Wave-equation Q tomography and least-squares migration

Thesis by

Gaurav Dutta

In Partial Fulfillment of the Requirements

For the Degree of

Doctor of Philosophy

King Abdullah University of Science and Technology, Thuwal,

Kingdom of Saudi Arabia

March, 2016

The thesis of Gaurav Dutta is approved by the examination committee

Committee Chairperson: Gerard T. Schuster

Committee Member: Paul R. Williamson

Committee Member: Tariq A. Alkhalifah

Committee Member: Daniel B. Peter

Committee Member: Ying Wu

Committee Member: George M. Turkiyyah

ABSTRACT

Wave-equation Q tomography and least-squares migration

Gaurav Dutta

This thesis designs new methods for Q tomography and Q-compensated prestack depth migration when the recorded seismic data suffer from strong attenuation. A motivation of this work is that the presence of gas clouds or mud channels in overburden structures leads to the distortion of amplitudes and phases in seismic waves propagating inside the earth. If the attenuation parameter Q is very strong, i.e., Q<30, ignoring the anelastic effects in imaging can lead to dimming of migration amplitudes and loss of resolution. This, in turn, adversely affects the ability to accurately predict reservoir properties below such layers.

To mitigate this problem, I first develop an anelastic least-squares reverse time migration (Q-LSRTM) technique. I reformulate the conventional acoustic least-squares migration problem as a viscoacoustic linearized inversion problem. Using linearized viscoacoustic modeling and adjoint operators during the least-squares iterations, I show with numerical tests that Q-LSRTM can compensate for the amplitude loss and produce images with better balanced amplitudes than conventional migration.

To estimate the background Q model that can be used for any Q-compensating migration algorithm, I then develop a wave-equation based optimization method that inverts for the subsurface Q distribution by minimizing a skeletonized misfit function ϵ. Here, ϵ is the sum of the squared differences between the observed and the pre-

dicted peak/centroid-frequency shifts of the early-arrivals. Through numerical tests on synthetic and field data, I show that noticeable improvements in the migration image quality can be obtained from Q models inverted using wave-equation Q tomography. A key feature of skeletonized inversion is that it is much less likely to get stuck in a local minimum than a standard waveform inversion method.

Finally, I develop a preconditioning technique for least-squares migration using a directional Gabor-based preconditioning approach for isotropic, anisotropic or anelastic least-squares migration. During the least-squares iterations, I impose sparsity constraints on the inverted reflectivity model in the local Radon domain. The forward and the inverse mapping of the reflectivity to the local Radon domain is done through 3D Fourier-based discrete Radon transform operators. Using numerical tests on synthetic and 3D field data, I demonstrate that the proposed preconditioning approach can discriminate against artifacts in the image resulting from irregular or insufficient acquisition and can produce images with improved signal-to-noise ratio when compared with standard migration.

ACKNOWLEDGEMENTS

I would first like to express my sincere gratitude to my advisor, Prof. Gerard T. Schuster, for his guidance throughout my PhD study at the King Abdullah University of Science and Technology (KAUST). He gave me the freedom to work on my dissertation and his scientific insights and encouragement were invaluable to me. His passion for science and his ability to explain very complex ideas in simplified terms will always motivate me.

I am also deeply grateful to the members of my dissertation committee: Dr. Paul R. Williamson, Prof. Tariq A. Alkhalifah, Prof. Daniel B. Peter, Prof. Ying Wu and Prof. George M. Turkiyyah. I thank them for taking their time, patience and for their insights and suggestions which benefitted my thesis tremendously.

I am indebted to TOTAL E&P France and TOTAL E&P USA for offering me several internships during my PhD. I am especially thankful to Dr. Paul R. Williamson, who is also my co-advisor, and a great teacher and mentor for me. I would also like to acknowledge the guidance that I received from Dr. Bertrand Duquet, Dr. Matteo Giboli, Dr. Cyril Agut, Dr. Piero Basini, Dr. Huimin Guan and Dr. Rami Nammour during my internships in TOTAL.

I also thank all of my CSIM colleagues of older and younger generations for their help and discussions during the last five years. I always learnt something from my discussions with Abdullah AlTheyab, Xin Wang, Wei Dai, Ge Zhan, Yunsong Huang, Mrinal Sinha, Bowen Guo and Kai Lu. I specially acknowledge the help of Dr. Sherif M. Hanafy, Dr. Ola AlHagan and Dr. Veronika Pelletier who helped me settle at CSIM during my first year.

Last but not the least, I thank my mother, my (late) father and my sister for their love, patience and support during all these years. They stood by me in all times and whatever I could accomplish in life was because of their support. I cannot express in words how much I owe them.

TABLE OF CONTENTS

Examination Committee Approval **2**

Copyright **3**

Abstract **4**

Acknowledgements **6**

List of Figures **11**

1 Introduction **16**
1.1 Linear viscoelasticity 16
1.2 Viscoelasticity and wave propagation 17
1.3 Motivation for attenuation compensation in seismic imaging 19
1.4 Previous works on Q compensation 22
1.5 Technical contributions 26
1.6 Thesis overview 27
1.6.1 Chapter 2: Q-LSRTM 27
1.6.2 Chapter 3: Wave-equation Q tomography 28
1.6.3 Chapter 4: LSRTM with directional Gabor-based preconditioning 28
1.6.4 Chapter 5: Conclusions 29

2 Attenuation compensation for least-squares reverse time migration using the viscoacoustic wave-equation **30**
2.1 Summary 30
2.2 Introduction 31
2.3 Theory of acoustic LSRTM 34
2.4 LSRTM using the viscoacoustic wave-equation 36
2.4.1 Q-LSRTM algorithm 41
2.5 Numerical results 43
2.5.1 Layered velocity model 43

2.5.2 Marmousi model 45
2.5.3 Sensitivity of Q-LSRTM to errors in the velocity model 49
2.5.4 Sensitivity of Q-LSRTM to errors in the Q model 52
2.5.5 Friendswood Crosswell Field Data 54
2.6 Discussion 58
2.7 Conclusions 62

3 Wave-equation Q tomography 63
3.1 Summary 63
3.2 Introduction 63
3.3 Theory of wave-equation Q inversion 66
3.3.1 Connective function 67
3.3.2 Misfit function 70
3.3.3 Fréchet derivative 71
3.3.4 WQ algorithm 75
3.4 Numerical Results 76
3.4.1 Crosswell Gaussian Q model 77
3.4.2 Surface-seismic model 77
3.4.3 Friendswood crosswell field data 79
3.5 Discussion 87
3.6 Conclusions 88

4 Least-squares reverse time migration with directional Gabor-based preconditioning 90
4.1 Summary 90
4.2 Introduction 91
4.3 Theory 94
4.4 Numerical results 98
4.4.1 Synthetic example 99
4.4.2 2D land data example 99
4.4.3 3D marine data example 102
4.5 Discussion 105
4.6 Conclusions 111

5 Conclusions and Future Work 114
5.1 Main Results 114
5.1.1 Q-LSRTM 114

5.1.2 Wave-equation Q tomography 115
5.1.3 LSRTM using Radon-based preconditioning 115
5.2 Future Research Work 116
5.2.1 Q-LSRTM and WQ using alternative time-domain Q formulations 116
5.2.2 Wave-equation Q tomography in the frequency domain 117
5.2.3 Wave-equation reflection Q tomography 117
5.2.4 LSRTM using Radon-based preconditioning in the presence of free-surface or internal multiples 118

References **121**

Appendices **130**

LIST OF FIGURES

1.1 (a) Hysteresis loss, (b) stress relaxation, and (c) creep for a viscoelastic solid. 17
1.2 Mechanical models used to characterize viscoelasticity: (a) Maxwell model where a spring is connected in series with a dashpot, (b) Kelvin-Voigt model where a spring is connected in parallel with a dashpot, and (c) a Zener or standard linear solid model where a spring is connected in series with a parallel combination of another spring and a dashpot ([Carcione, 2007]). 18
1.3 Snapshots showing an expanding wavefront at different instances of time for a medium with no attenuation (left panels), moderate attenuation (middle panels) and high attenuation (right panels). 19
1.4 Depth slices from Figure 1.3 showing the effect of attenuation on the amplitude and phase of a propagating seismic wave. 20
1.5 One-way wave-equation migration images from North Sea: (a) before Q compensation, and (b) after Q compensation ([Valenciano et al., 2012]). 22
1.6 Comparison between reverse time migration (RTM) images from the Alaminos Canyon in the Gulf of Mexico: (a) standard RTM image, and (b) Q-RTM image ([Zhang et al., 2010]). 23
1.7 Comparison between TTI prestack depth migration (PSDM) images from the Campos Basin in offshore Brazil: (a) standard TTI PSDM image, and (b) TTI PSDM image after Q compensation ([Zhou et al., 2011]). 23
1.8 Kirchhoff migration images from offshore Brunei: (a) before Q compensation, and (b) after Q prestack depth-migration ([Gamar et al., 2015]). 24

2.1 A layered model: (a) true velocity model, (b) true Q model, (c) migration velocity model, (d) Q model used for Q-RTM and Q-LSRTM. 44

2.2 Comparison between images from (a) acoustic RTM, (b) acoustic LSRTM, (c), Q-RTM, and (d) Q-LSRTM. The black arrows point to the reflectors below the strong attenuation layer where improvements from Q-LSRTM can be seen. 46
2.3 Magnified views of Figure 2.2 where all the images have been normalized such that the short reflector at the top of the Q anomaly have the same magnitude. 47
2.4 The modified Marmousi model: (a) true velocity model, (b) true Q model, (c) migration velocity model, (d) Q model used for Q-RTM and Q-LSRTM. 48
2.5 Comparison between images from (a) acoustic RTM, (b) acoustic LSRTM, (c), Q-RTM, and (d) Q-LSRTM. The black and blue boxes point to the areas for zoom views. 50
2.6 Magnified views of the black (left) and blue (right) boxes in Figure 2.5. (a), (d) True reflectivity models used only for comparison. (b), (e) acoustic LSRTM images. (c), (f) Q-LSRTM images. 51
2.7 Convergence curves for acoustic and Q-LSRTM for the Marmousi model. 52
2.8 Sensitivity of Q-LSRTM to errors in the migration velocity model. The figures in the left panel show the different velocity models used for Q-LSRTM. On the right are the corresponding Q-LSRTM images after 20 iterations. 53
2.9 Normalized data residual vs iteration number for Q-LSRTM with different percentage errors in the migration velocity model. 54
2.10 Sensitivity of Q-LSRTM to errors in the migration Q model. The figures in the left panel show the different Q models used for Q-LSRTM. The true model had a background Q value of 0.05. On the right are the corresponding Q-LSRTM images after 20 iterations for different background Q values. 55
2.11 Normalized data residual vs iteration number for Q-LSRTM with different migration Q models. 56
2.12 (a) A raw CSG from the Friendswood crosswell data. (b) The same CSG after applying all the data processing steps. 57
2.13 Background models used for Q-LSRTM: (a) Velocity tomogram estimated for the crosswell data using early-arrival waveform inversion. (b) Q tomogram estimated using wave-equation Q tomography. . . . 58

2.14 Comparison between images from (a) acoustic RTM, (b) acoustic LSRTM, (c) Q-RTM, and (d) Q-LSRTM after 10 iterations. The blue and black boxes point to the areas for zoom views. 59
2.15 Magnified views of the blue boxes in Figure 2.14. The black arrows point to the areas where improvements in amplitude balancing can be seen with Q-LSRTM. 60
2.16 Magnified views of the black boxes in Figure 2.14. The black arrows point to the areas where improvements in amplitude balancing can be seen with Q-LSRTM. 60

3.1 Variation of the parameters (a) τ_ϵ, (b) τ_σ, and (c) τ for different values of Q. The central frequency of the source wavelet is taken to be 20 Hz. 68
3.2 Comparison between the amplitude spectra of a single transmission arrival in a predicted and an observed trace. 69
3.3 Comparison between a predicted and an observed trace for a single transmission arrival. 69
3.4 Comparison between the amplitude spectra of predicted and observed traces for diving waves. 75
3.5 (a) True Q model, and (b) inverted Q model using WQ. 78
3.6 (a) True velocity, and (b) Q models used for generating the observed data, (c) velocity model used for WQ, and (d) inverted Q tomogram from WQ. 80
3.7 The peak-frequencies for different source-receiver pairs in (a) the observed data, and (b) the predicted data from WQ. 81
3.8 Acoustic (a) RTM, and (b) LSRTM images obtained from the visco-acoustic data. (c) Q-RTM, and (d) Q-LSRTM images for the same data using the tomogram obtained from WQ as the background Q model. The black boxes delineate the areas where improvements in imaging can be seen with WQ and Q-LSRTM. 82
3.9 (a) The first-arrival traveltime picks for the Friendswood data, and (b) the peak-frequencies for different source-receiver pairs. 83
3.10 (a) Velocity model used for WQ, (b) Q tomogram obtained from WQ. 84
3.11 The peak-frequencies for different source-receiver pairs for the predicted data (a) before WQ, and (b) after WQ. 85

3.12 Images from (a) acoustic RTM, (b) acoustic LSRTM, (c) Q-RTM, and (d) Q-LSRTM using the tomogram obtained from WQ as the background Q model. The black boxes delineate the areas where improvements in imaging can be seen with WQ and Q-LSRTM. 86
3.13 Q tomograms obtained from using different background velocity models. The panel on the left shows the velocity models used while the one on the right shows the corresponding Q tomograms obtained from WQ. 88

4.1 A 2D slice of the 3D SEG/EAGE salt model. 100
4.2 Comparison between images from (a) standard isotropic RTM, (b) standard isotropic LSRTM after 5 iterations, (c) preconditioned isotropic RTM, and (d) preconditioned isotropic LSRTM after 5 iterations. . . 101
4.3 Migration velocity model estimated using ray-based tomography for the 2D land dataset. 102
4.4 Comparison between images from (a) standard isotropic RTM, (b) standard isotropic LSRTM after 20 iterations, (c) preconditioned isotropic RTM, and (d) preconditioned isotropic LSRTM after 20 iterations. . . 103
4.5 Comparison between magnified views of the images from (a) standard isotropic RTM, (b) standard isotropic LSRTM, (c) preconditioned isotropic RTM, and (d) preconditioned isotropic LSRTM. The arrows indicate the areas where improvements in imaging can be seen from the Radon-based preconditioning. 104
4.6 Background (a) velocity, (b) epsilon, (c) delta, (d) inline dip, and (e) crossline dip models used for migration. 106
4.7 Comparison between images from (a) standard TTI RTM, and (b) preconditioned TTI RTM. 107
4.8 Comparison between images from (a) standard TTI LSRTM, and (b) preconditioned TTI LSRTM. 108
4.9 Comparison between magnified views of the images in the inline direction from (a) standard RTM, (b) standard LSRTM, (c) preconditioned RTM, and (d) preconditioned LSRTM. The black boxes indicate the areas where improvements in imaging can be seen from the Radon-based preconditioning. 109
4.10 Comparison between magnified views of the images in the crossline direction from (a) standard RTM, (b) standard LSRTM, (c) preconditioned RTM, and (d) preconditioned LSRTM. 110

4.11 Comparison between images from (a) standard isotropic RTM, (b) standard isotropic LSRTM after 5 iterations, (c) preconditioned isotropic RTM, and (d) preconditioned isotropic LSRTM after 5 iterations. The recorded data in this case had free-surface multiples. 112

A.1 Zener model or a standard linear solid (SLS) model ([Carcione, 2007]). 131

Chapter 1

Introduction

1.1 Linear viscoelasticity

Materials exhibit both viscous and elastic characteristics when undergoing deformation ([Christensen, 1982]). When a purely elastic material is subjected to an external stress, 100% of the mechanical energy is stored in the deformation. Once the stress is removed, an elastic material returns quickly to its original state. The stress-strain relation for a linear elastic medium is given by

$$\sigma = M_e \epsilon, \tag{1.1}$$

where σ and ϵ denote the instantaneous stress and the instantaneous strain of the material, respectively, and M_e is the elastic modulus. In such a material, an instantaneous stress σ produces an instantaneous strain ϵ scaled by the elastic modulus M_e.

However, a viscoelastic material resists the build up of strain when an external stress is applied. When an external load is applied to a viscous material, some of the mechanical energy is dissipated by friction and heating. The dissipation of energy or the hysteresis loss for such a material is shown in Figure 1.1(a). Under a fixed strain, a viscoelastic material relaxes to its internal stress state over a period of time when

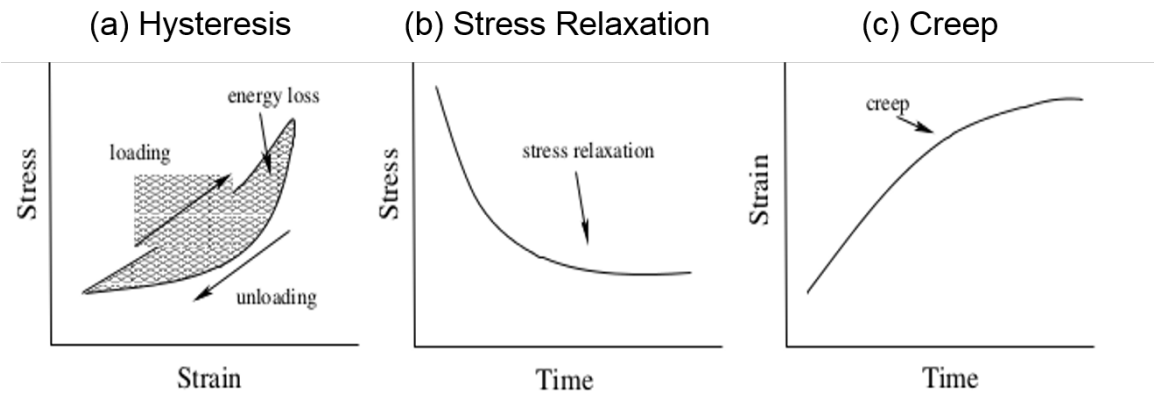

Figure 1.1: (a) Hysteresis loss, (b) stress relaxation, and (c) creep for a viscoelastic solid.

the external load is removed. Similary, under a fixed stress, the material creeps and reaches an asymptotic strain value. The stress relaxation and the creep behaviour are shown in Figures 1.1(b) and 1.1(c), respectively, while the stress-strain relation is given by

$$\sigma = G\frac{\partial \epsilon}{\partial t}. \tag{1.2}$$

Here, G is the viscosity variable or the relaxation function of the medium and $\frac{\partial \epsilon}{\partial t}$ is the strain rate.

1.2 Viscoelasticity and wave propagation

The real earth is anelastic and it distorts the amplitudes and the phases of the propagating seismic waves ([Aki and Richards, 1980]). Attenuation of P-waves can be quantified by a quality factor, Q, which accounts for the phase shift as a function of the frequency content of the propagating waves and the distance traveled. The quality factor, Q, quantifies dissipation and is defined as twice the time-averaged strain-energy density divided by the time-averaged dissipated energy-density ([Carcione, 2007]). Lower values of Q imply more energy loss of the wave per cycle or high attenuation.

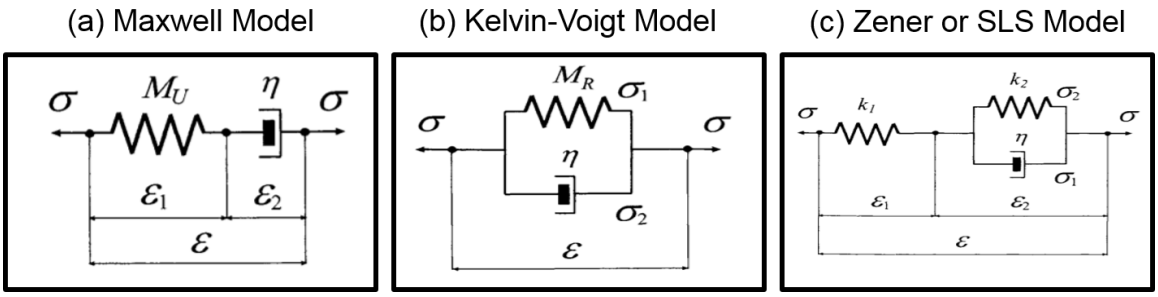

Figure 1.2: Mechanical models used to characterize viscoelasticity: (a) Maxwell model where a spring is connected in series with a dashpot, (b) Kelvin-Voigt model where a spring is connected in parallel with a dashpot, and (c) a Zener or standard linear solid model where a spring is connected in series with a parallel combination of another spring and a dashpot ([Carcione, 2007]).

The anelastic behavior of wave propagation can be described by viscoelastic/viscoacoustic wave equations based on mechanical models. An elastic solid can be described by only weightless springs that have no inertial effects. On the other hand, a viscoelastic material needs additional inertial elements, namely dashpots, that consist of loosely fitting pistons in cylinders filled with a viscous fluid. There are three different kinds of mechanical models that can be used to describe viscoelastic wave propagation. These models, shown in Figure 1.2, have different combinations of springs and dashpots in series or in parallel.

Using the standard linear solid (SLS) model in Figure 1.2(c), the effect of attenuation on the amplitude and phase of a propagating seismic wave is shown in Figure 1.3 for a homogeneous medium with a background velocity of 3000 m/s and for different values of Q. A Ricker wavelet with 20 Hz peak frequency is used as the source wavelet. The source is excited at the center of the model and the snapshots are taken at different intervals of time. Figure 1.4 compares the depth slices at different instances of time. It is evident from these figures that as the wave starts propagating, the amplitudes in the three cases are very similar. However, as the propagation time increases and the wave travels a greater distance, the wave amplitude in the case of Q= 20 is attenuated the most. The high-frequency components are distorted more compared

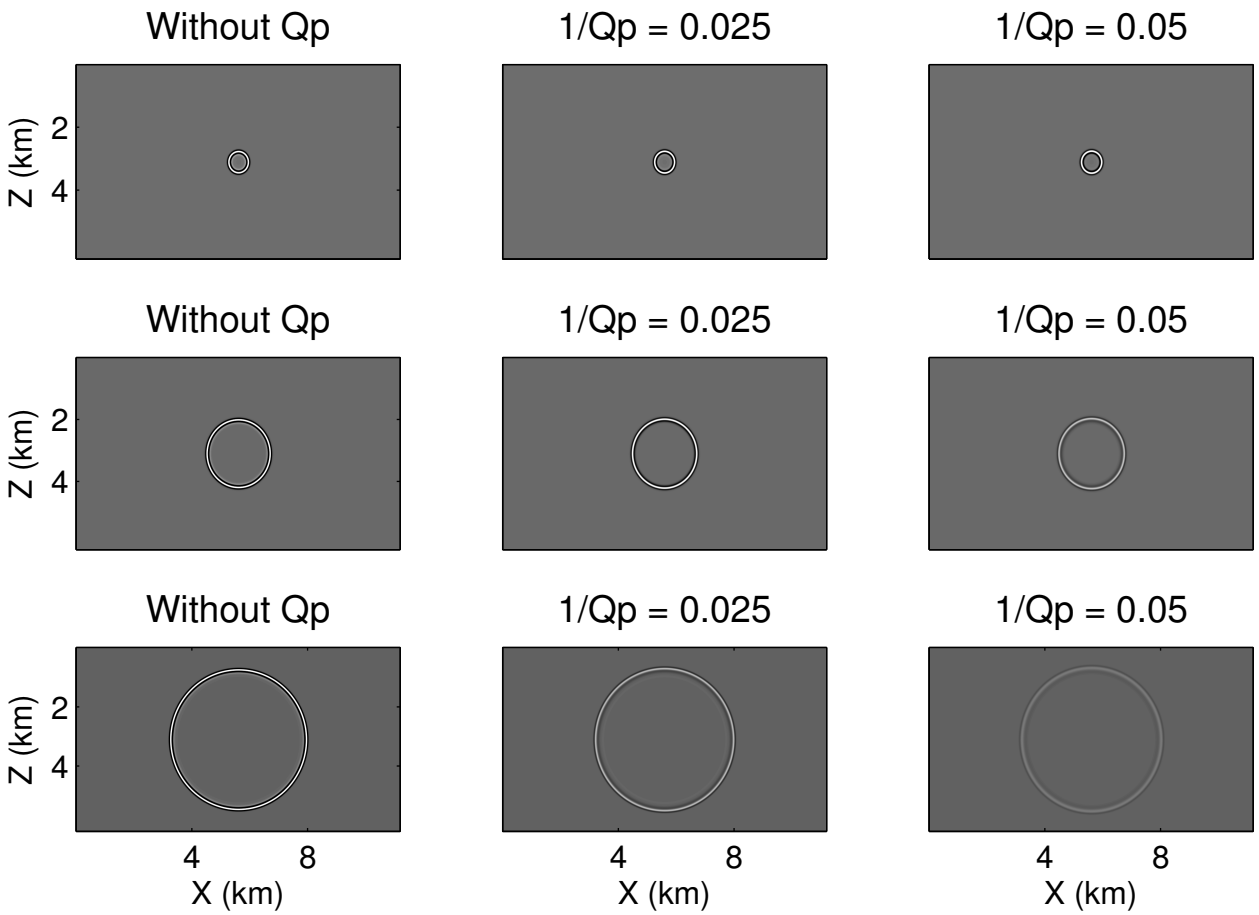

Figure 1.3: Snapshots showing an expanding wavefront at different instances of time for a medium with no attenuation (left panels), moderate attenuation (middle panels) and high attenuation (right panels).

to the low frequencies. For moderate-attenuation values ($Q \approx 40$), as shown by the red curves in Figure 1.4, the attenuation effect at different times is still significant compared to the blue curves which represent the case where there is no attenuation.

1.3 Motivation for attenuation compensation in seismic imaging

Gas clouds form in an overburden region when there is leakage of gas from deeper reservoirs. The gas migrates upwards through faults and fractures and get trapped

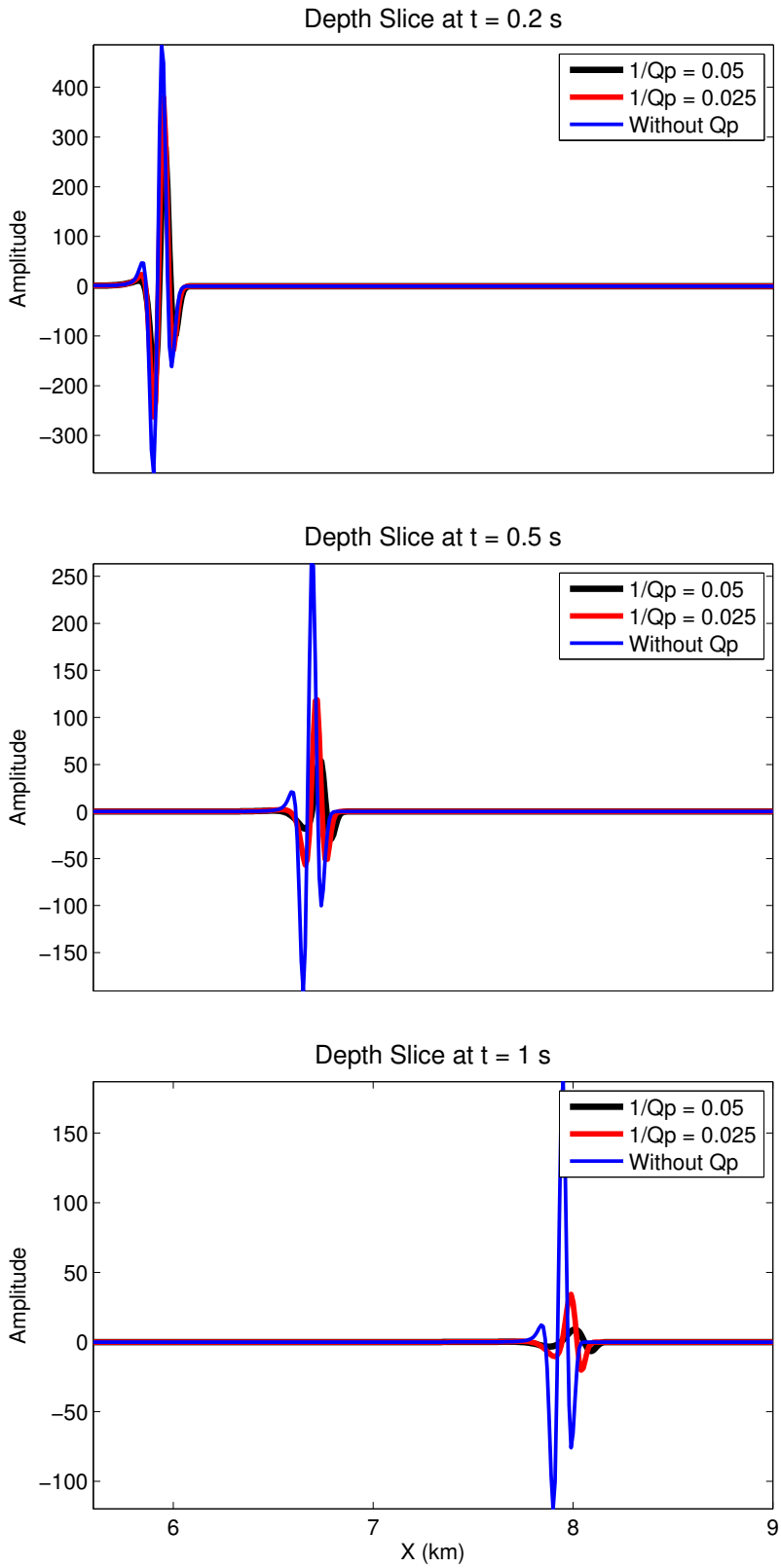

Figure 1.4: Depth slices from Figure 1.3 showing the effect of attenuation on the amplitude and phase of a propagating seismic wave.

in thin layered sediments. The presence of gas accumulations in the near-surface sediments is a good indicator of the presence of hydrocarbon reservoirs. If there is an interbedding of thin attenuative layers, it can lead to frequency-dependent seismic wave attenuation with direction or attenuation anisotropy ([Mavko and Nur, 1979, Zhu and Tsvankin, 2006, Zhu et al., 2007, Zhu and Tsvankin, 2007, Behura and Tsvankin, 2009]). Attenuation anisotropy contains useful information about pore-fluid content and properties and can distinguish between different fracturing and crystal orientation mechanisms that can lead to velocity anisotropy ([Carter and Kendall, 2006]).

However, the presence of gas anomalies in overburden structures possesses a significant challenge in seismic data processsing. The values of Q for unconsolidated gas-sandstones, gas clouds and shales are typically very low (Q $\approx$ 15-30). Fluids or mud channels in overburden structures cause wave dissipation and attenuation that severely affect the amplitude and phase of a propagating seismic wave. As a result, the imaging of the target oil and gas reservoirs below such formations is affected.

In places like the North Sea, the presence of gas anomalies leads to serious issues in imaging . The problems with conventional seismic processing based on the acoustic approximation can be seen in Figure 1.5(a). There is a dimming of the reflector amplitudes and loss of resolution in certain areas of the image because of the presence of shallow gas clouds. This is because conventional depth migration and tomography do not take into account the anelastic nature of wave propagation. Similar problems during migration can also be seen in the images from the Gulf of Mexico (Figure 1.6) offshore Brazil (Figure 1.7), offshore Brunei (Figure 1.8) and also in many other parts of the world ([Gherasim et al., 2010, Cavalca et al., 2011, Valenciano et al., 2011, Zhou et al., 2011, Valenciano et al., 2012, Teng et al., 2013, Valenciano and Chemingui, 2013, Bai et al., 2013, Gamar et al., 2015]) . Thus, imaging from data that suffers from strong attenuation requires migration and tomography techniques that take into account the physics of anelasticity in the forward modeling

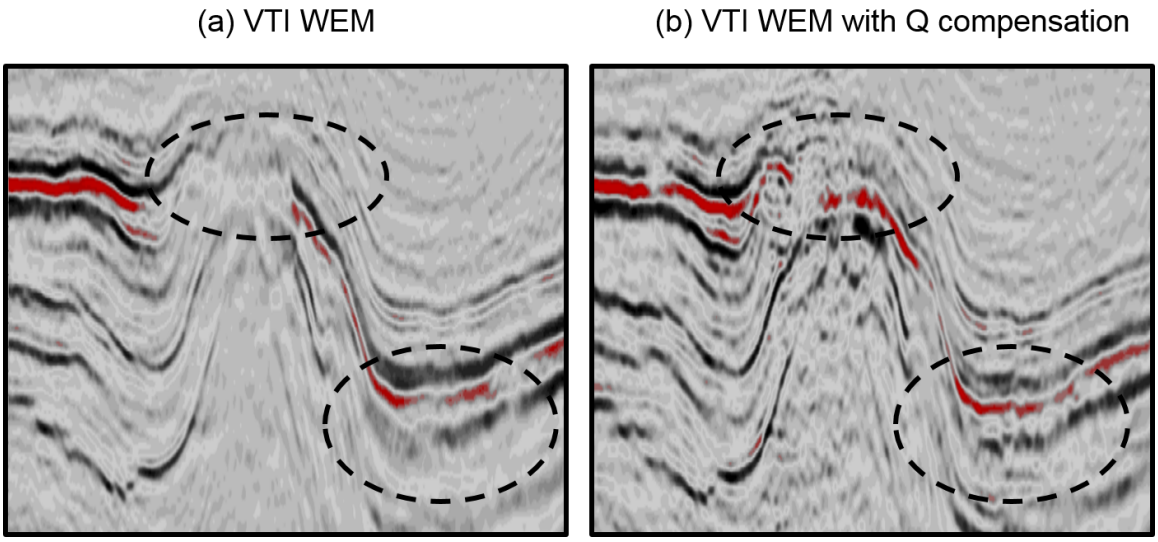

Figure 1.5: One-way wave-equation migration images from North Sea: (a) before Q compensation, and (b) after Q compensation ([Valenciano et al., 2012]).

and in the adjoint equations.

1.4 Previous works on Q compensation

From a data acquisition perspective, any effort to see through gas formations require using very expensive data acquisition techniques such as ocean-bottom recordings. However, deployment of ocean-bottom seismometers or ocean-bottom cables for 3D exploration purposes become prohibitively expensive. It is more feasible to compensate for the attenuation loss during the conventional seismic data processing workflow.

The earliest efforts to compensate for the attenuation loss were made in the data-domain using inverse Q-filtering based methods. [Bickel and Natarajan, 1985] used plane-wave Q deconvolution to compensate for the amplitude and high-frequency loss in prestack seismic data. They deconvolved the effects of the propagation path between a source and a receiver within the seismic band by using time-varying filters and doing a time reversal with Q replaced by $-$Q in the complex wavenumber domain. This changed absorption to gain in the wavenumber domain and each frequency

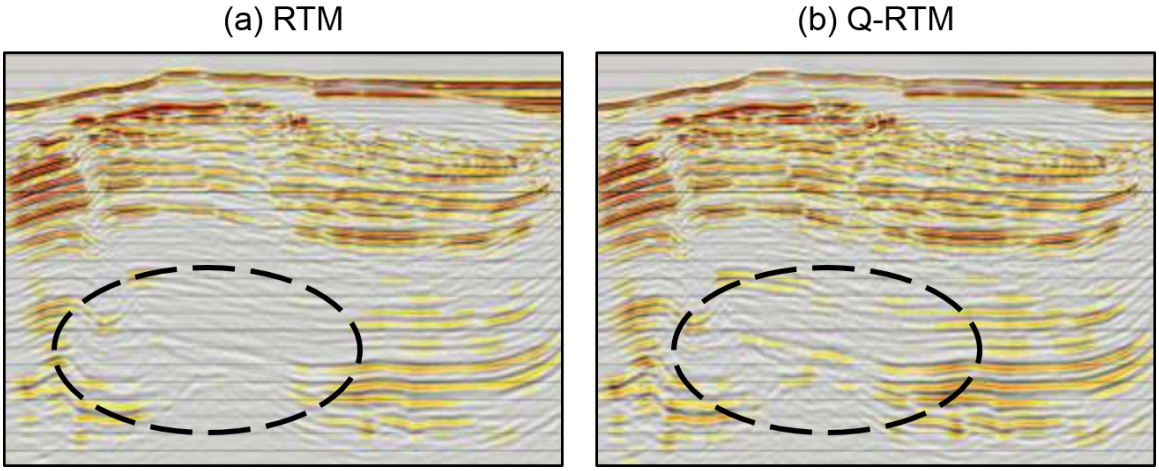

Figure 1.6: Comparison between reverse time migration (RTM) images from the Alaminos Canyon in the Gulf of Mexico: (a) standard RTM image, and (b) Q-RTM image ([Zhang et al., 2010]).

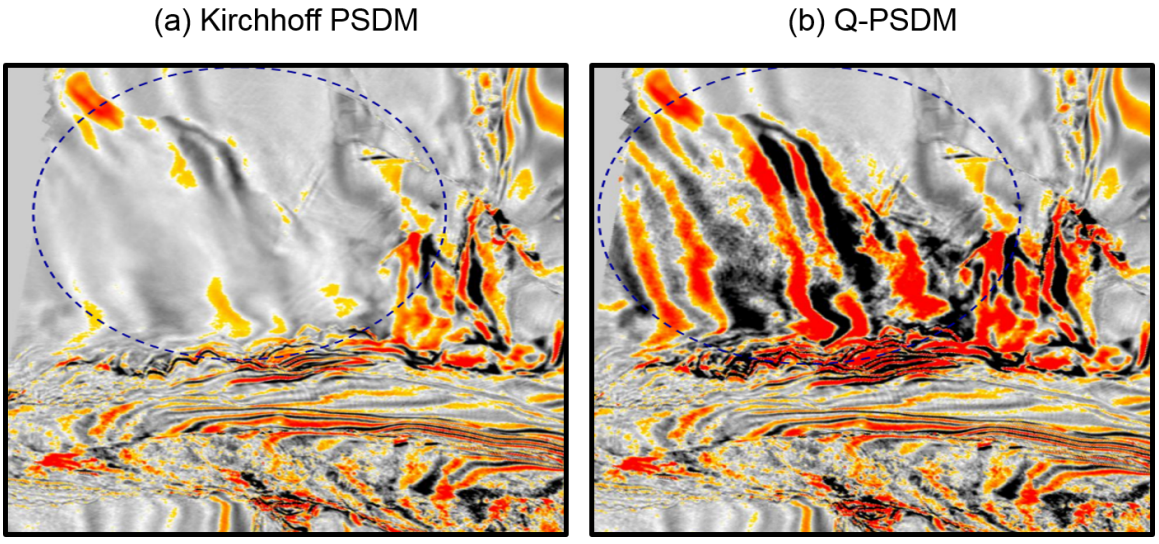

Figure 1.7: Comparison between TTI prestack depth migration (PSDM) images from the Campos Basin in offshore Brazil: (a) standard TTI PSDM image, and (b) TTI PSDM image after Q compensation ([Zhou et al., 2011]).

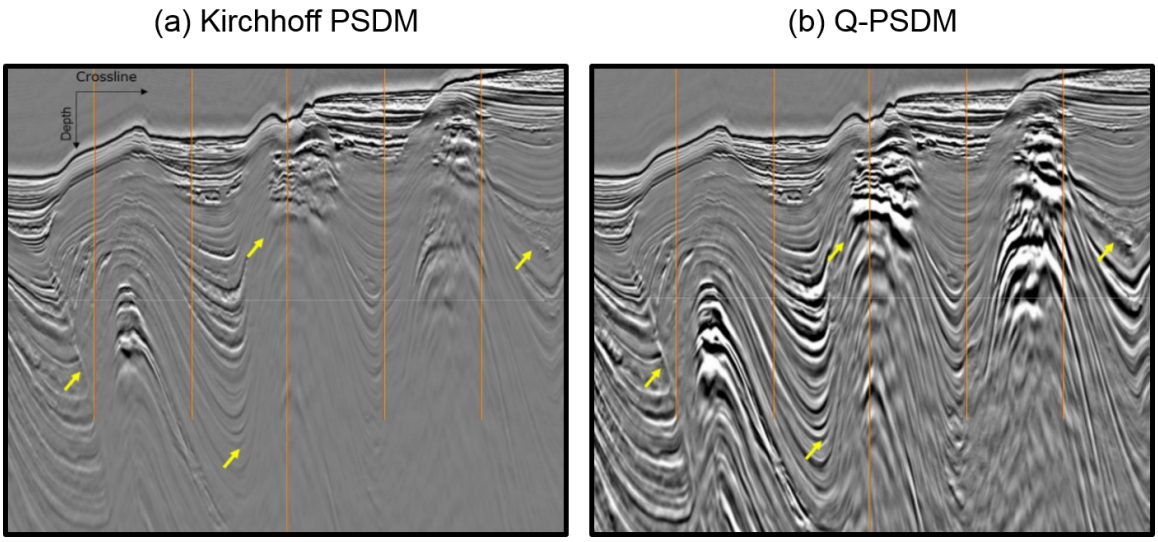

Figure 1.8: Kirchhoff migration images from offshore Brunei: (a) before Q compensation, and (b) after Q prestack depth-migration ([Gamar et al., 2015]).

component is restored to its original amplitude and phase. A similar data-domain approach was also used by [Hargreaves and Calvert, 1991] who performed inverse Q filtering for prestack data in the Fourier domain. They adopted a Stolt migration-type approach where they backpropagated the recorded data at the surface, removed the interval effect of Q at each depth step and retrieved the Q-filtered wavefield at time $t = 0$ by taking the inverse Fourier transform. Such data-domain approaches assume that the velocity is laterally constant.

However, in complex geological media where there are strong lateral and vertical variations in the subsurface properties, such data-domain filtering techniques were proved inadequate. Since attenuation loss occurs during wave propagation, Q compensation was gradually incorporated in the model building and depth migration workflows. Including Q during depth migration requires a migration Q model whereas Q estimation requires an accurate velocity model and a migration image. Such a dependency necessitates the development of accurate Q tomography and Q-compensated migration techniques that would enable interpreters to study the presence of fluids in oil and gas reservoirs with both low-intermediate wavenumber and high-wavenumber

resolution using Q tomography and migration, respectively.

In the context of prestack depth migration (PSDM), Kirchhoff and one-way wave-equation based migration techniques were earlier used to perform attenuation compensation ([Dai and West, 1994, Yu et al., 2002, Xie et al., 2009]). These approaches compensate for the attenuation loss during wavefield extrapolation by using anti-dissipation or amplitude amplification filters based on the interval Q model. However, the use of amplitude or high frequency amplification filters leads to numerical instabilities during wavefield extrapolation. The problem is excarberated in field data applications where there is noise in the recorded data.

