initial model

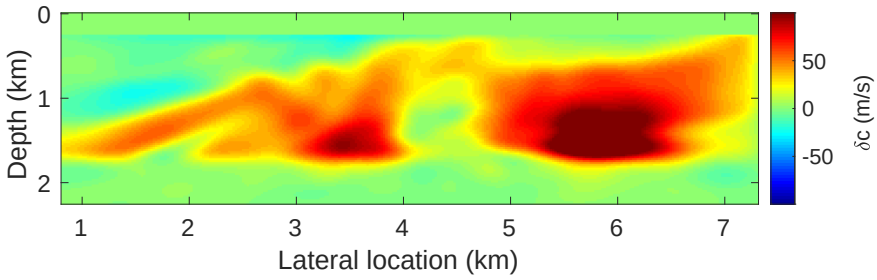
Figure 4.14: Marmousi2 model example. (a) True background velocity model. (b) Initial background velocity model.

offset at 2 km (MEMO), use the mid-to-far-offset residual data (|offsets| > 0.5 km) for tomography, and exclude the cycle-skipped far-offset data for tomography in each cycle. As we advance through cycles and update the background velocity, a growing number of offsets contribute to tomography as they are no longer cycle-skipped. This process continues until all offsets up to 4 km have been taken into account.

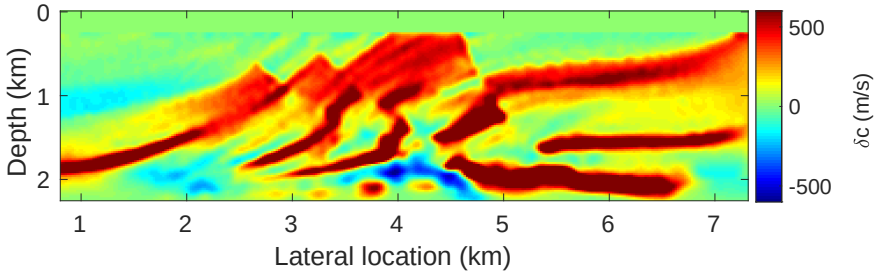
To invert the data within the 0-18 Hz frequency band, a multi-scaling strategy is employed, involving three frequency bands: 0-5 Hz, 0-11 Hz, and 0-18 Hz. For both 0-5 Hz and 0-11 Hz, the inversion process involves 20 cycles, each with 1x LS migration iteration and 1x LS tomography iteration. The image doesn't reset to zero after each cycle. For the 0-18 Hz band, the inversion process involves a total of 60 cycles, each with 1x

LS migration iteration and 1x LS tomography iteration, and no image reset after each cycle. Note that the reflectivity model is reset back to zero after each frequency scale. The estimated background velocity perturbations (cycles 1 and 100) and tomogram (cycle 100) through HR-ORWI are displayed in Figs 4.15 and 4.16, respectively.

The estimated background velocity perturbation after the first cycle (Fig. 4.15a) illustrates that the preconditioned tomographic gradient reliably points in the right direction even from the outset. The main reflections have been correctly interpreted. The tomo-



(a) Cycle 1



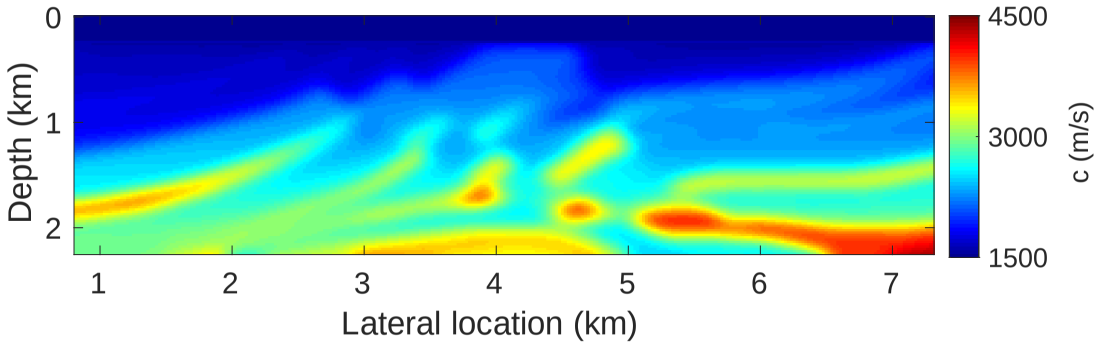
(b) Cycle 100

Figure 4.15: Background velocity perturbations associated with the Marmousi2 model example. (a) Background velocity perturbation estimated after the first cycle of HR-ORWI. (b) Background velocity perturbation estimated after the 100th cycle of HR-ORWI.

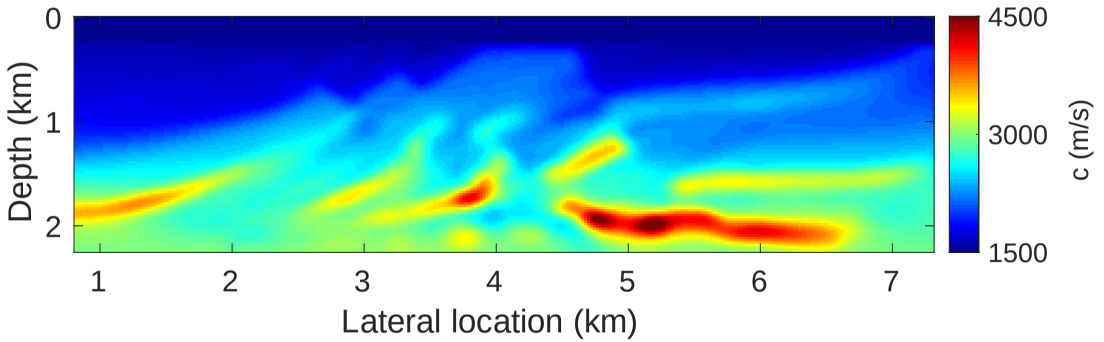
graphic update focuses primarily on reflected events within the depth range of 0-1.5 km. Fig. 4.15b shows the estimated background velocity perturbation after the 100th cycle. It is evident that the tomographic gradient of HR-ORWI successfully samples the low-to-intermediate model wavenumbers of the true background velocity properly in 100 cycles from shallow to deep, though with a depth-dependent decreasing resolution (Williamson, 1991).

The estimated tomogram after 100 cycles is shown in Fig. 4.16a. In addition, a smoothed version of the true background velocity model with the aid of a square Gaussian kernel with a standard deviation of 75 m and a water layer on top is generated and shown in Fig. 4.16a. A comparison of the two models presented in Fig. 4.16 indicates that the estimated tomogram has a resolution approximately similar to that of the smoothed true model.

To check the quality of the estimated tomogram, different vertical profiles located at

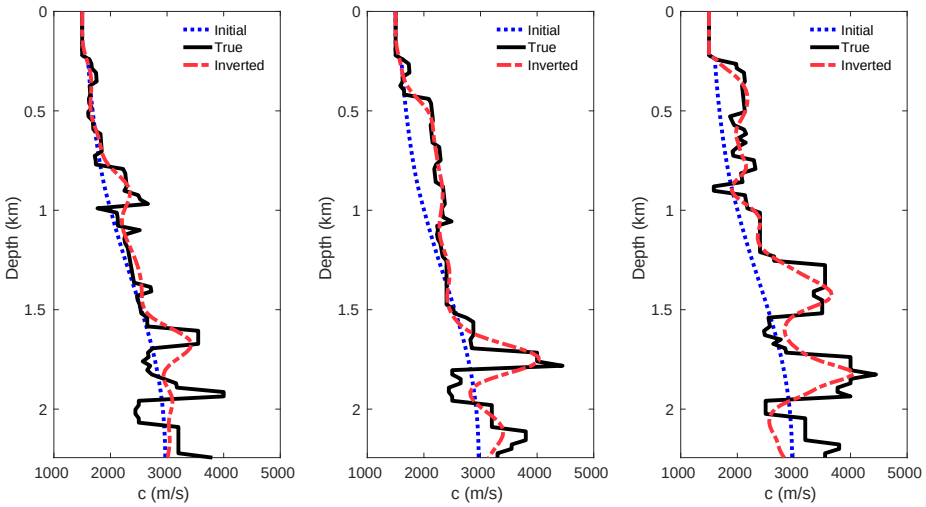


(a) Smoothed true background velocity model



(b) Estimated Model: Cycle 100

Figure 4.16: (a) Smoothed true background velocity model of Marmousi2 generated via a square Gaussian kernel with a standard deviation of 75 m and a water layer on top. (b) Estimated tomogram associated with the Marmousi2 model example after 100 cycles of HR-ORWI.



(a) Lateral location: 3000 m (b) Lateral location: 3750 m (c) Lateral location: 4600 m

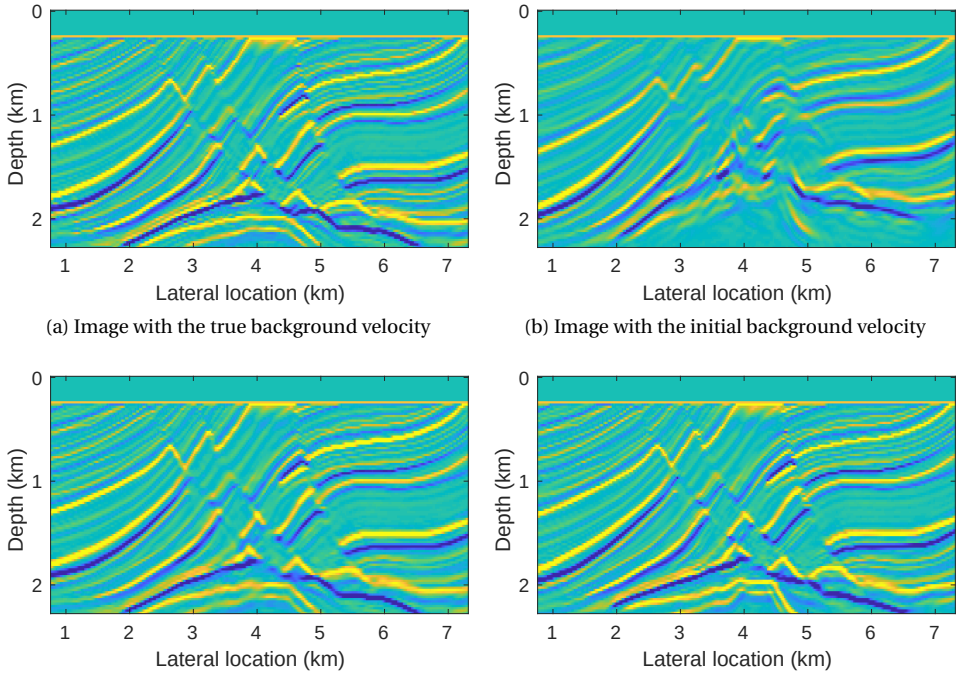
Figure 4.17: Vertical background velocity profiles associated with the Marmousi2 model example. A vertical profile of the estimated tomogram at the lateral location of (a) 3000 m, (b) 3750 m, and (c) 4600 m.

the lateral locations of 3000, 3750, and 4600 m are extracted and plotted out in Fig. 4.17. It can be seen that the estimated background velocity in all extracted lateral locations is smooth but significantly close to the true velocity.

To assess the extent of uplift in the background velocity from cycle 1 to cycle 100, various images using different background velocity models, including true, initial, smoothed, and estimated, are presented in Fig. 4.18. Upon comparing the images mapped with the initial and estimated background velocity models (Figs 4.18b and 4.18d), a significant kinematic uplift can be observed in Fig. 4.18d, attributed to the enhanced focus of the reflectors. After comparing the images mapped with the true and estimated background velocity models (Figs 4.18a and 4.18d), a significant kinematic uplift is again confirmed in Fig. 4.18d based on the reasonably accurate positioning of the reflectors. Comparing the images mapped with the smoothed and estimated background velocity models (Figs 4.18c and 4.18d) also suggests that a "maximum" kinematic uplift within the achievable resolution of RWI has occurred in Fig. 4.18d.

4.4. DISCUSSION

The one-way (up and down) wave-equation extrapolation technique has persisted in the seismic imaging industry despite its weaknesses (for details on weaknesses, see Mulder and Plessix, 2004), as it offers computational efficiency when contrasted with direct solutions to the Helmholtz equation. In this research, we demonstrated the potential use of one-way wave-equation forward modeling, i.e., PWMod, in efficient gradient preconditioning to enhance suboptimal images and tomograms resulting from RWI. In terms of constructing, inverting, or storing, depth-dependent approximate Hessians computed



(a) Image with the true background velocity

(b) Image with the initial background velocity

(c) Image with the smoothed background velocity

(d) Image with the estimated background velocity

Figure 4.18: Images using different background velocity models associated with the Marmousi2 model example. The mapped images correspond to the (a) true, (b) initial, (c) smoothed, and (d) estimated background velocity models.

based on PWMod are more feasible than doing so for preconditioners computed at once for all model parameters within the medium. Despite being feasible, it remains time-intensive compared to standard ORWI (i.e., HR-ORWI is six to nine times more expensive but requires half the iterations of standard ORWI, effectively making it three to five times more expensive overall), given the current implementation and the capabilities of the current computing resources. However, there are ways to reduce the time intensity. Among others, source subsampling in the computation of $\mathbf{H}^{\mathrm{a}}$ (Matharu and Sacchi, 2019) sounds promising. According to our in-house inquiries, HR-ORWI with source subsampling in constructing $\mathbf{H}^{\mathrm{a}}_{\omega}$—specifically, one source out of three—could retrieve high-resolution tomograms consistent with this study. Even though each iteration would still take two to three times longer, the retrieval is reached in half the iterations of standard ORWI, making it almost equally expensive overall.

