Vector interpolation and regularized elastic imaging of multicomponent seismic data

by

Kenneth Aaron Stanton

A thesis submitted in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

in

Geophysics

Department of Physics
University of Alberta

Abstract

Historically seismic data processing has relied on the acoustic approximation to process single component data under the simplifying assumption that the recorded wavefield consists mainly of compressional wave modes. With the advancement of multicomponent seismic technology there is an increased need for purpose-built processing tools to improve the quality of the data while considering its vector-elastic nature. Past research in this field resulted in the extension of several key processing steps from the scalar-acoustic to the vector-elastic case. These steps included elastic redatuming, noise attenuation, deconvolution, full waveform inversion and imaging. This thesis is focused on the extension of two other important processing steps to the vector-elastic case: regularization, which compensates for poor spatial sampling of receivers and sources at the earth's surface, and least squares imaging, which compensates for poor illumination of the earth's subsurface. These two topics approach the processing of multicomponent seismic data from two very different approaches. In the extension of scalar reconstruction to the vector case a mathematical representation of multicomponent data in the Fourier domain is achieved via hypercomplex numbers; more specifically, the quaternions. Following the algebraic rules of quaternions it is possible for Fourier regularization algorithms to be adapted to the vector case. A different approach is necessary to extend acoustic least squares migration to the elastic case. In this case the acoustic wave equation must be substituted for the elastic wave equation within the single scattering approximation. This allows elastic wave equation modelling and migration linear operators to be built, providing the necessary tools to approach elastic imaging as an inverse problem with the goal that the optimal image of the earth subsurface is one that best explains the data. Both vector regularization and elastic least squares imaging are found to be improvements over their scalar-acoustic counterparts. In the case of vector regularization the quality of the result is found to be insensitive to the orientation of the input measurements, while in elastic least squares imaging the minimization of wavefield crosstalk is improved by minimizing the least squares error between observed and predicted data components.

Preface

A version of the work in chapter 3 of this thesis has been published in the journal paper Stanton, A., and M. D. Sacchi, 2013, Vector reconstruction of multicomponent seismic data: Geophysics, 78. Additionally, a version of the work in chapter 4 of this thesis has been published in the journal paper Stanton, A., and M. D. Sacchi, 2017, Elastic least-squares one-way wave equation migration: Geophysics, 82. In these publications, I was responsible for designing the processing algorithms, generating the data examples and writing the manuscripts. Dr. Sacchi was the supervisory author and was involved in concept formulation and manuscript editing.

To Julie

Far better an approximate answer to the right question, which is often vague, than an exact answer to the wrong question, which can always be made precise.

John W. Tukey

Acknowledgements

I cannot thank Professor Mauricio Sacchi enough for taking me on as a student and supervising my studies. I'm often amazed by his remarkable ability to make short work of complicated problems; a clear forward problem, a slice of Occam's razor. I am also grateful for the students of the Signal Analysis and Imaging Group (SAIG), past and present, for sharing their ideas with me and for creating a lively research environment. I would also like to thank the members of my committee for taking the time to consider my work.

Over the course of my studies, I came to rely heavily on many existing tools in scientific computing. My introduction to C programming came from studying Seismic Unix (SU) programs, mainly written by the past and present students/staff of the Centre for Wave Phenomena (CWP) at the Colorado School of Mines. I am grateful to the creators of SU for its usefulness and straightforward design. Another software package that I made great use of was Madagascar; a package spearheaded by Sergey Fomel at the University of Texas at Austin, which uses an open source development framework. All of the elastic finite-difference wave simulations used in this thesis were created using this package. This package opened my eyes to the exciting world of reproducible research. I am also indebted to the creators of the Julia programming language, which I recently began using. This is a remarkable language for scientific computing, with Matlab-like syntax and C-like performance. The majority of the figures in this thesis were created in Julia, and many of the programs (whenever I could get away with it) were written in Julia. We at SAIG have contributed an open source seismic processing package written in the Julia language: Seismic.jl[1]. Thanks to Juan Sabbione for all his efforts improving this package. I am also grateful for the staff at Compute/Calcul Canada for providing the computational resources used in the more compute-intensive portions of my research.

I owe great thanks to Kurt Wikel and Daniel Trad for a joint internship, splitting my time between Petrobank and CGG in 2012; and Warren Ross, Craig Hyslop, Ramesh Neelamani,

---

[1] To install this package simply type Pkg.add(”Seismic”) on the Julia command line

Bill Curry and the management at the ExxonMobil Upstream Research Company for an internship in 2013. I would also like to thank Lee Bell and Richard Verm at Geokinetics for providing a field dataset which was extremely useful to test my ideas. Thanks to my colleagues at Key Seismic Solutions for supporting this research, for hiring me, and for encouraging my research in these last several months. Special thanks to Richard Bale for numerous discussions surrounding elastic wave propagation, his thesis often served as a roadmap to navigate this difficult topic.

I am so grateful to my family for being supportive and encouraging of my academic pursuits for all these years. I hope this thesis clarifies what I have been up to for so long!

Finally, I thank my wife Julie and my daughter Olive for holding my hands through all of the topographical extremes of a Ph.D. Their patience, love, and encouragement carried me through the hills and valleys, and their laughter and cheer have made the journey worthwhile.

Contents

1 Introduction **1**

1.1 Vector Interpolation 6

1.2 Regularized Elastic Imaging 9

1.3 Scope of this thesis 11

1.3.1 Main contributions 11

1.3.2 Organization 12

2 Elastic waves and vector measurements **13**

2.1 The elastic wave equation 13

2.2 Solving for polarizations and phase velocities 15

2.3 Example: a three layer model 22

3 Vector interpolation **26**

3.1 Theory 28

3.1.1 Separability of data components 29

3.1.2 Quaternions 30

3.1.3 Projection Onto Convex Sets (POCS) 35

3.2 Synthetic data example 36

3.3 Real data example 42

3.4 Discussion 53

3.5 Generalization of the approach 55

3.6 Conclusions 60

4 Regularized elastic imaging 61

4.1 Theory 61

4.1.1 Review of elastic wave propagation 61

4.1.2 Elastic shot-profile modeling and migration 66

4.1.3 Implementation 73

4.1.4 Least-squares formulation 73

4.2 A note on imaging conditions 75

4.3 Examples 77

4.3.1 Diffractor example 77

4.3.2 Flat layer example 79

4.3.3 BP 2.5D Model example 86

4.4 Conclusions 93

5 Regularized imaging of OBC data 94

5.1 Receiver profile modeling and migration 94

5.2 Velocity analysis 100

5.3 Application of adjoint operator and least squares migration 101

5.4 Remarks 104

5.5 Conclusions 107

6 Conclusions 108

Bibliography 111

A Elastic composition/ decomposition operators 121

B CG for multicomponent model and data vectors 123

C Angle gathers in wave equation migration 127

C.1 Poynting vector to angle formulation 127

C.2 Subsurface offset to angle formulation 130

D Seismic.jl: seismic data analysis in Julia 136

List of Tables

3.1 Parameters for the three layer model shown in Figure 3.3. Two scenarios are modelled. In the first scenario all three layers are isotropic and have constant velocity with azimuth. In the second scenario, the P and S-wave velocities for the second layer vary with azimuth and are computed using elastic coefficients given by C_{HTI} (specified in GPa) with a symmetry axis pointing to the east. This model was generated following Bale (2002). 29

D.1 Benchmark times relative to C (smaller is better, C performance = 1.0). Adapted from http://julialang.org. Please refer to the original for technical specifics. 136

List of Figures

1.1 A scalar processing workflow for multicomponent data. Shear wave processes are shown in red while compressional wave processes are shown in blue. Vector processes are shown in red with a blue outline. It is important to note that the inputs to this workflow, u_x, u_y, u_z represent vector-valued data, while the outputs, m_H, and m_V, represent models of subsurface reflectivity in the horizontal and vertical directions respectively. 3

1.2 A vector processing workflow for multicomponent data. Red squares with a blue outline indicate vector processes. The text on the right indicate publications corresponding to specific elements of the workflow, including the two topics that are the focus of this thesis. Here, it is important to note that the inputs to this workflow, u_x, u_y, u_z represent vector-valued data, while the output, $\boldsymbol{M}$, represents a multi-parameter model of subsurface consisting of (up to nine) distinct elastic reflection types. 4

2.1 Polarizations in the $k_x - k_y$ plane at $k_z = 0$ for an isotropic medium with v_p=1500m/s, v_s=700m/s, and ρ=1000kg/m^3. 20

2.2 Schematic showing the reorientation from an east-south-downward coordinate frame to a radial-transverse-downward coordinate frame. S and G represent the source and receiver locations respectively. 21

2.3 P and S wave velocities for a model containing three interfaces. The lateral source position is indicated above the model by a red star. 22

2.4 Single shot gather at x=5km generated via elastic finite difference modeling (the direct wave has been muted from both components). Notice the projection of earlier arriving PP waveforms mainly on the Z-component at near (< 2km) offsets, while the later arriving PS waveforms are projected mainly onto the X-component at near offsets. 23

2.5 FK amplitude spectra for X and Z components using only near (< 2km) offsets 24

2.6 FK amplitude spectra for X and Z components using all offsets 25

3.1 Single shot gather at x=5km generated via elastic finite difference modelling after randomly decimating the receiver sampling by 50%. 27

3.2 FK amplitude spectra for the randomly subsampled shot gather shown in figure 3.1. Notice the artefacts introduced by the decimated sampling. 27

3.3 Three layer model used for the generation of elastic finite-difference synthetic data examples. Two scenarios are modelled. In the first scenario all three layers are isotropic, while in the second scenario layer 2 has Horizontal Transversely Isotropic (HTI) symmetry. The parameters for each layer are given in table 3.1. The acquisition geometry consists of 18 receiver lines (blue) oriented about a single shot (red). 37

3.4 Elastic finite-difference synthetic shot gathers for Scenario 1 (azimuth 20°) (isotropy). Before reorientation: (a) east, (b) south, (c) downward, and after reorientation: (d) radial, (e) transverse, (f) downward. Notice that the energy in the horizontal plane can be placed into a single component by rotation. . . 38

3.5 Elastic finite-difference synthetic shot gathers for Scenario 2 (azimuth 20°) (Horizontal Transverse Isotropy). Before reorientation: (a) east, (b) south, (c) downward, and after reorientation: (d) radial, (e) transverse, (f) downward. Notice that the energy in the horizontal plane cannot be placed into a single component by rotation. 38

3.6 FK spectra of the horizontal components shown in Figure 3.5. (a) East component, (b) south component, (c) radial component, (d) transverse component, (e) quaternion FK spectrum for east and south components, (f) quaternion FK spectrum for radial and transverse components. Notice that the quaternion FK spectrum does not depend on orientation ((e) and (f) are identical). 39

3.7 Comparison of component by-component reconstruction to vector reconstruction for a synthetic 3D-2C common shot gather: Common-offset display (560m). The input has 60% missing traces and $SNR = 5$. (a) Component 1 input, (c) component 1 after single component reconstruction, (e) component 1 after vector reconstruction. (b), (d), and (f) show the same for component 2. 40

3.8 Comparison of component by-component reconstruction to vector reconstruction for a synthetic 3D two component common shot gather: Common-azimuth display (20°). The input has 60% missing traces and $SNR = 5$. (a) Component 1 input, (c) component 1 after single component reconstruction, (e) component 1 after vector reconstruction. (b), (d), and (f) show the same for component 2. 41

3.9 Quality for the reconstructions shown in Figures 3.7 & 3.8 when the percentage of missing traces is varied (a), and the QFT axis orientation is randomly oriented. The input data have an SNR of 5. 42

3.10 Shot (red) and receiver (blue) coordinates for a converted wave survey over a heavy oil reservoir. 43

3.11 total fold calculated over the offset interval 0-550m using $\gamma = 1$ and $\gamma = 3$ to calculate the bin location (Common Mid-Point). Notice the relatively even distribution of fold around source and receiver lines when $\gamma = 1$, compared to the clustering of high fold bin locations around receiver lines when $\gamma = 3$. . 44

3.12 Azimuth vs. Offset plotted for all traces in one 26x26 patch of ACP bins centered at ACP (150,150). The offset has been limited to the interval 0-550m. Azimuth is calculated as the angle from source to receiver counter-clockwise from east. The two missing lobes centered at azimuths 38° and 218° correspond to the crossline directions which limit the far offset due to the limited size of the recording patch. The maximum offset for the reconstruction was limited to 550m to mitigate the effect of missing azimuths at far offsets. . 44

3.13 Shot (red) and receiver (blue) coordinates for ACP bin (150,150) before 5D regularization (a), and after regularization (b). Black lines connect shot-receiver pairs. For 5D regularization the data are binned into 5x5m ACP bins, 50m offset bins, and 20° azimuth bins. For the ACP binning a value of $\gamma = 3$ was used. This positions the bin 3/4 the distance from source to receiver, giving an asymmetric pattern of sources and receivers for each ACP location. 46

3.14 5D reconstruction results shown for one ACP gather (150,150) sorted by offset, azimuth. (a) Radial component before reconstruction, (b) radial component after component-by-component reconstruction, (c) radial component after vector reconstruction. (d), (e), and (f) respectively show the same results for the transverse component. The area highlighted in red is shown in Figure 3.16 47

3.15 5D reconstruction results shown for one ACP gather (150,150) sorted by offset, azimuth, shown after shear wave splitting correction. (a) Radial component before reconstruction, (b) radial component after component-by-component reconstruction, (c) radial component after vector reconstruction. (d), (e), and (f) respectively show the same results for the transverse component. The area highlighted in red is shown in Figure 3.17 48

3.16 5D reconstruction results shown for offset 375m of ACP gather (150,150) sorted by azimuth. (a) Radial component before reconstruction, (b) radial component after component-by-component reconstruction, (c) radial component after vector reconstruction. (d), (e), and (f) respectively show the same results for the transverse component. 49

3.17 5D reconstruction results shown for offset 375m of ACP gather (150,150) sorted by azimuth, shown after shear wave splitting correction. (a) Radial component before reconstruction, (b) radial component after component-by-component reconstruction, (c) radial component after vector reconstruction. (d), (e), and (f) respectively show the same results for the transverse component. 50

3.18 5D reconstruction results shown for one stacked crossline (150) of the radial component sorted by inline. Shear wave splitting corrections have been applied. (a) Radial component before reconstruction, (b) radial component after component-by-component reconstruction, (c) radial component after vector reconstruction. 51

3.19 5D reconstruction results shown for one stacked crossline (150) of the transverse component sorted by inline. Shear wave splitting corrections have been applied. The complex envelope for each trace is input to the stacking process to avoid the effect of polarity reversals across azimuth. (a) Transverse component before reconstruction, (b) transverse component after component-by-component reconstruction, (c) transverse component after vector reconstruction. 52

3.20 decimated data (as shown in figure 3.1) interpolated using scalar POCS. The qualities of the reconstructions are 23.07 and 29.73 dB respectively. 56

3.21 FK amplitude spectra for the decimated data interpolated using scalar POCS. 57

3.22 decimated data (as shown in figure 3.1) interpolated using scalar MWNI. The qualities of the reconstructions are 27.49 and 32.72 dB respectively. 57

3.23 FK amplitude spectra for the data interpolated using scalar MWNI. The qualities of the reconstructions are and dB respectively. 58

3.24 decimated data (as shown in figure 3.1) interpolated using vector POCS. The qualities of the reconstructions are 23.76 and 30.81 dB respectively. 58

3.25 FK amplitude spectra for the data interpolated using vector POCS. 59

3.26 decimated data (as shown in figure 3.1) interpolated using vector MWNI. The qualities of the reconstructions are 27.53 and 33.41 dB respectively. . . . 59

3.27 FK amplitude spectra for the data interpolated using vector MWNI. 60

4.1 X and Z input data components for an elastic wave propagating in a homogeneous isotropic medium. The records were generated using elastic finite difference modeling. 65

4.2 Decomposed P and S components after applying the wavefield decomposition operator $\boldsymbol{Q}^{-1}$ to the data components U^x and U^y. 65

4.3 Recomposed X and Y components after applying the wavefield recomposition operator $\boldsymbol{Q}$ to the wavefield components U^p and U^s. 66

4.4 Born scattering of acoustic wavefields. 69

4.5 Migration of acoustic wavefields. 69

4.6 Born scattering of elastic wavefields. 70

4.7 Migration of elastic wavefields. 71

4.8 P wave velocity model used to create finite difference synthetic data. The background P-wave velocity is 2200 m/s, while the seven diffractors have a velocity of 2500 m/s. The S wave velocity model corresponds to this model by a scale factor of $1/\sqrt{3}$. 77

4.9 X and Z components for a synthetic shot gather at X = 2500m generated using elastic finite difference modeling. The data have been randomly decimated by 30% to simulate missing receivers. 78

4.10 PP image obtained by elastic migration (the adjoint operator) for a single shot at X = 2500m. 79

4.11 PS image obtained by elastic migration (the adjoint operator) for a single shot at X = 2500m. 80

4.12 PP image obtained by 20 iterations of ELSM for a single shot at X = 2500m. 80

4.13 PS image obtained by 20 iterations of ELSM for a single shot at X = 2500m. 81

4.14 Predicted X and Z components of the interpolated and wavefield separated PP wavefield for a shot gather at X = 2500m generated by the forward operator following 20 iterations of ELSM. 82

4.15 Predicted X and Z components of the interpolated and wavefield separated PS wavefield for a shot gather at X = 2500m generated by the forward operator following 20 iterations of ELSM. 83

4.16 Normalized misfit versus iteration number for 20 iterations of ELSM. 84

4.17 Velocity models used to create finite difference synthetic data. 84

4.18 X and Z components for a synthetic shot gather at X = 1250m generated using elastic finite difference modeling with random noise added. 85

4.19 Images obtained by elastic migration (the adjoint operator) for a single shot at X = 1250m 85

4.20 Images obtained by by 10 iterations of ELSM for a single shot at X = 1250m 86

4.21 Normalized misfit versus iteration number for 10 iterations of ELSM. The solid black line shows convergence when added random noise is present on the input data. The gray curves display trials of the algorithm using different combinations of velocity error. The solid gray line shows convergence using the true velocity, while the dashed lines show convergence when using different combinations of incorrect velocities. For a single shot experiment the convergence is relatively insensitive to small velocity errors. 87

4.22 A portion of the BP 2.5D P-wave velocity model. The S-wave velocity corresponds to this model by a scale factor of $1/\sqrt{3}$. 88

4.23 X and Z components for a synthetic shot gather at x = 2880m generated using elastic finite difference modeling. 89

4.24 PP image and an angle gather from x=4600m obtained by elastic migration (the adjoint operator). Arrows indicate regions of the model that have been poorly imaged. 90

4.25 PS image and an angle gather from x=4600m obtained by elastic migration (the adjoint operator). Arrows indicate regions of the model that have been poorly imaged. 90

4.26 PP image and an angle gather from x=4600m obtained by elastic migration (the adjoint operator). A smoothing filter has been applied across angles to attenuate dipping cross talk energy. Notice that this filter has little effect on the overall stack. Arrows indicate regions of the model that have been poorly imaged. 91

4.27 PS image and an angle gather from x=4600m obtained by elastic migration (the adjoint operator). A smoothing filter has been applied across angles to attenuate dipping cross talk energy. Notice that this filter has little effect on the overall stack. Arrows indicate regions of the model that have been poorly imaged. 91

4.28 PP image and an angle gather from x=4600m obtained by 10 iterations of ELSM. Arrows indicate regions of the model where the image has been improved. 92

4.29 PS image and an angle gather from x=4600m obtained by 10 iterations of ELSM. Arrows indicate regions of the model where the image has been improved. 92

4.30 Normalized misfit versus iteration number for 10 iterations of ELSM. 93

5.1 coordinates of sources (blue stars) and receivers (red triangles) for an ocean bottom dataset. Note sparse distribution of receivers compared to the dense distribution of shots. 95

5.2 X and Z components for a common source gather at X=6000m. The sampling in this domain is very sparse with a 500m receiver spacing. 95

5.3 X and Z component FK Amplitude spectra for a common source gather at X=6000m. Clearly, the data are highly aliased in this axis. 96

5.4 X and Z components for a common receiver gather at X=5521m. 97

5.5 X and Z component FK Amplitude spectra for a common receiver gather at X=5521m. 97

5.6 Schematic for P-wave source in a water layer converting to an S-wave reflection which is recorded by an ocean bottom receiver. 99

5.7 Correct use of the Maxwell-Betti reciprocal work theorem. An S-wave source embedded in the ocean bottom converting to a P-wave reflection which is recorded by a receiver in the water layer. 99

5.8 Incorrect use of the Maxwell-Betti reciprocal work theorem (the source and receiver impulses/responses were not interchanged). A P-wave source embedded in the ocean bottom converting to an S-wave reflection which is recorded by a receiver in the water layer. Clearly, the response will not be equivalent to that in figure 5.6. This differences will be kinematic (the ray-path is different), and dynamic (the S-wave cannot propagate in the water column). 100

5.9 Estimated source wavelet used for imaging. 101

5.10 PP angle gathers for X=4000m after elastic migration with -10% velocity perturbation (left), 0% velocity perturbation (middle), and +10% velocity perturbation (right). 102

5.11 PS angle gathers for X=4000m after elastic migration with -10% velocity perturbation (left), 0% velocity perturbation (middle), and +10% velocity perturbation (right). 102

5.12 P-wave velocity model. 103

5.13 S-wave velocity model. 103

5.14 PP image after elastic receiver gather migration. 104

5.15 PS image after elastic receiver gather migration. 105

5.16 PP image after 5 iterations of preconditioned elastic least squares receiver gather migration. 105

5.17 PS image after 5 iterations of preconditioned elastic least squares receiver gather migration. 106

5.18 Normalized cost as a function of iteration number. 106

C.1 Source and receiver wavefield Poynting vector conventions used when compute opening angle, θ, with resect to reflector normal, $\vec{\boldsymbol{n}}$. 129

C.2 Propagation angles for a source at X=5000m, computed using the Poynting vector method. 130

C.3 Near angle stack from the Poynting vector method using source side incidence angles between -10.0 and 10.0 degrees from vertical. 131

C.4 Angle gather at $X = 5000m$ generated for a single shot via the subsurface offset method. While the resolution of the angles and depths is slightly lower than for the equivalent Poynting vector angle gather (figure), the artefacts are fewer. 131

C.5 Image for a subsurface offsets of -250m (top), 0m (middle), and +250m (bottom). 134

C.6 Near angle stack from the subsurface offset method using half-opening angles from γ = -10.0 to 10.0 degrees. 135

C.7 Angle gather at $X = 5000m$ generated for a single shot via the subsurface offset method. While the resolution of the angles and depths is slightly lower than for the equivalent Poynting vector angle gather (figure), the artefacts are of lower amplitude. 135

D.1 decimated seismic data used as input to the SeisPOCS program. The data have 50% missing traces. 138

D.2 Seismic data interpolated using the SeisPOCS program. 138

D.3 A velocity model used to demonstrate the program SeisPWD. Computed normal vectors are plotted as purple arrows. 139

D.4 Shot and receiver coordinates extracted from the Teapot Dome dataset. . . . 140

D.5 Fold map computed from the Teapot Dome dataset. 141

CHAPTER 1

Introduction

A central problem in geophysics the estimation of the interior structure of the earth by attempting to fit observations with data synthesized from a physical theory. Mathematical approximations play an important role in solving these problems. They make solutions to intractable problems possible and often provide a reasonable starting point to begin making sense of observed data. For many years reflection seismology relied heavily on the acoustic approximation– the assumption that the earth behaves as a liquid supporting only stretching and squeezing wave motion. The acoustic approximation greatly simplifies seismic imaging, but ignores the fact that the earth is more accurately described as an elastic solid that supports both shear and compressional wave modes[1]. Over the past several decades the focus of exploration seismology has transitioned from structural mapping to the inversion of elastic properties for identifying rock types and the fluids that reside in their pore space. This is especially true in the analysis of resource plays where ideal drilling targets depend less on the structure of a rock than its permeability and ability to fracture. We now demand more from our data, including improved structural imaging, rock type, petrophysical properties, information about faults, fractures, stress regimes, fluid content, as well as rock and fluid changes with production (Stewart, 2009). Attempts to meet these demands have spurred the development of multicomponent seismic acquisition and processing technology (Stewart et al., 2003). A particular development that has increased the popularity of converted wave acquisition is the advent of Microelectromechanical Systems (MEMS) based receivers. The benefits of these receivers include direct digital output, improved vector fidelity, broadband linear phase and amplitude response, low harmonic distortion, measurement of sensor tilt, and reduced power consumption (Kendall, 2006). Importantly, these receivers triggered the movement away from receiver groups (used to attenuate surface waves via stacking), to the

---

[1] Or even more accurately as an anisotropic, poroelastic solid, supporting quasi-compressional and shear modes; or more accurately still, as an effective medium.

use of point receivers. Indeed, Hoffe et al. (2002) demonstrate that geophone groups are unnecessary for PP data[2] acquisition and are in fact detrimental to PS data[3] quality.

Industrial multicomponent seismic processing workflows have been slow to adapt to the times. Workflows typically consist of scalar algorithms applied to individual components of data. Figure 1.1 outlines a basic multicomponent processing flow aimed at taking multicomponent data recorded in the field to produce images that can assist in the exploration and production of hydrocarbons. This is by no means an exhaustive list and individual processes may be revisited several times over the course of a processing sequence. Nevertheless, scalar seismic processing flows typically include static corrections, aimed at compensating for time delays between adjacent measurements caused by rapid variations in the near surface (also called the weathering layer); noise attenuation, to remove energy that appears spatially uncorrelated (or correlated but undesirable, such as surface waves or multiply scattered reflections[4]); deconvolution, to compensate for the individual coupling and responses of sources and receivers; regularization, to compensate for the irregular spatial sampling of sources and receivers; and velocity analysis and imaging to map reflected energy from source-receiver-time coordinates to form an image of the earth's reflective layers.

Multicomponent processing flows typically begin by rotating survey oriented components into shot specific components of radial, transverse, and vertical as proxies for compressional and converted wavefields (Garotta, 1984). This is a vector process because it treats all data components simultaneously, although the assumption that the resulting components represent distinct wavefields assumes the incident wave arrives with a vertical angle of incidence; a requirement that is often violated in practice. Following rotation to a shot oriented frame of reference, the converted and compressional wave components are processed individually using scalar processing algorithms. An exception to this is shear wave splitting analysis and compensation, aimed at compensating for the effects of Horizontal Transverse Isotropy (HTI), which treats horizontal components as a vector quantity by rotating, shifting, and finally merging shear wave energy onto a common component before continuing with a scalar processing flow (Crampin, 1981; Silver and Chan, 1991). This is also a true vector process, but like the rotation of data components, it assumes vertical transmission through an anisotropic layer (Bale, 2006). Because of the simplicity of the assumptions used in a traditional multicomponent processing flow, it is highly likely that unwanted shear wave energy is will remain in the final vertical component image, as well as unwanted compressional wave energy in the horizontal component image. This cross-talk energy is improperly

---

[2] the abbreviation "PP" is commonly used to indicate compressional-wave to compressional-wave mode reflections.

[3] the abbreviation "PS" is commonly used to indicate compressional-wave to shear-wave mode reflections.

[4] Noise is often characterized as "whatever is undesirable." Consider single component processing where shear waves are treated as unwanted noise, compared with multicomponent processing, where every effort is made to preserve and enhance this energy.

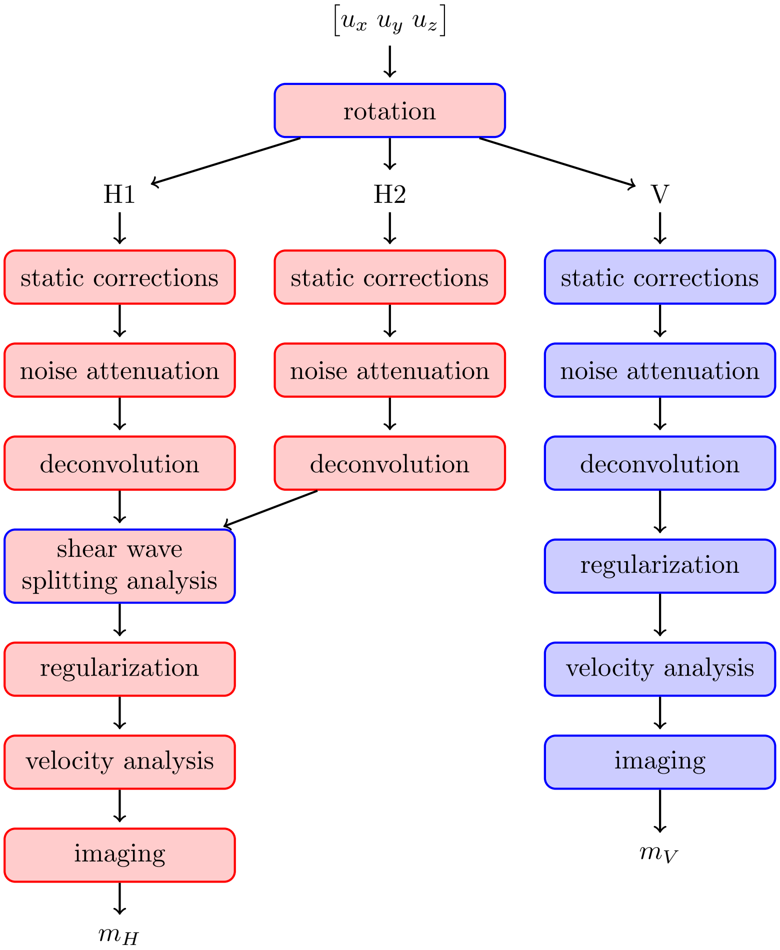


Figure 1.1: A scalar processing workflow for multicomponent data. Shear wave processes are shown in red while compressional wave processes are shown in blue. Vector processes are shown in red with a blue outline. It is important to note that the inputs to this workflow, u_x, u_y, u_z represent vector-valued data, while the outputs, m_H, and m_V, represent models of subsurface reflectivity in the horizontal and vertical directions respectively.

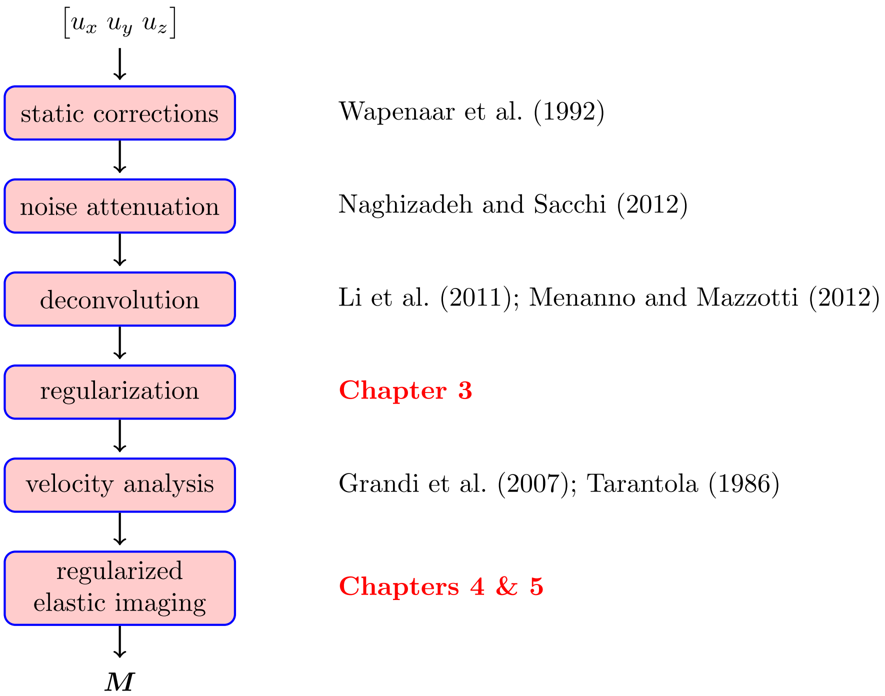


Figure 1.2: A vector processing workflow for multicomponent data. Red squares with a blue outline indicate vector processes. The text on the right indicate publications corresponding to specific elements of the workflow, including the two topics that are the focus of this thesis. Here, it is important to note that the inputs to this workflow, u_x, u_y, u_z represent vector-valued data, while the output, $\boldsymbol{M}$, represents a multi-parameter model of subsurface consisting of (up to nine) distinct elastic reflection types.

positioned and can interfere with the interpretation of correctly positioned reflections.

An alternative to an acoustic-scalar processing flow is a vector-elastic flow as shown in figure 1.2, with publications concerning individual steps listed on the right. In this flow, data samples are mathematically treated as a vector, and the mixture of wavefield potentials in data components is embraced as a consequence of the elastic wave equation. No doubt this adds some complexity to individual processing steps, but as this thesis will demonstrate, this increase in complexity can improve the accuracy of the final result.

To date, there has been limited research into individual items in the vector-elastic processing workflow. Wapenaar et al. (1992) provide a method that can compensate for the effects of the near surface via elastic redatuming[5]. In this approach, individual wavefield potentials are

---

[5] while conventional approaches to static corrections correct for spurious arrival times using ”static” time delays of individual traces, elastic redatuming must account for the distinct time delays of individual elastic wave modes which are mixed between data components.

extrapolated into the earth below the weathering layer before subsequent processing. While unseparated wavefield potentials can be extrapolated, the authors suggest performing a separation prior to this step. This could also present a problem as wavefield separation itself requires knowledge of near surface properties (chapter 2 will outline the physics of wavefield separation using the Christoffel equation). There are also mathematical approaches to wavefield separation using Independent Component Analysis (ICA) (van Der Baan, 2006), or via plane wave decomposition (Cary, 1998). These methods are attractive because they require less knowledge of near surface material properties. The second step in the workflow is vector noise attenuation. Recently Naghizadeh and Sacchi (2012) investigated the extension of FX autoregression noise attenuation to the vector case. They show an improvement over component-by-component noise attenuation using their approach. There has also been work on vector deconvolution of elastic data. Li et al. (2011) extended single component deconvolution to the two horizontal components of multicomponent seismic data. They find their method is insensitive to the orientation of input measurements, and that the resulting data are better conditioned for the vector process of shear wave splitting analysis and compensation. In a later publication Menanno and Mazzotti (2012) extended Wiener deconvolution to the multicomponent case by means of the quaternions. They found their approach to have superior performance to single component deconvolution, particularly when the input components were contaminated with different levels of noise (chapter 3 will make use of the quaternions in the extension of Fourier regularization to the multicomponent case). There have been many studies investigating the estimation of seismic velocities with multicomponent seismic data. Grandi et al. (2007) extend the attribute of semblance, commonly used in single component Root Mean Squared (RMS) velocity analysis, also making use of the quaternions. In a groundbreaking contribution Tarantola (1986) outlines the nonlinear elastic inversion of seismic data, more commonly referred to as Full Waveform Inversion (FWI). FWI could be considered as a means to obtain velocity information for a subsequent imaging step, or as an imaging algorithm in its own right. There are two notable gaps in the vector-elastic processing flow shown in figure 1.2, namely regularization and regularized imaging, which are the focus of this thesis.

It should be noted that while figure 1.2 leaves vector regularization and regularized imaging as items outstanding, there have been related studies in these fields to date. In the case of regularization Vassallo et al. (2010) made use of wavefield gradient measurements to better constrain the interpolation of marine streamer data in the sparsely sampled crossline direction via matching pursuit. This method processes multiple data types, as opposed to data components, in its formulation. It should also be noted that there has been a great deal of work done on elastic imaging. Sun and McMechan (1986) and Chang and McMechan (1987) performed elastic imaging where the target images represented correlations of wavefield displacements, while Sun and McMechan (2001) separated data components to wavefield potentials at a datum using Helmholtz decomposition, subsequently propagating and imaging the resulting scalar wavefield potentials. A more accurate approach was used by Yan and Sava (2008a). In this work elastic finite differences were used to propagate wavefield displacements in depth, leaving the separation to scalar potentials to immediate precede the imaging condition. In this thesis I draw a distinction between elastic imaging, which approximates the inverse of Born scattering by its adjoint; with regularized imaging, which better approximates the inverse of Born scattering via regularized least squares inversion. Approaching elastic imaging as an inverse problem offers the ability to improve the bandwidth of migrated images, compensate for the poor spatial sampling of sources and receivers at the earth's surface, as well as for poor illumination of the subsurface, and finally to minimize elastic wavefield crosstalk artefacts within individual images. There is considerable motivation to have a complete vector-elastic workflow. The extensive research into elastic imaging and FWI are not yet widely used in industry and will require extensive preconditioning of datasets to make them suitable for these processes. deviating from the vector-elastic workflow for an early processing step could destroy the vector nature of the data, thus preventing a return to the workflow for a later step.

1.1 Vector Interpolation

The goal of early seismic data processing steps is to prepare the data for prestack migration, a process which maps reflected energy from source-receiver acquisition coordinates and time to reflector positions in space. Migration consists of numerical approximations to integral equations. These approximations perform best when data are free of noise and have regular, densely sampled source-receiver acquisition coordinates. Regularization algorithms aim to provide a fully sampled noise free estimate of the data and often exploit the multidimensional nature of seismic data. They rely on a variety of signal processing strategies with the underlying assumption that there is sufficient simplicity in the observed wavefield for it to be represented using a finite number of basis functions. There are various strategies to tackle the reconstruction problem. Algorithms broadly fall under the following two categories: those based on wave equation principles (for example Ronen (1987)), and those based on signal processing principles. Methods based on signal processing are by far the most commonly used for interpolation and denoising of seismic data because they do not require velocity information to reconstruct the data. Such methods include prediction error filter techniques (Spitz, 1991; Naghizadeh and Sacchi, 2007a), transform based methods (Abma and Kabir, 2006; Liu and Sacchi, 2004), and rank reduction based methods (Trickett et al., 2010; Kreimer and Sacchi, 2011).

All methods for seismic data reconstruction assume a high level of parsimony in the wavefield that one desires to reconstruct. In other words, the observations must be represented via a simple regression in terms of known basis functions. The latter is the foundation of methods based on Fourier synthesis where one assumes that the observed seismic data can be represented in terms of a superposition of complex exponentials (Hindriks and Duijndam, 1999). The representation itself is not sufficient to properly honor the data and produce an accurate synthesis of unobserved data. This is why the decomposition must be in terms of a limited number of complex exponentials. Methods based on an assumption of sparsity have been successfully adopted to retrieve Fourier coefficients (Sacchi et al., 1998; Zwartjes and Gisolf, 2007) that permit the proper reconstruction of wavefields within multi-dimensional windows. Minimum weighted norm interpolation (MWNI), applies a less stringent constraint to the data and leads to solutions that are quite similar to those that one can estimate via sparse inversion (Liu and Sacchi, 2004; Trad, 2009; Cary, 2011b; Chiu, 2014)). However, an advantage of MWNI is that because the concept of sparsity is not invoked, it resorts to a simple algorithm based on quadratic regularization with pre-conditioning. In fact, MWNI should be the optimal regularization technique if one were to know the power spectral density of the data to be reconstructed. However, because the power spectral density is unknown, MWNI uses different strategies to bootstrap it from the data. The latter leads to an algorithm that is extremely similar to sparse inversion solutions via Iteratively Re-Weighted Least-squares (IRLS). Sparse Fourier inversion and MWNI, therefore, can be considered as similar ways to fit data via the superposition of a finite number of complex exponentials. They both can be generalized to multidimensional problems, and clearly, they both have problems with regular grids since they are affected by aliasing. However, patches of seismic prestack data often have enough irregularities to attenuate aliasing making MWNI a robust method for interpolating data on irregular grids (Trad, 2009).

Methods based on algorithms utilized in radio astronomy (the Clean Method of Schwarz (1978)) have also been proposed to regularize seismic data. Examples of the latter are the Antileakage Fourier Transform (ALFT) (Xu et al., 2005; Schonewille et al., 2009)) and Matching Pursuit Fourier reconstruction (MPFR) (Øzbek et al., 2009). These techniques find and retrieve one wavenumber at a time to synthesize a spatial plane wave. The latter is removed from the original data and the process is continued until no significant energy is observed in the residual signal.

In essence, the assumption of simplicity (sparsity) is also invoked by ALFT and MPFR methods as they also try to construct a synthesis by the superposition of a limited number of wave-numbers. Both methods have the ability to cope with data at irregular grid positions. When the data are on a regular grid, dip scanning methods can be applied to identify true wavenumbers from their associated alias components (Naghizadeh, 2012). At this point, it is important to mention that both Sparse Fourier inversion methods and MWNI can operate both with regular grids and irregular spatial positions. In the first case, the algorithms can be implemented via the FFT engine, whereas in the latter, it is necessary to utilize fast implementations of non-uniform DFTs (Jin, 2010). Finally, we would like to mention POCS, another Fourier method that synthesizes data in terms of, again, a parsimonious distribution of spectral coefficients (Abma and Kabir, 2006; Wang et al., 2010b; Stein et al., 2010; Stanton et al., 2013, 2015). POCS uses a Fourier domain amplitude threshold to represent the data also in terms of a simple distribution of spectral amplitudes. In general, POCS, ALFT, MPFR, and MWNI should produce results that are similar because, after, all they are developed under similar assumptions: simplicity (sparsity) of the data representation in terms of spatial plane waves (Stanton et al., 2012). Most differences between these techniques are probably attributed to implementation and developers ways to cope with geometries, input/output, and data preconditioning. A range of tricks are used to stabilize and bring these algorithms to the realm of industrial applications. For instance aliasing and its solutions for MWNI and POCS have discussed by Cary and Perz (2012) and Gao et al. (2013).