With the substantial increase in computing power and resources over the last decade, Q migration gained prominence with the more accurate two-way viscoacoustic wave-equation or viscoacoustic reverse time migration (RTM) ([Zhang et al., 2010, Suh et al., 2012, Fletcher et al., 2012, Bai et al., 2013, Zhu et al., 2014, Zhu and Harris, 2015]). Different viscoacoustic wave-equations were proposed with separate controls over phase and amplitude to compensate for the attenuation loss. Most of these approaches also relied on using high-cut filters to stabilize the propagation of the high-frequency components during wavefield extrapolation. Also, there were no efforts made to consider Q compensation during migration as part of an anelastic least-squares migration workflow since standard migration can be regarded as the first iteration of a least-squares migration algorithm ([Lailly, 1984, Nemeth et al., 1999]).

To estimate the background Q model required by these Q-based migration techniques, there are two categories of methods available:

- *Data-domain methods:* The data-domain techniques for Q tomography are mostly ray-based or full waveform inversion (FWI) based. In ray-based methods, the frequency-shifts between the predicted and the observed traces are smeared along raypaths ([Quan and Harris, 1997, He et al., 2013]). Such ray-based tomography methods are based on the high-frequency assumption where

it is assumed that the scale of variation in Q in the medium is much larger than the seismic wavelength. On the other hand, the FWI-based Q inversion approaches rely on minimizing the amplitude and the phase differences between the predicted and the observed traces to obtain the Q model ([Bai et al., 2013, Wang and Zhang, 2014]). Such approaches are susceptible to cycle-skipping problems that are typically associated with any FWI-like algorithm. However, the use of a robust skeletonized inversion approach where a skeletonized representation of the data is inverted using wave-equation operators has not been studied so far.

- *Image-domain methods:* The image-domain techniques for Q tomography rely on minimizing a residual image that is the difference between the image computed by a background Q model and an attenuation-free target image ([Zhou et al., 2011, Shen et al., 2014, Shen and Zhu, 2015, Shen et al., 2015]). The images can be picked horizons along which amplitude ratio maps or common image gathers are generated. The spectral difference between the observed and the target images are then minimized using ray-based or wave-equation-based migration velocity analysis operators. These approaches are computationally very expensive, especially for 3D problems.

In this dissertation, I develop novel time-domain Q-migration and Q-tomography techniques using the isotropic viscoacoustic wave-equation. The major original contributions of this dissertation are summarized below.

1.5 Technical contributions

I developed three novel imaging methods for migration and inversion of viscoacoustic data.

1. An anelastic least-squares reverse time migration (Q-LSRTM) technique to compensate for the attenuation loss during migration.

2. A wave-equation Q tomography approach using the frequency-shifts between the predicted and the observed traces. This is a skeletonized representation of the data that largely avoids getting stuck in local minima..

3. A directional Gabor-based preconditioning approach for least-squares reverse time migration (LSRTM) which can be used for acoustic, anisotropic and anelastic LSRTM.

1.6 Thesis overview

The dissertation is organized according to the following chapters:

1.6.1 Chapter 2: Q-LSRTM

In chapter 2, I propose an anelastic least-squares migration method, denoted as Q least-squares reverse time migration (Q-LSRTM), to compensate for the distortion of amplitudes and phases of seismic waves propagating inside the earth during migration. During the least-squares iterations, I use a linearized viscoacoustic modeling operator for forward modeling and use the corresponding adjoint equations, derived using the adjoint-state method, for backpropagating the residual data. The merit of this approach compared to conventional RTM and LSRTM is that Q-LSRTM compensates for the amplitude loss due to attenuation and can produce images with better balanced amplitudes below highly attenuative layers. Through numerical tests on synthetic and field data, I illustrate the advantages of Q-LSRTM over RTM and LSRTM when the recorded data have strong attenuation effects.

1.6.2 Chapter 3: Wave-equation Q tomography

In chapter 3, I present a wave-equation gradient optimization method that inverts for the subsurface Q distribution by minimizing a skeletonized misfit function which is the sum of the squared differences between the observed and the predicted peak/centroid-frequency shifts of the early arrivals. The gradient is computed by migrating the observed traces weighted by the frequency-shift residuals. The background Q model is perturbed until the predicted and the observed traces have the same peak frequencies or the same centroid frequencies. Using numerical tests on synthetic and field data, I show that an improved accuracy of the Q model by wave-equation Q tomography (WQ) leads to a noticeable improvement in the migration image quality.

1.6.3 Chapter 4: LSRTM with directional Gabor-based preconditioning

In chapter 4, I present a LSRTM method using directional Gabor-based preconditioning to overcome the low signal-to-noise (SNR) problem of noisy or severely undersampled data. I use a high-resolution local-Radon transform of the reflectivity and impose sparseness constraints on the inverted reflectivity in the local Radon domain during the least-squares iterations. The sparseness constraint is that the inverted reflectivity is sparse in the Radon domain and each location of the subsurface is represented by a limited number of geological dips. The forward and the inverse mapping of the reflectivity to the local Radon domain and vice versa is done through 3D Fourier-based discrete Radon transform operators. The weights for the preconditioning are chosen to be varying locally based on the amplitudes of the local dips or assigned using quantile measures. Using numerical tests on synthetic and 2D land and 3D marine field data, I validate the effectiveness of the proposed approach in producing images with good SNR and fewer aliasing artifacts when compared with standard RTM or

standard LSRTM.

1.6.4 Chapter 5: Conclusions

In chapter 5, I complete the dissertation with a summary of my results and a discussion of the advantages and limitations of the proposed methods. I also suggest some future research topics that can be built on my dissertation.

Chapter 2

Attenuation compensation for least-squares reverse time migration using the viscoacoustic wave-equation

2.1 Summary

Strong subsurface attenuation leads to distortion of amplitudes and phases of seismic waves propagating inside the earth. Conventional acoustic reverse time migration (RTM) and least-squares reverse time migration (LSRTM) do not account for this distortion which can lead to defocusing of migration images in highly attenuative geological environments. To correct for this distortion, I use a linearized inversion method, denoted as Q least-squares reverse time migration (Q-LSRTM). During the least-squares iterations, a linearized viscoacoustic modeling operator is used for forward modeling. The adjoint equations are derived using the adjoint-state method for backpropagating the residual wavefields. The merit of this approach compared to conventional RTM and LSRTM is that Q-LSRTM compensates for the amplitude loss

due to attenuation and can produce images with better balanced amplitudes below highly attenuative layers. Numerical tests on synthetic and field data illustrate the advantages of Q-LSRTM over RTM and LSRTM when the recorded data have strong attenuation effects. Similar to standard LSRTM, the sensitivity tests for background velocity and Q errors show that the liability of this method is the requirement for smooth and accurate migration velocity and attenuation models.

2.2 Introduction

Fluids trapped in overburden structures cause strong attenuation of P-waves that can decrease the resolution of migration images. This can be attributed to the fact that the real earth is anelastic and therefore distorts the amplitudes and phases of the propagating seismic waves ([Aki and Richards, 1980]). If the subsurface attenuation is very strong, ignoring it during migration can lead to blurring of migration amplitudes below these layers.

Attenuation of P-waves can be quantified by a quality factor, Q, which accounts for the phase shift as a function of the frequency content of the propagating waves and the distance traveled. Lower values of Q imply more energy loss of the wave per cycle or high attenuation. For example, the values of Q for unconsolidated gas-sandstones and shales are typically very low ($Q \approx 15-30$) which necessitates the need to account for Q during imaging. Another example is the presence of gas pockets in North Sea sediments that distort the migrated amplitudes of the underlying reflectors.

The earliest efforts to compensate for attenuation loss in seismic data were performed in the data domain using an inverse Q-filtering method ([Bickel and Natarajan, 1985]; [Hargreaves and Calvert, 1991]). These data domain methods partially correct for the attenuation loss because attenuation loss occurs during wave propagation and therefore Q compensation is required during migration.

For pre-stack depth migration, [Xin et al., 2008] and [Xie et al., 2009] performed attenuation compensation using ray-tracing methods. [Dai and West, 1994], [Yu et al., 2002], [Wang, 2008] and [Valenciano et al., 2011] used one-way wave-equation migration methods in the frequency domain for attenuation compensation. These approaches compensate for the attenuation loss during wavefield extrapolation by using anti-dissipation or amplitude amplification filters based on the interval Q model. The more accurate two-way wave-equation migration or reverse time migration (RTM) ([Baysal et al., 1983]; [McMechan, 1983]; [Whitmore, 1983]) does not take into account attenuation because the lossless acoustic wave-equation is used for wavefield extrapolation. Similarly, for least-squares RTM (LSRTM) ([Plessix and Mulder, 2004]; [Dai and Schuster, 2010a]; [Dai et al., 2011]; [Wong et al., 2011]), if the gradient computation at each iteration is done by RTM of the data residuals, then the reflector image below the attenuation zone will be degraded in quality.

In the context of RTM, [Zhang et al., 2010] used the dispersion relation for a viscoacoustic medium ([Kjartansson, 1979]) and derived a pseudo-differential equation with separate controls over phase and amplitude to model and migrate viscoacoustic waves. They required a regularization process to stabilize the back-propagating wavefield since the high frequency amplitudes increase with time and can lead to numerical instability. [Suh et al., 2012] extended the work of [Zhang et al., 2010] for a VTI medium where they ignored the phase effects due to attenuation and only compensated for the amplitude loss. Their approach suffered from the same instability problem during back-propagating the receiver wavefield as in [Zhang et al., 2010]. They applied a high-cut filter to the receiver wavefield to stabilize the amplification of high frequency components. [Bai et al., 2013] derived a new viscoacoustic wave-equation without any memory variable. They accounted for attenuation by incorporating a pseudo-differential operator in the time and space domains. Similar to [Zhang et al., 2010] and [Suh et al., 2012], they applied a high-cut filter to sta-

bilize the wavefield propagation. [Fletcher et al., 2012] proposed a stable approach where they applied separate phase and amplitude filters to the source and receiver wavefields to compensate for amplitude and phase effects. These filters were applied prior to imaging and were estimated based upon running acoustic and viscoacoustic propagators twice to estimate attenuated amplitudes along wavepaths.

In this work, I use the standard time-domain viscoacoustic wave-equation for wavefield extrapolation and use the linearized least-squares inversion method to compensate for the attenuation loss. For least-squares migration, a linearized viscoacoustic wave-equation modeling operator based on the perturbation of the bulk modulus is derived, and then the appropriate adjoint equations and imaging condition are used for the least-squares iterations. The advantage of this approach over other existing methods is that one does not need to modify and regularize the adjoint wave-equations during the receiver-side wavefield extrapolation to compensate for the attenuation loss and hence, it is always stable. Accounting for Q in the modeling and adjoint operations coupled with the least-squares iterations compensates for the amplitude loss due to attenuation. Numerical tests on synthetic and field data show that the amplitudes below highly attenuative layers are better balanced in the inverted images from Q-LSRTM compared to standard RTM and LSRTM, and the reflectors are focused at the right locations. The disadvantage is that this method is expensive (computational cost per iteration is more than six times that of standard RTM) and the computational cost is proportional to the number of least-squares iterations. However, similar to standard LSRTM, the cost of Q-LSRTM can be significantly reduced by using a multisource encoded strategy ([Dai and Schuster, 2009]; [Tang, 2009]; [Dai et al., 2010]; [Huang and Schuster, 2012]). Another liability of this method is that besides the input requirement of a smooth and accurate migration velocity model, a good estimate of the background Q model is also needed.

This chapter is organized into five sections. After the introduction, the second

section describes the theory of standard LSRTM and then the time-domain implementaton of Q-LSRTM is presented in the third section. Numerical results on synthetic and field data are presented in the fourth section. Dicussions based on the numerical simulations are in the fifth section and the conclusions are in the last section.

2.3 Theory of acoustic LSRTM

Conventional least-squares migration seeks to reconstruct the earth's reflectivity image from the recorded waveform data under the Born approximation ([Tarantola, 1984]; [Lailly, 1984]; [Nemeth et al., 1999]; [Duquet et al., 2000]; [Plessix and Mulder, 2004]; [Wong et al., 2011]; [Dai et al., 2012]). In this section, we will briefly review the theory of acoustic LSRTM in the time-domain.

For a given background velocity model, $v_0(\mathbf{x})$, the pressure wavefield, $p_0(\mathbf{x},t)$, satisfies the acoustic wave-equation with constant density,

$$\frac{1}{v_0(\mathbf{x})^2}\frac{\partial^2 p_0(\mathbf{x},t)}{\partial t^2} - \nabla^2 p_0(\mathbf{x},t) = S(\mathbf{x}_s,t), \tag{2.1}$$

where $S(\mathbf{x}_s,t)$ represents a bandlimited point source function at $\mathbf{x}=\mathbf{x}_s$.

For LSRTM, we seek to find the perturbation in the wavefield, $\delta p(\mathbf{x},t)$, related to the perturbation in the velocity, $\delta v(\mathbf{x})$. Under the Born approximation, the perturbed wavefield can be calculated as ([Stolt and Benson, 1986]; [Plessix and Mulder, 2004]; [Dai et al., 2012])

$$\frac{1}{v_0(\mathbf{x})^2}\frac{\partial^2 \delta p(\mathbf{x},t)}{\partial t^2} - \nabla^2 \delta p(\mathbf{x},t) = \frac{m(\mathbf{x})}{v_0(\mathbf{x})^2}\frac{\partial^2 p_0(\mathbf{x},t)}{\partial t^2}, \tag{2.2}$$

where the reflectivity image $m(\mathbf{x})$ is defined as $m(\mathbf{x}) = \frac{2\delta v(\mathbf{x})}{v_0(\mathbf{x})}$. To find the perturbed wavefield numerically, two finite-difference simulations are required, one for the background wavefield $p_0(\mathbf{x},t)$ given by equation 2.1, and the other for the perturbed

wavefield $\delta p(\mathbf{x}, t)$ given by equation 2.2. In the context of least-squares migration, the solution to equation 2.2 can be represented as the matrix-vector operation $\mathbf{d} = \mathbf{Lm}$, where $\mathbf{d}$ is the data, $\mathbf{L}$ is a linear modeling operator and $\mathbf{m}$ represents the migration image which in this case is related to the perturbation in velocity.

The receiver side residual wavefield $r(\mathbf{x}, t)$ can be computed by reverse time propagation ([McMechan, 1983, Whitmore, 1983]) of the residual data $\Delta d(\mathbf{x}_g, t; \mathbf{x}_s)$ and is equivalent to the adjoint state solution defined by [Lailly, 1984] and [Plessix, 2006]. The backprojected field, $r(\mathbf{x}, t)$ is the solution to

$$\frac{1}{v_0(\mathbf{x})^2}\frac{\partial^2 r(\mathbf{x}, t)}{\partial t^2} - \nabla^2 r(\mathbf{x}, t) = \Delta d(\mathbf{x}_g, t; \mathbf{x}_s), \tag{2.3}$$

where the data residual at a geophone location, $\mathbf{x} = \mathbf{x}_g$ and a source at $\mathbf{x} = \mathbf{x}_s$ is given by $\Delta d(\mathbf{x}_g, t; \mathbf{x}_s) = p_{obs}(\mathbf{x}_g, t; \mathbf{x}_s) - \delta p(\mathbf{x}_g, t; \mathbf{x}_s)$. $p_{obs}(\mathbf{x}_g, t; \mathbf{x}_s)$ is the recorded data for the same source-geophone pair. The gradient, $g(\mathbf{x})$, can be computed by taking a zero-lag cross-correlation of the source and residual wavefields as

$$g(\mathbf{x}) = \sum_t \frac{1}{v_0(\mathbf{x})^2}\frac{\partial^2 p_0(\mathbf{x}, t)}{\partial t^2} r(\mathbf{x}, t). \tag{2.4}$$

In matrix-vector notation, the solutions to equations 2.3 and 2.4 is equivalent to the operation $\mathbf{g} = \mathbf{L}^T \Delta \mathbf{d}$. The reflectivity distribution, $m(\mathbf{x})$, can be iteratively estimated using any gradient based method such as the steepest descent method ([Nemeth et al., 1999])

$$\begin{aligned}
\mathbf{g}^{(i)} &= \mathbf{L}^T \left[\mathbf{L}\mathbf{m}^{(i)} - \mathbf{d}^{obs}\right], \\
\alpha &= \frac{\left(\mathbf{g}^{(i)}\right)^T \mathbf{g}^{(i)}}{\left(\mathbf{L}\mathbf{g}^{(i)}\right)^T \left(\mathbf{L}\mathbf{g}^{(i)}\right)}, \\
\mathbf{m}^{(i+1)} &= \mathbf{m}^i + \alpha \mathbf{g}^{(i)},
\end{aligned} \tag{2.5}$$

where i represents the iteration index for the least-squares inversion and $\mathbf{g}^{(i)}$ and α represent the gradient and the step-length, respectively. The gradient at each iteration is computed by RTM of the data residuals, $\mathbf{Lm}^{(i)} - \mathbf{d}^{obs}$. In acoustic LSRTM, the operators $\mathbf{L}$ and $\mathbf{L}^T$ depend on the background velocity model, $v_0(\mathbf{x})$, which remains fixed at every iteration.

The theory of time-domain Q-LSRTM in now formulated in the next section.

2.4 LSRTM using the viscoacoustic wave-equation

For an isotropic viscoelastic solid, the current value of the stress tensor σ_{ij} depends upon the complete past history of the components of the strain tensor ϵ_{ij}. The relation between the stress-strain tensors can be represented in the following integral form ([Christensen, 1982]):

$$\sigma_{ij}(t) = \int_0^t G_{ijkl}(t-\tau)\frac{d\epsilon_{kl}(\tau)}{d\tau}d\tau, \tag{2.6}$$

where the indices i, j have the range 1,2,3 and repeated indices imply the summation convention. G_{ijkl} is a fourth order relaxation tensor such that $G_{ijkl} = 0$ for $-\infty < t < 0$. For a pressure field P acting on the medium, the stress tensor can also be written as

$$\sigma_{ij}(t) = -P(t)\delta_{ij}, \tag{2.7}$$

where δ_{ij} is the Kronecker delta. Substituting equation 2.7 into equation 2.6, an explicit expression for the pressure field can be written as

$$P(t) = -\int_0^t G(t-\tau)\frac{de(\tau)}{d\tau}d\tau,$$
$$P(t) = -M_R\int_0^t \dot{e}(\tau)\left[1-\sum_{l=1}^{L}\left(1-\frac{\tau_{\epsilon l}}{\tau_{\sigma l}}\right)exp\left(-\frac{t-\tau}{\tau_{\sigma l}}\right)\right]d\tau, \tag{2.8}$$

where e denotes the trace of the strain tensor matrix ϵ_{ij} or the dilatation, $\tau_{\sigma l}$ and $\tau_{\epsilon l}$ denote material relaxation times for the l-th mechanism, L is the number of relaxation mechanisms for a standard linear solid model and M_R is the relaxed modulus of the medium. For a standard linear solid (SLS) model, the relaxation function G is derived in Appendix A.

The equation of motion can be written as

$$-\frac{1}{\rho}\nabla P = \frac{\partial \mathbf{v}}{\partial t} \tag{2.9}$$

where ρ represents the density, P represents the pressure wavefield and $\mathbf{v}$ represents the particle velocity vector. Equations (2.8) and (2.9) together describe the deformation in a viscoacoustic medium.

Equation 2.8 is expensive to solve by numerical modeling because of the associated convolution operation. [Robertsson et al., 1994] simplified the convolution term by introducing a memory variable term, r_p. [Blanch et al., 1995] and [Zhu et al., 2013] later showed that only one relaxation mechanism ($L = 1$) is sufficient for practical purposes. Thus, for practical numerical modeling applications, the equations of motion for a 2D viscoacoustic medium can be written as ([Christensen, 1982,

Carcione et al., 1988, Blanch and Symes, 1995])

$$\begin{aligned}
&\frac{\partial P}{\partial t} + K\frac{\tau_\epsilon}{\tau_\sigma}\left(\nabla\cdot\mathbf{v}\right) + r_p = S(\mathbf{x}_s, t),\\
&\frac{\partial \mathbf{v}}{\partial t} + \frac{1}{\rho}\nabla P = 0,\\
&\frac{\partial r_p}{\partial t} + \frac{1}{\tau_\sigma}\left(r_p + K\left(\frac{\tau_\epsilon}{\tau_\sigma} - 1\right)\left(\nabla\cdot\mathbf{v}\right)\right) = 0.
\end{aligned} \tag{2.10}$$

Here, $\mathbf{v} = \{v_x, v_y, v_z\}$ represents the particle velocity vector, K represents the bulk modulus of the medium and $S(\mathbf{x}_s, t)$ represents a bandlimited point source function at $\mathbf{x} = \mathbf{x}_s$. A detailed derivation of these equations is shown in Appendix B.

The stress and strain relaxation parameters, τ_σ and τ_ϵ, are related to the quality factor, Q, and the reference angular frequency, ω, usually chosen to be the central frequency of the source wavelet, as (see Appendix C)

$$\begin{aligned}
\tau_\sigma &= \frac{\sqrt{1+\frac{1}{Q^2}} - \frac{1}{Q}}{\omega},\\
\tau_\epsilon &= \frac{1}{\omega^2\tau_\sigma} = \frac{\sqrt{1+\frac{1}{Q^2}} + \frac{1}{Q}}{\omega}.
\end{aligned} \tag{2.11}$$

Similar to standard acoustic LSRTM, for Q-LSRTM, we seek to find the perturbed wavefield, $\delta p(\mathbf{x}, t)$, related to the perturbation in the medium parameters $K, \rho, \tau, \tau_\sigma$. For algebraic simplicity, the following substitution is made in equation 2.10,

$$\tau = \frac{\tau_\epsilon}{\tau_\sigma} - 1 = \frac{2}{Q}\left(\frac{1}{Q} + \sqrt{1+\frac{1}{Q^2}}\right),$$

so that equation 2.10 becomes

$$\begin{aligned}
&\frac{\partial P}{\partial t} + K(\tau+1)\left(\nabla\cdot\mathbf{v}\right) + r_p = S(\mathbf{x}_s, t),\\
&\frac{\partial \mathbf{v}}{\partial t} + \frac{1}{\rho}\nabla P = 0,\\
&\frac{\partial r_p}{\partial t} + \frac{1}{\tau_\sigma}\left(r_p + \tau K\left(\nabla\cdot\mathbf{v}\right)\right) = 0.
\end{aligned} \tag{2.12}$$

Let ρ_0, K_0, τ_{σ_0} and τ_0 be the background medium parameters. Perturbing them by an amount $\delta\rho$, δK, $\delta\tau_\sigma$ and $\delta\tau$, respectively, gives the new medium parameters as

$$\begin{aligned}
\rho &= \rho_0 + \delta\rho,\\
K &= K_0 + \delta K,\\
\tau_\sigma &= \tau_{\sigma_0} + \delta\tau_\sigma,\\
\tau &= \tau_0 + \delta\tau.
\end{aligned} \tag{2.13}$$

The perturbed wavefields can thus be written as

$$\begin{aligned}
&\frac{\partial \delta P}{\partial t} + K(\tau+1)\left(\nabla\cdot\delta\mathbf{v}\right) + \delta r_p = -\delta K(\tau+1)\left(\nabla\cdot\mathbf{v}\right) - K\delta\tau\left(\nabla\cdot\mathbf{v}\right),\\
&\frac{\partial \delta\mathbf{v}}{\partial t} + \frac{1}{\rho}\nabla\delta P = \frac{\delta\rho}{\rho^2}\nabla P,\\
&\frac{\partial \delta r_p}{\partial t} + \frac{1}{\tau_\sigma}\left(\delta r_p + \tau K\left(\nabla\cdot\delta\mathbf{v}\right)\right) = -\frac{1}{\tau_\sigma}\left(K\delta\tau + \tau\delta K\right)\left(\nabla\cdot\mathbf{v}\right) + \frac{\delta\tau_\sigma}{\tau_\sigma^2}\left(r_p + \tau K\left(\nabla\cdot\mathbf{v}\right)\right).
\end{aligned} \tag{2.14}$$

To invert only for the perturbations in the bulk modulus δK, we make the following assumptions in equation 2.14:

- The density is constant, i.e., $\delta\rho = 0$.
- The material relaxation parameters are constant, i.e., $\delta\tau = 0$ and $\delta\tau_\sigma = 0$.

Equation 2.14 thus simplifies to

$$\begin{aligned}
&\frac{\partial \delta P}{\partial t} + K(\tau+1)\left(\nabla \cdot \delta \mathbf{v}\right) + \delta r_p = -\delta K(\tau+1)\left(\nabla \cdot \mathbf{v}\right), \\
&\frac{\partial \delta \mathbf{v}}{\partial t} + \frac{1}{\rho}\nabla \delta P = 0, \\
&\frac{\partial \delta r_p}{\partial t} + \frac{1}{\tau_\sigma}\left(\delta r_p + \tau K\left(\nabla \cdot \delta \mathbf{v}\right)\right) = -\frac{\tau}{\tau_\sigma}\delta K\left(\nabla \cdot \mathbf{v}\right).
\end{aligned} \tag{2.15}$$

In the context of Q-LSRTM, equation 2.15 is equivalent to the matrix-vector operation $\mathbf{d} = \mathbf{Lm}$. Here $\mathbf{d}$ represents the Born-modeled data with attenuation and is given by the solution of the linearized equations in equation 2.15, $\mathbf{L}$ is a linear modeling operator and $\mathbf{m}$ is related to the bulk modulus of the medium.

The adjoint equations for equation 2.12 can be derived using the adjoint-state method (shown in Appendix D) and is given by

$$\begin{aligned}
&\frac{\partial q}{\partial t} + \nabla \cdot \left(\frac{1}{\rho}\mathbf{u}\right) = -\Delta d(\mathbf{x}_g, t; \mathbf{x}_s), \\
&\frac{\partial \mathbf{u}}{\partial t} + \left[\nabla\left(K(1+\tau)q\right) + \nabla\left(\frac{1}{\tau_\sigma}K\tau s\right)\right] = 0, \\
&\frac{\partial s}{\partial t} - \frac{s}{\tau_\sigma} - q = 0.
\end{aligned} \tag{2.16}$$

Here $(q, \mathbf{u}, s)$ are the adjoint-state variables of the state variables $(P, \mathbf{v}, r_p)$ and $\Delta d(\mathbf{x}_g, t; \mathbf{x}_s)$ represents the residual pressure seismogram. For Q-LSRTM, $\Delta d(\mathbf{x}_g, t; \mathbf{x}_s)$ represents the data residual between the predicted and the observed pressure data at every iteration.

The perturbation in the image, δm, is related to the perturbation in the bulk modulus, δK, which in turn can be obtained by zero-lag cross-correlation of the adjoint fields from equation 2.16 with the background wavefields from equation 2.12

as

$$\delta m \approx \delta K = \int_0^T (1+\tau)(\nabla \cdot \mathbf{v})q + \frac{\tau}{\tau_\sigma}(\nabla \cdot \mathbf{v})s \; dt. \tag{2.17}$$

The solution to equation 2.17 is equivalent to the $\mathbf{g} = \mathbf{L}^T \Delta \mathbf{d}$ step in Q-LSRTM. A detailed derivation of the adjoint equations and the gradient is shown in Appendix D.

The next sub-section describes the algorithm for numerical implementation of Q-LSRTM.

2.4.1 Q-LSRTM algorithm

The following steps are carried out for numerically implementing Q-LSRTM by a preconditioned conjugate gradient method ([Nocedal and Wright, 1999]), where a diagonal preconditioning matrix $\mathbf{C}$ is assumed. In this work, source-side illumination ([Plessix and Mulder, 2004]) is used as the diagonal preconditioner.

- Form the misfit function, ϵ as

$$\epsilon = \frac{1}{2}||\mathbf{L}\mathbf{m}^{(i+1)} - \mathbf{d}^{obs}||^2,$$

  where $\mathbf{L}$ represents a linear modeling operator and $\mathbf{L}\mathbf{m}^{(i+1)}$ is the predicted data given by the solution to equation 2.15, $\mathbf{d}^{obs}$ represents the recorded pressure seismogram, $\mathbf{m}$ represents the reflectivity image and i represents the iteration index.

- Compute the gradient given by

$$\mathbf{g}^{(i+1)} = \mathbf{L}^T \left[\mathbf{L}\mathbf{m}^{(i+1)} - \mathbf{d}^{obs}\right] = \mathbf{L}^T \Delta \mathbf{d}^{(i+1)},$$

  where $\Delta \mathbf{d}$ represents the data residual between the predicted and the observed

data, which is backpropagated by using the adjoint equations in 2.16. The adjoint wavefields are cross-correlated with the background fields, given in equation 2.12, to give the perturbation in bulk modulus in equation 2.17 at each iteration. This perturbation can then be suitably scaled to give the perturbation in the reflectivity image, δm as

$$K = \rho v^2 \Rightarrow \delta K = 2\rho v \delta v \ (\because \delta\rho = 0)$$
$$\therefore \delta m = \frac{\delta v}{v} = \frac{\delta K}{2\rho v^2} = \frac{\delta K}{2K}.$$

- Update the gradient using the conjugate gradient formula as

$$\mathbf{dk}^{(i+1)} = \mathbf{C}\mathbf{g}^{(i+1)} + \beta \mathbf{dk}^{(i)},$$

where β is given by

$$\beta = \frac{\left(\mathbf{g}^{(i+1)}\right)^T \mathbf{C}\mathbf{g}^{(i+1)}}{\left(\mathbf{g}^{(i)}\right)^T \mathbf{C}\mathbf{g}^{(i)}}.$$

- Compute the step length α,

$$\alpha = \frac{\left(\mathbf{dk}^{(i+1)}\right)^T \mathbf{g}^{(i+1)}}{\left(\mathbf{L}\mathbf{dk}^{(i+1)}\right)^T \left(\mathbf{L}\mathbf{dk}^{(i+1)}\right)}.$$

- Iteratively update the reflectivity image,

$$\mathbf{m}^{(i+2)} = \mathbf{m}^{(i+1)} + \alpha \mathbf{dk}^{(i+1)},$$

until the length of the residual vector falls below a specified threshold.

2.5 Numerical results

The effectiveness of Q-LSRTM is now demonstrated with both synthetic and field data records from a crosswell experiment in Friendswood, Texas. The synthetic examples are for two models with strong attenuation: 1) a layered model with a shallow velocity and Q anomaly, and 2) the Marmousi model with embedded Q anomalies.

In the synthetic examples, the observed data are generated by a O(2,8) time-space domain staggered grid solution of the viscoacoustic wave-equation in equation 2.10. A Ricker wavelet with a peak frequency of 20 Hz is chosen as the source wavelet. The data are then migrated by using RTM, LSRTM, Q-RTM and Q-LSRTM. Here, RTM and Q-RTM refer to the first iteration of LSRTM and Q-LSRTM, respectively. Source-side illumination is used as the pre-conditioning factor during the least-squares iterations for LSRTM and Q-LSRTM. The standard RTM and Q-RTM images are also illumination compensated.

2.5.1 Layered velocity model

We first demonstrate the effect of strong subsurface attenuation on migrated images using a simple example of a flat layered model. Figure 2.1 shows a layered velocity model with a shallow Q anomaly. The Q value in the anomaly is 20, implying very strong attenuation for a wave propagating through this layer. To generate the synthetic data, equation 2.10 is solved for 100 shots evenly spaced at 40 m on the surface. 200 receivers, evenly distributed at 20 m on the surface, recorded the data.

Figures 2.2(a) and 2.2(b) compare the RTM and LSRTM images when the data having strong attenuation are migrated under the acoustic approximation. The LSRTM image has fewer artifacts and better balanced amplitudes in the shallow layers compared to the RTM image. However, the deeper layers below the anomaly (shown by the black arrows in Figure 2.2) have very weak amplitudes in both the

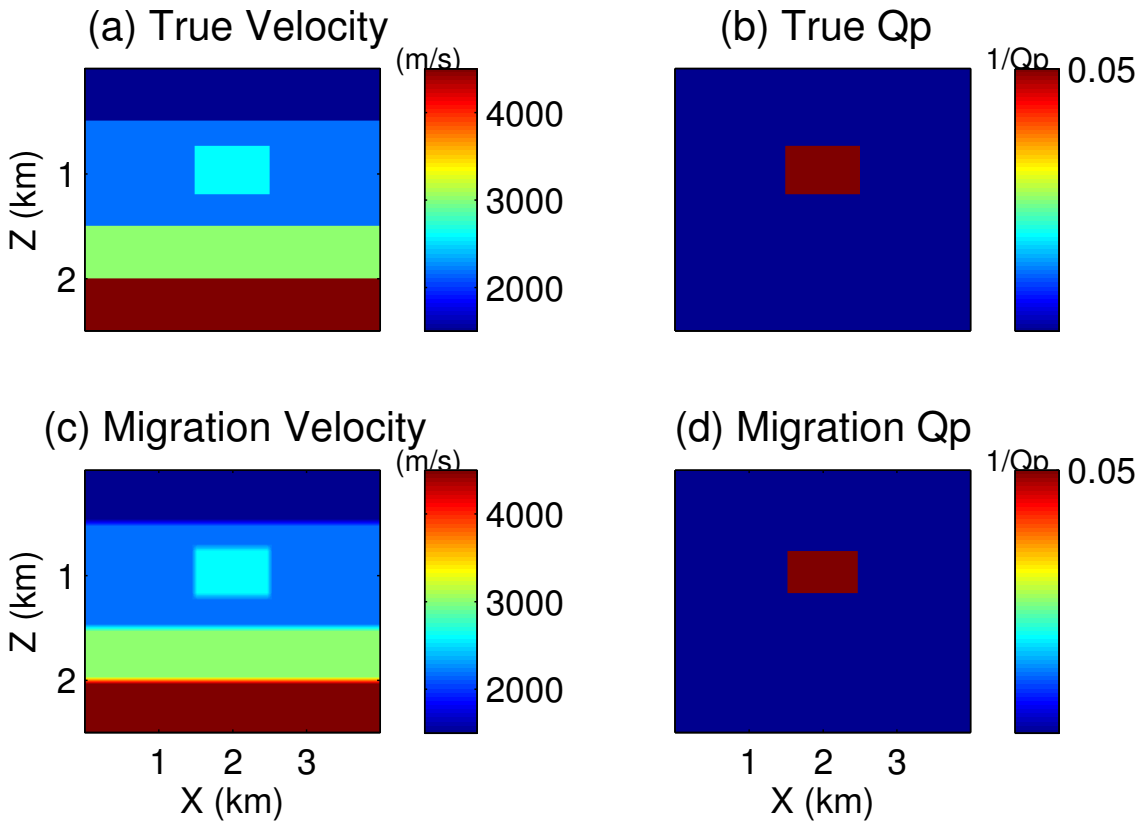

Figure 2.1: A layered model: (a) true velocity model, (b) true Q model, (c) migration velocity model, (d) Q model used for Q-RTM and Q-LSRTM.

RTM and the LSRTM images. There is also a slight mispositioning in their locations. This is because strong attenuation affects both the amplitudes and the phases of the propagating waves.

The Q-RTM and Q-LSRTM images after 20 iterations are shown in Figures 2.2(c) and 2.2(d), respectively. The shallow Q-LSRTM image is similar to that for standard LSRTM in terms of artifact mitigation and better balancing of reflector amplitudes. However, improvements with Q-LSRTM are evident at the base of the anomaly and the reflectors directly beneath it (shown by the black arrows in Figure 2.2 and in the zoom views in Figure 2.3). The reflectors are imaged at the correct locations and the migration amplitudes of these layers are more accurate than those obtained by standard LSRTM. As shown in Figure 2.3, Q-LSRTM has corrected for the amplitude loss and the phase distortion in the deeper layers due to the overlying Q anomaly.

2.5.2 Marmousi model

The Q-LSRTM method is now tested on the more complex Marmousi model. Figures 2.4(a) and 2.4(b) show the true velocity and Q models, respectively, used for generating the observed data. The migration velocity and Q models are shown in Figures 2.4(c) and 2.4(d), respectively, and the Q model is chosen such that the attenuation layers are overlying the targeted deeper anticlines. The observed synthetic data are generated with a fixed spread geometry where 200 shots are excited with a 40 m shot interval at a depth of 10 m. Each shot is recorded with 400 receivers and a 20 m receiver interval with a recording time of around 8 seconds.

Conventional acoustic RTM and LSRTM images are displayed in Figures 2.5(a) and 2.5(b), respectively. Similar to the previous example, the LSRTM image shows better resolution and fewer artifacts in the shallow layers compared to the standard RTM image. However, in the deeper layers, the amplitudes of the images from these two imaging methods are very weak. The reflectors and the anticlines cannot be

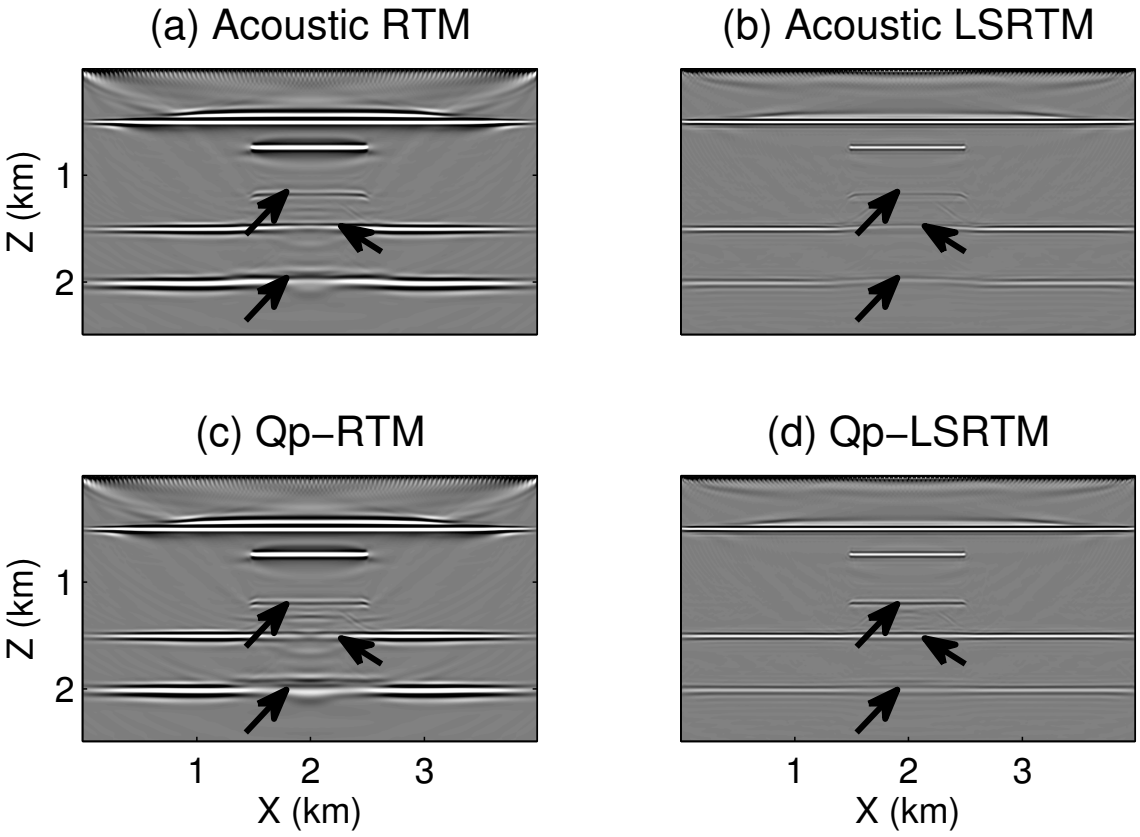

Figure 2.2: Comparison between images from (a) acoustic RTM, (b) acoustic LSRTM, (c), Q-RTM, and (d) Q-LSRTM. The black arrows point to the reflectors below the strong attenuation layer where improvements from Q-LSRTM can be seen.

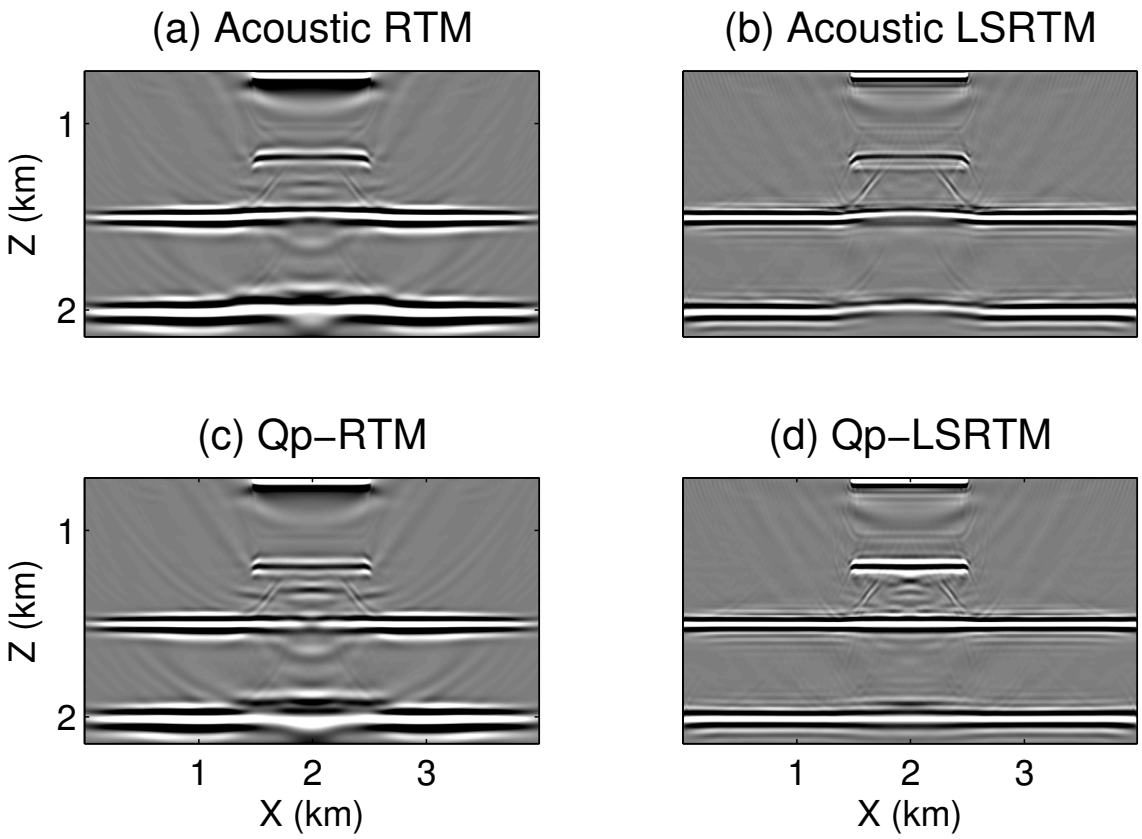

Figure 2.3: Magnified views of Figure 2.2 where all the images have been normalized such that the short reflector at the top of the Q anomaly have the same magnitude.