Selecting one technology over another often involves evaluating the cost-versus-quality trade-off. While cost-effectiveness must remain a key priority, compromising accuracy due to budget constraints in exploration seismology can raise significant concerns. In the reservoir model example presented in this paper, we clearly observed how choosing a budget-conscious technology (i.e., standard ORWI) can lead to incorrect subsurface

illumination, resulting in misinterpretations of subsurface properties and potentially far-reaching consequences. Using cheaper technology can indeed be a double-edged sword; while it saves money upfront, the outcome could be irreversible. This portrays the complexity of decision-making dynamics in exploration seismology.

Although we did not use any model-based regularization in this study, integrating approaches like structure-oriented regularization guided by the reflectivity model may further enhance the fidelity of tomographic updates in HR-ORWI (Masaya and Verschuur, 2018; Yao et al., 2019; Provenzano et al., 2023).

Real data applications require careful consideration of several additional factors. The negative potential impact of source wavelet estimation errors and the absence of low-frequency data on the stability of HR-ORWI should be carefully explored. Furthermore, addressing amplitude versus offset (AVO) effects in the ORWI process is crucial. Qu and Verschuur (2021) recommend minimizing AVO effects before applying ORWI. In this study, through inverse-crime numerical examples, we assumed reduced AVO effects in the data, facilitating a thorough examination of ORWI convergence. Challenges may also arise in accommodating anisotropy effects and integrating the Q-effect into ORWI, as highlighted in other studies (Alshuhail and Verschuur, 2019; Safari and Verschuur, 2023).

4.5. CONCLUSIONS

Limited or unresolved resolution in both migration and reflection tomography components of RWI can lead to suboptimal velocity estimation. To address this within the context of ORWI, we introduced a preconditioning strategy that enhanced both the migration and tomography loops, resulting in high-resolution updates for both reflectivity and background velocity. We named this approach HR-ORWI, where diagonal blocks of the mono-frequency Gauss-Newton Hessian operators, each corresponding to a specific virtual depth level, are extracted and inverted to precondition their corresponding mono-frequency gradients. This approach performs a partial deconvolution of the gradient while keeping memory and computational requirements manageable. Through three synthetic examples, we demonstrated that HR-ORWI generates superior background velocity estimations compared to standard ORWI.

REFERENCES

Abolhassani, S. and E. Verschuur (2022). “Fast Gauss-Newton full-wavefield migration”. In: SEG/AAPG Second International Meeting for Applied Geoscience & Energy, Expanded Abstracts.

Abolhassani, S. and D. J. Verschuur (2024). “Efficient preconditioned least-squares wave-equation migration”. In: *Geophysics* 89.3, S275–S288.

Alkhalifah, T. (2014). “Scattering-angle based filtering of the waveform inversion gradients”. In: *Geophysical Journal International* 200.1, pp. 363–373.

Almomin, A. and B. Biondi (2012). “Tomographic full waveform inversion: Practical and computationally feasible approach”. In: SEG Annual International Meeting, Expanded Abstracts.

Alshuhail, A. A. and D. J. Verschuur (2019). "Robust estimation of vertical symmetry axis models via joint migration inversion: Including multiples in anisotropic parameter estimation". In: *Geophysics* 84.1, pp. C57–C74.

Aoki, N. and G. T. Schuster (2009). "Fast least-squares migration with a deblurring filter". In: *Geophysics* 74 (6), WCA83–WCA93.

Asnaashari, A. et al. (2013). "Regularized seismic full waveform inversion with prior model information". In: *Geophysics* 78.2, R25–R36.

Assis, C. A. M. et al. (2022). "Tomographic Hessian-Based Inversion Velocity Analysis: a First Evaluation". In: 83rd EAGE Annual Conference & Exhibition, Extended Abstracts.

Assis, C. A. and J. Schleicher (2021). "Introduction of the Hessian in joint migration inversion and improved recovery of structural information using image-based regularization". In: *Geophysics* 86 (6), R777–R793.

Audebert, F. and D. Ortiz-Rubio (2018). "Separation of scales in FWI and RFWI. Filling the gap? Or exploiting it?" In: SEG Annual International Meeting, Expanded Abstracts.

Berkhout, A. J. (1982). *Seismic migration, imaging of acoustic energy by wave field extrapolation, A: theoretical aspects*. Elsevier.

— (2012). "Combining full wavefield migration and full waveform inversion, a glance into the future of seismic imaging". In: *Geophysics* 77.2, S43–S50.

— (2014). "Review paper: An outlook on the future seismic imaging, part III: Joint Migration Inversion". In: *Geophysical Prospecting* 62.5, pp. 950–971.

Berkhout, A. (1981). "Wave field extrapolation techniques in seismic migration, a tutorial". In: *Geophysics* 46.12, pp. 1638–1656.

Biondi, B. and W. W. Symes (2004). "Angle-domain common-image gathers for migration velocity analysis by wavefield-continuation imaging". In: *Geophysics* 69.5, pp. 1283–1298.

Brittan, J. and I. Jones (2019). "FWI evolution — From a monolith to a toolkit". In: *The Leading Edge* 38.3, pp. 179–184.

Brossier, R., S. Operto, and J. Virieux (2009). "Seismic imaging of complex onshore structures by 2D elastic frequency-domain full-waveform inversion". In: *Geophysics* 74.6, WCC105–WCC118.

Bunks, C. et al. (1995). "Multiscale seismic waveform inversion Geophysics". In: *Geophysics* 60.5, pp. 1457–1473.

Chauris, H. and M. Noble (2001). "Two-dimensional velocity macro model estimation from seismic reflection data by local differential semblance optimization: applications to synthetic and real data sets". In: *Geophysical Journal International* 144.1, pp. 14–26.

Chauris, H. and E. Cocher (2017). "From migration to inversion velocity analysis". In: *Geophysics* 82 (3), S207–S223.

Chavent, G. and R. E. Plessix (1999). "An optimal true-amplitude least-squares prestack depth-migration operator". In: *Geophysics* 64.2, pp. 508–515.

Choi, Y., D.-J. Min, and C. Shin (2008). "Frequency-domain elastic full waveform inversion using the new pseudo-Hessian matrix: Experience of elastic Marmousi-2 synthetic data". In: *Bulletin of the Seismological Society of America* 98.5, pp. 2402–2415.

Clément, F., G. Chavent, and S. Gómez (2001). "Migration-based traveltime waveform inversion of 2-D simple structures: A synthetic example". In: *Geophysics* 66.3, pp. 845–860.

Fletcher, R. P. et al. (2016). “Least-squares migration — Data domain versus image domain using point spread functions”. In: *The Leading Edge* 35.2, pp. 157–162.
Gomes, A. and Z. Yang (2017). “Improving reflection FWI reflectivity using LSRTM in the curvelet domain”. In: SEG Annual International Meeting, Expanded Abstracts.
Guitton, A. (2004). “Amplitude and kinematic corrections of migrated images for nonunitary imaging operators”. In: *Geophysics* 69.4, pp. 1017–1024.
Guitton, A. (2017). “Fast 3D least-squares RTM by preconditioning with nonstationary matching filters”. In: SEG Annual International Meeting, Expanded Abstracts.
Guo, S. and H. Wang (2020). “Image domain least-squares migration with a Hessian matrix estimated by non-stationary matching filters”. In: *Journal of Geophysics and Engineering* 17.1, pp. 148–159.
Hou, J. and W. W. Symes (2015). “An approximate inverse to the extended Born modeling operator”. In: *Geophysics* 80 (6), R331–R349.
Hu, J., G. T. Schuster, and P. Valasek (2001). “Poststack migration deconvolution”. In: *Geophysics* 66.3, pp. 939–952.
Jang, U., D.-J. Min, and C. Shin (2009). “Comparison of scaling methods for waveform inversion”. In: *Geophysical Prospecting* 57.1, pp. 49–59.
Jun, H., E. Park, and C. Shin (2015). “Weighted pseudo-Hessian for frequency-domain elastic full waveform inversion”. In: *Journal of Applied Geophysics* 123, pp. 1–17.
Lameloise, C.-A., H. Chauris, and M. Noble (2015). “Improving the gradient of the image-domain objective function using quantitative migration for a more robust migration velocity analysis”. In: *Geophysical Prospecting* 63.2, pp. 391–404.
Lecomte, I. (2008). “Resolution and illumination analyses in PSDM: A ray-based approach”. In: *The Leading Edge* 27.5, pp. 650–663.
Li, A. et al. (2024). “An accurate propagator for heterogeneous media in full wavefield migration”. In: *IEEE Transactions on Geoscience and Remote Sensing*.
Lu, S. et al. (2018). “Least-squares full-wavefield migration”. In: *The Leading Edge* 37.1, pp. 46–51.
Martin, G. S., R. Wiley, and K. J. Marfurt (2006). “Marmousi2: An elastic upgrade for Marmousi”. In: *The leading edge* 25.2, pp. 156–166.
Masaya, S. and D. Verschuur (2018). “Iterative reflectivity-constrained velocity estimation for seismic imaging”. In: *Geophysical Journal International* 214.1, pp. 1–13.
Matharu, G. and M. Sacchi (2019). “A subsampled truncated-Newton method for multiparameter full-waveform inversion”. In: *Geophysics* 84.3, R333–R340.
Métivier, L. et al. (2013). “Full waveform inversion and the truncated Newton method”. In: *SIAM Journal on Scientific Computing* 35.2, B401–B437.
Mora, P. (1989). “Inversion = migration + tomography”. In: *Geophysics* 54.12, pp. 1575–1586.
Mulder, W. A. and R.-E. Plessix (2004). “A comparison between one-way and two-way wave-equation migration”. In: *Geophysics* 69.6, pp. 1491–1504.
Oh, J.-W. and D.-J. Min (2013). “Weighting technique using backpropagated wavefields incited by deconvolved residuals for frequency-domain elastic full waveform inversion”. In: *Geophysical Journal International* 194.1, pp. 322–347.
Plessix, R. E. and W. A. Mulder (2004). “Frequency-domain finite-difference amplitude-preserving migration”. In: *Geophysical Journal International* 157.3, pp. 975–987.

Plessix, R. E. et al. (2013). "Multiparameter full waveform inversion: marine and land examples". In: *The Leading Edge* 32.9, pp. 1030–1038.

Pratt, R. G., C. Shin, and G. J. Hicks (1998). "Gauss–Newton and full Newton methods in frequency-space seismic waveform inversion". In: *Geophysical Journal International* 133.2, pp. 341–362.

Provenzano, G., R. Brossier, and L. Métivier (2023). "Robust and efficient waveform-based velocity-model building by optimal transport in the pseudotime domain: Methodology". In: *Geophysics* 88.2, U49–U70.

Qu, S. and D. Verschuur (2021). "An effective scheme of pseudo-time Joint migration inversion with an AVO mitigating workflow". In: 82th EAGE Annual Conference & Exhibition, Extended Abstracts.

Safari, M. and J. Verschuur (2023). "Joint Migration Inversion including Q effects: towards Q estimation". In: 84th EAGE Annual Conference & Exhibition, Extended Abstracts.

Sava, P. and B. Biondi (2004). "Wave-equation migration velocity analysis. I. Theory". In: *Geophysical Prospecting* 52.6, pp. 593–606.

Shen, P. (2005). "Wave equation migration velocity analysis by differential semblance optimization". PhD thesis. Rice University.

Shen, P. and W. W. Symes (2008). "Automatic velocity analysis via shot profile migration". In: *Geophysics* 73.5, VE49–VE59.

Shen, P., W. W. Symes, and C. C. Stolk (2003). "Differential semblance velocity analysis by wave-equation migration". In: SEG Annual International Meeting, Expanded Abstracts.

Shin, C., C. S. Jang, and D.-J. Min (2001). "Improved amplitude preservation for prestack depth migration by inverse scattering theory". In: *Geophysical Prospecting* 49.5, pp. 592–606.

Staal, X. R. (2015). "Combined imaging and velocity estimation by Joint Migration Inversion". PhD thesis. Delft University of Technology.

Sun, Verschuur, and Qu (2019). "Research note: Derivations of gradients in angle-independent joint migration inversion". In: *Geophysical Prospecting* 67.3, pp. 572–579.

Sun, Y. and E. Verschuur (2020). "Full wavefield modeling using rigorous one-way propagation, reflection and transmission operators". In: *SEG International Exposition and Annual Meeting*. SEG.

Symes, W. W. (2008). "Migration velocity analysis and waveform inversion". In: *Geophysical Prospecting* 56.6, pp. 765–790.

Tang, Y. (2009). "Target-oriented wave-equation least-squares migration/inversion with phase-encoded Hessian". In: *Geophysics* 74.6, WCA95–WCA107.

Tarantola, A. (1984). "Inversion of seismic reflection data in the acoustic approximation". In: *Geophysics* 49.8, pp. 1259–1266.

ten Kroode, F. (2014). "A Lie group associated to seismic velocity estimation". In: Inverse Problems from Theory to Applications Conference, Institute of Physics Publishing, 142-146.

Thorbecke, J. W., K. Wapenaar, and G. Swinnen (2004). "Design of one-way wavefield extrapolation operators, using smooth functions in wlsq optimization". In: *Geophysics* 69.4, pp. 1037–1045.