A superposition of complex exponentials (the assumption made by Fourier reconstruction methods) leads to the linear prediction model. In other words, the superposition of complex exponentials immersed in noise can be represented via autoregressive (AR) models. These models are the basis of linear prediction theory, where one states that observations at one channel are a linear combination of observations of adjacent channels. This is also the principle behind FX noise attenuation via prediction filters. The superposition of complex sinusoids in FX (linear events in TX) leads to predictability via AR models in X for a given frequency F. This model is astutely exploited for denoising and for reconstruction (Spitz, 1991; Gülünay, 2003; Naghizadeh and Sacchi, 2007b). The advantage of FX prediction filter methods is that they can cope with aliasing. In addition, they can be generalized to N-D cases (Naghizadeh and Sacchi, 2010). Again, the assumption is that the data can be represented via a superposition of plane waves that admits representation in terms of an autoregressive model. For this to be true, one needs to segment the data in small spatiotemporal windows to guarantee the validity of the plane wave assumption. This is also a problem we encountered when reconstructing data with Fourier reconstruction methods.

In general, all reconstruction methods are based on the assumption of simplicity of the representation of the waveform. For instance, a superposition of linear events has a sparse representation in the frequency-wavenumber domain, similarly, a superposition of linear events is predictable in the frequency-space domain. In this regard, one can claim the linear event model is behind all reconstruction methods for multidimensional signals. In practice, a linear superposition of events can be safely assumed by windowing the data or by using localized transforms (Hennenfent and Herrmann, 2006, 2008; Wang et al., 2010a).

A difficulty in applying these methods to multicomponent data is that they are inherently nonlinear, that is, the coefficients necessary to predict missing data are derived from the data themselves. As such, applying scalar interpolation algorithms to components independently is highly likely to destroy the vector relationship between these components. A solution of this problem is to make the interpolation itself a vector process, by adapting the linear operators within Fourier interpolation methods (the FK transform and its adjoint) to the vector case. This thesis extends two popular Fourier regularization techniques, POCS and MWNI, to the multicomponent case by utilizing the quaternions to represent vector data in the Fourier domain.

1.2 Regularized Elastic Imaging

While seismic data interpolation compensates for irregular surface sampling, least squares imaging compensates for irregular subsurface illumination of the data. Least-squares migration seeks a reflectivity model that fits the observed data. It is used to compensate for acquisition noise, the poor sampling of sources and receivers on the surface, as well as poor illumination of the subsurface. To date, least-squares migration has been restricted to the imaging of acoustic wavefields. This thesis presents an extension of wave equation least-squares migration for elastic wavefields in isotropic media. For P-wave seismology, it is common to simulate wave propagation using the scalar wave equation. This approximation is often sufficient to produce reasonable P-wave structural images. Extending this approximation to converted wave imaging disregards the polarization of wave modes within data components. Migration algorithms based on the elastic wave equation more accurately model wave propagation and allow multiple wavefields to be imaged simultaneously. There are a variety of approaches to elastic migration. Kuo and Dai (1984) developed an elastic implementation of Kirchhoff migration. Chang and McMechan (1987) applied Reverse Time Migration (RTM) to multicomponent data using an elastic finite-difference algorithm, resulting in horizontal and vertical component images. Dellinger and Etgen (1990) proposed the application of Helmholtz decomposition to elastic data using a Fourier domain operator. Later implementations of elastic RTM use this idea to provide distinct PP and PS images (see for example Yan and Sava (2008b)). Recently, Cheng et al. (2016) use a low-rank approximation to simulate decoupled elastic waves in anisotropic media, and Guo and Alkhalifah (2016) perform reflection waveform inversion using elastic data. Extensive work has also been done on one-way wave equation based methods to migrate elastic data. Xie and Wu (2005) used an elastic version of split step migration to downward continue elastic data, while Bale (2006) modified phase shift migration to handle anisotropic elastic wavefields. One-way wave equation based methods are an attractive option because they perform more accurately than ray-based methods, and are much less computationally expensive than methods that employ finite differences. For example, in time domain RTM many wavefield snapshots for all time samples must be generated and saved prior to imaging, while in one-way wave equation migration frequency slices are treated independently and the wavefield is recursively updated with depth. One-way wave equation migration operators are also computationally efficient in both memory and operation count, perform accurately in moderately complex geological settings, and result in relatively few numerical artefacts. There are also many promising methods to improve on traditional one-way wave equation migration that can make its results competitive with RTM (for example Shan and Biondi (2008)).

While in acoustic migration the image represents the partitioning of energy between incident and reflected waves that share a common mode, in elastic migration a multi-parameter image is created that represents the partitioning of energy between all combinations of incident and reflected wave modes. The relative strength of these images as a function of angle of incidence are related to the physical properties of the interface by the Zoeppritz equations (Aki and Richards, 2002). Because P waves propagate by compression and dilation, their corresponding reflection strengths are greatly influenced by pore fluids. S-wave reflection strengths, on the other hand, are less influenced by the presence of pore fluids and instead depend more heavily on changes in the rock matrix. As an example, MacLeod et al. (1999) delineate a sandstone reservoir surrounded in shale by comparing PP and PS reflection strengths. A counteraction of impedance associated with a change in lithology and an oil-water contact made the reservoir difficult to identify in PP images, while the PS image was able to isolate the change in lithology necessary to identify the reservoir.

The migration of converted waves comes with some additional complications compared to P-wave migration. Typically PS data have a lower signal to noise ratio than PP data, and the lower propagation velocity of the receiver side wavefield moves the point of reflection away from the midpoint and toward the receiver. While lower propagation velocity implies more restrictive aliasing criteria, it also presents an opportunity: given the same acquisition geometry, converted waves can image the subsurface with a wider source side aperture than P-wave data. Furthermore, ignoring the effects of attenuation, a given frequency of converted wave data images the earth with a higher resolution than P-wave data (Stewart et al., 2002). Realistically attenuation is always a factor (especially in the near surface), which challenges the resolution of both PP and PS depth images (Bale and Stewart, 2002). Some of these challenges can be dealt with before migration. For example, regularization can be applied to reduce the effects of acquisition footprint on the migrated image (Cary, 2011a). In this thesis, I aim to address some of the challenges facing converted wave migration via least-squares inversion.

Least-squares migration, in its various forms, has been an active field of research for many years. Lambaré et al. (1992) used an iterative data fitting approach to solve for reflectivity. Nemeth et al. (1999) applied least-squares migration with a Kirchhoff operator to image in the presence of poor spatial sampling, while Kühl and Sacchi (2003) used regularized least-squares one-way wave equation migration to generate angle gathers that are unaffected by acquisition footprint. Wang et al. (2005) and Wang (2005) explore the application of 3D one-way wave equation least-squares migration with constraints to a real data set of the Western Canadian Sedimentary Basin. In a subsequent contribution, Wang and Sacchi (2007) also explore the incorporation of a sparsity constraint to least-squares migration to the ubiquitous problem of vertical resolution enhancement. Least-squares migration of structurally complex data has been also being investigated via the utilization of preconditioning operators synthesized using prediction error filters (Wang and Sacchi, 2009). More recently, there has been a great deal of research incorporating two-way wave equation migration operators into least-squares migration (Dai and Schuster, 2013; Zhang et al., 2015; Dutta and Schuster, 2014a; Ji, 2009; Wong et al., 2015). A common theme of all least-squares migration algorithms is data fitting which depends greatly on the ability of the migration operator to accurately propagate energy from source to receiver. Extending least-squares migration from the acoustic to the elastic case is a natural progression in data fitting.

Least-squares migration is an iterative method that requires a forward and an adjoint operator. In elastic least-squares wave equation migration, the forward operator generates data components from multiparameter images by recursive wavefield decomposition, extrapolation, and recomposition. Conversely, the adjoint operator generates multiparameter images from data components by recursively applying the adjoint of the wavefield recomposition, extrapolation, and wavefield decomposition operators. An extended imaging condition is used, defining the image as a function of reflection angle, and the inversion is regularized by applying a smoothing filter on the depth-angle axes of each common image point gather to reduce the effect of source/receiver sampling, noise, and crosstalk artefacts. Numerical examples show that elastic least-squares migration can be used for interpolation, wavefield separation, and illumination compensation of multicomponent seismic data.

1.3 Scope of this thesis

1.3.1 Main contributions

This thesis adds two new processes to the growing list of vector-elastic processing tools, namely vector interpolation and elastic least squares migration. The main contributions of this thesis are to

1. provide a new method to represent multicomponent seismic data in the Fourier domain using the quaternion Fourier transform, where all components share a common amplitude spectrum and three phase terms
2. make use of the quaternion Fourier transform to extend Fourier interpolation to the multicomponent case to improve the interpolation of multicomponent seismic data
3. provide a new method to model elastic data via shot profile one-way elastic wave equation modeling using an extended model domain
4. demonstrate that the adjoint of elastic modeling is elastic migration, and use these two operators to improve the imaging of elastic data via regularized least squares inversion

1.3.2 Organization

This thesis is organized as follows. Chapter two reviews the physics of elastic wave propagation, including a discussion on the relationship between wavefield potentials and particle motion in three components as governed by the Christoffel equation, and the projection of particle motion onto a three-component receiver. Chapter three outlines Fourier regularization, a class of methods used to improve the spatial sampling of scalar datasets, describes a vector representation of multicomponent seismic data in the Fourier domain via the quaternions, and extends Fourier regularization to the multicomponent case. Synthetic examples demonstrate that vector regularization consistently improves the quality of the regularization independent of the orientation of the input components. A field data example demonstrates that the approach better prepares the data for shear wave splitting analysis and compensation. Chapter four develops extended shot profile one-way elastic wave equation modeling. Taking the adjoint of this linear operator, a cost function for elastic least squares imaging is defined and minimized via conjugate gradients. A variety of synthetic data examples demonstrate that elastic least squares migration minimizes crosstalk artefacts, mitigates the effects of poor spatial sampling, and compensate for non-uniform illumination of the subsurface.

CHAPTER 2

Elastic waves and vector measurements

Elastic wave propagation introduces additional complexity into recorded seismic data, namely the orientation of particle motion in the subsurface as a function of the direction of propagation for individual wave modes, the material parameters in which the waves are propagating, and the subsequent projection of these waves onto an array of vector receivers at the earth's surface. Considering these complexities, multicomponent recordings must be treated as vector quantities (as detailed in chapter 3), or as a superposition of elastic wave modes (as detailed in chapters 4 and 5). This chapter will outline the physics of elastic wave propagation, the separation of elastic wavefields via the Christoffel equation, and the projection of wave modes onto vector receivers at the earth's surface.

2.1 The elastic wave equation

A force acting on a small region of a solid will result in three mutually orthogonal wave modes known as elastic waves. To understand why this is the case we begin with Newton's second law of motion relating force and acceleration

$$F_i = m\ddot{u}_i, \tag{2.1}$$

where

$$m = \rho dx_1 dx_2 dx_3, \tag{2.2}$$

and

$$F_i = \partial_j \tau_{ij} dx_1 dx_2 dx_3 + f_i. \tag{2.3}$$

where τ_{ij} indicates the force acting in the direction $\hat{\imath}$ on the face of an infinitesimal cube with normal vector $\hat{\jmath}$, leading to the equation of motion:

$$\rho \ddot{u}_i = \partial_j \tau_{ij} + f_i. \tag{2.4}$$

Next, we relate stress and strain using Hooke's law

$$\tau_{ij} = C_{ijkl} e_{kl}, \tag{2.5}$$

where C_{ijkl} is the 81 term tensor of elastic moduli [1], where the strain is given by

$$e_{kl} = \frac{1}{2}(\partial_k u_l + \partial_l u_k). \tag{2.6}$$

Combining equations 2.4, 2.5, and 2.6 while ignoring external forces f_i we obtain the wave equation

$$\rho \ddot{u}_i = C_{ijkl} \partial_j \partial_l u_k \tag{2.7}$$

This equation already makes the assumption that elastic parameters vary gradually at the scale of elastic wave propagation, suggesting the measured data would be insensitive to this variability.

Displacement can be represented as a complex harmonic (plane wave), $u_k = U_k e^{i\omega(t - s_j x_j)}$, where s_j is the slowness vector that is the inverse of the phase velocity. The phase velocity is related to the slowness via $s_j = n_j/V$, where n_j indicates the direction of wave propagation. The direction of wave propagation is just the normalized wavenumber vector, $n_j = k_j/k$, where $k = \sqrt{k_x + k_y + k_z}$. Substituting a plane wave into the wave equation we obtain the Christoffel equation

$$\rho U_i = C_{ijkl} s_j s_l U_k, \tag{2.8}$$

More commonly written in the form

---

[1] Actually, symmetries of the elastic moduli tensor allows for it to be written using Voigt notation as a symmetric 6x6 matrix following the rules $C_{ijkl} = C_{IJ}$, where $11 \rightarrow 1$, $22 \rightarrow 2$, $33 \rightarrow 3$, $23 \rightarrow 4$, $13 \rightarrow 5$, and $12 \rightarrow 6$. Further symmetries of the material parameters can reduce the number of free parameters from 21 (triclinic) to as little as 2 (isotropy).

$$\begin{bmatrix} G_{11}-\rho V^2 & G_{12} & G_{13} \\ G_{12} & G_{22}-\rho V^2 & G_{23} \\ G_{13} & G_{23} & G_{33}-\rho V^2 \end{bmatrix} \begin{bmatrix} U_1 \\ U_2 \\ U_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \tag{2.9}$$

where the 3x3 Christoffel matrix is defined by

$$G_{ij} = C_{ijkl} n_j n_i \tag{2.10}$$

2.2 Solving for polarizations and phase velocities

The Christoffel equation is a 3 x 3 eigenvalue - eigenvector problem. Given a tensor of elastic moduli the eigenvalues correspond to the phase velocities of three distinct wave modes, and the eigenvectors correspond to the polarizations of these modes. In the context of elastic migration, the eigenvectors provide the polarization information necessary for wavefield separation, and the eigenvalues provide the phase velocity necessary for extrapolation.

The first step in solving the Christoffel equation is to determine the elements of the Christoffel matrix given a tensor of elastic moduli of some assumed symmetry. The elements of the Christoffel matrix (equation 4.4) for arbitrary symmetry expand to

$$\begin{aligned} G_{11} &= C_{11}n_1n_1 + C_{16}n_1n_2 + C_{15}n_1n_3 \\ &+ C_{61}n_2n_1 + C_{66}n_2n_2 + C_{65}n_2n_3 \\ &+ C_{51}n_3n_1 + C_{56}n_3n_2 + C_{55}n_3n_3 \end{aligned} \tag{2.11}$$

$$\begin{aligned} G_{22} &= C_{66}n_1n_1 + C_{62}n_1n_2 + C_{64}n_1n_3 \\ &+ C_{26}n_2n_1 + C_{22}n_2n_2 + C_{24}n_2n_3 \\ &+ C_{46}n_3n_1 + C_{42}n_3n_2 + C_{44}n_3n_3 \end{aligned} \tag{2.12}$$

$$\begin{aligned} G_{33} &= C_{55}n_1n_1 + C_{54}n_1n_2 + C_{53}n_1n_3 \\ &+ C_{45}n_2n_1 + C_{44}n_2n_2 + C_{43}n_2n_3 \\ &+ C_{35}n_3n_1 + C_{34}n_3n_2 + C_{33}n_3n_3 \end{aligned} \tag{2.13}$$

$$\begin{aligned} G_{12} &= C_{16}n_1n_1 + C_{12}n_1n_2 + C_{14}n_1n_3 \\ &+ C_{66}n_2n_1 + C_{62}n_2n_2 + C_{54}n_2n_3 \\ &+ C_{56}n_3n_1 + C_{52}n_3n_2 + C_{54}n_3n_3 \end{aligned} \tag{2.14}$$

$$\begin{aligned} G_{13} &= C_{15}n_1n_1 + C_{14}n_1n_2 + C_{13}n_1n_3 \\ &+ C_{65}n_2n_1 + C_{64}n_2n_2 + C_{63}n_2n_3 \\ &+ C_{55}n_3n_1 + C_{54}n_3n_2 + C_{53}n_3n_3 \end{aligned} \tag{2.15}$$

$$\begin{aligned} G_{23} &= C_{65}n_1n_1 + C_{64}n_1n_2 + C_{63}n_1n_3 \\ &+ C_{25}n_2n_1 + C_{24}n_2n_2 + C_{23}n_2n_3 \\ &+ C_{45}n_3n_1 + C_{44}n_3n_2 + C_{43}n_3n_3 \end{aligned} \tag{2.16}$$

Applying the "cookie-cutter" of transverse isotropy to the matrix of elastic moduli, many terms disappear leading to

$$G_{11} = C_{11}n_1n_1 + C_{66}n_2n_2 + C_{55}n_3n_3 \tag{2.17}$$

$$G_{22} = C_{66}n_1n_1 + C_{22}n_2n_2 + C_{44}n_3n_3 \tag{2.18}$$

$$G_{33} = C_{55}n_1n_1 + C_{44}n_2n_2 + C_{33}n_3n_3 \tag{2.19}$$

$$G_{12} = C_{12}n_1n_2 + C_{66}n_2n_1 \tag{2.20}$$

$$G_{13} = C_{13}n_1n_3 + C_{55}n_3n_1 \tag{2.21}$$

$$G_{23} = C_{23}n_2n_3 + C_{44}n_3n_2 \tag{2.22}$$

Simplifying further still by assuming isotropy:

$$C_{IJ} = \begin{bmatrix} \lambda+2\mu & \lambda & \lambda & 0 & 0 & 0 \\ \lambda & \lambda+2\mu & \lambda & 0 & 0 & 0 \\ \lambda & \lambda & \lambda+2\mu & 0 & 0 & 0 \\ 0 & 0 & 0 & \mu & 0 & 0 \\ 0 & 0 & 0 & 0 & \mu & 0 \\ 0 & 0 & 0 & 0 & 0 & \mu \end{bmatrix}, \tag{2.23}$$

and substituting into equation 4.3 considering propagation in the $[x_1, x_3]$ plane we obtain the following Christoffel equation

$$\begin{bmatrix} (\lambda+2\mu)n_1^2 + \mu n_3^2 - \rho V^2 & 0 & (\lambda+\mu)n_1 n_3 \\ 0 & \mu(n_1^2 + n_3^2) - \rho V^2 & 0 \\ (\lambda+\mu)n_1 n_3 & 0 & \mu n_1^2 + (\lambda+2\mu)n_3^2 - \rho V^2 \end{bmatrix} \begin{bmatrix} U_1 \\ U_2 \\ U_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \tag{2.24}$$

To find the eigenvalues corresponding to the three wave modes we set the determinant equal to zero

$$\begin{vmatrix} (\lambda+2\mu)n_1^2 + \mu n_3^2 - \rho V^2 & 0 & (\lambda+\mu)n_1 n_3 \\ 0 & \mu(n_1^2 + n_3^2) - \rho V^2 & 0 \\ (\lambda+\mu)n_1 n_3 & 0 & \mu n_1^2 + (\lambda+2\mu)n_3^2 - \rho V^2 \end{vmatrix} = 0 \tag{2.25}$$

Substituting $\boldsymbol{n} = \begin{bmatrix} 0 & 0 & 1 \end{bmatrix}$ (i.e. assuming a plane wave traveling vertically downward), we obtain the cubic equation

$$0 = (\mu - \rho V^2)(\mu - \rho V^2)(\lambda + 2\mu - \rho V^2) \tag{2.26}$$

with solutions $V_1 = \sqrt{(\lambda+2\mu)/\rho}$, $V_2 = \sqrt{\mu/\rho}$, and $V_3 = \sqrt{\mu/\rho}$ which are the familiar relations for P, SV, and SH phase velocities respectively. Substituting the largest eigenvalue into the Christoffel equation and assuming isotropy we obtain

$$\begin{aligned}
&\begin{bmatrix} G_{11}-(\lambda+2\mu) & G_{12} & G_{13} \\ G_{12} & G_{22}-(\lambda+2\mu) & G_{23} \\ G_{13} & G_{23} & G_{33}-(\lambda+2\mu) \end{bmatrix} \begin{bmatrix} U_1^P \\ U_2^P \\ U_3^P \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \\
\Rightarrow &\begin{bmatrix} \hat{n_1}^2-1 & \hat{n_1}\hat{n_2} & \hat{n_1}\hat{n_3} \\ \hat{n_1}\hat{n_2} & \hat{n_2}^2-1 & \hat{n_2}\hat{n_3} \\ \hat{n_1}\hat{n_3} & \hat{n_2}\hat{n_3} & \hat{n_3}^2-1 \end{bmatrix} \begin{bmatrix} U_1^P \\ U_2^P \\ U_3^P \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \qquad (2.27) \\
\Rightarrow &\begin{bmatrix} \hat{n_3} & 0 & -\hat{n_1} \\ 0 & \hat{n_3} & -\hat{n_2} \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} U_1^P \\ U_2^P \\ U_3^P \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}
\end{aligned}$$

Which has the solution

$$\begin{bmatrix} U_1^P \\ U_2^P \\ U_3^P \end{bmatrix} = \begin{bmatrix} \hat{n_1} \\ \hat{n_2} \\ \hat{n_3} \end{bmatrix}. \qquad (2.28)$$

Repeating the process for the second largest eigenvalue we obtain

$$\begin{aligned}
&\begin{bmatrix} G_{11}-\mu & G_{12} & G_{13} \\ G_{12} & G_{22}-\mu & G_{23} \\ G_{13} & G_{23} & G_{33}-\mu \end{bmatrix} \begin{bmatrix} U_1^{S1} \\ U_2^{S1} \\ U_3^{S1} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \\
\Rightarrow &\begin{bmatrix} \hat{n_1}^2 & \hat{n_1}\hat{n_2} & \hat{n_1}\hat{n_3} \\ \hat{n_1}\hat{n_2} & \hat{n_2}^2 & \hat{n_2}\hat{n_3} \\ \hat{n_1}\hat{n_3} & \hat{n_2}\hat{n_3} & \hat{n_3}^2 \end{bmatrix} \begin{bmatrix} U_1^{S1} \\ U_2^{S1} \\ U_3^{S1} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \qquad (2.29) \\
\Rightarrow &\begin{bmatrix} \hat{n_1} & \hat{n_2} & \hat{n_3} \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} U_1^{S1} \\ U_2^{S1} \\ U_3^{S1} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}
\end{aligned}$$

which is further constrained by $\hat{n}_1^2+\hat{n}_2^2+\hat{n}_3^2=1$ giving

$$\begin{bmatrix} \hat{n_1} & \hat{n_2} & \sqrt{1-\hat{n_1}^2-\hat{n_2}^2} \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} U_1^{S1} \\ U_2^{S1} \\ U_3^{S1} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}. \qquad (2.30)$$

Letting $a^2 = \hat{n_1}^2 + \hat{n_2}^2$ we get the solution

$$\begin{bmatrix} U_1^{S1} \\ U_2^{S1} \\ U_3^{S1} \end{bmatrix} = \begin{bmatrix} -\hat{n}_2/a \\ \hat{n}_1/a \\ 0 \end{bmatrix}. \tag{2.31}$$

Because the eigenvectors of a 3x3 symmetric matrix form a orthonormal basis in $\Re^3$ we can use the relation $\boldsymbol{U^{S2}} = \boldsymbol{U^P} \times \boldsymbol{U^{S1}}$ to get

$$\begin{bmatrix} U_1^{S2} \\ U_2^{S2} \\ U_3^{S2} \end{bmatrix} = \begin{bmatrix} \hat{n}_1\hat{n}_3/a \\ \hat{n}_2\hat{n}_3/a \\ -a \end{bmatrix}. \tag{2.32}$$

For more general symmetries of the stiffness tensor the eigensolutions are found numerically. Figure 2.1 demonstrates a numerical solution to the Christoffel equation for an isotropic medium with v_p=1500m/s, v_s=700m/s, and ρ=1000kg/m^3.

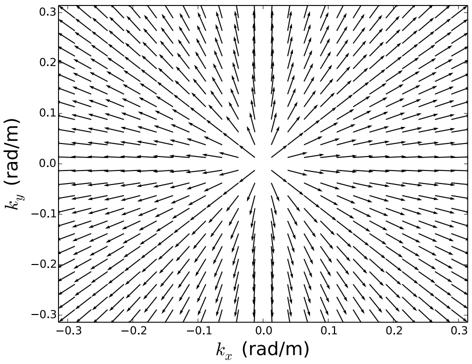
The importance of these relations become apparent in chapter 4 where they are used for the extrapolation of elastic wavefields.

From the above discussion, it is apparent that elastic waves have an amplitude and polarization that depends on the velocity of the wave and the elastic properties of the medium. To measure these waves we place many 3-component [2] sensors at the surface of the earth and the incident wavefields are projected onto the X, Y, and Z components of each receiver.

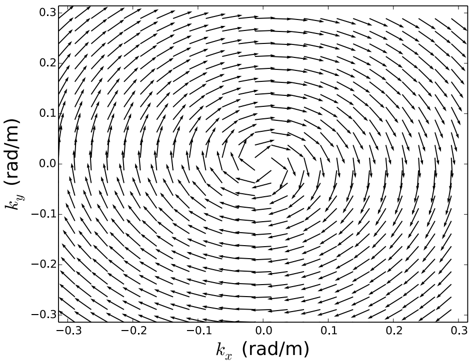
Because vector measurements are sensitive to the orientation of the receivers the orientation of multicomponent geophones within a survey are ideally kept constant across the survey area, or alternatively, the orientation of the geophone must be measured precisely. Three angles are required to describe the orientation of an axis in three dimensions: its roll (γ, rotation about the X-axis), pitch (β, rotation about the Y-axis), and yaw (α, rotation about the Z-axis). From these angles, a multicomponent recording can be re-oriented to a common reference frame such as an east, south, and downward. The choice of reference frame is arbitrary, although a right-handed system is desirable if applying rotation matrices. To reorient a multicomponent sample by counter-clockwise angles (γ, β, α) three consecutive matrix-vector multiplications are applied:

---

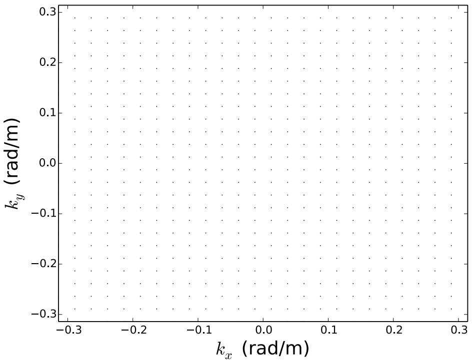
[2] Four component measurements are also common in ocean bottom surveying, where the fourth "component" is pressure. Other measurements are also possible including curls and spatial gradients



(a) Phase 1



(b) Phase 2



(c) Phase 3

Figure 2.1: Polarizations in the $k_x - k_y$ plane at $k_z = 0$ for an isotropic medium with v_p=1500m/s, v_s=700m/s, and ρ=1000kg/m^3.

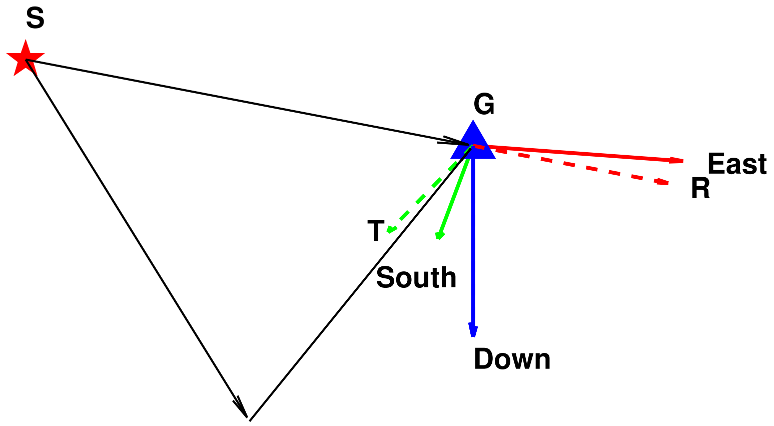


Figure 2.2: Schematic showing the reorientation from an east-south-downward coordinate frame to a radial-transverse-downward coordinate frame. S and G represent the source and receiver locations respectively.

$$\begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} = R_x(\gamma)R_y(\beta)R_z(\alpha)\begin{pmatrix} x \\ y \\ z \end{pmatrix} \tag{2.33}$$

$$= \begin{pmatrix} 1 & 0 & 0 \\ 0 & cos\gamma & -sin\gamma \\ 0 & sin\gamma & cos\gamma \end{pmatrix}\begin{pmatrix} cos\beta & 0 & sin\beta \\ 0 & 1 & 0 \\ -sin\beta & 0 & cos\beta \end{pmatrix}\begin{pmatrix} cos\alpha & -sin\alpha & 0 \\ sin\alpha & cos\alpha & 0 \\ 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix} \tag{2.34}$$

To reverse a rotation the inverse (≡ transpose) of the rotation matrix is used. Typically a multicomponent processing sequence will involve rotating from a survey-wide orientation to a shot-specific orientation of radial (pointing from source to receiver), transverse (normal to the radial component in the horizontal plane), and downward (otherwise called the Vertical component). The conventions for orientation of multicomponent seismic data are described in Brown et al. (2002). Figure 2.2 shows a schematic of this reorientation for a source-receiver pair.

In converted wave recording a P-wave source is used to generate both P-P and P-S mode-converted wavefields which are then recorded using 3C (geophone) or 4C (geophone + hydrophone) receivers. For deep reflections, nearly vertical arrival angles are observed at the receiver. As a result, P-wave or quasi-P-wave oscillation (in the case of anisotropy) is mainly recorded by the vertical component, while S-wave or quasi-S-wave oscillation is mainly recorded by the horizontal components.

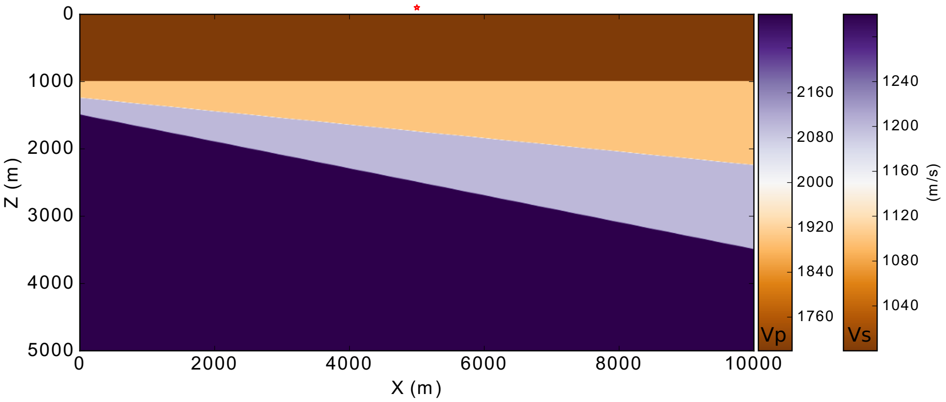


Figure 2.3: P and S wave velocities for a model containing three interfaces. The lateral source position is indicated above the model by a red star.

2.3 Example: a three layer model

To demonstrate the projection of a wavefield onto data components consider the velocity model pictured in figure 2.3. Elastic finite difference modeling for a single source with this model results in the data shown in figure 2.4 (the direct wave has been muted from both components). Because these reflections were generated relatively deep in the model ($\geq$ 1km) and the shallow velocity is slow, near offset reflections arrive at the receiver approximately vertically, resulting in mainly PP reflections on the vertical component and mainly PS reflections on the horizontal component. Nevertheless, we see a non-negligible amount of PP energy in the horizontal component and a complete violation of this assumption at far offsets.

Another way to visualize the projection of wavefields onto data components is via FK amplitude spectra. Figure 2.5 shows FK amplitude spectra for the near offsets ($<$ 2km) for the shot gather shown in figure 2.4. For this range of offsets very little PS energy is visible in the vertical amplitude spectrum, making it approximately valid to process limited offset data via scalar acoustic methods. For the horizontal component, a noticeable amount of PP reflections are visible, even for such a limited range of offsets, making the assumption that only PS reflections are recorded by the horizontal component a less valid approximation than the PP reflection for the vertical component.

Calculating the FK amplitude spectrum including all offsets (figure 2.6), we see that PP and PS reflections are clearly mixed between the two components. It is for this reason that

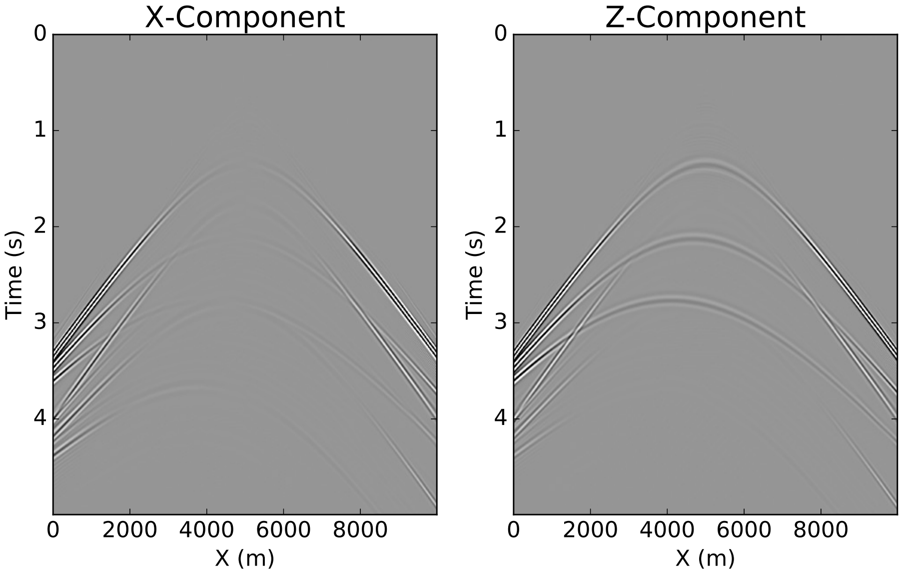


Figure 2.4: Single shot gather at x=5km generated via elastic finite difference modeling (the direct wave has been muted from both components). Notice the projection of earlier arriving PP waveforms mainly on the Z-component at near ($<$ 2km) offsets, while the later arriving PS waveforms are projected mainly onto the X-component at near offsets.

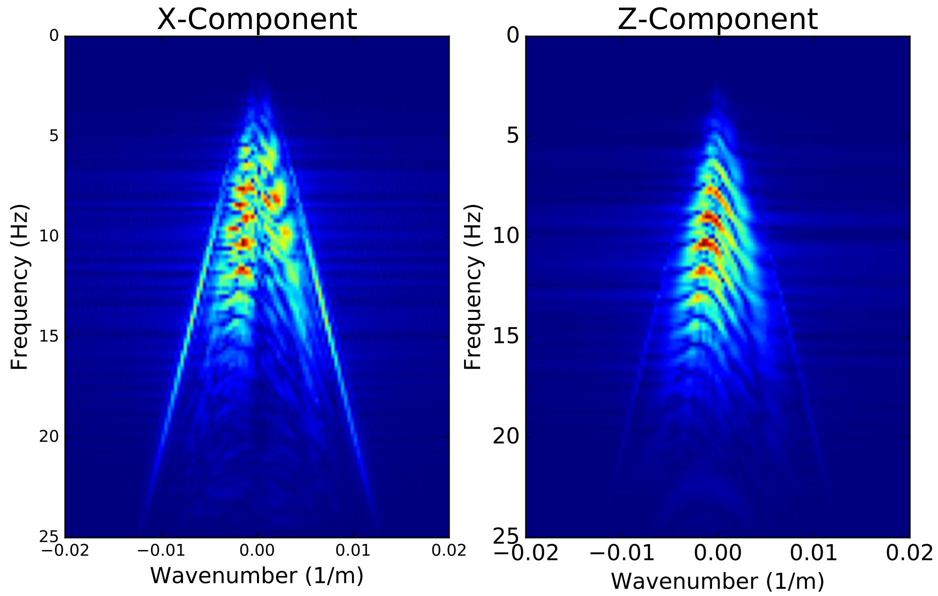


Figure 2.5: FK amplitude spectra for X and Z components using only near (< 2km) offsets

vector processing of multicomponent seismic data is an important consideration for signal preservation and vector fidelity.

The next chapter discusses the effect of subsampling the spatial axes of these data, and how a vector processing strategy is useful in preserving the vector fidelity of multicomponent seismic data.

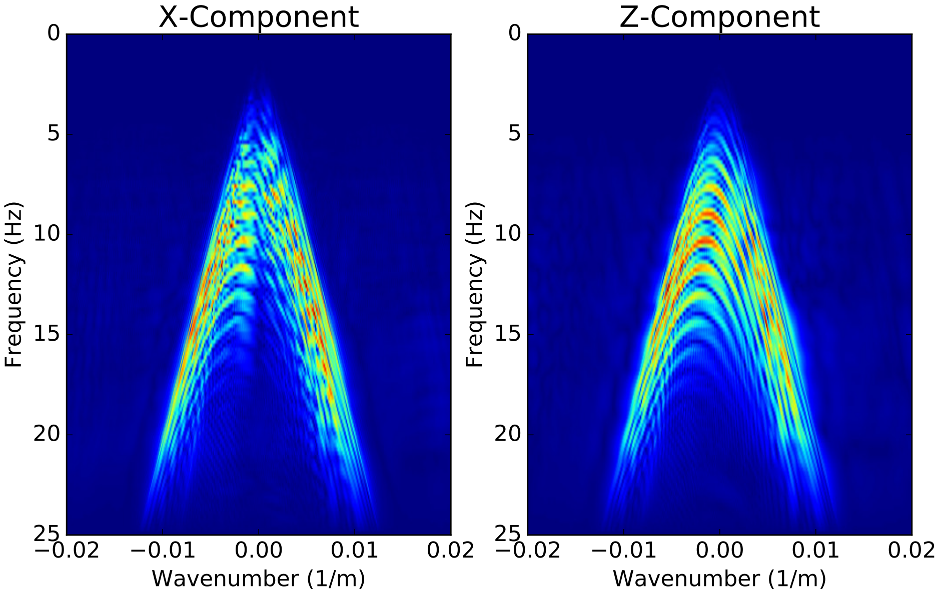


Figure 2.6: FK amplitude spectra for X and Z components using all offsets

CHAPTER 3

Fourier interpolation and its extension to the multicomponent case[1]

The spatial sampling of seismic data is rarely as dense or as regular as is required to generate a reliable image of the subsurface. Seismic data regularization is often applied within a processing workflow to estimate a regular spatial grid of data using either surface (source and receiver), or subsurface (midpoint, offset and azimuth or vector offset) coordinates. A popular family of algorithms to regularize seismic data rely on the Fourier transform. Fourier regularization methods attempt to regularize the data by extracting high amplitude coefficients, or "signal," from the Fourier spectrum while removing Fourier artefacts generated by irregular spatial sampling, or "noise". As an example, consider the multicomponent common shot gather from the previous chapter and its "clean" FK amplitude spectrum. By randomly subsampling the receiver grid, we obtain the data shown in figure 3.1, with FK amplitude spectra shown in figure 3.2.

The decimation of random traces gives the FK spectra shown in figure 3.2 a noisy appearance compared with the spectrum in figure 2.6. Fourier interpolation algorithms exploit this difference via a sparsity constraint on the FK spectrum of the data to predict missing traces.

Reconstruction is typically not applied to the horizontal components of multicomponent seismic data. If the process is done at all, it is typically applied only to the radial component which neglects the vector relationship between the horizontal components. This paper introduces a method to reconstruct two components simultaneously using the quaternion Fourier transform and Projection Onto Convex Sets (POCS). Quaternions have been used

---

[1] Parts of this chapter have been published in Stanton and Sacchi (2013), Stanton and Sacchi (2012), and Stanton and Sacchi (2011).

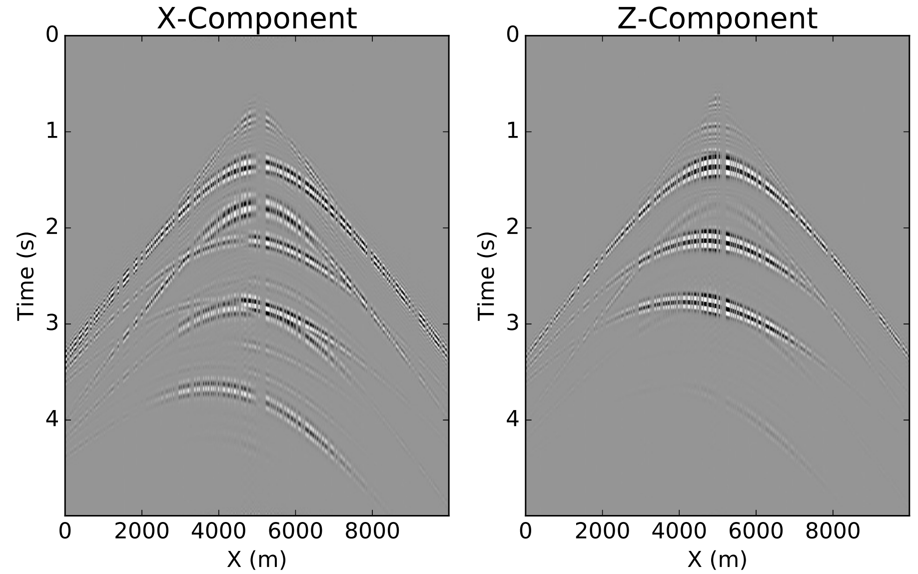


Figure 3.1: Single shot gather at x=5km generated via elastic finite difference modelling after randomly decimating the receiver sampling by 50%.