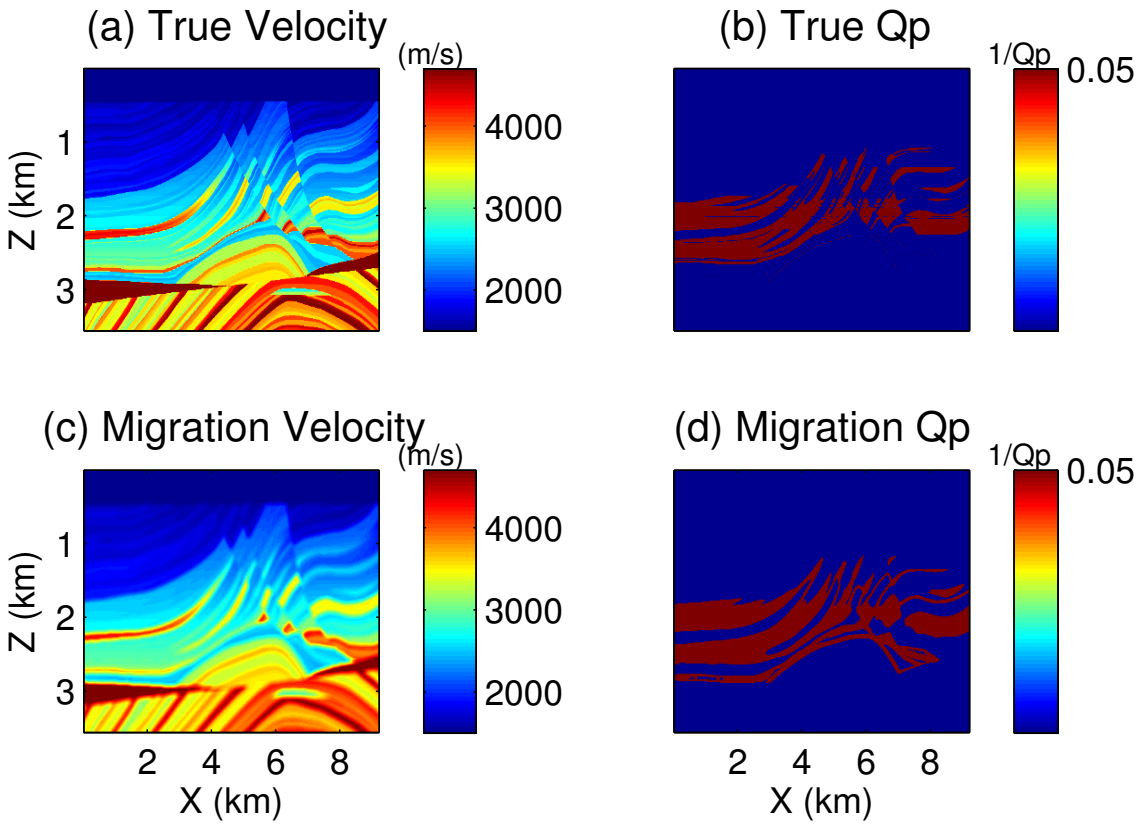

Figure 2.4: The modified Marmousi model: (a) true velocity model, (b) true Q model, (c) migration velocity model, (d) Q model used for Q-RTM and Q-LSRTM.

properly delineated in spite of using a very accurate velocity model for migration. This distortion can be attributed to the strong attenuation above these layers.

The Q-RTM image, shown in Figure 2.5(c), shows no improvements in the deeper layers compared to the standard RTM image in Figure 2.5(a). However, significant improvements can be seen in the Q-LSRTM image shown in Figure 2.5(d). In the shallow layers, the acoustic LSRTM and Q-LSRTM images are very similar in terms of the image quality. However, the amplitudes of the deeper reflectors and the anticlines are significantly better balanced in the Q-LSRTM image than in the standard RTM and LSRTM images. The zoomed views of the black and blue boxes in Figures 2.5(b) and 2.5(d) are shown in Figure 2.6. The black arrows point to the areas where noticeable improvements in the image quality can be seen with Q-LSRTM.

The residual as a function of iteration number for acoustic and Q-LSRTM is plotted in Figure 2.7. The convergence rate for Q-LSRTM is better than that for acoustic LSRTM because the correct physics of attenuation is accounted for in the forward and adjoint operators.

2.5.3 Sensitivity of Q-LSRTM to errors in the velocity model

The sensitivity of Q-LSRTM to errors in the migration velocity model are now tested. For these numerical simulations, we assume that an accurate estimate of Q is available. The errors are introduced into the migration velocity model by applying a triangle smoothing filter with increasing window lengths to the true velocity model in Figure 2.4(a). The Q-LSRTM images for the different velocity models are shown in Figure 2.8(b). It becomes evident from these figures that the blurring in the Q-LSRTM image increases with depth for velocity errors exceeding around 8%. The convergence curves, shown in Figure 2.9, show that the convergence of Q-LSRTM decreases with increasing errors in the velocity model.

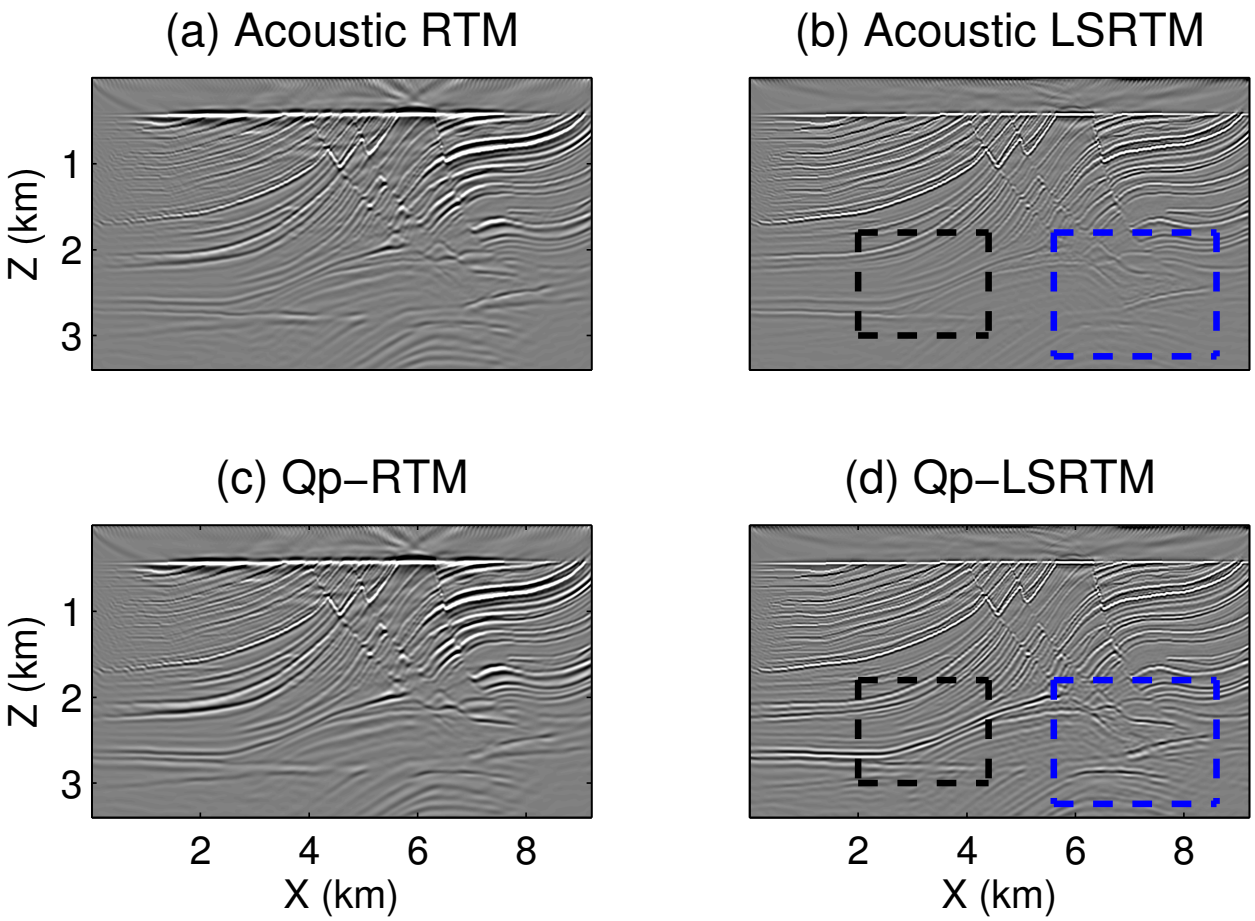

Figure 2.5: Comparison between images from (a) acoustic RTM, (b) acoustic LSRTM, (c), Q-RTM, and (d) Q-LSRTM. The black and blue boxes point to the areas for zoom views.

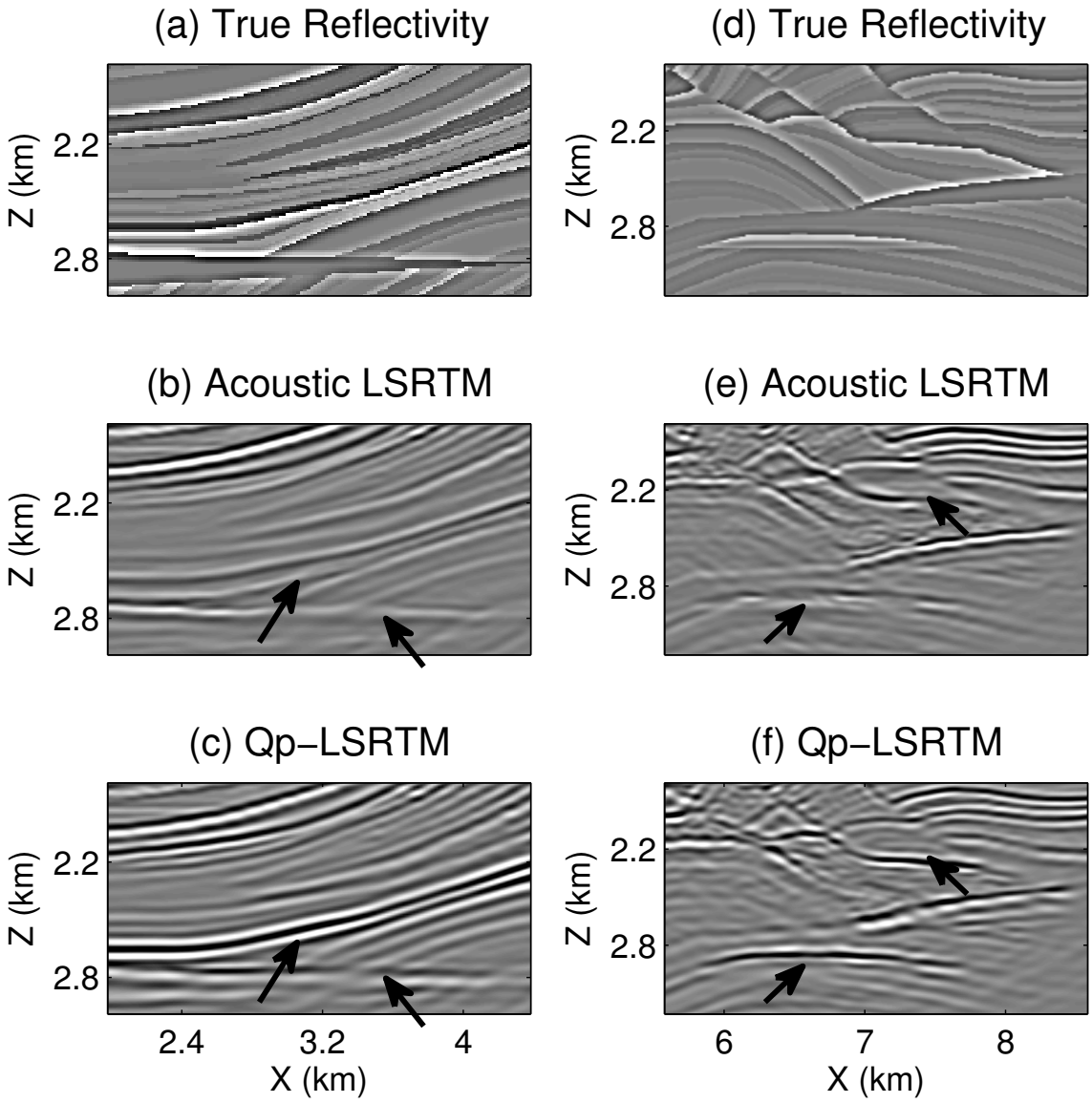

Figure 2.6: Magnified views of the black (left) and blue (right) boxes in Figure 2.5. (a), (d) True reflectivity models used only for comparison. (b), (e) acoustic LSRTM images. (c), (f) Q-LSRTM images.

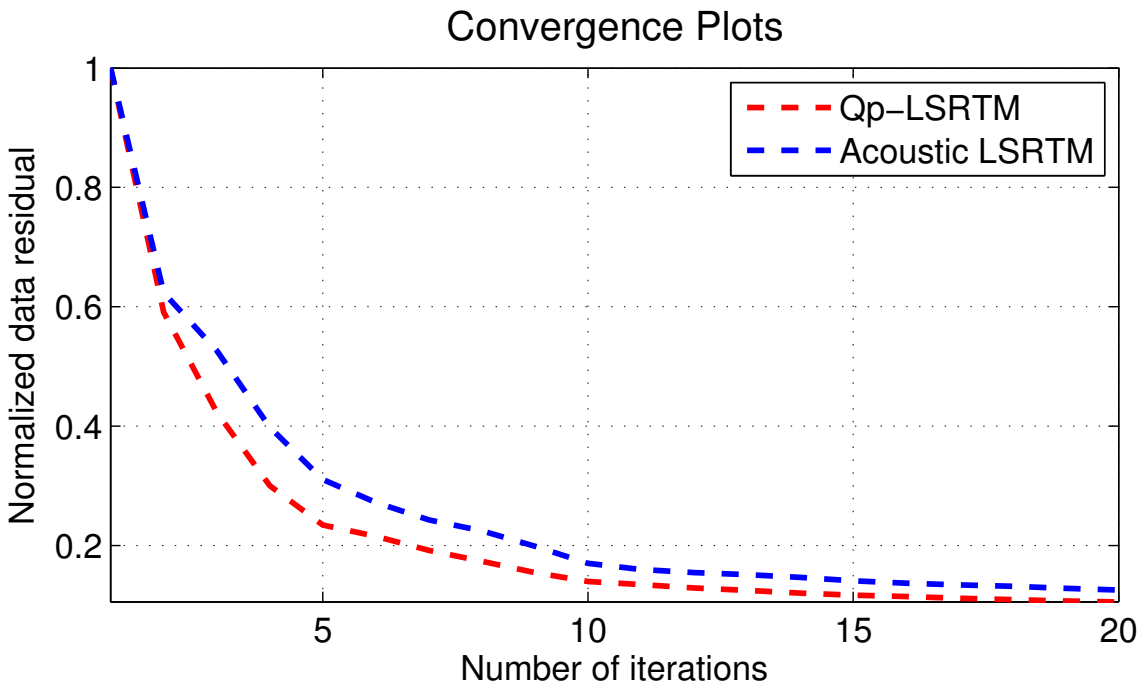

Figure 2.7: Convergence curves for acoustic and Q-LSRTM for the Marmousi model.

2.5.4 Sensitivity of Q-LSRTM to errors in the Q model

To quantify the sensitivity of Q-LSRTM to errors in the Q migration model, the Q-LSRTM images for different attenuation models are shown in Figures 2.10(b). It is evident from these figures that Q-LSRTM shows improvements over standard LSRTM when the error in the migration Q model is around 50%. For attenuation errors exceeding 50%, the amplitude loss is not compensated for by Q-LSRTM and the inverted images are very similar to the acoustic LSRTM image shown in Figure 2.5(b). The convergence curves for the different cases, shown in Figure 2.11, suggest that Q-LSRTM converges quickly even when there are significant errors in the migration Q model. This is contrary to the convergence curves shown in Figure 2.9 where the convergence of Q-LSRTM becomes slower with increasing errors in the migration velocity model. These curves suggest that the convergence rate of Q-LSRTM is more sensitive to errors in the velocity model than to errors in the Q model. However, with increasing errors in the migration Q model, the amplitude loss due to atten-

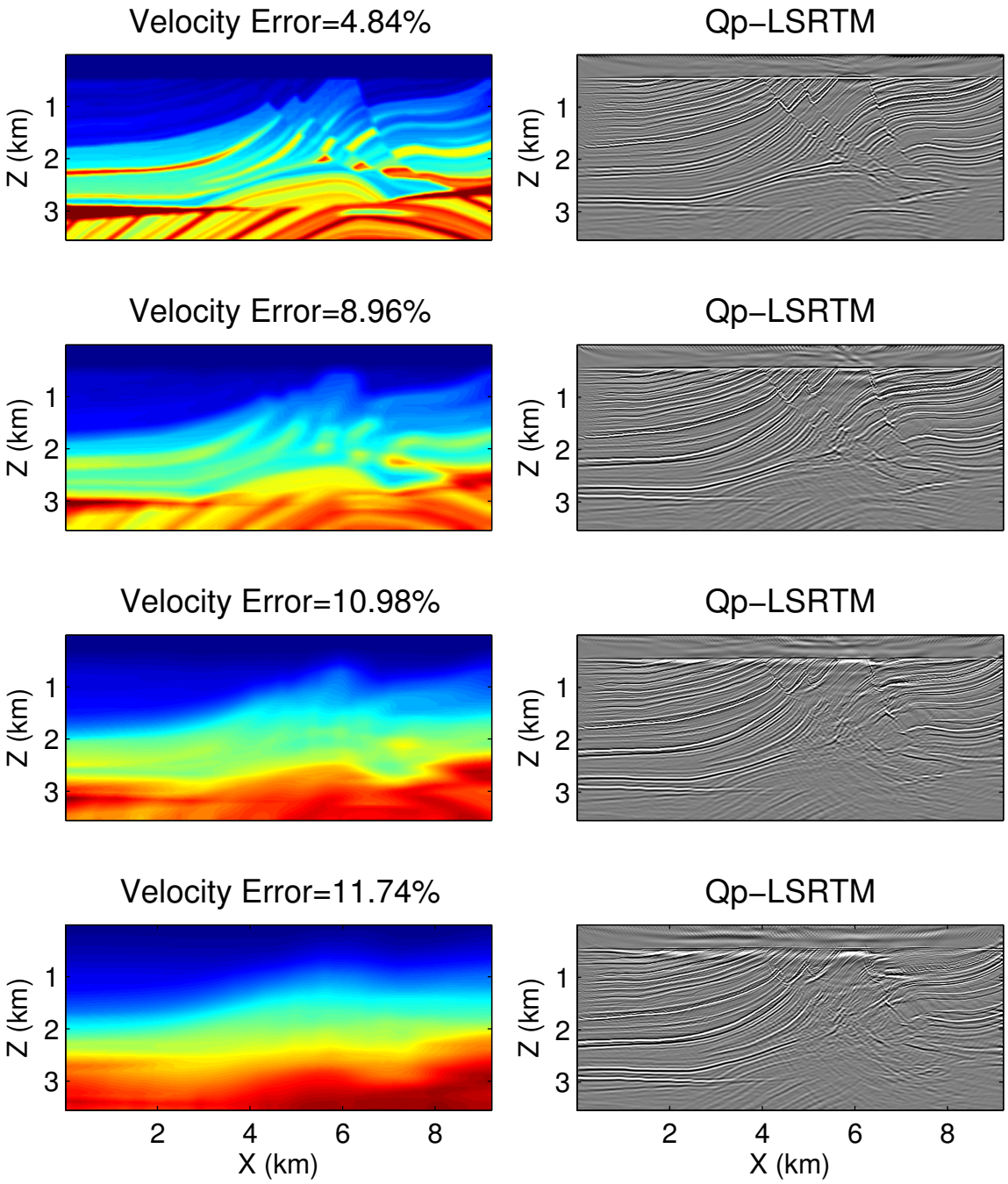

Figure 2.8: Sensitivity of Q-LSRTM to errors in the migration velocity model. The figures in the left panel show the different velocity models used for Q-LSRTM. On the right are the corresponding Q-LSRTM images after 20 iterations.

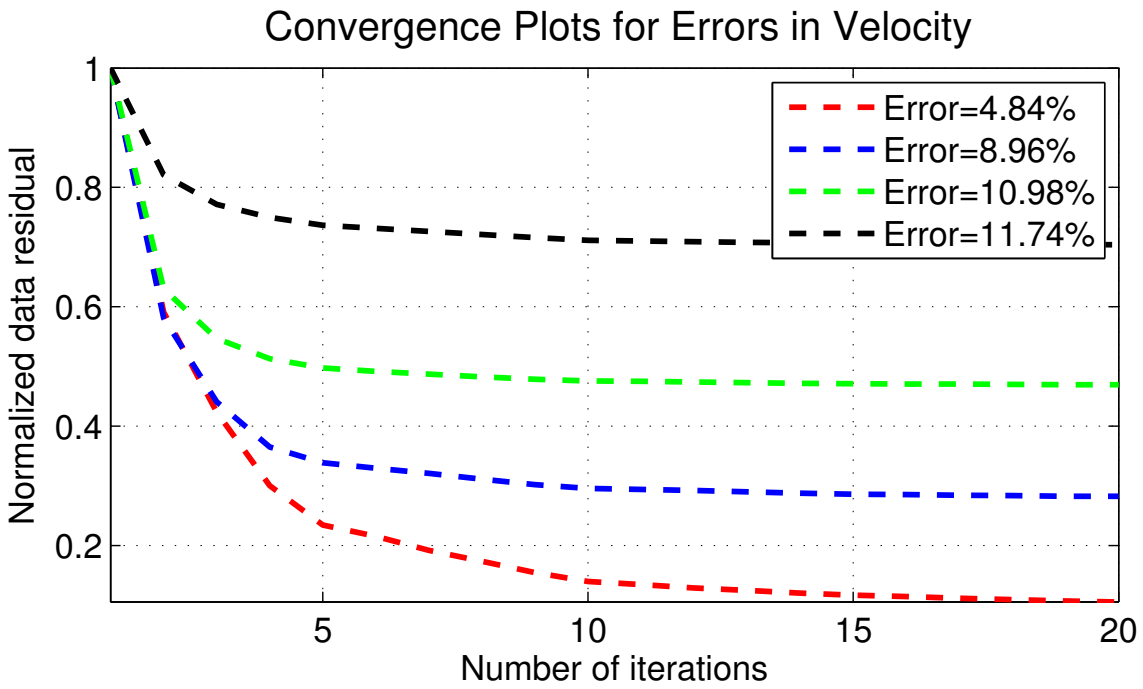

Figure 2.9: Normalized data residual vs iteration number for Q-LSRTM with different percentage errors in the migration velocity model.

uation is not corrected for by Q-LSRTM. Thus, a fairly accurate estimation of the Q distribution is required to see noticeable improvements in the image quality with Q-LSRTM.

2.5.5 Friendswood Crosswell Field Data

As a final example, Q-LSRTM is applied to the Friendswood crosswell data ([Chen et al., 1990]). Two 305 m deep cased wells separated by 183 m were used as the source and receiver wells. Downhole explosive sources of 10 g charges were fired at intervals of 3 m from 305 m to 9 m in the source well and the receiver well had 96 receivers placed at depths ranging from 293 m to 3 m. The data were recorded with a sampling interval of 0.25 ms for a total recording time of 0.375 s. The following processing steps were first applied to the data:

- The recorded data are corrected from 3D to 2D format by scaling the amplitudes

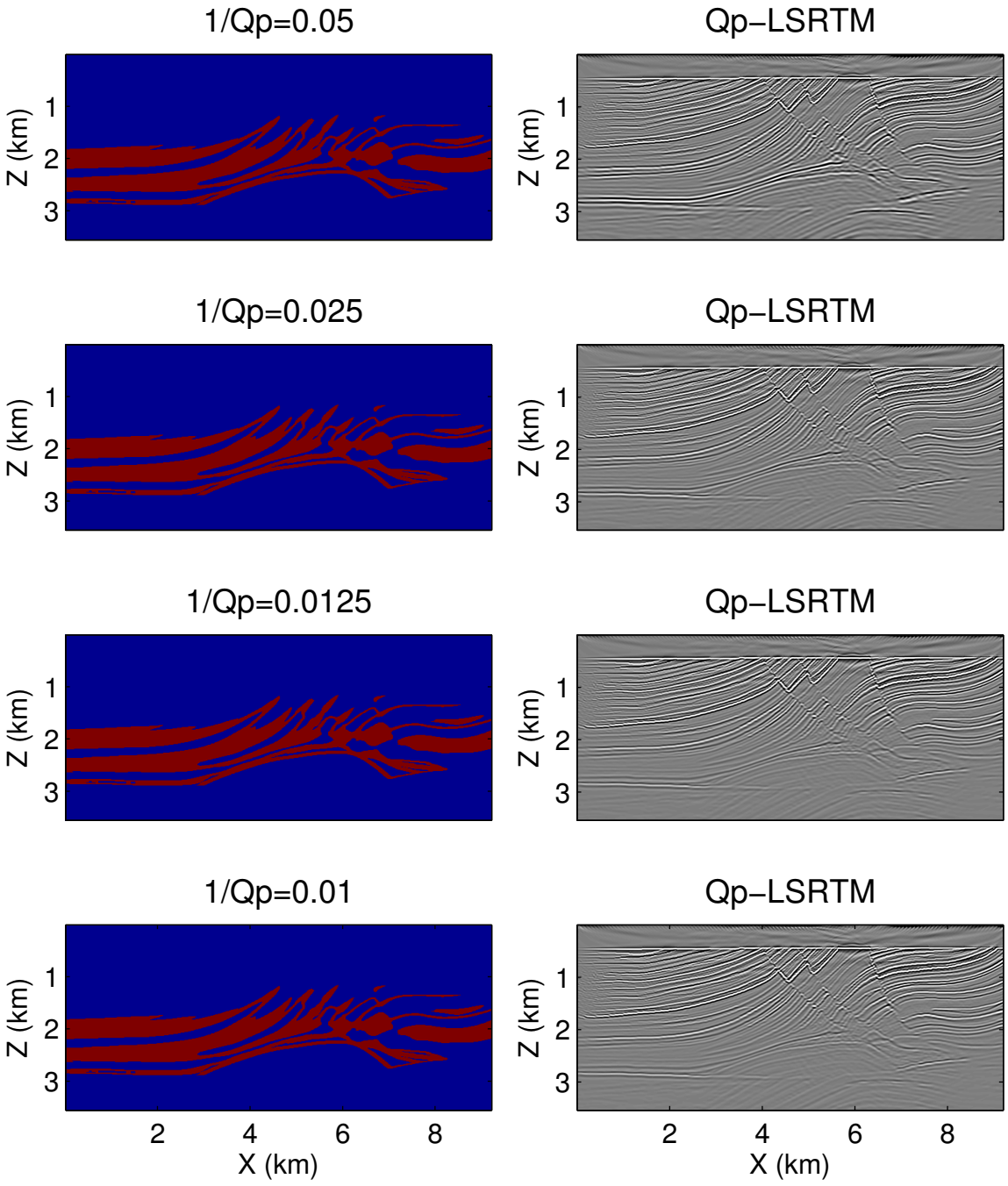

Figure 2.10: Sensitivity of Q-LSRTM to errors in the migration Q model. The figures in the left panel show the different Q models used for Q-LSRTM. The true model had a background Q value of 0.05. On the right are the corresponding Q-LSRTM images after 20 iterations for different background Q values.

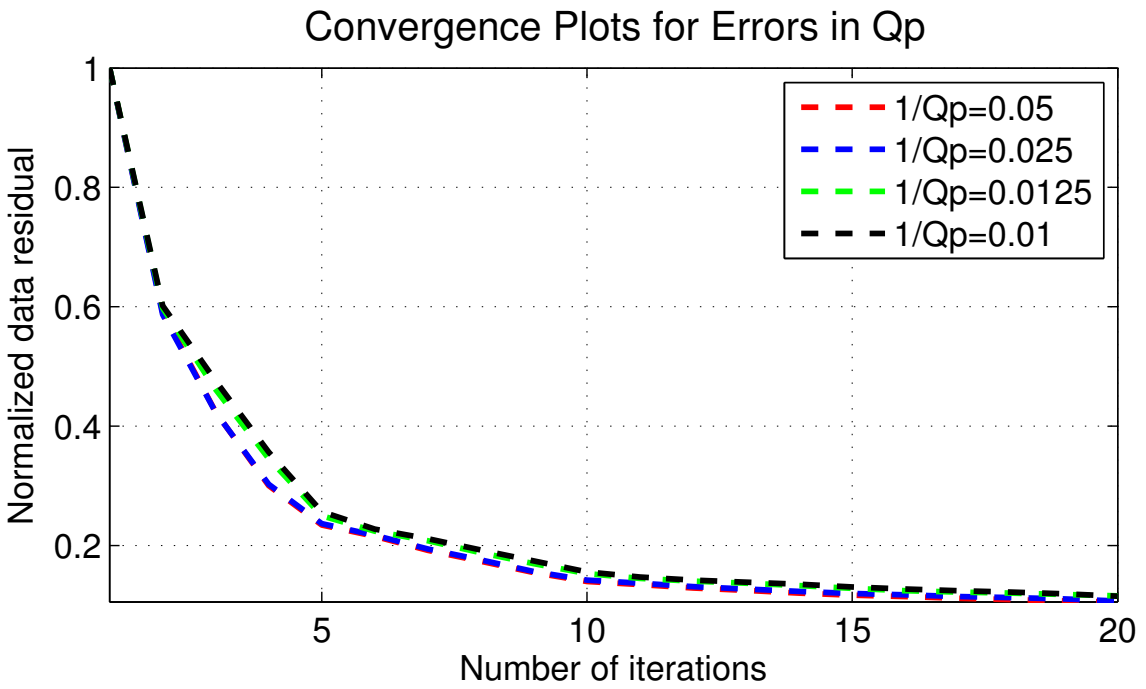

Figure 2.11: Normalized data residual vs iteration number for Q-LSRTM with different migration Q models.

by $\sqrt{t}$ to approximate geometrical spreading. A phase correction is applied by multiplying the spectrum of the observed seismogram with the filter $\sqrt{i/\omega}$ ([Zhou et al., 1995]).

- A directional 9-*point* median filter is used to eliminate the tube waves which are seen as linear events in the common shot gather (CSG) shown in Figure 2.12(a). To implement this filter, a linear moveout correction is applied to the CSG to flatten the tube wave arrivals so that they arrive at the same time. For every time sample, a 9-*trace* median filter is applied to the traces so that the output traces have highly amplified tube waves and diminished reflection events around the onset of the tube waves. These predicted tube waves are then subtracted from the original data.

- A bandpass filter of 80-600 Hz is applied to the data to remove any extreme noise from the data. The final processed CSG after applying all these processing

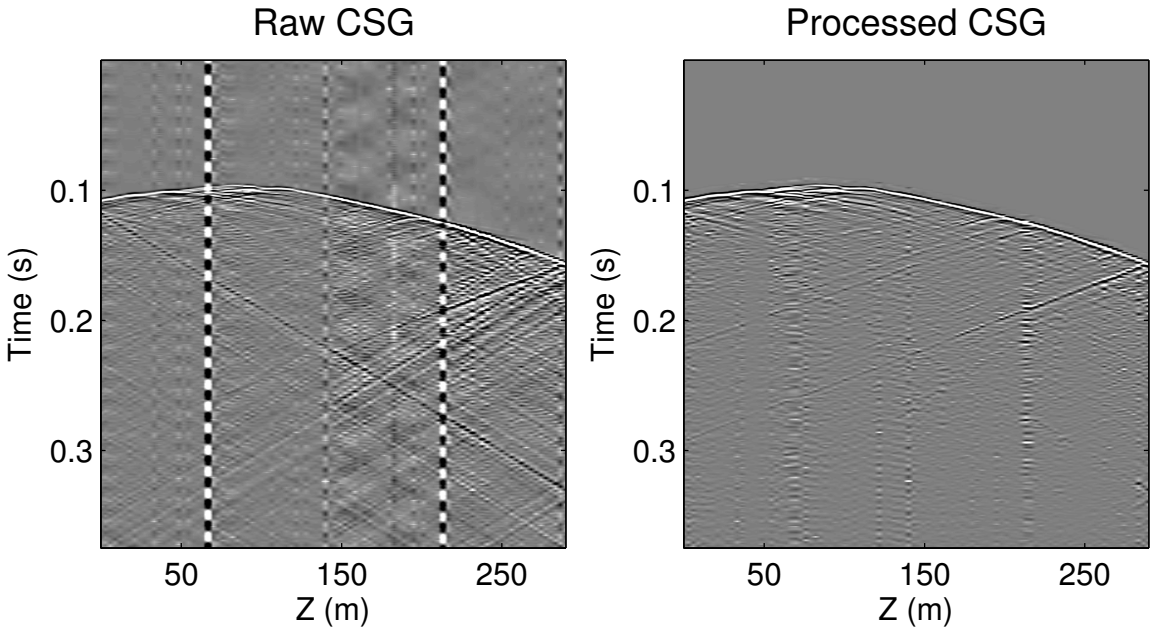

Figure 2.12: (a) A raw CSG from the Friendswood crosswell data. (b) The same CSG after applying all the data processing steps.

steps is shown in Figure 2.12(b).

- Since the bandwidth of 80-600 Hz is too broad for waveform tomography and Q-LSRTM, the data are Wiener filtered to transform the original wavelet to a Ricker wavelet with a 200 Hz peak frequency. This bandwidth is chosen based on the frequency content of the data where most of the signal is concentrated between 150-250 Hz.

- The first-arrival traveltimes are then picked from all the CSGs and inverted to get a starting velocity model for early-arrival waveform inversion. The early-arrival waveform tomogram, shown in Figure 2.13(a), is used as the migration velocity model. The wave-equation Q tomography method, discussed in the next chapter, is used to get a reference Q model for Q-RTM and Q-LSRTM. The estimated Q tomogram is shown in Figure 2.13(b).

The images from acoustic LSRTM and Q-LSRTM after 10 iterations are shown in Figures 2.14(b) and 2.14(d), respectively. The decrease in residual was around

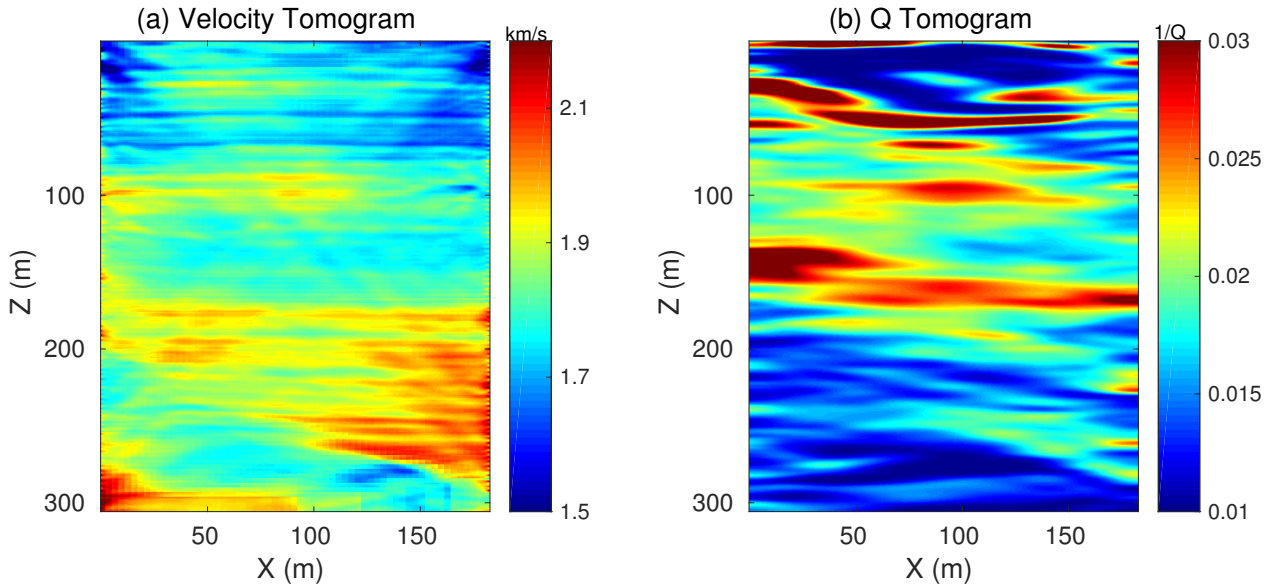

Figure 2.13: Background models used for Q-LSRTM: (a) Velocity tomogram estimated for the crosswell data using early-arrival waveform inversion. (b) Q tomogram estimated using wave-equation Q tomography.

25-30% for both the inversions. In the shallow part of the image between 25-120 m, the amplitudes of the reflectors have improved in the Q-LSRTM image compared to the acoustic LSRTM image. In the deeper part between 170-200 m, the layers have been better delineated in the Q-LSRTM image. The lateral continuity of the imaged reflectors have also improved in this area. Magnified views of the blue and black boxes in Figure 2.14, shown in Figures 2.15 and 2.16 respectively, further emphasize the improvement in the image quality from Q-LSRTM.

2.6 Discussion

The adjoint propagators used for Q-LSRTM are also attenuative and they damp the high frequency components of the residual wavefield at every iteration. The effect of this extra damping during the residual wavefield extrapolation can be seen in the magnified views of the Q-LSRTM images in Figures 2.15 and 2.16. Even though the Q-LSRTM image has better balanced amplitudes than the acoustic RTM and LSRTM

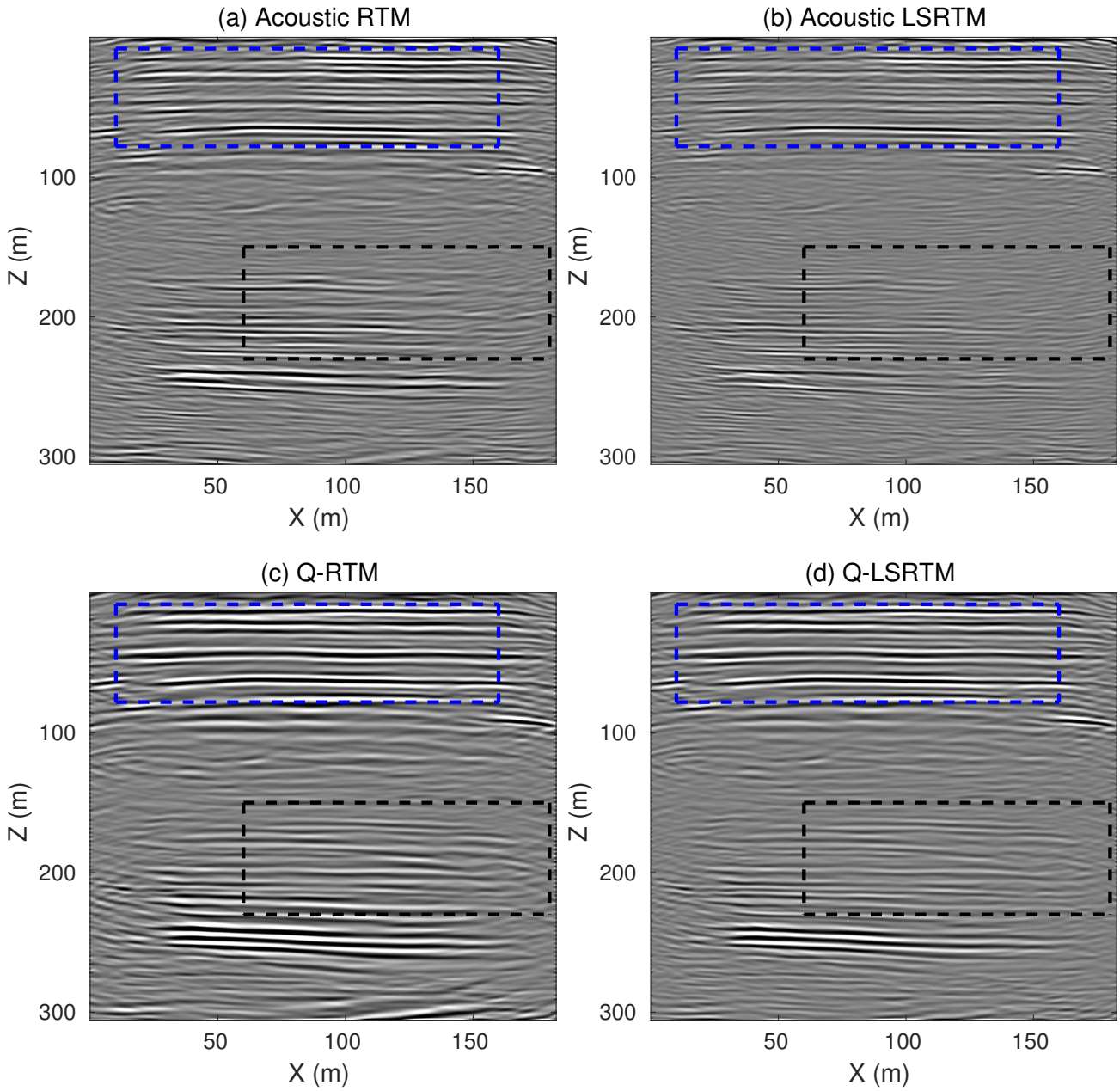

Figure 2.14: Comparison between images from (a) acoustic RTM, (b) acoustic LSRTM, (c) Q-RTM, and (d) Q-LSRTM after 10 iterations. The blue and black boxes point to the areas for zoom views.