Valenciano, A. A., B. Biondi, and A. Guitton (2006). "Target-oriented wave-equation inversion". In: *Geophysics* 71 (4), A35–A38.
Vigh, D., J. Kapoor, and H. Li (2011). "Full-waveform inversion application in different geological settings". In: SEG Annual International Meeting, Expanded Abstracts.
Wang, S. et al. (2013). "Reflection-based full waveform inversion (RFWI) in the frequency domain". In: SEG Annual International Meeting, Expanded Abstracts.
Williamson, P. (1991). "A guide to the limits of resolution imposed by scattering in ray tomography". In: *Geophysics* 56.2, pp. 202–207.
Wu, Z. and T. Alkhalifah (2015). "Simultaneous inversion of the background velocity and the perturbation in full-waveform inversion". In: *Geophysics* 80.6, R317–R329.
Xu, S. et al. (2012). "Inversion on reflected seismic wave". In: SEG Annual International Meeting, Expanded Abstracts.
Yang, J. et al. (2021). "Approximating the Gauss-Newton Hessian using a space-wavenumber filter and its applications in least-squares seismic imaging". In: *IEEE Transactions on Geoscience and Remote Sensing* 60, pp. 1–13.
Yang, J. et al. (2022). "An Efficient and Stable High-Resolution Seismic Imaging Method: Point-Spread Function Deconvolution". In: *Journal of Geophysical Research: Solid Earth* 127.7, e2021JB023281.
Yao, G., D. Wu, and S. Wang (2020). "A review on reflection-waveform inversion". In: *Petroleum Science* 17, pp. 334–351.
Yao, G., N. V. da Silva, and D. Wu (2019). "Reflection-waveform inversion regularized with structure-oriented smoothing shaping". In: *Pure and Applied Geophysics* 176.12, pp. 5315–5335.
Yu, J. et al. (2006). "Prestack migration deconvolution". In: *Geophysics* 71 (2), S53–S62.
Zhou, H., L. Amundsen, and G. Zhang (2012). "Fundamental Issues in Full Waveform Inversion". In: SEG Annual International Meeting, Expanded Abstracts.
Zhou, W. et al. (2015). "Full waveform inversion of diving waves for velocity model building with impedance inversion based on scale separation". In: *Geophysical Journal International* 202.3, pp. 1535–1554.
Zou, Z. et al. (2014). "Refraction full-waveform inversion in a shallow water environment". In: 76th EAGE Conference & Exhibition, Extended Abstracts.

APPENDIX

4.A. OFFSET-SELECTION STRATEGY

Despite its conceptual appeal and aside from the resolution challenges, ORWI could face other limitations due to other factors: (1) Full-wave inconsistency in the short-offset residual data for tomography due to inconsistent reflectivity and velocity models (Valensi and Baina, 2021; Liang et al., 2022; Provenzano et al., 2023). (2) The standard practice in the RWI community for mitigating the adverse imprint of inconsistent reflectivity and velocity models often involves using short-offset or near-zero offset data for migration (Provenzano et al., 2023). However, this can degrade the illumination and signal-to-noise ratio of the reflectivity model and result in an increased number of least-squares migration iterations. (3) Adverse contribution of cycle-skipped long-offset reflection data in tomography (e.g., Abolhassani and Verschuur, 2022; Provenzano et al., 2023).

4.A.1. TOMOGRAPHY OFFSET: SHORT OFFSETS

To minimize the impact of inconsistent reflectivity and velocity models on the tomographic gradient of ORWI, we propose muting their relevant erroneous tomographic wavepaths in the residual data gathers, which are typically established at short offsets.

4.A.2. MIGRATION OFFSET

Building upon the muting of short-offset residual waveforms for tomography, we extend the migration offset to improve both the illumination and the signal-to-noise ratio of the reflectors. We follow Algorithm 4.A.1, aiming for an extended offset for migration (rather than short offset) where a significant reduction in the misfit function is impracticable. The misfit indeed remains unchanged as the amplitude fit fails to better itself while mapping more out-of-phase reflection data into the model space. This denotes the maximum effective migration offset (MEMO), lying averagely between the short offset and the maximum uncycle-skipped offset in the data. Algorithm 4.A.1 is applied as a quality control measure prior to ORWI. This algorithm calculates the misfit function value for various offset ranges up to the maximum offset within the acquisition. Following this, on the misfit-offset graph, we pick an offset right before the curve starts to become nearly horizontal on a logarithmic scale.

Algorithm 4.A.1: Migration offset analysis pseudo-code

```
for h ← 0 to max(offset) do
    image ← 1x LS migration with h
    d⁻_mod ← forward modeling with image and background velocity model
    Δd⁻ ← d⁻_obs − d⁻_mod        ▷ Δd⁻ is the full-offset residual data
    store the misfit function value associated with h
```

4.A.3. TOMOGRAPHY OFFSET: LONG OFFSETS

A given pair of modeled and observed seismic traces in the time domain are called cycle-skipped traces if their time distance is larger than half a cycle time shift (Virieux and Operto, 2009).

Algorithm 4.A.2: Cycle-skipping check pseudo-code

```
▷ D_obs and D_mod are 3D matrices representing the observed and modeled data (full-shot).
▷ n_t × n_g × n_s is the size of D_obs.
▷ n_t is the total number of time samples.
▷ n_g and n_s are the total number of receivers and the total number of shots.
▷ ℓ_w and ℓ_r are the correlation-window length and reference lag.
▷ ω_c is the current maximum frequency in D.
▷ ω_m is the peak frequency of the seismic wavelet.
▷ win_c is the current correlation window.
function CycleSkippingCheck(D_obs, D_mod, n_g, n_s, n_t, ℓ_w, ℓ_r, ω_c, ω_m)
    TIMETABLE ← a matrix of ones of size: n_g × n_s
    for ishot ← 1 to n_s do
        φ ← determine the anchor offset index in the current shot gather
        ι ← φ
        κ ← φ − 1
        for m ← [1, 2] do
            flag ← true
            while (m = 1 and flag = true and ι ≤ n_g) or (m = 2 and flag = true and κ ≤ 1) do
                d_obs, n_obs ← current observed trace and the associated first-break sample
                d_mod ← current modeled trace
                if ω_c > ω_m then
                    d_obs ← d_obs[n_obs : n_t]
                    d_mod ← d_mod[n_obs : n_t]
                else
                    d_obs ← d_obs[1 : n_t]
                    d_mod ← d_mod[1 : n_t]
                calculate cross-correlation between d_obs and d_mod        ▷ global cross-correlation
                if max correlation lag < ℓ_r then
                    win ← [1 : ℓ_w : n_t]
                    b ← length(win)
                    for τ ← 1 to b do
                        win_c ← [1 : win(τ + 1)]
                        calculate cross-correlation between d_obs and d_mod in win_c        ▷ local cross-correlation
                        if max correlation lag < ℓ_r then
                            if τ = b then
                                tmp ← n_t
                            else
                                tmp ← n_obs + win_c(end)
                            if m = 1 then
                                TIMETABLE[ι, ishot] ← tmp
                            else
                                TIMETABLE[κ, ishot] ← tmp
                        else
                            flag ← false
                            break
                else
                    flag ← false
                if m = 1 then
                    ι = ι + 1
                else
                    κ = κ − 1
    return TIMETABLE
```

Inspired by van Leeuwen and Mulder (2010), we defined Algorithm 4.A.2 as a time-domain two-step data-selection algorithm excluding the contribution of the cycle-skipped long-offset residual data in the tomography loop of ORWI to obtain a tomographic update free of the damaging effect of cycle-skipping in each iteration. In Algorithm 4.A.2, we identify a given pair of time-domain modeled and observed seismic traces as cycle-skipped if their maximum correlation lag exceeds a reference lag determined by the dominant period of the observed trace. Based on this criterion, our time-domain data-selection algorithm relies on the cross-correlation of these traces within a sliding window (referred

to as local cross-correlation), safeguarded by a preliminary global cross-correlation of the given traces. This global cross-correlation functions as a protection, protecting subsequent local cross-correlations from irregularities in the modeled waveform. Without this safeguard, the outputs of the local cross-correlations would not meet the required accuracy.

5

A Comparative Study of Three Approximate Hessians in Wave-Equation Migration

"The most certain way to succeed is always to try just one more time."

– Thomas Edison

Enhanced pre-stack depth migration, characterized by improved resolution and amplitudes, ensures a more accurate representation of the subsurface, proving essential for reducing the likelihood of geological misinterpretations and facilitating informed decision-making in seismic exploration. However, obtaining high-resolution images with preserved amplitudes through standard depth migration could face several hurdles known as migration artifacts. Iterative least-squares migration (LSM) was developed to address these migration artifacts. However, the convergence rate of LSM using a gradient descent approach tends to be slow. Several researchers have attempted to achieve computational efficiency in linearized LSM through gradient preconditioning. In the context of iterative least-squares wave-equation migration, this extended abstract compares three approaches for minimizing an error function. This comparison is carried out through two numerical examples, one with an inverse-crime scenario and the other with a non-inverse-crime scenario.

This chapter is based on the following extended abstract: **Abolhassani, S.**, Verschuur, E. (2024). A Comparative Study of Three approximate Hessians in Wave-Equation Migration. 85th EAGE Annual Conference & Exhibition.

5.1. INTRODUCTION

Pre-stack depth migration in exploration seismology transforms seismic data into a subsurface depth image, aiming for accurate positioning of seismic amplitudes. However, obtaining high-resolution images with preserved amplitudes through standard pre-stack depth migration faces several hurdles associated with irregular acquisition geometry and organized noises, among others, degrading image accuracy, known as migration artifacts (Jones, 2018).

Iterative least-squares migration (LSM) represents an advanced variant of pre-stack depth migration, where the subsurface reflectivity model (image) receives iterative updates through an iterative migration-demigration cycle to converge towards an "ideal representation" of the reflectivity model, aiming to mitigate the migration artifacts (Nemeth et al., 1999). However, the convergence rate of LSM using a gradient descent approach tends to be rather slow. It may require quite a few iterations to get higher accuracy due to the band-limited nature of seismic data, making it a demanding computational effort.

Linearized LSM, where the forward problem is linearized through the Born approximation, aims to accelerate the convergence rate and achieve the highest possible accuracy within the resolution power of the acquired seismic data. It minimizes S, a quadratic function of $\delta\mathbf{r}^{\cup}$,

$$S(\delta\mathbf{r}^{\cup}) = \frac{1}{2}\int_{\omega} d\omega \sum_{s} \|\delta\mathbf{d}_{s,\omega} - \mathbf{J}_{s,\omega}\delta\mathbf{r}^{\cup}\|_2^2, \tag{5.1}$$

with the minimizer (Lines and Treitel, 1984),

$$\frac{\partial S}{\partial \delta\mathbf{r}^{\cup}} = 0 \quad \rightarrow \quad \delta\mathbf{r}^{\cup} = -\frac{\mathbf{Re}\Big\{\int_{\omega} d\omega \sum_{s} \mathbf{J}^{\dagger}_{s,\omega}\delta\mathbf{d}_{s,\omega}\Big\}}{\mathbf{Re}\Big\{\int_{\omega} d\omega \sum_{s} \mathbf{J}^{\dagger}_{s,\omega}\mathbf{J}_{s,\omega}\Big\}}, \tag{5.2}$$

in which $\delta\mathbf{r}^{\cup}$ is the upward reflectivity model perturbation vector, ω is an angular frequency component, s means a shot, $\delta\mathbf{d}$ denotes the data residual vector, $\dagger$ means adjoint, $\mathbf{J}$ stands for the partial derivatives of the modeled wavefield, known as the Jacobian matrix, and $\mathbf{J}^{\dagger}\mathbf{J}$ is known as the approximate Hessian matrix ($\mathbf{H}^{a}$). The upward reflectivity model can be then updated iteratively via $\mathbf{r}^{\cup}_{k+1} = \mathbf{r}^{\cup}_{k} + \alpha_k \delta\mathbf{r}^{\cup}_{k}$, where k and α are the iteration number and the minimization step length, respectively. If one assumes that S is a quadratic function of the $\delta\mathbf{r}^{\cup}_{\omega}$ (frequency-dependent reflectivity model perturbation), the minimizer transforms to (Jang et al., 2009; Oh and Min, 2013),

$$\delta\mathbf{r}^{\cup} = -\int_{\omega} d\omega\, \delta\mathbf{r}^{\cup}_{\omega} = -\int_{\omega} d\omega \frac{\mathbf{Re}\Big\{\sum_{s} \mathbf{J}^{\dagger}_{s,\omega}\delta\mathbf{d}_{s,\omega}\Big\}}{\mathbf{Re}\Big\{\sum_{s} \mathbf{J}^{\dagger}_{s,\omega}\mathbf{J}_{s,\omega}\Big\}}, \tag{5.3}$$

in which, an effective deconvolution effect on the source signature is facilitated through the sequential application of the denominator (approximate Hessian) to the gradient at each frequency, compared to Equation 5.2.

Over time, the replacement of the denominator in Equations 5.2 and 5.3 (gradient preconditioning) with cost-effective alternatives has been pursued by several researchers

Table 5.1: Investigated approaches

Approach	Error Function	Update Equation	Denominator in the Update Equation	Explanation
AP-c	$S(\delta\mathbf{r}^{\cup})$	2	"Equation 27" in Plessix and Mulder (2004) (i.e., diagonal approximation of $\mathbf{H}^{\mathrm{a}}$)	–
AP-1	$S(\delta\mathbf{r}^{\cup}_{\omega})$	3	depth-dependent version of $\sum_s \mathbf{J}^{\dagger}_{s,\omega}\mathbf{J}_{s,\omega}$ (i.e., block-diagonal approximation of $\mathbf{H}^{\mathrm{a}}$)	–
AP-2	$S(\delta\mathbf{r}^{\cup}_{\omega})$	3	depth-dependent version of $(\sum_s \mathbf{J}_{s,\omega})^{\dagger}\sum_s \mathbf{J}_{s,\omega}$ (i.e., block-diagonal approximation of $\mathbf{H}^{\mathrm{a}}$)	it includes monochromatic source interference in the denominator to be cheaper than the AP-1 approach

to ensure the computational efficiency of linearized LSM in large-scale applications (e.g., Shin et al., 2001; Plessix and Mulder, 2004; Choi et al., 2008; Lu et al., 2018). In the context of space-frequency linearized least-squares one-way wave-equation migration, which uses angle-independent reflection/transmission coefficients and involves upward/downward wavefield extrapolation between virtual data levels within the medium (Berkhout, 2014a), Abolhassani and Verschuur (2022) introduced a depth-dependent $\mathbf{H}^{\mathrm{a}}$. They partition $\mathbf{H}^{\mathrm{a}}$ into smaller operators, each capturing the correlation of partial derivative wavefields at identical depth levels. The depth-dependent approximate Hessian, constructed and applied (along with its inverse) directly at each virtual data level within the medium, efficiently reduces computational expenses. This efficiency arises because the number of model parameters per virtual data level constitutes only a small fraction of the total. Combined into a single matrix, these operators form a block-diagonal approximation for $\mathbf{H}^{\mathrm{a}}$.