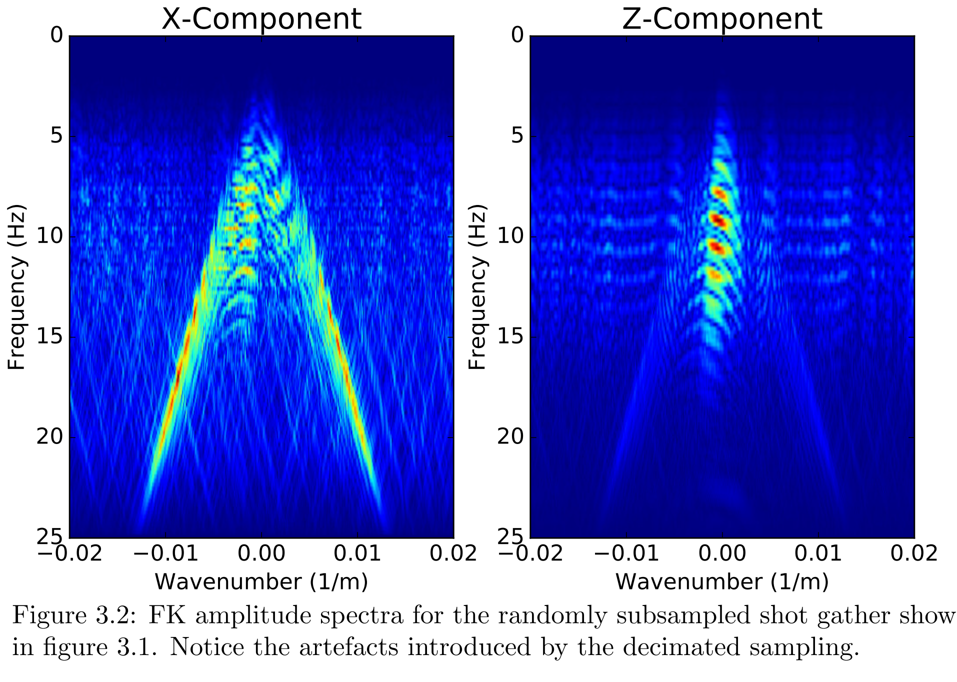


Figure 3.2: FK amplitude spectra for the randomly subsampled shot gather shown in figure 3.1. Notice the artefacts introduced by the decimated sampling.

for other applications in seismic data processing such as the computation of spectral attributes (Bihan and Mars, 2001), time-lapse analysis and boundary detection (Witten and Shragge, 2006), multicomponent velocity analysis (Grandi et al., 2007), and multicomponent deconvolution (Menanno and Mazzotti, 2012). Quaternions were first introduced by Sir William Rowan Hamilton while investigating how to extend the complex numbers into three dimensions. He knew how to add and multiply three-dimensional numbers, but was struggling to find a way to divide them. In 1843 Hamilton discovered that to allow division a fourth dimension is necessary (Hamilton, 1866)[2]. A vector representation of two component data in the frequency domain can be achieved by putting the real and imaginary elements of each component into the arguments of a quaternion. This allows for operations on both components to be carried out simultaneously. The quaternions are transformed to the frequency-wavenumber domain using the quaternion Fourier transform and a single amplitude spectrum for both components is defined using a polar representation of quaternions (Sangwine and Ell, 1999). Reconstruction of missing or noisy traces is carried out using Projection Onto Convex Sets (Abma and Kabir, 2006). The method has the distinct advantages that the orthogonality of the input components is maintained (the signals are not mixed), and the similarities between the components serve to improve the quality of the reconstruction.

3.1 Theory

Multicomponent seismic data suffer from the same spatial sampling and SNR problems as single component data as well as some other complications. Shear waves are slower than compressional wave velocities requiring higher spatial sampling to overcome aliasing. In the case of converted wave acquisition, the conversion point of the ray-path lies nearer to the receiver, giving an irregular pattern of fold.

In the presence of Horizontal Transverse Isotropy (HTI), shear wave splitting occurs (Crampin, 1981). Pre-stack analysis of shear wave splitting has been shown to benefit from a good distribution of azimuths at every bin location (Gaiser, 2000; Bale et al., 2005). For pre-stack analyses superbinning strategies are typically used to borrow missing azimuths from neighboring bin locations (Cary, 2006), but this has the effect of smearing parameter estimates over the area of the superbinning. 5D reconstruction methods can be used to fully populate missing spatial samples in the subsurface geometry (bin$_x$, bin$_y$, offset, azimuth), but reconstruction of the components individually could be damaging to their vector relationship.

If a linear dependence exists between components then attempts should be made to ensure

---

[2] While the discovery of quaternions is typically attributed to Hamilton, unpublished works of Gauss indicate his discovery of the quaternions as early as 1819.

Parameter	ρ (kg/m^3)	V_p (m/s)	V_s (m/s)
Layer 1	2000	2000	1000
Layer 2	2100	2500*	1250*
Layer 3	2500	2800	1500

$$C_{HTI} = \begin{pmatrix} 11.1 & 3.3 & 3.3 & 0 & 0 & 0 \\ 3.3 & 11.7 & 4.3 & 0 & 0 & 0 \\ 3.3 & 4.3 & 11.7 & 0 & 0 & 0 \\ 0 & 0 & 0 & 3.7 & 0 & 0 \\ 0 & 0 & 0 & 0 & 2.9 & 0 \\ 0 & 0 & 0 & 0 & 0 & 2.9 \end{pmatrix}$$

Table 3.1: Parameters for the three layer model shown in Figure 3.3. Two scenarios are modelled. In the first scenario all three layers are isotropic and have constant velocity with azimuth. In the second scenario, the P and S-wave velocities for the second layer vary with azimuth and are computed using elastic coefficients given by C_{HTI} (specified in GPa) with a symmetry axis pointing to the east. This model was generated following Bale (2002).

they are preserved. Traditionally this is done by rotating components to a reference frame in which the components are linearly independent (Li and Yuan, 1999). At this point, each component can be processed separately with no loss of fidelity. Unfortunately, it is not always possible to separate wavefields in this way as the next section demonstrates. In these cases, a vector processing strategy must be used, which is the focus of this paper.

3.1.1 Separability of data components

We consider two scenarios in which the recorded wavefields in converted wave exploration have different degrees of separability with regard to the reference frame.

The two scenarios consider elastic wave propagation over a model consisting of three homogeneous layers as shown in Figure 3.3. In the first scenario, all three layers are isotropic. The densities (ρ), P, and S-wave velocities (V_p, V_s) are provided in Table 3.1.

In the second scenario, the middle layer is replaced with a Horizontal Transversely Isotropic layer, and the P and S-wave velocities are replaced by using elastic coefficients given by the stiffness tensor C_{HTI} in Table 3.1. This model is generated following Bale (2002). A second-order elastic finite-difference simulation (Boyd, 2006) is used to generate one synthetic 3C shot gather for each scenario. For each simulation, the receivers lie along a constant azimuth with 120 offsets (8-meter increments) with a 1ms sample rate. The source is represented by a vertical displacement with a Ricker wavelet of 15Hz peak frequency. A 3D shot gather is generated by repeating 2D simulations for each azimuth in 10-degree increments to produce a 3D shot gather with azimuths ranging from 0 to 350 degrees.

The first scenario we consider is the case of 3C acquisition over an isotropic medium (shown in Figure 3.4). Here we record P-P, P-S, and S-S wavefields (neglecting multiples and higher order mode conversions). A rotation to a radial, transverse, and downward reference frame places the majority of the P-S wavefield in the radial component, zero in the transverse component, and the majority of the P-P wavefield in the downward component. In this case, the transverse component can be neglected as it contains no signal.

In the second scenario, we consider 3C acquisition over an HTI medium (shown in Figure 3.5). Here we again record P-P, P-S, and S-S wavefields, but the HTI anisotropy leads to a secondary P-S wavefield that is orthogonal to the first (Crampin, 1981). Two P-S wavefields exist on the horizontal components: P-S1, and P-S2 (this could be compounded with HTI effects from multiple layers), where P-S1 and P-S2 indicate converted waves polarized in the fast and slow directions respectively. A rotation to a radial, transverse, and downward reference frame places the P-S1 and P-S2 wavefields in the two horizontal components, and the P-P wavefield in the downward component. Neglecting the interaction between the radial and Vertical components, the two horizontal components can be represented by a two-component vector, while the Vertical component can be approximated by a scalar.

Our synthetic example demonstrates that all three components share common wavefields, but the majority of the P-S energy is contained in the horizontal components. This is in agreement with real data examples that we have studied. In this paper, we mainly focus on the mixing of P-S wavefields between the two horizontal components, but the algorithm we present could also be easily extended to 3C or 4C seismic data.

3.1.2 Quaternions

Hamilton clearly recognized the connection between quaternions and the four dimensions of space-time, famously stating "and how the one of time, of space the three, might in the chain of symbols girdled be." A quaternion is defined as

$$q = a + bi + cj + dk \tag{3.1}$$

where, $i^2 = j^2 = k^2 = ijk = -1$, $ij = -ji = k$, $jk = -kj = i$, and $ki = ik = j$. A quaternion is defined as *pure* when $a = 0$.

Two component data in the frequency domain have the form

$$D1(\omega, \vec{x}) = D1_{real}(\omega, \vec{x}) + D1_{imag}(\omega, \vec{x})i \tag{3.2}$$

$$D2(\omega, \vec{x}) = D2_{real}(\omega, \vec{x}) + D2_{imag}(\omega, \vec{x})i \tag{3.3}$$

The quaternions can represent both components as

$$Q(\omega,\vec{x}) = D1_{real}(\omega,\vec{x}) + D1_{imag}(\omega,\vec{x})i + D2_{real}(\omega,\vec{x})j + D2_{imag}(\omega,\vec{x})k. \tag{3.4}$$

It is often useful in the Fourier domain to use a polar representation of data in terms of amplitude and phase. The polar representation of quaternions is given by

$$q = |q|e^{\mu\phi} \tag{3.5}$$

where,

$$|q| = \sqrt{a^2+b^2+c^2+d^2}, \tag{3.6}$$

$$\mu = \frac{bi+cj+dk}{\sqrt{b^2+c^2+d^2}}, \tag{3.7}$$

$$\phi = cos^{-1}\frac{a}{|q|}, \tag{3.8}$$

and μ is a pure unit quaternion ($|\mu| = 1$) which is referred to as the quaternion's eigenaxis. Another representation of equation 3.5 is (Ell and Sangwine, 2007)

$$q = |q|e^{\mu\phi} = |q|(cos\phi + \mu sin\phi). \tag{3.9}$$

This equation is important because it allows for the amplitude of all elements of the quaternion to be expressed as a single term, $|q|$. Care has to be taken when carrying out algebra with quaternions. Multiplication is not commutative, and as a result, multiplication by an inverse function, or division, has distinct left and right sided representations. This also implies that $e^{ib+jc} \neq e^{ib}e^{jc}$ (Sangwine, 1998). The two-dimensional forward and inverse continuous quaternion Fourier transforms given by Ell (1992, 1993) can be adapted to four dimensions giving

$$\begin{aligned} Q(\omega,\nu_1,\nu_2,\nu_3,\nu_4) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} [\quad & e^{-2\pi j\nu_3 x_3}e^{-2\pi j\nu_1 x_1}q(\omega,x_1,x_2,x_3,x_4) \\ & \cdot e^{-2\pi k\nu_2 x_2}e^{-2\pi k\nu_4 x_4}]dx_1dx_2dx_3dx_4 \end{aligned} \tag{3.10}$$

$$q(\omega, x_1, x_2, x_3, x_4) = \frac{1}{4\pi^2} \int\limits_{-\infty}^{\infty} \int\limits_{-\infty}^{\infty} \int\limits_{-\infty}^{\infty} \int\limits_{-\infty}^{\infty} \left[e^{2\pi j \nu_3 x_3} e^{2\pi j \nu_1 x_1} Q(\omega, \nu_1, \nu_2, \nu_3, \nu_4) \cdot e^{2\pi k \nu_2 x_2} e^{2\pi k \nu_4 x_4} \right] d\nu_1 d\nu_2 d\nu_3 d\nu_4 \tag{3.11}$$

where ω is temporal frequency and ν_n, $n \in (1, 4)$, are the spatial wavenumbers. It should be noted that the selection of the imaginary terms i, j, k in the exponentials is not uniquely defined. Also, because multiplication of quaternions is non-commutative, the placement of the exponential terms on the left or right side of the function being transformed is not uniquely defined. In the above notation, we follow the definition of the QFT used by Ell (1992, 1993). It should also be noted that the expansions of equations 3.10 and 3.11 to a summation over each of the four elements of the quaternions can be shown using equation 3.9 and following the multiplication rules of quaternions provided in this section.

The formulation for the forward discrete transform (Sangwine and Ell, 2000) can be extended to four dimensions in a similar manner giving

$$Q(\omega, \nu_1, \nu_2, \nu_3, \nu_4) = \sum_{x_4=0}^{N_4-1} \sum_{x_3=0}^{N_3-1} \sum_{x_2=0}^{N_2-1} \sum_{x_1=0}^{N_1-1} \left[e^{-j2\pi(\frac{x_3\nu_3}{N_3})} e^{-j2\pi(\frac{x_1\nu_1}{N_1})} q(\omega, x_1, x_2, x_3, x_4) \cdot e^{-k2\pi(\frac{x_2\nu_2}{N_2})} e^{-k2\pi(\frac{x_4\nu_4}{N_4})} \right]. \tag{3.12}$$

In this formulation the basis of the QFT are formed by the imaginary axes (i, j, k). Sangwine and Ell (2000) show that the basis of the QFT can be written more generally:

$$Q(\omega, \nu_1, \nu_2, \nu_3, \nu_4) = \sum_{x_4=0}^{N_4-1} \sum_{x_3=0}^{N_3-1} \sum_{x_2=0}^{N_2-1} \sum_{x_1=0}^{N_1-1} e^{-\mu_1 2\pi(\frac{x_1\nu_1}{N_1} + \frac{x_2\nu_2}{N_2} + \frac{x_3\nu_3}{N_3} + \frac{x_4\nu_4}{N_4})} q(\omega, x_1, x_2, x_3, x_4) \tag{3.13}$$

where μ_1 is any unit pure quaternion and defines the axis of the transform. The standard complex Fourier transform is a special case of this transform and occurs when $\mu_1 = i$ and the function being transformed is complex (Ell and Sangwine, 2007).

In the above formulation, μ_1 can also be thought of as a versor (a quaternion of norm one) which can be used to represent a rotation around an arbitrary axis, equivalent to a change of basis in $\Re^3$. Writing the quaternion μ_1 as the first row in the change of basis operator

$$\boldsymbol{\mu} = \begin{pmatrix} \ldots & \mu_1 & \ldots \\ \ldots & \mu_2 & \ldots \\ \ldots & \mu_3 & \ldots \end{pmatrix} = \begin{pmatrix} \mu_{11} & \mu_{12} & \mu_{13} \\ \mu_{21} & \mu_{22} & \mu_{23} \\ \mu_{31} & \mu_{32} & \mu_{33} \end{pmatrix} \tag{3.14}$$

a quaternion $q = a + bi + cj + dk$ can be decomposed into its symplectic form (as a pair of complex numbers)

$$q' = (a' + b'\mu_1) + (c' + d'\mu_1)\mu_2 \tag{3.15}$$

where the first and second terms of q' are called its simplex and perplex parts respectively, and a', b', c', and d' are found using the change of basis

$$\begin{aligned} a' &= a \\ b' &= b\mu_{11} + c\mu_{12} + d\mu_{13} \\ c' &= b\mu_{21} + c\mu_{22} + d\mu_{23} \\ d' &= b\mu_{31} + c\mu_{32} + d\mu_{33} \end{aligned} \tag{3.16}$$

(Ell and Sangwine, 2007). Given a choice of the first axis of the basis, μ_1, the initial choice of μ_2 is arbitrary, but must not be parallel to μ_1. μ_3 can be calculated by taking the cross product of μ_1 and μ_2 and setting its modulus to one. Finally, the cross product of μ_1 and μ_3 is taken to provide an updated value of μ_2 that completes an orthonormal basis. This allows for the quaternion Fourier transform of equation 3.15 to be written as

$$\begin{aligned} Q'(\omega, \vec{\nu}) &= \mathcal{F}[a'(\omega, \vec{x}) + b'(\omega, \vec{x})i] + \mathcal{F}[c'(\omega, \vec{x}) + d'(\omega, \vec{x})i]j \\ &= A'(\omega, \vec{\nu}) + B'(\omega, \vec{\nu})i + C'(\omega, \vec{\nu})j + D'(\omega, \vec{\nu})k \end{aligned} \tag{3.17}$$

where $\mathcal{F}$ is the standard Fourier transform. The change of basis is then reversed to give

$$\begin{aligned} A &= A' \\ B &= B'\mu_{11} + C'\mu_{21} + D'\mu_{31} \\ C &= B'\mu_{12} + C'\mu_{22} + D'\mu_{32} \\ D &= B'\mu_{13} + C'\mu_{23} + D'\mu_{33}. \end{aligned} \tag{3.18}$$

This allows for existing FFT codes to be used in calculating the QFT (Ell and Sangwine, 2007).

The following are the steps to carry out the N-dimensional quaternion Fourier transform:

1. Begin with quaternion $q = a + bi + cj + dk$
2. Apply a change of basis to the quaternion using equation 3.16 to get $q' = a' + b'i + c'j + d'k$.
3. Write $q'_1 = a' + b'i$ and $q'_2 = c' + d'i$
4. Apply two N-dimensional complex FFTs of q'_1 and q'_2 giving $Q'_1 = A' + B'i$ and $Q'_2 = C' + D'i$.
5. Construct the quaternion in the Fourier domain: $Q' = A' + B'i + C'j + D'k$
6. Reverse the change of basis using equation 3.18 to give $Q = A + Bi + Cj + Dk$. The inverse QFT is done in a similar manner.

The Fourier transform of a quaternion is itself a quaternion which can be written in the polar representation (equation 3.5).

Next we show the importance of the selection of the QFT axis, μ_1. Choosing $\mu_1 = i$ gives a new basis of (i, j, k), equivalent to performing a change of basis using the identity matrix:

$$\boldsymbol{\mu} = \begin{pmatrix} ... & i & ... \\ ... & j & ... \\ ... & k & ... \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = I \tag{3.19}$$

Which maintains a quaternion in the basis (i, j, k). Returning to our vector representation of two component data in the frequency domain (equation 3.4), the choice of QFT axis should be considered. In the Fourier analysis of color images Ell and Sangwine (2007) choose $\mu_1 = (i, j, k)/\sqrt{3}$ (a unit vector pointing with equal length along each axis in $\Re^3$). This choice of axis comes naturally in the analysis of (R, G, B) images as this comprises one axis of a basis in which the simplex part represents intensity (luminance), while the perplex part represents the color information (chrominance). The choice of the axis can also be seen to represent the amount of mixing between elements of the quaternions in the Fourier domain. We do not see an analogy to luminance and chrominance in the representation of multicomponent seismic data. A synthetic data reconstruction example shown later in this paper demonstrates that the choice of QFT axis has little impact on the quality of reconstruction results. For these reasons, we make our selection of QFT axis based on the computational advantage of choosing $\mu_1 = i$, which leads to a change of basis using the identity matrix as shown in equation 3.19.

This greatly simplifies the workflow for the N-dimensional quaternion Fourier transform to:

1. Begin with quaternion $q = a + bi + cj + dk$
2. Write $q_1 = a + bi$ and $q_2 = c + di$
3. Apply two N-dimensional complex FFTs of q_1 and q_2 giving $Q_1 = A + Bi$ and $Q_2 = C + Di$.
4. Construct the quaternion in the Fourier domain: $Q = A + Bi + Cj + Dk$

In our reconstruction algorithm, this effectively eliminates the need for considering quaternion algebra, except in the computation of amplitude (equation 3.6).

To demonstrate the use of our vector representation of data we show FK spectra for the horizontal components of a common shot gather over an HTI medium in Figure 3.6. The east (a) and south (b) components appear nearly identical, but after rotation to radial (c) and transverse (d) components we uncover the weak signal of shear wave splitting in the transverse component. For this reason, the reconstruction of components individually clearly depends on the orientation of the input data. The quaternion FK spectrum for the east-south (e), and radial-transverse (f) configurations, on the other hand, do not depend on the orientation of the receivers, giving identical FK spectra in both cases. For this reason, the reconstruction of components together by a vector representation in the FK domain is invariant to the orientation of the receivers. Invariance to receiver orientation is an important distinction between scalar and vector processing strategies. This distinction was first identified by Cary (1995) to perform deconvolution prior to shear wave splitting analysis.

The main contribution of this paper is to modify the computation of the FK amplitude spectrum needed by our reconstruction algorithm to include more than one component of seismic data. The next section shows how the quaternion FK amplitude spectrum (equation 3.6) can be used to reconstruct more than one component of data simultaneously.

3.1.3 Projection Onto Convex Sets (POCS)

Projection Onto Convex Sets (POCS) can be used to reconstruct seismic data by iteratively thresholding the amplitude spectrum while using a reinsertion operator to control which traces are reconstructed. It is shown to be an effective method for seismic data reconstruction

(Abma and Kabir, 2006; Gao et al., 2012; Wang et al., 2010c; Stein et al., 2010), but typically requires many iterations to achieve good results. A data-driven thresholding schedule has been shown to give a significant improvement in the number of iterations required while still achieving good results (Gao et al., 2012). A modification to the reinsertion operator allows for the algorithm to also be used for denoising of seismic data (Gao et al., 2012).

For a given temporal frequency, ω, a quaternion of data in the $(\omega, m_x, m_y, h, az)$ domain at the n^{th} iteration of POCS is given by

$$Q^n = \alpha Q^{obs} + (1 - \alpha S)\mathcal{F}_Q^{-1} T \mathcal{F}_Q Q^{n-1}, n = 1, ..., N, \tag{3.20}$$

where Q^{obs} is the quaternion representation of the original data with missing traces, and $\mathcal{F}_Q$ and $\mathcal{F}_Q^{-1}$ are the forward and inverse 4D quaternion Fourier transforms respectively. In this notation $Q^n(\omega, k_{m_x}, k_{m_y}, k_h, k_{az}) = \mathcal{F}_Q Q^n(\omega, m_x, m_y, h, az)$, and T is an iteration dependent quaternion threshold operator. The sampling operator is given by S and is equal to one for points with existing traces and zero for points with unrecorded observations. The scaling factor $\alpha \leq 1$ can be used to simultaneously denoise the data. A choice of $\alpha = 1$ reinserts the noisy original data at each iteration, whereas a lower value of α will denoise the volume by taking an average of the original and reconstructed data. The threshold operator, T, is designed using an amplitude distribution derived from the input data. The details for deriving the threshold operator are found in Gao et al. (2012).

3.2 Synthetic data example

The synthetic data we consider for our reconstruction tests is the elastic finite-difference synthetic generated using an HTI model (Scenario 2), in the radial-transverse coordinate frame. Figure 3.7 shows an example of the reconstruction of a synthetic 3D-2C common shot gather with $SNR = 5$ shown for a constant offset of 560m, which has been decimated to 40% of the original number of traces (left column). The data are reconstructed using both component-by-component reconstruction (middle column), and vector reconstruction (right column). The quality of the reconstruction is improved with the use of vector reconstruction. Figure 3.8 shows the same plots for a constant azimuth of 20°. Figure 3.9 (a) displays a plot of the quality of the reconstruction versus the percentage of traces decimated. The quality of the reconstruction is measured using

$$R - SNR = 10 \log \frac{||\boldsymbol{d_o}||_2^2}{||\boldsymbol{d} - \boldsymbol{d_o}||_2^2}, (dB) \tag{3.21}$$

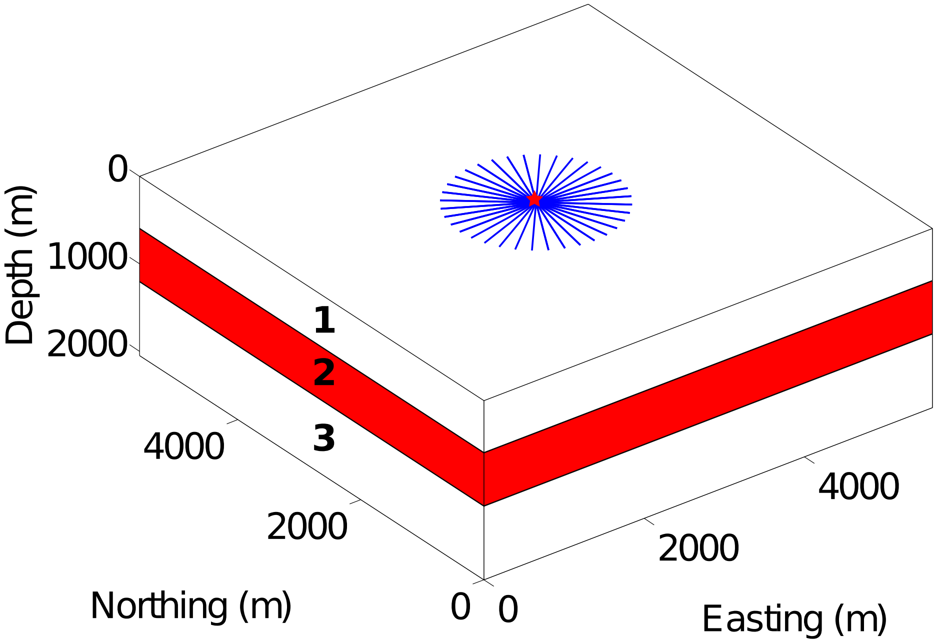


Figure 3.3: Three layer model used for the generation of elastic finite-difference synthetic data examples. Two scenarios are modelled. In the first scenario all three layers are isotropic, while in the second scenario layer 2 has Horizontal Transversely Isotropic (HTI) symmetry. The parameters for each layer are given in table 3.1. The acquisition geometry consists of 18 receiver lines (blue) oriented about a single shot (red).

where $\boldsymbol{d_o}$ and $\boldsymbol{d}$ are the true and reconstructed data respectively. The results are plotted for both traditional component by component POCS reconstruction (dashed lines), and vector POCS reconstruction (solid lines). The quality of the reconstructions for both components is noticeably improved through the use of vector reconstruction, although the transverse component sees the greatest improvement. This is likely due to the fact that the transverse component has a lower SNR compared to the radial component. During the course of our tests, we have found that traditional component-by-component reconstruction provides slightly better results in the east-south coordinate frame than in the radial-transverse coordinate frame. This is because the east-south coordinate frame mixes the weak transverse wavefield with the much stronger radial wavefield between two components. The overlap of the two components provides a higher amplitude value in the FK domain (in contrast to the artifacts in this domain related to noise or missing traces), leading to an improved reconstruction (this effect could be termed "spectral overlap"). Vector reconstruction shows an improved reconstruction for a similar reason, spectral overlap, but in this case, the quaternion FK spectrum is invariant to the orientation of the input data.

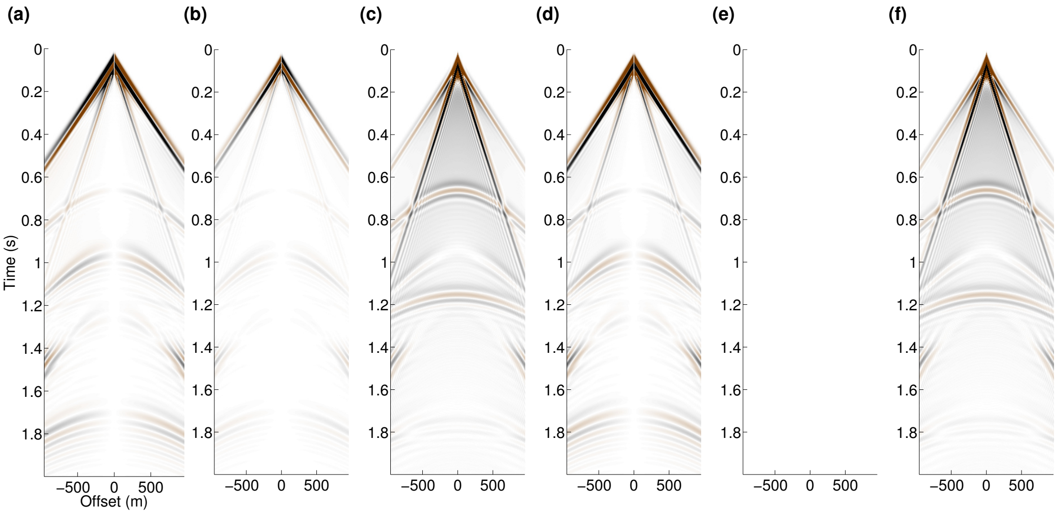


Figure 3.4: Elastic finite-difference synthetic shot gathers for Scenario 1 (azimuth 20°) (isotropy). Before reorientation: (a) east, (b) south, (c) downward, and after reorientation: (d) radial, (e) transverse, (f) downward. Notice that the energy in the horizontal plane can be placed into a single component by rotation.

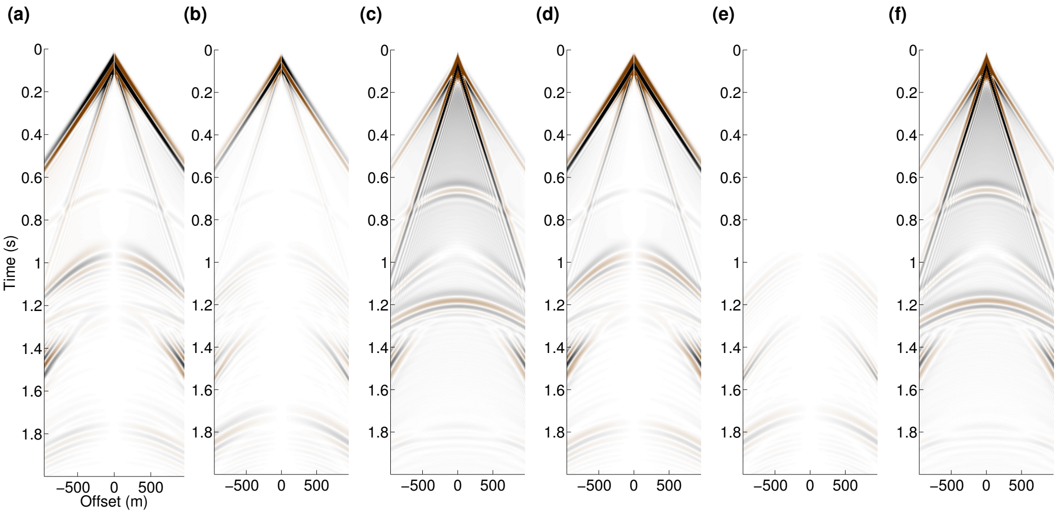


Figure 3.5: Elastic finite-difference synthetic shot gathers for Scenario 2 (azimuth 20°) (Horizontal Transverse Isotropy). Before reorientation: (a) east, (b) south, (c) downward, and after reorientation: (d) radial, (e) transverse, (f) downward. Notice that the energy in the horizontal plane cannot be placed into a single component by rotation.

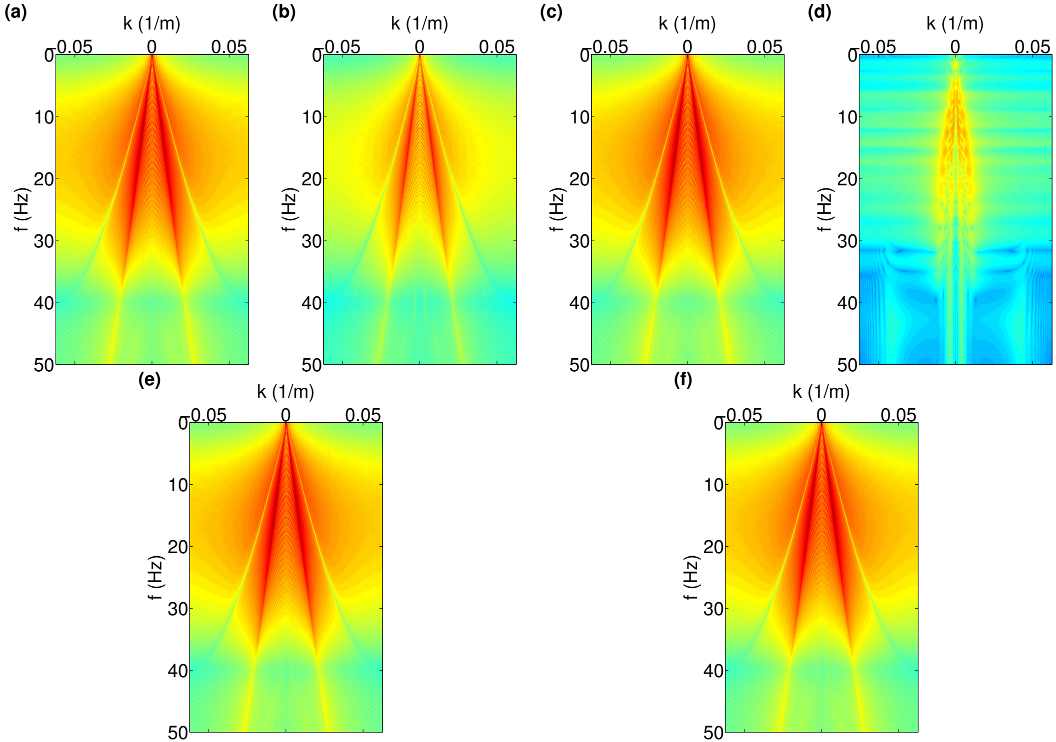


Figure 3.6: FK spectra of the horizontal components shown in Figure 3.5. (a) East component, (b) south component, (c) radial component, (d) transverse component, (e) quaternion FK spectrum for east and south components, (f) quaternion FK spectrum for radial and transverse components. Notice that the quaternion FK spectrum does not depend on orientation ((e) and (f) are identical).

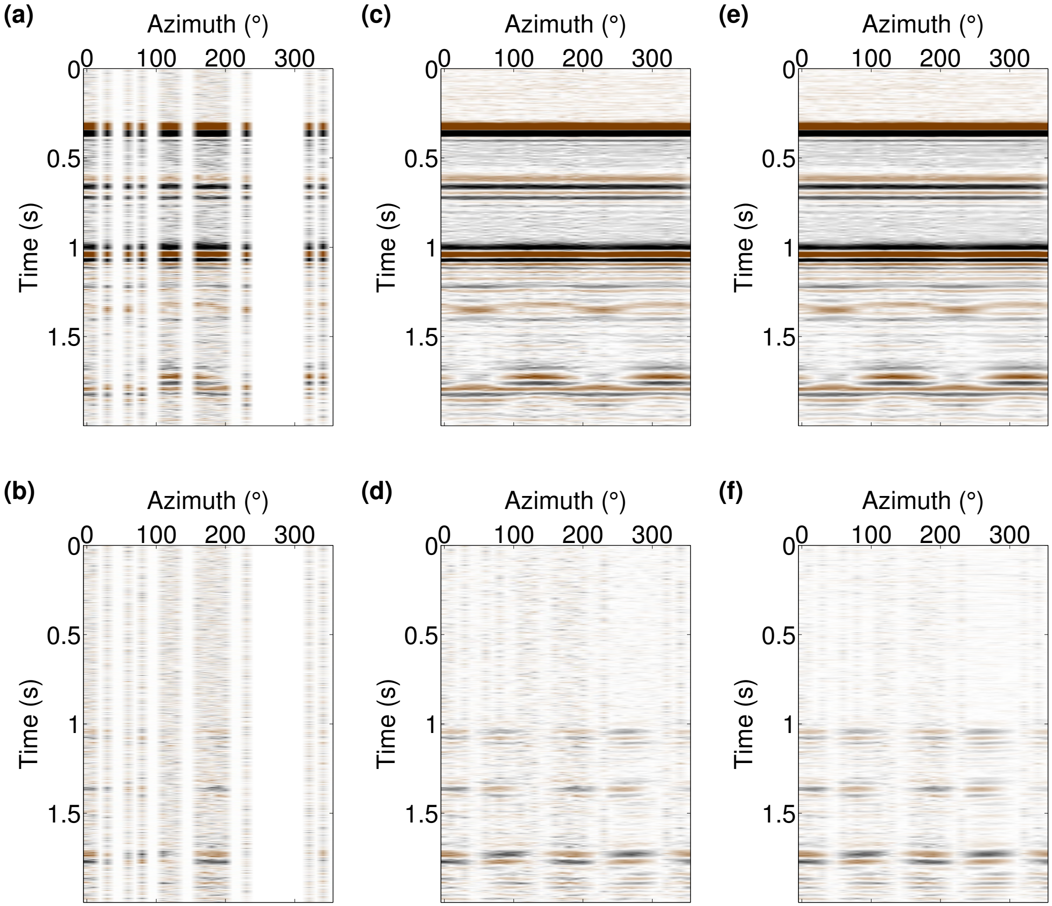


Figure 3.7: Comparison of component by-component reconstruction to vector reconstruction for a synthetic 3D-2C common shot gather: Common-offset display (560m). The input has 60% missing traces and $SNR = 5$. (a) Component 1 input, (c) component 1 after single component reconstruction, (e) component 1 after vector reconstruction. (b), (d), and (f) show the same for component 2.

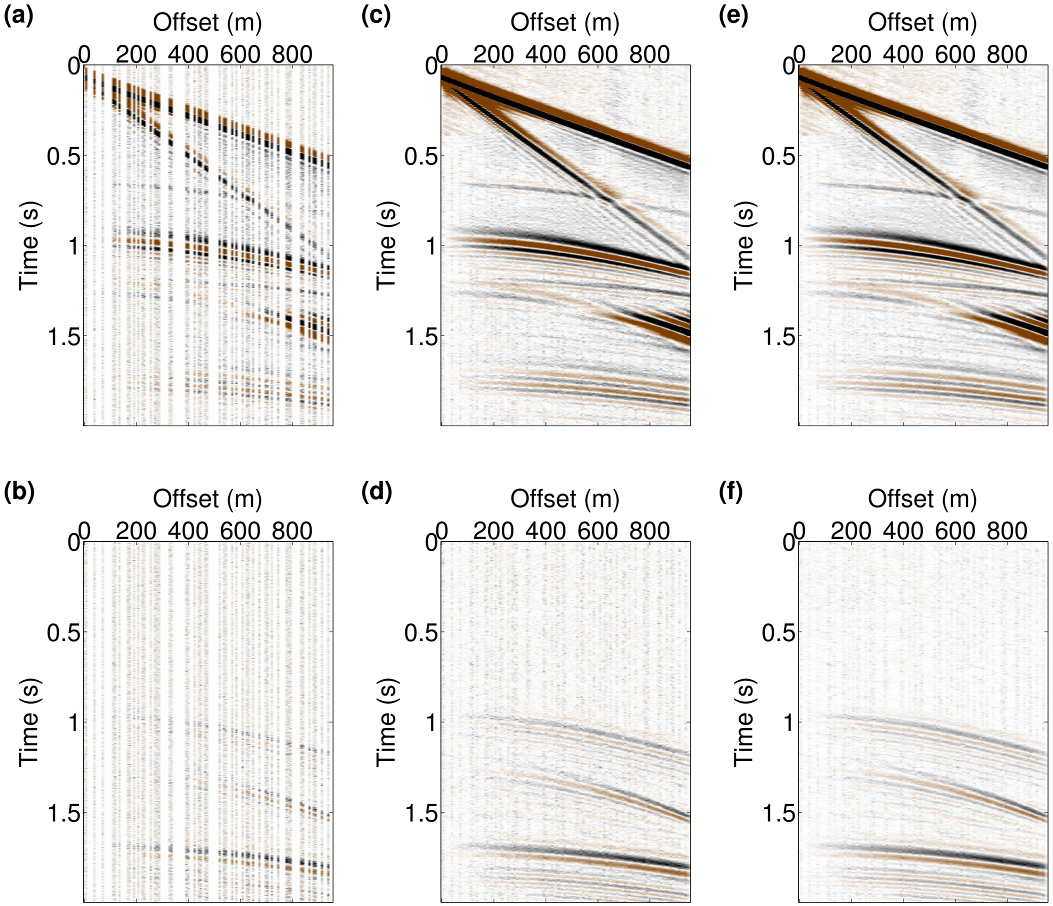
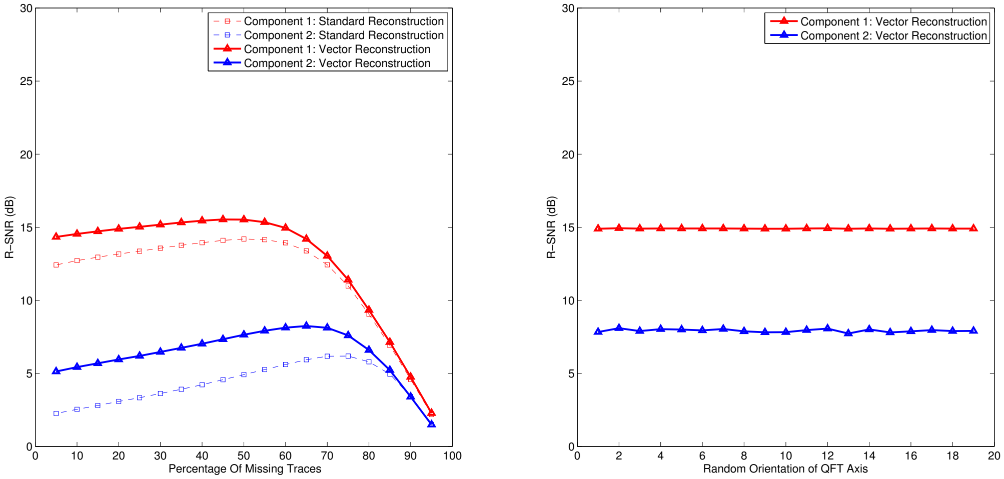


Figure 3.8: Comparison of component by-component reconstruction to vector reconstruction for a synthetic 3D two component common shot gather: Common-azimuth display (20°). The input has 60% missing traces and $SNR = 5$. (a) Component 1 input, (c) component 1 after single component reconstruction, (e) component 1 after vector reconstruction. (b), (d), and (f) show the same for component 2.