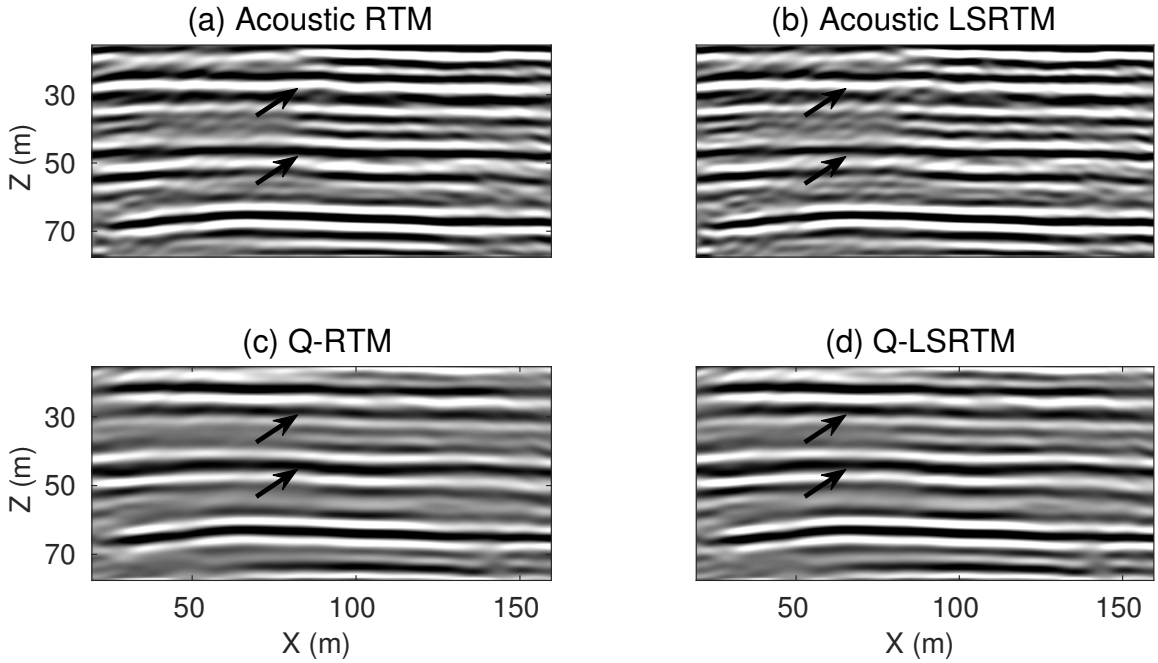

Figure 2.15: Magnified views of the blue boxes in Figure 2.14. The black arrows point to the areas where improvements in amplitude balancing can be seen with Q-LSRTM.

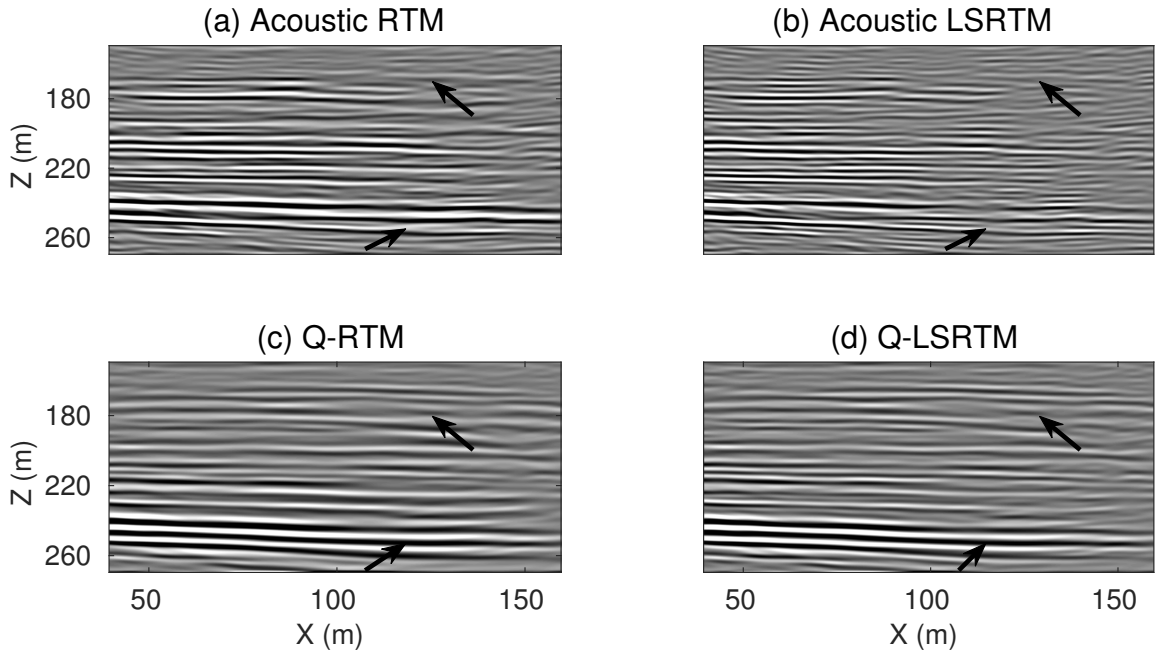

Figure 2.16: Magnified views of the black boxes in Figure 2.14. The black arrows point to the areas where improvements in amplitude balancing can be seen with Q-LSRTM.

images in areas where there is strong attenuation, the acoustic LSRTM image is richer in high wavenumber content and has better resolution than the Q-LSRTM image. To compensate for this loss in resolution, a deblurring filter ([Aoki and Schuster, 2009, Dai et al., 2011]) can be used as a preconditioner during the Q-LSRTM iterations.

Throughout the numerical tests, a single relaxation mechanism is assumed which is valid if the bandwidth of the data is narrow. For a wide bandwidth of the data, as in the case of the Friendswood crosswell data, a single relaxation mechanism may not be sufficient to accurately model the effect of Q. However, using more than one relaxation mechanism will significantly increase the computational cost of Q-LSRTM. For typical exploration problems for RTM where the bandwidth of the data used is around 5-30 Hz, use of a single relaxation mechanism should suffice ([Blanch et al., 1995]; [Zhu et al., 2013]). Also, in the numerical simulations, it is assumed that the recorded data are free from shear waves. Similar to standard LSRTM, if there are shear waves present in the data, Q-LSRTM will produce images with strong shear wave artifacts. Also, the Q-LSRTM formulation in this chapter uses the isotropic viscoacoustic wave-equation. However, the real earth is anisotropic and interbedding of thin thin attenuative layers can lead fo attenuation anisotropy ([Mavko and Nur, 1979, Carter and Kendall, 2006, Zhu and Tsvankin, 2006, Zhu et al., 2007, Zhu and Tsvankin, 2007, Behura and Tsvankin, 2009]). Studying the effect of anisotropy on Q-LSRTM will be a topic of future research.

A disadvantage of Q-LSRTM is that the computational cost per iteration is more than six times that of standard RTM. The cost increases linearly with the number of least-squares iterations. As in the case of standard least-squares migration, Q-LSRTM can be made more efficient using multisource encoded migration ([Dai and Schuster, 2009]; [Tang, 2009]; [Dai et al., 2010]; [Huang and Schuster, 2012]) where several shot gathers are randomly shifted in time to form a phase-encoded supergather.

2.7 Conclusions

A time-domain least-squares reverse time migration method is presented that uses the viscoacoustic wave-equation to compensate for the distortion in amplitudes and phases of seismic waves propagating in highly attenuative layers. Numerical results on synthetic and field data validate that if the recorded data have strong attenuation, then conventional acoustic RTM and LSRTM cannot correct for the attenuation loss. However, if the linearized viscoacoustic wave-equation and its adjoint equations are used for LSRTM, the attenuation loss can be compensated during the iterations. Results with synthetic and field data for strongly attenuative media show that LSRTM with Q compensation produces images with better balanced amplitudes and accurately positioned reflectors compared to acoustic RTM and LSRTM. Similar to standard LSRTM, Q-LSRTM is also sensitive to errors in the migration velocity model. Another input requirement for this method is an accurate estimate of the smoothly varying Q distribution in the subsurface. Accurate estimation of Q and the linearized Q-LSRTM method proposed in this chapter have the potential for accurate imaging in highly attenuative geological environments such as gas-sandstones and shales.

Chapter 3

Wave-equation Q tomography

3.1 Summary

A wave-equation gradient optimization method is presented that inverts for the subsurface Q distribution that minimizes a skeletonized misfit function ϵ. Here, ϵ is the sum of the squared differences between the observed and the predicted peak/centroid-frequency shifts of the early arrivals. The gradient is computed by migrating the observed traces weighted by the frequency-shift residuals. The background Q model is perturbed until the predicted and the observed traces have the same peak frequencies or the same centroid frequencies. Numerical tests show that an improved accuracy of the Q model by wave-equation Q inversion (WQ) leads to a noticeable improvement in the migration-image quality.

3.2 Introduction

The real earth is anelastic and distorts the amplitude and the phase of a propagating seismic wave ([Aki and Richards, 1980]). Attenuation of P-waves can be quantified by a quality factor Q which accounts for the phase shift as a function of the frequency content of the propagating waves and the distance traveled. Lower values of Q imply more energy loss of the wave per cycle or higher attenuation.

The amplitude loss due to attenuation is often compensated for during prestack depth migration (PSDM). For example, [Xin et al., 2008] and [Xie et al., 2009] compensated for the attenuation loss by ray-tracing methods. [Dai and West, 1994], [Yu et al., 2002], [Wang, 2008] and [Valenciano et al., 2011] used one-way wave-equation migration in the frequency domain for attenuation compensation. For reverse time migration, [Zhang et al., 2010], [Suh et al., 2012], [Fletcher et al., 2012], [Bai et al., 2013], [Zhu et al., 2014] and [Zhu and Harris, 2015] proposed different visco-acoustic wave equations with separate controls over phase and amplitude to compensate for the attenuation loss. [Dutta and Schuster, 2014] and [Sun et al., 2015a] used anelastic least-squares RTM schemes to successfully compensate for the amplitude loss and phase distortion because of Q during imaging.

Besides velocity, an additional input requirement for these Q-based migration algorithms is a reliable estimate of the background Q model. A Q model can be either estimated using data-domain or image-domain based tomographic techniques. In the data-domain, [Brzostowski and McMechan, 1992] used the attenuation of amplitudes as the input data for Q tomography. In contrast, [Quan and Harris, 1997] used the frequency-shifts between the predicted and the observed traces and smeared the shifts along raypaths to update the Q model. A similar adaptive centroid frequency-shift approach was also used by [He et al., 2013]. These frequency-shift methods rely on the high-frequency assumption made in classical ray-based tomography. Alternative data-domain approaches include using a FWI-like algorithm where an objective function is set up to invert for a Q model that minimizes the L2 norm of the residual between the observed and the predicted data ([Bai and Yingst, 2013]; [Wang and Zhang, 2014]).

For the image-domain techniques, a residual image is usually minimized, which is the difference between the image computed by the background Q model and a target image, which is attenuation-free ([Shen et al., 2014]; [Shen and Zhu, 2015]; [Shen et al., 2015]). The image perturbation is then related to the perturbation in Q

using wave-equation based tomography operators. An interpretative image-domain based Q tomography technique was developed by [Zhou et al., 2011] where they generated amplitude ratio maps by comparing the reflection amplitudes associated with a lossy layer with that of a reference horizon that remains unaffected. Through ray-tracing, they accumulate the attenuation effects along raypaths and a Q volume is then estimated from tomographic inversion.

In this chapter, I present a new skeletonized wave-equation Q inversion method that is based on minimizing the difference between the peak frequencies of the observed and the predicted transmission arrivals. The peak frequencies are obtained from the amplitude spectra of the traces, which are a skeletonized representation of the data. The Fréchet derivative is derived using the implicit function theorem and the gradient is numerically obtained by a zero-lag cross-correlation between the forward propagated source wavefield and the backpropagated residual traces. A residual trace is obtained by weighting the observed trace by the frequency shift between the trace and its corresponding predicted trace. Unlike conventional ray-based Q tomography, the residuals in wave-equation Q tomography are smeared along transmission wavepaths ([Woodward, 1992]) computed from finite-difference solutions to the time-domain visco-acoustic wave equation characterized by the standard linear solid (SLS) mechanism ([Christensen, 1982]; [Carcione et al., 1988]; [Blanch et al., 1995]). The proposed approach has no high-frequency assumptions about the data unlike ray-based tomography methods. It is also less susceptible to cycle-skipping problems associated with any FWI-like algorithm where the amplitude and the phase differences between the predicted and the observed traces are minimized to obtain the Q model.

This chapter is organized into five sections. After the introduction, the second section describes the theory of wave-equation Q tomography. Numerical results on synthetic and field data are then presented in the third section. The limitations of

the proposed method are discussed in the fourth section and the conclusions are in the last section.

3.3 Theory of wave-equation Q inversion

The key steps in wave-equation Q inversion are similar to the wave-equation traveltime inversion algorithm proposed in [Luo and Schuster, 1991a] and [Luo and Schuster, 1991b]. These steps can be generalized for any type of skeletonized data and are the following: (1) define a connective function that connects the frequency-shift residual with the pressure seismogram, (2) define a phase-misfit function, and (3) derive the perturbation of the misfit function with respect to Q using the connective function and the visco-acoustic wave equation.

In my analysis, I use the 2D time-domain visco-acoustic wave-equation in chapter 2 given by

$$\begin{aligned}
&\frac{\partial P}{\partial t} + K(\tau+1)\left(\nabla\cdot\mathbf{v}\right) + r_p = S(\mathbf{x}_s,t),\\
&\frac{\partial \mathbf{v}}{\partial t} + \frac{1}{\rho}\nabla P = 0,\\
&\frac{\partial r_p}{\partial t} + \frac{1}{\tau_\sigma}\left(r_p + \tau K\left(\nabla\cdot\mathbf{v}\right)\right) = 0.
\end{aligned} \tag{3.1}$$

Here, the variables have the same definition as in the previous chapter and the relation between τ and Q is given by

$$\tau = \frac{2}{Q}\left(\frac{1}{Q} + \sqrt{1+\frac{1}{Q^2}}\right). \tag{3.2}$$

Figures 3.1(a)-3.1(c) show the variation of the relaxation parameters τ_σ, τ_ϵ and τ with Q. It can be seen from these figures that high values of τ imply strong attenuation while low values indicate weak attenuation. Also, for realistic geological models, Q varies from 20-200. In this range, τ has a wider variation in its values than τ_σ and τ_ϵ.

Thus, for the parameterization, τ is used since it is quite sensitive to small changes in Q.

In the next sub-section, we define a connective function that connects the change in peak frequency of an arrival with the observed and the predicted pressure seismograms.

3.3.1 Connective function

Let $\tilde{P}_f(\mathbf{x}_r, t; \mathbf{x}_s)$ denote a predicted event for a given background Q model recorded at the receiver location $\mathbf{x}_r$ due to a source excited at time $t = 0$ and at location $\mathbf{x}_s$. f is the peak frequency of this event that can be obtained from its amplitude spectrum (shown by the red curve in Figure 3.2). Similarly, let $P_{f-\Delta f}(\mathbf{x}_r, t; \mathbf{x}_s)$ denote the same event in the observed data (the spectrum of this event is shown by the blue curve in Figure 3.2). Δf is the shift between the peak-frequencies of the predicted and the observed traces because of Q.

For the right background velocity model, the similarity between the amplitude-normalized observed and predicted traces in Figure 3.3 can be written as

$$\begin{aligned} F_f(\mathbf{x}_r, t; \mathbf{x}_s) &= \int dt \;\; \frac{P_{f-\Delta f}(\mathbf{x}_r, t; \mathbf{x}_s)}{A_1(\mathbf{x}_r; \mathbf{x}_s)} \frac{\tilde{P}_f(\mathbf{x}_r, t; \mathbf{x}_s)}{A_2(\mathbf{x}_r; \mathbf{x}_s)}, \\ &= \int dt \;\; \frac{P_{f-\Delta f}(\mathbf{x}_r, t; \mathbf{x}_s)}{A(\mathbf{x}_r; \mathbf{x}_s)} \tilde{P}_f(\mathbf{x}_r, t; \mathbf{x}_s). \end{aligned} \tag{3.3}$$

Here $A_1(\mathbf{x}_r; \mathbf{x}_s)$ and $A_2(\mathbf{x}_r; \mathbf{x}_s)$ are the amplitude normalization factors for the observed and the predicted events, respectively, and $A(\mathbf{x}_r; \mathbf{x}_s) = A_1(\mathbf{x}_r; \mathbf{x}_s) A_2(\mathbf{x}_r; \mathbf{x}_s)$. These factors normalize the events to a maximum amplitude of 1 such that only the phase mismatch or the shift in peak-frequency is emphasized during the inversion.

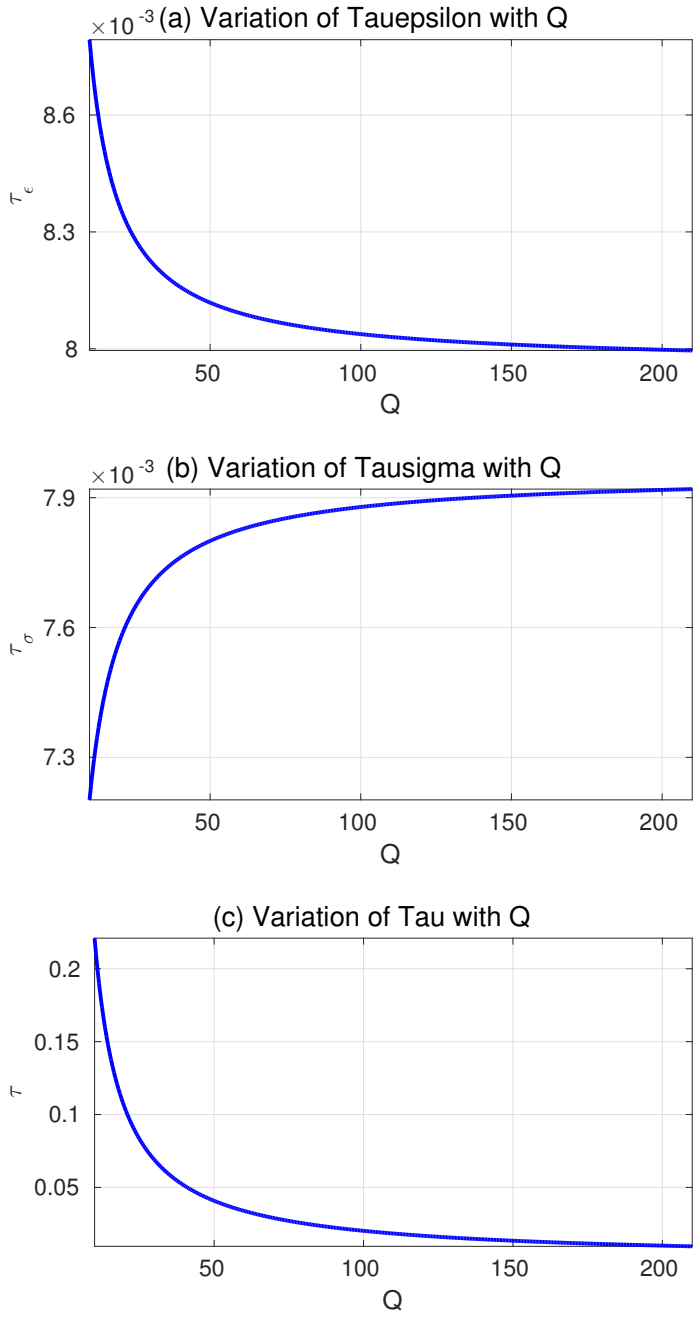

Figure 3.1: Variation of the parameters (a) τ_ϵ, (b) τ_σ, and (c) τ for different values of Q. The central frequency of the source wavelet is taken to be 20 Hz.

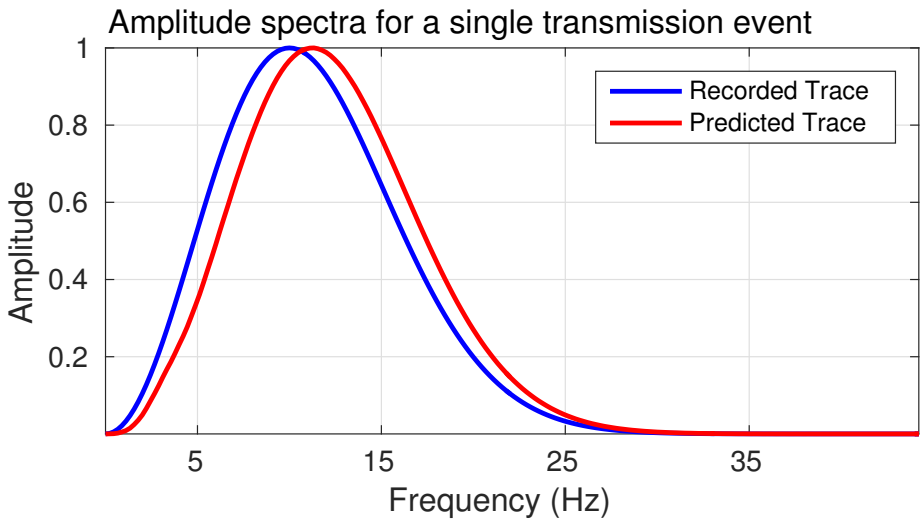

Figure 3.2: Comparison between the amplitude spectra of a single transmission arrival in a predicted and an observed trace.

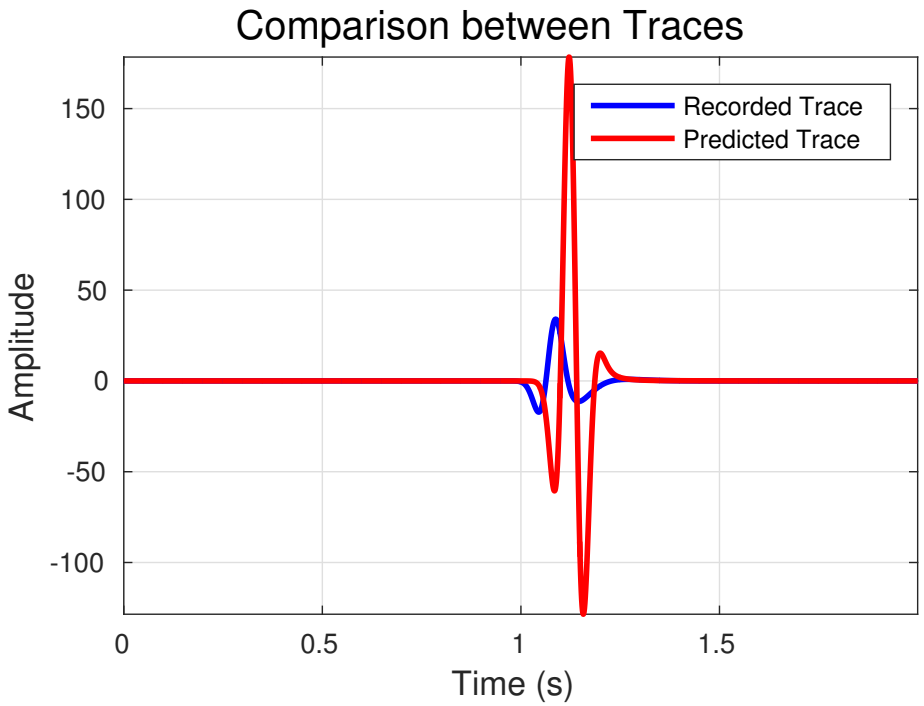

Figure 3.3: Comparison between a predicted and an observed trace for a single transmission arrival.

In WQ, the peak-frequency shift between an observed and a predicted trace is minimized. For the right background Q model, the predicted and the observed arrivals will have the same peak frequency, i.e. $\Delta f = 0$, and the normalized cross-correlation

function in equation 3.3 is maximized. The derivative of $F_f(\mathbf{x}_r, t; \mathbf{x}_s)$ with respect to f should then be zero at $f = \Delta f$. Thus,

$$\begin{aligned}\dot{F}_{\Delta f} &= \left[\frac{\partial F_f(\mathbf{x}_r, t; \mathbf{x}_s)}{\partial f}\right]_{f=\Delta f} \\ &= \int dt \;\; \frac{\dot{P}_{f-\Delta f}(\mathbf{x}_r, t; \mathbf{x}_s)}{A(\mathbf{x}_r; \mathbf{x}_s)} \tilde{P}_f(\mathbf{x}_r, t; \mathbf{x}_s) = 0, \end{aligned} \tag{3.4}$$

where $\dot{P}_f(\mathbf{x}_r, t; \mathbf{x}_s) = \partial P_f(\mathbf{x}_r, t; \mathbf{x}_s)/\partial f$. The derivatives over the amplitude normalization terms have been ignored here and will be ignored in all subsequent steps since only the peak-frequency shifts are used to update the background Q model. Equation 3.4 is the connective function which will be later used to derive the Fréchet derivative of τ.

3.3.2 Misfit function

The WQ method attempts to invert for a Q model or an equivalent τ model which predicts pressure seismograms $\tilde{P}_f(\mathbf{x}_r, t; \mathbf{x}_s)$ that minimize the misfit function

$$\epsilon = \frac{1}{2} \sum_s \sum_r \Delta f(\mathbf{x}_r, \mathbf{x}_s)^2, \tag{3.5}$$

where Δf is defined in the previous section and the summation in equation 3.5 is over all sources and receivers. The gradient $\gamma(\mathbf{x})$ is given by

$$\gamma(\mathbf{x}) = -\frac{\partial \epsilon}{\partial \tau(\mathbf{x})} = -\sum_s \sum_r \frac{\partial \Delta f}{\partial \tau(\mathbf{x})} \Delta f(\mathbf{x}_r, \mathbf{x}_s). \tag{3.6}$$

From equation 3.4 we can get the following 3 equations,

i)

$$\begin{aligned}
&\dot{F}_{\Delta f}(\Delta f, \tau(\mathbf{x})) = 0, \\
\Rightarrow &\frac{\partial \dot{F}_{\Delta f}}{\partial \Delta f}\frac{\partial \Delta f}{\partial \tau(\mathbf{x})} + \frac{\partial \dot{F}_{\Delta f}}{\partial \tau(\mathbf{x})} = 0, \\
\Rightarrow &\frac{\partial \Delta f}{\partial \tau(\mathbf{x})} = -\frac{\frac{\partial \dot{F}_{\Delta f}}{\partial \tau(\mathbf{x})}}{\frac{\partial \dot{F}_{\Delta f}}{\partial \Delta f}},
\end{aligned} \tag{3.7}$$

ii)

$$\frac{\partial \dot{F}_{\Delta f}}{\partial \Delta f} = \int dt \frac{\ddot{P}_{f-\Delta f}(\mathbf{x}_r, t; \mathbf{x}_s)}{A(\mathbf{x}_r; \mathbf{x}_s)} \tilde{P}_f(\mathbf{x}_r, t; \mathbf{x}_s), \tag{3.8}$$

iii)

$$\frac{\partial \dot{F}_{\Delta f}}{\partial \tau(\mathbf{x})} = \int dt \frac{\dot{P}_{f-\Delta f}(\mathbf{x}_r, t; \mathbf{x}_s)}{A(\mathbf{x}_r; \mathbf{x}_s)} \frac{\partial \tilde{P}_f(\mathbf{x}_r, t; \mathbf{x}_s)}{\partial \tau(\mathbf{x})}. \tag{3.9}$$

Using equation 3.7, the gradient in equation 3.6 can be written as

$$\gamma(\mathbf{x}) = \sum_s \sum_r \frac{\frac{\partial \dot{F}_{\Delta f}}{\partial \tau(\mathbf{x})}}{\frac{\partial \dot{F}_{\Delta f}}{\partial \Delta f}} \Delta f(\mathbf{x}_r, \mathbf{x}_s). \tag{3.10}$$

The Fréchet derivative $\frac{\partial \tilde{P}_f(\mathbf{x}_r, t; \mathbf{x}_s)}{\partial \tau(\mathbf{x})}$ is now derived in the next subsection.

3.3.3 Fréchet derivative

To obtain the Fréchet derivative of the pressure field with respect to the perturbation in $\tau(\mathbf{x})$, I linearize the visco-acoustic wave equation in equation 3.1. A perturbation of $\tau \rightarrow \tau + \delta\tau$ will produce perturbed wavefields $\delta\tilde{P}_f$, $\delta\mathbf{v}$ and δr_p which satisfy the

linearized visco-acoustic wave equation given by ([Dutta and Schuster, 2014])

$$\begin{aligned}
&\frac{\partial \delta \tilde{P}_f}{\partial t} + K(\tau+1)\left(\nabla \cdot \delta \mathbf{v}\right) + \delta r_p = -K\delta\tau\left(\nabla \cdot \mathbf{v}\right), \\
&\frac{\partial \delta \mathbf{v}}{\partial t} + \frac{1}{\rho}\nabla \delta \tilde{P}_f = 0, \\
&\frac{\partial \delta r_p}{\partial t} + \frac{1}{\tau_\sigma}\left(\delta r_p + \tau K\left(\nabla \cdot \delta \mathbf{v}\right)\right) = -\frac{K}{\tau_\sigma}\delta\tau\left(\nabla \cdot \mathbf{v}\right).
\end{aligned} \tag{3.11}$$

Using the Green's functions $g_P(\mathbf{x}_r, t; \mathbf{x}, 0)$ and $g_{r_p}(\mathbf{x}_r, t; \mathbf{x}, 0)$, equation 3.11 can also be expressed as

$$\delta \tilde{P}_f(\mathbf{x}_r, t; \mathbf{x}_s) = -\left(K\left(\mathbf{x}\right)(g_P(\mathbf{x}_r, t; \mathbf{x}, 0) * \nabla \cdot \mathbf{v}(\mathbf{x}, t; \mathbf{x}_s)) + \frac{K(\mathbf{x})}{\tau_\sigma(\mathbf{x})}\left(g_{r_p}(\mathbf{x}_r, t; \mathbf{x}, 0) * \nabla \cdot \mathbf{v}(\mathbf{x}, t; \mathbf{x}_s)\right)\right) \delta\tau(\mathbf{x}), \tag{3.12}$$

where $*$ denotes convolution in time. Dividing by $\delta\tau(\mathbf{x})$ on both sides we get,

$$\frac{\partial \tilde{P}_f(\mathbf{x}_r, t; \mathbf{x}_s)}{\partial \tau(\mathbf{x})} = -\left(K(\mathbf{x})(g_P(\mathbf{x}_r, t; \mathbf{x}, 0) * \nabla \cdot \mathbf{v}(\mathbf{x}, t; \mathbf{x}_s)) + \frac{K(\mathbf{x})}{\tau_\sigma(\mathbf{x})}(g_{r_p}(\mathbf{x}_r, t; \mathbf{x}, 0) * \nabla \cdot \mathbf{v}(\mathbf{x}, t; \mathbf{x}_s))\right). \tag{3.13}$$

Here, $g_p(\mathbf{x}, t; \mathbf{x}_s)$ and $g_{r_p}(\mathbf{x}, t; \mathbf{x}_s)$ are the pressure and the memory variable Green's functions, respectively, and $\mathbf{v}$ satisfies the system of equations $\mathbf{Sw} = \mathbf{F}$ in equation 3.1. Equation 3.9 can now be written as

$$\begin{aligned}
\frac{\partial \dot{F}_{\Delta f}}{\partial \tau(\mathbf{x})} &= \int dt \;\; \frac{\dot{P}_{f-\Delta f}(\mathbf{x}_r, t; \mathbf{x}_s)}{A(\mathbf{x}_r; \mathbf{x}_s)} \frac{\partial \tilde{P}_f(\mathbf{x}_r, t; \mathbf{x}_s)}{\partial \tau(\mathbf{x})} \\
&= -\int dt \;\; K(\mathbf{x})\left(g_P(\mathbf{x}_r, t; \mathbf{x}, 0) * \nabla \cdot \mathbf{v}(\mathbf{x}, t; \mathbf{x}_s) + \frac{1}{\tau_\sigma(\mathbf{x})} g_{r_p}(\mathbf{x}_r, t; \mathbf{x}, 0) * \nabla \cdot \mathbf{v}(\mathbf{x}, t; \mathbf{x}_s)\right) \frac{\dot{P}_{f-\Delta f}(\mathbf{x}_r, t; \mathbf{x}_s)}{A(\mathbf{x}_r; \mathbf{x}_s)}.
\end{aligned} \tag{3.14}$$

Substituting equations 3.8 and 3.14 into equation 3.10, the gradient $\gamma(\mathbf{x})$ can be expressed as

$$\begin{aligned}
\gamma(\mathbf{x}) &= \sum_s \sum_r \frac{\frac{\partial \dot{F}_{\Delta f}}{\partial \tau(\mathbf{x})}}{\frac{\partial \dot{F}_{\Delta f}}{\partial \Delta f}} \Delta f(\mathbf{x}_r, \mathbf{x}_s) \\
&= -\sum_s \sum_r \frac{\int dt \;\; K(\mathbf{x}) \left(g_P(\mathbf{x}_r, t; \mathbf{x}, 0) * \nabla \cdot \mathbf{v}(\mathbf{x}, t; \mathbf{x}_s) + \frac{1}{\tau_\sigma(\mathbf{x})} g_{r_p}(\mathbf{x}_r, t; \mathbf{x}, 0) * \nabla \cdot \mathbf{v}(\mathbf{x}, t; \mathbf{x}_s) \right)}{\int dt \frac{\ddot{P}_{f-\Delta f}(\mathbf{x}_r, t; \mathbf{x}_s)}{A(\mathbf{x}_r; \mathbf{x}_s)} \tilde{P}_f(\mathbf{x}_r, t; \mathbf{x}_s)} \\
&\qquad \frac{\dot{P}_{f-\Delta f}(\mathbf{x}_r, t; \mathbf{x}_s)}{A(\mathbf{x}_r; \mathbf{x}_s)} \Delta f(\mathbf{x}_r, \mathbf{x}_s) \\
&= -\frac{K(\mathbf{x})}{E} \sum_s \sum_r \int dt \;\; \left(g_P(\mathbf{x}_r, t; \mathbf{x}, 0) * \nabla \cdot \mathbf{v}(\mathbf{x}, t; \mathbf{x}_s) + \frac{1}{\tau_\sigma(\mathbf{x})} g_{r_p}(\mathbf{x}_r, t; \mathbf{x}, 0) * \nabla \cdot \mathbf{v}(\mathbf{x}, t; \mathbf{x}_s) \right) \\
&\qquad \dot{P}_{f-\Delta f}(\mathbf{x}_r, t; \mathbf{x}_s) \Delta f(\mathbf{x}_r, \mathbf{x}_s),
\end{aligned} \tag{3.15}$$

where,

$$E = \int dt \;\; \ddot{P}_{f-\Delta f}(\mathbf{x}_r, t; \mathbf{x}_s) \tilde{P}_f(\mathbf{x}_r, t; \mathbf{x}_s). \tag{3.16}$$

Let $\Delta P_f(\mathbf{x}_r, t; \mathbf{x}_s) = \dot{P}_{f-\Delta f}(\mathbf{x}_r, t; \mathbf{x}_s) \Delta f(\mathbf{x}_r, \mathbf{x}_s)$ denote the data residual obtained by weighting the observed traces with the residual frequency-shifts. In my present implementation, the frequency derivative over $P_{f-\Delta f}(\mathbf{x}_r, t; \mathbf{x}_s)$ is ignored while calculating the data residual. Equation 3.15 now becomes

$$\begin{aligned}
\frac{\partial \epsilon}{\partial \tau(\mathbf{x})} = -\frac{K(\mathbf{x})}{E} \sum_s \sum_r \int dt \;\; &\left(g_P(\mathbf{x}_r, t; \mathbf{x}, 0) * \nabla \cdot \mathbf{v}(\mathbf{x}, t; \mathbf{x}_s) + \frac{1}{\tau_\sigma(\mathbf{x})} g_{r_p}(\mathbf{x}_r, t; \mathbf{x}, 0) * \nabla \cdot \mathbf{v}(\mathbf{x}, t; \mathbf{x}_s) \right) \\
&\Delta P_f(\mathbf{x}_r, t; \mathbf{x}_s).
\end{aligned} \tag{3.17}$$

Using the identity,

$$\int dt\,[f(t) * g(t)]\,h(t) = \int dt\; g(t)\,[f(-t) * h(t)]\,, \tag{3.18}$$

equation 3.17 can be re-written as

$$\begin{aligned}
\frac{\partial \epsilon}{\partial \tau(\mathbf{x})} &= -\frac{K(\mathbf{x})}{E}\sum_s\sum_r\int dt\;\; (\overbrace{\nabla\cdot\mathbf{v}(\mathbf{x},t;\mathbf{x}_s)}^{source}\overbrace{(g_P(\mathbf{x}_r,-t;\mathbf{x},0)*\Delta P_f(\mathbf{x}_r,t;\mathbf{x}_s))}^{backpropagated\ residual} \\
&\qquad + \frac{1}{\tau_\sigma(\mathbf{x})}\overbrace{\nabla\cdot\mathbf{v}(\mathbf{x},t;\mathbf{x}_s)}^{source}\overbrace{\big(g_{r_p}(\mathbf{x}_r,-t;\mathbf{x},0)*\Delta P_f(\mathbf{x}_r,t;\mathbf{x}_s)\big)}^{backpropagated\ residual}) \\
&= -\frac{K(\mathbf{x})}{E}\sum_s\int dt\;\; (\overbrace{\nabla\cdot\mathbf{v}(\mathbf{x},t;\mathbf{x}_s)}^{source}\sum_r\overbrace{(g_P(\mathbf{x}_r,-t;\mathbf{x},0)*\Delta P_f(\mathbf{x}_r,t;\mathbf{x}_s))}^{backpropagated\ residual} \\
&\qquad + \frac{1}{\tau_\sigma(\mathbf{x})}\overbrace{\nabla\cdot\mathbf{v}(\mathbf{x},t;\mathbf{x}_s)}^{source}\sum_r\overbrace{\big(g_{r_p}(\mathbf{x}_r,-t;\mathbf{x},0)*\Delta P_f(\mathbf{x}_r,t;\mathbf{x}_s)\big)}^{backpropagated\ residual}) \\
&= -\frac{K(\mathbf{x})}{E}\sum_s\int dt\;\; \left(\nabla\cdot\mathbf{v}(\mathbf{x},t;\mathbf{x}_s)q(\mathbf{x},t;\mathbf{x}_s) + \frac{1}{\tau_\sigma(\mathbf{x})}\nabla\cdot\mathbf{v}(\mathbf{x},t;\mathbf{x}_s)s(\mathbf{x},t;\mathbf{x}_s)\right) \\
&= -\frac{K(\mathbf{x})}{E}\sum_s\int dt\;\; \nabla\cdot\mathbf{v}(\mathbf{x},t;\mathbf{x}_s)\left(q(\mathbf{x},t;\mathbf{x}_s) + \frac{s(\mathbf{x},t;\mathbf{x}_s)}{\tau_\sigma(\mathbf{x})}\right).
\end{aligned} \tag{3.19}$$

Here, q and s are the adjoint-state variables of P and r_p, respectively. The gradient in equation 3.19 can be numerically computed by a zero-lag cross-correlation of a forward propagated source wavefield term $\nabla\cdot\mathbf{v}(\mathbf{x},t;\mathbf{x}_s)$ and backpropagated residual wavefield terms q and s. The residual wavefield is computed by weighting the observed traces with their corresponding frequency shifts with the predicted traces. These weighted traces can be used as an adjoint source from the receiver side. An alternative derivation of the gradient for WQ using the adjoint-state method is shown in Appendix E.

For diving waves, where there are more than one event present, the frequency spectra for an observed trace and a predicted trace are shown in Figure 3.4. It is evident that in such cases, the peak-frequency for the arrivals cannot be accurately

estimated since each arrival has its own peak frequency. In such cases, the objective function in equation 3.5 can be modified to minimize the centroid frequency shifts between the predicted and the observed traces as

$$\begin{aligned}\epsilon &= \frac{1}{2}\sum_s\sum_r \Delta f_{centroid}(\mathbf{x}_r,\mathbf{x}_s)^2,\\ &= \frac{1}{2}\sum_s\sum_r \left(\frac{\int_0^f f_p A(f_p)}{\int_0^f A(f_p)} - \frac{\int_0^f f_o A(f_o)}{\int_0^f A(f_o)}\right)^2.\end{aligned} \tag{3.20}$$

The subscripts o and p stand for observed and predicted, respectively, and $A(f)$ is the amplitude for a frequency f. The connective function, the Fréchet derivative and the gradient can be similarly derived as shown in the previous sub-sections.

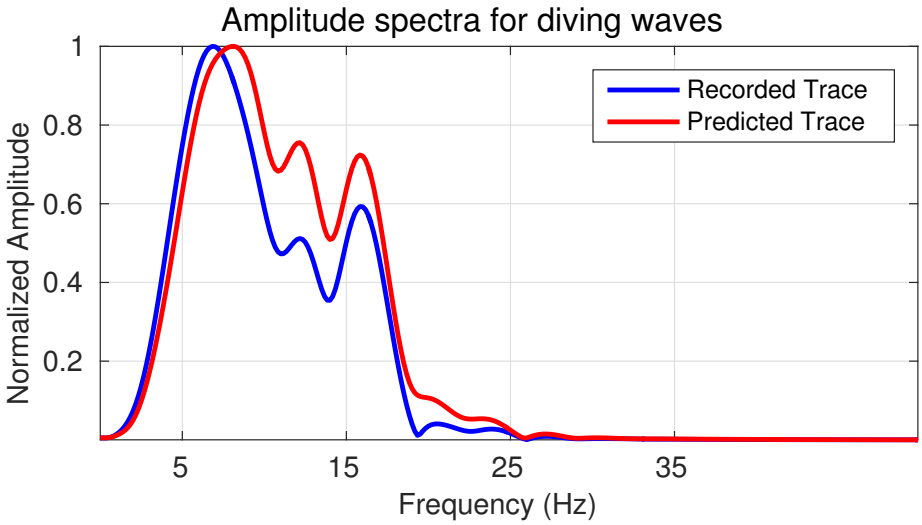

Figure 3.4: Comparison between the amplitude spectra of predicted and observed traces for diving waves.