In this extended abstract, we initially compare three minimization approaches, as listed in Table 5.1, using an "inverse-crime" scenario. Subsequently, given the similar performance of AP-1 and AP-2, but with AP-2 being cheaper, we further compare AP-c and AP-2 using a full-wavefield "non-inverse-crime" scenario. We end the paper with conclusions.

5.2. THEORY

In LS-WEM, we use a space-frequency one-way wavefield extrapolation scheme called primary-wavefield modeling (PWMod) (Berkhout, 2014b). This scheme models angle-independent two-way primary reflection data, including reflection and transmission effects,

$$\mathbf{p}^{-}_{\mathrm{mod}}(z_0,\omega) = \sum_{m=N}^{1} \overbrace{\left[\left[\prod_{n=1}^{m-1}\mathbf{W}^{-}_{z_{n-1};z_n}\mathbf{T}^{-}(z_n)\right]\mathbf{W}^{-}_{z_{m-1};z_m}\right]}^{\mathbf{L}^{-}_{z_0;z_m}}\left(\mathbf{r}^{\cup}(z_m)\circ\mathbf{p}^{+}_{\mathrm{mod}}(z_m,\omega)\right) \tag{5.4}$$

$$\mathbf{p}^{+}_{\mathrm{mod}}(z_m,\omega) = \left[\left[\prod_{n=m-1}^{1}\mathbf{W}^{+}_{z_{n+1};z_n}\mathbf{T}^{+}(z_n)\right]\mathbf{W}^{+}_{z_1;z_0}\right]\mathbf{s}^{+}(z_0,\omega), \tag{5.5}$$

in which $\mathbf{p}^{-}_{\mathrm{mod}}(z_0,\omega)$ is the monochromatic upgoing wavefield modeled at z_0, $\mathbf{p}^{+}_{\mathrm{mod}}(z_m,\omega_f)$ is the monochromatic downgoing wavefield modeled at z_m, $\mathbf{W}^{-}_{z_{n-1};z_n}$ is the upward phase-shift extrapolation matrix-operator from z_n to z_{n-1}, $\mathbf{W}^{+}_{z_{n+1};z_n}$ is the downward phase-shift

extrapolation matrix-operator from z_n to z_{n+1}, $\mathbf{r}^{\cup}(z_m)$ represents the angle-independent upward reflectivity vector-operator at z_m, $\mathbf{T}^{\pm}(z_n)$ denotes the downward/upward transmission diagonal-matrix-operator $\left(\mathbf{T}^{\pm}(z_n) = \mathbf{I} \pm \text{diag}(\mathbf{r}^{\cup}(z_n))\right.$, $\mathbf{s}^{+}\left(z_0, \omega_f\right)$ shows the monochromatic downgoing physical source at the Earth's surface, N is the total number of virtual depth levels, and the symbol $\circ$ means element-wise product.

Compared to PWMod, FWMod (full-wavefield modeling) takes into account the multiple scattering by replacing Equation 5.5 with (Berkhout, 2014b),

$$\mathbf{P}^{+}_{\text{mod}}(z_m,\omega) = \sum_{m=1}^{N} \mathbf{L}^{+}_{z_m;z_0}\left(\mathbf{s}^{+}(z_0,\omega) + \mathbf{r}^{\cap}(z_m) \circ \mathbf{p}^{-}_{\text{mod}}(z_m,\omega)\right), \tag{5.6}$$

where $\mathbf{r}^{\cap}(z_m)$ is the angle-independent downward reflectivity vector-operator at z_m with $\mathbf{r}^{\cap} = -\mathbf{r}^{\cup}$. Now, running the recursive summation in Equation 5.4 for multiple iterations results in multiple scattering.

In the introduction, we outlined three minimization approaches for linearized least-squares one-way wave-equation migration, listed in Table 5.1, to be examined and compared in this study, namely AP-c, AP-1, and AP-2, where AP stands for amplitude preserved. The nominator in all update equations ("gradient") for a pair of source and frequency components and a given depth level (z_m) reads,

$$\mathbf{g}_s(z_m,\omega) = \mathbf{Re}\left\{\left(\mathbf{p}^{+}_{\text{mod},s}(z_m,\omega)\right)^{*} \circ \overbrace{\left(\mathbf{L}^{-\dagger}_{z_0;z_m}\delta\mathbf{d}^{-}_{s}(\omega)\right)}^{\text{back-propagated wavefield}}\right\}, \tag{5.7}$$

where $*$ denotes the complex conjugate transpose.

To calculate the denominators for AP-c, AP-1, and AP-2, the only absent component is the Jacobian matrix. Abolhassani and Verschuur (2022) presented the expression of $\mathbf{J}_s(z_m,\omega)$ as,

$$\begin{aligned}\mathbf{J}_s(z_m,\omega) = \Big[& p^{+}_{1\,\text{mod},s}\left(z_m,\omega_f\right)\mathbf{L}^{-}_{*,1\,z_0;z_m} \quad p^{+}_{2\,\text{mod},s}\left(z_m,\omega_f\right)\mathbf{L}^{-}_{*,2\,z_0;z_m} \\ & \ldots \quad p^{+}_{n_x\,\text{mod},s}\left(z_m,\omega_f\right)\mathbf{L}^{-}_{*,n_x\,z_0;z_m}\Big],\end{aligned} \tag{5.8}$$

in which $\mathbf{L}^{-}_{*,j\,z_0;z_m}$ denotes the jth column of $\mathbf{L}^{-}_{z_0;z_m}$, j denotes the lateral location of a model parameter at z_m, n_x denotes the number of model parameters at z_m.

5.3. NUMERICAL EXAMPLES

In the first example, our true model is a section of the BP 2004 model, including gas clouds. To enhance the model, we introduce three point diffractors to the model, as shown in Figure 5.1. The model consists of a grid with 351 × 326 points in both horizontal and vertical directions. On top of the model, we have positioned 51 shot points with a spacing of 175 meters. For each shot, 351 receivers spaced 25 meters apart record reflection data. The source function used in this example is a Ricker wavelet with a dominant frequency of 10 Hz. Our forward modeling tool for generating the observed dataset is PWMod, i.e., only primaries, applied on the true model ("inverse-crime" scenario), with a recording time of 4.092 seconds. Employing an accurate migration velocity model (without diffractors),

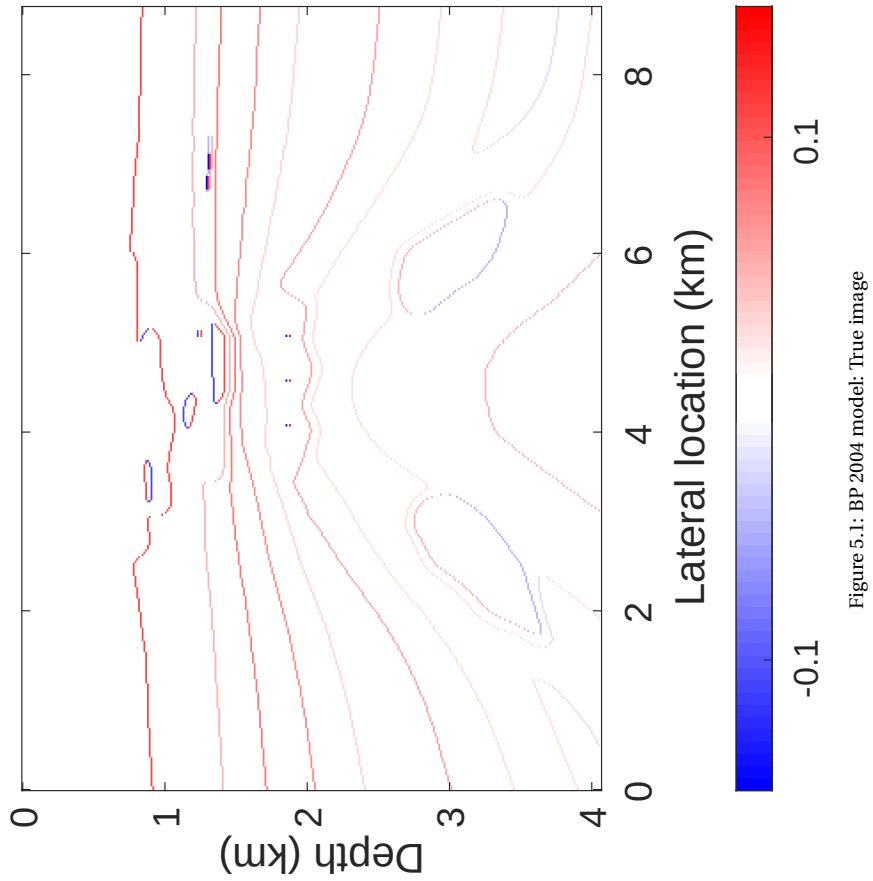


Figure 5.1: BP 2004 model: True image

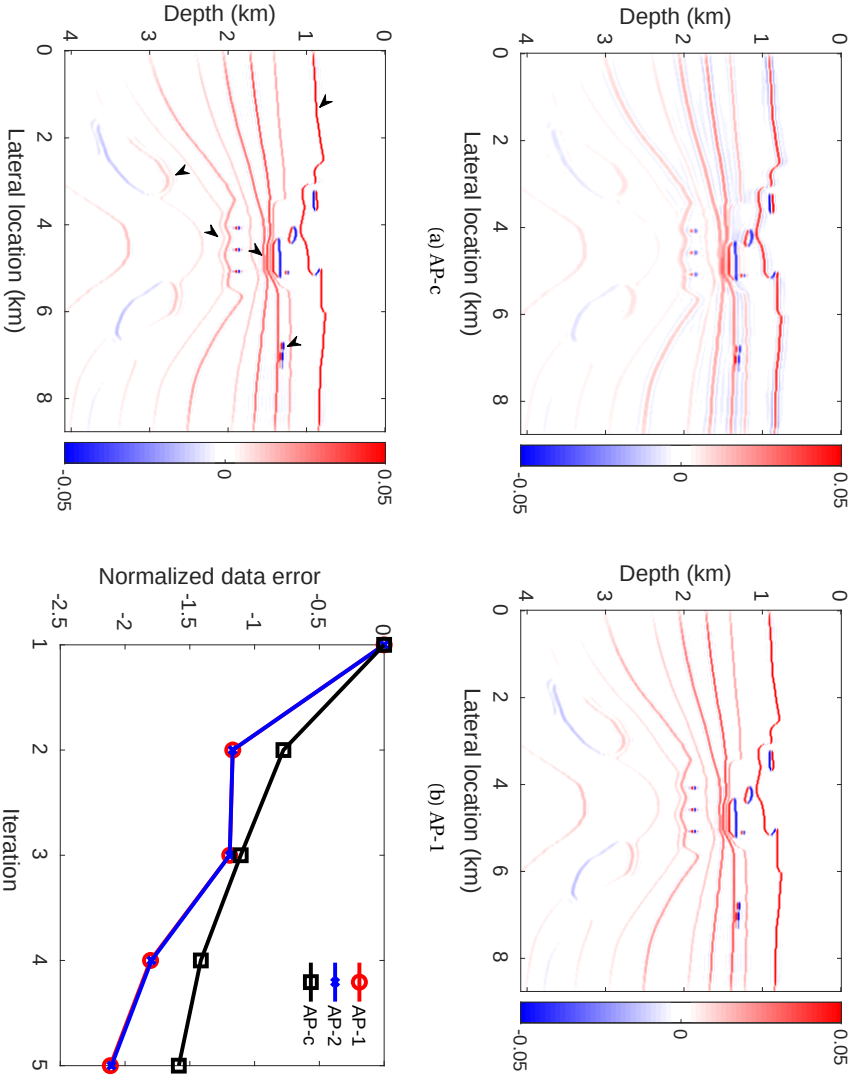


Figure 5.2: BP 2004 model example: Estimated images after 5 iterations and convergence history

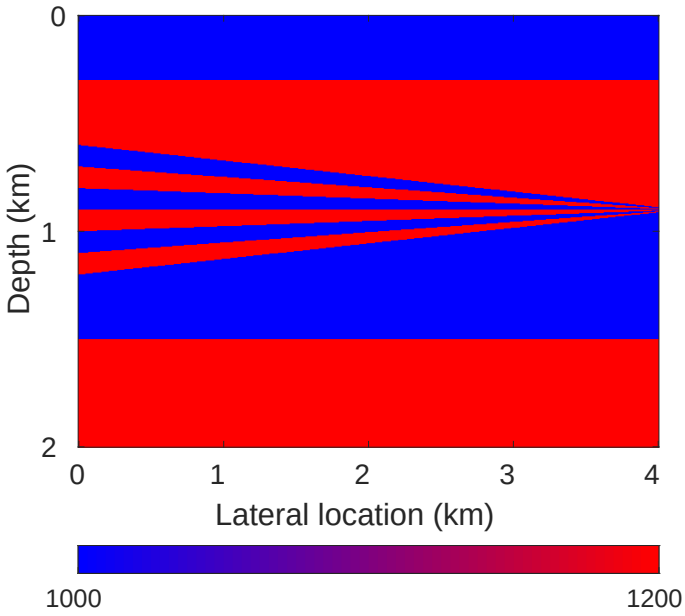


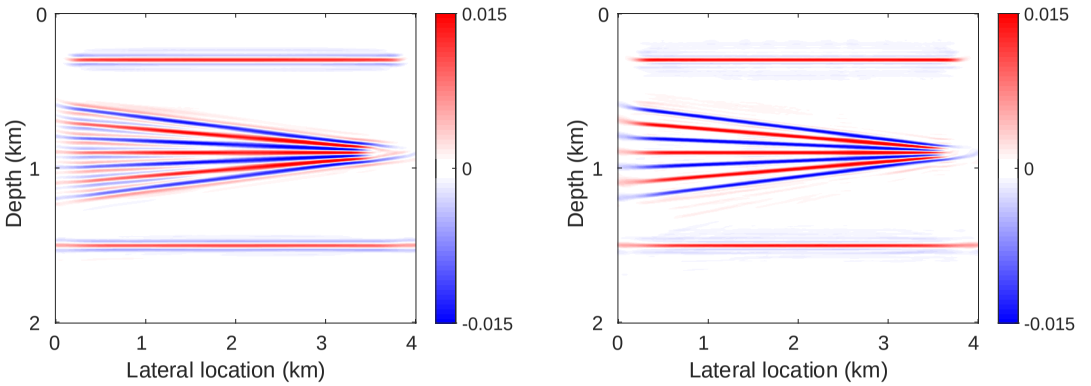
Figure 5.3: Wedge model: True density

we evaluate and compare AP-c, AP-1, and AP-2. Figures 5.2a, 5.2b, and 5.2c compare the results after 5 iterations.