(a) Reconstruction quality as a function of the percentage of missing traces.

(b) Reconstruction quality as a function of randomly oriented QFT axis.

Figure 3.9: Quality for the reconstructions shown in Figures 3.7 & 3.8 when the percentage of missing traces is varied (a), and the QFT axis orientation is randomly oriented. The input data have an SNR of 5.

Next, we demonstrate that the choice of QFT axis, μ_1, has little effect on the quality of the reconstruction result. We repeat the reconstruction of the data shown in Figures 3.7 and 3.8 with twenty different randomly oriented choices of QFT axis. Figure 3.9 (b) shows that for each choice of μ_1 the reconstruction quality is relatively constant. This gives us confidence that we can choose $\mu_1 = i$, giving an important computational improvement by eliminating the change of basis operation for every forward and inverse quaternion Fourier transform.

3.3 Real data example

For a real data example we consider the 5D reconstruction of a converted wave land dataset acquired over a heavy oil reservoir. The reservoir is being produced by a fluid injection technique, and time-lapse analyses demonstrate changes in shear wave splitting effects as a result of production. While shear wave splitting is commonly associated with thickness and intensity of fractured zones, in this case the changes in shear wave splitting are thought to be associated with changes in horizontal stresses as a result of fluid injection (Wikel et al., 2012). The shot and receiver locations for the acquisition are shown in Figure 3.10. There are significant surface obstructions leading to patches of low fold in the compressional wave (P-P) data as shown in Figure 3.11a. The converted wave (P-S) data suffer from the additional problem of asymmetric binning. Asymptotic Conversion Point (ACP) binning

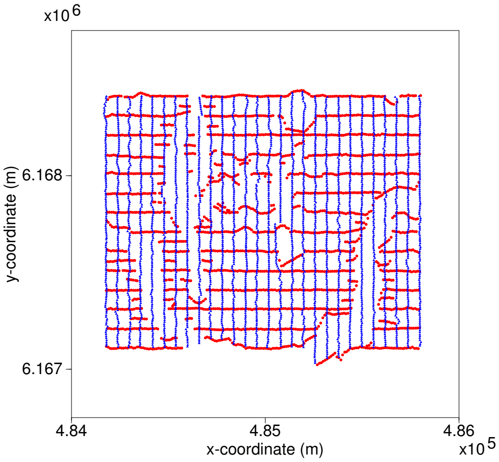
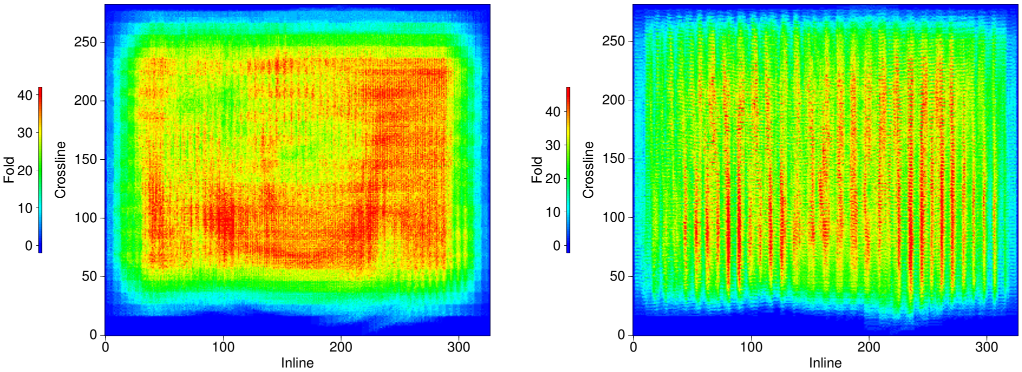


Figure 3.10: Shot (red) and receiver (blue) coordinates for a converted wave survey over a heavy oil reservoir.

(Fromm et al., 1985) with $\gamma = 3$ (where γ is defined as an average P and S-wave RMS velocity ratio) gives a fold map with high fold clustered around receiver lines as shown in Figure 3.11b. In addition to constraints on acquisition, 5D reconstruction is also beneficial in this case due to the need for dense sampling of azimuths for shear wave splitting analysis and compensation. In the absence of HTI media, the radial component is expected to contain all of the P-S energy, while the transverse component should appear free from signal. In the presence of HTI media the P-S wavefield is split into fast and slow components which appear as characteristic patterns in the radial and transverse components (Li, 1998; Cary, 2002). Generally speaking, the travel-times of the radial component display an apparent sinusoidal variation across azimuths, while the transverse component displays polarity reversals across azimuths. Without sharing traces from neighboring bins (superbinning) it is difficult to identify these patterns.

Figure 3.12 displays offset vs. azimuth for all traces in the survey in the offset range of 0-550m. This figure demonstrates that over the entire survey nearly all offsets and azimuths are represented, while Figure 3.13 (a) demonstrates that for a given bin location (150,150) many offsets and azimuths are missing. The basis of 5D regularization is that these data can be used to predict traces at individual bin locations. Figure 3.13 (b) shows bin (150,150) after 5D regularization. The data are binned into 5x5m Asymptotic Conversion Point (ACP) bins, 50m offset bins, and 20° azimuth bins. Unlike standard P-P wavefield regularization, source-receiver reciprocity cannot be used in the case of PS acquisition, requiring azimuths to be defined on the interval 0-359° and leading to a higher degree of sparsity. In this example, the offset is limited to 550m to mitigate the effects of the limited



(a) $\gamma = 1$ (b) $\gamma = 3$

Figure 3.11: total fold calculated over the offset interval 0-550m using $\gamma = 1$ and $\gamma = 3$ to calculate the bin location (Common Mid-Point). Notice the relatively even distribution of fold around source and receiver lines when $\gamma = 1$, compared to the clustering of high fold bin locations around receiver lines when $\gamma = 3$.

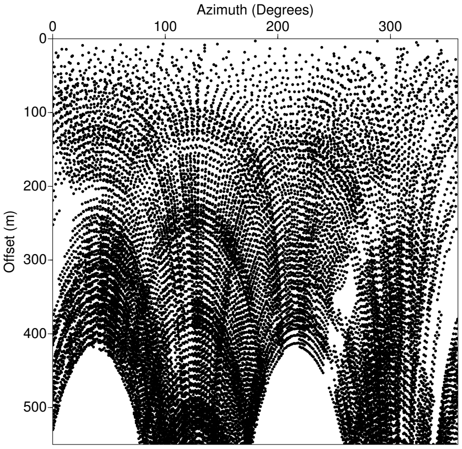
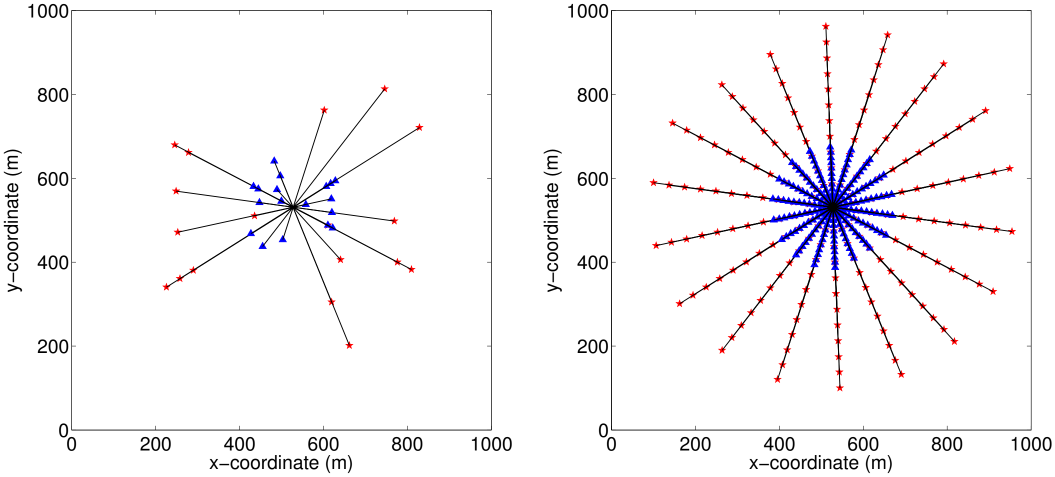


Figure 3.12: Azimuth vs. Offset plotted for all traces in one 26x26 patch of ACP bins centered at ACP (150,150). The offset has been limited to the interval 0-550m. Azimuth is calculated as the angle from source to receiver counter-clockwise from east. The two missing lobes centered at azimuths 38° and 218° correspond to the crossline directions which limit the far offset due to the limited size of the recording patch. The maximum offset for the reconstruction was limited to 550m to mitigate the effect of missing azimuths at far offsets.

size of the recording patch in the crossline direction. The reconstructions are carried out on a 26-crossline-wide region of data in patches of 26x26 ACPs with an overlap of 13 ACPs. Each patch of data has approximately 92% of traces missing, but of the 8% "alive" traces within each patch, approximately 35% are redundant. This redundancy can be reduced by decreasing bin increments, but at the expense of decreasing the percentage of "alive" traces, and increasing the need for computational memory (Trad, 2009). Due to the low SNR of the input data we find best results when using the denoising parameter $\alpha = 0.1$ with a relatively high number (500) of iterations. Both the standard and vector reconstructions are parameterized identically for comparison purposes.

The results of component-by-component reconstruction and vector reconstruction are shown in Figure 3.14. Radial and transverse components sorted by offset and azimuth are shown for ACP gather (150,150). The radial and transverse input gathers (a) and (d) have many missing offsets and azimuths, making it difficult to identify the effects of shear wave splitting. The data after component-by-component reconstruction (b) and (e) make it easier to identify these effects. The radial component shows a sinusoidal-like pattern of travel-times across azimuths at approximately 800ms, while the transverse component shows characteristic polarity reversals across azimuths at this same event. The vector reconstructions (c) and (f) show a slightly improved reconstruction of the radial component. For the transverse component, the improvement is more noticeable, making the effects of shear wave splitting easier to identify. Figure 3.16 shows a zoomed-in display of these data at an offset of 275m. The shear wave splitting effect at approximately 800ms is more easily identified in the transverse component after vector reconstruction (f) compared with the result after component-by-component reconstruction (e).

To better evaluate the reconstruction results shear wave splitting corrections are applied to the data. The inversion of the splitting parameters is done via a grid search with a cost function aimed at minimizing the energy on the transverse component. This cost function is given by equation (8) of Simmons (2009). The goal of shear wave splitting analysis is to find two parameters. The first parameter, Δt, is the travel time difference between a shear wave traveling in the fast direction and the slow direction. The second parameter, ϕ_{S1}, is the azimuth of the fast direction of the anisotropy. A layer stripping approach is applied for layers at times of 400, 600, and 800ms. The inversion of the splitting parameters was carried out for each of the inputs individually, resulting in the following parameter estimates for ACP gather (150,150). The input data before reconstruction provide the splitting parameters $\phi^1_{S1} = 103°$, $\Delta t^1 = 0$ms, $\phi^2_{S1} = 107°$, $\Delta t^2 = 1$ms, $\phi^3_{S1} = 100°$, and $\Delta t^3 = 1$ms, where ϕ^1_{S1} and Δt^1 indicate the azimuth of the fast direction and time lag for layer 1 respectively. The data after component-by-component reconstruction provide the parameters $\phi^1_{S1} = 104°$, $\Delta t^1 = 1$ms, $\phi^2_{S1} = 105°$, $\Delta t^2 = 1$ms, $\phi^3_{S1} = 103°$, and $\Delta t^3 = 1$ms, while the data after vector reconstruction provide the parameters $\phi^1_{S1} = 103°$, $\Delta t^1 = 1$ms,



(a) Before

(b) After

Figure 3.13: Shot (red) and receiver (blue) coordinates for ACP bin (150,150) before 5D regularization (a), and after regularization (b). Black lines connect shot-receiver pairs. For 5D regularization the data are binned into 5x5m ACP bins, 50m offset bins, and 20° azimuth bins. For the ACP binning a value of $\gamma = 3$ was used. This positions the bin 3/4 the distance from source to receiver, giving an asymmetric pattern of sources and receivers for each ACP location.

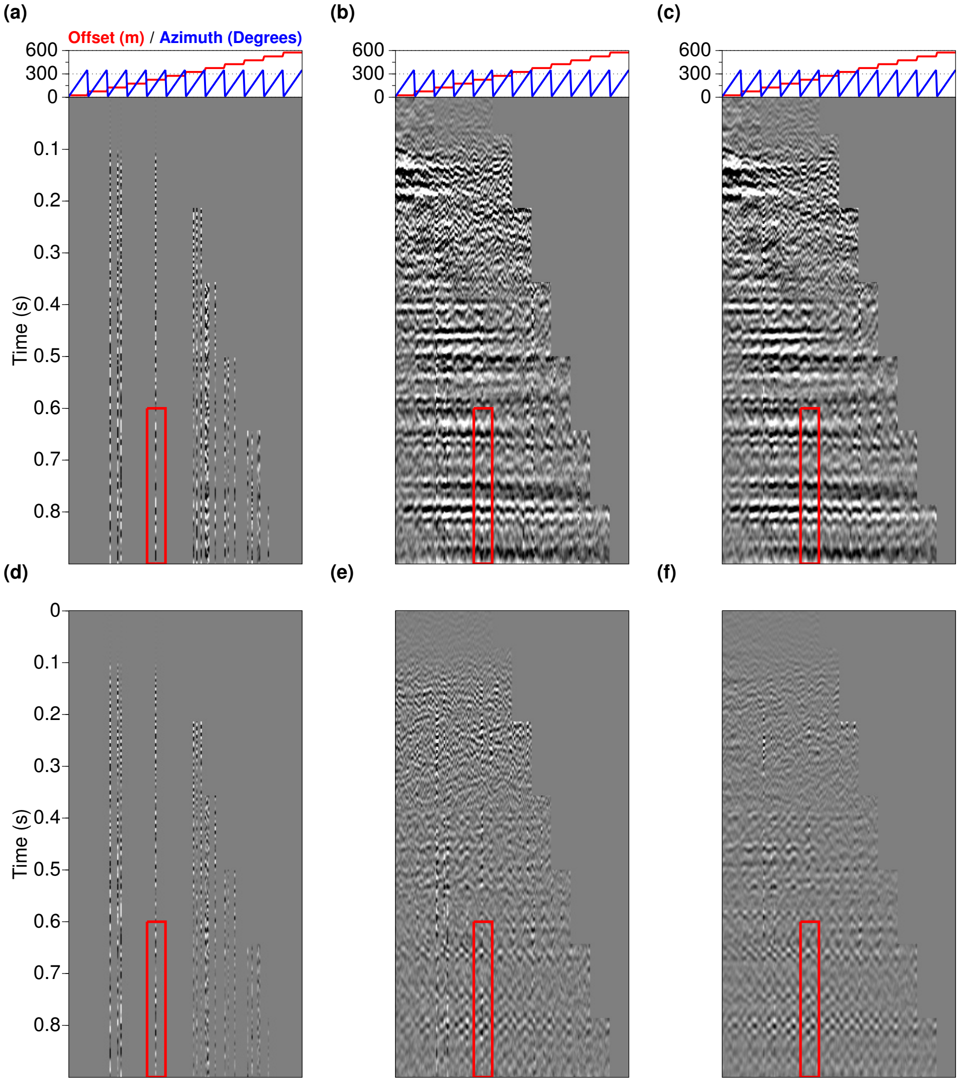


Figure 3.14: 5D reconstruction results shown for one ACP gather (150,150) sorted by offset, azimuth. (a) Radial component before reconstruction, (b) radial component after component-by-component reconstruction, (c) radial component after vector reconstruction. (d), (e), and (f) respectively show the same results for the transverse component. The area highlighted in red is shown in Figure 3.16

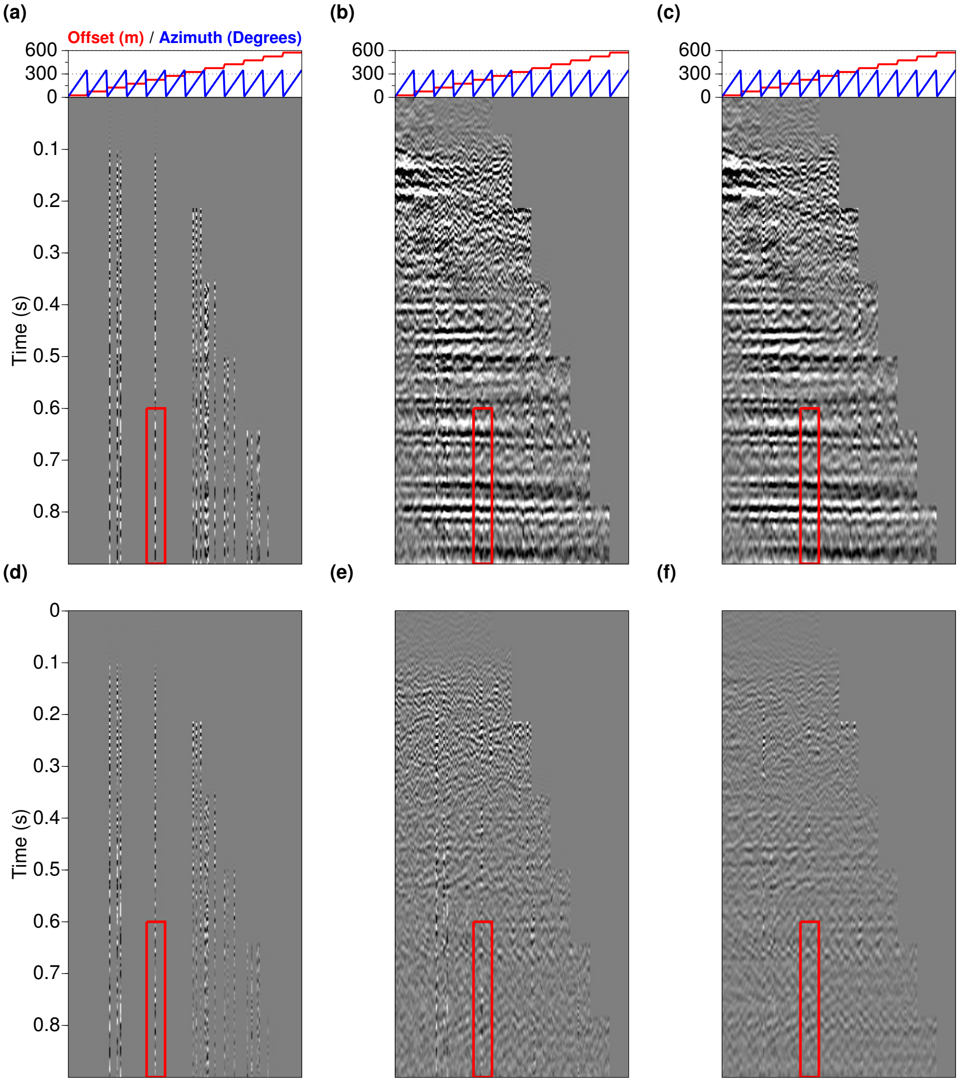


Figure 3.15: 5D reconstruction results shown for one ACP gather (150,150) sorted by offset, azimuth, shown after shear wave splitting correction. (a) Radial component before reconstruction, (b) radial component after component-by-component reconstruction, (c) radial component after vector reconstruction. (d), (e), and (f) respectively show the same results for the transverse component. The area highlighted in red is shown in Figure 3.17

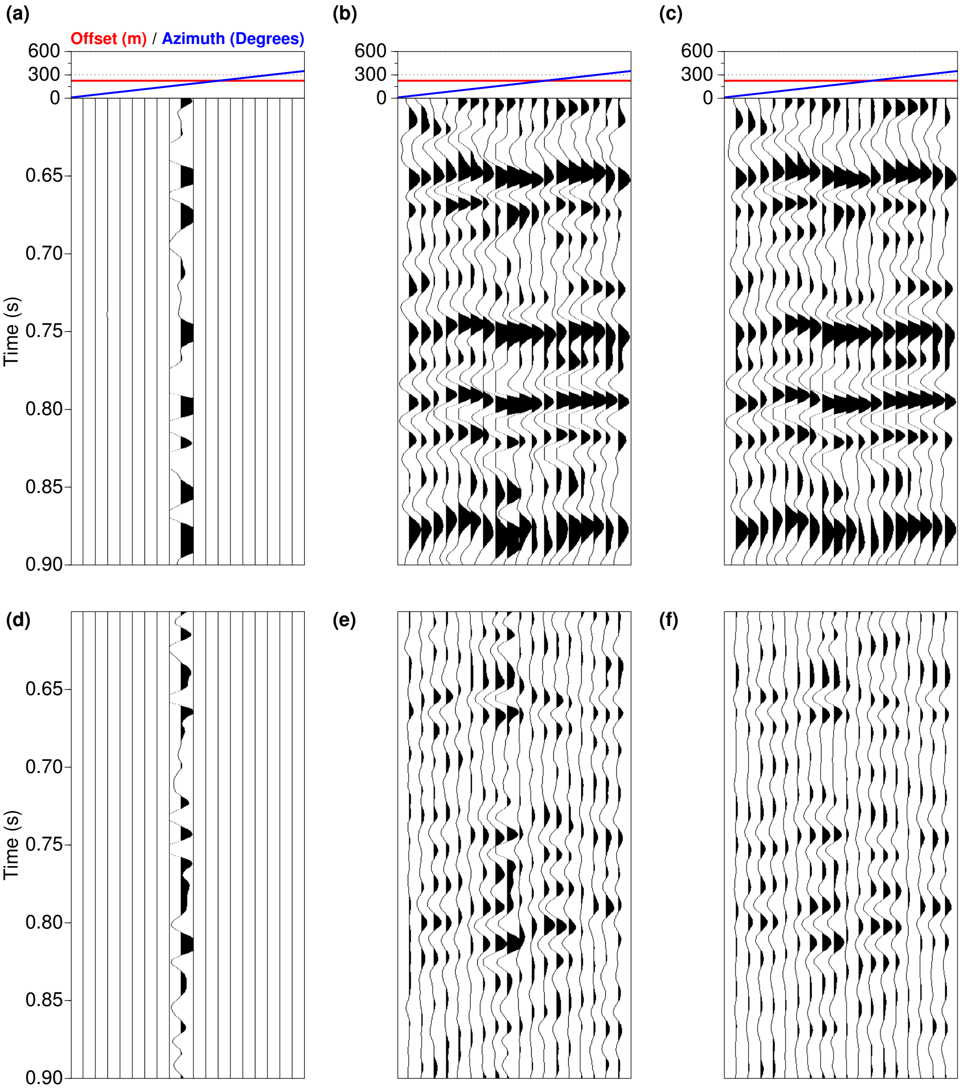


Figure 3.16: 5D reconstruction results shown for offset 375m of ACP gather (150,150) sorted by azimuth. (a) Radial component before reconstruction, (b) radial component after component-by-component reconstruction, (c) radial component after vector reconstruction. (d), (e), and (f) respectively show the same results for the transverse component.

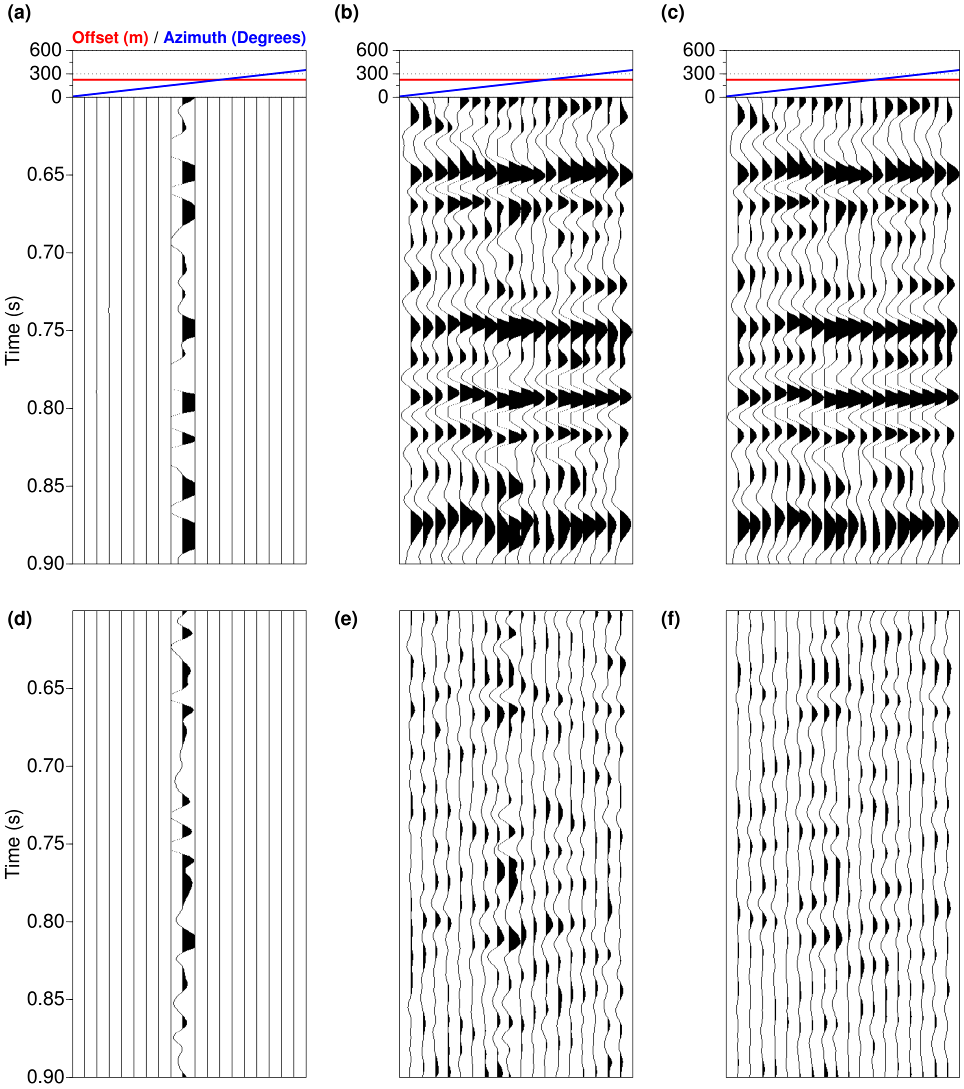


Figure 3.17: 5D reconstruction results shown for offset 375m of ACP gather (150,150) sorted by azimuth, shown after shear wave splitting correction. (a) Radial component before reconstruction, (b) radial component after component-by-component reconstruction, (c) radial component after vector reconstruction. (d), (e), and (f) respectively show the same results for the transverse component.

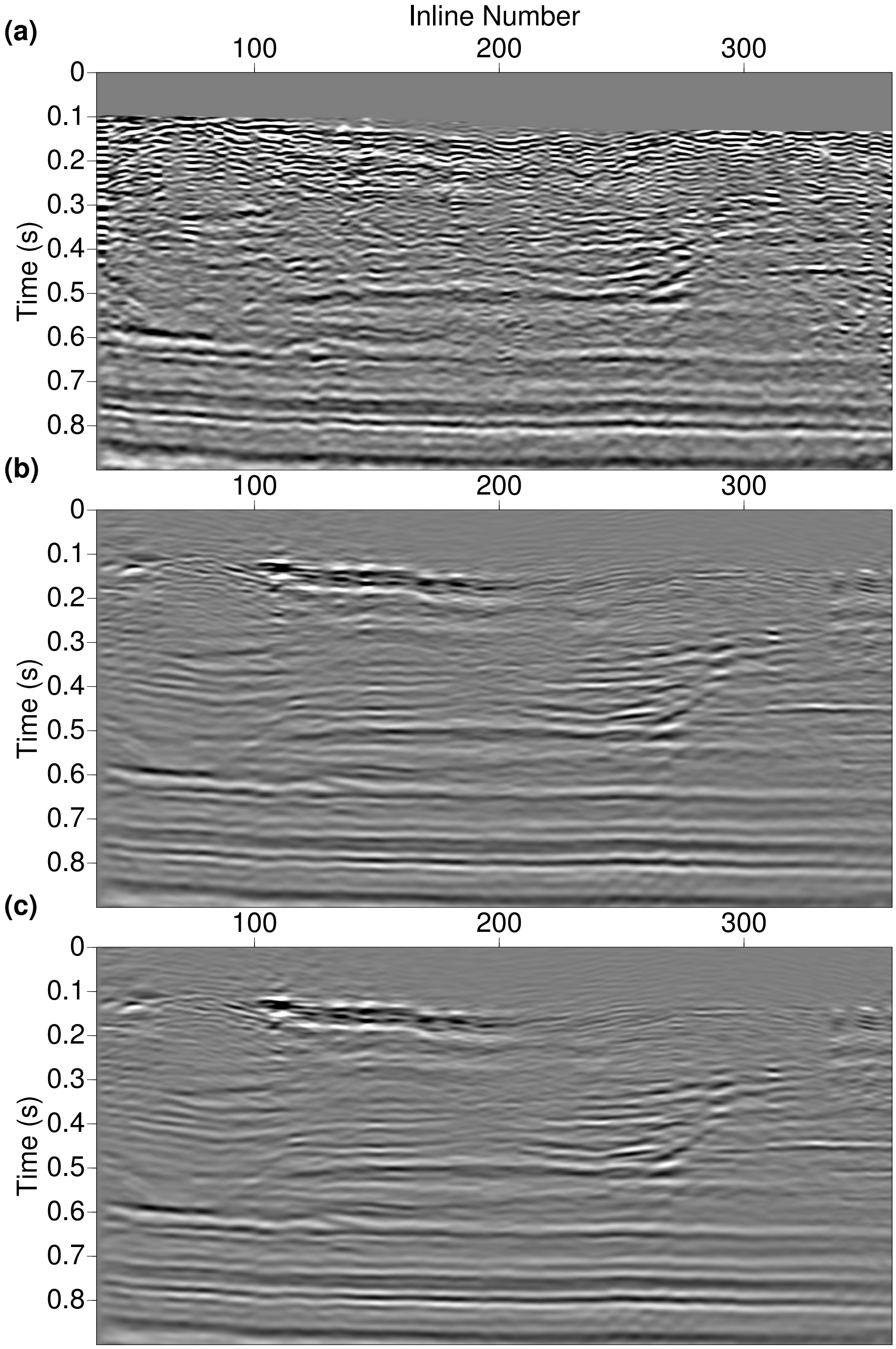


Figure 3.18: 5D reconstruction results shown for one stacked crossline (150) of the radial component sorted by inline. Shear wave splitting corrections have been applied. (a) Radial component before reconstruction, (b) radial component after component-by-component reconstruction, (c) radial component after vector reconstruction.

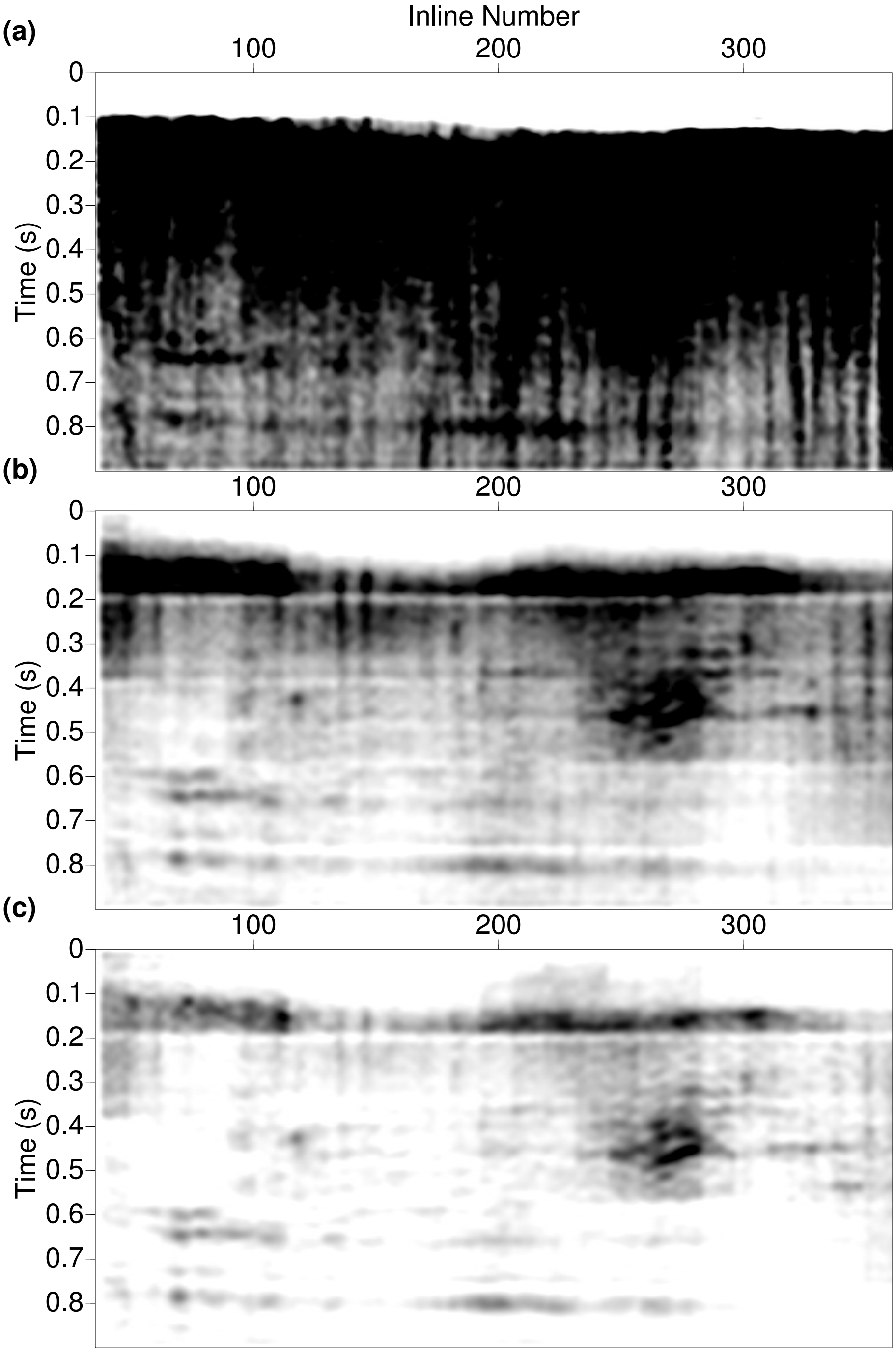


Figure 3.19: 5D reconstruction results shown for one stacked crossline (150) of the transverse component sorted by inline. Shear wave splitting corrections have been applied. The complex envelope for each trace is input to the stacking process to avoid the effect of polarity reversals across azimuth. (a) Transverse component before reconstruction, (b) transverse component after component-by-component reconstruction, (c) transverse component after vector reconstruction.

$\phi^2_{S1} = 106°$, $\Delta t^2 = 1$ms, $\phi^3_{S1} = 105°$, and $\Delta t^3 = 1$ms. All three inputs provide similar values for this particular gather. Figure 3.15 shows the results of shear wave splitting analysis and compensation for ACP gather (150,150). On the input gathers it is difficult to see any improvement by the process, while on the component-by-component reconstruction and vector reconstruction results it is apparent that the radial component is more flat, and the transverse component has a more dramatic reduction in energy. In the zoomed displays of Figure 3.17 at an offset of 275m the apparent sinusoid pattern on the radial component appears flatter and the energy of the transverse component appears to be lessened by the process. For the transverse component, the vector reconstruction results show a lower level of amplitude compared with the component-by-component reconstruction result.

Stack sections of the radial component after shear wave splitting correction for crossline 150 are shown in Figure 3.18. The input stack (Figure 3.18a) suffers from irregular low fold. Component-by-component reconstruction (Figure 3.18b) provides an improvement to the quality of the stack, while vector reconstruction (Figure 3.18c) offers a very minor improvement to the quality of the radial stack. Notice the improved flatness and amplitude of the horizon at 800ms and inline number 270. A greater improvement is seen in the transverse stack (Figure 3.19). A reduction in amplitude in the transverse component after application of shear wave splitting corrections is commonly used as a measure of the quality of the chosen parameters. This stack was generated following Simmons (2009), where the envelope of complex traces are stacked to demonstrate the level of energy in the data by avoiding the amplitude canceling effect of the polarity reversals on the transverse component. Notice the improved reduction in amplitude of the horizon at 800ms after vector reconstruction (Figure 3.19c) compared with the result after component-by-component reconstruction (Figure 3.19b).

3.4 Discussion

We find an improvement in the reconstruction of two component 5D data when using a vector reconstruction method that considers both components simultaneously compared to component-by-component reconstruction. The improvement is due to overlap of the components in the frequency-wavenumber domain which is advantageous in vector POCS reconstruction. We have also shown that our vector representation of multicomponent seismic data in the frequency-wavenumber domain is invariant to the orientation of the input data. A drawback of the method is that if the components do not overlap in the frequency-wavenumber domain there is little benefit in considering the components simultaneously. During the course of our research, we first attempted to represent three component data by forming the quaternion of data in the time domain (Stanton and Sacchi, 2011). In the reconstruction of real data, we find that three component representations produce slightly worse results than component-by-component reconstructions. This is likely due to the fact that the Vertical component contains mainly the P-P wavefield while the horizontal components mainly contain the P-S wavefields. This results in a decrease in spectral overlap which is a disadvantage in vector POCS reconstruction. Reconstruction of three component seismic data also adds a further complication in the context of binning geometry. In our real data example, we reconstruct 5D data to a regular subsurface grid (Asymptotic Conversion Point, offset, azimuth). Including the Vertical component in this type of reconstruction would require binning the Vertical component onto a different subsurface grid (Common Mid-Point, offset, azimuth), giving a multi-valued sampling operator. An alternative would be to perform 5D reconstruction in surface geometry (source, receiver), which is common for all components. Since the analysis of shear wave splitting is typically done considering subsurface binning this type of reconstruction does not fit our objective. Despite these complications, our proposed algorithm can be easily extended to a higher number of components by modifying equation 3.4 to give a time domain representation of multiple components $D(t,\vec{x}) = H(t,\vec{x}) + X(t,\vec{x})i + Y(t,\vec{x})j + Z(t,\vec{x})k$, where $H(t,\vec{x})$ is a scalar component such as a hydrophone in the time domain that is invariant to rotation, and $X(t,\vec{x})$, $Y(t,\vec{x})$, $Z(t,\vec{x})$ are the three vector components with an arbitrary orientation. The proposed algorithm for vector reconstruction should then be modified to include a quaternion Fourier transform along the time axis to bring the data into the $(\omega,\vec{x})$ domain before reconstruction. This transform does not have the conjugate symmetry along the frequency axis that our current 2C implementation has (leading to an increase in the memory and computational time required), but the rest of the algorithm remains unchanged. An implementation of our algorithm for a higher number of components is a subject of future research.

An alternative to a quaternion formulation is to use the complex numbers to provide a two component representation of data in the time domain. A drawback of this approach is that it destroys the conjugate symmetry along the frequency axis, requiring reconstruction on both halves of the axis[3]. We prefer a two component representation in the frequency domain using the quaternions as this preserves the conjugate symmetry along the frequency axis. The number of computational operations required when using complex numbers and quaternions is equivalent, but the quaternion approach offers a reduced computational memory requirement due to conjugate symmetry. Another advantage is that the sampling operator, T, in the case of a complex representation can not handle different sampling for each component, while the quaternion representation in the frequency domain can handle different sampling for each component. In the course of this study, we also find a connection of the quaternion formulation to that of group sparsity (Rodriguez et al., 2012), which points to

---

[3] Loading two real-valued traces into the real and imaginary part of a complex trace is an often used trick in digital signal processing to speed up the computation of the Fourier transform for two real valued traces by a factor of two (M. van der Baan, personal communication, April 7, 2017).

another strategy to handle a higher number of components.