3.3.4 WQ algorithm

The following steps are carried out for numerically implementing WQ using a preconditioned gradient-based method, where the preconditioner is a source-side illumination factor.

- Form the misfit function ϵ as

$$\epsilon = \frac{1}{2}\sum_s \sum_r \Delta f(\mathbf{x}_r, \mathbf{x}_s)^2. \tag{3.21}$$

- Compute the gradient by

$$\frac{\partial \epsilon}{\partial \tau(\mathbf{x})} = -\frac{K(\mathbf{x})}{E}\sum_s \int dt \;\; (\overbrace{\nabla \cdot \mathbf{v}(\mathbf{x}, t; \mathbf{x}_s)}^{source}) \sum_r \overbrace{(g_P(\mathbf{x}_r, -t; \mathbf{x}, 0) * \nabla \cdot \mathbf{v}(\mathbf{x}, t; \mathbf{x}_s))}^{backpropagated\ residual}$$
$$+ \frac{1}{\tau_\sigma(\mathbf{x})} \overbrace{\nabla \cdot \mathbf{v}(\mathbf{x}, t; \mathbf{x}_s)}^{source} \sum_r \overbrace{\left(g_{r_p}(\mathbf{x}_r, -t; \mathbf{x}, 0) * \nabla \cdot \mathbf{v}(\mathbf{x}, t; \mathbf{x}_s)\right)}^{backpropagated\ residual}). \tag{3.22}$$

- Estimate the step-length α by any backtracking line-search method ([Nocedal and Wright, 199

- Update the tau model $\tau(\mathbf{x})$ using the iterative steepest descent formula:

$$\tau(\mathbf{x})^{(k+1)} = \tau(\mathbf{x})^{(k)} - \alpha P(\mathbf{x}) \frac{\partial \epsilon}{\partial \tau(\mathbf{x})}, \tag{3.23}$$

  where k represents the iteration index and $P(\mathbf{x})$ is the preconditioning factor. At every iteration, the background $\tau(\mathbf{x})$ is updated and the update in $\tau(\mathbf{x})$ is then mapped to $Q(\mathbf{x})$ using equation 3.2.

3.4 Numerical Results

The effectiveness of WQ is now demonstrated with synthetic and field data records from a crosswell experiment in Friendswood, Texas. The synthetic examples are for two models with strong attenuation: (1) a crosswell example with a Gaussian Q anomaly, and (2) a surface-seismic example where there are shallow Q anomalies that

hinder the imaging of the reflectors below the anomaly.

In the synthetic examples, the observed data are generated by an $O(2,8)$ time-space-domain staggered-grid solution of the viscoacoustic-wave equation in equation 3.1. A Ricker wavelet with a peak frequency of 15 Hz is used as the source wavelet.

3.4.1 Crosswell Gaussian Q model

Figure 3.5(a) shows a homogeneous model with a velocity of 2 km/s. A Gaussian Q anomaly is embedded at the center of the model. The maximum Q at the center of the anomaly is 40 and the source and the receiver wells are offset by 4 km. There are 60 evenly spaced sources in the source well and 200 evenly spaced receivers in the receiver well. For WQ, the starting Q model is taken to be homogeneous with $Q = 1000$. Figure 3.5(b) shows the final Q tomogram obtained from WQ. It is evident that the Gaussian Q anomaly is successfully reconstructed by WQ.

3.4.2 Surface-seismic model

The WQ method is now tested on a more complex 2D section of the 3D SEG/EAGE overthrust model. Figures 3.6(a) and 3.6(b) show the true velocity and Q models, respectively, used for generating the observed data. A smooth version of the true velocity model, shown in Figure 3.6(c), is used as the background velocity model for WQ. The observed data are generated by 200 shots evenly distributed on the surface. 400 receivers at an offset of 30 m on the surface recorded the data for 6 seconds.

For WQ, the diving waves in the predicted and the observed traces are separated using the water velocity. The objective function in equation 3.20 is used to update the background Q model. The inverted Q tomogram after 35 iterations is shown in Figure 3.6(d). It is evident that WQ can successfully recover the low-intermediate wavenumber details of the background Q model. Comparison between the peak-

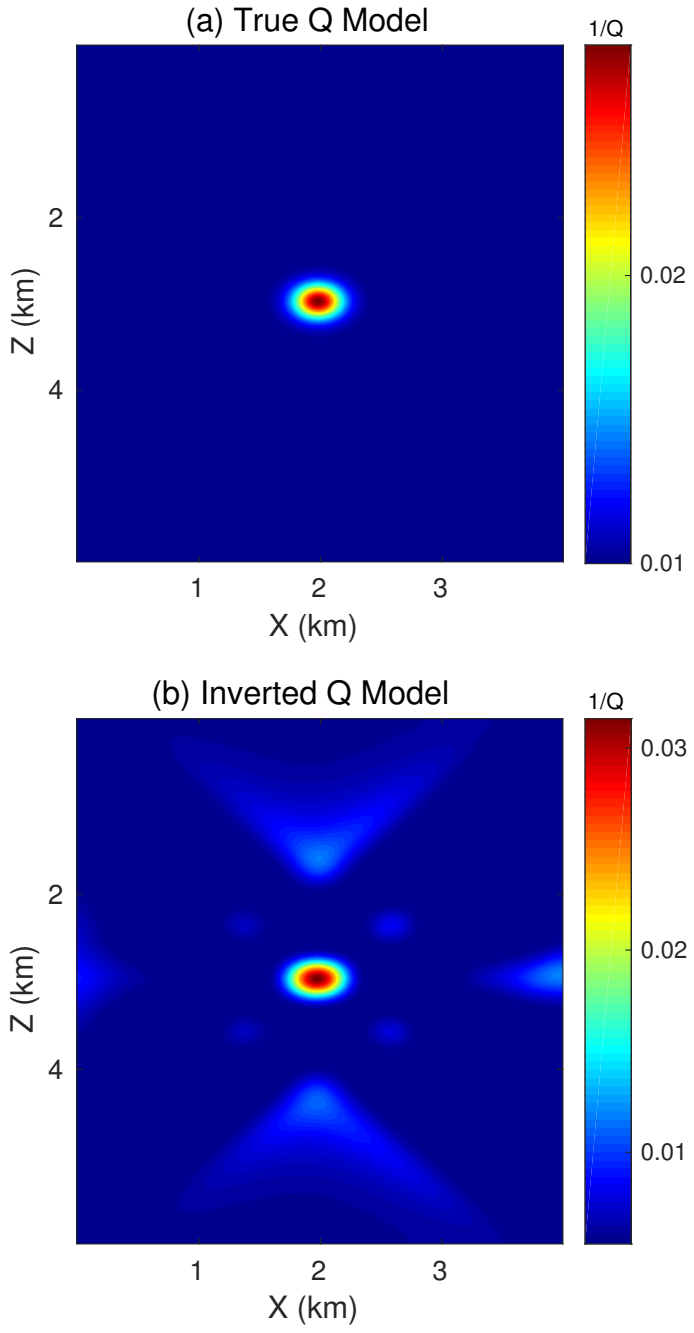

Figure 3.5: (a) True Q model, and (b) inverted Q model using WQ.

frequencies for different source-receiver pairs in the observed and the predicted traces after WQ, shown in Figure 3.7, also validate the accuracy of the inverted Q tomogram.

Q-LSRTM ([Dutta and Schuster, 2014]) is now used to check the fidelity of the inverted Q tomogram. The acoustic RTM and LSRTM images are shown in Figures 3.8(a) and 3.8(b), respectively. The black boxes in these figures indicate the areas that have been affected by the shallow Q anomalies. The amplitudes as well as the phases of the events in these areas have been distorted because of attenuation. The Q-RTM and Q-LSRTM images are shown in Figures 3.8(c) and 3.8(d), respectively. The amplitude and the phase distortions are corrected in the Q-LSRTM image. This example demonstrates that the inverted Q tomogram from WQ can be used as the background Q model for any Q-PSDM algorithm to obtain images with better resolution and better balanced amplitudes than standard migration techniques.

3.4.3 Friendswood crosswell field data

As a final example, WQ is applied to the Friendswood crosswell data ([Chen et al., 1990]). Two 305 m deep cased wells separated by 183 m were used as the source and receiver wells. Downhole explosive sources of 10 g charges were fired at intervals of 3 m from 305 m to 9 m in the source well and the receiver well had 96 receivers placed at depths ranging from 293 m to 3 m. The data were recorded with a sampling interval of 0.25 ms for a total recording time of 0.375 s. The same processing steps as described in Chapter 2 are used to process the data for WQ. The picked first-arrival traveltimes from the recorded data are shown in Figure 3.9(a). The early-arrival waveform tomogram, shown in Figure 3.10(a), is used as the background velocity model for WQ.

Figures 3.9(a)-3.9(b) show the traveltime picks and the peak frequency picks for the raw field data set. It can be seen that the low peak frequencies along the diagonal elements in this figure correlate with the high traveltime/low velocity regions in

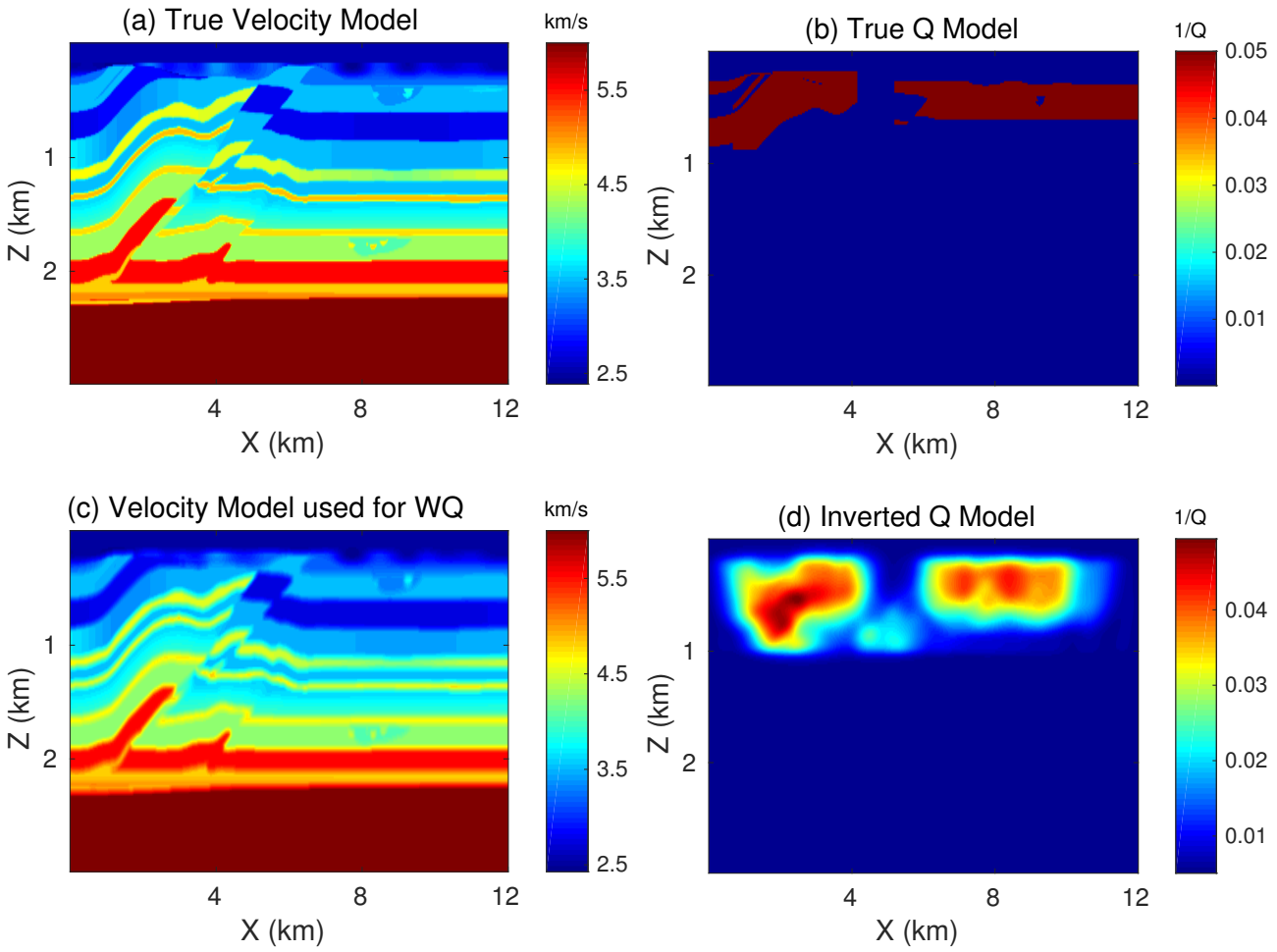

Figure 3.6: (a) True velocity, and (b) Q models used for generating the observed data, (c) velocity model used for WQ, and (d) inverted Q tomogram from WQ.

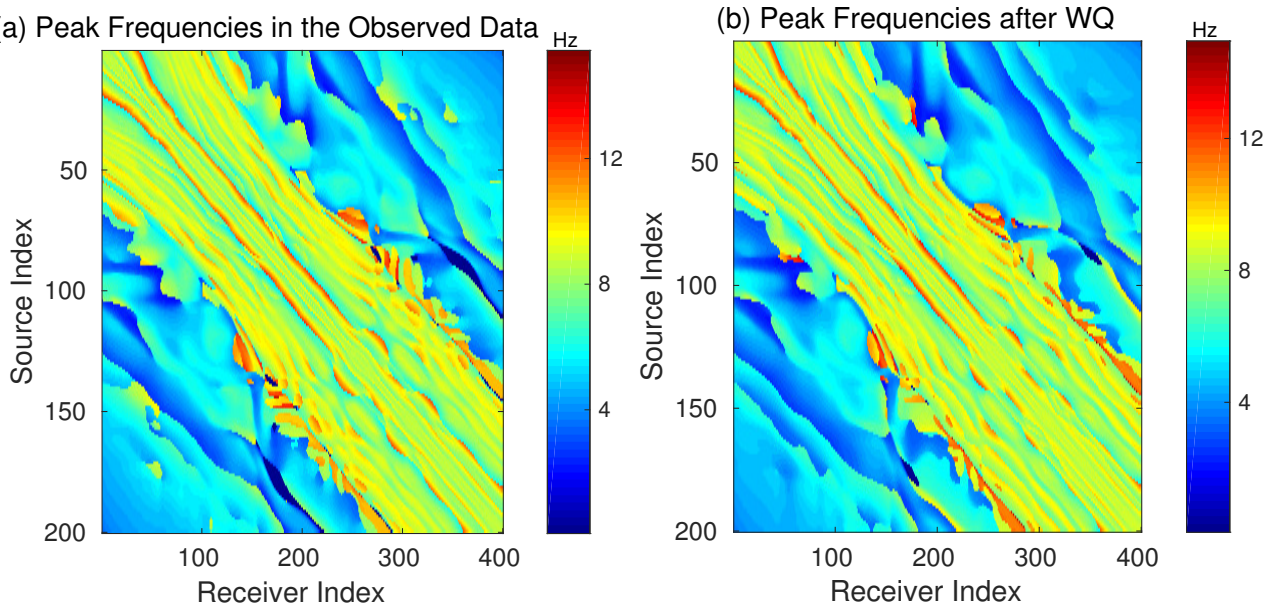

Figure 3.7: The peak-frequencies for different source-receiver pairs in (a) the observed data, and (b) the predicted data from WQ.

the model. This suggests that the low velocity formations in this model have high attenuation or low Q values.

For WQ, the starting Q model is taken to be homogeneous with Q=1000. The Q tomogram obtained after 30 iterations is shown in Figure 3.10(b). There is a good agreement geologically between the velocity and the Q tomograms. The high attenuation regions in the Q tomogram correspond to the low velocity regions in the FWI tomogram. This is also consistent with the traveltime picks and the peak-frequency picks shown in Figures 3.9(a) and 3.9(b), respectively. There is also a reasonable agreement between the peak-frequencies of the observed and the predicted traces, as can be seen from Figures 3.9(b) and 3.11(b).

The Q-LSRTM image, shown in Figure 3.12(d), is obtained by using the velocity and the WQ tomograms in Figure 3.10. Similar to the synthetic example, the Q-LSRTM image using the WQ tomogram has events with better balanced amplitudes than the standard RTM and LSRTM images in Figures 3.12(a) and 3.12(b), respectively.

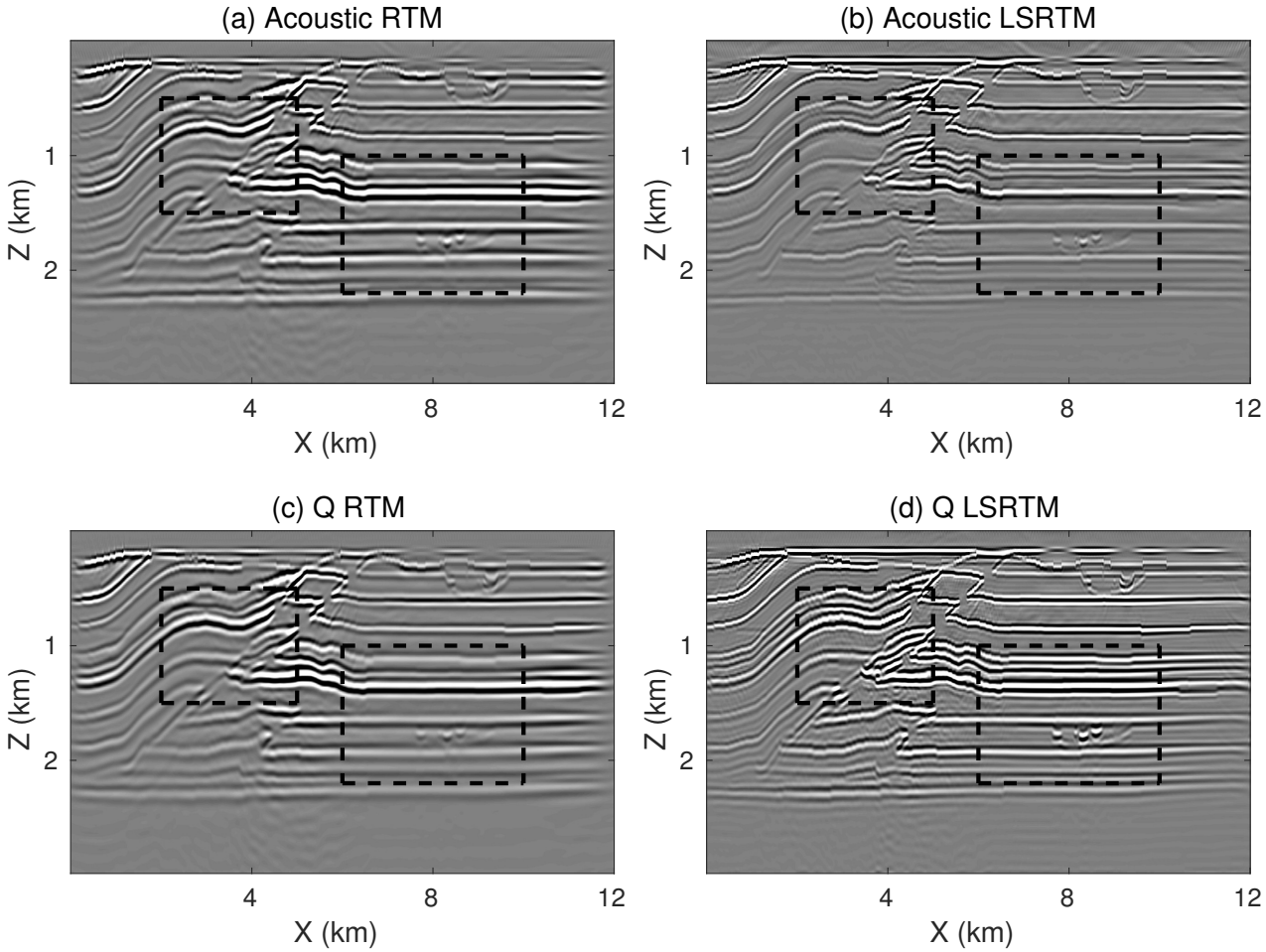

Figure 3.8: Acoustic (a) RTM, and (b) LSRTM images obtained from the visco-acoustic data. (c) Q-RTM, and (d) Q-LSRTM images for the same data using the tomogram obtained from WQ as the background Q model. The black boxes delineate the areas where improvements in imaging can be seen with WQ and Q-LSRTM.

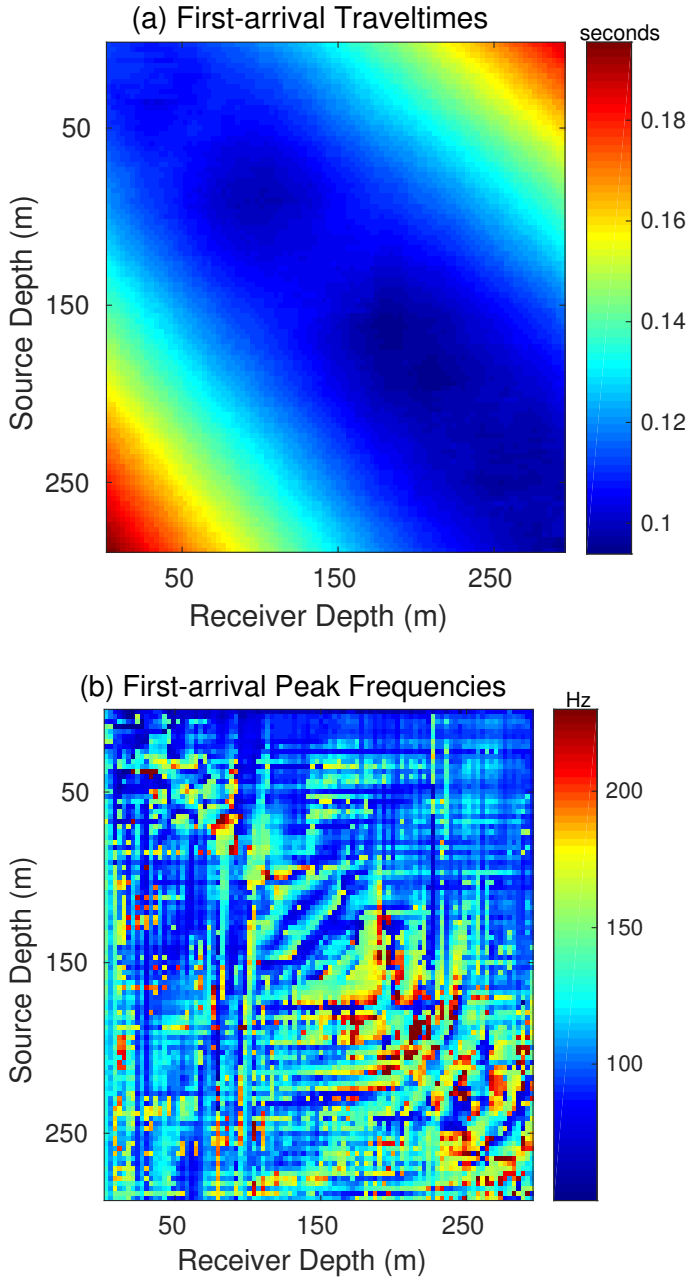

Figure 3.9: (a) The first-arrival traveltime picks for the Friendswood data, and (b) the peak-frequencies for different source-receiver pairs.

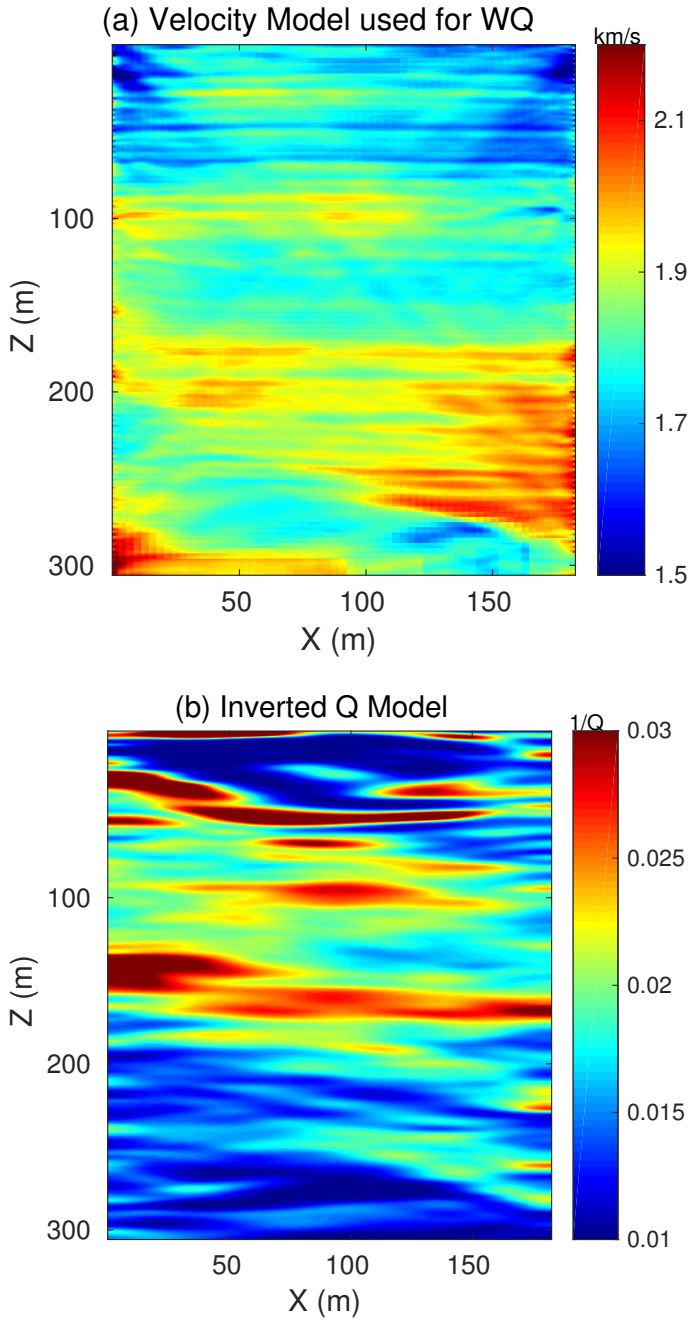

Figure 3.10: (a) Velocity model used for WQ, (b) Q tomogram obtained from WQ.

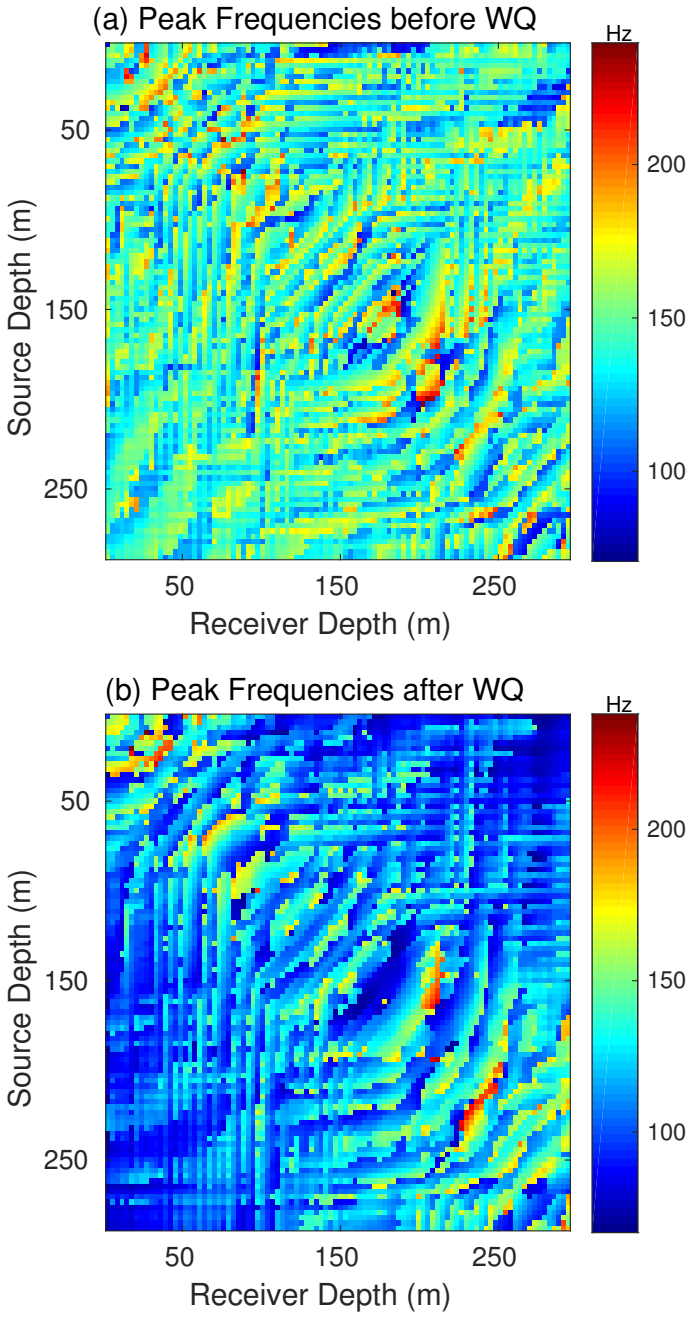

Figure 3.11: The peak-frequencies for different source-receiver pairs for the predicted data (a) before WQ, and (b) after WQ.

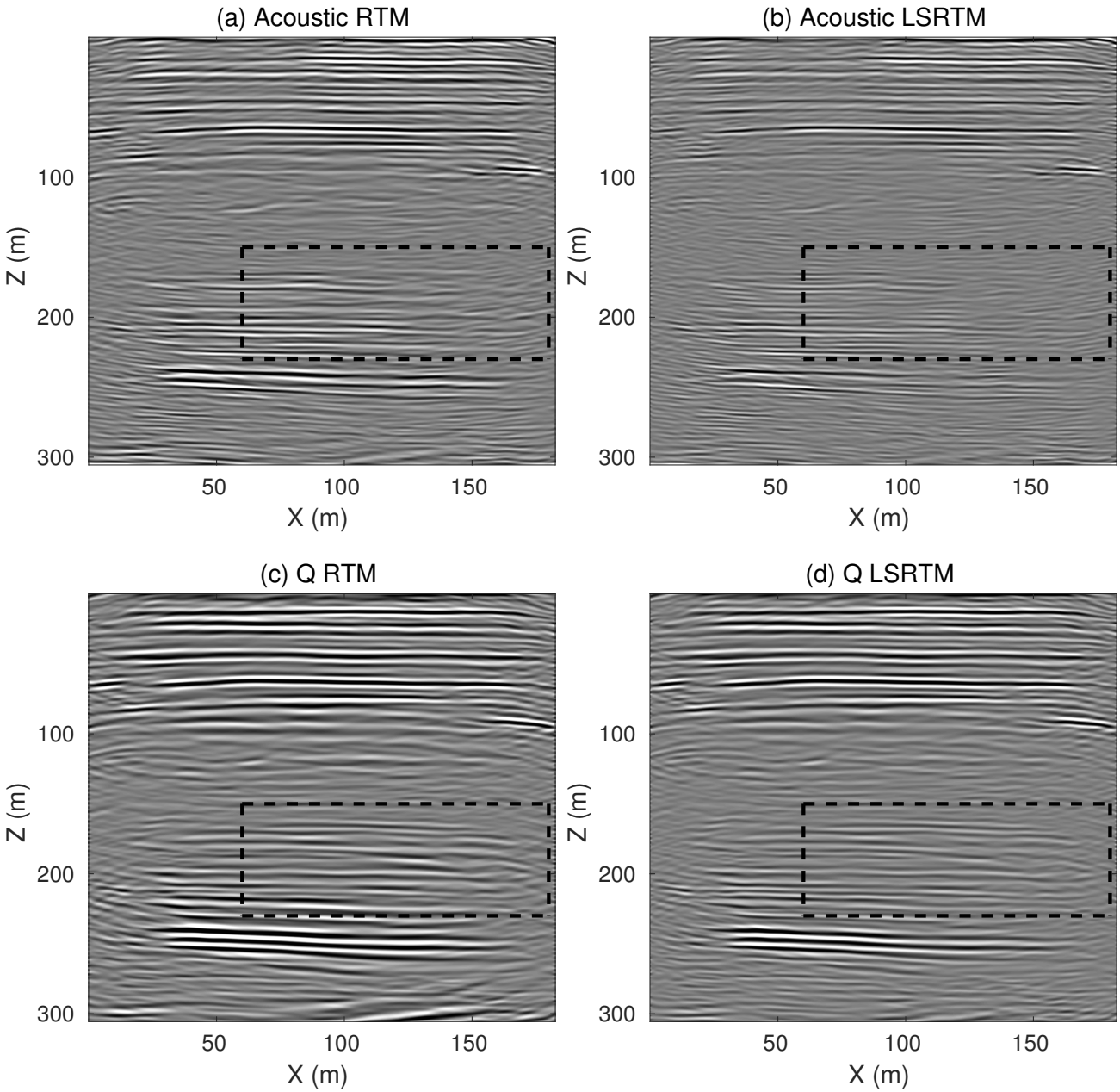

Figure 3.12: Images from (a) acoustic RTM, (b) acoustic LSRTM, (c) Q-RTM, and (d) Q-LSRTM using the tomogram obtained from WQ as the background Q model. The black boxes delineate the areas where improvements in imaging can be seen with WQ and Q-LSRTM.

3.5 Discussion

Synthetic and field data examples demonstrate that WQ can be used to invert for Q tomograms, which can be used as the background Q model for any Q-PSDM algorithm. The background velocity model is needed as an input to compute the visco-acoustic Green's functions for forward propagating the source and for backprojecting the weighted data residuals. For the synthetic example in Figure 3.6, the velocity model for WQ was obtained by smoothing the true slowness model whereas, for the field data example, full waveform inversion was used to obtain the velocity model. Thus, in all the examples, it is ensured that a sufficiently accurate background velocity model is used during WQ. If the background velocity has significant errors, it is likely to lead to significant inaccuracies in the inverted Q tomogram.

To illustrate this point, the synthetic example in Figure 3.6 is repeated again. However, errors are now introduced in the background velocity by applying a triangular smoothing filter with increasing window lengths to the true model in Figure 3.6(a). The WQ tomograms for the different velocity models are shown in the right panel in Figure 3.13. It is evident from these tomograms that as the errors in the background velocity model increase, the Q anomalies are not delineated at the right locations.

For field data applications, if the recorded data are contaminated by noise, the spectra can be quite rough, as seen in Figure 3.9(b). Accurate picking of the peak-frequencies is not trivial in such cases. To mitigate this problem, the amplitude spectra can be smoothed before the peak-frequencies for different source-receiver pairs are estimated. The inverted Q tomogram is then expected to be more accurate.

In all the examples, the SLS-based time-domain visco-acoustic wave equation was used for computing the Green's functions. The visco-acoustic wave equation without memory variables proposed by [Bai et al., 2013] or the decoupled Q equations involving fractional Laplacians and its low-rank formulation ([Zhu and Harris, 2014, Sun et al., 2015b]) can also be used to compute the Green's functions. The WQ

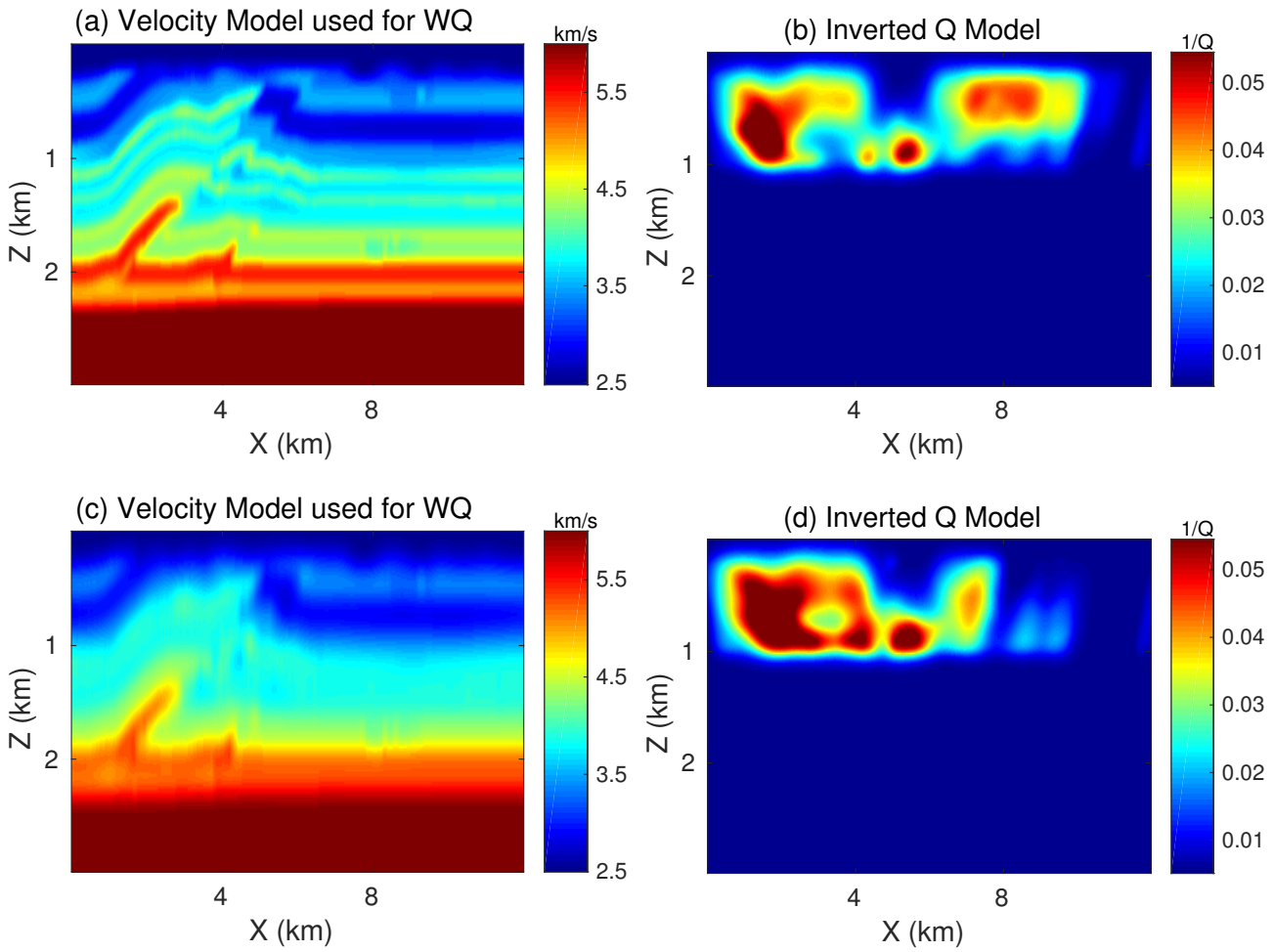

Figure 3.13: Q tomograms obtained from using different background velocity models. The panel on the left shows the velocity models used while the one on the right shows the corresponding Q tomograms obtained from WQ.

method proposed in this paper is generic and can be used to obtain reliable Q tomograms, as long as the modeling equations satisfy a constant Q behavior within the same frequency band ([Kjartansson, 1979]).

3.6 Conclusions

A novel wave-equation Q inversion method is presented where a skeletonized representation of the data, i.e., the difference between the peak frequencies or the centroid frequencies of the observed and the predicted arrivals are inverted to estimate

the background Q model. The gradient for WQ is derived using the implicit function theorem and is numerically obtained by a zero-lag cross-correlation between the forward propagated source wavefield and the backprojected observed pressure seismograms that are weighted by the frequency shifts. Numerical results on synthetic and a crosswell field dataset demonstrate that if the recorded data suffer from strong attenuation, the WQ method can be used to estimate the background Q model. The inverted Q model can be used with any Q-PSDM algorithm to obtain images with balanced amplitudes and high resolution in areas where there is strong attenuation. An input requirement for WQ is an accurate estimate of the background velocity model, which can be obtained by conventional velocity model building algorithms.

Chapter 4

Least-squares reverse time migration with directional Gabor-based preconditioning

4.1 Summary

In this chapter, I present a least-squares reverse time migration (LSRTM) method using directional Gabor-based preconditioning to overcome the low signal-to-noise (SNR) problem of noisy or severely undersampled data. A high resolution local Radon transform of the reflectivity is used and sparseness constraints are imposed on the inverted reflectivity in the local Radon domain. The sparseness constraint is that the inverted reflectivity is sparse in the Radon domain and each location of the subsurface is represented by limited number of geological dips. The forward and the inverse mapping of the reflectivity to the local Radon domain and vice versa is done through 3D Fourier-based discrete Radon transform operators. The weights for the preconditioning are chosen to be varying locally based on the relative amplitudes of the local dips or assigned using quantile measures. Numerical tests on synthetic and field data validate the effectiveness of the proposed approach in producing images

with good SNR and fewer aliasing artifacts when compared with standard RTM or standard LSRTM.

4.2 Introduction

Least-squares migration (LSM) or linearized inversion ([Lailly, 1984]) has been shown to produce images with balanced amplitudes, better resolution and fewer artifacts than standard migration ([Schuster, 1993]; [Nemeth et al., 1999]; [Duquet et al., 2000]; [Plessix and Mulder, 2004]; [Tang, 2009]; [Dai and Schuster, 2010b]; [Wong et al., 2011]). Besides a migration of the data residual, every iteration of LSM involves Born modeling to estimate the step-length and to update the data residual. The potential of LSM, especially LSRTM, to produce images of superior quality than any other conventional migration technique has been well studied in recent years.