AP-1 and AP-2 outperform AP-c in resolution and amplitude preservation. Figure 5.2c highlights significant disparities. Notably, interference between neighboring mono-frequency sources in AP-2 doesn't degrade the image. Figure 5.2d confirms the faster convergence of AP-1 and AP-2 over AP-c. Given the similar performance of AP-1 and AP-2, the second example exclusively compares AP-2 with AP-c, as AP-2 is cheaper than AP-1 due to requiring fewer matrix-matrix multiplications.

In the second example, we use a density-driven wedge model (Figure 5.3) with alternating reflection coefficients ($r^{\cup} = \pm 0.1$), where the true velocity model is a homogeneous velocity of 2000 m/s. Acoustic finite-difference modeling with a 1 m grid and a 15 Hz Ricker wavelet generates observed reflection data. For 2.92 s, 401 receivers, planted on top of the model, at 10 m spacing record 31 shots at 120 m spacing from 200 m to 3800 m, triggered sequentially on top of the model. For inversion, with the same source-receiver setup and the true homogeneous velocity model, FWMod with a 10 m horizontal grid and a 4 m vertical grid models the reflection data in each iteration (non-inverse-crime scenario). With this experimental setup, we evaluate and compare AP-c and AP-2. Figure 5.4 displays the estimated images after 5 and 10 iterations.

As seen in Figure 5.4, AP-2 exhibits superior resolution and significantly better-preserved amplitudes in half the iterations of AP-c. Once again, and now in a non-inverse-crime scenario, we observe that the interference between adjacent mono-frequency sources in AP-2 doesn't degrade the image. Figure 5.5 depicts the data residual of AP-c and AP-2 after 10 iterations for selected sources. It confirms that AP-2 reaches smaller



(a) AP-c: Iteration 5

(b) AP-2: Iteration 5

(c) AP-c: Iteration 10

(d) AP-2: Iteration 10

Figure 5.4: Wedge model example: Estimated images after 5 and 10 iterations

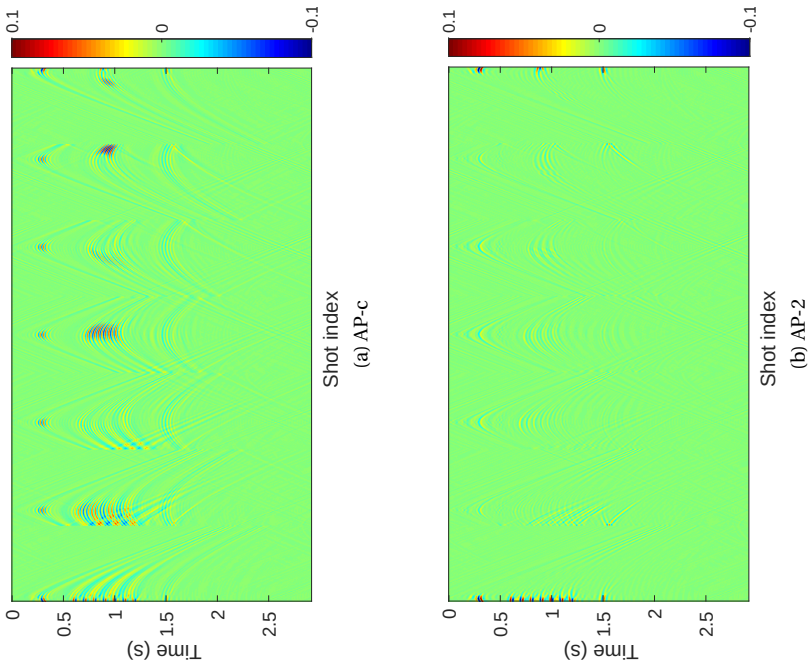


Figure 5.5: Wedge model example: Data residuals after 10 iterations for selected sources.

data errors (faster convergence in terms of iterations).

5.4. CONCLUSIONS

This study highlights the notable impact of the depth-dependent approximate Hessian in large-scale linearized least-squares migration, with AP-2 distinguishing itself in resolution power and amplitude perseverance. The example with "non-inverse-crime" scenario pinpoints the convergence benefit (in terms of iterations) of AP-2 over AP-c, achieving superior images and less data errors in half the iterations. The resistance of AP-2 to the interference between sources adds further to its appeal. These findings make clear the potential of the depth-dependent approximate Hessian to enhance seismic imaging, fostering improved subsurface characterization and interpretation.

REFERENCES

Abolhassani, S. and E. Verschuur (2022). "Fast Gauss-Newton full-wavefield migration". In: SEG/AAPG Second International Meeting for Applied Geoscience & Energy, Expanded Apbstract.

Berkhout, A. J. (2014a). "Review paper: An outlook on the future of seismic imaging, part II: Full-wavefield migration". In: *Geophysical Prospecting* 62.5, pp. 931–949.

— (2014b). "Review paper: An outlook on the future seismic imaging, part I: forward and reverse modelling". In: *Geophysical Prospecting* 62.5, pp. 911–930.

Choi, Y., D.-J. Min, and C. Shin (2008). "Frequency-domain elastic full waveform inversion using the new pseudo-Hessian matrix: Experience of elastic Marmousi-2 synthetic data". In: *Bulletin of the Seismological Society of America* 98.5, pp. 2402–2415.

Jang, U., D.-J. Min, and C. Shin (2009). "Comparison of scaling methods for waveform inversion". In: *Geophysical Prospecting* 57.1, pp. 49–59.

Jones, I. F. (2018). *Velocities, imaging, and waveform inversion*. EAGE.

Lines, L. and S. Treitel (1984). "A review of least-squares inversion and its application to geophysical problems". In: *Geophysical Prospecting* 32.2, pp. 159–186.

Lu, S. et al. (2018). "Least-squares full-wavefield migration". In: *The Leading Edge* 37.1, pp. 46–51.

Nemeth, T., C. Wu, and G. T. Schuster (1999). "Least-squares migration of incomplete reflection data". In: *Geophysics* 64.1, pp. 208–221.

Oh, J.-W. and D.-J. Min (2013). "Weighting technique using backpropagated wavefields incited by deconvolved residuals for frequency-domain elastic full waveform inversion". In: *Geophysical Journal International* 194.1, pp. 322–347.

Plessix, R. E. and W. A. Mulder (2004). "Frequency-domain finite-difference amplitude-preserving migration". In: *Geophysical Journal International* 157.3, pp. 975–987.

Shin, C., C. S. Jang, and D.-J. Min (2001). "Improved amplitude preservation for prestack depth migration by inverse scattering theory". In: *Geophysical Prospecting* 49.5, pp. 592–606.

6

JOINT EWI AND ORWI VIA THE SUBSPACE GRADIENT METHOD

"Shoot for the moon. Even if you miss, you'll land among the stars."

– NORMAN VINCENT PEALE

The successful application of deep full waveform inversion relies on the availability of broadband and 360° *seismic data. However, acquiring such data can be infeasible. Reflection waveform inversion was born to sample deep targets with short-offset seismic reflection data via model scale separation. In Chapter 1, we explored the complementary characteristics of transmission and transmission-after-reflection wavepaths. While the transmission wavepath samples mostly vertical low local model wavenumbers, the transmission-after-reflection wavepaths sample mostly horizontal low local model wavenumbers. To benefit from both and retrieve a broadband subsurface model, in this chapter, we propose to combine early-arrival waveform inversion (EWI) and one-way reflection waveform inversion (ORWI) using the subspace gradient method.*

6.1. INTRODUCTION

Almost all continuous linear inverse problems in geophysics have the general form of (Hansen, 1998)

$$\int_{\Omega} \text{input} \times \text{system}\, d\Omega = \text{output}, \tag{6.1}$$

where the output (data) is given, and the input (model/solution) is unknown. Continuous inverse problems in geophysics are almost always a subset of ill-posed problems (Kabanikhin, 2011) in the sense of Hadamard (1902). After discretization, the discrete form of the problem inherits the quality of being ill-posed, i.e., it transforms to a discrete ill-posed

problem. For instance, if the solution to the continuous inverse problem lacks stability, its discrete counterpart inherits this instability, manifesting as an ill-conditioned system operator. Alternatively, in cases where the solution to the continuous inverse problem is not unique, its discrete counterpart inherits this non-uniqueness, manifesting as a rank-deficient system operator. Indeed, non-uniqueness spans a range from ill-conditioning to singularity in discrete ill-posed problems. Note that non-uniqueness may also occur when the system operator displays rank deficiency, attributed to computational "machinery," including floating-point and finite-precision arithmetic systems used in the digital world (Hansen, 1998).

To tackle a non-unique inverse problem, we need to restrict the solution space, ensuring a unique solution. This process, known as regularization, can be implemented in various ways (Jackson, 1979), such as adding a positivity constraint, using a good reference solution, looking for a desired degree of smoothness or sparseness of the solution, or adding more geophysical observations to the problem through joint inversion methods. Different approaches to classical geophysical joint inversion are illustrated in Figure 6.1 (Stefano et al., 2011; Dell'Aversana, 2014; Moorkamp et al., 2016). While multi-physics joint inversion relies on multiple geophysical datasets, each sensitive to a different model parameter, single-physics joint inversion relies on multiple geophysical datasets, all sensitive to one model parameter, such as P-wave velocity. Their complementary nature helps mitigate non-uniqueness, thereby enhancing the model reconstruction.

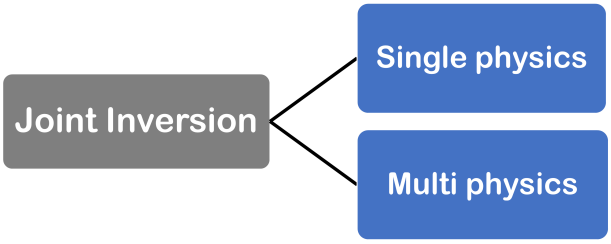


Figure 6.1: Different types of geophysical joint inversion.

In Chapter 1, we examined the sampling pattern of various wavepaths within the FWI gradient. We observed that these patterns do not intersect; rather, they show a complementary nature. Specifically, the transmission wavepath—constructed by diving and refracted or early-arrival energies—samples low vertical local model wavenumbers for P-wave velocity. On the other hand, the pair of transmission-after-reflection wavepaths—constructed by reflected energies—samples low horizontal local model wavenumbers for P-wave velocity. This observation explains the potential advantage of a joint inversion incorporating both reflected and early-arrival waveforms.

Several studies have explored the effectiveness of incorporating early-arrival waveform inversion (EWI) and reflection waveform inversion (RWI) in one inversion framework. This incorporation aims to guide the search direction, facilitating the retrieval of a wide low local-model-wavenumber spectrum for the P-wave velocity model. Notable among these studies are the works of Zhou et al. (2015), Wu and Alkhalifah (2015), and Davydenko and Verschuur (2019), where, the objective function is considered as a weighted sum of

early-arrival and reflection data. Moreover, a version of simultaneous EWI-ORWI was also proposed by Davydenko and Verschuur (2020), where the total gradient is considered as a weighted sum of the EWI ($\mathbf{g}_{\text{div}}$) and ORWI ($\mathbf{g}_{\text{refl}}$) gradients via constant scalars

$$\mathbf{g}_{\text{tot}} = (1-\theta)\,\mathbf{g}_{\text{div}} + \theta\,\mathbf{g}_{\text{refl}}, \quad \text{where } \theta \in [0,1]. \tag{6.2}$$

Davydenko and Verschuur (2020) examine their proposed gradient (Equation 6.2) using a true velocity model including a positive velocity contrast within the model that generates diving waves (Figure 6.2a). Observed and modeled data are generated using a finite-difference technique with a free-surface boundary condition. They examine the proposed gradient with three different sets of weights. At first, by setting the ORWI gradient weight to zero, exclusively activating EWI, the proposed gradient reaches a local minimum compared to the true velocity model (See Figure 6.3a). Second, by setting the EWI gradient weight to zero, exclusive activation of ORWI, the proposed gradient reaches another local minimum in comparison to the true velocity model (see Figure 6.3b). Third, by giving equal weights to the EWI and ORWI gradients, where both EWI and ORWI are active with equal importance, the proposed gradient reaches a desired minimum compared to the true velocity model (See Figure 6.3c). The results suggest that joint EWI-ORWI shows promise, yet further improvement may be achieved by automatically assigning optimized weights to the gradients.

This chapter will investigate the potential application of the subspace gradient method in single-physics joint inversion between reflected and early-arrival waveforms (joint EWI-ORWI), both sensitive to P-wave velocity, to achieve better model reconstruction.