3.5 Generalization of the approach

While the above formulation of vector regularization makes use of the POCS method, the core concept of vector regularization is readily extended to a wider class of Fourier regularization techniques. An often used Fourier regularization technique in the industry is Minimum Weighted Norm Interpolation (MWNI) (Liu and Sacchi, 2004). The method begins with the objective

$$J = ||\boldsymbol{T}(\boldsymbol{F}_q\boldsymbol{m} - \boldsymbol{d})||_2^2 + ||\boldsymbol{\mu}^T\boldsymbol{m}||_2^2 \tag{3.22}$$

where $\boldsymbol{F}_q$ is the forward quaternion Fourier transform over the spatial axes, $\boldsymbol{T}$ is a diagonal sampling operator with 1's in place of observations and 0's in place of missing traces, $\boldsymbol{\mu}$ is a multicomponent trade-off parameter to control the level to fit the data versus the level to minimize the $l-2$ norm of the model, $\boldsymbol{m}$ are vector-valued Fourier coefficients for the fully sampled noise free data components (the unknown model we seek), and $\boldsymbol{d}$ is the frequency slice of recorded vector data binned onto a regular spatial grid. The minimum norm solution to this problem is

$$\boldsymbol{m} = (\boldsymbol{F}_q^T\boldsymbol{T}\boldsymbol{F}_q + \boldsymbol{\Upsilon})^{-1}\boldsymbol{F}_q^T\boldsymbol{T}\boldsymbol{d} \tag{3.23}$$

where $\boldsymbol{\Upsilon} = \begin{bmatrix} \mu_1 & 0 & 0 \\ 0 & \mu_2 & 0 \\ 0 & 0 & \mu_3 \end{bmatrix}$. Because $\boldsymbol{F}_q$ is an orthogonal operator ($\boldsymbol{F}_q^T = \boldsymbol{F}_q^{-1}$), this solution reduces to

$$\boldsymbol{m} = \boldsymbol{F}_q\boldsymbol{d}, \tag{3.24}$$

which implies that

$$\boldsymbol{d}_{interp} = \boldsymbol{d}, \tag{3.25}$$

which is not a helpful result to say the least. A solution is to impose a sparsity constraint on the spectrum of the interpolated data:

$$J = ||\boldsymbol{T}(\boldsymbol{F}_q\boldsymbol{m} - \boldsymbol{d})||_2^2 + ||\boldsymbol{\mu}^T\boldsymbol{W}_m\boldsymbol{m}||_2^2 \tag{3.26}$$

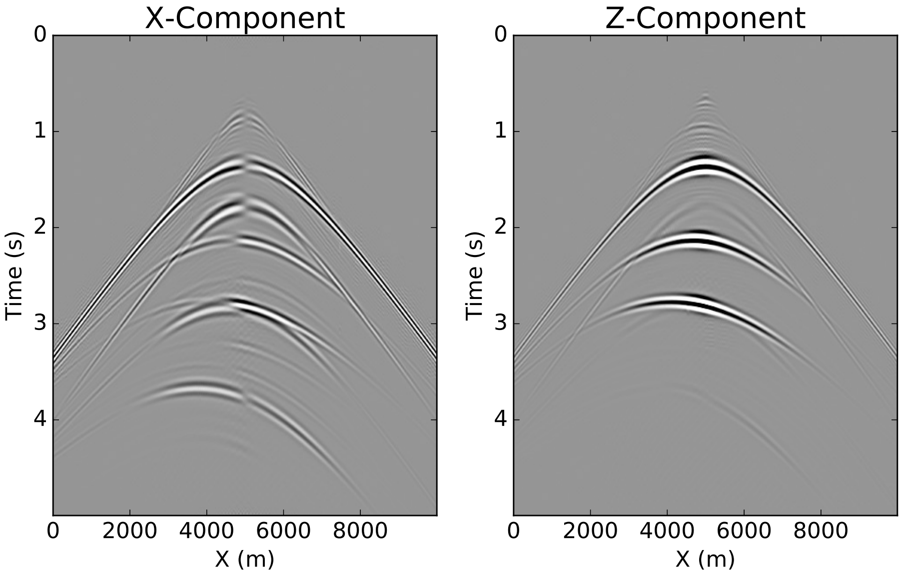


Figure 3.20: decimated data (as shown in figure 3.1) interpolated using scalar POCS. The qualities of the reconstructions are 23.07 and 29.73 dB respectively.

The weighting in equation 3.26 is designed to promote high amplitudes, corresponding to spectral coefficients of the true undecimated data, while weighting down low amplitude sampling artefacts. This equation can be solved via Iteratively Reweighted Least Squares (IRLS), with an inner loop minimized via conjugate gradients. Because equation 3.26 involves the use of vector valued model, data, and regularization parameters, a scalar version of conjugate gradients will not suffice. Appendix B discusses the application of the Conjugate Gradients algorithm to multicomponent-valued model and data vectors. This solver is re-used in chapters 4 and 5 to extend acoustic least squares migration to the elastic case. These algorithms were implemented using the *multiple dispatch* feature of the Julia programming language, and publicly released as part of an official Julia package: Seismic.jl. An overview of this package is provided in Appendix D.

To validate the extension of MWNI to the vector case I next the algorithm to the decimated elastic finite difference synthetic shot gather shown in figure 3.1. Figures 3.20 and 3.21 show the result of applying scalar POCS interpolation, while figures 3.22 and 3.23 show the result of applying scalar MWNI interpolation. For such a simple case we expect both algorithms to perform similarly well. Figures 3.24 and 3.25 show the result of applying vector POCS interpolation, while figures 3.26 and 3.27 show the result of applying vector MWNI interpolation. Again, both approached provide moderately better results with virtually identical computational cost.

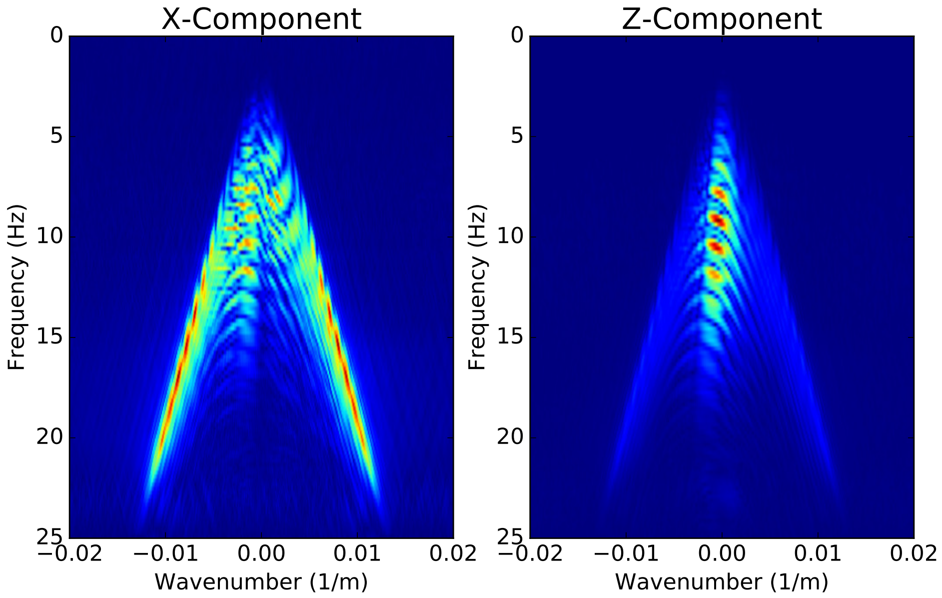


Figure 3.21: FK amplitude spectra for the decimated data interpolated using scalar POCS.

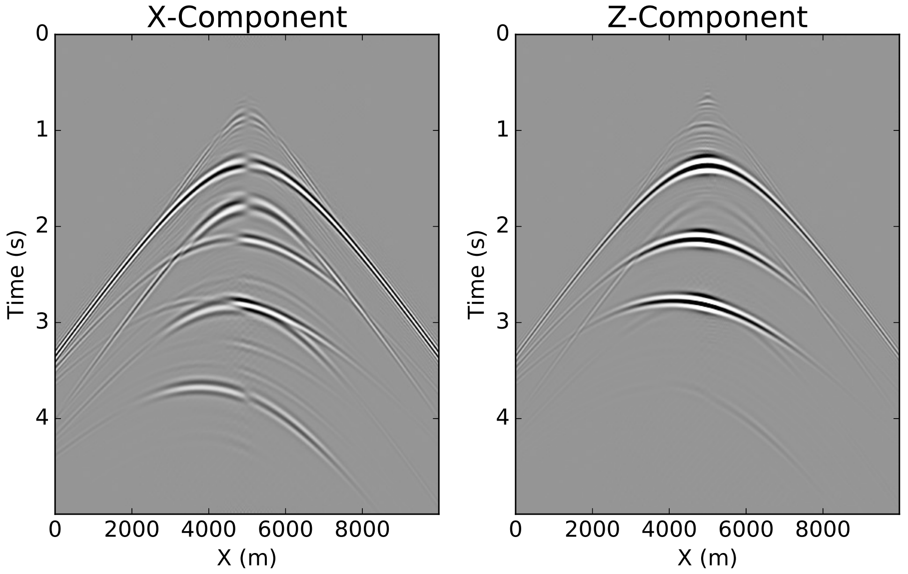


Figure 3.22: decimated data (as shown in figure 3.1) interpolated using scalar MWNI. The qualities of the reconstructions are 27.49 and 32.72 dB respectively.

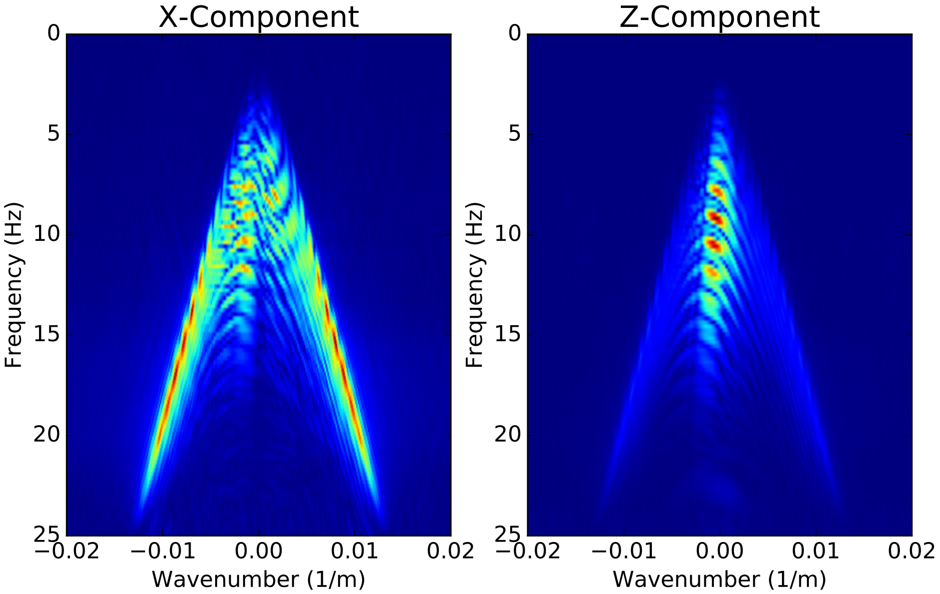


Figure 3.23: FK amplitude spectra for the data interpolated using scalar MWNI. The qualities of the reconstructions are and dB respectively.

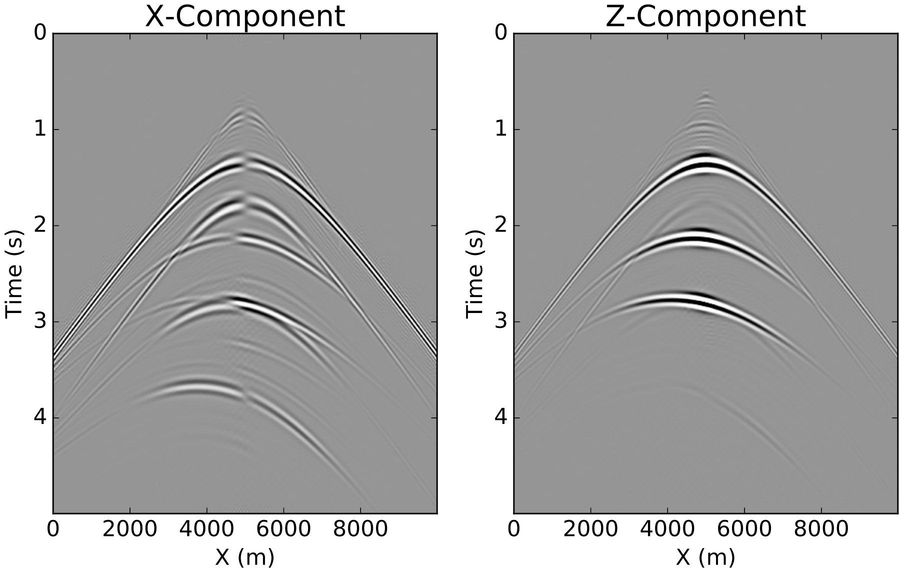


Figure 3.24: decimated data (as shown in figure 3.1) interpolated using vector POCS. The qualities of the reconstructions are 23.76 and 30.81 dB respectively.

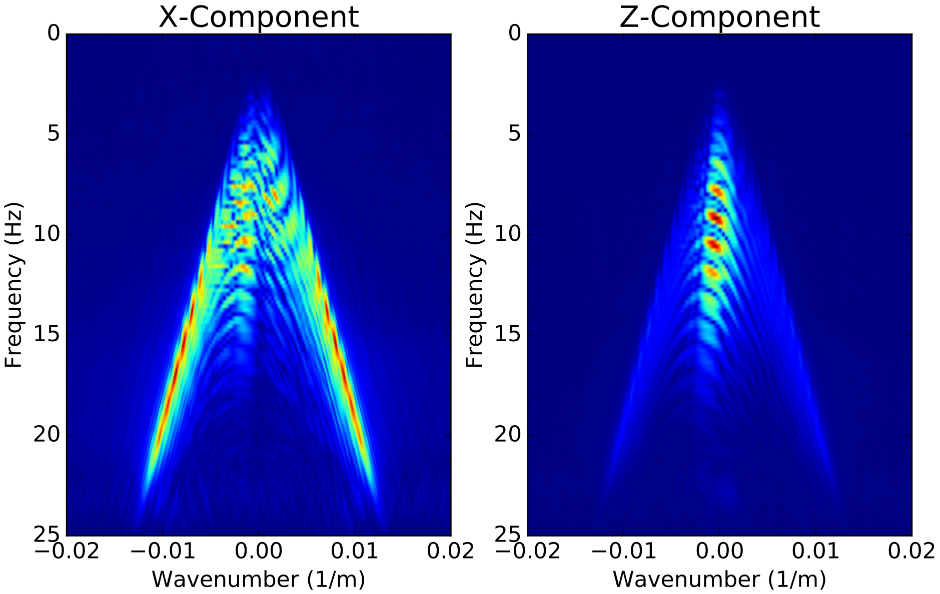


Figure 3.25: FK amplitude spectra for the data interpolated using vector POCS.

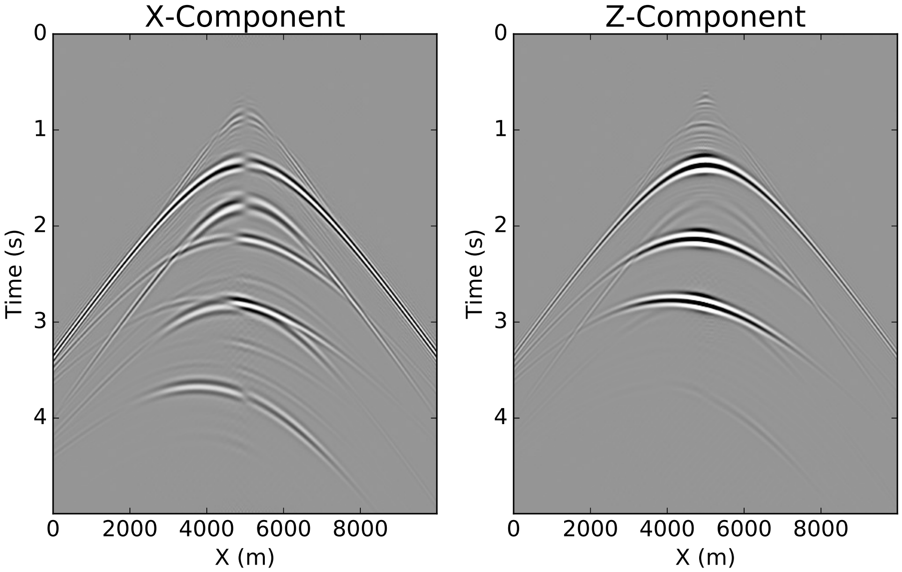


Figure 3.26: decimated data (as shown in figure 3.1) interpolated using vector MWNI. The qualities of the reconstructions are 27.53 and 33.41 dB respectively.

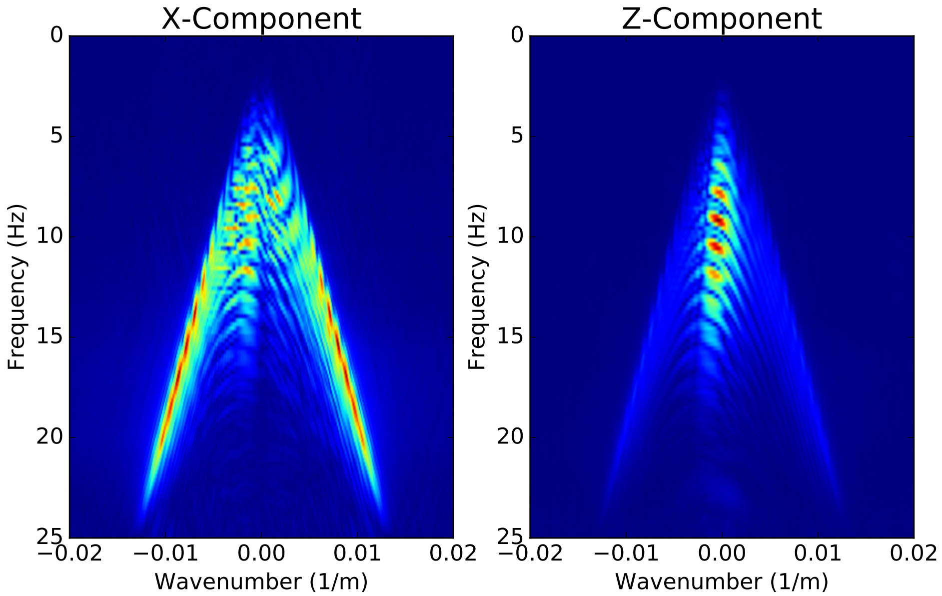


Figure 3.27: FK amplitude spectra for the data interpolated using vector MWNI.

3.6 Conclusions

This chapter introduced a method to reconstruct two components of a multicomponent seismic volume simultaneously using the quaternion Fourier transform and POCS. The algorithm improves the quality of the reconstruction while preserving the orthogonality of the input components. The algorithm can also be used for noise attenuation of seismic volumes. The improvement is due to the fact that one amplitude spectrum is calculated for the two components which is beneficial when the components have significant spectral overlap. A drawback of the method is that if the components do not share significant overlap in the frequency-wavenumber domain the benefit will not be as great. For this reason, we choose to represent the two horizontal components using a vector representation in the frequency domain while treating the Vertical component as a scalar quantity. In the 5D reconstruction of a land dataset displaying the effects of shear wave splitting, we find an improvement in reconstruction quality when treating the radial and transverse components as a vector quantity. By using a symplectic representation of quaternions and choosing an appropriate axis for the Quaternion Fourier Transform (QFT) the formulation for the QFT can be greatly simplified and existing FFT algorithms can be used.

CHAPTER 4

Regularized elastic imaging of multicomponent data[1]

While the previous chapter dealt with the sampling of multicomponent seismic data in surface coordinates, the ultimate goal of reflection seismology is to transform the data into subsurface coordinates, that is, to map reflected energy to its scattering point.

This chapter is organized as follows. First we review the theory of elastic wave propagation and derive the necessary elements for elastic one-way wave propagation as eigensolutions to the Christoffel equation. Next, we integrate these elements into the first order Born approximation, setting up the forward problem; to simulate elastic wavefields from multiparameter images. We then pose elastic imaging as an inverse problem where the ideal model best fits the observed data in a least-squares sense. Finally, we demonstrate the ability of Elastic Least-Squares Migration (ELSM) to improve wavefield separation, mitigate the effects of poor spatial sampling, and compensate for non-uniform illumination.

4.1 Theory

4.1.1 Review of elastic wave propagation

The mechanics necessary for wavefield separation and extrapolation of vector data are provided by the elastic wave equation

$$\rho \ddot{u}_i = C_{ijkl}\partial_j \partial_l u_k, \tag{4.1}$$

---

[1] Parts of this chapter have been published in Stanton and Sacchi (2017), and Stanton and Sacchi (2015).

where u_i is displacement in the direction i, ρ is density and C_{ijkl} is the tensor of elastic moduli (Bale, 2006). In this section Einstein notation is used, implying summation over repeated indices. Displacement can be represented as a complex harmonic (plane wave), $u_k = U_k e^{i\omega(t-s_j x_j)}$, where s_j is the slowness vector that is the inverse of the phase velocity. The phase velocity is related to the slowness via $s_j = n_j/V$, where n_j indicates the direction of wave propagation. The direction of wave propagation is just the normalized wavenumber vector, $n_j = k_j/k$, where $k = \sqrt{k_1^2 + k_2^2 + k_3^2}$. Substituting a plane wave into the wave equation we obtain the Christoffel equation

$$\rho U_i = C_{ijkl} s_j s_l U_k, \tag{4.2}$$

more commonly written in the form

$$\begin{bmatrix} G_{11} - \rho V^2 & G_{12} & G_{13} \\ G_{12} & G_{22} - \rho V^2 & G_{23} \\ G_{13} & G_{23} & G_{33} - \rho V^2 \end{bmatrix} \begin{bmatrix} U_1 \\ U_2 \\ U_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \tag{4.3}$$

where the 3x3 Christoffel matrix is defined by

$$G_{ij} = C_{ijkl} n_j n_i \tag{4.4}$$

The Christoffel equation is a 3 x 3 eigenvalue - eigenvector problem. Given a tensor of elastic moduli the eigenvalues correspond to the phase velocities of three distinct wave modes, and the eigenvectors correspond to the polarizations of these modes (Tsvankin, 2012). In the context of elastic migration, the eigenvectors provide the polarization information necessary for wavefield separation, and the eigenvalues provide the phase velocity necessary for extrapolation. Assuming isotropy and substituting into equation 4.3 considering propagation in the $[x_1, x_3]$ plane we obtain the following Christoffel equation

$$\begin{bmatrix} (\lambda + 2\mu) n_1^2 + \mu n_3^2 - \rho V^2 & 0 & (\lambda + \mu) n_1 n_3 \\ 0 & \mu(n_1^2 + n_3^2) - \rho V^2 & 0 \\ (\lambda + \mu) n_1 n_3 & 0 & \mu n_1^2 + (\lambda + 2\mu) n_3^2 - \rho V^2 \end{bmatrix} \begin{bmatrix} U_1 \\ U_2 \\ U_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \tag{4.5}$$

where λ and μ are Lamé's first and second parameters respectively. To find the eigenvalues corresponding to the three wave modes we set the determinant equal to zero and substitute $\boldsymbol{n} = \begin{bmatrix} 0 & 0 & 1 \end{bmatrix}$ (i.e. assuming a plane wave traveling vertically downward) leading to the cubic equation

$$0 = (\mu - \rho V^2)(\mu - \rho V^2)(\lambda + 2\mu - \rho V^2) \tag{4.6}$$

with solutions $V_1 = \sqrt{(\lambda + 2\mu)/\rho}$, $V_2 = \sqrt{\mu/\rho}$, and $V_3 = \sqrt{\mu/\rho}$ which are the familiar relations for P, SV, and SH phase velocities respectively. Solving for the corresponding eigenvectors we obtain

$$\boldsymbol{U}^P = \begin{bmatrix} \hat{n_1} \\ \hat{n_2} \\ \hat{n_3} \end{bmatrix} \quad , \quad \boldsymbol{U}^{S1} = \begin{bmatrix} -\hat{n}_2/a \\ \hat{n}_1/a \\ 0 \end{bmatrix}, \quad \text{and} \quad \boldsymbol{U}^{S2} = \begin{bmatrix} -\hat{n}_1\hat{n}_3/a \\ -\hat{n}_2\hat{n}_3/a \\ a \end{bmatrix}, \tag{4.7}$$

where $a^2 = \hat{n_1}^2 + \hat{n_2}^2$.

While we simplify our calculations greatly by assuming isotropy, lower symmetry material parameters can also be incorporated into the Christoffel equation. For example, Bale (2006) solves the Christoffel equation assuming transverse isotropy.

Considering one-way wave propagation in laterally homogeneous media, we can propagate elastic wavefields vertically by wavefield decomposition, a vertical phase shift, and wavefield composition. This is represented by

$$\vec{\boldsymbol{u}}(\omega, x, y, z + \Delta z) = \boldsymbol{P}\vec{\boldsymbol{u}}(\omega, x, y, z), \tag{4.8}$$

where

$$\boldsymbol{P} = \boldsymbol{Q}\boldsymbol{R}\boldsymbol{Q}^{-1}. \tag{4.9}$$

In laterally homogeneous media the operator $\boldsymbol{R}$ corresponds to a vertical phase shift (Gazdag, 1978) applied to each wavefield potential

$$\boldsymbol{R} = \begin{bmatrix} \boldsymbol{\mathcal{F}}_{\boldsymbol{x}}^{-1} & 0 & 0 \\ 0 & \boldsymbol{\mathcal{F}}_{\boldsymbol{x}}^{-1} & 0 \\ 0 & 0 & \boldsymbol{\mathcal{F}}_{\boldsymbol{x}}^{-1} \end{bmatrix} \begin{bmatrix} e^{i\omega k_z(v_p)\Delta z} & 0 & 0 \\ 0 & e^{i\omega k_z(v_s)\Delta z} & 0 \\ 0 & 0 & e^{i\omega k_z(v_s)\Delta z} \end{bmatrix} \begin{bmatrix} \boldsymbol{\mathcal{F}}_{\boldsymbol{x}} & 0 & 0 \\ 0 & \boldsymbol{\mathcal{F}}_{\boldsymbol{x}} & 0 \\ 0 & 0 & \boldsymbol{\mathcal{F}}_{\boldsymbol{x}} \end{bmatrix}, \tag{4.10}$$

and the wavefield composition operator is given by

$$\boldsymbol{Q} = \begin{bmatrix} \boldsymbol{\mathcal{F}}_{\boldsymbol{x}}^{-1} & 0 & 0 \\ 0 & \boldsymbol{\mathcal{F}}_{\boldsymbol{x}}^{-1} & 0 \\ 0 & 0 & \boldsymbol{\mathcal{F}}_{\boldsymbol{x}}^{-1} \end{bmatrix} \begin{bmatrix} \boldsymbol{U}^P & \boldsymbol{U}^{S1} & \boldsymbol{U}^{S2} \end{bmatrix} \begin{bmatrix} \boldsymbol{\mathcal{F}}_{\boldsymbol{x}} & 0 & 0 \\ 0 & \boldsymbol{\mathcal{F}}_{\boldsymbol{x}} & 0 \\ 0 & 0 & \boldsymbol{\mathcal{F}}_{\boldsymbol{x}} \end{bmatrix}, \tag{4.11}$$

where $\boldsymbol{\mathcal{F}_x}$ and $\boldsymbol{\mathcal{F}_x^{-1}}$ represent forward and inverse Fourier transforms over spatial axes. For isotropic media the composition operator becomes

$$\boldsymbol{Q} = \begin{bmatrix} \boldsymbol{\mathcal{F}_x^{-1}} & 0 & 0 \\ 0 & \boldsymbol{\mathcal{F}_x^{-1}} & 0 \\ 0 & 0 & \boldsymbol{\mathcal{F}_x^{-1}} \end{bmatrix} \begin{bmatrix} \hat{n}_1 & -\hat{n}_2/a & -\hat{n}_1\hat{n}_3/a \\ \hat{n}_2 & \hat{n}_1/a & -\hat{n}_2\hat{n}_3/a \\ \hat{n}_3 & 0 & a \end{bmatrix} \begin{bmatrix} \boldsymbol{\mathcal{F}_x} & 0 & 0 \\ 0 & \boldsymbol{\mathcal{F}_x} & 0 \\ 0 & 0 & \boldsymbol{\mathcal{F}_x} \end{bmatrix} \tag{4.12}$$

with wavefield decomposition given by

$$\boldsymbol{Q}^{-1} = \begin{bmatrix} \boldsymbol{\mathcal{F}_x^{-1}} & 0 & 0 \\ 0 & \boldsymbol{\mathcal{F}_x^{-1}} & 0 \\ 0 & 0 & \boldsymbol{\mathcal{F}_x^{-1}} \end{bmatrix} \begin{bmatrix} \hat{n}_1 & \hat{n}_2 & \hat{n}_3 \\ -\hat{n}_2/a & \hat{n}_1/a & 0 \\ -\hat{n}_1\hat{n}_3/a & -\hat{n}_2\hat{n}_3/a & a \end{bmatrix} \begin{bmatrix} \boldsymbol{\mathcal{F}_x} & 0 & 0 \\ 0 & \boldsymbol{\mathcal{F}_x} & 0 \\ 0 & 0 & \boldsymbol{\mathcal{F}_x} \end{bmatrix}, \tag{4.13}$$

where $\hat{n}_1$, $\hat{n}_1$, and $\hat{n}_1$ can be replaced by the normalized wavenumbers $\hat{k}_1$, $\hat{k}_2$, and $\hat{k}_3$.

While theoretically, the wavefield combination operator is orthonormal, numerically the computation is complicated by the fact that the vertical wavenumber must be computed from the material parameters corresponding to each wave mode and the horizontal wavenumbers via the dispersion relation. This issue is further discussed in appendix A.

To demonstrate the action of the composition ($\boldsymbol{Q}$) and decomposition ($\boldsymbol{Q}^{-1}$) operators we consider the decomposition and recomposition of an elastic wavefield in a homogeneous, isotropic 2D medium as shown in figure 4.1. The model consists of a P-wave velocity of 1700m/s and an S-wave velocity of 981m/s. Here X and Z components are shown for a time slice at 0.35s after a displacement source was fired in the center of the model. In figure 4.2 the separated P and S wavefields are shown after application of the decomposition operator, while the recomposed wavefields are shown in figure 4.3.

It is interesting to compare the operator $\boldsymbol{Q}^{-1}$ with the Helmholtz decomposition operator which separates an elastic wavefield into curl free ($\frac{\partial u_1}{\partial x_1} + \frac{\partial u_2}{\partial x_2} + \frac{\partial u_3}{\partial x_3}$) and divergence free components $((\frac{\partial u_3}{\partial x_2} - \frac{\partial u_2}{\partial x_3})i + (\frac{\partial u_1}{\partial x_3} - \frac{\partial u_3}{\partial x_1})j + (\frac{\partial u_2}{\partial x_1} - \frac{\partial u_1}{\partial x_2})k)$. Writing this as a linear operator we have

$$\boldsymbol{H}^{-1} = \boldsymbol{\mathcal{F}_x^{-1}} \begin{bmatrix} ik_1 & ik_2 & ik_3 \\ 0 & -ik_3 & ik_2 \\ ik_3 & 0 & -ik_1 \\ -ik_2 & ik_1 & 0 \end{bmatrix} \boldsymbol{\mathcal{F}_x}, \tag{4.14}$$

which (aside from being a rectangular matrix) is somewhat similar to an unnormalized version of the decomposition operator derived from the Christoffel equation. This normalization

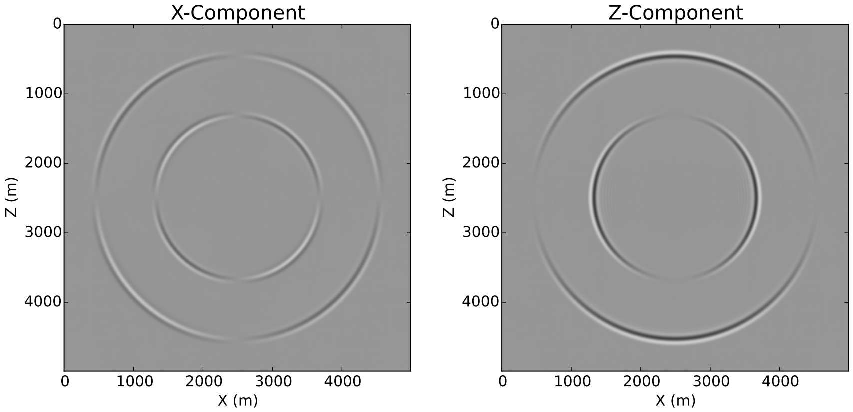


Figure 4.1: X and Z input data components for an elastic wave propagating in a homogeneous isotropic medium. The records were generated using elastic finite difference modeling.

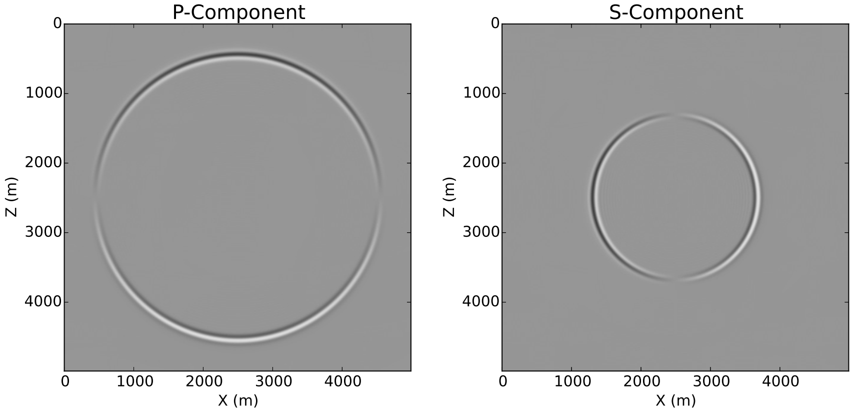


Figure 4.2: Decomposed P and S components after applying the wavefield decomposition operator $\boldsymbol{Q}^{-1}$ to the data components U^x and U^y.

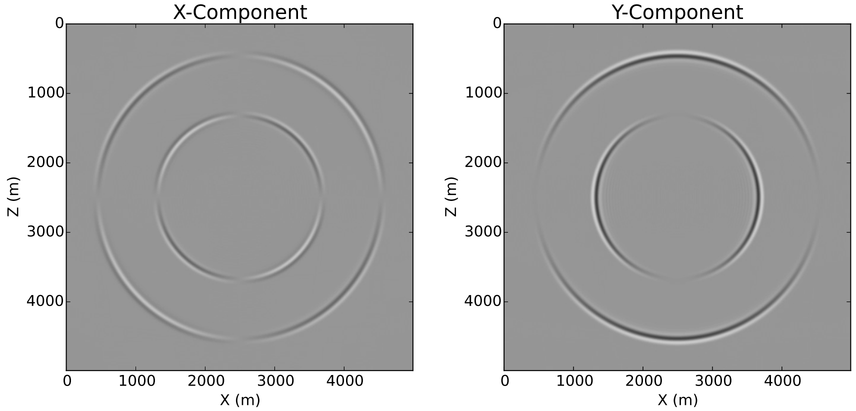


Figure 4.3: Recomposed X and Y components after applying the wavefield recomposition operator $\boldsymbol{Q}$ to the wavefield components U^p and U^s.

is important for stability when we consider the recursive application of the operator and its inverse during propagation. Sun et al. (2011) examine the changes in amplitude introduced by separating elastic wavefields via Helmholtz decomposition. This amplitude change is a significant problem because it alters the amplitude ratio between PP and PS images which can greatly affect any subsequent interpretation. The fact that the Christoffel formulation results in a separation operator with columns that form an orthonormal basis means that the decomposition and recomposition steps are unitary operations that leave the relative amplitudes between wave modes unaltered. Another important distinction is that for three component data in three spatial dimensions Helmholtz decomposition provides a single component curl free wavefield and a three component divergence free wavefield, whereas the Christoffel based formulation provides three mutually orthogonal scalar wavefields, even for anisotropic media. This property makes the imaging condition straightforward to design– we can correlate every combination of source and receiver side wave-mode to form a 9 component image or, as is done in this study, choose to only correlate the source side P-wavefield with the receiver side P and S-wavefields to form a 2 component image.

4.1.2 Elastic shot-profile modeling and migration

We incorporate the wavefield decomposition, extrapolation and recomposition operators derived in the previous section to model elastic wavefields. Our approach closely follows the derivation of forward and adjoint one-way wave equation operators for acoustic least-squares migration given by Rickett (2003) and Kaplan et al. (2010). Here we extend shot-profile modeling and migration to the elastic case. We begin with the first order Born approximation

$$\vec{\boldsymbol{u}}(\boldsymbol{x}^g,\omega) \sim \omega^2 \int_{-\infty}^{\infty} G_0(\boldsymbol{x}^g|\boldsymbol{x})\boldsymbol{M}(\boldsymbol{x})G_0(\boldsymbol{x}|\boldsymbol{x}^s)d\boldsymbol{x}, \tag{4.15}$$

where $\boldsymbol{x}$ indicates all subsurface positions, $\boldsymbol{x}^g$ indicates all receiver positions, and $\boldsymbol{x}^s$ indicates the position of a source. Here $G_0(\boldsymbol{x}|\boldsymbol{x}^s)$ is the 3x1 Green's function response at all subsurface positions due to an impulse at the position of the source (representing all modes of incident wavefield), $\boldsymbol{M}(\boldsymbol{x})$ is a 3x3 matrix of scattering potentials, and $G_0(\boldsymbol{x}^g|\boldsymbol{x})$ is the 3x3 Green's function response at the receiver positions due to an impulse generated at all subsurface positions (scattering all combinations of incident and scattered wave modes). The recorded data components at all receiver locations for a given frequency, ω, are denoted by $\vec{\boldsymbol{u}}(\boldsymbol{x}^g,\omega)$. A complete description of the perturbations in physical properties that comprise the 3x3 scattering potential $\boldsymbol{M}(\boldsymbol{x})$ is provided by Stolt and Weglein (2012). Practically equation 4.15 implies downward continuation of the source wavefield, multiplication with the reflectivity, followed by upward continuation of the scattered wavefield to the receiver datum.

It is perhaps more helpful to represent equation 4.15 in terms of a series of discrete linear operations that can be readily programmed in a computer. To derive the forward operator for a single source and single frequency we begin by downward continuing the source wavefield by

$$\begin{bmatrix} \vec{\boldsymbol{\psi}}_1 \\ \vec{\boldsymbol{\psi}}_2 \\ \vec{\boldsymbol{\psi}}_3 \\ ... \\ \vec{\boldsymbol{\psi}}_{N_z} \end{bmatrix} = \begin{bmatrix} \boldsymbol{Q}_1^{-1} & 0 & 0 & ... & 0 \\ 0 & \boldsymbol{Q}_2^{-1}\boldsymbol{P}_1 & 0 & ... & 0 \\ 0 & 0 & \boldsymbol{Q}_3^{-1}\boldsymbol{P}_2\boldsymbol{P}_1 & ... & 0 \\ ... & ... & ... & ... & ... \\ 0 & 0 & 0 & ... & \boldsymbol{Q}_{N_z}^{-1}\boldsymbol{P}_{N_z-1}...\boldsymbol{P}_1 \end{bmatrix} \begin{bmatrix} \vec{\boldsymbol{u}}_1^s \\ \vec{\boldsymbol{u}}_1^s \\ \vec{\boldsymbol{u}}_1^s \\ ... \\ \vec{\boldsymbol{u}}_1^s \end{bmatrix}, \tag{4.16}$$

where $\boldsymbol{P}_i = \boldsymbol{Q}_i\boldsymbol{R}_i\boldsymbol{Q}_i^{-1}$ represents elastic propagation using the medium properties of layer i, $\vec{\boldsymbol{u}}_1^s$ represents one frequency of the vector-valued source wavelet at the surface of the earth, and $\vec{\boldsymbol{\psi}}_i = [\boldsymbol{\psi}_i^p \ \boldsymbol{\psi}_i^{s_1} \ \boldsymbol{\psi}_i^{s_2}]^T$ are the source-side wavefield potentials for a single frequency at depth interval i of the model. To obtain data at the surface of the earth we propagate the scattered wavefield upward by

$$\vec{\boldsymbol{u}}_1^g = \begin{bmatrix} \boldsymbol{Q}_1 & \boldsymbol{P}_1^{-1}\boldsymbol{Q}_2 & ... & \boldsymbol{P}_1^{-1}...\boldsymbol{P}_{N_z-1}^{-1}\boldsymbol{Q}_{N_z} \end{bmatrix} \begin{bmatrix} \boldsymbol{M}_1\vec{\psi}_1 \\ \boldsymbol{M}_2\vec{\psi}_2 \\ ... \\ \boldsymbol{M}_{N_z}\vec{\psi}_{N_z} \end{bmatrix} \tag{4.17}$$

where

$$\boldsymbol{M}_i = \begin{bmatrix} \boldsymbol{m}_i^{pp} & \boldsymbol{m}_i^{s_1 p} & \boldsymbol{m}_i^{s_2 p} \\ \boldsymbol{m}_i^{ps_1} & \boldsymbol{m}_i^{s_1 s_1} & \boldsymbol{m}_i^{s_2 s_1} \\ \boldsymbol{m}_i^{ps_2} & \boldsymbol{m}_i^{s_1 s_2} & \boldsymbol{m}_i^{s_2 s_2} \end{bmatrix} \tag{4.18}$$

are the multi-parameter image amplitudes at depth interval i of the model. For the adjoint operator, we begin by propagating the receiver wavefield (data) into the earth by

$$\begin{bmatrix} \vec{\boldsymbol{\phi}}_1 \\ \vec{\boldsymbol{\phi}}_2 \\ \ldots \\ \vec{\boldsymbol{\phi}}_{N_z} \end{bmatrix} = \begin{bmatrix} \boldsymbol{Q}_1^\dagger & 0 & 0 & 0 \\ 0 & \boldsymbol{Q}_2^\dagger(\boldsymbol{P}_1^{-1})^\dagger & 0 & 0 \\ \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & 0 & \boldsymbol{Q}_{N_z}^\dagger(\boldsymbol{P}_{N_z-1}^{-1})^\dagger \ldots (\boldsymbol{P}_1^{-1})^\dagger \end{bmatrix} \begin{bmatrix} \vec{\boldsymbol{u}}_1^g \\ \vec{\boldsymbol{u}}_1^g \\ \vec{\boldsymbol{u}}_1^g \\ \vec{\boldsymbol{u}}_1^g \end{bmatrix}, \tag{4.19}$$

where $\vec{\boldsymbol{\phi}}_i = [\boldsymbol{\phi}_i^p \ \boldsymbol{\phi}_i^{s_1} \ \boldsymbol{\phi}_i^{s_2}]^T$ are the receiver-side wavefield potentials for a single frequency at depth interval i of the model. We then correlate using the outer product of the source and receiver wavefields

$$\begin{bmatrix} \tilde{\boldsymbol{M}}_1 \\ \tilde{\boldsymbol{M}}_2 \\ \ldots \\ \tilde{\boldsymbol{M}}_{N_z} \end{bmatrix} = \begin{bmatrix} \vec{\boldsymbol{\psi}}_1^* & 0 & 0 & 0 \\ 0 & \vec{\boldsymbol{\psi}}_2^* & 0 & 0 \\ \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & 0 & \vec{\boldsymbol{\psi}}_{N_z}^* \end{bmatrix} \begin{bmatrix} \vec{\boldsymbol{\phi}}_1^T \\ \vec{\boldsymbol{\phi}}_2^T \\ \ldots \\ \vec{\boldsymbol{\phi}}_{N_z}^T \end{bmatrix} \tag{4.20}$$

to obtain a multiparameter image $\tilde{\boldsymbol{m}}$.