The problems associated with LSRTM can be broadly grouped into three major categories: 1) inadequate physics taken into account by the modeling and the adjoint equations, and 2) errors in the migration velocity model, and 3) the computational cost. For reducing the computational cost, phase-encoded migration ([Morton and Ober, 1998]; [Romero et al., 2000]) was proposed that was later extended to multisource LSRTM by [Dai et al., 2010], [Schuster et al., 2011], [Dai et al., 2012] and several other authors. They showed that by an iterative migration of supergathers, multisource LSRTM can produce more accurate reflectivity images than standard RTM and at a fraction of the cost of standard LSRTM. A similar approach was adopted by [Boonyasiriwat and Schuster, 2010] and [Herrmann and Li, 2012] as they used a combination of randomized dimensionality-reduction and divide-and-conquer-techniques to decimate the LSM problem as a series of smaller sub-problems where each sub-problem involved iterating on a small randomized subset of the data. [Herrmann and Li, 2012] also combined their approach with compressive sensing ([Candes et al., 2006b]) and

curvelet-domain sparse recovery ([Candes et al., 2006a]) to mitigate the crosstalk noise.

Besides the computational cost and errors in the migration velocity model, the other major problem related to LSM is insufficient physics taken into account by the modeling and the adjoint equations. This is because the real earth is anisotropic and anelastic and estimation of all the subsurface parameters that have significant effect on wave propagation is not trivial. Limitations in our acquisition capabilities, especially related to data sampling and illumination also remain an issue. Thus, for field-data applications of LSM, it becomes important to incorporate some form of regularization or preconditioning into the least-squares inversion that would allow for a more accurate representation of the subsurface model parameters and mitigate some of these problems. The inverted image should also be consistent with any prior information that is available for them.

However, in the absence of well logs, getting reliable prior information that can be used as constraints to guide the inversion is not straightforward. A more popular choice is to impose constraints on the property of a image that is desired. For example, [Wang and Sacchi, 2007] use a cost function for one way wave-equation based LSM with regularization constraints for smoothness along offset-domain common image gathers (CIGs) and reflectivity sparseness in depth. [Cabrales-Vargas and Marfurt, 2013] also formulated a regularized least-squares Kirchhoff migration problem where they used a penalty function that controls the amount of roughness in common reflection point gathers (CRPGs). They used a three-point mean filter in every CRPG to remove the aliasing artifacts. [Aldawood et al., 2014] used a compressed sensing based approach for Kirchhoff migration and imposed sparsity constraints on the inverted reflectivity image by solving a Lasso problem. Total variation regularization based approaches ([Anagaw and Sacchi, 2012]; [Lin and Lianjie, 2015]) have also been used with LSM to obtain images with sharp interfaces and discontinuities.

Another approach is to use a change of basis for the reflectivity using linear sparse transforms or some form of model reparameterization ([Harlan, 1995]; [Fomel and Guitton, 2006]). Sparseness constraints can then be imposed on the image in the new domain since the true reflectivity is never sparse in the depth domain. The choice of a suitable transform is dependent on 1) perfect reconstruction of the parameters after forward and inverse transforms, 2) suitability for use with the conjugate gradient method, 3) efficient computation, and 4) minimal redundancy ([Kingsbury, 2001]). [Miller et al., 2005] used the Dual Tree Complex Wavelet Transform as a basis for the reflectivity and demonstrated that such a change of basis leads to a better reduction of noise and migration artifacts while, at the same time, the discontinuities are preserved better than standard LSM for very sparse data. [Herrmann et al., 2009] and [Herrmann and Li, 2012] used curvelets and [Dutta, 2015] used seislets as basis functions for the reflectivity and showed that with sparsity promoting imaging techniques, it is possible to recover high-quality images from undersampled or noisy data.

In this chapter, we use a directional Gabor-based preconditioning approach for LSRTM to overcome some of the problems associated with LSRTM when the recorded data are noisy and the background velocity model is inaccurate. We use the prior information that the inverted reflectivity is sparse in the local Radon domain and is represented by limited number of dips. The LSRTM problem is then posed as an optimization problem that minimizes the L1-norm of the local Radon transform of the image subject to the condition that the data misfit is minimized to an acceptable tolerance level. The forward mapping to the local Radon domain and its inverse is done through 3D discrete Radon transform operators in the Fourier domain. Using model reparameterization and re-weighting, the problem is solved as a preconditioning problem where the weights are chosen based on the amplitudes of the events in the dip domain. Numerical tests on synthetic and on 2D land and 3D marine data show

that Radon-based preconditioning can produce images with fewer migration artifacts and better SNR than standard RTM or LSRTM in a few iterations. Even at greater depths where migration velocity errors accumulate, the preconditioning approach is shown to produce images with better focusing and fewer artifacts than standard RTM and LSRTM.

This chapter is organized into five sections. After the introduction, the second section describes the theory of LSRTM using directional Gabor-based preconditioning. Numerical results with synthetic and field data are then presented in the third section. The limitations of the proposed method are discussed in the fourth section and the conclusions are in the last section.

4.3 Theory

For a regularized L2-norm inverse problem, we seek to find a solution that best explains the observed data and is consistent with the prior knowledge that is available before any observations are made. This can be obtained by minimizing the misfit function over the image space $\mathbf{M}$ given by

$$\underset{\mathbf{m}\in\mathbf{M}}{min}\left[\phi(\mathbf{m}) = \frac{1}{2}||\mathbf{d}-\mathbf{L}\mathbf{m}||_2^2 + \lambda\Re(\mathbf{m})\right]. \tag{4.1}$$

Here, $\mathbf{d}$ represents the observed data, $\mathbf{m}$ represents the reflectivity model and $\Re(\mathbf{m})$ is a discrete regularizer/prior that imposes constraints on the solution $\mathbf{m}$. These constraints can be such that $\Re(\mathbf{m})$ should be sparse or the reflectors in $\mathbf{m}$ should be sharp. $\mathbf{L}$ represents a linear modeling operator and $\lambda > 0$ controls the strength of the regularization term. If $\lambda = 0$, equation 4.1 is equivalent to solving an unconstrained optimization problem where the emphasis is purely on minimizing the misfit between the predicted and the observed data. For incomplete, undersampled or noisy data, emphasizing only the data misfit can lead to images that are degraded

in quality with iterations because of over-fitting the noise. In the case of LSM, errors in the migration velocity model can also lead to defocussing of images with iterations ([Dutta et al., 2014]). Thus, for real data applications, it is important to incorporate some sort of regularization term into the inversion that allows for a more accurate representation of the subsurface model parameters.

If $\lambda \neq 0$, equation 4.1 solves a constrained optimization problem where the tradeoff between the data fit and the regularization term is controlled by the damping factor λ. This damping parameter can be estimated by trial-and-error, or a more rigorous procedure is to obtain it from L-shaped curves constructed by a log-log plot of the length of the model vector vs the length of the residual vector for different choices of λ ([Calvetti et al., 2000]). Such an approach, however, is not feasible for industrial-scale 3D seismic inversion problems.

An alternative approach is to reformulate the problem in equation 4.1 as

$$\underset{\mathbf{m}\in\mathbf{M}}{min}||\mathbf{Rm}||_1 \text{ subject to } ||\mathbf{d}-\mathbf{Lm}||_2 < \epsilon, \tag{4.2}$$

where $\mathbf{Rm}$ can be any linear transform of the model vector $\mathbf{m}$ that minimizes the L1-norm of the model in the transformed domain and ϵ is the tolerance level to which the L2-norm data misfit is minimized. In this case, $\mathbf{R}$ is taken to be the directional Gabor transform or the local Radon transform of the image which is computed using a 3D Fourier-based discrete Radon transform algorithm (See Appendix F). Using a weighting matrix $\mathbf{W}$, equation 4.2 can be written as

$$\underset{\mathbf{m}\in\mathbf{M}}{min}||\mathbf{WRm}||_2 \text{ subject to } ||\mathbf{d}-\mathbf{Lm}||_2 < \phi. \tag{4.3}$$

Here, $\mathbf{WRm}$ is a weighted linear radon transform operation that makes the model vector $\mathbf{m}$ sparse in the transformed domain after the weighting. In our work, we have chosen the weights as a function of the amplitudes of the local dip events. This can be

done by weighting the dips based on the strength of their amplitudes. The diagonal elements of the weighting matrix $\mathbf{W}$ are given by $diag(\mathbf{W}) = |\mathbf{Rm}|^{-1}$. Alternatively, some quantile-based measure can be used to select a threshold level beyond which the weaker dips or the noise are down-weighted.

The normal equations corresponding to the misfit function in equation 4.3 are given by

$$\left(\lambda \mathbf{R^T W^T W R} + \mathbf{L^T L}\right) \mathbf{m} = \mathbf{L^T d}, \tag{4.4}$$

which can be solved using any conjugate-gradient based algorithm ([Nocedal and Wright, 1999]). The solution $\mathbf{m}$ for the normal equations in 4.4 spans over the Krylov subspace

$$\begin{aligned} \kappa^{\mathbf{L^T L}+\lambda \mathbf{R^T W^T W R}} = \text{span}\{&\mathbf{L^T d}, (\mathbf{L^T L} + \lambda \mathbf{R^T W^T W R})\mathbf{L^T d}, \ldots, \\ &(\mathbf{L^T L} + \lambda \mathbf{R^T W^T W R})^{i_{\max}-1}\mathbf{L^T d}\} \subset \mathbf{M}, \end{aligned} \tag{4.5}$$

where $i_{\max}$ is the limit of the Krylov subspace or the number of iterations carried out. The solutions in equation 4.5 are controlled by the damping parameter λ and i_{max}. Since the optimum estimation of λ is not trivial, a model reparameterization approach is required for the reflectivity $\mathbf{m}$. If $\mathbf{W}$ and $\mathbf{R}$ are invertible, we can write

$$\mathbf{p} = \mathbf{WRm}, \ \mathbf{m} = \mathbf{R^{-1} W^{-1} p} \tag{4.6}$$

such that equation 4.3 gives

$$\underset{\mathbf{p} \in \mathbf{WRM}}{min} ||\mathbf{p}||_2 \text{ subject to } ||\mathbf{d} - \mathbf{L R^{-1} W^{-1} p}||_2 < \phi. \tag{4.7}$$

The normal equations corresponding to the misfit function in equation 4.7 are given

by

$$\left(\lambda \mathbf{I} + \mathbf{W^{-T}R^{-T}L^{T}LR^{-1}W^{-1}}\right)\mathbf{p} = \mathbf{W^{-T}R^{-T}L^{T}d}. \tag{4.8}$$

Setting $\lambda = 0$ in equation 4.8 we get the new normal equations

$$\left(\mathbf{W^{-T}R^{-T}L^{T}LR^{-1}W^{-1}}\right)\mathbf{p} = \mathbf{W^{-T}R^{-T}L^{T}d}. \tag{4.9}$$

The solution $\mathbf{p}$ for the normal equations in equation 4.9 spans over the Krylov subspace

$$\begin{aligned}\kappa^{\mathbf{W^{-T}R^{-T}L^{T}LR^{-1}W^{-1}}} = \text{span}\{&\mathbf{W^{-T}R^{-T}L^{T}d}, (\mathbf{W^{-T}R^{-T}L^{T}LR^{-1}W^{-1}})\mathbf{W^{-T}R^{-T}L^{T}d}, \\ &\ldots (\mathbf{W^{-T}R^{-T}L^{T}LR^{-1}W^{-1}})^{\mathbf{i}_{\max}-1}\mathbf{W^{-T}R^{-T}L^{T}d}\} \subset \mathbf{WRM}.\end{aligned} \tag{4.10}$$

On comparing the Krylov spaces in equations 4.5 and 4.10 it can be seen that sparseness constraints are imposed on the solutions in equation 4.10 by the action of the Radon transform and the weighting operators. In contrast, the constraints in equation 4.5 are controlled by the parameter λ. Thus, we can set $\lambda = 0$ in equation 4.8 and let the number of internal iterations in the conjugate gradient algorithm play the role of the regularizer. Even though the system of equations is partially solved at every external iteration because the conjugate gradient algorithm is stopped before the solution is complete, the total matrix operator resulting from the preconditioning is diagonalized enough so that one can avoid using a damping factor for the smaller eigenvalues ([Trad et al., 2003]). Also, for different values of λ, equation 4.10 remains invariant whereas equation 4.5 needs to be recomputed from the beginning for different values of λ.

Such a preconditioning approach is useful because it provides a way of incorporat-

ing the prior information (that the inverted reflectivity should be sparse in the local Radon domain) directly into the forward operator. Otherwise, the solution would require a large number of iterations in the counterpart regularized problem. A large number of iterations will be required before the small eigenvalues become dominant and the corresponding features start showing up in the space $\kappa^{\mathbf{L^T L}+\lambda\mathbf{R^T W^T W R}}$. On the other hand, the vectors in the Krylov space $\kappa^{\mathbf{W^{-T} R^{-T} L^T L R^{-1} W^{-1}}}$ are directly influenced by the transformation and the weighting operators acting on the gradients at every iteration. Thus, the images from the very first iteration are expected to be sparser in the local Radon domain and cleaner with good SNR and fewer migration artifacts in the image domain for the preconditioned problem in equation 4.7.

4.4 Numerical results

The effectiveness of LSRTM using Gabor-based preconditioning is now demonstrated with synthetic and field data examples. The synthetic example is the 3D SEG/EAGE salt model whereas the field data examples are 1) 2D land, and 2) 3D marine datasets.

In the synthetic example, the observed data are generated by an O(2,8) time-space-domain solution of the constant-density second-order acoustic wave equation without a free-surface. A Ricker wavelet with a 15-Hz peak frequency is chosen as the source wavelet. The data are then migrated using RTM, LSRTM, preconditioned RTM and preconditioned LSRTM. Here, RTM and preconditioned RTM refer to the first iteration of standard LSRTM and the proposed preconditioned LSRTM, respectively. Source-side illumination compensation ([Plessix and Mulder, 2004]) is used as an additional preconditioner for both standard and preconditioned LSRTM . All the images have also been filtered using a classical bandpass filter using the same low-cut and high-cut frequencies. 20 iterations are ran for the 2D example whereas only 5 iterations are carried out for the 3D cases.

4.4.1 Synthetic example

I first demonstrate the effect of using Gabor-based preconditioned LSRTM on the 3D SEG/EAGE salt model. There are only 45 shots recorded on a 5×9 source grid with a 960 m shot and shot-line separation. Each shot is recorded by receivers on a 201×201 receiver grid with a 50 m spacing between the receivers. The observed data are recorded using the true velocity model (a 2D slice is shown in Figure 4.1) for around 8 seconds with a 4 ms sampling rate. For migration, the true slowness model is smoothed using a triangle smoothing filter with a smoothing radius of 50 m in each of the 3 directions.

Figures 4.2(a)-4.2(d) compare 2D slices of the images from standard RTM, standard LSRTM, preconditioned RTM and preconditioned LSRTM, respectively. The RTM image in Figure 4.2(a) suffers from very strong backscattering noise because of the presence of the salt body. The slices show strong acquisition footprints because of the sparse acquisition geometry and the reflector amplitudes are also very weak. The image contains significant high-frequency noise because of using severely undersampled data. The LSRTM image in Figure 4.2(b) shows some improvements over the standard RTM image. The reflector amplitudes are better balanced and the acquisition footprints are mitigated. However, the aliasing noise is still prominent and is severe below the salt body. The preconditioned LSRTM image, shown in Figure 4.2(d), is free from aliasing noise and the subsalt images are much cleaner when compared to the standard RTM and LSRTM images. The preconditioned images also have a much better SNR than the standard RTM and LSRTM images.

4.4.2 2D land data example

The preconditioning approach is now tested on a 2D land dataset. There are 251 shots buried at a depth of 3 m and a shot interval of 60 m. Each shot is recorded by 1000 receivers at a spacing of 15 m on the surface for a total recording time of 8 seconds.

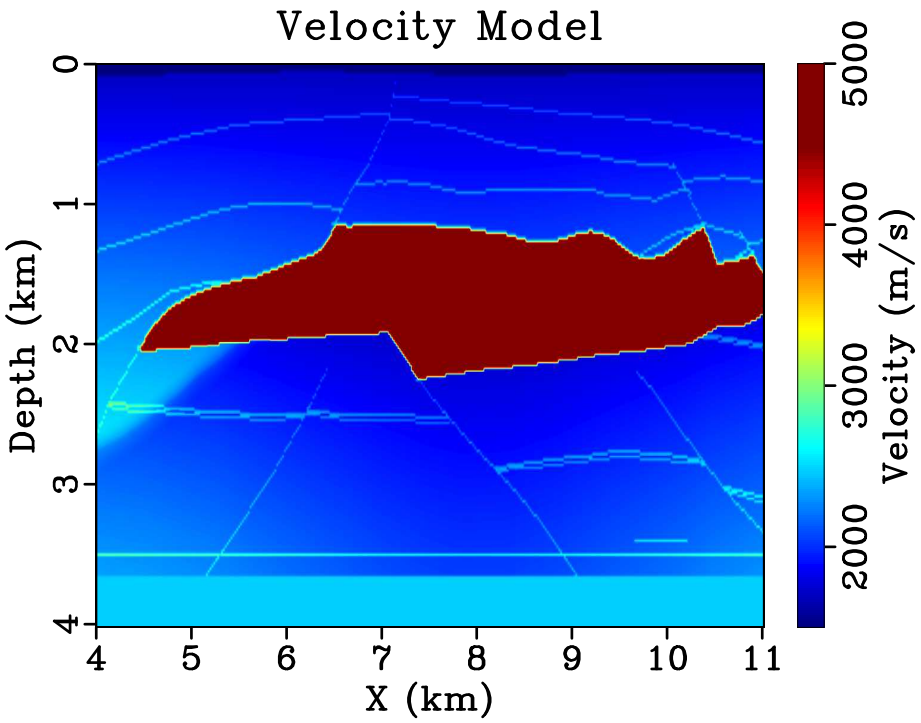

Figure 4.1: A 2D slice of the 3D SEG/EAGE salt model.

During data processing, refraction statics, residual statics and F-X noise attenuation are applied in both the shot and the receiver domains. The velocity model, shown in Figure 4.3, is obtained by ray-based refraction tomography and interpretative model building.

Figures 4.4(a) and 4.4(b) compare the standard RTM and the standard LSRTM images, respectively. The images are quite noisy and it is hard to delineate the truncation of the high-velocity feature. This can be seen from the magnified views of these images in Figures 4.5(a) and 4.5(b). The preconditioned RTM and the preconditioned LSRTM images, shown in Figures 4.4(c) and 4.4(d), respectively, have a better SNR than the standard images at both shallow and deeper depths. The truncation of the high velocity build up and the base of it can be clearly delineated in the magnified views of these images in Figures 4.5(c) and 4.5(d).

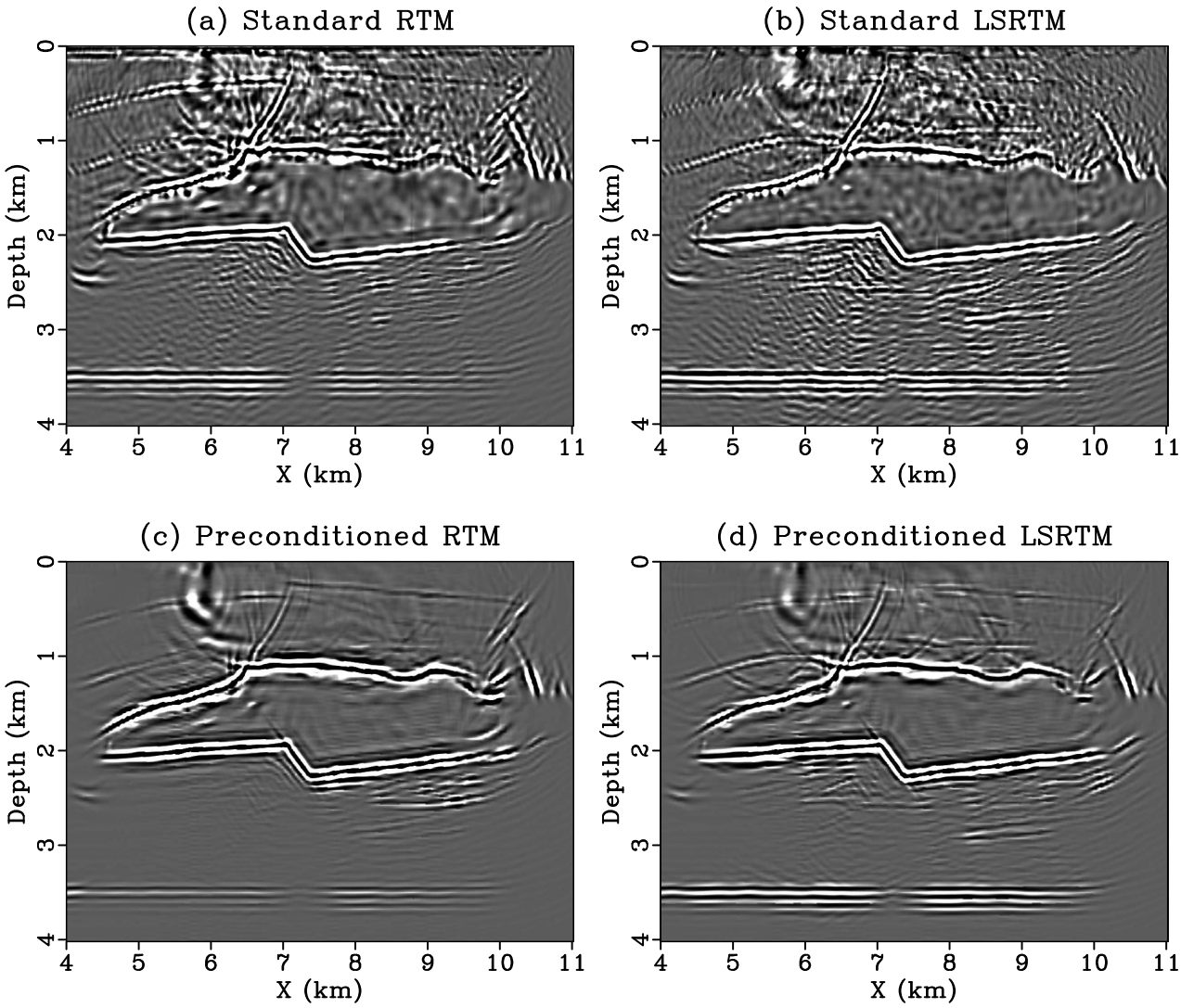

Figure 4.2: Comparison between images from (a) standard isotropic RTM, (b) standard isotropic LSRTM after 5 iterations, (c) preconditioned isotropic RTM, and (d) preconditioned isotropic LSRTM after 5 iterations.

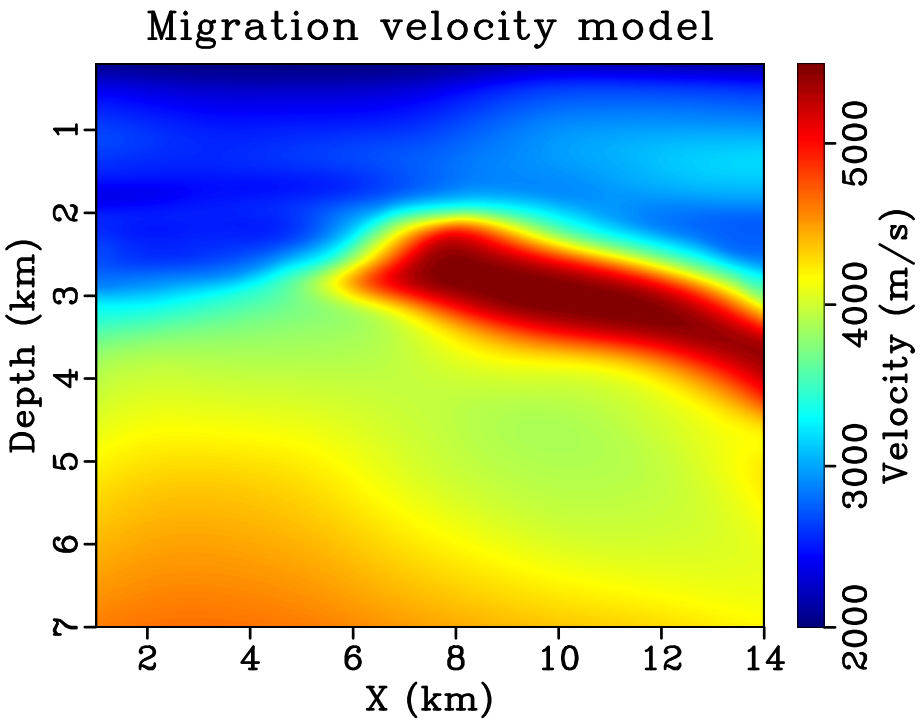

Figure 4.3: Migration velocity model estimated using ray-based tomography for the 2D land dataset.

4.4.3 3D marine data example

As a final example, preconditioned LSRTM is applied to a 3D marine dataset from offshore West Africa. The acquisition is broadband with a shot interval of 18.75 m (flip-flop) and a source centre separation of 25 m. The receiver spread has 12 cables with an active length of 6 km. There are 480 receiver groups with a group interval of 12.5 m. The sampling interval is 2 ms with a total recording time of 8 s. Before tomography and migration, standard processing steps like swell noise attenuation, linear noise attenuation and receiver deghosting are carried out. 3D SRME and Radon de-multiple are also used to remove the free-surface multiples.

The background velocity, epsilon, delta, inline dip and crossline dip models are shown in Figures 4.6(a)-4.6(e), respectively. The velocity model is obtained using ray-based tomography and full waveform inversion, whereas ray-based tomography is used to obtain the epsilon model. The delta and dip models are constructed using

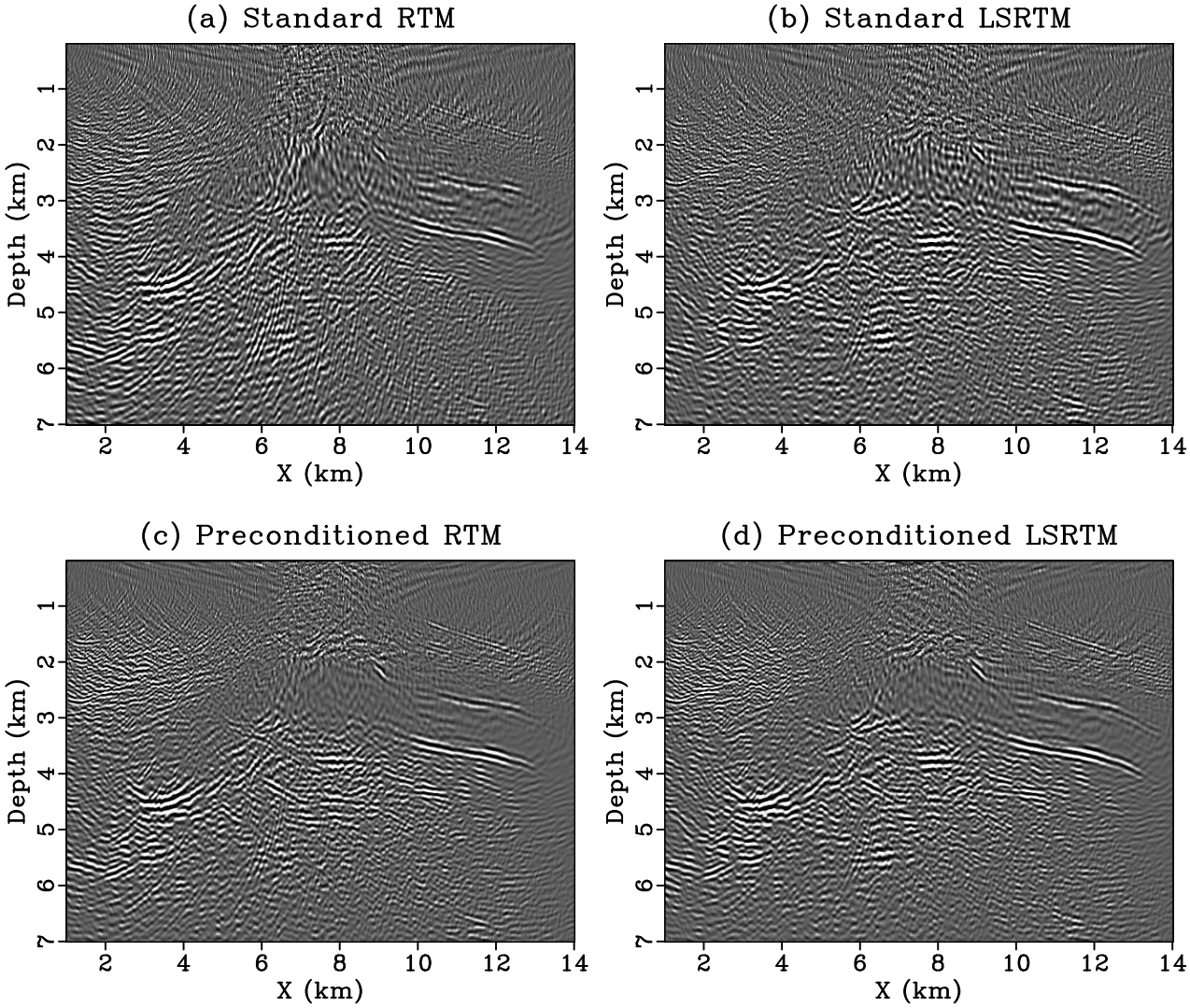

Figure 4.4: Comparison between images from (a) standard isotropic RTM, (b) standard isotropic LSRTM after 20 iterations, (c) preconditioned isotropic RTM, and (d) preconditioned isotropic LSRTM after 20 iterations.

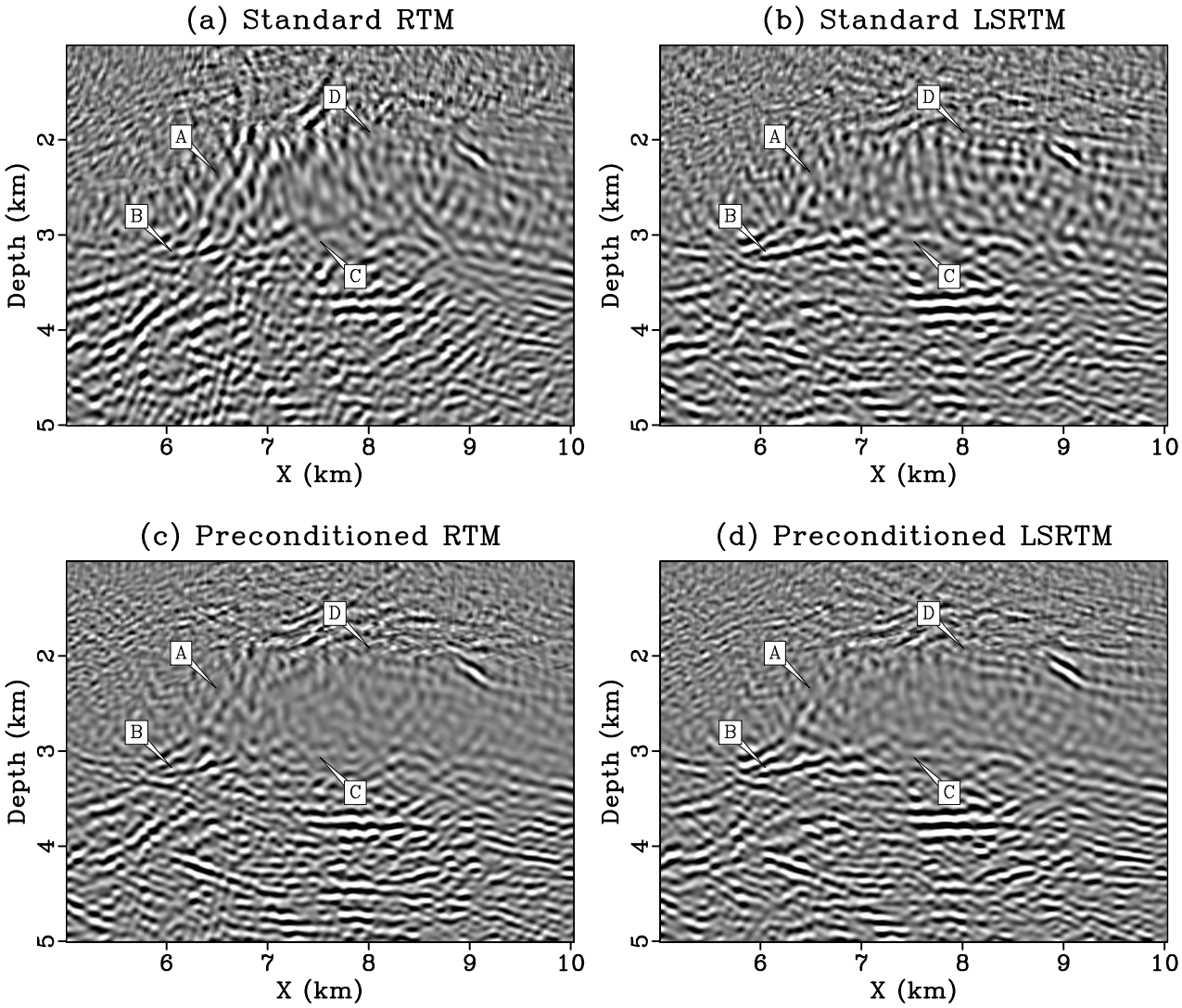

Figure 4.5: Comparison between magnified views of the images from (a) standard isotropic RTM, (b) standard isotropic LSRTM, (c) preconditioned isotropic RTM, and (d) preconditioned isotropic LSRTM. The arrows indicate the areas where improvements in imaging can be seen from the Radon-based preconditioning.

information from wells in the study area.

For migration, there are 10,241 shots used and a bandpass filter of 3-5-25-30 Hz is applied to the traces. The standard TTI RTM and TTI LSRTM images are shown in Figures 4.7(a) and 4.8(a), respectively. It is evident from these images that standard LSRTM improves the image quality (in terms of improved resolution and balanced amplitudes) in the shallow parts. However, the image quality of the carbonate turtlebacks at depths of 3-5 kms, shown in Figures 4.9(a) and 4.9(b), are degraded in both the standard RTM and the standard LSRTM images. We suspect that inaccuracies in the background models or presence of noise in the data might lead to such degradation. The crossline sections of these images, shown in Figures 4.10(a) and 4.10(b), also suffer from strong aliasing artifacts and migration smiles due to large aperture.

The preconditioned TTI RTM and TTI LSRTM images are shown in Figures 4.7(b) and 4.8(b), respectively. Magnified views of these images in the inline and crossline directions are shown in Figures 4.9(c)-4.9(d) and Figures 4.10(c)-4.10(d), respectively. It is evident from these images that the turtle backs are imaged with an improved SNR and they can be clearly delineated. The aliasing noise has also been significantly mitigated in the crossline images.

4.5 Discussion

The Gabor-based preconditioning approach provides significant uplift in the quality of the inverted images when the desired features are planar and the data used for migration have been carefully pre-processed such that free-surface multiples or internal multiples are eliminated. However, the presence of free-surface or internal multiples in the data will lead to false reflection events in the image which are not discriminated against by the present approach. Unless a very large number of iterations of

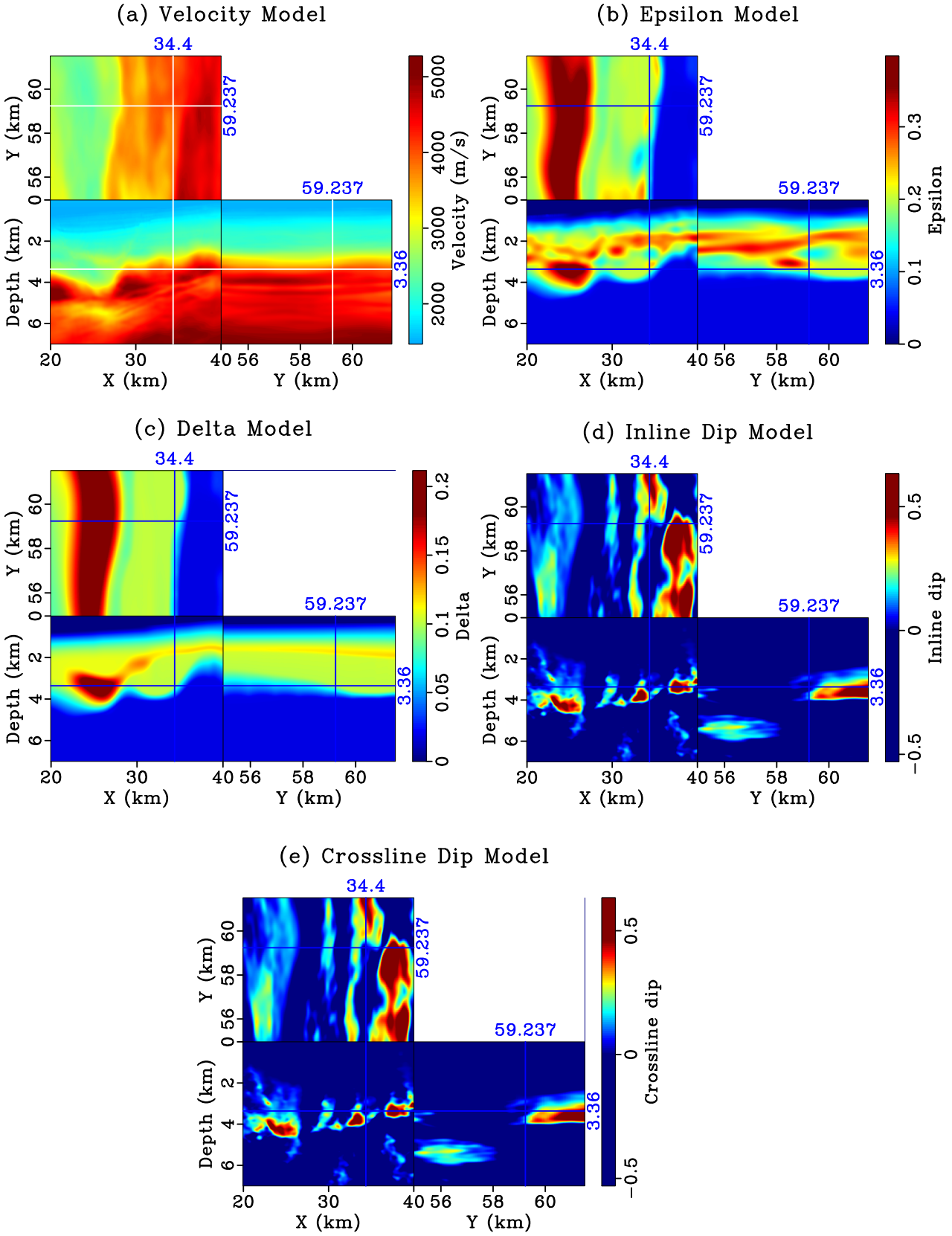

Figure 4.6: Background (a) velocity, (b) epsilon, (c) delta, (d) inline dip, and (e) crossline dip models used for migration.

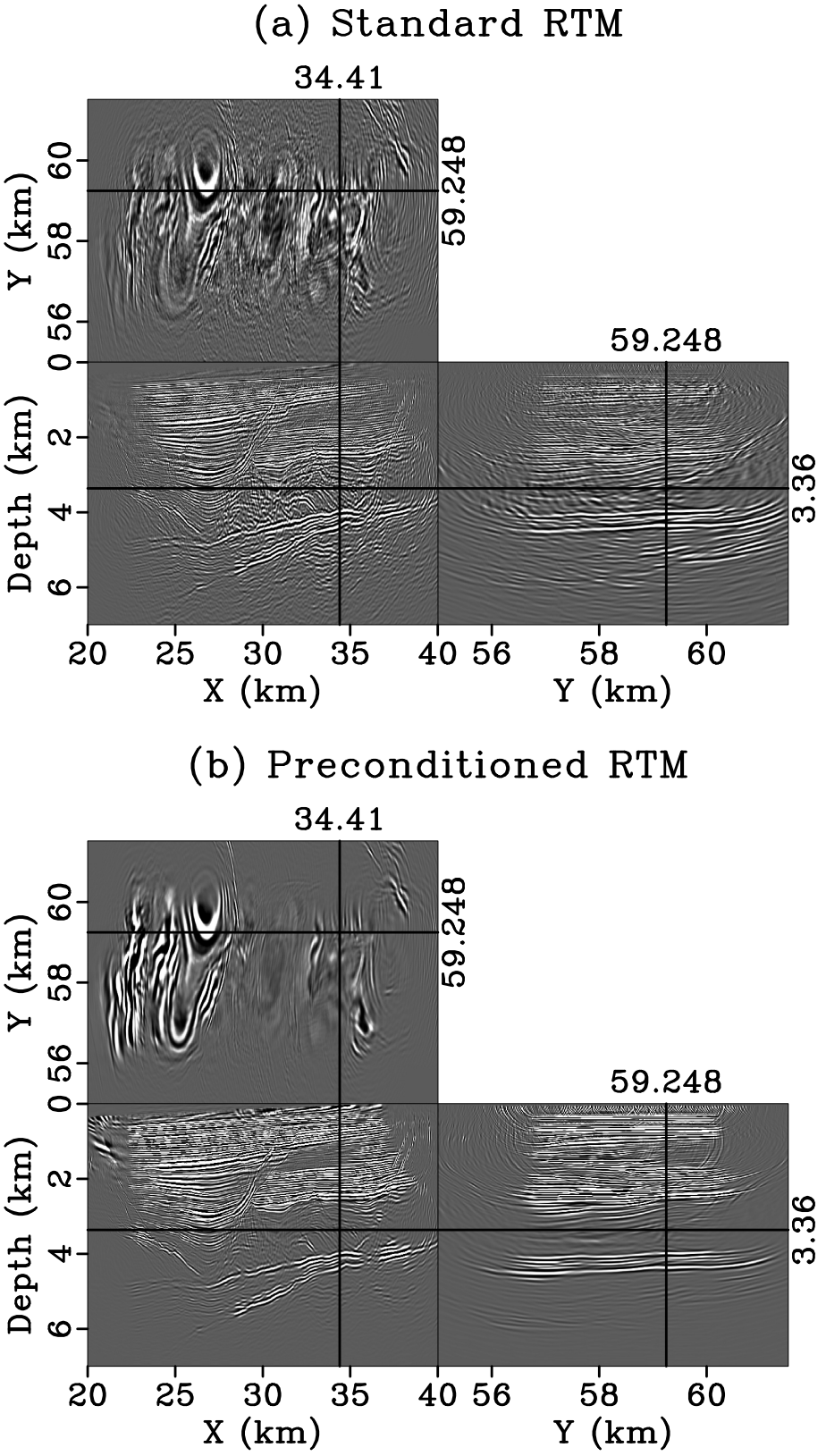

Figure 4.7: Comparison between images from (a) standard TTI RTM, and (b) preconditioned TTI RTM.