6.2. SUBSPACE GRADIENT METHOD

In the subspace gradient method, initially introduced by Kennett et al. (1988), the idea is to limit the model perturbation to our chosen subspaces. To do so, the model perturbation is written as a linear combination of a couple of basis vectors

$$\delta\mathbf{m} = \sum_{j=1}^{n_b} \delta\theta_j \mathbf{a}^{(j)}, \tag{6.3}$$

where $\delta\theta_j$ is the coefficient for the basis vector $\mathbf{a}^{(j)}$ and n_b is the number of basis vectors. Ordering the basis vectors in a matrix called the projection matrix ($\mathbf{A}$) whose columns are the basis vectors gives the following:

$$\mathbf{A} = \begin{bmatrix} a_1^{(1)} & a_1^{(2)} & \dots & a_1^{(n_b)} \\ a_2^{(1)} & a_2^{(2)} & \vdots & a_2^{(n_b)} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n_m}^{(1)} & a_{n_m}^{(2)} & \dots & a_{n_m}^{(n_b)} \end{bmatrix}_{n_m \times n_b}. \tag{6.4}$$

In geophysics, nonlinear inverse problems are often simplified by implicitly assuming that the forward modeling operator behaves linearly in the vicinity of the current

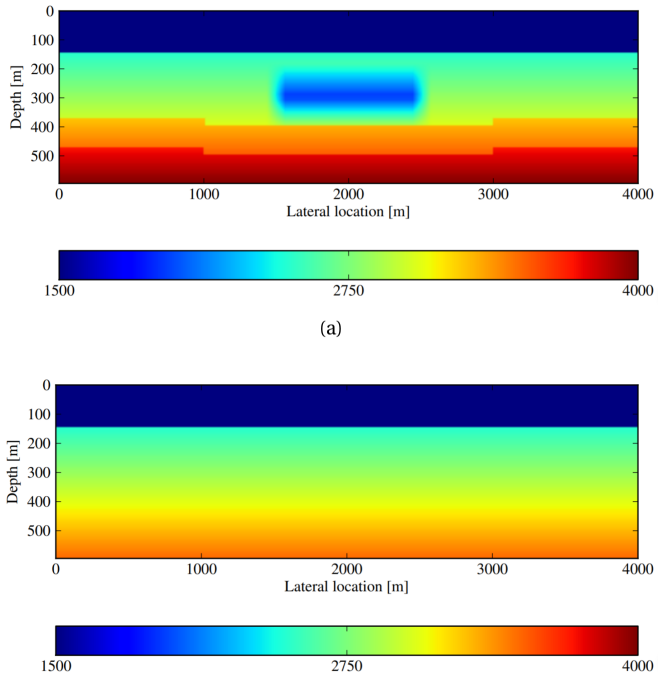


Figure 6.2: True and initial velocity model used by Davydenko and Verschuur (2020). (a) True velocity model, where a positive velocity contrast, placed within the model, produces diving waves. (b) 1D initial velocity model.

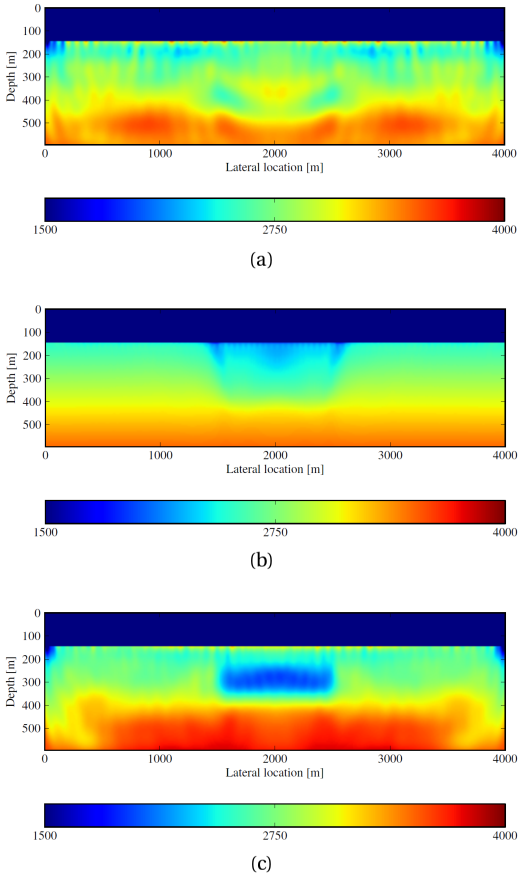


Figure 6.3: Inverted models for the true and initial velocity models shown in Figure 6.2 using the gradient proposed by Davydenko and Verschuur (2020). (a) A local minimum is reached by the proposed gradient, where the ORWI gradient weight is set to zero. (b) Another local minimum is reached by the proposed gradient, where the EWI gradient weight is set to zero. (c) A desired minimum is reached by the proposed gradient, where the EWI and ORWI gradients are set to be equal.

model parameter vector. This enables the objective function (E) to be approximated as a quadratic function (E^Q) around the current model parameter vector ($\mathbf{m}_c$) using a Taylor series expansion:

$$E^Q(\mathbf{m}_c + \delta\mathbf{m}) = E(\mathbf{m}_c) + \mathbf{g}^T\delta\mathbf{m} + \frac{1}{2}\delta\mathbf{m}^T\mathbf{H}\delta\mathbf{m}, \tag{6.5}$$

where $\mathbf{g}$ and $\mathbf{H}$ are the gradient and Hessian of the objective function, respectively, and T denotes the matrix transpose operator. Replacing the model perturbation in Equation 6.5 with Equation 6.3 gives

$$E^Q\left(\mathbf{m}_c + \sum_{j=1}^{n_b}\delta\theta_j\mathbf{a}^{(j)}\right) = E(\mathbf{m}_c) + \sum_{j=1}^{n_b}\delta\theta_j\mathbf{g}^T\mathbf{a}^{(j)} + \frac{1}{2}\sum_{j=1}^{n_b}\sum_{k=1}^{n_m}\delta\theta_j\delta\theta_k\mathbf{a}^{(k)T}\mathbf{H}\mathbf{a}^{(j)}. \tag{6.6}$$

Instead of minimizing the quadratic approximation of the objective function with respect to the current model, if the objective function is minimized with respect to the coefficients of the basis vectors (first-order optimality condition)

$$\frac{\partial E^Q}{\partial\delta\theta_j} = 0 \quad\rightarrow\quad \delta\boldsymbol{\theta} = -\left(\mathbf{A}^T\mathbf{H}\mathbf{A}\right)^{-1}\mathbf{A}^T\mathbf{g}, \tag{6.7}$$

where $\delta\boldsymbol{\theta}$ is called the coefficient perturbation vector.

To back project into the initial model domain, it suffices to apply the projection matrix on the coefficient perturbation vector (Equation 6.3)

$$\delta\mathbf{m} = -\mathbf{A}\left(\mathbf{A}^T\mathbf{H}\mathbf{A}\right)^{-1}\mathbf{A}^T\mathbf{g}. \tag{6.8}$$

6.3. JOINT EWI AND ORWI USING THE SUBSPACE GRADIENT METHOD

To bring the simultaneous EWI-ORWI in the framework of the subspace approach, the objective function is defined as a function of two basis coefficients, one corresponding to the reflection data and one corresponding to the refraction/diving wave data

$$E\left(\delta\theta_{\text{div}}, \delta\theta_{\text{refl}}\right) = \frac{1}{2}\left\|\mathbf{d}_{\text{div}}^{\text{obs}} - \mathbf{d}_{\text{div}}^{\text{mod}}\right\|_2^2 + \frac{1}{2}\left\|\mathbf{d}_{\text{refl}}^{\text{obs}} - \mathbf{d}_{\text{refl}}^{\text{mod}}\right\|_2^2, \tag{6.9}$$

where $\mathbf{d}_{\text{refl}}^{\text{obs}}$ and $\mathbf{d}_{\text{refl}}^{\text{mod}}$ denote the observed and modeled reflection data vectors, respectively, and $\mathbf{d}_{\text{div}}^{\text{obs}}$ and $\mathbf{d}_{\text{div}}^{\text{mod}}$ denote the observed and modeled refraction/diving wave data vectors, respectively.

Following the objective function introduced in Equation 6.9, the total gradient vector reads

$$[\mathbf{g}]_{n_m\times 1} = \mathbf{g}_{\text{div}} + \mathbf{g}_{\text{refl}}, \tag{6.10}$$

where n_m represents the number of model parameters for inversion.

After defining the projection matrix as

$$\mathbf{A} = \left[\mathbf{g}_{\text{div}} \quad \mathbf{g}_{\text{refl}}\right]_{n_m\times 2}, \tag{6.11}$$

the coefficient perturbation vector ($\delta\boldsymbol{\theta}$) is given by

$$\delta\boldsymbol{\theta} = \begin{bmatrix} \delta\theta_{\text{div}} \\ \delta\theta_{\text{refl}} \end{bmatrix} = - \begin{bmatrix} \mathbf{g}_{\text{div}}^T \mathbf{H}_{\text{div div}} \mathbf{g}_{\text{div}} & \mathbf{g}_{\text{div}}^T \mathbf{H}_{\text{div refl}} \mathbf{g}_{\text{refl}} \\ \mathbf{g}_{\text{refl}}^T \mathbf{H}_{\text{refl div}} \mathbf{g}_{\text{div}} & \mathbf{g}_{\text{refl}}^T \mathbf{H}_{\text{refl refl}} \mathbf{g}_{\text{refl}} \end{bmatrix}_{2\times 2}^{-1} \mathbf{A}^T \mathbf{g}. \tag{6.12}$$

Replacing the Hessian matrices ($\mathbf{H}$) in Equation 6.12 with identity matrices yields

$$\delta\boldsymbol{\theta} \approx - \begin{bmatrix} \mathbf{g}_{\text{div}}^T \mathbf{g}_{\text{div}} & \mathbf{g}_{\text{div}}^T \mathbf{g}_{\text{refl}} \\ \mathbf{g}_{\text{ref}}^T \mathbf{g}_{\text{div}} & \mathbf{g}_{\text{refl}}^T \mathbf{g}_{\text{refl}} \end{bmatrix}_{2\times 2}^{-1} \begin{bmatrix} \mathbf{g}_{\text{div}}^T \mathbf{g} \\ \mathbf{g}_{\text{refl}}^T \mathbf{g} \end{bmatrix}_{2\times 1}, \tag{6.13}$$

that could be solved iteratively through conjugate gradient (CG) iterations.

Referring to Equation 6.13, one can infer that the quest for a search direction in the subspace simultaneous EWI-ORWI is conducted within a plane defined by the introduced basis vectors. In this process, at each iteration, two suitable weights of the basis vectors are combined, aiming to minimize the unified objective function. In the subspace approach, after acquiring the necessary basis vectors, a Gram-Schmidt orthogonalization process may be used to avoid inter-dependency (Rawlinson and Sambridge, 2003). The basis vectors can also be extended to account for potential leakage between gradient components (Geng et al., 2020).

The prospect of this methodology seems bright, as its goal is to reconstruct a broadband subsurface model through a strategic combination of the strengths of EWI and ORWI technologies, thereby effectively mitigating the individual limitations of each (highlighted in Chapter 1). The innovative aspect of this approach lies in its ability to pinpoint specific subsurface areas where both technologies can collaborate synergistically, maximizing their combined potential for greater performance.

6.4. CONCLUSION

As the sampling of high-reflective wavenumbers nonlinearly depends on sampling low-propagative wavenumbers, it is essential to estimate them accurately first. The simultaneous EWI-ORWI approach with equal weights for both EWI and ORWI gradients is effective, but it could be improved significantly by the subspace approach, where two proper weights, rather than two constant weights, are estimated for the basis vectors to estimate a minimizer at each iteration.

REFERENCES

Davydenko, M. and D. J. Verschuur (2019). "Using the full wavefield both in FWI and wavefield tomography". In: SEG Annual International Meeting, Expanded Abstracts.

Davydenko, M. and E. Verschuur (2020). "Combining JMI and FWI methods". In: *DELPHI The Reservoir Characterization and Monitoring Annual Report*, 10–1 – 10–9.

Dell'Aversana, P. (2014). *Integrated geophysical models: Combinining rock physics with seismic, electromagnetic and gravity data*. EAGE.

Geng, Y., K. A. Innanen, and W. Pan (2020). "Subspace method for multi-parameter FWI". In: *CREWES Research Report*.

Hadamard, J. (1902). “Sur les problèmes aux dérivés partielles et leur signification physique”. In: *Princeton University Bulletin* 13, pp. 49–52.

Hansen, P. C. (1998). *Rank-deficient and discrete ill-posed problems: numerical aspects of linear inversion*. SIAM.

Jackson, D. D. (1979). “The use of a priori data to resolve non-uniqueness in linear inversion”. In: *Geophysical Journal International* 57.1, pp. 137–157.

Kabanikhin, S. I. (2011). *Inverse and ill-posed problems: theory and applications*. Walter De Gruyter.

Kennett, B., M. Sambridge, and P. Williamson (1988). “Subspace methods for large inverse problems with multiple parameter classes”. In: *Geophysical Journal International* 94.2, pp. 237–247.

Moorkamp, M. et al. (2016). *Integrated imaging of the earth: Theory and applications*. John Wiley & Sons.

Rawlinson, N. and M. Sambridge (2003). “Seismic travel time tomograpgy of the crust and lithosphere”. In: *Advances in Geophysics* 46, pp. 81–197.

Stefano, M. D. et al. (2011). “Multiple-domain, simultaneous joint inversion of geophysical data with application to subsalt imaging”. In: *Geophysics* 76.3, pp. 69–80.

Wu, Z. and T. Alkhalifah (2015). “Simultaneous inversion of the background velocity and the perturbation in full-waveform inversion”. In: *Geophysics* 80.6, R317–R329.

Zhou, W. et al. (2015). “Full waveform inversion of diving waves for velocity model building with impedance inversion based on scale separation”. In: *Geophysical Journal International* 202.3, pp. 1535–1554.

7

CONCLUSIONS AND RECOMMENDATIONS

"Every exit is an entry somewhere else."

– TOM STOPPARD

7.1. CONCLUSIONS

As the main conclusion, it can be stated that one-way reflection waveform inversion (ORWI) or any waveform inversion scheme embodying transmission-after-reflection wavepaths, demands considerable care in order to achieve an optimal background velocity model of the subsurface. Despite its conceptual appeal, ORWI faces several limitations, including low-resolution images with unpreserved amplitudes, low-resolution tomographic wavepaths, imprint of inconsistent reflectivity and velocity models, and cycle skipping in long-offset tomography data. The aim of Chapters 2, 3, 4, and 5 was to address these limitations that hinder the effectiveness of ORWI.