Figures 4.4 and 4.5 outline the workflow for acoustic modeling and migration, while figures 4.6 and 4.7 show the extension to the elastic case outlined above.

To avoid confusing notation equations 4.16 to 4.20 only consider vertical propagation in a 1D earth. Extending these equations in the X and Y dimensions is achieved by extending each sub-vector by a factor of N_{xy}, and block diagonalizing each sub-matrix increasing their rank by a factor of N_{xy}. Furthermore, to extend the equations to the case of multiple frequencies and multiple shots the forward operator begins by spraying the reflectivities to N_ω frequencies and N_{shot} shots, while the adjoint ends by summing over all frequencies and shots (Rickett, 2003).

To extend the extrapolation operator, $\boldsymbol{R}$, to the case of laterally inhomogeneous velocity we incorporate the Phase Shift Plus Interpolation (PSPI) and Split Step corrections (Gazdag and Sguazzero, 1984; Stoffa et al., 1990; Xie and Wu, 2005). An important consideration for one-way extrapolation is the accuracy of the operator versus propagation angle. Biondi (2002) provides an analysis of the phase error for several pseudo-differential operators as a function of propagation angle. Notably, the Split Step correction is accurate (< 0.2 radians

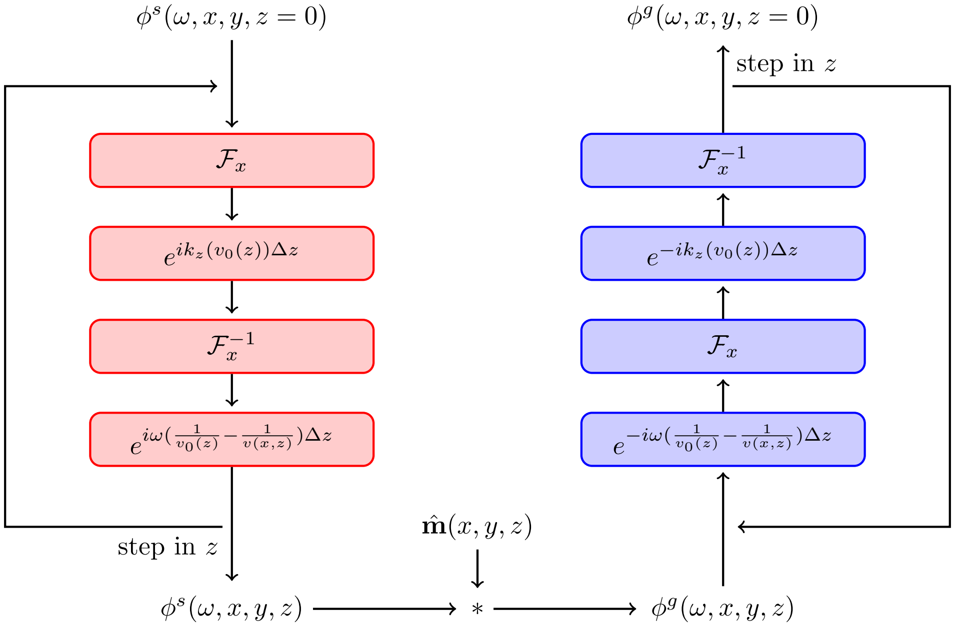


Figure 4.4: Born scattering of acoustic wavefields.

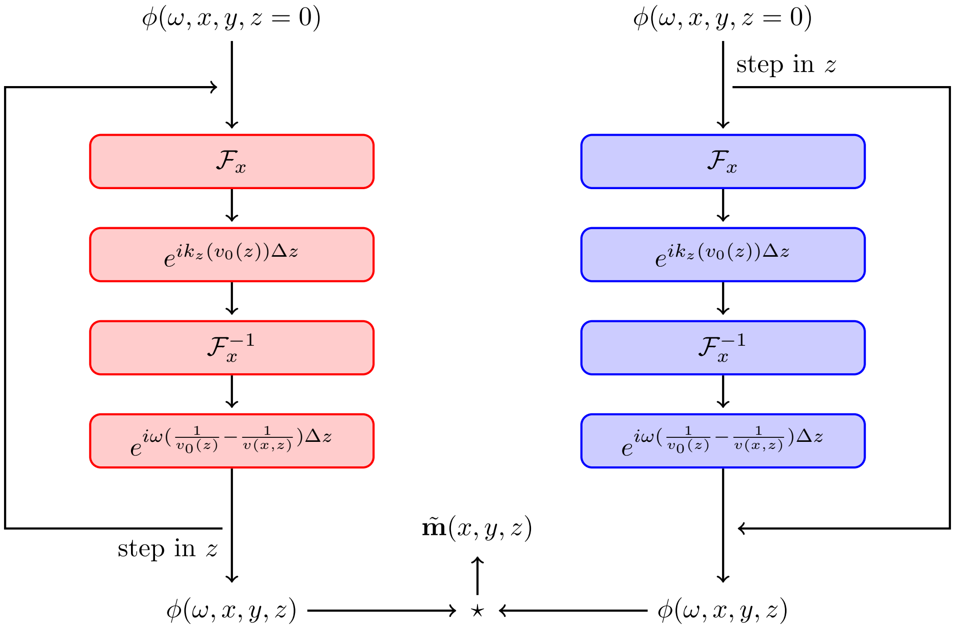


Figure 4.5: Migration of acoustic wavefields.

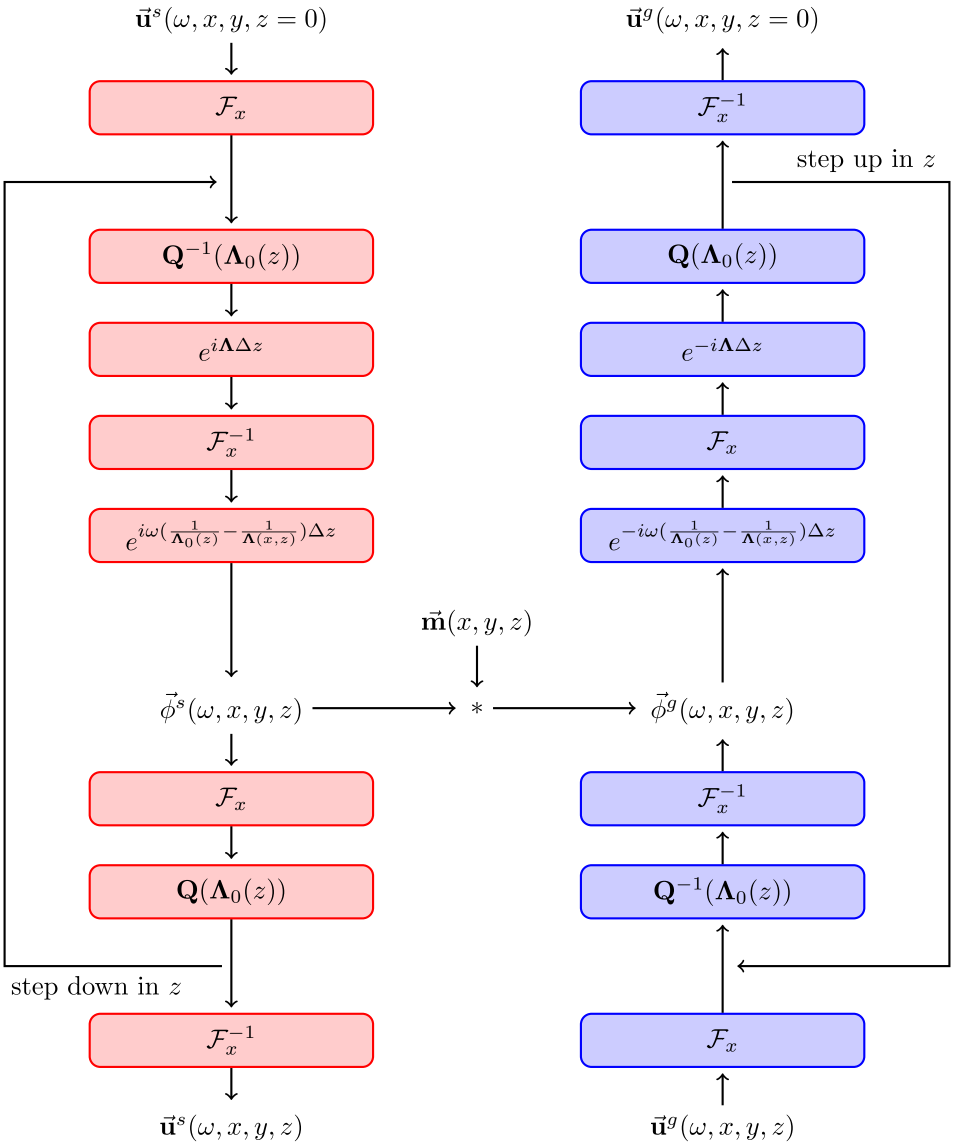


Figure 4.6: Born scattering of elastic wavefields.

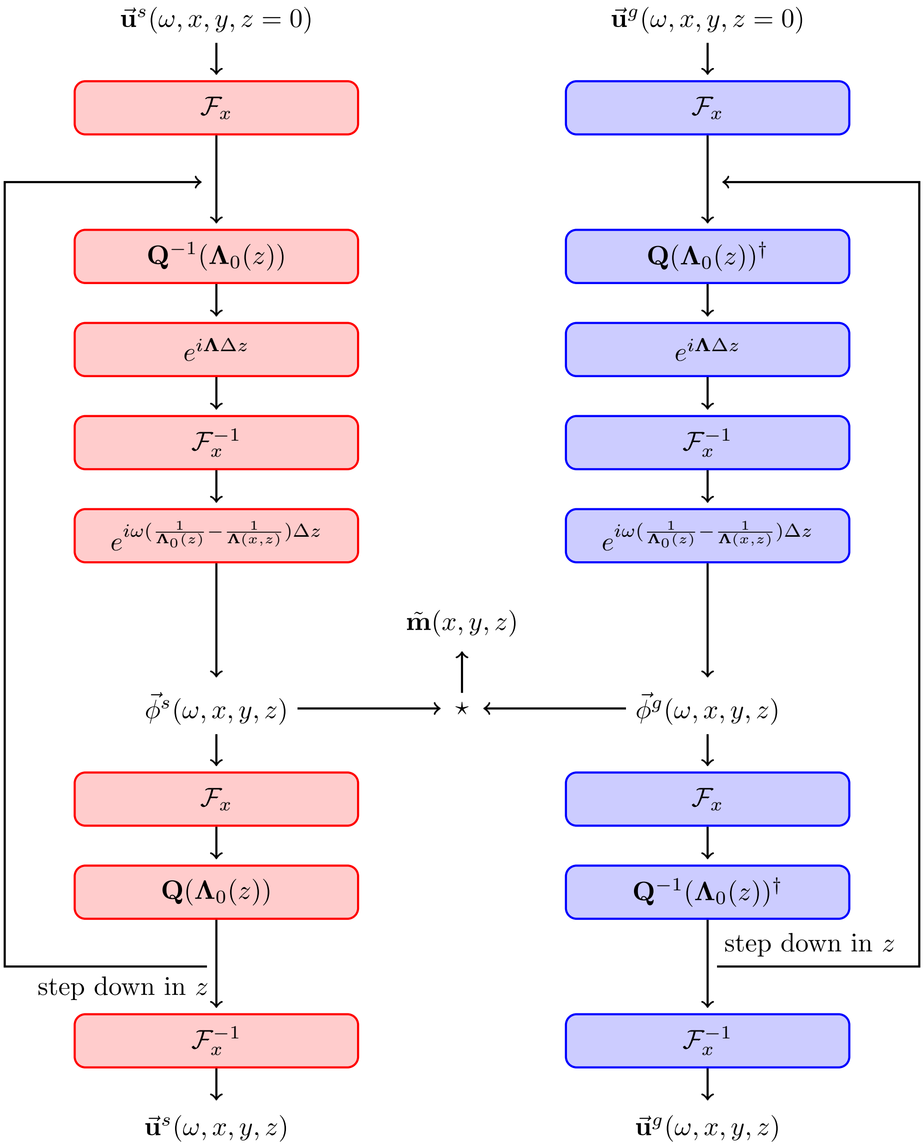


Figure 4.7: Migration of elastic wavefields.

phase error) to approximately 45 degrees for a reference velocity with an error of 200m/s. By combining the Split Step and PSPI corrections we are able to narrow the range of reference velocity error significantly. The examples in this chapter use 5 reference velocities. A detailed derivation of forward and adjoint acoustic shot profile split-step migration operators is provided in Kaplan et al. (2010).

To apply the separation operator, $\boldsymbol{Q}^{-1}$, in laterally inhomogeneous media we use the average P and S-wave velocity for a given layer. The accuracy of the separation could be further improved via a spatial windowing approach following Bale and Margrave (2004).

It is important to note that in this work we do not consider the effect of the free surface. In our numerical examples, we use data generated using elastic finite differences with an absorbing boundary condition above the recording surface. Effectively, this assumes that the recorded wavefield consists only of upgoing energy. To compensate for the effects of upgoing and downgoing energy at the recording surface a projection onto the downgoing wavefield must be made as an initial step (Bale, 2006). A further complication in least-squares migration is the estimation of the source wavelet. In our numerical examples, we consider the case of a known source wavelet. A method to estimate an unknown source wavelet from the data can be found in Zhang et al. (2016).

In the following discussion of the least-squares formulation we refer to the forward and adjoint operations more simply as $\boldsymbol{u} = \boldsymbol{L}\boldsymbol{m}$ and $\tilde{\boldsymbol{m}} = \boldsymbol{L}^{\dagger}\boldsymbol{u}$, respectively, where it is understood that vector $\boldsymbol{m}$ consists of sub-matrices, and vector $\boldsymbol{u}$ consists of sub-vectors.

Finally, to model the complex amplitude behavior of elastic data, including polarity reversal of PS images at an incidence angle of zero, we use an extended imaging condition. Of the numerous methods to compute extended images we selected a mapping from subsurface offset image gathers to opening angle image gathers via the radial trace transform as outlined in Rickett and Sava (2002) as we find it provides smooth, artefact free angle gathers. Appendix C details the method used to calculate propagation angles. The subsurface offset method parameterizes the images as a function of the half opening angle between the source and receiver wavefields. For PP reflections the equivalence of source and receiver side velocity means the opening angle is bisected by the reflector normal vector, making the half opening angle equivalent to the source side incidence angle with respect to reflector normal. For PS reflections there is no such equivalence, and the half opening angle does not adequately describe the zero incidence angle where polarity reversal is expected (except for the special case of a flat-lying reflector). While a post-processing step can be used to correctly position the polarity reversal for PS reflections prior to stacking (Rosales et al., 2007), we observe that simply using the half opening angle provides an adequate extended imaging condition for elastic least-squares migration.

4.1.3 Implementation

While our derivation of forward and adjoint operators are written in such a way to accommodate 3D, anisotropic wave propagation, we make some simplifying assumptions in our implementation. These simplifications are to avoid unnecessary complications that deserve more detailed and focused explanation in future research, and to make our numerical examples efficient and clear. First, we assume isotropic symmetry of material parameters and only consider wave propagation in a 2D plane. These simplifications imply that only X and Z data components are required, and only the P and S2 elastic wave modes need be considered. Isotropy also implies that within the propagator the vertical phase shift is independent of the propagation angle and azimuth. Finally, we assume a purely compressional source-side wavefield, which implies that we record only PP and PS reflections in the recorded wavefield. This is achieved using an acoustic propagation code to simulate the source side wavefield. Taking these simplifying assumptions into consideration our multiparameter image given in equation 4.18 simplifies to:

$$\boldsymbol{M}_i = \begin{bmatrix} \boldsymbol{m}_i^{pp} & 0 & 0 \\ 0 & 0 & 0 \\ \boldsymbol{m}_i^{ps_2} & 0 & 0 \end{bmatrix}. \tag{4.21}$$

While this final assumption effectively makes the code focused on converted wave exploration, there could be some benefit to including additional wave modes, even in the case of converted wave acquisition, where the source side wavefield is generally assumed to be largely free of useful shear wave energy. In fact, research suggests that vertical vibrator generated land seismic data contains a large amount of useful shear wave energy[2] that could be used to generate SS, or SP images (R. Stewart, personal communication, April 7, 2017). An additional simplification is that we do not take attenuation into consideration. Incorporating attenuation into acoustic least squares migration was studied by Dutta and Schuster (2014b); making use of the visco-acoustic wave equation. A valuable future contribution would be to extend this formulation to make use of the visco-elastic wave equation.

4.1.4 Least-squares formulation

We pose elastic imaging as an inverse problem with the objective function

$$J = ||\boldsymbol{T}(\boldsymbol{L}\boldsymbol{m} - \boldsymbol{u})||_2^2 + ||\boldsymbol{\mu}^T \boldsymbol{m}||_2^2 \tag{4.22}$$

---

[2] In fact Miller and Pursey (1955) suggest that the *majority* of elastic wave energy propagating away from a vertical point force should be in the form of shear waves.

where $\boldsymbol{T}$ is a diagonal sampling operator with 1's in place of observations and 0's in place of missing traces, $\boldsymbol{m}$ represents the multicomponent image, $\boldsymbol{u}$ represents the multicomponent data, and $\boldsymbol{\mu}$ is a multicomponent trade-off parameter to control the level to fit the data versus the level to minimize the $l-2$ norm of the model. The components of $\boldsymbol{\mu}$ can be fixed for all components of the image (effectively treating $\boldsymbol{\mu}$ as a scalar), or can be tailored to individually constrain components.

Since wavefield cross-talk artefacts are migrated at an incorrect velocity, they will appear as dipping energy in angle gathers. For this reason, we can regularize the inversion with an operator that weights up this energy in each angle gather

$$J = ||\boldsymbol{T}(\boldsymbol{Lm} - \boldsymbol{u})||_2^2 + ||\boldsymbol{\mu}^T \boldsymbol{Dm}||_2^2, \tag{4.23}$$

or, by a change of variables $\boldsymbol{z} = \boldsymbol{Dm}$ we write

$$J = ||\boldsymbol{T}(\boldsymbol{LSz} - \boldsymbol{u})||_2^2 + ||\boldsymbol{\mu}^T \boldsymbol{z}||_2^2 \tag{4.24}$$

such that $\boldsymbol{S}$ ($\approx \boldsymbol{D}^{-1}$) is a lateral smoother that removes strongly dipping energy in the angle gathers (equation 4.24 is an approximate form of equation 4.23). We minimize equation 4.24 using CG with forward operator $\boldsymbol{TLS}$ and adjoint operator $\boldsymbol{S}^\dagger \boldsymbol{L}^\dagger \boldsymbol{T}$. After the final iteration we substitute $\boldsymbol{m} = \boldsymbol{Sz}$.

To solve equation 4.24 it is necessary to use a vectorized modified form of Conjugate Gradients that uses vectorized inner products. This requirement for a specialized solver is due to the fact that the linear operator $\mathbf{L}$ in equation 4.24 mixes model space vectors to form data space vectors. If $\mathbf{L}$ were to map independent images to independent data components then equation 4.24 could be solved equivalently via

$$\begin{aligned} J &= ||\mathbf{T}(\mathbf{LSz} - \mathbf{u})||_2^2 + ||\boldsymbol{\mu}^T \mathbf{z}||_2^2 && (4.25) \\ &= \left\| \begin{array}{c} ||\mathbf{T}(\mathbf{LS}z_{pp} - u_1)||_2^2 + \mu_1 ||z_{pp}||_2^2 \\ ||\mathbf{T}(\mathbf{LS}z_{ps_1} - u_2)||_2^2 + \mu_2 ||z_{ps_1}||_2^2 \\ ||\mathbf{T}(\mathbf{LS}z_{ps_2} - u_3)||_2^2 + \mu_3 ||z_{ps_2}||_2^2 \end{array} \right\|_2^2, && (4.26) \end{aligned}$$

which would be equivalent to solving three independent sub-problems. Since this is not the case we must use a vectorized form of Conjugate Gradients. Appendix B discusses the application of the Conjugate Gradients algorithm to multicomponent-valued model and data vectors.

Equation 4.24 poses imaging as an inverse problem from the perspective of data-fitting. While this is an intuitive approach, there are other ways to set up the problem. Considering the fact that the migrated image is generated by applying the adjoint operator to the data, $\boldsymbol{L}^T\boldsymbol{d}$, equation 4.22 may be rewritten as

$$J = ||(\boldsymbol{L}^T\boldsymbol{T}\boldsymbol{L}\boldsymbol{m} - \boldsymbol{L}^T\boldsymbol{u})||_2^2 + ||\boldsymbol{\mu}^T\boldsymbol{m}||_2^2 \tag{4.27}$$

or as

$$J = ||(\boldsymbol{H}\boldsymbol{m} - \boldsymbol{L}^T\boldsymbol{u})||_2^2 + ||\boldsymbol{\mu}^T\boldsymbol{m}||_2^2 \tag{4.28}$$

where $\boldsymbol{H} = \boldsymbol{L}^T\boldsymbol{T}\boldsymbol{L}$ is the Hessian, or blurring kernel. Viewed this way, least-squares migration aims to deconvolve the Hessian from the migrated image to recover the true, unblurred image. Hu and Schuster (1998) call this process migration deconvolution, while Fletcher et al. (2015) refers to model domain vs. data domain least-squares migration. Equation 4.28 is solved using Conjugate Gradients, where the migrated image, $\boldsymbol{L}^T\boldsymbol{m}$, is substituted for the input data, and the Hessian, $\boldsymbol{H} = \boldsymbol{L}^T\boldsymbol{T}\boldsymbol{L}$ and its adjoint, $\boldsymbol{H}^T = \boldsymbol{L}\boldsymbol{T}\boldsymbol{L}^T$, are substituted for the modeling and migration operators respectively. While this increases the cost of a single iteration of Conjugate Gradients, the number of iterations necessary to converge to a good solution can often be reduced by this formulation (Fletcher et al., 2015). This approach opens the possibility of different approximate methods to least-squares migration. Rickett (2003) approximate smoothed illumination compensation scalars by taking the ratio of RMS amplitudes of windows of an initial model and the same model after application of the Hessian. This ratio is then multiplied against the migrated image to boost amplitudes in poorly illuminated regions of the model. Sacchi et al. (2007) approximate the diagonal of the blurring kernel via a stochastic process in which random realizations of models are passed through the blurring filter, are multiplied against the input model, and then are summed with previous realizations. Their approach will eventually converge to the true diagonal after many iterations, but practically a few iterations provide a reasonable scaling. Fletcher et al. (2016) make use of point spread functions to approximate the inverse diagonal of the Hessian. Recently Wang et al. (2016) incorporate nonstationarity of the compensation filter by calculating and applying the filter in the curvelet domain (Starck et al., 2002).

4.2 A note on imaging conditions

It is interesting to note that unlike traditional imaging methods that begin with the definition of an imaging condition, least-squares migration begins with the definition of a modeling operator, in this case first order Born scattering. Beginning with Claerbout (1971), imaging algorithms have incorporated a mechanism to quantify the magnitude of the interaction of source and receiver wavefields, and the position where this interaction takes place. There are a variety of ways this could be achieved. A popular and often used method is the *zero lag cross-correlation* imaging condition, which is based on the idea that source and receiver wavefields should be maximally correlated at their reflection point. Interestingly, for the case of least-squares migration, the adjoint of first order Born modeling is identical to this imaging condition, although in least-squares migration it is a consequence rather than a condition.

An alternative imaging condition can be parameterized as the ratio between upgoing and downgoing wavefields at a given location: $m(x,y,z) = \phi^U(x,y,z)/\phi^D(x,y,z)$. When no interface is present no energy is backscattered and the ratio is approximately zero, while an impedance contrast will result in a perturbation to the ratio. Treating the downward continued data $\phi^U(x,y,z,t)$ as a known quantity, we can write this as a linear inverse problem

$$J = ||\phi^D \mathbf{m} - \phi^U|| + \mu||\mathbf{m}||_2^2. \tag{4.29}$$

This is a localized inversion for a migrated image which has the solution

$$\mathbf{m} = \frac{\phi^{D*}\phi^U}{\mu + \phi^{D*}\phi^D} \tag{4.30}$$

where μ controls the level of damping. A high value of μ lends itself to the so-called cross-correlation imaging condition, while a low value of damping has the effect of deconvolving the source illumination giving the so-called deconvolution imaging condition. Again, this imaging condition can also be found in the formulation of least-squares migration. Since the source side wavefield propagation is equivalent for the forward and adjoint Born scattering operator (as indicated in the left-hand columns of figures 4.6 and 4.7), the magnitude of the source wavefield can be taken outside of the propagation operator and written as a diagonal weighting operator $\mathbf{W}_s$, recasting the original cost function for least-squares migration (4.23) as

$$J = ||\mathbf{T}(\mathbf{L}\mathbf{m} - \mathbf{u})||_2^2 + \mu||\mathbf{W}_s\mathbf{m}||_2^2, \tag{4.31}$$

Which incorporates the source illumination into the penalty of the model parameter. Rewriting via the change of variables, $\zeta = \mathbf{W}_s\mathbf{m}$, as

$$J = ||\mathbf{T}(\mathbf{L}\mathbf{W}_s^{-1}\zeta - \mathbf{u})||_2^2 + \mu||\zeta||_2^2, \tag{4.32}$$

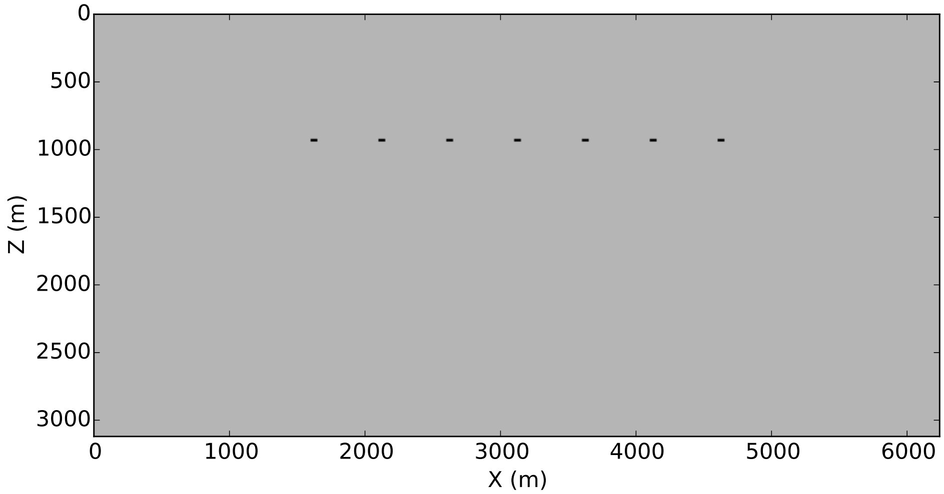


Figure 4.8: P wave velocity model used to create finite difference synthetic data. The background P-wave velocity is 2200 m/s, while the seven diffractors have a velocity of 2500 m/s. The S wave velocity model corresponds to this model by a scale factor of $1/\sqrt{3}$.

substituting $\mathbf{m} = \mathbf{W}_s^{-1}\zeta$ after the final iteration. In this way, the deconvolution imaging condition can work its way into least-squares migration in the form of a preconditioner.

4.3 Examples

4.3.1 Diffractor example

As a first example, we demonstrate the ability of ELSM to interpolate and separate wavefield components using the velocity model shown in figure 4.8.

A single shot gather, shown in figure 4.9, was generated using finite difference modeling and 30% of the traces were zeroed to simulate an irregular receiver geometry. In all examples, a compressional source wavelet was used for both numerical modeling and imaging.

Next, we applied the adjoint operator to obtain the migrated images shown in figures 4.10 and 4.11. The adjoint operator is not able to perfectly separate elastic wavefields, leaving some wavefield cross-talk artefacts in both images as well as artefacts related to the irregular receiver grid. In the PP image the PS cross-talk artefacts are propagated with too high of

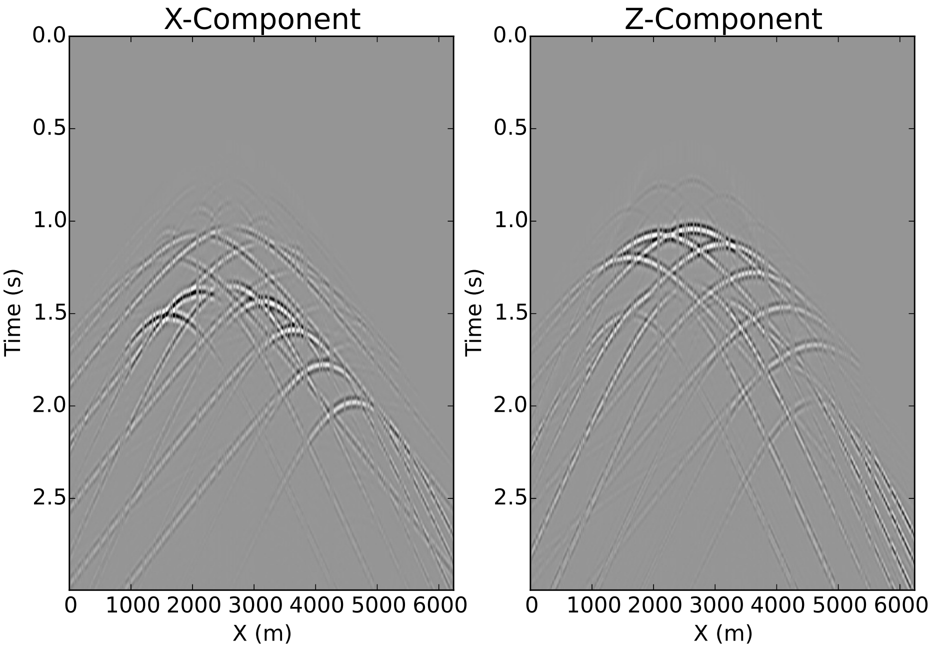


Figure 4.9: X and Z components for a synthetic shot gather at X = 2500m generated using elastic finite difference modeling. The data have been randomly decimated by 30% to simulate missing receivers.

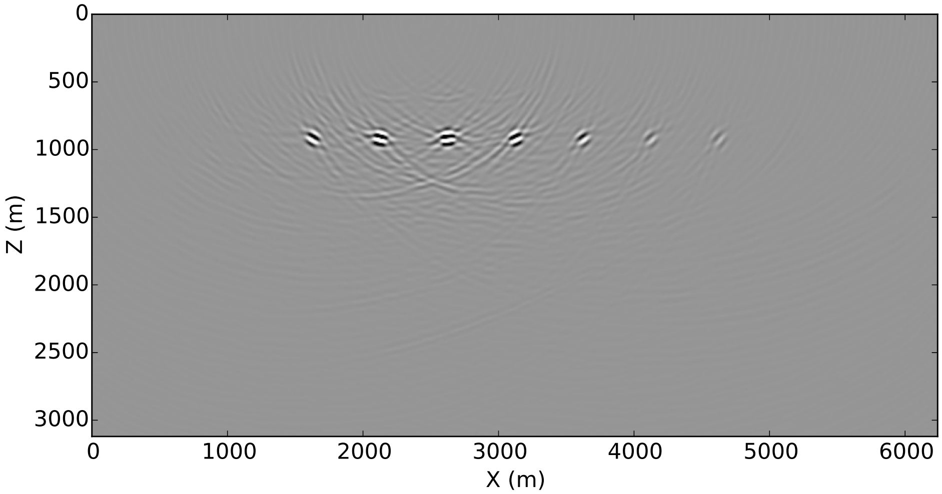


Figure 4.10: PP image obtained by elastic migration (the adjoint operator) for a single shot at X = 2500m.

a velocity, sending them too deep into the earth. Conversely, in the PS image the PP cross-talk artefacts are propagated with too low of a velocity, giving them an imaging depth that is too shallow.

For the next step we performed 20 iterations of ELSM for one shot with a purely data fitting objective function ($\boldsymbol{\mu} = \mathbf{0}$ in equation 4.24). The results are shown in figures 4.12 and 4.13.

Even without any regularization, the data fitting is able to reduce cross-talk artefacts in the images, especially in the PS image. The ELSM algorithm presents an opportunity to simultaneously interpolate and separate elastic wave modes. As an example, after first fitting the data in the exercise above we can then forward model a pure reflected P wavefield by $\boldsymbol{L}[\boldsymbol{m}_{pp}\ \mathbf{0}\ \mathbf{0}]^T$ as shown in figure 4.14, or similarly forward model a pure reflected S wavefield by $\boldsymbol{L}[\mathbf{0}\ \mathbf{0}\ \boldsymbol{m}_{ps}]^T$ as shown in figure 4.15. Interestingly, a purely data-fitting based approach such as this is not able to perfectly separate the wavefields. Additional constraints that penalize cross-talk energy might benefit the solution.

Finally, figure 4.16 shows the normalized misfit versus iteration number.

4.3.2 Flat layer example

Next, we consider a model consisting of flat layers with a depth variable v_p/v_s ratio. Figure 4.17 shows the P and S wave velocity models.

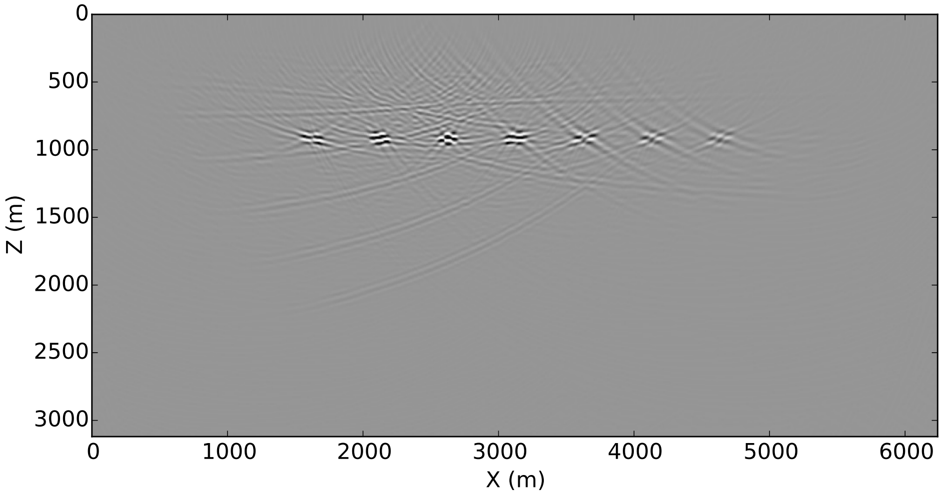


Figure 4.11: PS image obtained by elastic migration (the adjoint operator) for a single shot at X = 2500m.

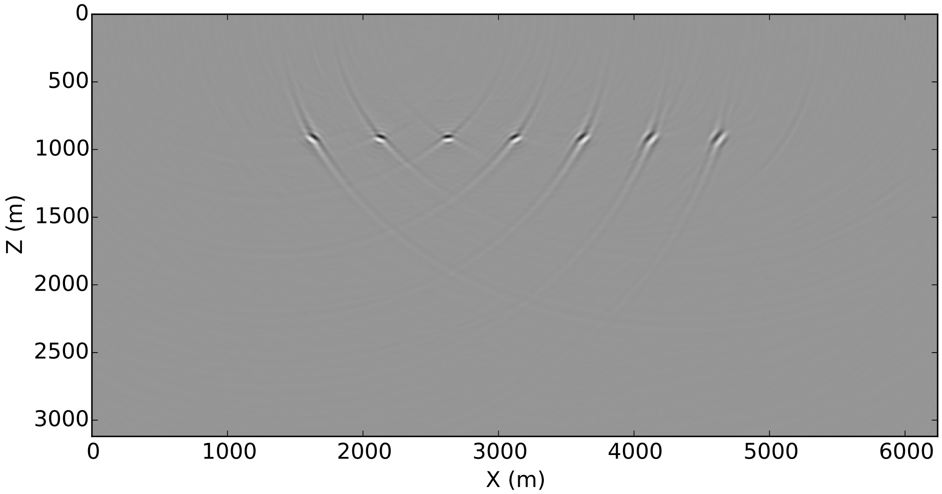


Figure 4.12: PP image obtained by 20 iterations of ELSM for a single shot at X = 2500m.

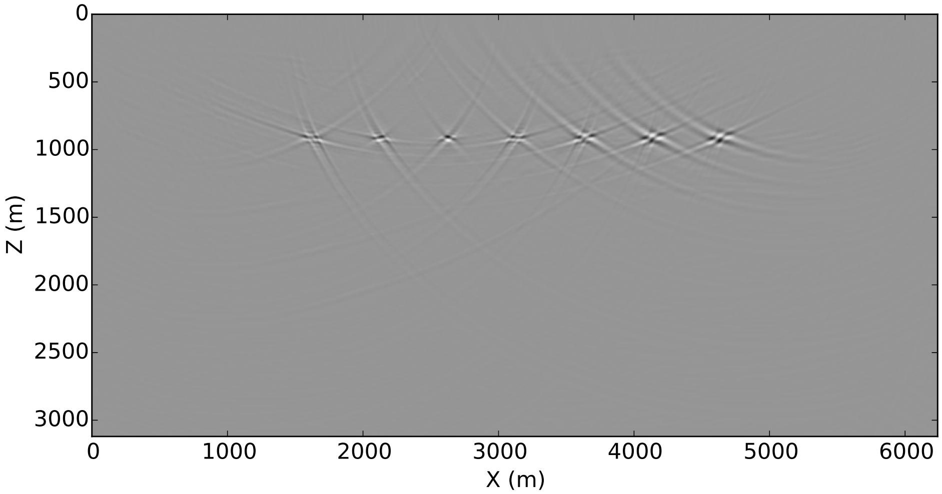


Figure 4.13: PS image obtained by 20 iterations of ELSM for a single shot at X = 2500m.

A shot gather, shown in figure 4.18, was generated using finite difference modeling and random noise was added to the data.

Next, the adjoint operator was applied to obtain the PP and PS images shown in figure 4.19.

Cross-talk artefacts are clearly visible in the adjoint images. In the PP image, the PS energy has been extrapolated using too high of a velocity (placing it too deep), while in the PS image the PP energy has been extrapolated using too low of a velocity (placing it too shallow). The results of 10 iterations of ELSM using $\boldsymbol{\mu} = [0.001\ 0.001\ 0.001]^T$ are shown in figure 4.20.

In both images, the resolution has been improved and the cross talk artefacts have been largely attenuated. The PP cross-talk artefact at approximately Z = 1000m in the adjoint PS image (figure 4.19) could be mistakenly interpreted as a true PS reflection corresponding to the event at a similar depth in the PP adjoint image. This example demonstrates the ability of ELSM to mitigate cross-talk artefacts that could potentially harm an interpretation.

Figure 4.21 shows the misfit as a function of iteration number for the case of added random noise (the solid black line), as well as for four trials considering velocity perturbations from the true models using noise-free input data. In the first trial (the solid gray line), the correct velocity was used, while in the other trials the velocities were varied by ±10% from the true models. It is interesting to note that for a single shot experiment the convergence is largely

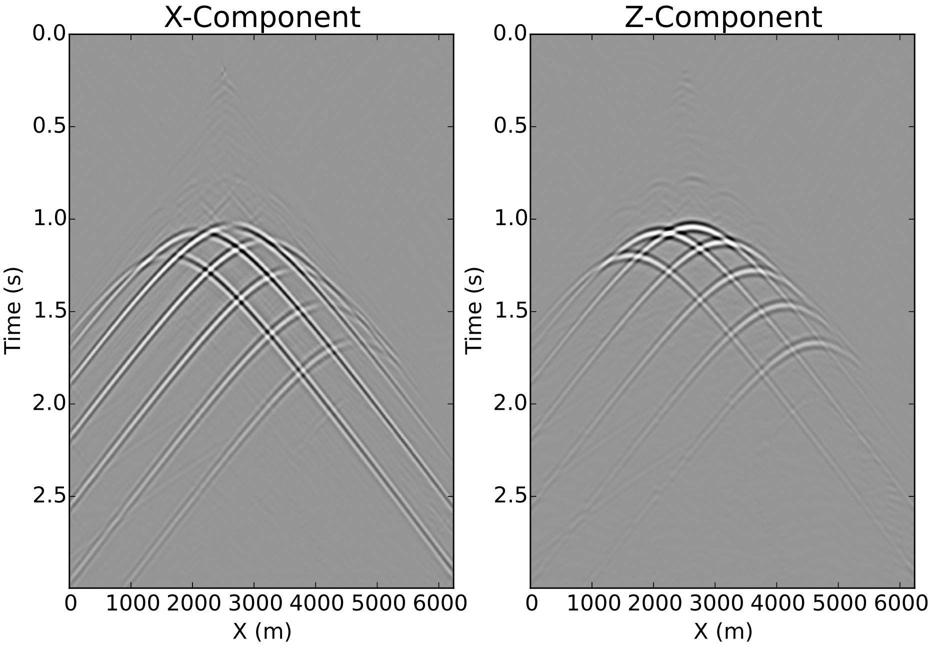


Figure 4.14: Predicted X and Z components of the interpolated and wavefield separated PP wavefield for a shot gather at X = 2500m generated by the forward operator following 20 iterations of ELSM.