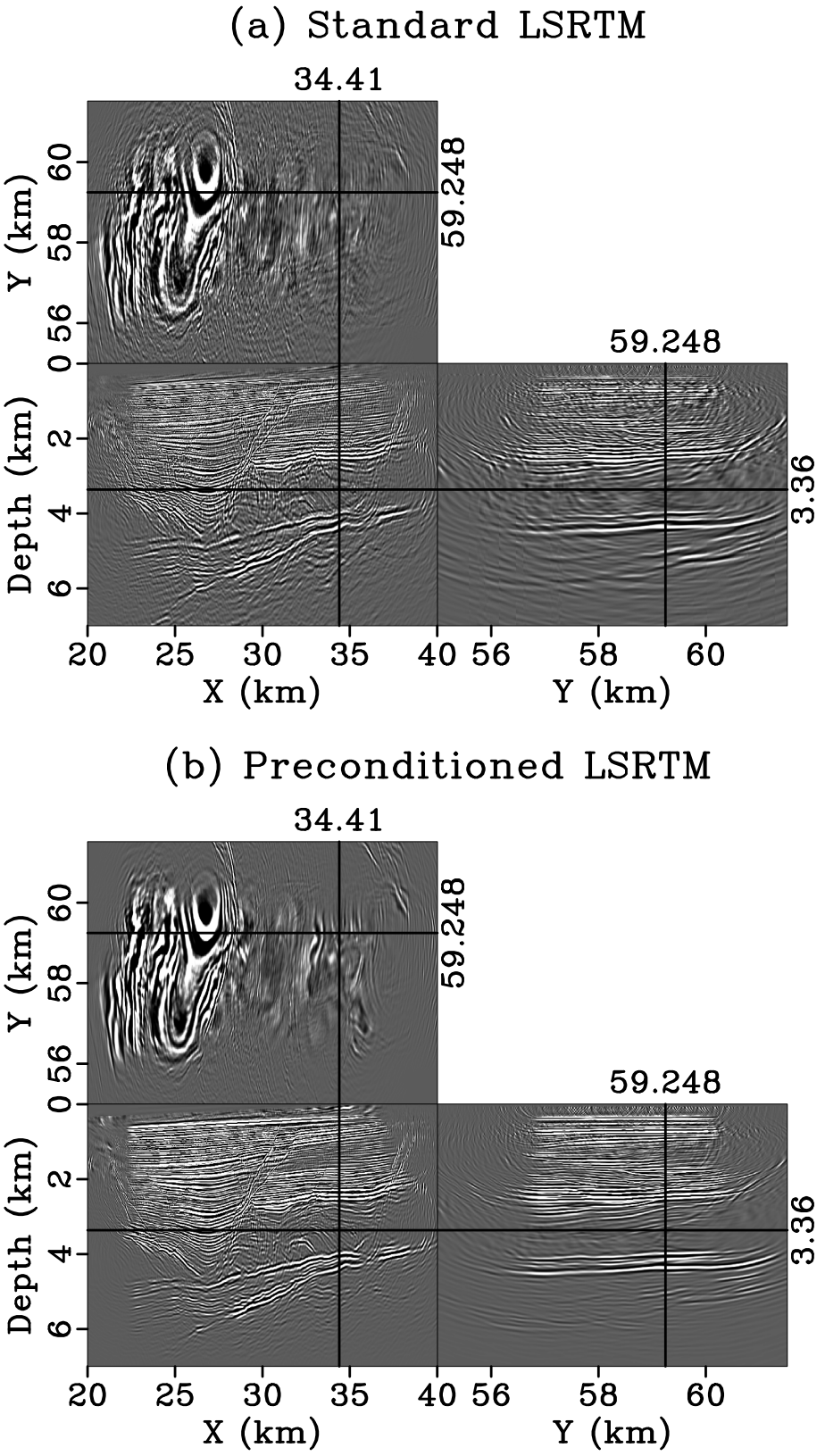

Figure 4.8: Comparison between images from (a) standard TTI LSRTM, and (b) preconditioned TTI LSRTM.

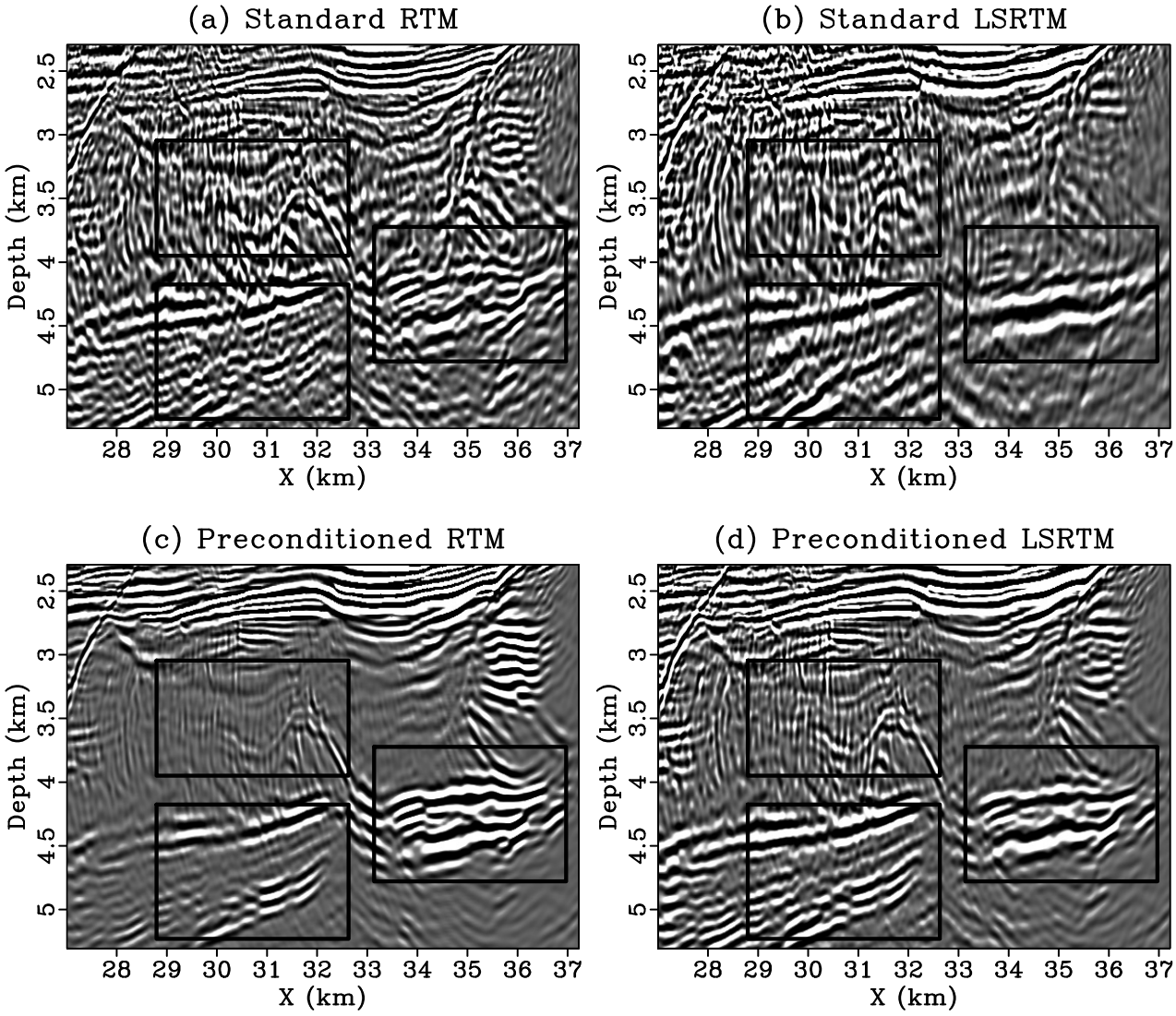

Figure 4.9: Comparison between magnified views of the images in the inline direction from (a) standard RTM, (b) standard LSRTM, (c) preconditioned RTM, and (d) preconditioned LSRTM. The black boxes indicate the areas where improvements in imaging can be seen from the Radon-based preconditioning.

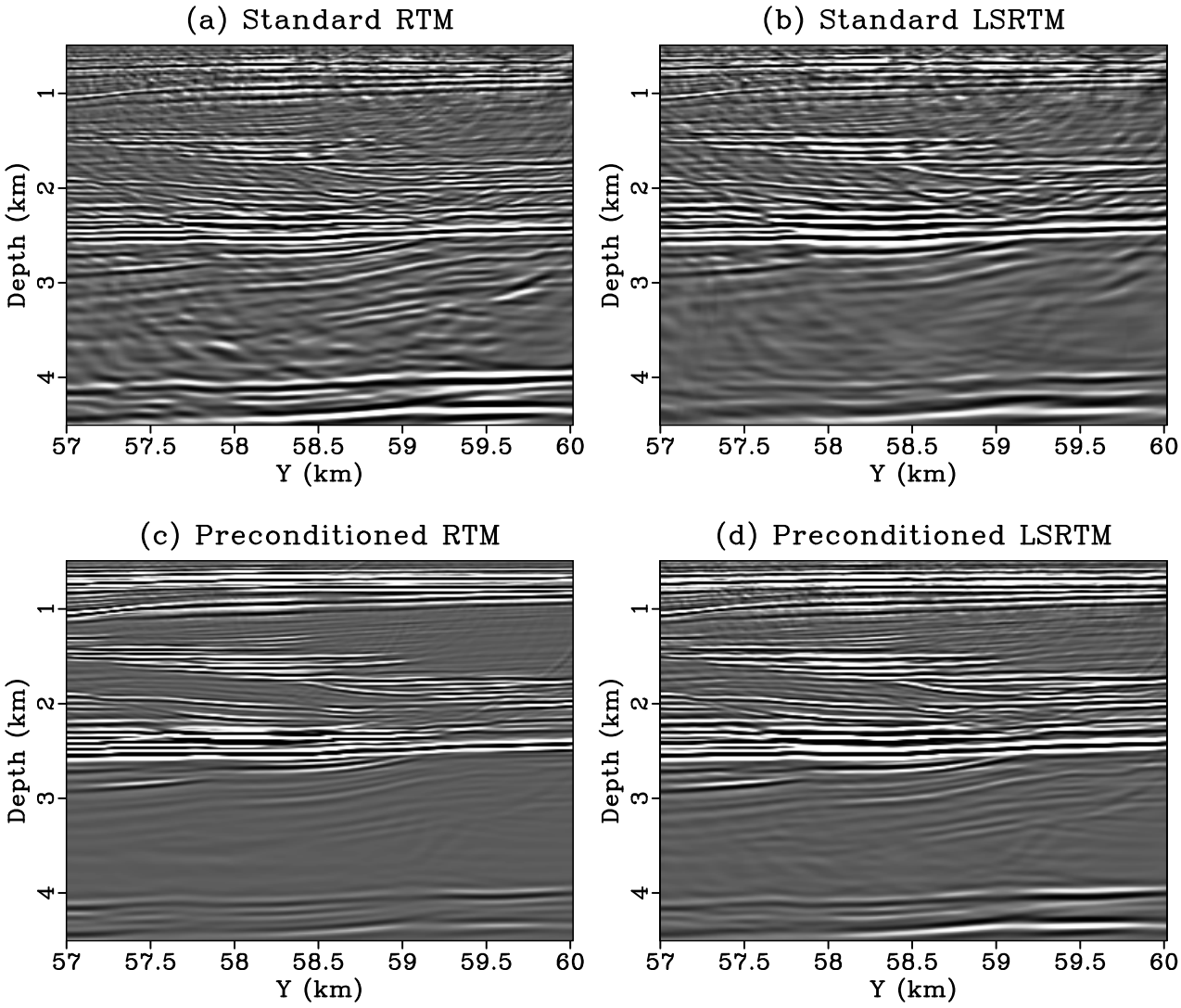

Figure 4.10: Comparison between magnified views of the images in the crossline direction from (a) standard RTM, (b) standard LSRTM, (c) preconditioned RTM, and (d) preconditioned LSRTM.

least-squares is carried out, these events will still be quite strong in the image and they will have strong amplitude signatures in the local Radon domain. Filtering out these events in the extended domain is not trivial.

To illustrate our point, the synthetic example shown earlier is repeated again. However, the observed data is now modeled with a free-surface. The standard RTM and LSRTM images are shown in Figures 4.11(a) and 4.11(b), respectively. The effect of free-surface multiples and internal multiples is evident specially below the sub-salt areas. The preconditioned RTM and LSRTM images, shown in Figures 4.11(c)-4.11(d), are cleaner above the salt but they have many false events below the salt. This can be verified by comparing these images with the true velocity model is Figure 4.1.

Another disadvantage of this method is the additional computational cost incurred during the preconditioning. Solving the normal equations in equation 4.9 involves storing an extended model to memory at every iteration. For a 2D image $m(x, z)$, a local Radon transform of it will lead to an extended image $\hat{m}(x, z, p_x)$ whereas for a 3D image $m(x, y, z)$, the transfomation will lead to a 5D volume $\hat{m}(x, y, z, p_x, p_y)$. The extended models require the computational nodes to have sufficient local memory to avoid using I/O from the disk. Also, there is an additional computational overhead because of the forward and inverse Radon transforms. This cost will vary depending on the size of the model. For the 3D field data example, each iteration of preconditioned LSRTM took roughly 3-4 times as long as standard LSRTM.

4.6 Conclusions

A directional Gabor-based preconditioning approach is presented for LSRTM that mitigates the problems of low SNR data and aliasing. Sparseness constraints are imposed on the reflectivity in the local Radon domain and the forward and the inverse mapping of the reflectivity to the local Radon domain is done using 3D Fourier-

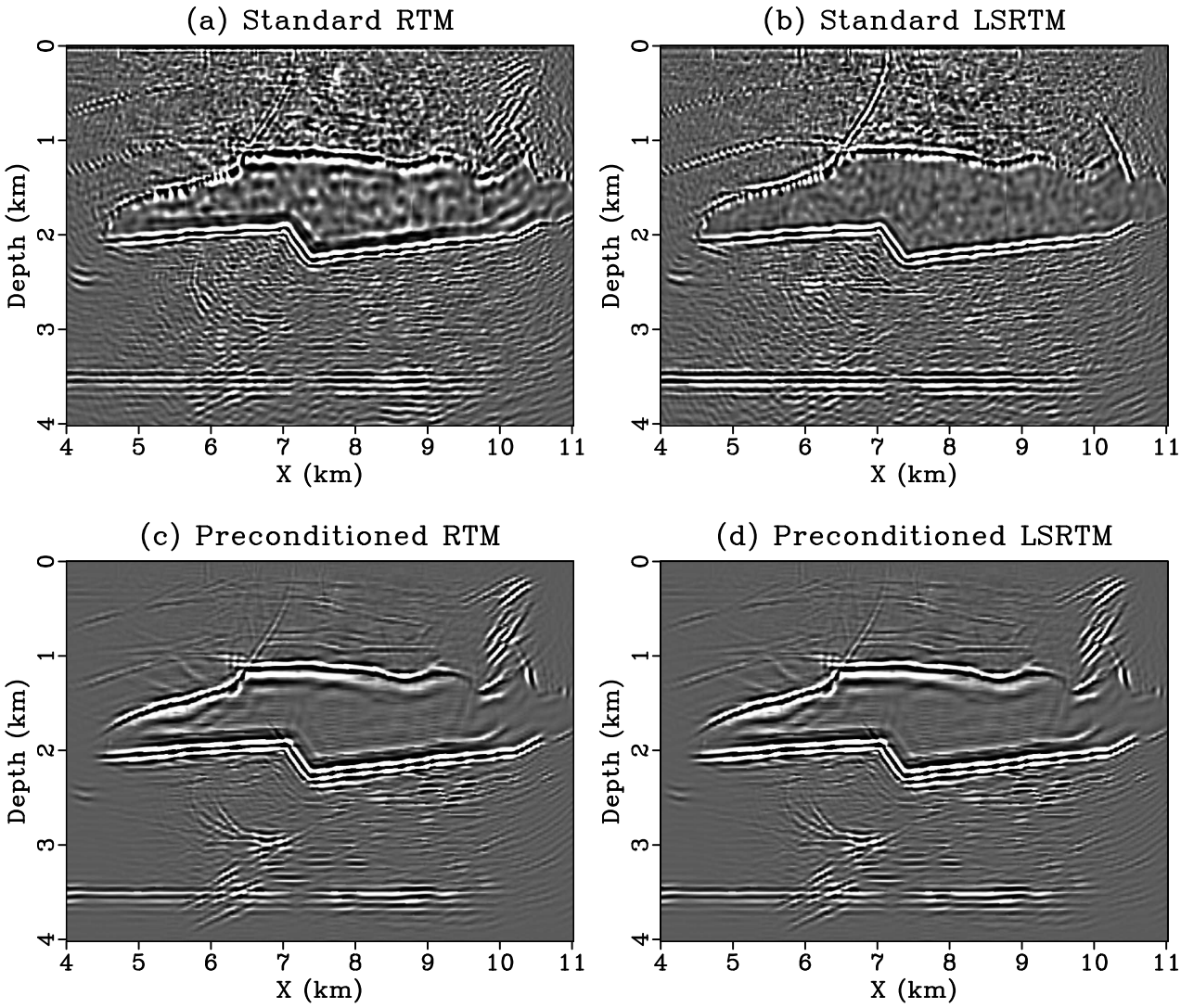

Figure 4.11: Comparison between images from (a) standard isotropic RTM, (b) standard isotropic LSRTM after 5 iterations, (c) preconditioned isotropic RTM, and (d) preconditioned isotropic LSRTM after 5 iterations. The recorded data in this case had free-surface multiples.

based Radon operators. Numerical tests on synthetic, 2D land and 3D marine data show that if the recorded data are noisy or undersampled, the standard RTM and LSRTM images suffer from poor SNR and strong aliasing artifacts. The Gabor-based preconditioning approach is shown to produce reliable images with good SNR and fewer artifacts than standard RTM or standard LSRTM both at shallow and at deeper depths. A disadvantage of this method is that it is quite sensitive to the presence of free-surface multiples or internal multiples in the data. This is because the false reflection events from these multiples have strong local dip amplitudes and as such, they are difficult to filter out. For field data applications, it is important to eliminate these multiples from the data before the proposed preconditioned LSRTM algorithm is applied. Investigating the selection of an appropriate weighting strategy to take into account the effect of multiples is a topic of future research.

Chapter 5

Conclusions and Future Work

5.1 Main Results

In this thesis, I developed novel migration and tomography methods for compensation of attenuation during imaging and estimation of the subsurface attenuation parameter, Q. The main results of my thesis are summarized below.

5.1.1 Q-LSRTM

In chapter 2, I proposed to use the time-domain visco-acoustic wave-equation in the context of LSRTM to compensate for the distortion in amplitudes and phases of seismic waves propagating in highly attenuative layers. The linear modeling operator, $\mathbf{L}$, used to compute the perturbed wavefields during Q-LSRTM is obtained by linearizing the visco-acoustic wave-equation w.r.t perturbations in the bulk modulus or the reflectivity. The adjoint equations, $\mathbf{L}^T$, used to back-propagate the data residual at every iteration and the gradient for Q-LSRTM are then derived using the adjoint-state method. Using numerical tests on synthetic and field data, I showed that Q-LSRTM can produce images with better balanced amplitudes and accurately positioned reflectors compared to acoustic RTM and LSRTM.

5.1.2 Wave-equation Q tomography

In chapter 3, I developed a novel wave-equation Q tomography method that can be used to estimate the background Q model, which can be used with any Q-PSDM algorithm to obtain images with balanced amplitudes and high resolution in areas where there is strong attenuation. I used a skeletonized misfit function which is the sum of the squared differences between the observed and the predicted peak/centroid frequency-shifts of the early-arrivals in the observed and in the predicted traces.

Using a connective function, which is the normalized zero-lag cross-correlation between an observed and a predicted trace, and the implicit function theorem, I then showed that the gradient for WQ is obtained by a zero-lag cross-correlation between the forward propagated source wavefield and the backpropagated residual traces, obtained by weighting the observed traces by the frequency-shifts between them and the corresponding predicted traces. Unlike conventional ray-based Q tomography which has high-frequency assumptions about the data, the residuals in WQ are smeared along transmission wavepaths computed from finite-difference solutions to the time-domain visco-acoustic wave equation characterized by the standard linear solid (SLS) mechanism. Through numerical tests on synthetic and a crosswell field dataset, I demonstrated that an improved accuracy of the Q model by wave-equation Q tomography led to a noticeable improvement in the migration image quality.

5.1.3 LSRTM using Radon-based preconditioning

In chapter 4, I investigated the use of a directional Gabor-based preconditioning approach for LSRTM. I formulated the LSRTM problem as a sparse optimization problem that minimizes the L-1 norm of the local Radon transform of the image subject to the condition that the data residual is minimized to an acceptable tolerance level. Using a model re-parameterization and re-weighting approach, I reformulated the sparse optimization problem as a preconditioning problem, where the weights used

to sparsify the image in the local Radon domain are chosen based on the amplitudes of the events in the dip domain. I use 3D discrete Radon transform operators in the Fourier domain for forward mapping of the reflectivity to the local Radon domain and back. Through numerical tests on synthetic and on 2D land and 3D marine data, I showed that the Radon-based preconditioning approach can produce images with fewer migration artifacts and better SNR than standard RTM or LSRTM in a few iterations. The proposed approach is also shown to produce images with better focusing and fewer artifacts at greater depths where migration velocity errors accumulate.

5.2 Future Research Work

I list below some potential research directions from my thesis that can be looked into for future work.

5.2.1 Q-LSRTM and WQ using alternative time-domain Q formulations

In Chapters 2 and 3, I used the time-domain visco-acoustic wave-equation characterized by the standard linear solid mechanism with one relaxation function to compute the visco-acoustic Green's functions. For typical exploration problems, where the bandwidth of the data used for migration and tomography is approximately 3-30 Hz, the use of a single relaxation mechanism has been shown to be sufficient to accurately model the effect of Q ([Blanch et al., 1995]; [Zhu et al., 2013]). Using more than one relaxation mechanism will significantly increase the cost of Q-LSRTM and WQ. The use of alternative time-domain formulations of the visco-acoustic wave-equation without memory variables ([Bai et al., 2013]) or the decoupled Q equations involving fractional Laplacians and its low-rank formulation ([Zhu and Harris, 2014]) can be looked into to compute the visco-acoustic Green's functions in the context of

Q-LSRTM or WQ.

5.2.2 Wave-equation Q tomography in the frequency domain

A correct modeling scheme for Q in the time-domain should yield a constant Q behavior over the range of frequencies being investigated. Ensuring such a constant Q behavior can be difficult in the time-domain when the bandwidth of the observed data is very large. However, the formulation is slightly easier in the frequency domain, where Q can be taken to be bandwidth dependent, depending on the range of frequencies/bandwidths being inverted. For a given frequency, ω, Q(ω) is defined as the ratio of the real part of the complex bulk modulus to the imaginary part of the bulk modulus (see equation C.10 in Appendix C). Thus, the WQ method proposed in Chaper 3, can be extended to the frequency-domain where the background Q model can be estimated for a given frequency bandwidth.

5.2.3 Wave-equation reflection Q tomography

In Chapter 3, wave-equation Q tomography is formulated for transmission arrivals where the peak-frequency shifts between observed and predicted traces are smeared along transmission wavepaths. Unlike diving waves, reflection waves provide important information about the deeper parts of the subsurface. To invert for deeper Q anomalies in the subsurface which are beyond the reach of diving waves, the proposed WQ method can be extended to reflection-based wave-equation Q tomography where the centroid frequency shifts between predicted and observed traces are smeared along reflection wavepaths, computed by time-domain finite-difference solutions of the visco-acoustic wave-equation. Since reflection events are not generally isolated in the data, the reflection events can be windowed in the time-domain. The length of the time-window can be chosen based on the dominant period of the arrivals.

5.2.4 LSRTM using Radon-based preconditioning in the presence of free-surface or internal multiples

In Chapter 4, I showed that the local Gabor-based preconditioning approach can provide significant uplifts in the quality of the inverted images when the recorded data are free from free-surface or internal multiples. In the present implementation, the weights for the preconditioning are chosen locally based on the amplitudes of the events in the dip-domain. However, if the recorded data are contaminated by multiples, the false reflection events generated by these multiples will have strong amplitude signatures in the dip domain. To mitigate the effect of multiples, alternative strategies for choosing the weights could be investigated.

6 Papers published and submitted

Chapters 2-4 of the thesis have been published or submitted to Geophysics:

- **Dutta, G.** and G. T. Schuster, 2014, Attenuation compensation for least-squares reverse time migration using the viscoacoustic-wave equation: *Geophysics*, 79, S251-S262 (Chapter 2).
- **Dutta, G.** and G. T. Schuster, 2016, Wave-equation Q tomography: submitted to *Geophysics* (Chapter 3).
- **Dutta, G.**, M. Giboli, C. Agut, P. Williamson, and G. T. Schuster, Least-squares reverse time migration with directional Gabor-based preconditioning: submitted to *Geophysics* (Chapter 4).

In addition to the publications listed above, the following publications in journals and conferences have also been a part of my PhD work:

- **Dutta, G.**, A. AlTheyab, A. Tarhini, S. Hanafy, and G. T. Schuster, 2016, Extracting 220 Hz information from 55 Hz field data by near-field superresolution imaging: *Geophysical Journal International* (accepted subject to revision).
- **Dutta, G.**, 2015, Sparse least-squares reverse time migration using seislets: *SEG Technical Program Expanded Abstracts*, 810, 4232-4237.

- **Dutta, G.**, M. Sinha, and G. T. Schuster, 2014, A cross-correlation objective function for least-squares migration and visco-acoustic imaging: *SEG Technical Program Expanded Abstracts*, 761, 3985-3990.

- **Dutta, G.**, Y. Huang, W. Dai, X. Wang, and G.T. Schuster, 2014, Making the most out of the least (squares migration): *SEG Technical Program Expanded Abstracts*, 840, 4405-4410.

- **Dutta, G.** and G. T. Schuster, 2013, Multisource early-arrival waveform inversion of crosswell data: *SEG Workshop on Full Waveform Inversion: From Near Surface to Deep, Oman.*

REFERENCES

[Aki and Richards, 1980] Aki, K. and P. G. Richards, 1980, Quantitative seismology: Freeman Publication Co., San Francisco.

[Aldawood et al., 2014] Aldawood, A., I. Hoteit, and T. Alkhalifah, 2014, The possibilities of compressed-sensing-based kirchhoff prestack migration: Geophysics, **79**, S113–S120.

[Anagaw and Sacchi, 2012] Anagaw, A. Y. and M. D. Sacchi, 2012, Edge-preserving seismic imaging using the total variation method: Journal of Geophysics and Engineering, **9**, 138.

[Aoki and Schuster, 2009] Aoki, N. and G. T. Schuster, 2009, Fast least-squares migration with a deblurring filter: Geophysics, **74**, WCA83–WCA93.

[Averbuch and Shkolnisky, 2003] Averbuch, A. and Y. Shkolnisky, 2003, 3D fourier based discrete radon transform: Applied and Computational Harmonic Analysis, **15**, 33–69.

[Bai et al., 2013] Bai, J., G. Chen, D. Yingst, and J. Leveille, 2013, Attenuation compensation in viscoacoustic reverse time migration: SEG Technical Program Expanded Abstracts, **740**, 3825–3830.

[Bai and Yingst, 2013] Bai, J. and D. Yingst, 2013, Q estimation through waveform inversion: EAGE Technical Program Expanded Abstracts.

[Baysal et al., 1983] Baysal, E., D. D. Kosloff, and J. W. Sherwood, 1983, Reverse time migration: Geophysics, **48**, 1514–1524.

[Behura and Tsvankin, 2009] Behura, J. and I. Tsvankin, 2009, Estimation of interval anisotropic attenuation from reflection data: Geophysics, **74**, A69–A74.

[Bickel and Natarajan, 1985] Bickel, S. and R. Natarajan, 1985, Plane wave Q deconvolution: Geophysics, **50**, 1426–1439.

[Blanch and Symes, 1995] Blanch, J. and W. Symes, 1995, Efficient iterative viscoacoustic linearized inversion: SEG Technical Program Expanded Abstracts, **171**, 627–630.

[Blanch et al., 1995] Blanch, J. O., J. O. Robertsson, and W. W. Symes, 1995, Modeling of a constant Q: methodology and algorithm for an efficient and optimally inexpensive viscoelastic technique: Geophysics, **60**, 176–184.

[Boonyasiriwat and Schuster, 2010] Boonyasiriwat, C. and G. Schuster, 2010, 3D multisource full-waveform inversion using dynamic quasi-monte carlo phase encoding: EGU General Assembly Conference Abstracts, **12**, 7298.

[Brzostowski and McMechan, 1992] Brzostowski, M. A. and G. A. McMechan, 1992, 3D tomographic imaging of near-surface seismic velocity and attenuation: Geophysics, **57**, 396–403.

[Cabrales-Vargas and Marfurt, 2013] Cabrales-Vargas, A. and K. J. Marfurt, 2013, Amplitude-preserving imaging of aliased data using preconditioned kirchhoff least-squares depth migration: SEG Technical Program Expanded Abstracts, **721**, 3726–3730.

[Calvetti et al., 2000] Calvetti, D., S. Morigi, L. Reichel, and F. Sgallari, 2000, Tikhonov regularization and the l-curve for large discrete ill-posed problems: Journal of computational and applied mathematics, **123**, 423–446.

[Candes et al., 2006a] Candes, E., L. Demanet, D. Donoho, and L. Ying, 2006a, Fast discrete curvelet transforms: Multiscale Modeling & Simulation, **5**, 861–899.

[Candes et al., 2006b] Candes, E. J., J. K. Romberg, and T. Tao, 2006b, Stable signal recovery from incomplete and inaccurate measurements: Communications on pure and applied mathematics, **59**, 1207–1223.

[Carcione, 2007] Carcione, J. M., 2007, Wave fields in real media: Wave propagation in anisotropic, anelastic, porous and electromagnetic media, volume **38**: Elsevier.

[Carcione et al., 1988] Carcione, J. M., D. Kosloff, and R. Kosloff, 1988, Wave propagation simulation in a linear viscoacoustic medium: Geophysical Journal International, **93**, 393–401.

[Carter and Kendall, 2006] Carter, A. J. and J.-M. Kendall, 2006, Attenuation anisotropy and the relative frequency content of split shear waves: Geophysical Journal International, **165**, 865–874.

[Cavalca et al., 2011] Cavalca, M., I. Moore, L. Zhang, S. L. Ng, R. Fletcher, and M. Bayly, 2011, Raybased tomography for Q estimation and Q compensation in complex media: SEG Technical Program Expanded Abstracts, **785**, 3989–3993.

[Chen et al., 1990] Chen, S., L. Zimmerman, and J. Tugnait, 1990, Subsurface imaging using reversed vertical seismic profiling and crosshole tomographic methods: Geophysics, **55**, 1478–1487.

[Christensen, 1982] Christensen, R. M., 1982, Theory of viscoelasticity: An introduction: New York: Academic Press.

[Dai and West, 1994] Dai, N. and G. West, 1994, Inverse Q migration: SEG Technical Program Expanded Abstracts, **389**, 1418–1421.

[Dai et al., 2010] Dai, W., C. Boonyasiriwat, and G. T. Schuster, 2010, 3D multi-source least-squares reverse time migration: SEG Technical Program Expanded Abstracts, **29**, 3120–3124.

[Dai et al., 2012] Dai, W., P. Fowler, and G. T. Schuster, 2012, Multisource least-squares reverse time migration: Geophysical Prospecting, **60**, 681–695.

[Dai and Schuster, 2009] Dai, W. and G. T. Schuster, 2009, Least-squares migration of simultaneous sources data with a deblurring filter: SEG Technical Program Expanded Abstracts, **28**, 2090–2994.

[Dai and Schuster, 2010a] ——, 2010a, Multi-source wave equation least-squares migration with a deblurring filter: EAGE Technical Program Expanded Abstracts, 276.

[Dai and Schuster, 2010b] ——, 2010b, Multi-source wave equation least-squares migration with a deblurring filter: EAGE Technical Program Expanded Abstracts, 276.

[Dai et al., 2011] Dai, W., X. Wang, and G. T. Schuster, 2011, Least-squares migration of multisource data with a deblurring filter: Geophysics, **76**, R135–R146.

[Duquet et al., 2000] Duquet, B., K. J. Marfurt, and J. A. Dellinger, 2000, Kirchhoff modeling, inversion for reflectivity, and subsurface illumination: Geophysics, **65**, 1195–1209.

[Dutta, 2015] Dutta, G., 2015, Sparse least-squares reverse time migration using seislets: SEG Technical Program Expanded Abstracts, **810**, 4232–4237.

[Dutta et al., 2014] Dutta, G., Y. Huang, W. Dai, X. Wang, and G. T. Schuster, 2014, Making the most out of the least (squares migration): SEG Technical Program Expanded Abstracts, **840**, 4405–4410.

[Dutta and Schuster, 2014] Dutta, G. and G. T. Schuster, 2014, Attenuation compensation for least-squares reverse time migration using the viscoacoustic-wave equation: Geophysics, **79**, S251–S262.

[Fletcher et al., 2012] Fletcher, R., D. Nichols, and M. Cavalca, 2012, Wavepath-consistent effective Q estimation for Q-compensated reverse time migration: EAGE Technical Program Expanded Abstracts.

[Fomel and Guitton, 2006] Fomel, S. and A. Guitton, 2006, Regularizing seismic inverse problems by model reparameterization using plane-wave construction: Geophysics, **71**, A43–A47.

[Gamar et al., 2015] Gamar, F., D. Carotti, P. Guillaume, A. Gacha, and L. L. CGG, 2015, Success of high-resolution volumetric Q-tomography in the automatic detection of gas anomalies on offshore brunei data: 2015 SEG Annual Meeting.

[Gherasim et al., 2010] Gherasim, M., R. Crider, S. Davis, E. Shoshitaishvili, M. Trout, Y. Quist, and C. Casler, 2010, Q-compensation study at the thunder horse field, gulf of mexico: The Leading Edge, **29**, 408–413.

[Hargreaves and Calvert, 1991] Hargreaves, N. and A. Calvert, 1991, Inverse Q filtering by fourier transform: Geophysics, **56**, 519–527.

[Harlan, 1995] Harlan, W. S., 1995, Regularization by model reparameterization: Citeseer.

[He et al., 2013] He, Y., K. Xin, and Y. Xie, 2013, 3D Q tomographic inversion using adaptive centroid frequency shift for compensating absorption and dispersion.

[Herrmann et al., 2009] Herrmann, F., C. Brown, Y. Erlangga, and P. Moghaddam, 2009, Curvelet-based migration preconditioning and scaling: Geophysics, **74**, A41–A46.

[Herrmann and Li, 2012] Herrmann, F. J. and X. Li, 2012, Efficient least-squares imaging with sparsity promotion and compressive sensing: Geophysical prospecting, **60**, 696–712.

[Huang and Schuster, 2012] Huang, Y. and G. T. Schuster, 2012, Multisource least-squares migration of marine streamer data with frequency-division encoding: Geophysical Prospecting, **60**, 663–680.

[Kingsbury, 2001] Kingsbury, N., 2001, Complex wavelets for shift invariant analysis and filtering of signals: Applied and computational harmonic analysis, **10**, 234–253.

[Kjartansson, 1979] Kjartansson, E., 1979, Constant Q-wave propagation and attenuation: Journal of Geophysical Research, **84**, 4737–4748.

[Lailly, 1984] Lailly, P., 1984, Migration methods: partial but efficient solutions to the seismic inverse problem: Inverse problems of acoustic and elastic waves, **51**, 1387–1403.

[Lin and Lianjie, 2015] Lin, Y. and L. Lianjie, 2015, Least-squares reverse-time migration with modified total-variation regularization: SEG Technical Program Expanded Abstracts, **816**, 4264–4269.

[Liu et al., 1976] Liu, H.-P., D. L. Anderson, and H. Kanamori, 1976, Velocity dispersion due to anelasticity; implications for seismology and mantle composition: Geophysical Journal International, **47**, 41–58.

[Luo and Schuster, 1991a] Luo, Y. and G. T. Schuster, 1991a, Wave equation inversion of skeletalized geophysical data: Geophysical Journal International, **105**, 289–294.

[Luo and Schuster, 1991b] ———, 1991b, Wave-equation traveltime inversion: Geophysics, **56**, 645–653.

[Mavko and Nur, 1979] Mavko, G. M. and A. Nur, 1979, Wave attenuation in partially saturated rocks: Geophysics, **44**, 161–178.

[McMechan, 1983] McMechan, G., 1983, Migration by extrapolation of time-dependent boundary values: Geophysical Prospecting, **31**, 413–420.

[Miller et al., 2005] Miller, M. A., N. G. Kingsbury, and R. W. Hobbs, 2005, Seismic image reconstruction using complex wavelets: Electronic Imaging, 27–35.

[Morton and Ober, 1998] Morton, S. A. and C. C. Ober, 1998, Faster shot record depth migrations using phase encoding: SEG Technical Program Expanded Abstracts, **298**, 1131–1134.

[Nemeth et al., 1999] Nemeth, T., C. Wu, and G. T. Schuster, 1999, Least-squares migration of incomplete reflection data: Geophysics, **64**, 208–221.

[Nocedal and Wright, 1999] Nocedal, J. and S. Wright, 1999, Numerical optimization: Springer Verlag.

[Plessix, 2006] Plessix, R., 2006, A review of the adjoint-state method for computing the gradient of a functional with geophysical applications: Geophysical Journal International, **167**, 495–503.

[Plessix and Mulder, 2004] Plessix, R.-E. and W. A. Mulder, 2004, Frequency-domain finite-difference amplitude-preserving migration: Geophysical Journal International, **157**, 975–987.

[Quan and Harris, 1997] Quan, Y. and J. M. Harris, 1997, Seismic attenuation tomography using the frequency shift method: Geophysics, **62**, 895–905.

[Robertsson et al., 1994] Robertsson, J. O., J. O. Blanch, and W. W. Symes, 1994, Viscoelastic finite-difference modeling: Geophysics, **59**, 1444–1456.

[Romero et al., 2000] Romero, L. A., D. C. Ghiglia, C. C. Ober, and S. A. Morton, 2000, Phase encoding of shot records in prestack migration: Geophysics, **65**, 426–436.

[Schuster, 1993] Schuster, G. T., 1993, Least-squares cross-well migration: SEG Technical Program Expanded Abstracts, **12**, 110–113.

[Schuster et al., 2011] Schuster, G. T., X. Wang, Y. Huang, W. Dai, and C. Boonyasiriwat, 2011, Theory of multisource crosstalk reduction by phase-encoded statics: Geophysical Journal International, **184**, 1289–1303.

[Shen et al., 2015] Shen, Y., B. Biondi, and R. Clapp, 2015, Wave-equation based Q tomography from angle-domain common image gathers: SEG Technical Program Expanded Abstracts, **829**, 4334–4338.

[Shen et al., 2014] Shen, Y., B. Biondi, R. Clapp, and D. Nichols, 2014, Wave-equation migration Q analysis (WEMQA): SEG Technical Program Expanded Abstracts, **718**, 3757–3762.

[Shen and Zhu, 2015] Shen, Y. and T. Zhu, 2015, Image-based Q tomography using reverse time Q migration: SEG Technical Program Expanded Abstracts, **708**, 3694–3698.

[Stolt and Benson, 1986] Stolt, R. H. and A. K. Benson, 1986, Seismic migration: Theory and practice, volume **5**: Geophysical Press London.

[Suh et al., 2012] Suh, S., K. Yoon, J. Cai, and B. Wang, 2012, Compensating visco-acoustic effects in anisotropic resverse-time migration: SEG Technical Program Expanded Abstracts, **649**, 1–5.

[Sun et al., 2015a] Sun, J., S. Fomel, and T. Zhu, 2015a, Preconditioning least-squares RTM in viscoacoustic media by Q-compensated RTM: SEG Technical Program Expanded Abstracts, **758**, 3959–3965.

[Sun et al., 2015b] Sun, J., T. Zhu, and S. Fomel, 2015b, Viscoacoustic modeling and imaging using low-rank approximation: Geophysics, **80**, A103–A108.

[Tang, 2009] Tang, Y., 2009, Target-oriented wave-equation least-squares migration/inversion with phase-encoded Hessian: Geophysics, **74**, WCA95–WCA107.

[Tarantola, 1984] Tarantola, A., 1984, Linearized inversion of seismic reflection data: Geophysical prospecting, **32**, 998–1015.

[Teng et al., 2013] Teng, K., J. Zhou, X. Wu, Y. Zhou, T. Brothers, S. Bergler, and C. Gibson, 2013, Compensating attenuation due to gas cloud through QPSDM: Case study from offshore brunei: EAGE Workshop on Seismic Attenuation.

[Trad et al., 2003] Trad, D., T. Ulrych, and M. Sacchi, 2003, Latest views of the sparse radon transform: Geophysics, **68**, 386–399.

[Valenciano and Chemingui, 2013] Valenciano, A. and N. Chemingui, 2013, Tomographic Q estimation for viscoacoustic imaging: EAGE Technical Program Expanded Abstracts.

[Valenciano et al., 2012] Valenciano, A., N. Chemingui, et al., 2012, Viscoacoustic imaging: Tomographic Q estimation and migration compensation: 2012 SEG Annual Meeting.

[Valenciano et al., 2011] Valenciano, A., N. Chemingui, D. Whitmore, and S. Brandsberg-Dahl, 2011, Wave equation migration with attenuation and anisotropy compensation: SEG Technical Program Expanded Abstracts, **45**, 232–236.

[Wang and Sacchi, 2007] Wang, J. and M. D. Sacchi, 2007, High-resolution wave-equation amplitude-variation-with-ray-parameter (avp) imaging with sparseness constraints: Geophysics, **72**, S11–S18.

[Wang, 2008] Wang, Y., 2008, Inverse Q- filtered migration: Geophysics, **73**, S1–S6.