Chapter 2 introduced a cost-effective preconditioned least-squares wave-equation migration (PLS-WEM) algorithm for high-resolution seismic imaging. By leveraging a one-way wavefield extrapolation technique, PLS-WEM decomposes and reduces the linear Hessian operator (also known as the approximate or Gauss-Newton Hessian operator) for the entire medium into multiple smaller sub-operators, each relevant to a specific virtual depth level within the medium. Specifically, PLS-WEM extracts and inverts the diagonal blocks of the mono-frequency Gauss-Newton Hessian operator to precondition their respective mono-frequency gradients. Chapter 2 demonstrated how the proposed depth-dependent preconditioners effectively deconvolve the seismic source signature from the reflectors while preserving the reflector amplitudes.

Chapter 2 also discussed how PLS-WEM could improve tomography in ORWI. This is because ORWI relies on migrated reflectors for its tomographic updates, so a higher-resolution, amplitude-preserved image is expected to yield more accurate and consistent tomograms. Chapter 3 confirmed this discussion. In the ORWI context, Chapter 3 demonstrated that generating strong, in-phase transmission-after-reflection wavepaths relies on mapping reflectors with the truest possible amplitude and highest possible resolution, considering inaccuracies in the background velocity model. This chapter addressed the issue of inconsistent reflectivity and velocity models by muting short-offset waveforms in the residual data gathers for tomography. By muting these short-offset residuals, this chapter adopted an extended migration offset (MEMO) instead of short offsets, thereby mitigating limitations in reflection tomography related to poor signal-to-noise ratio and inadequate illumination of the reflectors. Additionally, this chapter introduced a data-selection algorithm to exclude the damaging effect of cycle-skipped long-offset data from the tomographic gradient in ORWI. With all the introduced additions to ORWI, This chapter finally verified the successful application of the proposed ORWI. It concluded that, while in scenarios with non-complex layering ("flattish layers") the proposed ORWI reliably estimates final tomograms that closely mirror reality across all depth levels, capturing details from shallow to deep, it may struggle to accurately estimate deep structures in scenarios with complex layering ("extreme faulting"), despite the additional aids.

Chapter 4 transformed ORWI into a fully high-resolution yet computationally feasible version (HR-ORWI), aiming to enhance the non-optimal background velocity updates attributed to the resolution issue in both migration and tomography loops. This was established after developing the required mathematical groundwork to incorporate the depth-dependent preconditioning concept, which was introduced in Chapter 2 for migration, also for background velocity reconstruction. Chapter 4 concluded that HR-ORWI equipped with all the additional aids proposed in Chapter 3, proves itself as a promising tool for accurate reflection tomography from shallow to deep even in the case of extreme faulting (structural complexity) in the medium.

Chapter 5 investigated the effectiveness of depth-dependent gradient preconditioning in the context of least-squares wave-equation migration through different update equations: conventional (AP-c), free of interference between neighboring mono-frequency sources (AP-1), and including interference between neighboring mono-frequency sources (AP-2). Chapter 5 revealed the appeal of AP-2 through two numerical examples. It showed that while AP-1 and AP-2 converge faster in iterations and exhibit superior resolution compared to AP-c, AP-2 provides equal image quality as AP-1 at a lower computational time cost due to requiring fewer matrix-matrix multiplications.

Both vertical and horizontal model wavenumbers are essential to construct a broadband subsurface model. While the early-arrival waveform inversion (EWI) technique mostly samples vertical low local model wavenumbers, ORWI mostly samples horizontal low local model wavenumbers. As a result, Chapter 6 presented a math foundation using the subspace gradient method to combine EWI and ORWI for optimal outcomes, to be validated in future studies.

7.2. RECOMMENDATIONS

Further research is recommended to gain a deeper understanding of the following points.

7.2.1. GRADIENT PRECONDITIONING AND MULTIPLE SCATTERING

While better-focused reflectors were expected to improve the modeling of multiple scattering and lead to reduced cross-talk noise, the full-wavefield migration experiment in Chapter 2 only yielded minor cross-talk noise reductions. We attributed the underperformance to the following factors:

- Preconditioning could indeed act as a double-edged sword in the case of full-wavefield migration. It enhances resolution for both real and ghost reflectors in modeling, yet this enhancement may result in conflicts between the two during relevant data fitting, potentially leading to a local minimum.
- The resonant wedge experiment is highly nonlinear.
- FWMod has convergence problems with spatially inhomogeneous media (McMaken, 1986).

It is recommended to conduct additional investigations to quantify the contribution of each factor and determine the main driver of underperformance.

7.2.2. MODEL-BASED REGULARIZATION IN ORWI

In this study, we chose not to include any model-based regularization in our tomography misfit function. However, evidence suggests that incorporating such regularization, particularly a structure-oriented approach guided by the reflectivity model, could steer the tomographic updates positively (Masaya and Verschuur, 2018; Yao et al., 2019; Provenzano et al., 2023). Therefore, exploring and demonstrating this potential benefit is recommended.

7.2.3. IMPRINT OF INCONSISTENT REFLECTIVITY AND VELOCITY MODELS IN ORWI

Our proposed solution for handling the imprint of inconsistent reflectivity and velocity models in ORWI (due to the neglect of reflectivity-velocity coupling) averagely attenuates the imprint but may not be optimal. As a potential improvement, training a deep learning model is recommended for effectively learning this task. Additionally, three other approaches have been proposed in the literature to address the imprint of inconsistent reflectivity and velocity models in RWI: simultaneous inversion of reflectivity and velocity (e.g., Wu and Alkhalifah, 2015), ensuring arrival-time consistency in the data by taking into account the coupling effect in the RWI gradient (Valensi and Baina, 2021), and pseudo-time domain RWI (e.g., Qu et al., 2020; Provenzano et al., 2023). Given the availability of pseudo-time ORWI within the Delphi research group as an in-house technology (Qu et al., 2020), it is worth investigating how adopting this approach could enhance the effectiveness and overall convergence of HR-ORWI.

7.2.4. INVERSE-CRIME ORWI VERSUS ORWI WITH FD DATA

A comparative study on various synthetic examples, ranging from “flattish” to “extreme faulting” layering, “weak” to “strong” multiple scattering, and “weak” to “strong” AVO effect between an inverse-crime ORWI and ORWI with finite-difference observed data

is recommended. This study will offer valuable insights into the discrepancies of ORWI when explaining finite-difference observed data. Furthermore, through validating ORWI with field data and comparing the outcomes with insights from synthetic examples based on finite-difference observed data, we can further deepen our understanding of its performance. This will not only aid in refining the ORWI technology but also serve as a blueprint for future research and development efforts aimed at optimizing ORWI and PWMod/FWMod for improved subsurface imaging and exploration.

7.2.5. APPLICATION TO FIELD DATA

Studying the application of the proposed technologies to field data is crucial for several reasons. First, HR-ORWI or PLS-WEM cannot directly model amplitude variation with offset (AVO) due to their reliance on angle-independent reflectivities. Pre-processing steps, as recommended by Qu and Verschuur (2021), are essential to mitigate the AVO effect before applying PLS-WEM or HR-ORWI. Second, ensuring accurate estimation of the source signature signal in real-world data applications is crucial for maintaining high imaging resolution and tomographic accuracy when using depth-dependent preconditioners. Third, another challenge with field data is the potential presence of anisotropy and the Q-effect. Accounting for anisotropy and the Q-effect in PLS-WEM and HR-ORWI, whether forward or inverse-wise, enhances our capability to explain such complexities in real-world seismic data (Alshuhail and Verschuur, 2019; Safari and Verschuur, 2023).

7.2.6. 3D IMPLEMENTATION OF HR-ORWI

We believe that expanding HR-ORWI to a 3D scenario, where wavefields are observed along both inline and crossline directions, is feasible. In this case, the Hessian approximation matrix for the entire medium takes the form of a 3D block-diagonal matrix. Each block, relevant to a specific virtual data plane, contains non-zero elements centered around main diagonals in three directions, resulting in a 3D band-diagonal matrix for each virtual data plane. While it is acknowledged that working in 3D requires significantly more resources in terms of computing power compared to 2D, advancements in computing technology make it increasingly viable to tackle such computational challenges. Moreover, with the growing demand for higher-resolution subsurface imaging, the benefits of transitioning to a 3D framework outweigh the associated resource requirements.

7.2.7. COMBINING EWI AND ORWI FOR OPTIMAL OUTCOMES

According to Chapter 6, future research is advised to validate the joint EWI-ORWI framework proposed through the subspace gradient method. This approach aims to develop a broadband subsurface model with proper gradient weighting, where EWI targets vertical model wavenumbers and ORWI targets horizontal ones.

REFERENCES

Alshuhail, A. A. and D. J. Verschuur (2019). “Robust estimation of vertical symmetry axis models via joint migration inversion: Including multiples in anisotropic parameter estimation”. In: *Geophysics* 84.1, pp. C57–C74.

Masaya, S. and D. Verschuur (2018). "Iterative reflectivity-constrained velocity estimation for seismic imaging". In: *Geophysical Journal International* 214.1, pp. 1–13.

McMaken, H. (1986). "On the convergence of the Bremmer series for the Helmholtz equation in 2-D". In: *Wave Motion* 8.3, pp. 277–283.

Provenzano, G., R. Brossier, and L. Métivier (2023). "Robust and efficient waveform-based velocity-model building by optimal transport in the pseudotime domain: Methodology". In: *Geophysics* 88.2, U49–U70.

Qu, S. and D. Verschuur (2021). "An effective scheme of pseudo-time Joint migration inversion with an AVO mitigating workflow". In: 82th EAGE Annual Conference & Exhibition, Extended Abstracts.

Qu, S., Y. van den Brule, and D. J. Verschuur (2020). "A stable scheme of joint migration inversion in the pseudo-time domain". In: EAGE Annual Conference & Exhibition Online, Extended Abstracts.

Safari, M. and J. Verschuur (2023). "Joint Migration Inversion including Q effects: towards Q estimation". In: 84th EAGE Annual Conference & Exhibition, Extended Abstracts.

Valensi, R. and R. Baina (2021). "A time consistent waveform inversion (TWIN) method". In: 82nd EAGE Annual Conference & Exhibition, Extended Abstracts.

Wu, Z. and T. Alkhalifah (2015). "Simultaneous inversion of the background velocity and the perturbation in full-waveform inversion". In: *Geophysics* 80.6, R317–R329.

Yao, G., N. V. da Silva, and D. Wu (2019). "Reflection-waveform inversion regularized with structure-oriented smoothing shaping". In: *Pure and Applied Geophysics* 176.12, pp. 5315–5335.

ACKNOWLEDGEMENTS

This dissertation represents a significant chapter in my life, shaped by the support and guidance of many incredible people. I am thankful to everyone who stood by me, offered wisdom, and believed in my efforts. Your presence made this journey both meaningful and memorable.

I would like to extend my heartfelt gratitude to my promoter, Dr. **Eric Verschuur**, whose presence transcends that of a mere mentor; an angel of unparalleled wisdom and kindness, not from this Earth, and incomparable to anyone I have ever had the privilege to meet. Eric, your unique perspective, exemplary commitment, and unwavering support have been a great source of inspiration and encouragement throughout this journey. You are a beacon of wisdom and humility—someone who exemplifies the idea that "the higher we are placed, the more humbly we should walk." You taught me to embrace curiosity, to learn from every encounter—be it from supporters or adversaries—and to approach challenges with honesty, ethics, and courage. I openly confess that your guidance and demeanor have had a profound impact on me and have shaped me into a better scientist, researcher, and human.

I am unconditionally and forever thankful to the love of my life and the soul of my every endeavor, my beautiful wife, **Sanaz**. Your unwavering support, boundless patience, and unconditional love have been my rock through every storm and my light in the darkest moments. You have been my anchor, grounding me when challenges felt overwhelming, and my greatest source of joy, constantly reminding me of the beauty in every step of this journey. Your quiet sacrifices, your unshakable faith in me, and your ability to rekindle resilience in me, even when I doubted, have made this accomplishment as much yours as it is mine. For your true love, your grace, and everything you have given me, I forever cherish and admire you and am proud beyond words to share this life with you for a lifetime and a day.

I would like to sincerely thank Prof. **Femke Vossepoel** for her invaluable mentorship throughout our collaboration. Femke, your absolute confidence in my technical skills has had a lasting influence on my professional growth, for which I will always be grateful. You put in a lot of effort, took my hand and navigated me through the infinite ocean of research, teaching me the courage to dive into challenges without overthinking. Through your mentorship, I took in the subtle art of communication and the true meaning of professionalism in academia. Both Sanaz and I cherish our friendship with you.

I sincerely appreciate Prof. **Evert Slob** for his exceptional mentorship throughout our collaboration. Evert, beyond being an outstanding PI, you have been a true friend—always ready to lend a helping hand to anyone in need. Your support has had a significant impact on my career, helping me align with a more fitting track, for which I will be forever grateful.

Despite your immense intellect, you have the rare gift of walking alongside each individual at their own pace—what I felt fortunate to experience with you. Through your mentorship, I learned to think critically, communicate with simplicity and directness, and confidently advocate for my rights. Your impact on me will stay with me forever.

I would like to express my sincere gratitude to Prof. **Guus Berkhout**, Dr. **Xander Staal**, and Dr. **Mikhail Davydenko** for their invaluable contributions to the development of our in-house imaging and tomography technologies. Their knowledge, dedication, and commitment to excellence have been instrumental in advancing our work and pushing the boundaries of what we can achieve within the Delphi Research Consortium.

I would like to extend my sincere appreciation to Prof. **Sjoerd Stallinga** for his support throughout my PhD journey. I am also thankful to Prof. **Bernd Rieger**, Dr. **Jeroen Kalkman**, and Dr. **Miriam Menzel** for generously sharing their knowledge, insights, and feedback during our weekly research seminars.