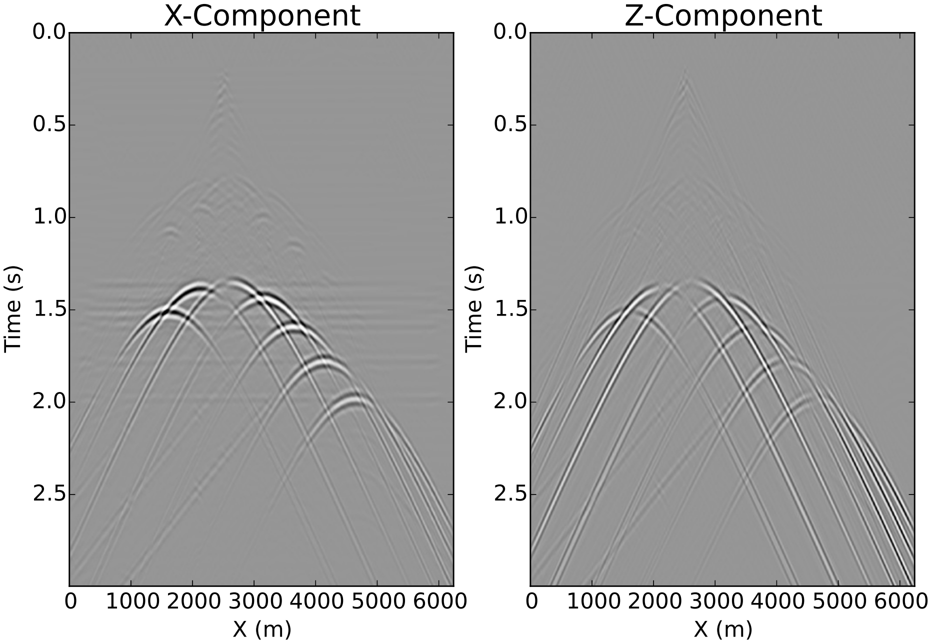


Figure 4.15: Predicted X and Z components of the interpolated and wavefield separated PS wavefield for a shot gather at X = 2500m generated by the forward operator following 20 iterations of ELSM.

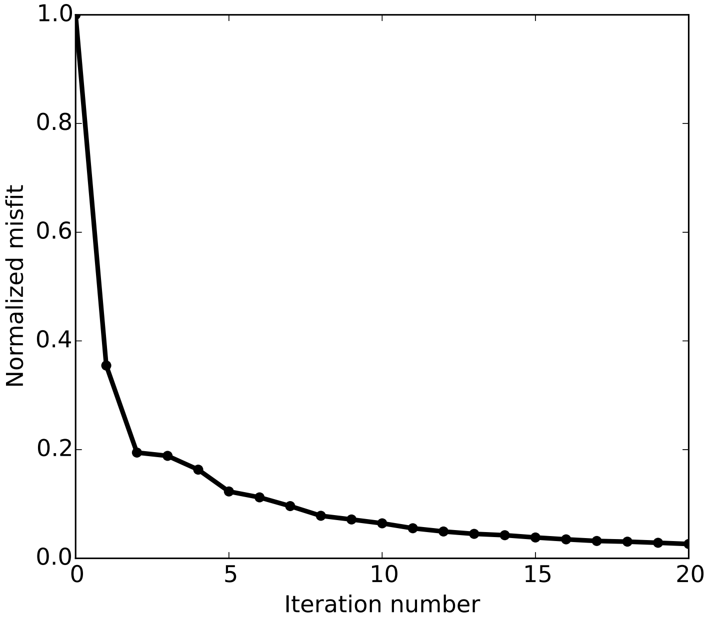
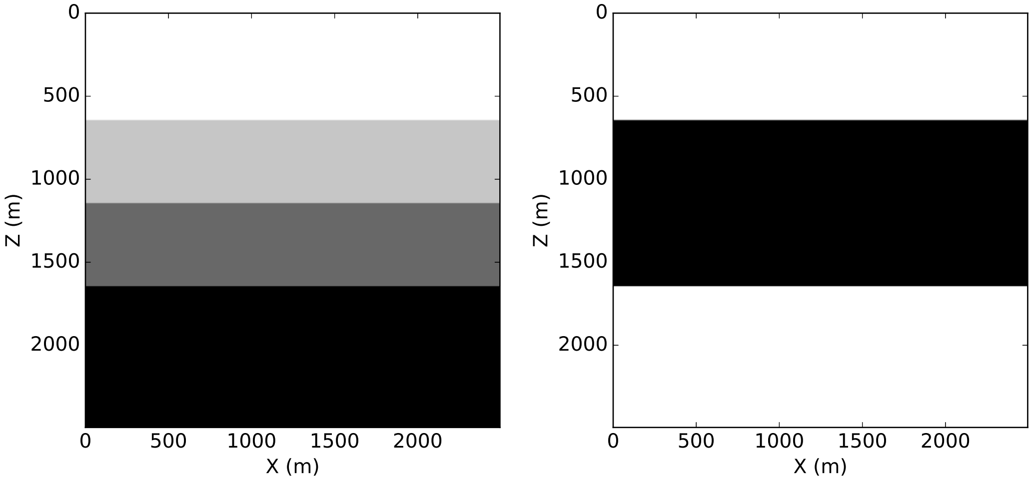


Figure 4.16: Normalized misfit versus iteration number for 20 iterations of ELSM.



(a) P-wave velocity

(b) S-wave velocity

Figure 4.17: Velocity models used to create finite difference synthetic data.

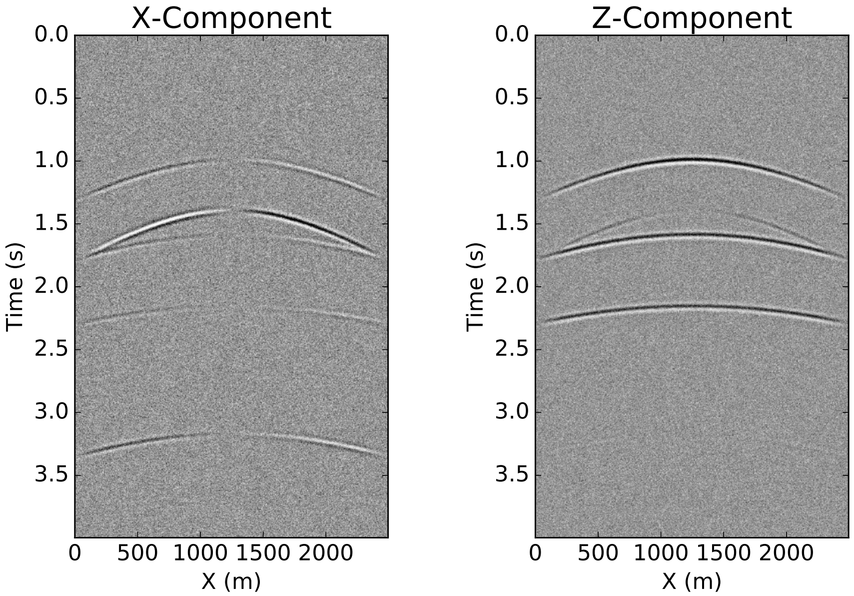
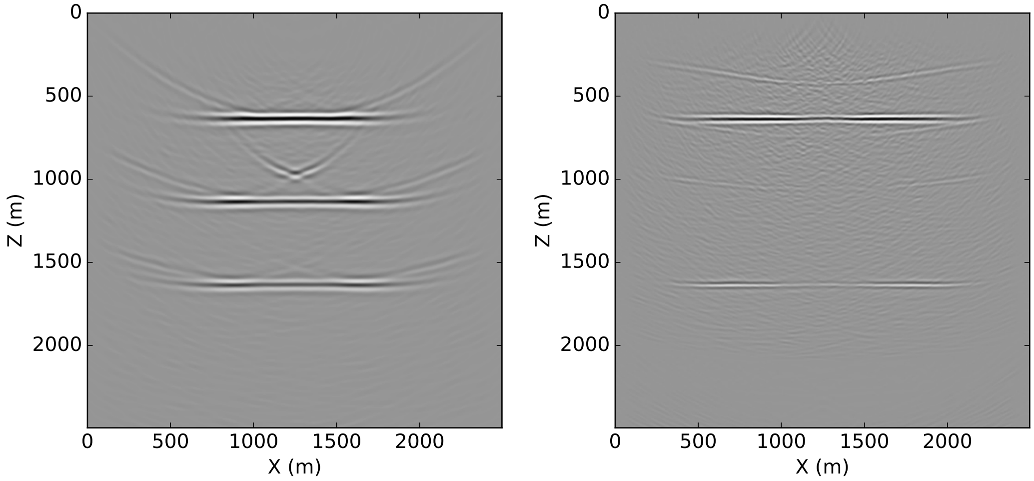


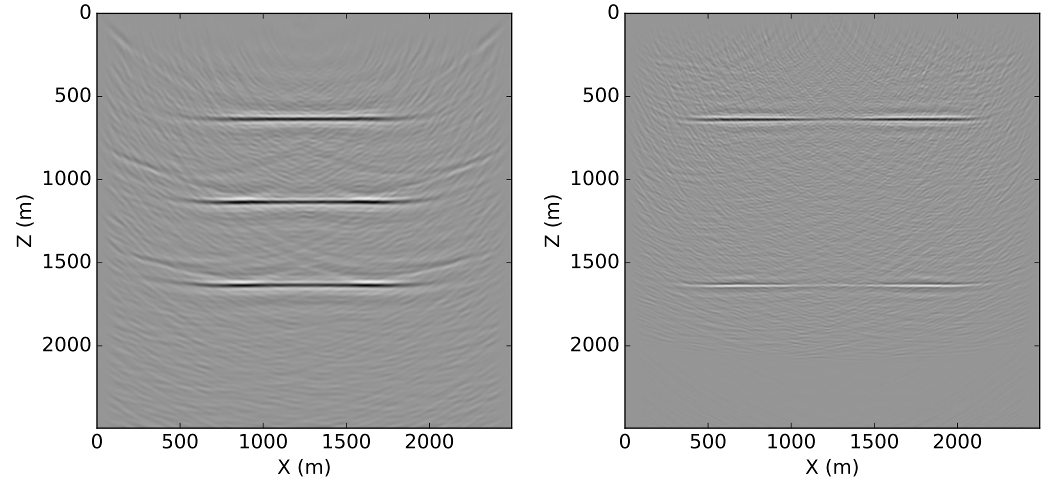
Figure 4.18: X and Z components for a synthetic shot gather at X = 1250m generated using elastic finite difference modeling with random noise added.



(a) PP

(b) PS

Figure 4.19: Images obtained by elastic migration (the adjoint operator) for a single shot at X = 1250m



(a) PP

(b) PS

Figure 4.20: Images obtained by by 10 iterations of ELSM for a single shot at X = 1250m

insensitive to these perturbations in the velocity models. When multiple sources are used in shot-profile LSM the convergence is more sensitive to the accuracy of the velocity model as a consequence of stacking/spraying multiple improperly migrated/demigrated shot gathers. This dependence is relaxed through the use of an extended image.

4.3.3 BP 2.5D Model example

For a more complicated example we next consider a modified version of the BP 2.5D Model (Etgen and Regone, 1998) (the model increments were changed from dx=12.5m, dz=12.5m to dx=8m, dz=8m prior to modeling and an S-wave velocity model was created as $1/\sqrt{3}$ times the P-wave velocity).

The P-wave velocity model is shown in figure 4.22. This model contains high-velocity contrasts that challenge the accuracy of our wavefield separation and propagation operators, and contains highly variable illumination patterns that make least-squares migration an ill-posed inverse problem. We generated 130 shot gathers using elastic finite-difference modeling. The horizontal and vertical components for a common shot gather at x = 2880m are shown in figure 4.23.

The results of applying the adjoint operator and stacking are shown in figures 4.24 and 4.25.

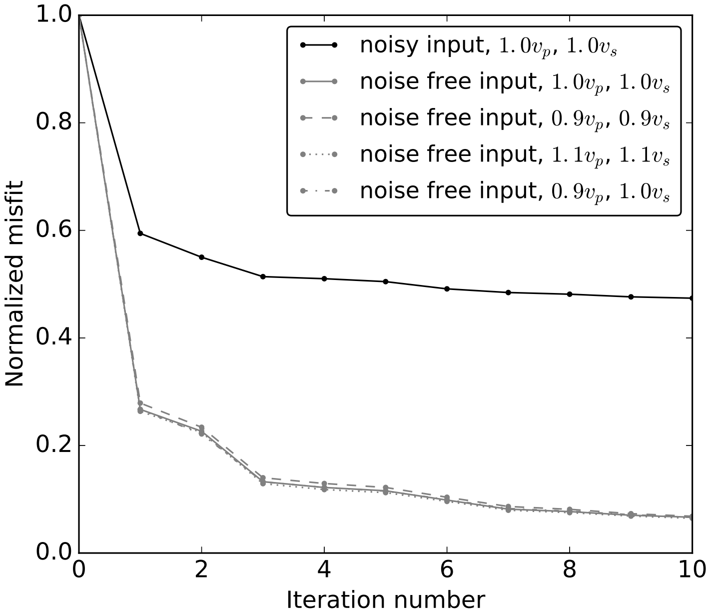


Figure 4.21: Normalized misfit versus iteration number for 10 iterations of ELSM. The solid black line shows convergence when added random noise is present on the input data. The gray curves display trials of the algorithm using different combinations of velocity error. The solid gray line shows convergence using the true velocity, while the dashed lines show convergence when using different combinations of incorrect velocities. For a single shot experiment the convergence is relatively insensitive to small velocity errors.

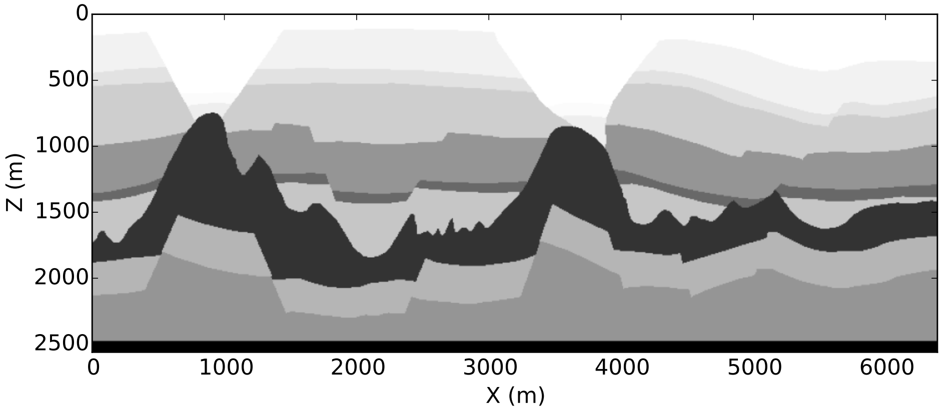


Figure 4.22: A portion of the BP 2.5D P-wave velocity model. The S-wave velocity corresponds to this model by a scale factor of $1/\sqrt{3}$.

To demonstrate the effect of the preconditioning operator written in equation 4.24 we applied it to the adjoint image volumes. The result of this filter are shown in figures 4.26 and 4.27.

It is interesting to note that while the smoothing filter was able to attenuate much of the dipping energy within each angle gather, it produced little effect on the overall stack. This implies that lateral smoothing across reflection angle attenuates little coherent, properly migrated (i.e. flat), reflection energy. The results of applying 10 iterations of preconditioned ELSM with $\boldsymbol{\mu} = [0.1\ 0.1\ 0.1]^T$ are shown in figures 4.28 and 4.29, with the cost function shown in figure 4.30.

Ideally, the amplitude of the bottom-most reflector should be constant, but variations in illumination cause the amplitude to fluctuate. Comparing the adjoint and ELSM PP images (figures 4.24 and 4.28), we see that the larger amplitude variations of this bottom reflector have been partially corrected, and improvements to the spatial resolution can be seen throughout the image. Comparing the adjoint and ELSM PS images (figures 4.25 and 4.29), we see a similar improvement to the illumination of the base reflector, especially the continuity of the bottom-most reflector at x=1200m. Other improvements include the attenuation of wavefield cross-talk artefacts within the high-velocity body, as well as improvements to the spatial resolution of the surrounding structure. Rickett (2003) apply a purely data-fitting based least-squares migration to the BP 2.5D Model, showing an increase in the noise level with least-squares migration and instead advocate using an approximation to the diagonal

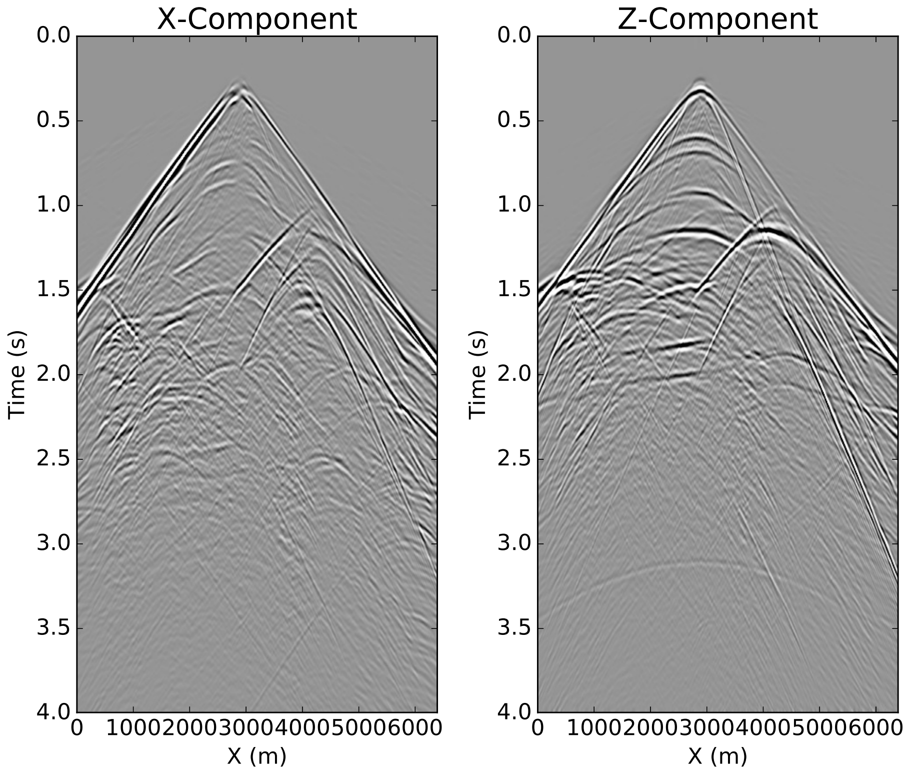


Figure 4.23: X and Z components for a synthetic shot gather at x = 2880m generated using elastic finite difference modeling.

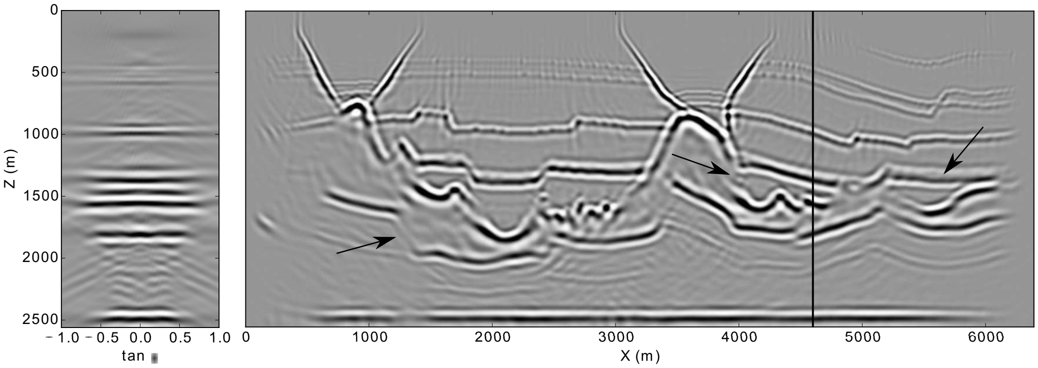


Figure 4.24: PP image and an angle gather from x=4600m obtained by elastic migration (the adjoint operator). Arrows indicate regions of the model that have been poorly imaged.

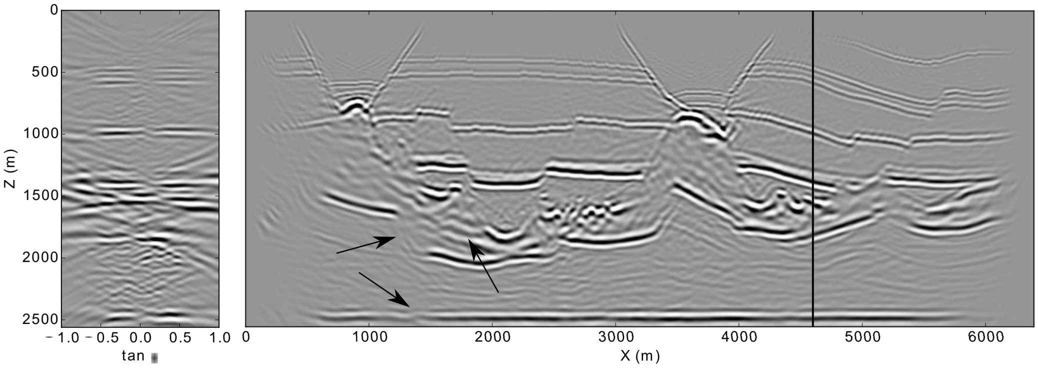


Figure 4.25: PS image and an angle gather from x=4600m obtained by elastic migration (the adjoint operator). Arrows indicate regions of the model that have been poorly imaged.

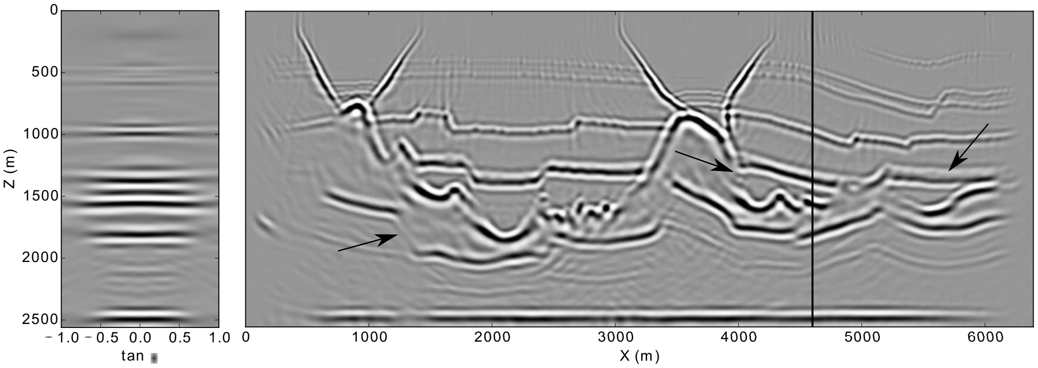


Figure 4.26: PP image and an angle gather from x=4600m obtained by elastic migration (the adjoint operator). A smoothing filter has been applied across angles to attenuate dipping cross talk energy. Notice that this filter has little effect on the overall stack. Arrows indicate regions of the model that have been poorly imaged.

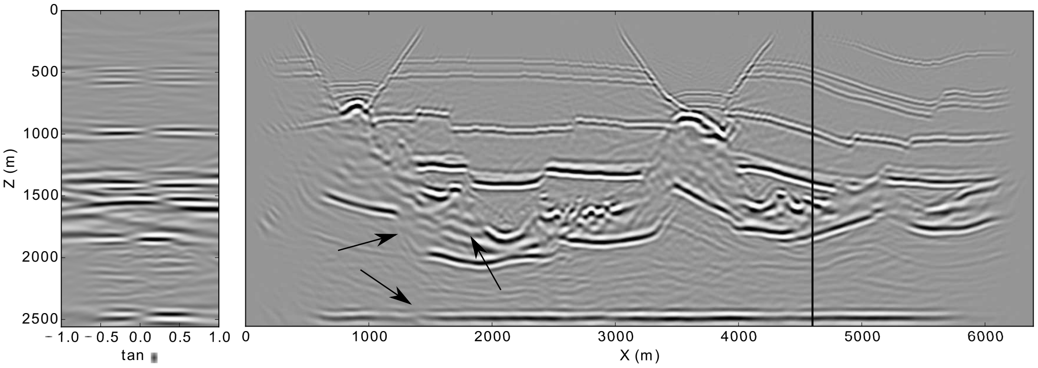


Figure 4.27: PS image and an angle gather from x=4600m obtained by elastic migration (the adjoint operator). A smoothing filter has been applied across angles to attenuate dipping cross talk energy. Notice that this filter has little effect on the overall stack. Arrows indicate regions of the model that have been poorly imaged.

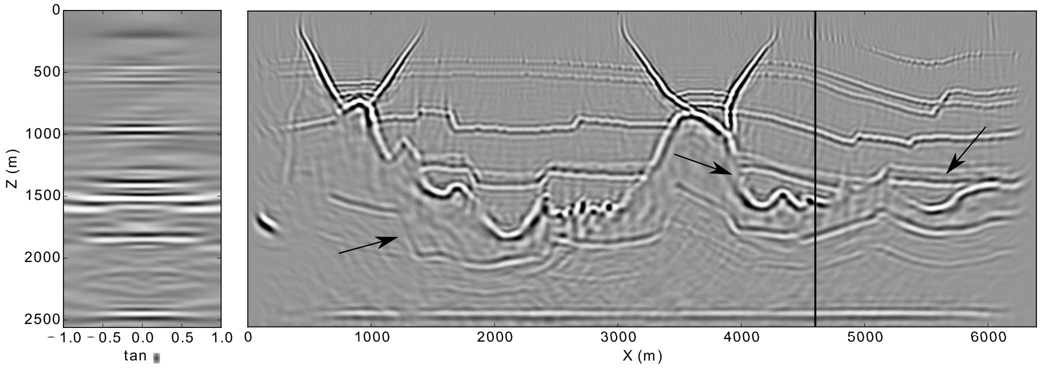


Figure 4.28: PP image and an angle gather from x=4600m obtained by 10 iterations of ELSM. Arrows indicate regions of the model where the image has been improved.

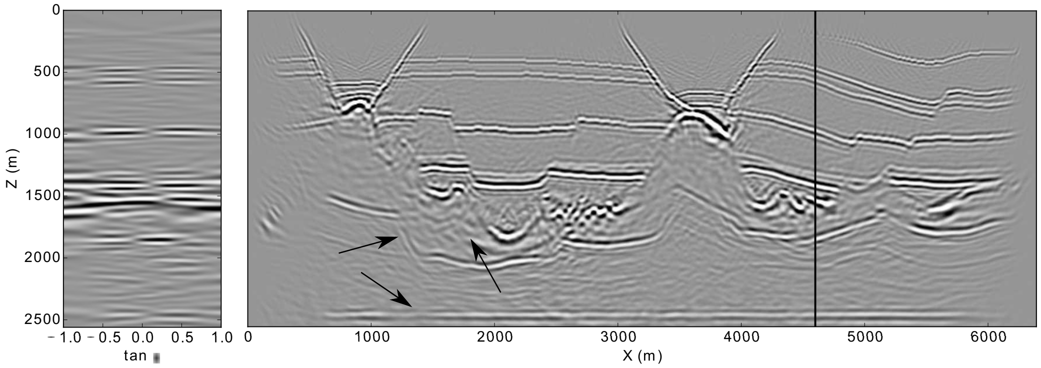


Figure 4.29: PS image and an angle gather from x=4600m obtained by 10 iterations of ELSM. Arrows indicate regions of the model where the image has been improved.

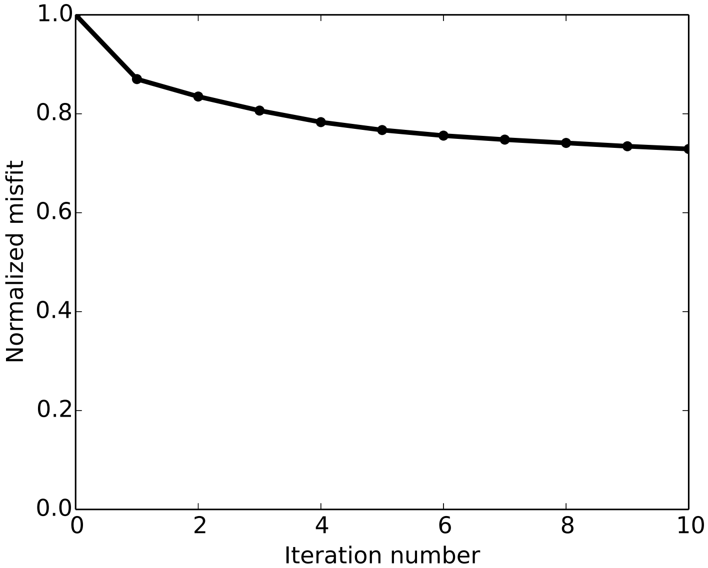


Figure 4.30: Normalized misfit versus iteration number for 10 iterations of ELSM.

of the Hessian using the forward and adjoint operators common to least-squares migration. Illumination compensation such as this could also be applied using the elastic forward and adjoint operators described in this paper.

4.4 Conclusions

Least-squares migration algorithms attempt to fit recorded data with predictions generated from a migrated image. By improving the accuracy of the migration operator to include elastic wave propagation we expect to improve the ability of least-squares migration to fit reflection amplitudes. Our examples demonstrate that least-squares migration can improve the imaging of multicomponent seismic data and that it can also be used for missing trace interpolation and wavefield decomposition. A significant challenge in the migration of elastic wavefields is the mitigation of cross-talk artefacts. While pure data fitting is able to partially resolve the problem, incorporating additional constraints could further improve the performance of the method.

CHAPTER 5

Regularized imaging of OBC data

5.1 Receiver profile modeling and migration

The theoretical foundation for least squareS-wave equation migration is first order Born scattering. This consists of propagation of a wavefield away from a source, interaction of the source wavefield with reflectors, followed by propagation of the scattered wavefield toward an array of receivers. In other words, the forward operator is set up to generate common source gathers. In many situations this is a convenient formulation; for example in land seismic acquisition receivers are inexpensive to place while sources require either expensive planning and drilling in the case of dynamite, or time-consuming sweeps in the case of vibrators. In other situations, receivers are far more costly than sources. Such is the case in ocean bottom acquisition where receivers must be deployed and recovered from the sea floor. Figure 5.1 shows the source and receiver sampling for a 2D line taken from an Ocean Bottom Cable (OBC) survey acquired offshore Northwest Australia in the Dampier sub-basin. The acquisition target in this region are deep (∼3700m in depth) Triassic gas bearing sands. The imaging in the area is complicated by the presence of shallow carbonates and clastic intervals which provide rapid velocity variations. Note the sparse distribution of receivers (500m spacing) compared to the dense distribution of shots (25m spacing). This $11,975m$ 2D line is a subset of a larger 3D dataset.

Clearly, in this case, we have well sampled common receivers gathers rather than well sampled common source gathers. A look at a common source gather (figure 5.2) reveals just how poorly sampled data are in this domain; distinct reflection events are barely identifiable. Looking at the F-K spectrum for these data (figure 5.3) we see the sparse sampling has aliased the events severely.

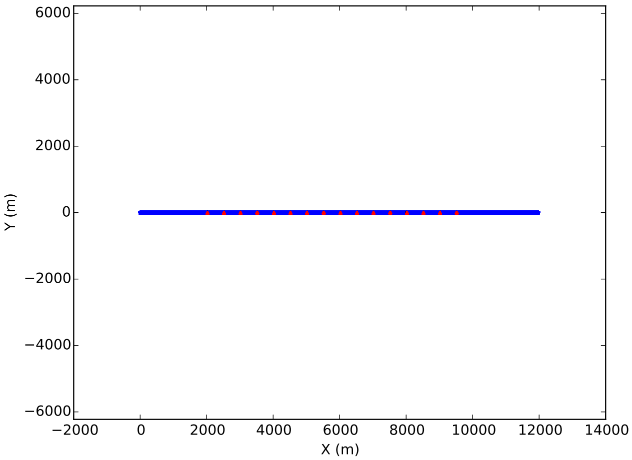


Figure 5.1: coordinates of sources (blue stars) and receivers (red triangles) for an ocean bottom dataset. Note sparse distribution of receivers compared to the dense distribution of shots.

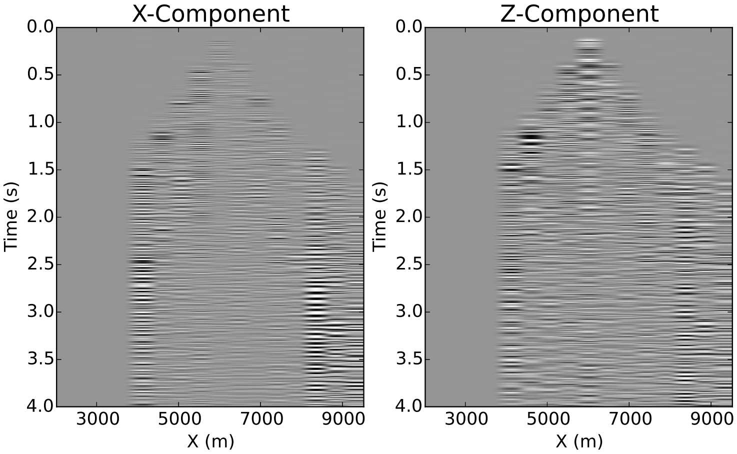


Figure 5.2: X and Z components for a common source gather at X=6000m. The sampling in this domain is very sparse with a 500m receiver spacing.

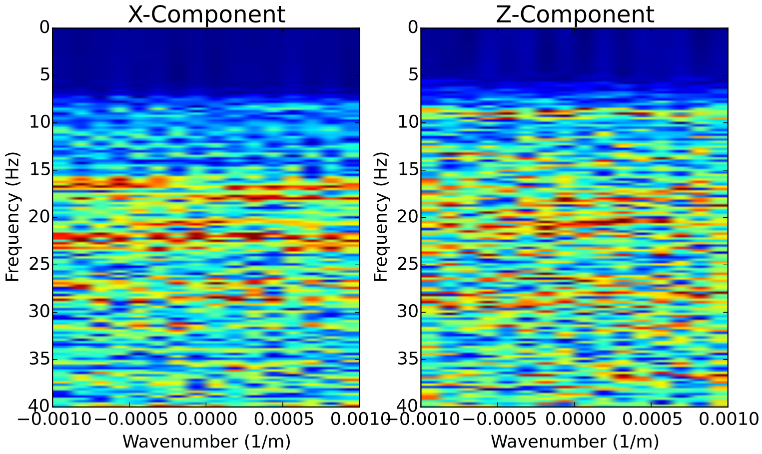


Figure 5.3: X and Z component FK Amplitude spectra for a common source gather at X=6000m. Clearly, the data are highly aliased in this axis.

A solution to this problem is to invoke source-receiver reciprocity. Organizing the data into common receiver gathers (figure 5.4), we see the trace spacing is improved from $500m$ to $25m$. The FK spectrum of these data is shown in figure 5.5. Compared with the common source gather FK spectrum in figure 5.3, it is clear that the aliasing of the data has been greatly reduced.

Source-receiver reciprocity is a consequence of Newton's third law of motion: "for every action, there is an equal and opposite reaction[1]." In the case of acoustic wave propagation, it is obvious that a wave traveling from a source to a receiver will trace the same path as a wave traveling from a receiver to a source (see Fenati and Rocca (1984) for a demonstration on field data). The wave will arrive with the same travel time and with the same amplitude, making reciprocal waveforms both kinematically and dynamically identical. This principle is often invoked to perform acoustic common receiver gather migration of ocean bottom data. Less obvious is the application of source-receiver reciprocity to the case of elastic wave propagation.

The Maxwell-Betti reciprocal work theorem provides a way to extend source-receiver reciprocity to the elastic case. Considering displacement fields $\boldsymbol{u}_1$ and $\boldsymbol{u}_2$ that are the result of body forces $\boldsymbol{f}_1$, and $\boldsymbol{f}_2$ plus tractions $\boldsymbol{t}_1$, and $\boldsymbol{t}_2$, the Maxwell-Betti reciprocal work theorem states that

$$\int_S \boldsymbol{t}_2 \cdot \boldsymbol{u}_1 dS + \int_V \boldsymbol{f}_2 \cdot \boldsymbol{u}_1 dV = \int_S \boldsymbol{t}_1 \cdot \boldsymbol{u}_2 dS + \int_V \boldsymbol{f}_1 \cdot \boldsymbol{u}_2 dV. \tag{5.1}$$

---

[1] Consider the reciprocal statement: "For every reaction, there is an equal and opposite action," which has the same basic meaning.

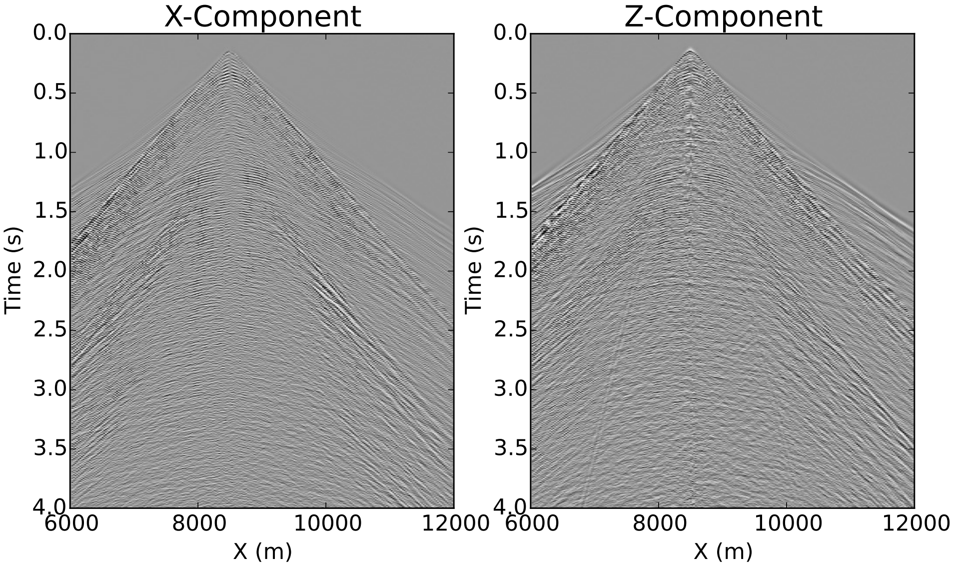


Figure 5.4: X and Z components for a common receiver gather at X=5521m.

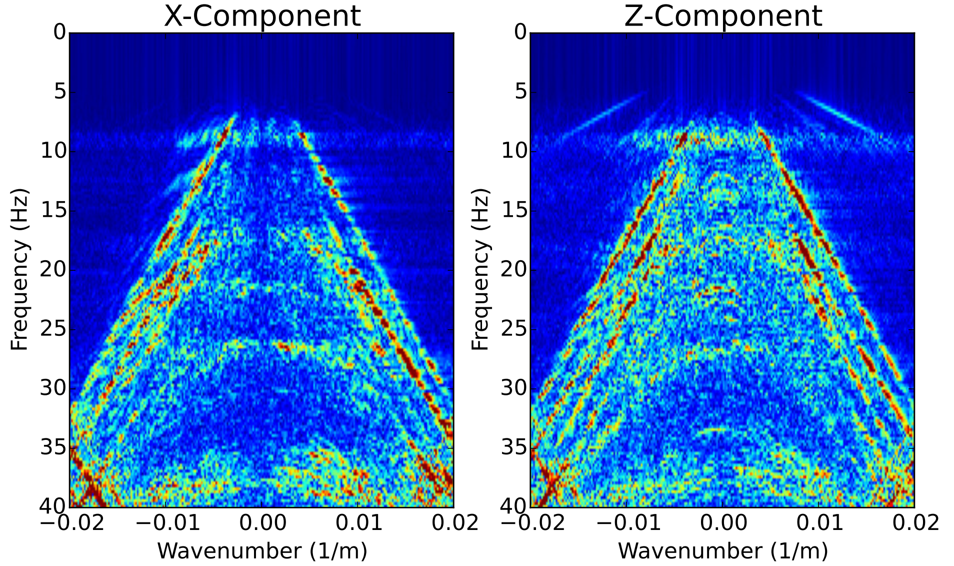


Figure 5.5: X and Z component FK Amplitude spectra for a common receiver gather at X=5521m.

Taking positions 1 and 2 to mean the position of a source or receiver, this equation relates the displacement fields that result from to a force applied at these positions. Knopoff and Gangi (1959) demonstrate the Maxwell-Betti reciprocal work theorem for wave propagation an inhomogeneous, anisotropic viscoelastic medium using transducers located on a large aluminum sheet with a brass object acting as a scatterer. A useful property of equation 5.1 is that it implies the reciprocity of Green's functions

$$G_{ij}(\boldsymbol{x}_1|\boldsymbol{x}_2) = G_{ji}(\boldsymbol{x}_2|\boldsymbol{x}_1) \tag{5.2}$$

(Tarantola, 1987), implying the response at position x_1 in direction i due to a source at position x_2 in direction j, is equal to the response at position x_2 in direction j due to a source at position x_1 in direction i. Put more simply, if you invoke reciprocity you must also switch the source/receiver impulses/responses[2]. Figure 5.6 illustrates this concept with wave modes. Consider an explosive source in a water layer converting to an S-wave reflection which is recorded by an ocean bottom receiver. A correct use of the Maxwell-Betti reciprocal work theorem is shown in figure 5.7. Switching the source/receiver impulses/responses will provide an identical response as indicated by equation 5.2. An incorrect use of the Maxwell-Betti reciprocal work theorem is shown in figure 5.8. Here the source/receiver impulses/responses were not interchanged. Clearly, the response will not be equivalent to that in figure 5.6. This differences will be both kinematic (the ray-path is different, and hence the travel-time will likely be different), as well as dynamic (the S-wave cannot propagate in the water column).