[Wang and Zhang, 2014] Wang, Y. and J. Zhang, 2014, Pseudo 2D elastic waveform inversion for Q factor in the near surface: SEG Technical Program Expanded Abstracts, **386**, 2019–2023.

[Whitmore, 1983] Whitmore, N., 1983, Iterative depth migration by backward time propagation: SEG Technical Program Expanded Abstracts, **180**, 382–385.

[Wong et al., 2011] Wong, M., S. Ronen, and B. Biondi, 2011, Least-squares reverse time migration/inversion for ocean bottom data: A case study: SEG Technical Program Expanded Abstracts, **30**, 2369–2373.

[Woodward, 1992] Woodward, M. J., 1992, Wave-equation tomography: Geophysics, **57**, 15–26.

[Xie et al., 2009] Xie, Y., K. Xin, J. Sun, C. Notfors, A. Biswal, and M. Balasubramaniam, 2009, 3D prestack depth migration with compensation for frequency dependent absorption and dispersion: SEG Technical Program Expanded Abstracts, **586**, 2919–2923.

[Xin et al., 2008] Xin, K., B. Hung, S. Birdus, and J. Sun, 2008, 3D tomographic amplitude inversion for compensating amplitude attenuation in the overburden: SEG Technical Program Expanded Abstracts, **652**, 3239–3243.

[Yu et al., 2002] Yu, Y., R. Lu, and M. Deal, 2002, Compensation for the effects of shallow gas attenuation with viscoacoustic wave equation migration: SEG Technical Program Expanded Abstracts, **523**, 2062–2065.

[Zhang et al., 2010] Zhang, Y., P. Zhang, and H. Zhang, 2010, Compensating for visco-acoustic effects in reverse-time migration: SEG Technical Program Expanded Abstracts, **619**, 3160–3164.

[Zhou et al., 1995] Zhou, C., W. Cai, Y. Luo, G. T. Schuster, and S. Hassanzadeh, 1995, Acoustic wave-equation traveltime and waveform inversion of crosshole seismic data: Geophysics, **60**, 765–773.

[Zhou et al., 2011] Zhou, J., S. Birdus, B. Hung, K. H. Teng, Y. Xie, D. Chagalov, A. Cheang, D. Wellen, and J. Garrity, 2011, Compensating attenuation due to shallow gas through Q tomography and Q-PSDM, a case study in Brazil: SEG Technical Program Expanded Abstracts, **650**, 3332–3336.

[Zhu et al., 2013] Zhu, T., J. M. Carcione, and J. M. Harris, 2013, Approximating constant-q seismic propagation in the time domain: Geophysical prospecting, **61**, 931–940.

[Zhu and Harris, 2014] Zhu, T. and J. M. Harris, 2014, Modeling acoustic wave propagation in heterogeneous attenuating media using decoupled fractional laplacians: Geophysics, **79**, T105–T116.

[Zhu and Harris, 2015] ———, 2015, Improved seismic image by Q-compensated reverse time migration: Application to crosswell field data, west Texas: Geophysics, **80**, B61–B67.

[Zhu et al., 2014] Zhu, T., J. M. Harris, and B. Biondi, 2014, Q-compensated reverse-time migration: Geophysics, **79**, S77–S87.

[Zhu and Tsvankin, 2006] Zhu, Y. and I. Tsvankin, 2006, Plane-wave propagation in attenuative transversely isotropic media: Geophysics, **71**, T17–T30.

[Zhu and Tsvankin, 2007] ———, 2007, Plane-wave attenuation anisotropy in orthorhombic media: Geophysics, **72**, D9–D19.

[Zhu et al., 2007] Zhu, Y., I. Tsvankin, P. Dewangan, and K. van Wijk, 2007, Physical modeling and analysis of p-wave attenuation anisotropy in transversely isotropic media: Geophysics, **72**, D1–D7.

APPENDICES

A Relaxation function for a SLS model

In this appendix, I derive the relation between the relaxation function G for a visco-acoustic medium represented by a standard linear solid (SLS) mechanical model. For the SLS model shown in Figure A.1, the stress-strain relation for the individual elements can be written as,

$$\begin{aligned} \sigma &= k_1 \epsilon_1, \\ \sigma_1 &= \eta \frac{\partial \epsilon_2}{\partial t}, \\ \sigma_2 &= k_2 \epsilon_2, \end{aligned} \tag{A.1}$$

and for the whole model as,

$$\begin{aligned} \sigma &= \sigma_1 + \sigma_2, \\ \epsilon &= \epsilon_1 + \epsilon_2. \end{aligned} \tag{A.2}$$

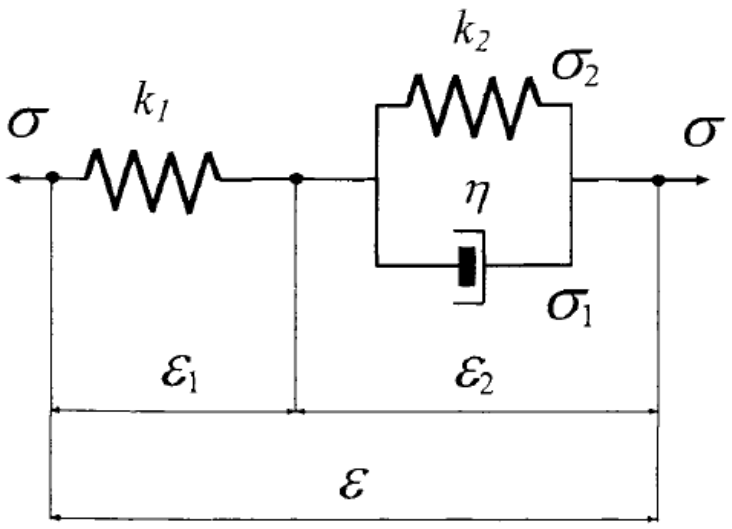

Figure A.1: Zener model or a standard linear solid (SLS) model ([Carcione, 2007]).

From equations A.1 and A.2, we get the stress-strain relation as,

$$\sigma + \tau_\sigma \frac{\partial \sigma}{\partial t} = M_R(\epsilon + \tau_\epsilon \frac{\partial \epsilon}{\partial t}). \tag{A.3}$$

Here M_R is the relaxed modulus, τ_σ and τ_ϵ are the relaxation times given by

$$\begin{aligned} M_R &= \frac{k_1 k_2}{k_1 + k_2}, \\ \tau_\sigma &= \frac{\eta}{k_1 + k_2}, \\ \tau_\epsilon &= \frac{\eta}{k_2}. \end{aligned} \tag{A.4}$$

The relaxation function $G(t)$ for any medium can be obtained by measuring the stress after imposing a rapidly constant unit strain in a relaxed sample of the medium. If

the strain is given by $\epsilon(t) = H(t)$, where $H(t)$ is a Heaviside function, then

$$\begin{aligned}
\sigma(t) &= \frac{\partial G(t)}{\partial t} * \epsilon(t), \\
\sigma(t) &= G(t) * \frac{\partial \epsilon}{\partial t}, \\
\sigma(t) &= G(t) * \delta(t), \\
\sigma(t) &= G(t). \qquad \text{(A.5)}
\end{aligned}$$

Thus, from equation A.3 we get,

$$\begin{aligned}
G(t) + \tau_\sigma \frac{\partial G(t)}{\partial t} &= M_R \left(H(t) + \tau_\epsilon \delta(t) \right), \\
G(t) &= M_R \left(1 - \left(1 - \frac{\tau_\epsilon}{\tau_\sigma} \right) e^{-t/\tau_\sigma} \right) H(t). \qquad \text{(A.6)}
\end{aligned}$$

For a series of L standard linear solids connected in parallel, as shown in Figure A.2, equation A.6 has the general form,

$$G(t) = M_R \left(1 - \sum_{l=1}^{L} \left(1 - \frac{\tau_{\epsilon l}}{\tau_{\sigma l}} \right) e^{-t/\tau_{\sigma l}} \right) H(t). \qquad \text{(A.7)}$$

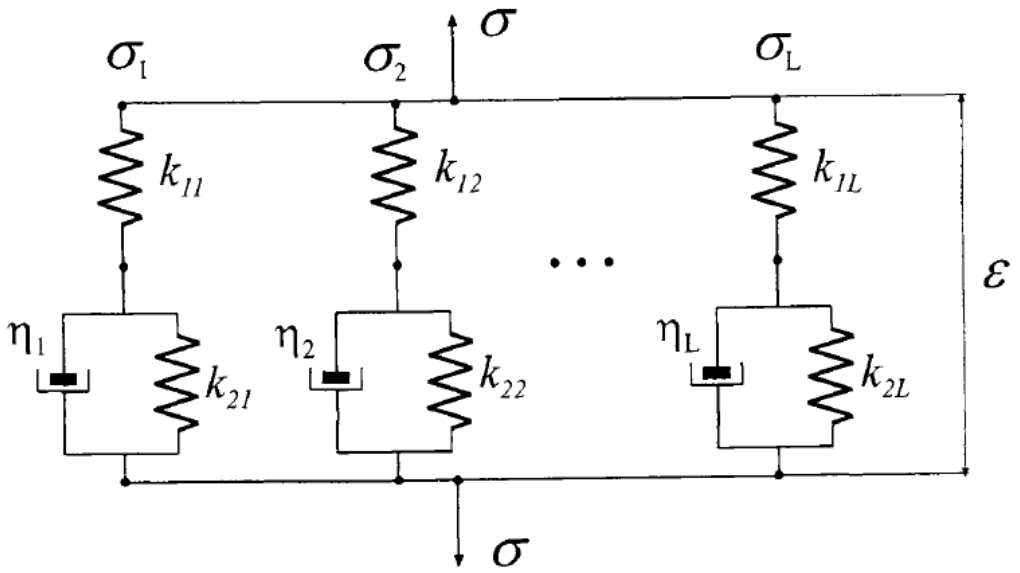

Figure A.2: Generalized Zener model with L standard linear solid models connected in parallel ([Carcione, 2007]).

B Equations of motion for a visco-acoustic medium

For an anelastic medium, the current value of the stress tensor depends upon the history of the strain tensor. The stress-strain relation in such a medium can be expressed as ([Liu et al., 1976]; [Christensen, 1982]; [Carcione et al., 1988]; [Robertsson et al., 1994])

$$\sigma = G * \dot{\epsilon} = \dot{G} * \epsilon \tag{B.1}$$

Here $*$ denotes convolution in time, $G(t)$ is the relaxation function, that transforms a strain history, $\epsilon(t)$, into the corresponding stress history, $\sigma(t)$. The relaxation function determines the behavior of a material. For a lossless medium, equation B.1 becomes

$$\sigma = M_e \epsilon, \tag{B.2}$$

where M_e is the elastic modulus. To construct a mechanical model to model visco-elasticity, two basic elements are required: weightless springs (where no inertial effects are present) to represent a elastic solid and dashpots (consisting of loosely fitting pistons in cylinders) to represent a visco-elastic fluid. The simplest cases are the Maxwell, Kelvin-Voigt and the Zener of standard linear solid (SLS) models. $G(t)$ in equation B.1 varies depending on the type of model used. For a SLS model with L standard linear solids connected in parallel, G has the general form ([Liu et al., 1976];

[Carcione et al., 1988]; [Robertsson et al., 1994]),

$$G(t) = K\left(1 - \sum_{l=1}^{L}\left(1 - \frac{\tau_{\epsilon l}}{\tau_{\sigma l}}\right)e^{-t/\tau_{\sigma l}}\right)H(t). \tag{B.3}$$

Here, K is the relaxed modulus of the medium and $H(t)$ is the heaviside function. $\tau_{\sigma l}$ and $\tau_{\epsilon l}$ are the stress and strain relaxation times for the l-th mechanism, respectively.

From the definition of pressure and strain we know that for a 1-D medium,

$$\sigma = -p, \tag{B.4}$$

and,

$$v_x = \frac{\partial v}{\partial x} = \dot{\epsilon}, \tag{B.5}$$

where v_x is the particle velocity in the X-direction. From equations B.1, B.3, B.4 and B.5 we have

$$\begin{aligned} -\dot{p} &= \dot{G} * v_x \\ -\dot{p} &= K\left(1 - \sum_{l=1}^{L}\left(1 - \frac{\tau_{\epsilon l}}{\tau_{\sigma l}}\right)\right)v_x + K\left(\sum_{l=1}^{L}\frac{1}{\tau_{\sigma l}}\left(1 - \frac{\tau_{\epsilon l}}{\tau_{\sigma l}}\right)e^{-t/\tau_{\sigma l}}\right)H(t) * v_x. \end{aligned} \tag{B.6}$$

[Carcione et al., 1988] simplified the convolution term in equation B.6 by introducing a memory variable term r_l. Equation B.6 then becomes

$$-\dot{p} \;=\; K\left(1 - \sum_{l=1}^{L}\left(1 - \frac{\tau_{\epsilon l}}{\tau_{\sigma l}}\right)\right)v_x + \sum_{l=1}^{L} r_l. \tag{B.7}$$

Here,

$$r_l = K\left(\frac{1}{\tau_{\sigma l}}\left(1-\frac{\tau_{\epsilon l}}{\tau_{\sigma l}}\right)e^{-t/\tau_{\sigma l}}\right)H(t) * v_x, \quad 1 \leq l \leq L. \tag{B.8}$$

Taking the time derivative of equation B.8, we get

$$\begin{aligned}
\dot{r_l} &= -\frac{1}{\tau_{\sigma l}}K\left(\frac{1}{\tau_{\sigma l}}\left(1-\frac{\tau_{\epsilon l}}{\tau_{\sigma l}}\right)e^{-t/\tau_{\sigma l}}\right)H(t) * v_x + K\left(\frac{1}{\tau_{\sigma l}}\left(1-\frac{\tau_{\epsilon l}}{\tau_{\sigma l}}\right)e^{-t/\tau_{\sigma l}}\right)\delta(t) * v_x, \\
\dot{r_l} &= -\frac{1}{\tau_{\sigma l}}r_l + K\frac{1}{\tau_{\sigma l}}\left(1-\frac{\tau_{\epsilon l}}{\tau_{\sigma l}}\right)v_x, \quad 1 \leq l \leq L.
\end{aligned} \tag{B.9}$$

From Newton's second law,

$$\dot{v_x} = -\frac{1}{\rho}P. \tag{B.10}$$

Equations B.9 and B.10 together describe the wave propagation in a 1D visco-acoustic medium represented by L sets of SLS's connected in parallel. In a more general form, equations B.9 and B.10 can be written as

$$\begin{aligned}
\frac{\partial \mathbf{v}}{\partial t} &= -\frac{1}{\rho}\nabla P, \\
\frac{\partial P}{\partial t} &= -K\left(1-\sum_{l=1}^{L}\left(1-\frac{\tau_{\epsilon l}}{\tau_{\sigma l}}\right)\right)(\nabla \cdot \mathbf{v}) - \sum_{l=1}^{L} r_l, \\
\frac{\partial r_l}{\partial t} &= -\frac{1}{\tau_{\sigma l}}r_l + K\frac{1}{\tau_{\sigma l}}\left(1-\frac{\tau_{\epsilon l}}{\tau_{\sigma l}}\right)(\nabla \cdot \mathbf{v}), \quad 1 \leq l \leq L,
\end{aligned} \tag{B.11}$$

where $\mathbf{v}$ is the particle velocity vector given by $\mathbf{v} = \begin{bmatrix} v_x & v_y & v_z \end{bmatrix}^T$. [Blanch et al., 1995] and [Zhu et al., 2013] showed that only one relaxation mechanism is sufficient for practical exploration purposes and Q can be assumed to be constant over a limited frequency bandwitdh of 5-30 Hz. Thus, for a single relaxation mechanism, i.e. $l = 1$,

equation B.11 becomes

$$\begin{aligned}
\frac{\partial \mathbf{v}}{\partial t} &= -\frac{1}{\rho}\nabla P, \\
\frac{\partial P}{\partial t} &= -K\frac{\tau_\epsilon}{\tau_\sigma}\left(\nabla.\mathbf{v}\right) - r, \\
\frac{\partial r}{\partial t} &= -\frac{1}{\tau_\sigma}\left(r + K\left(\frac{\tau_\epsilon}{\tau_\sigma} - 1\right)\left(\nabla.\mathbf{v}\right)\right).
\end{aligned} \tag{B.12}$$

Equation B.12 represents the time-domain visco-acoustic wave-equations based on the SLS model with a single relaxation mechanism.

C Relation between Q and the τ's

Hooke's law in a lossless medium is given by

$$\sigma = M_e \epsilon, \tag{C.1}$$

where M_e is the elastic modulus. For a lossy medium, the stress-strain relation becomes,

$$\sigma = G * \dot{\epsilon} = \dot{G} * \epsilon \tag{C.2}$$

Here $*$ denotes convolution in time, $G(t)$ is the relaxation function, that transforms a strain history, $\epsilon(t)$, into the corresponding stress history, $\sigma(t)$. The Fourier transform of C.2 gives

$$F(\sigma(\omega)) = M(\omega)F(\epsilon(\omega)), \tag{C.3}$$

where F is the Fourier transform operator and $M(\omega) = F(\dot{G}(\omega))$. $M(\omega)$ is also known as the complex modulus and can be represented as

$$M(\omega) = M_1(\omega) + iM_2(\omega). \tag{C.4}$$

The time-averaged strain-energy density can be written as

$$\langle V \rangle = \frac{1}{2} \langle \mathrm{Re}(\epsilon)\mathrm{Re}(M)\mathrm{Re}(\epsilon) \rangle = \frac{1}{4}\mathrm{Re}(\epsilon M \epsilon^*) = \frac{1}{4} M_1 |\epsilon|^2. \tag{C.5}$$

Similarly, the time-averaged dissipated energy density is given by

$$\langle D \rangle = \frac{1}{2} M_2 |\epsilon|^2. \tag{C.6}$$

The quality factor, Q, quantifies dissipation and is defined as twice the time-averaged strain-energy density divided by the time-averaged dissipated energy-density, i.e.,

$$Q = \frac{2\langle V \rangle}{\langle D \rangle} = \frac{M_1}{M_2} = \frac{\mathrm{Re}(M)}{\mathrm{Im}(M)}. \tag{C.7}$$

For a SLS model with a single relaxation mechanism, the stress-strain relation is given by

$$\sigma + \tau_\sigma \frac{\partial \sigma}{\partial t} = M_R(\epsilon + \tau_\epsilon \frac{\partial \epsilon}{\partial t}), \tag{C.8}$$

where $\sigma = M\epsilon$. Taking the Fourier transform of C.8, we get

$$M(\omega) = \frac{M_R(1 + i\omega\tau_\epsilon)}{1 + i\omega\tau_\sigma}. \tag{C.9}$$

Thus, from equation C.7, Q can be written as

$$Q(\omega) = \frac{\mathrm{Re}(M)}{\mathrm{Im}(M)} = \frac{1 + \omega^2 \tau_\sigma \tau_\epsilon}{\omega(\tau_\epsilon - \tau_\sigma)}. \tag{C.10}$$

The model has a relaxation peak at

$$\omega_0 = \frac{1}{\sqrt{\tau_\sigma \tau_\epsilon}}, \tag{C.11}$$

which can be verified by setting $\frac{dQ}{d\omega} = 0$ in equation C.10. ω_0 here is also defined as the center frequency of the source.

For practical exploration purposes, constant Q models are used to parameterize attenuation in rocks since it is usually hard to quantify the frequency dependence of Q. Also Q has been shown to be constant in many frequency bands. Thus, a more physical parameterization of τ_σ and τ_ϵ can be obtained at the center frequency ω_0. From equation C.10, at $\omega = \omega_0$ we get,

$$Q_0 = \frac{2}{\omega_0(\tau_\epsilon - \tau_\sigma)}. \tag{C.12}$$

τ_σ and τ_ϵ can be obtained from equations C.11 and C.12 as

$$\begin{aligned} \tau_\sigma &= \frac{\sqrt{1 + \frac{1}{Q_0^2}} - \frac{1}{Q_0}}{\omega_0}, \\ \tau_\epsilon &= \frac{1}{\omega_0^2 \tau_\sigma}. \end{aligned} \tag{C.13}$$

D Adjoint equations and gradient for Q-LSRTM

In matrix-vector notation, equation 2.12 can be rewritten as

$$\begin{bmatrix} \frac{\partial}{\partial t} & K(1+\tau)\nabla\cdot & 1 \\ \frac{1}{\rho}\nabla & \frac{\partial}{\partial t} & 0 \\ 0 & \frac{\tau}{\tau_\sigma}K\nabla\cdot & \frac{\partial}{\partial t}+\frac{1}{\tau_\sigma} \end{bmatrix} \begin{bmatrix} P \\ \mathbf{v} \\ r_p \end{bmatrix} = \begin{bmatrix} f \\ 0 \\ 0 \end{bmatrix}, \tag{D.1}$$

or in a more compact form,

$$\mathbf{S}\mathbf{w} = \mathbf{F}. \tag{D.2}$$

where

$$\mathbf{S} = \begin{bmatrix} \frac{\partial}{\partial t} & K(1+\tau)\nabla\cdot & 1 \\ \frac{1}{\rho}\nabla & \frac{\partial}{\partial t} & 0 \\ 0 & \frac{\tau}{\tau_\sigma}K\nabla\cdot & \frac{\partial}{\partial t}+\frac{1}{\tau_\sigma} \end{bmatrix}, \mathbf{w} = \begin{bmatrix} P \\ \mathbf{v} \\ r_p \end{bmatrix}, \text{ and}$$

$$\mathbf{F} = \begin{bmatrix} f \\ 0 \\ 0 \end{bmatrix}. \tag{D.3}$$

Here, $\mathbf{w}$ represents the state variables and $\mathbf{S}$ represents the forward modeling operator. The adjoint operator, $\mathbf{S}^*$, of $\mathbf{S}$ in this particular case is given by

$$\mathbf{S}^* = \begin{bmatrix} -\frac{\partial}{\partial t} & -\nabla \cdot \frac{1}{\rho} & 0 \\ -\nabla K(1+\tau) & -\frac{\partial}{\partial t} & -\nabla \frac{\tau}{\tau_\sigma} K \\ 1 & 0 & -\frac{\partial}{\partial t} + \frac{1}{\tau_\sigma} \end{bmatrix}. \tag{D.4}$$

The least-squares misfit functional, $J(\mathbf{m})$ for a model parameter $\mathbf{m}$ can be written as

$$J(\mathbf{m}) = \frac{1}{2}||\mathbf{w}(\mathbf{m}) - \mathbf{d}||^2 = \frac{1}{2}\left\langle \mathbf{w}(\mathbf{m}) - \mathbf{d}, \mathbf{w}(\mathbf{m}) - \mathbf{d} \right\rangle, \tag{D.5}$$

where $\mathbf{w}(\mathbf{m})$ and $\mathbf{d}$ represent the modeled and recorded data vectors, respectively. For a visco-acoustic medium, the model parameter $\mathbf{m}$ can be K, ρ, τ_σ or τ. The gradient of J is given by

$$\frac{\partial J(\mathbf{m})}{\partial \mathbf{m}} = \left\langle \frac{\partial \mathbf{w}(\mathbf{m})}{\partial \mathbf{m}}, \mathbf{w}(\mathbf{m}) - \mathbf{d} \right\rangle. \tag{D.6}$$

Now, for the system of equations,

$$\begin{aligned} &\mathbf{S}(\mathbf{m})\mathbf{w}(\mathbf{m}) = \mathbf{F} \\ \Rightarrow &\frac{\partial \mathbf{S}(\mathbf{m})}{\partial \mathbf{m}}\mathbf{w}(\mathbf{m}) + \mathbf{S}(\mathbf{m})\frac{\partial \mathbf{w}(\mathbf{m})}{\partial \mathbf{m}} = 0 \\ \Rightarrow &\frac{\partial \mathbf{w}(\mathbf{m})}{\partial \mathbf{m}} = -\mathbf{S}^{-1}(\mathbf{m})\frac{\partial \mathbf{S}(\mathbf{m})}{\partial \mathbf{m}}\mathbf{w}(\mathbf{m}). \end{aligned} \tag{D.7}$$

Inserting equation D.7 into equation D.6 we get

$$\begin{aligned}\frac{\partial J(\mathbf{m})}{\partial \mathbf{m}} &= -\left\langle \mathbf{S}^{-1}(\mathbf{m})\frac{\partial \mathbf{S}(\mathbf{m})}{\partial \mathbf{m}}\mathbf{w}(\mathbf{m}), \mathbf{w}(\mathbf{m}) - \mathbf{d} \right\rangle \\ &= -\left\langle \frac{\partial \mathbf{S}(\mathbf{m})}{\partial \mathbf{m}}\mathbf{w}(\mathbf{m}), \left(\mathbf{S}(\mathbf{m})^{-1}\right)^{*}\left(\mathbf{w}(\mathbf{m}) - \mathbf{d}\right) \right\rangle \\ &= -\left\langle \frac{\partial \mathbf{S}(\mathbf{m})}{\partial \mathbf{m}}\mathbf{w}(\mathbf{m}), \left(\mathbf{S}(\mathbf{m})^{-1}\right)^{*}\Delta\mathbf{d} \right\rangle \\ &= -\left\langle \frac{\partial \mathbf{S}(\mathbf{m})}{\partial \mathbf{m}}\mathbf{w}(\mathbf{m}), \mathbf{w}^{*}(\mathbf{m}) \right\rangle, \end{aligned} \tag{D.8}$$

where $*$ denotes the adjoint, $\Delta\mathbf{d}$ denotes the residual data vector and $\mathbf{w}^{*}$ denotes the adjoint or the residual wavefield, which is obtained by solving the adjoint equations,

$$\mathbf{S}^{*}(\mathbf{m})\mathbf{w}^{*}(\mathbf{m}) = \Delta\mathbf{d}. \tag{D.9}$$

$\mathbf{w}^{*}$ is also known as the adjoint state variable of $\mathbf{w}$. Let $\mathbf{w}^{*}$ be denoted as

$$\mathbf{w}^{*} = \begin{bmatrix} q & \mathbf{u} & s \end{bmatrix}^{T}, \tag{D.10}$$

where q is the adjoint state variable of the pressure wavefield P, $\mathbf{u}$ is the adjoint of the particle velocity vector $\mathbf{v}$ and s is the adjoint of the memory variable r_p. Assuming that we only record pressure seismograms, the residual vector $\Delta\mathbf{d}$ will have only one component, i.e.,

$$\Delta\mathbf{d} = \begin{bmatrix} \Delta d & 0 & 0 \end{bmatrix}^{T}. \tag{D.11}$$

Using equations D.4, D.9, D.10 and D.11, we get the adjoint-state equations for a

visco-acoustic medium as

$$\begin{bmatrix} -\frac{\partial}{\partial t} & -\nabla\cdot\frac{1}{\rho} & 0 \\ -\nabla K(1+\tau) & -\frac{\partial}{\partial t} & -\nabla\frac{\tau}{\tau_\sigma}K \\ 1 & 0 & -\frac{\partial}{\partial t}+\frac{1}{\tau_\sigma} \end{bmatrix} \begin{bmatrix} q \\ \mathbf{u} \\ s \end{bmatrix} = \begin{bmatrix} \Delta d \\ 0 \\ 0 \end{bmatrix}$$

$$\begin{aligned} &\frac{\partial q}{\partial t} + \nabla\cdot\left(\frac{1}{\rho}\mathbf{u}\right) = -\Delta d(\mathbf{x}_g, t; \mathbf{x}_s), \\ &\frac{\partial \mathbf{u}}{\partial t} + \left[\nabla K(1+\tau)q + \nabla\left(\frac{1}{\tau_\sigma}K\tau s\right)\right] = 0, \\ &\frac{\partial s}{\partial t} - \frac{s}{\tau_\sigma} - q = 0. \end{aligned} \tag{D.12}$$

For $\mathbf{m} = K$, the gradient in equation D.8 can be written as

$$\begin{aligned} \frac{\partial J}{\partial K} &= -\left\langle \frac{\partial \mathbf{S}}{\partial K}\mathbf{w}, \mathbf{w}^* \right\rangle \\ &= -\left\langle \begin{bmatrix} 0 & (1+\tau)\nabla\cdot & 0 \\ 0 & 0 & 0 \\ 0 & \frac{\tau}{\tau_\sigma}\nabla\cdot & 0 \end{bmatrix} \begin{bmatrix} P \\ \mathbf{v} \\ r_p \end{bmatrix}, \begin{bmatrix} q \\ \mathbf{u} \\ s \end{bmatrix} \right\rangle \\ &= -\int_0^T (1+\tau)(\nabla\cdot\mathbf{v})q + \frac{\tau}{\tau_\sigma}(\nabla\cdot\mathbf{v})s \;\; dt \;\; . \end{aligned} \tag{D.13}$$

E Adjoint-state method for wave-equation Q tomography

In matrix vector notation, equation 3.1 can be written as

$$\begin{bmatrix} \frac{\partial}{\partial t} & K(1+\tau)\nabla\cdot & 1 \\ \frac{1}{\rho}\nabla & \frac{\partial}{\partial t} & 0 \\ 0 & \frac{\tau}{\tau_\sigma}K\nabla\cdot & \frac{\partial}{\partial t}+\frac{1}{\tau_\sigma} \end{bmatrix} \begin{bmatrix} P \\ \mathbf{v} \\ r_p \end{bmatrix} = \begin{bmatrix} f \\ 0 \\ 0 \end{bmatrix}, \tag{E.1}$$

or in a more compact form,

$$\mathbf{S(m)w(m)} = \mathbf{F}. \tag{E.2}$$

where,

$$\mathbf{S} = \begin{bmatrix} \frac{\partial}{\partial t} & K(1+\tau)\nabla\cdot & 1 \\ \frac{1}{\rho}\nabla & \frac{\partial}{\partial t} & 0 \\ 0 & \frac{\tau}{\tau_\sigma}K\nabla\cdot & \frac{\partial}{\partial t}+\frac{1}{\tau_\sigma} \end{bmatrix}, \quad \mathbf{w} = \begin{bmatrix} P \\ \mathbf{v} \\ r_p \end{bmatrix},$$

$$\mathbf{F} = \begin{bmatrix} f \\ 0 \\ 0 \end{bmatrix}, \quad \mathbf{m} = \begin{bmatrix} K & \rho & \tau & \tau_\sigma \end{bmatrix}^T. \tag{E.3}$$

The perturbation of the state variable $\mathbf{w}$ w.r.t to the model variable $\mathbf{m}$ is given

by $\frac{\partial \mathbf{w}(\mathbf{m})}{\partial \mathbf{m}}$ and can be obtained from

$$\begin{aligned} & \mathbf{S}(\mathbf{m})\mathbf{w}(\mathbf{m}) = \mathbf{F} \\ \Rightarrow & \frac{\partial \mathbf{S}(\mathbf{m})}{\partial \mathbf{m}}\mathbf{w}(\mathbf{m}) + \mathbf{S}(\mathbf{m})\frac{\partial \mathbf{w}(\mathbf{m})}{\partial \mathbf{m}} = 0 \\ \Rightarrow & \frac{\partial \mathbf{w}(\mathbf{m})}{\partial \mathbf{m}} = -\mathbf{S}^{-1}(\mathbf{m})\frac{\partial \mathbf{S}(\mathbf{m})}{\partial \mathbf{m}}\mathbf{w}(\mathbf{m}). \end{aligned} \tag{E.4}$$

Ignoring the amplitude normalization factor in equation 3.4, the connective function for WQ can also be written as

$$\dot{f}_{\Delta f} = \langle \mathbf{S}_r \tilde{\mathbf{w}}_f(\mathbf{x}, t; \mathbf{x}_s), \dot{\mathbf{w}}_{f-\Delta f}(\mathbf{x}_r, t; \mathbf{x}_s) \rangle \,. \tag{E.5}$$

Here, $\mathbf{S}_r$ is a restriction operator onto the receiver position and it depends on the spatial coordinates. The restriction operator samples the wavefield at the receiver locations while the adjoint of it, $\mathbf{S}_r{}^*$, sprays the recorded data from the receiver coordinates to the model coordinates. $\tilde{\mathbf{w}}_f$ denotes the predicted event for a given background Q model recorded at the receiver location $\mathbf{x}_r$ due to a source excited at time $t = 0$ and at location $\mathbf{x}_s$. $\mathbf{w}_{f-\Delta f}$ denotes the same event in the observed data.

From equation 3.5, the misfit function for WQ is given by

$$J = \frac{1}{2}\sum_s \sum_r \Delta f(\mathbf{x}_r, \mathbf{x}_s)^2. \tag{E.6}$$

The gradient of J is given by

$$\begin{aligned}
\frac{\partial J}{\partial \tau(\mathbf{x})} &= -\sum_s \sum_r \frac{\partial \Delta f(\mathbf{x}_r, \mathbf{x}_s)}{\partial \tau(\mathbf{x})} \Delta f(\mathbf{x}_r, \mathbf{x}_s), \\
&= \frac{1}{E} \sum_s \sum_r \frac{\partial \dot{f}_{\Delta f}}{\partial \tau(\mathbf{x})} \Delta f(\mathbf{x}_r, \mathbf{x}_s), \quad \text{(Using equation 3.7)} \\
&= \frac{1}{E} \sum_s \sum_r \frac{\partial}{\partial \tau(\mathbf{x})} \left\langle \mathbf{S}_r \tilde{\mathbf{w}}_f(\mathbf{x}, t; \mathbf{x}_s), \dot{\mathbf{w}}_{f-\Delta f}(\mathbf{x}_r, t; \mathbf{x}_s) \right\rangle \Delta f(\mathbf{x}_r, \mathbf{x}_s), \quad \text{(Using equation E.5)} \\
&= \frac{1}{E} \sum_s \sum_r \left\langle \mathbf{S}_r \frac{\partial \tilde{\mathbf{w}}_f(\mathbf{x}, t; \mathbf{x}_s)}{\partial \tau(\mathbf{x})}, \dot{\mathbf{w}}_{f-\Delta f}(\mathbf{x}_r, t; \mathbf{x}_s) \Delta f(\mathbf{x}_r, \mathbf{x}_s) \right\rangle, \\
&= -\frac{1}{E} \sum_s \sum_r \left\langle \mathbf{S}_r \mathbf{S}^{-1} \frac{\partial \mathbf{S}}{\partial \tau(\mathbf{x})} \tilde{\mathbf{w}}_f(\mathbf{x}, t; \mathbf{x}_s), \dot{\mathbf{w}}_{f-\Delta f}(\mathbf{x}_r, t; \mathbf{x}_s) \Delta f(\mathbf{x}_r, \mathbf{x}_s) \right\rangle, \text{(Using equation E.4)} \\
&= -\frac{1}{E} \sum_s \left\langle \frac{\partial \mathbf{S}}{\partial \tau(\mathbf{x})} \tilde{\mathbf{w}}_f(\mathbf{x}, t; \mathbf{x}_s), \left(\mathbf{S}^{-1}\right)^* \sum_r \left({\mathbf{S}_r}^* \dot{\mathbf{w}}_{f-\Delta f}(\mathbf{x}_r, t; \mathbf{x}_s) \Delta f(\mathbf{x}_r, \mathbf{x}_s) \right) \right\rangle, \\
&= -\frac{1}{E} \sum_s \left\langle \frac{\partial \mathbf{S}}{\partial \tau(\mathbf{x})} \tilde{\mathbf{w}}_f(\mathbf{x}, t; \mathbf{x}_s), \mathbf{w}^*(\mathbf{x}, t; \mathbf{x}_s) \right\rangle. \qquad \text{(E.7)}
\end{aligned}$$

Here, $\mathbf{w}^* = \begin{bmatrix} q & \mathbf{u} & s \end{bmatrix}^T$ is the adjoint-state variable of $\mathbf{w} = \begin{bmatrix} P & \mathbf{v} & r_p \end{bmatrix}^T$ and is numerically obtained by a finite-difference solution of the adjoint-state equation in D.12,

$$\begin{aligned}
&\frac{\partial q}{\partial t} + \nabla \cdot \left(\frac{1}{\rho} \mathbf{u} \right) = -\sum_r {\mathbf{S}_r}^* \left(P(\mathbf{x}_r, t; \mathbf{x}_s)_{obs} \Delta f(\mathbf{x}_r, \mathbf{x}_s) \right), \\
&\frac{\partial \mathbf{u}}{\partial t} + \left[\nabla K(1+\tau) q + \nabla \left(\frac{1}{\tau_\sigma} K \tau s \right) \right] = -\sum_r {\mathbf{S}_r}^* \left(\mathbf{v}(\mathbf{x}_r, t; \mathbf{x}_s)_{obs} \Delta f(\mathbf{x}_r, \mathbf{x}_s) \right), \\
&\frac{\partial s}{\partial t} - \frac{s}{\tau_\sigma} - q = -\sum_r {\mathbf{S}_r}^* \left(r_p(\mathbf{x}_r, t; \mathbf{x}_s)_{obs} \Delta f(\mathbf{x}_r, \mathbf{x}_s) \right). \qquad \text{(E.8)}
\end{aligned}$$

From equation E.3, we have

$$\mathbf{S} = \begin{bmatrix} \frac{\partial}{\partial t} & K(1+\tau)\nabla\cdot & 1 \\ \frac{1}{\rho}\nabla & \frac{\partial}{\partial t} & 0 \\ 0 & \frac{\tau}{\tau_\sigma}K\nabla\cdot & \frac{\partial}{\partial t}+\frac{1}{\tau_\sigma} \end{bmatrix},$$

$$\frac{\partial \mathbf{S}}{\partial \tau} = \begin{bmatrix} 0 & K\nabla\cdot & 0 \\ 0 & 0 & 0 \\ 0 & \frac{K\nabla\cdot}{\tau_\sigma} & 0 \end{bmatrix}. \tag{E.9}$$

Using equations E.3 and E.9 we can get,

$$\begin{aligned} \frac{\partial J}{\partial \tau(\mathbf{x})} &= -\frac{1}{E}\sum_s \left\langle \frac{\partial \mathbf{S}}{\partial \tau(\mathbf{x})}\tilde{\mathbf{w}}_f(\mathbf{x},t;\mathbf{x}_s), \mathbf{w}^*(\mathbf{x},t;\mathbf{x}_s)\right\rangle, \\ &= -\frac{1}{E}\sum_s \left\langle \begin{bmatrix} 0 & K\nabla\cdot & 0 \\ 0 & 0 & 0 \\ 0 & \frac{K\nabla\cdot}{\tau_\sigma} & 0 \end{bmatrix} \begin{bmatrix} P \\ \mathbf{v} \\ r_p \end{bmatrix}, \begin{bmatrix} q \\ \mathbf{u} \\ s \end{bmatrix} \right\rangle, \\ &= -\frac{1}{E}\int_s dt \;\; K(\mathbf{x})(\nabla\cdot\mathbf{v}(\mathbf{x},t;\mathbf{x}_s))\left(q(\mathbf{x},t;\mathbf{x}_s)+\frac{s(\mathbf{x},t;\mathbf{x}_s)}{\tau_\sigma(\mathbf{x})}\right). \end{aligned} \tag{E.10}$$

F 3D Fourier based discrete Radon transform

For a discrete image, I, of size $n \times n \times n$, the Radon transform, R, for the x-plane is given by ([Averbuch and Shkolnisky, 2003])

$$R(I, \{x = s_1 y + s_2 z + t\}) = \sum_{v=-n/2}^{n/2-1} \sum_{w=-n/2}^{n/2-1} \tilde{I}^1(s_1 v + s_2 w + t, v, w), \tag{F.1}$$

where

$$\tilde{I}(x, v, w) = \sum_{u=-n/2}^{n/2-1} I(u, v, w) D_m(x - u), \quad v, w \in \left\{-\frac{n}{2}, \ldots, \frac{n}{2} - 1\right\}. \tag{F.2}$$

Here s_1 and s_2 are the slopes along the x-plane and t is the intercept. D_m is the Dirichlet kernel given by

$$D_m = \frac{\sin(\pi t)}{m \sin(\pi t/m)}, \quad m = 3n + 1, \tag{F.3}$$

where m is the length of the Dirichlet kernel.

Similarly for the y- and z-planes, the Radon transform is given by

$$\begin{aligned} R(I, \{y = s_1 x + s_2 z + t\}) &= \sum_{u=-n/2}^{n/2-1} \sum_{w=-n/2}^{n/2-1} \tilde{I}^2(u, s_1 u + s_2 w + t, w), \\ R(I, \{z = s_1 x + s_2 y + t\}) &= \sum_{u=-n/2}^{n/2-1} \sum_{v=-n/2}^{n/2-1} \tilde{I}^3(u, v, s_1 u + s_2 v + t), \end{aligned} \tag{F.4}$$

where

$$\tilde{I}^2(u,y,w) = \sum_{v=-n/2}^{n/2-1} I(u,v,w) D_m(y-v), \quad u,w \in \left\{-\frac{n}{2},\ldots,\frac{n}{2}-1\right\},$$
$$\tilde{I}^3(u,v,z) = \sum_{w=-n/2}^{n/2-1} I(u,v,w) D_m(z-w), \quad u,v \in \left\{-\frac{n}{2},\ldots,\frac{n}{2}-1\right\}. \tag{F.5}$$

Using the Fourier slice theorem, the slopes (s_1, s_2) of an x-plane can be obtained by

$$\hat{RI}_x(s_1,s_2,k) = \sum_{v,w=-n/2}^{n/2-1} \sum_{u=-n/2}^{n/2-1} I(u,v,w) e^{(-2\pi ik/m)(u-s_1v-s_2w)}, \tag{F.6}$$

where $k = \left\{-\frac{3n}{2},\ldots,\frac{3n}{2}\right\}$. Similarly, the slopes (s_1, s_2) of a y-plane and a z-plane can be obtained by

$$\hat{RI}_y = \sum_{u,w=-n/2}^{n/2-1} \sum_{v=-n/2}^{n/2-1} I(u,v,w) e^{(-2\pi ik/m)(v-s_1u-s_2w)},$$
$$\hat{RI}_z = \sum_{u,v=-n/2}^{n/2-1} \sum_{w=-n/2}^{n/2-1} I(u,v,w) e^{(-2\pi ik/m)(w-s_1u-s_2v)}. \tag{F.7}$$

The adjoint of the forward transform used during the preconditioning is the numerical adjoint of the forward transform steps described above.