I sincerely thank the sponsors of the **Delphi Research Consortium** for their insightful feedback, engaging discussions, and support during our biannual meetings. Your interest in the technologies we developed has been motivating and appreciated.

I am also grateful to the Delphi supervisory team and Delphi advisory board, especially Dr. **Gerrit van Dijk**, Dr. **Deyan Draganov**, Dr. **Koen van Dongen**, Dr. **Jan Thorbecke**, Dr. **Roald van Borselen** (Shearwater), and Dr. **Walter Rietveld** (BP). Dear **Walter** and **Roald**, your inspiring and motivational speeches will forever be ingrained in my memory. I will make every effort to pass them on to future generations.

I would like to sincerely thank Dr. **Maria Sovago** for the wonderful energy and uplifting vibe she always shared with me, her willingness to offer help whenever I reached out, and her practical tips in both research and personal life.

The support staff in the Imaging Physics department have been quite helpful throughout this journey. A special thank you to **Annelies** and **Angela** for their kindness and assistance. Without you, this journey would have been far more challenging. I would also like to extend my gratitude to **Henry** and **Ronald** for their help with technical cluster-related issues. You both have always been so friendly to me, brightening every moment with your warm smiles. I also extend my heartfelt thanks to **Gerrie** for her tireless efforts in ensuring that the Delphi meetings ran smoothly for everyone involved.

The support staff in the Geoscience and Engineering department is composed of incredibly kind, helpful, and awesome individuals. Dear **Raf**, **Lydia**, **Marlijn**, **Ralf**, **Barbara**, **Felicia**, and **Chenoa**, I appreciate your friendly demeanor and willingness to assist in any condition. Your positive energy has made working here a joy.

I wish to extend my heartfelt thanks to all my teachers prior to my PhD, whose guidance and knowledge shaped my academic success. I am especially grateful to my master's program supervisors at the University of Tehran, Dr. **Navid Amini** and Prof. **Hamidreza Siahkoohi**, for their mentorship and for laying a strong scientific foundation that has been central to my growth and achievements in research.

I would like to extend my warm appreciation to the following brilliant minds in geophysics for their consistent support: Prof. **Tariq Alkhalifah** (Kaust), Dr. **Ivan Vasconcelos** (Shearwater), Dr. **Stefan Carpentier** (TNO), Dr. **Matteo Ravasi** (Shearwater), Dr. **Nihed Alloucheh** (Shearwater), Dr. **Gabrio Rizzuti** (Shearwater), and Dr. **Milad Farshad** (CGG), and Dr. **Carlos André Martins de Assis** (CGG). Your support, from EAGE and Delphi events to social media interactions have been a great source of motivation. Your belief in my potential and your generosity in promoting my contributions have expanded my growth, visibility, and confidence as a geophysicist.

I owe a great deal of heartfelt thanks to my dear uncle, Prof. **Bahman Abolhassani**, whose belief in me has been a great source of inspiration. Your mentorship, coupled with the invaluable opportunity to work alongside you in your lab in Iran before my move to the Netherlands, were transformative experiences. Those moments shaped my growth and set me on the path to achieving my dreams, for which I will always be thankful. While working in the lab, I had the privilege of learning from Dr. **Hadi Jamali-Rad** (Shell), whose outstanding academic journey inspires me to this day. Your advice and encouragement have been central in setting me on the path to my PhD at TU Delft. Thank you so much, Hadi!

I would like to give a big thank you to my closest friends and colleagues for their warm presence, endless laughter, advice, and the countless unforgettable moments we shared. Without you, this journey would have been far less colorful and much more serious. **Aydin**, thank you for always lending a listening ear, being a true loyal friend, creating memorable moments, collaborative thinking, and sharing your unique perspective on life with me. **Azin**, I appreciate your genuine friendship, your generosity in always seeing the best in me and reminding me of my technical qualities, and your willingness to help me in any way you can; you are awesome! **Ali**, **Dong**, and **Leo** (Delphi 0.0 in alphabetic order), my dear friends and officemates at the Imaging Physics department, we shared countless unforgettable moments—playing board games, dining out, sharing secrets, enjoying lunch and coffee breaks, and chatting about everything from fun to serious topics. I truly cherished your presence and hope to keep this true friendship forever. **Dieter**, I will never forget your kindness, your positive attitude, your warm presence, and your true friendship. As my open-source enthusiast friend, I learned so much from you. You and **Liliana** left the Netherlands, and this left all of us with a sense of missing your positive vibes here. **Arash**, **Farid**, and **Masoud**, we shared countless memorable moments, from playing Mafia and volleyball to enjoying Persian dinner together. Your kindness and willingness to help in various situations made a significant difference during challenging times. Even though we don't see each other as often now, our bond remains strong, and our friendship continues to bring light into my life. I want to sincerely thank you for everything. **Milad**, getting to know you has been a great pleasure. The positive energy and vibe around you is exemplary. Remember, if Aydin ever swims against the current (which he often does), you can count on me to join you in giving him a hard time! **Amin**, **Fardad**, **Mohammad**, and **Alireza**, thank you so much for your unconditional friendship, kindness, and support whenever I turned to you. You have been central in my mind and your presence made this journey feel less challenging. **Malihe** and **Kiarash** (Pashna), I'm genuinely grateful to have such wonderful officemates. Our shared laughter,

endless Persian chats, and cozy coffee breaks have brought so much brightness to my days. Genuine friendships like ours are rare, and I feel fortunate to have you both by my side. **Aparajita**, I genuinely appreciate the positive comments you have always shared about my work with generosity. Your encouragement and support has been a great source of inspiration for me. **Shan** and **Ulaş**, I cannot thank you enough for always being so kind to me and generously supporting me on my wild adventure toward the banking sector. **Ines**, **Joost**, **Anyu**, and **Nick**, although we were officemates for only a short time, you were indeed wonderful companions, and I thank you for your presence.

Special thanks go to the rest of my Delphi colleagues, with whom I've shared so many happy moments, either daily or while traveling to conferences. Your friendship and companionship have made this journey truly unforgettable. I appreciate the time spent with **Sixue**, **Billy**, **Camile**, **Sverre**, **Andreas**, **Tiexing**, **Boris**, **Shotaro**, **Matteo**, **Ayush**, **Andrea**, **Runhai**, **Siddarth**, **Apostolos**, **Junhai**, **Abdulrahman**, **Bouchaib**, **Hannes**, **Naveed**, **Hussain**, **Jan-Willem**, **Shogo**, **Özkan**, and **Sijmen**. Thank you so much!

I would like to thank my dear colleagues and friends in our data assimilation group, including **Hamed**, **Samantha**, **Mohsan**, **Arundhuti**, **Gabriel**, **Jenny**, **Celine**, and **Liang**. It has been a pleasure learning from all of you.

I would like to extend a warm appreciation to the following friends and colleagues from the Geoscience and Engineering department. Your friendship and kindness made this journey enjoyable and memorable: **Myrna**, **Sian** (Jones), **Entela**, **David**, **Ali** (Golchin), **Maria**, **Mahmoud**, **Selvican**, **Chris**, **Florencia**, **Joeri**, **Kiarash** (Mansour Pour), **Bahareh**, **Mohsen** (Mirzaie Yegane), **Faezeh**, **Sadegh**, **Kai**, **Marc** (Friebel), **Qin**, **Ana**, **Tianqi**, **Tina**, **Mahtab**, **Parvin**, **Eli**, **Mahnaz**, **Sepideh** (Harajchi), **Moein**, **Eddy**, **Mosab**, **Jingming**, **Zhenja**, **Shihao**, **Aleks**, **Ilshat**, **Max**, **Giovanni**, **Herminio**, **Sajjad**, **Nazife**, **Emilio**, **Aukje**, **Johno**, **Ilshat**, **Jasper**, **Kishan**, **Sara**, **Marat**, **Menno**, and **Iban**. Thank you all!

I owe a great deal of my success to the companionship of my friends and colleagues in the department of Imaging Physics, including **Qingru**, **Wenxiu**, **Peter**, **Isabel**, **Sobhan**, **Vahid**, **Alim**, **Konstantine**, **Spozmai**, and **Loes**. Your friendship and presence have meant so much to me throughout this challenging journey. Thank you for making this experience memorable.

A huge thank you goes to my friends and colleagues in **D.O.G.S.**, **EAGE Local chapter Netherlands**, and **MV** for their efforts in organizing engaging talks, workshops, and events.

I will carry the happy memories I've made with **Nima**, **Sahar**, **Amir**, **Neda**, **Afshin**, **Negar**, **Francis**, **Behnaz**, and **Sepideh** with me forever. I'm so grateful that you have been a part of my life.

My incredible parents, **Parvin and Farhad**, you are the heart of my story, the compass that has guided me through every twist and turn of life. You have taught me, from childhood, that honesty, humanity, and ethics are not merely values to admire but principles to live by. These timeless lessons became the foundation upon which I approached every stage of my life, guiding me from the wonder of childhood curiosity to the perseverance

required for the demanding journey of earning my PhD. Your sacrifices, selflessness, and love have been the silent forces propelling me forward in life, and for that, I will remain forever grateful. You didn't just teach me to aim high; you showed me how to climb with integrity, kindness, and purpose. I owe every step of my success to you, and words will never suffice to express the depth of my gratitude and love. Being your child is the greatest blessing of my life.

My amazing parents-in-law, **Simin and Mahmoud**, from the very beginning, you welcomed me not just into your family, but into your hearts, with a warmth and love that I will always treasure. Your kindness and love has been a constant source of comfort, wrapping me in a sense of belonging that words can scarcely capture. In your presence, I have felt seen, valued, and embraced, and it means the world to me. Being your child is the most precious gift I've ever received.

My beloved siblings and siblings-in-law, **Parisa, Sara, Mohammad Saeed, and Saeed,** You are not just my siblings; you are my heroes, my missing pieces, and my lifelong gifts. I am endlessly grateful for your presence in my life and carry your love with me in all that I do.

CURRICULUM VITÆ

Siamak ABOLHASSANI

EDUCATION

2020–2024 **PhD in Applied Geophysics** – *Expected Graduation: January 23, 2025*

Delft University of Technology, Netherlands

Dissertation: One-Way Reflection Waveform Inversion

Promoter: Dr. Dirk Jacob (Eric) Verschuur

2013–2016 **MSc in Applied Geophysics**

University of Tehran, Iran

Thesis: Mono and Multi-Parameter Full Waveform Inversion

Supervisor: Dr. Navid Amini (Shearwater GeoServices)

EXPERIENCE

2024–Present **Postdoctoral Researcher | Full-time**

Delft University of Technology, Netherlands

Project: Software development for seismic imaging and tomography using C

2020–2024 **PhD Researcher | Full-time**

Delft University of Technology, Netherlands

Project: Improving Delphi in-house migration and tomography technologies

2018–2020 **Researcher | Full-time**

Delft University of Technology, Netherlands

Project: Reducing monitoring uncertainties using multi-physics data inversion

2015–2016 **Reservoir Geophysicist | Internship**

Well Services of Iran, Iran

Project: Borehole seismic data processing

ACADEMIC VOLUNTEERING AND LEADERSHIP

2024 **Chair of Technical Sessions**

EAGE Annual Meeting, Oslo

2023 **Chair of Technical Sessions**

EAGE Annual Meeting, Vienna

2022–2023 **Teaching Assistant in Inverse Modeling Course (AESM1330)**

Delft University of Technology

2023–Present **Technical Reviewer**

Geophysical Prospecting Journal

2022–Present **Technical Reviewer**

Extended Abstracts for the SEG and EAGE Annual Meetings

List of Publications

Journal Papers

1. **S. Abolhassani**, and D. J. Verschuur, *Efficient preconditioned least-squares wave-equation migration.* Geophysics, 89(3), S275-S288 (2024).

2. **S. Abolhassani**, L. Hoogerbrugge, and D. J. Verschuur, *One-way reflection waveform inversion with depth-dependent gradient preconditioning.* Geophysical Journal International, 240(1), 652–672 (2025).

3. **S. Abolhassani**, and D. J. Verschuur, *Improved one-way reflection waveform inversion and strategies for optimal offset selection.* Submitted to Geophysical Prospecting (2024).

4. A. Li, D. J. Verschuur, X. Liu, and **S. Abolhassani**, *An accurate propagator for heterogeneous media in full wavefield migration.* IEEE Transactions on Geoscience and Remote Sensing (2024).

Conference Papers

1. **S. Abolhassani** and E. Verschuur, *A Comparative Study of Three Hessian Approximations in Wave-Equation Migration.* in 85th EAGE Annual Conference & Exhibition (2024).

2. **S. Abolhassani**, and D. J. Verschuur, *High-resolution one-way reflection waveform inversion.* in 84th EAGE Annual Conference & Exhibition, Extended Abstract (2023).

3. A. Li, D. J. Verschuur, and **S. Abolhassani**, *Full wavefield migration based on eigen-decomposition propagation operators.* in 84th EAGE Annual Conference & Exhibition (2023).

4. **S. Abolhassani**, and D. J. Verschuur, *Fast Gauss-Newton full-wavefield migration*, in Second International Meeting for Applied Geoscience & Energy (IMAGE), Expanded Abstract (2022).

5. **S. Abolhassani** and E. Verschuur, *Towards a more robust joint migration inversion.* in 83rd EAGE Annual Conference & Exhibition, Extended Abstract (2022).

6. **S. Abolhassani**, E. Slob, and F. Vossepoel, *Probabilistic estimation of reservoir parameters using the complementary nature of seismic and mCSEM data*, in 82nd EAGE Annual Conference & Exhibition, Extended Abstract (2021).

7. S. M. A. Shoja, **S. Abolhassani**, and N. Amini, *A comparison between time-domain and frequency-domain full waveform inversion*, in 80th EAGE Conference & Exhibition, Extended Abstract (2018).