Returning to the case of elastic wave propagation in OBC acquisition, with an acoustic (water-contained) source wavelet, and elastic (ocean bottom-contained) response at the receivers, after invoking source-receiver reciprocity the first order Born approximation (see 4.15) is written

$$\vec{\boldsymbol{u}}(\boldsymbol{x}^s, \omega) \sim \omega^2 \int_{-\infty}^{\infty} G_0(\boldsymbol{x}^s|\boldsymbol{x}) \boldsymbol{M}(\boldsymbol{x}) G_0(\boldsymbol{x}|\boldsymbol{x}^g) d\boldsymbol{x}, \tag{5.3}$$

where $\boldsymbol{x}$ indicates all subsurface positions, $\boldsymbol{x}^s$ indicates all source positions, and $\boldsymbol{x}^g$ indicates the position of a receiver, $G_0(\boldsymbol{x}|\boldsymbol{x}^g)$ is the 3x1 Green's function response at all subsurface positions due to an impulse at the position of the receiver (representing all modes of received wavefield), $\boldsymbol{M}(\boldsymbol{x})$ is a 3x3 matrix of scattering potentials, and $G_0(\boldsymbol{x}^s|\boldsymbol{x})$ is the 3x3 Green's function response at the source positions due to an impulse generated at all subsurface positions (scattering all combinations of incident and scattered wave modes). The

---

[2] practically speaking equation 5.2 considers only displacements, not potentials. As an example consider a downward impulse at the location of a source resulting in a sideways impulse at the location of a receiver. Equation 5.2 implies that a sideways impulse at the location of a source will result in a downward impulse at the location of a receiver. To apply this useful relationship in practice it is necessary to translate from wavefield potentials to displacements using equation 4.12.

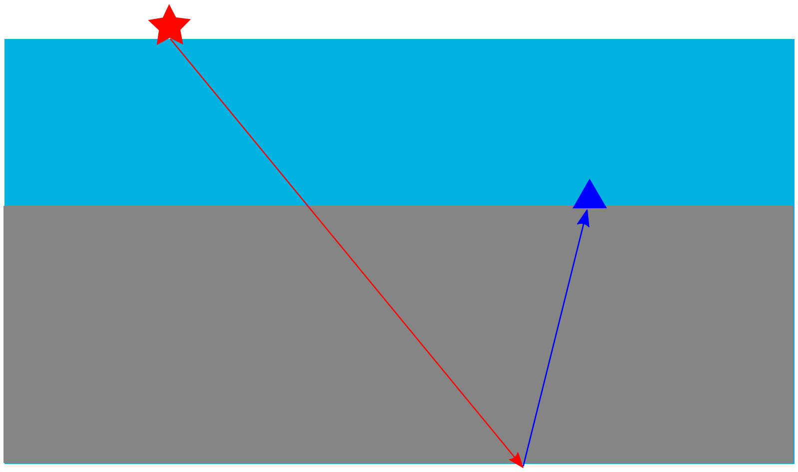


Figure 5.6: Schematic for P-wave source in a water layer converting to an S-wave reflection which is recorded by an ocean bottom receiver.

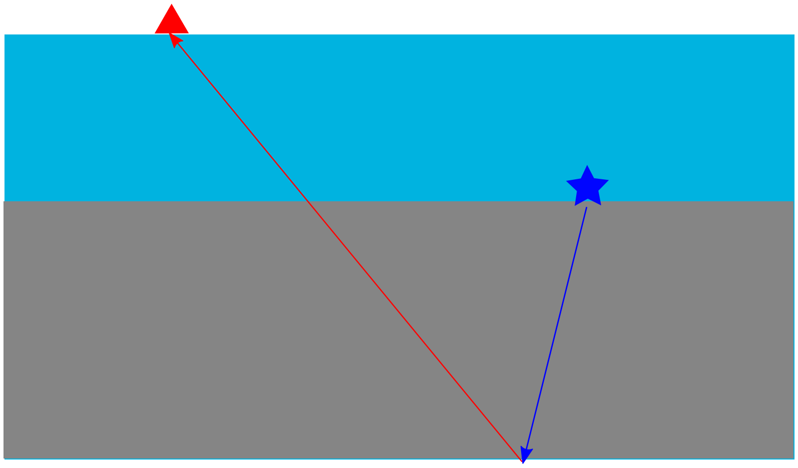


Figure 5.7: Correct use of the Maxwell-Betti reciprocal work theorem. An S-wave source embedded in the ocean bottom converting to a P-wave reflection which is recorded by a receiver in the water layer.

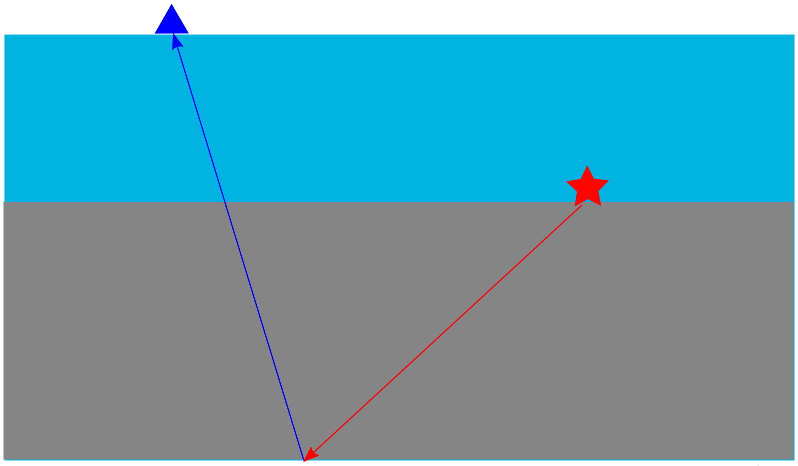


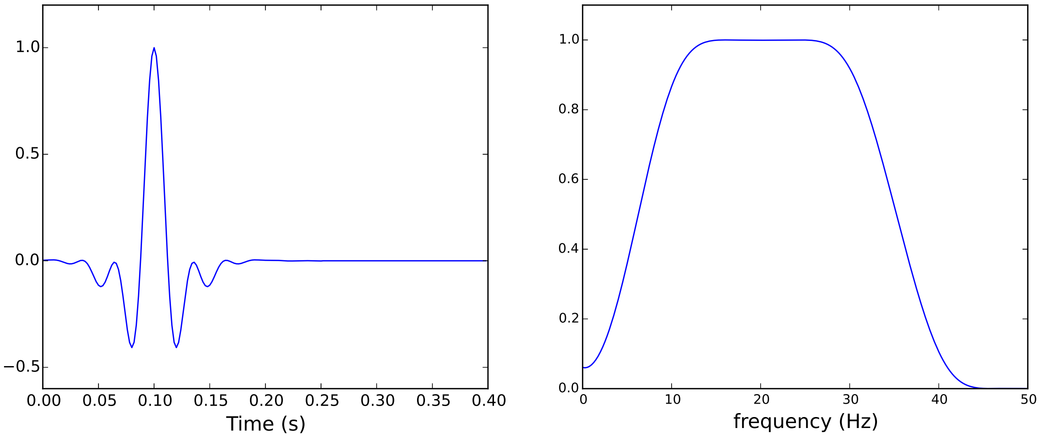
Figure 5.8: Incorrect use of the Maxwell-Betti reciprocal work theorem (the source and receiver impulses/responses were not interchanged). A P-wave source embedded in the ocean bottom converting to an S-wave reflection which is recorded by a receiver in the water layer. Clearly, the response will not be equivalent to that in figure 5.6. This differences will be kinematic (the ray-path is different), and dynamic (the S-wave cannot propagate in the water column).

propagated data components at all source locations for a given frequency, ω, are denoted by $\vec{\boldsymbol{u}}(\boldsymbol{x}^s, \omega)$. In this chapter, I use this relation to formulate elastic least squares receiver profile wave equation migration and apply the algorithm to image multicomponent ocean bottom receiver gathers.

5.2 Velocity analysis

A first step in the migration of elastic data is the estimation of a source wavelet. Figure 5.9 shows the wavelet estimated for these data and its associated amplitude spectrum. The wavelet was estimated by considering an average water depth in the survey area of $80m$ (as evidenced by the approximately 50ms (one-way) onset time of near-offset traces), and adding a $50ms$ delay to both the wavelet and the data to ensure the wavelet is causal. The wavelet type is an Ormsby wavelet with corner frequencies of 2,10,30, and 40Hz respectively.

Next P and S-wave velocity models needed to be estimated. A straightforward manual approach to estimate a 1D velocity model was used, making use of angle super-gathers. Groups of 100 neighboring angle gathers were stacked to form "super-gathers" and elastic migration and manual velocity model updates were made to improve the flatness of the gathers. PP and PS supergathers gathers for $X = 4000m$ after iterating this procedure several times are shown in figures 5.10 and 5.11, with 10% velocity perturbations to show the



(a) Wavelet

(b) Amplitude spectrum

Figure 5.9: Estimated source wavelet used for imaging.

deviation from the estimated velocity. This procedure has flattened most events reasonably well. For both the PP and PS angle gathers the −10% velocity perturbation tends to under-migrate events (not pushing events deep enough at far angles and making them point upward in a ”∪-shape” within angle gathers), while the +10% velocity perturbation tends to over-migrate events (pushing events too deep at far angles and making them point downward in a ”∩-shape” within angle gathers). This provides a reasonable confidence interval for the chosen 1D velocity model.

Unlike traditional velocity analysis, velocity analysis in elastic wave propagation influences not only the kinematics of the angle gathers, but also the decomposition and composition of wave-modes at a given depth interval. The velocity models resulting from migration velocity analysis are shown in figures 5.12 and 5.13.

5.3 Application of adjoint operator and least squares migration

After velocity analysis the adjoint (migration) operator was applied to the data, resulting in the stacked images shown in figures 5.14 and 5.15. Because of the coarse spatial sampling of receivers, there are significant gaps in the shallow imaging. The contributions of each of the 16 common receiver gathers are easily identified in the shallow region of the images. In the PP common image gather (figure 5.14) there is significant over-migrated PS cross-talk

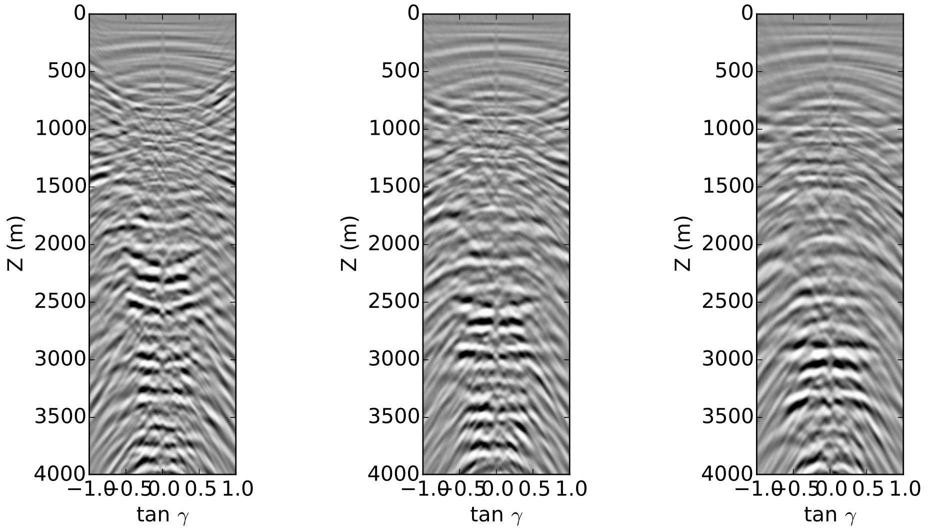


Figure 5.10: PP angle gathers for X=4000m after elastic migration with -10% velocity perturbation (left), 0% velocity perturbation (middle), and +10% velocity perturbation (right).

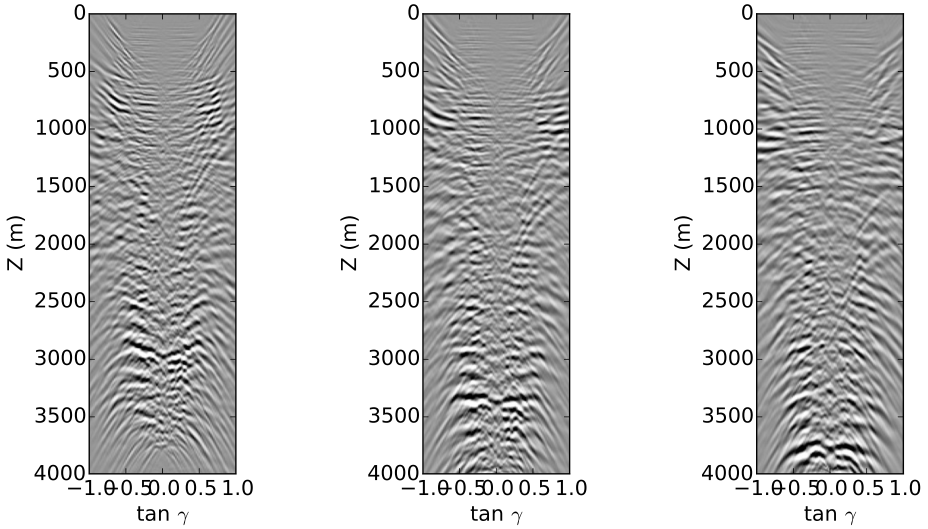


Figure 5.11: PS angle gathers for X=4000m after elastic migration with -10% velocity perturbation (left), 0% velocity perturbation (middle), and +10% velocity perturbation (right).

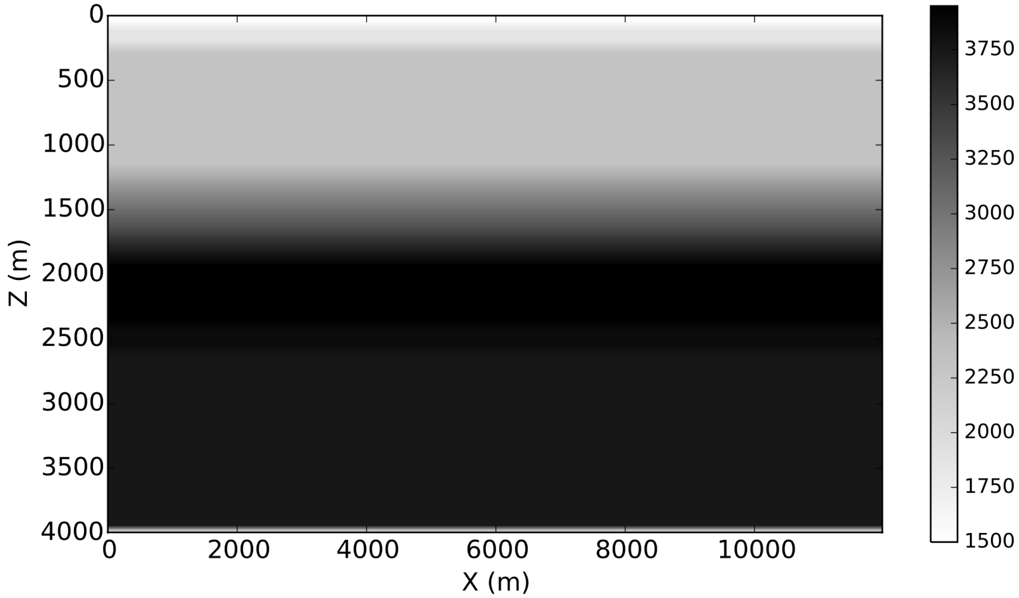


Figure 5.12: P-wave velocity model.

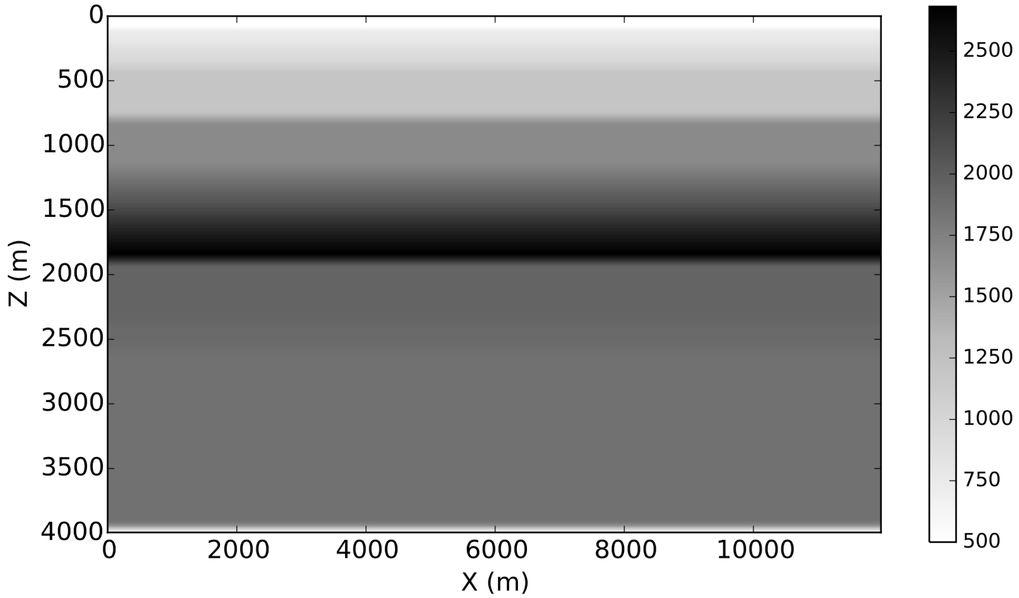


Figure 5.13: S-wave velocity model.

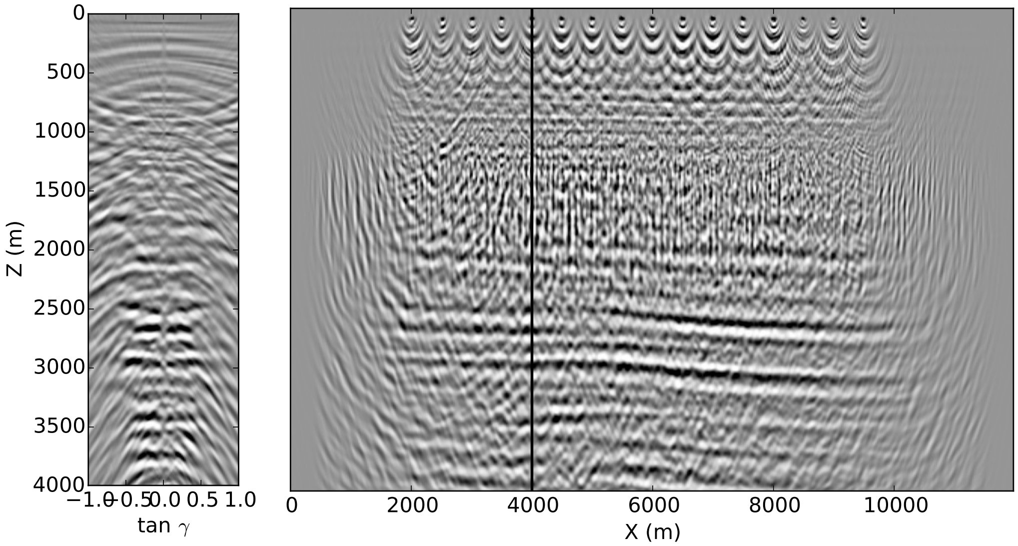


Figure 5.14: PP image after elastic receiver gather migration.

energy. Looking at the PS image (figure 5.15) there is significant PP crosstalk energy in the common image gather that appears under-migrated. Since this energy is dipping, much of it stacks out in the adjoint images, but the images are still degraded by the mispositioned energy.

Performing 5 iterations of elastic receiver-profile least squares imaging results in the stack sections shown in figures 5.16 and 5.17.

In the PP image, it is clear that elastic least squares migration has improved the shallow portion of the image considerably. A reduction of dipping energy in the common image gathers is also apparent, suggesting the approach has reduced the effects of wavefield crosstalk artefacts on the final images. The cost function is shown in figure 5.18.

5.4 Remarks

While elastic least squares migration was partially able to compensate for the poor spatial sampling of receivers, these data might have benefitted from the Fourier interpolation methods described in chapter 3 of this thesis. An often encountered problem with interpolation algorithms is that the axes that could benefit the most from interpolation are typically the same axes that are the most difficult to interpolate. The receiver spacing of 500m makes this axis severely undersampled and prone to spatial aliasing (as indicated by the FK amplitude spectrum in figure 5.3). Efforts to interpolate this axis are challenged by severe aliasing

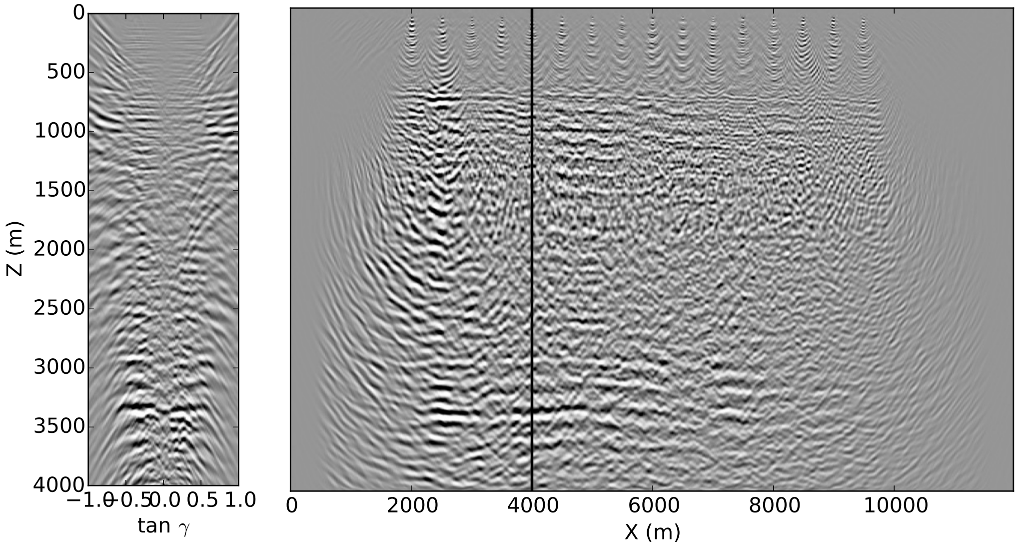


Figure 5.15: PS image after elastic receiver gather migration.

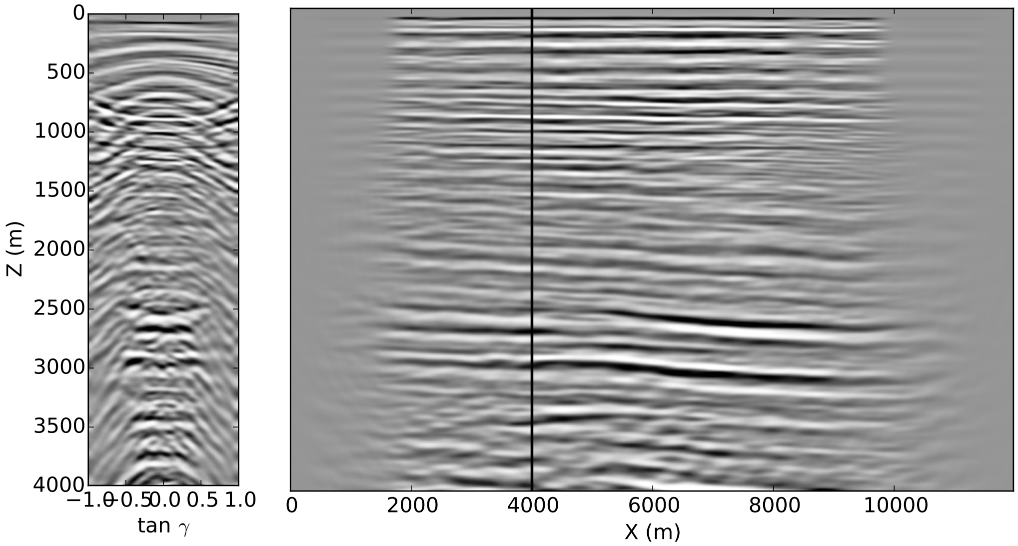


Figure 5.16: PP image after 5 iterations of preconditioned elastic least squares receiver gather migration.

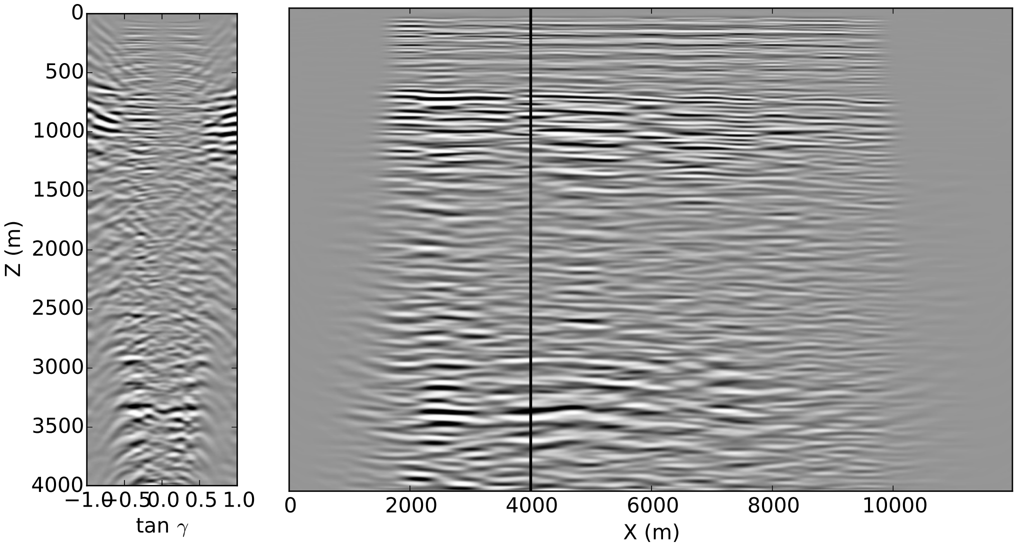


Figure 5.17: PS image after 5 iterations of preconditioned elastic least squares receiver gather migration.

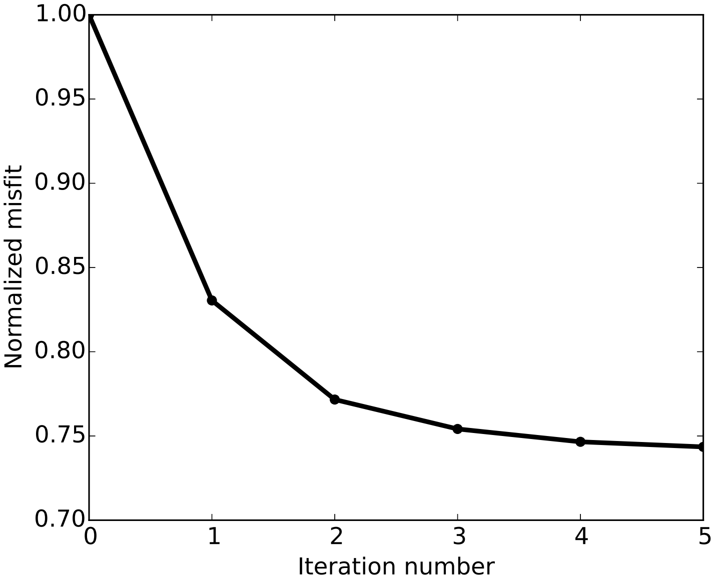


Figure 5.18: Normalized cost as a function of iteration number.

and the regularity of the receiver spacing (making the amplitude of spectral coefficients associated with aliasing equal in amplitude to spectral coefficients associated with signal). An approach that could potentially work well is to use least-squares receiver-gather migration as an interpolation engine (as described in chapter 4 of this thesis). A more robust approach would be to design ocean bottom acquisition parameters with these goals in mind. While it is often not practical to use high enough channel counts to avoid spectral aliasing along receiver axes, other measures could be put in place to satisfy sampling criteria. For example, 5D interpolation technology allows "swapping" of sources and receivers to achieve sampling patterns that are well sampled in 4 spatial dimensions. Practically this means that up to 50% of source expenditure could perhaps be better used to improve receiver sampling. Additionally, money spent placing sources and receivers on an almost perfectly regular grid would be better spent on using slightly more sources and receivers, while allowing their positions to fall on a random or semi-randomized grid (Hennenfent and Herrmann, 2008; Naghizadeh, 2015a,b).

5.5 Conclusions

This chapter introduced the concept of applying the reciprocity of Green's functions to apply least squares migration to common receiver gathers, with application to an ocean bottom seismic dataset. Least squares imaging was able to improve the shallow portion of the imaging, primarily for the PP section. While the objective of this acquisition were deeper targets ($\sim$3700m in depth), the improved imaging of the shallow structure could provide uplift in the identification of potential drilling hazards.

CHAPTER 6

Conclusions

Vector-elastic processing strategies can improve the performance of multicomponent seismic data processing tools over their scalar-acoustic counterparts. Research into vector-elastic processing has thus far lacked the necessary tools for two important steps, namely interpolation and regularized imaging. This thesis aimed to address these gaps in the workflow by introducing two new algorithms; vector interpolation, to compensate for irregular surface sampling of multicomponent seismic sources and receivers; and regularized elastic least-squares imaging, to compensate for irregular subsurface illumination of elastic wavefields in the imaging process.

Vector interpolation begins by treating multicomponent seismic data as a vector quantity in the Fourier domain using the quaternion Fourier transform. This approach makes it possible to extend a popular Fourier regularization algorithm, Projection Onto Convex Sets (POCS), to the vector case. Vector regularization improves the quality of an interpolation while preserving the orthogonality of the input components. The improvement stems from the fact that one amplitude spectrum is calculated for multiple components of data. This overlap is beneficial when the components have significant spectral overlap because it increases the overall sparsity of the spectrum, but can be a drawback if the components do not share significant overlap in the frequency-wavenumber domain. By using a symplectic representation of quaternions and choosing an appropriate axis for the Quaternion Fourier Transform (QFT) the formulation for the QFT can be greatly simplified and existing FFT algorithms can be used, allowing this approach to be applied to an arbitrary number of components. In chapter 3 this method was applied to the horizontal components of a land dataset displaying the effects of shear wave splitting. Interpolating the data in four spatial dimensions using a vector approach better preserved the azimuthal signature of shear wave splitting and allowed for an improved correction of the horizontal components. As a generalization of the approach another popular Fourier regularization algorithm, Minimum Weighted Norm Interpolation (MWNI), was extended to the vector case. Applying the two algorithms to a synthetic dataset it is clear that both the POCS and MWNI algorithms enjoy a similar improvement in reconstruction quality when they are extended to the vector case. While POCS tackles the interpolation problem from the perspective of convex optimization by alternating between a vector-valued sparsity-promoting projection and an imputation of known vector-valued data samples, MWNI takes an iteratively re-weighted least-squares (IRLS) approach to the problem, requiring Conjugate Gradients (CG) to be adapted to the case of vector-valued data, model, and gradient values. Generalizing CG in this way allows for it to be reused for a wide variety of vector processing problems. Indeed, the solver developed in chapter 3 to extend MWNI to the vector case was later repurposed in chapter 4 to extend acoustic least-squares migration to the elastic case. While this thesis made use of the quaternion Fourier transform to extend Fourier regularization to the vector case, I expect the same approach could be used to similarly extend other processing steps, for example, the linear and parabolic radon transforms.

Regularized least-squares migration has been of great benefit to the imaging of acoustic wavefields, mainly in compensating for irregular illumination of the subsurface, incomplete sampling of sources and receivers, and extending the bandwidth of migrated images. Great motivation to extend least-squares imaging to the elastic case was provided by the fact that multicomponent data suffer from many of these same issues. Similar to acoustic least-squares migration, elastic least-squares migration begins with a single scattering approximation. Small perturbations from the background elastic parameters scatter an incident elastic source wavefield. Of the numerous possible strategies for the propagation of elastic wavefields, this thesis made use of a one-way wave equation propagation for its practicality and efficiency. In this framework, elastic wavefields are propagated recursively, one depth level at a time. Extrapolating from one depth level to the next is achieved by decomposing from displacements into scalar potentials associated with individual elastic wave modes, a vertical phase shift of each mode, followed by the reversal of the decomposition (or recomposition) from scalar potentials to displacements. In chapter 2 these elements are shown to come from eigensolutions of the Christoffel equation. With these linear operators at hand, elastic migration is found to be equivalent to the adjoint of first order shot-profile elastic scattering. Furthermore, making use of the vector-CG algorithm used in chapter 3, the system of equations for elastic modeling can be solved to find a migrated image that best fits the observed data in a least-squares sense. Numerical examples in chapter 4 demonstrate that this approach can improve wavefield separation, mitigate the effects of poor spatial sampling, and compensate for non-uniform illumination for elastic wavefields.

In chapter 5 regularized elastic imaging was applied to an ocean bottom dataset comprised of coarsely sampled receivers with finely sampled sources, contrary to the typical shot-based approach used for elastic imaging. The reciprocity of elastic Green's functions was employed to adapt elastic shot-profile least-squares imaging to elastic receiver-profile least-squares imaging. An important distinction from reciprocity in the acoustic case is the requirement that the source and receiver impulses and responses must also be interchanged. This approach results in a clear improvement over conventional elastic imaging, particularly in the continuity of shallow reflectors and the reduction of wavefield cross-talk artefacts.

This thesis offers two additions to a vector-elastic processing workflow for multicomponent seismic data. First, a vector representation of multicomponent seismic data in the Fourier domain using the quaternion Fourier transform which enabled Fourier regularization to be extended to the multicomponent case. Second, a new method to model elastic data via shot-profile one-way elastic wave equation modeling using an extended model domain. This allows for elastic imaging to be approached as an inverse problem, where the best image is chosen to be the one that minimizes the misfit between observed and predicted data. These two approaches are connected by their use of data-fitting constraints that honor the vector-elastic nature of the data, with secondary constraints that enforce simplicity in the model estimate. In the case of vector interpolation, the optimal solution is one that shares a common level of sparsity for all components, while in the case of regularized elastic imaging the optimal images have uniform illumination.

A robust vector-elastic processing workflow that is suitable for industrial application will likely soon become a reality. Some topics that require future research include the robust estimation and removal of near-surface effects in elastic data, further research into vector noise attenuation strategies, extending multiple attenuation to the multicomponent case (for example Radon demultiple and Surface Related Multiple Attenuation (SRME)), robust wavefield separation for offline processing of elastic data (to derive attributes which can aid in the main vector-elastic processing workflow), and making more use of elastic wave modes that reflection seismology typically treats as noise (for example Raleigh waves). These changes will direct our efforts away from simplifying assumptions and toward treating the full recorded wavefield as useful signal in the prediction of subsurface properties.

Bibliography

Abma, R., and N. Kabir, 2006, 3d interpolation of irregular data with a pocs algorithm: Geophysics, **71**, E91–E97.

Aki, K., and P. G. Richards, 2002, Quantitative seismology: University Science Books, **1**.

Bale, R., 2002, Modelling 3D anisotropic elastic data using the pseudospectral approach: CREWES Research Report, **14**.

Bale, R., and G. Margrave, 2004, Elastic wave-equation migration for laterally varying isotropic and HTI media: Technical report, CREWES research report.

Bale, R., and R. Stewart, 2002, The impact of attenuation on the resolution of multicomponent seismic data: Technical report, CREWES Research Report.

Bale, R. A., 2006, Elastic wave-equation depth migration of seismic data for isotropic and azimuthally anisotropic media: PhD thesis, University of Calgary.

Bale, R. A., J. Li, and B. Mattocks, 2005, Robust estimation of fracture directions from 3D converted waves: SEG Technical Program Expanded Abstracts, Society of Exploration Geophysicists, 889–892.

Bihan, N. L., and J. Mars, 2001, New 2D attributes based on complex and hypercomplex analytic signal: SEG Technical Program Expanded Abstracts, Society of Exploration Geophysicists, 1961–1964.

Biondi, B., 2002, Stable wide-angle fourier finite-difference downward extrapolation of 3-d wavefields: Geophysics, **67**, 872–882.

Boyd, O. S., 2006, An efficient Matlab script to calculate heterogeneous anisotropically elastic wave propagation in three dimensions: Comput. Geosci., **32**, 259–264.

Brown, R. J., R. R. Stewart, and D. C. Lawton, 2002, A proposed polarity standard for multicomponent seismic data: Geophysics, **67**, 1028–1037.

Cary, P., 2011a, True-amplitude ps prestack time migration via 5D interpolation: Presented at the CSEG conference and exhibition, Canadian Society of Exploration Geophysicists.

Cary, P., and M. Perz, 2012, 5d leakage: measuring what 5d interpolation misses: SEG Technical Program Expanded Abstracts, Society of Exploration Geophysicists, 1–5.

Cary, P. W., 1995, A strategy for layer stripping 3D converted wave data: SEG Technical Program Expanded Abstracts, Society of Exploration Geophysicists, 356–359.

———, 1998, P/s wavefield separation in the presence of statics: CREWES Res. Report, **10**, 30–1.

———, 2002, Detecting false indications of shear-wave splitting: SEG Technical Program Expanded Abstracts, Society of Exploration Geophysicists, 1014–1017.

———, 2006, 3-D converted-wave processing in the presence of large S-wave splitting and statics: SEG Technical Program Expanded Abstracts, Society of Exploration Geophysicists, 1203–1207.

———, 2011b, Aliasing and 5d interpolation with the mwni algorithm: SEG Technical Program Expanded Abstracts, Society of Exploration Geophysicists, 3080–3084.

Chang, W.-F., and G. A. McMechan, 1987, Elastic reverse-time migration: Geophysics, **52**, 1365–1375.

Cheng, J., T. Alkhalifah, Z. Wu, P. Zou, and C. Wang, 2016, Simulating propagation of decoupled elastic waves using low-rank approximate mixed-domain integral operators for anisotropic media: Geophysics, **81**, T63–T77.

Chiu, S. K., 2014, Multidimensional interpolation using a model-constrained minimum weighted norm interpolation: Geophysics, **79**, V191–V199.

Claerbout, J. F., 1971, Toward a unified theory of reflector mapping: Geophysics, **36**, 467–481.

———, 1992, Earth soundings analysis: Processing versus inversion: Blackwell Scientific Publications Cambridge, Massachusetts, USA, **6**.

Crampin, S., 1981, A review of wave motion in anisotropic and cracked elastic-media: Wave Motion, **3**, 343 – 391.

Dai, W., and G. T. Schuster, 2013, Plane-wave least-squares reverse-time migration: Geophysics, **78**, S165–S177.

Dellinger, J., and J. Etgen, 1990, Wave field separation in two dimensional anisotropic media: Geophysics, **55**, 914–919.

Dickens, T., and G. Winbow, 2011, Rtm angle gathers using poynting vectors: Presented at the SEG Technical Program Expanded Abstracts, Society of Exploration Geophysicists.

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

———, 2014b, Attenuation compensation for least-squares reverse time migration using the viscoacoustic-wave equation: Geophysics, **79**, S251–S262.

Ell, T., 1992, Hypercomplex spectral transformations: PhD thesis, University of Minnesota.

Ell, T. A., 15-17 Dec 1993, Quaternion-Fourier transforms for analysis of two-dimensional linear time-invariant partial differential systems: Decision and Control, 1993., Proceedings of the 32nd IEEE Conference on, 1830–1841 vol.2.

Ell, T. A., and S. J. Sangwine, Jan. 2007, Hypercomplex Fourier transforms of color images: Image Processing, IEEE Transactions on, **16**, 22–35.

Etgen, J., and C. Regone, 1998, Strike shooting, dip shooting, widepatch shooting – Does prestack migration care? A model study.: 68th Annual Internat. Mtg., Soc. Expl. Geophys., SEG Technical Program Expanded Abstracts, Society of Exploration Geophysicists, 66–69.

Fenati, D., and F. Rocca, 1984, Seismic reciprocity field tests from the italian peninsula: Geophysics, **49**, 1690–1700.

Fletcher, R., D. Nichols, R. Bloor, and R. Coates, 2015, Least-squares migration: data domain versus image domain: Presented at the 77th EAGE Conference and Exhibition 2015, EAGE.

Fletcher, R. P., D. Nichols, R. Bloor, and R. T. Coates, 2016, Least-squares migrationdata domain versus image domain using point spread functions: The Leading Edge, **35**, 157–162.

Fromm, G., T. Krey, and B. Wiest, 1985, Static and dynamic corrections: Handbook of Geophysical Exploration, Geophysical Press, 191–225.

Gaiser, J. E., 2000, Advantages of 3-D PS-wave data to unravel S-wave birefringence for fracture detection: SEG Technical Program Expanded Abstracts, Society of Exploration Geophysicists, 1201–1204.

Gao, J., A. Stanton, M. Naghizadeh, M. D. Sacchi, and X. Chen, 2012, Convergence improvement and noise attenuation considerations for beyond alias projection onto convex sets reconstruction: Geophysical Prospecting, 1–14.

———, 2013, Convergence improvement and noise attenuation considerations for beyond alias projection onto convex sets reconstruction: Geophysical Prospecting, **61**, 138–151.

Garotta, R., 1984, Two-component acquisition as a routine procedure: SEG Technical Program Expanded Abstracts, Society of Exploration Geophysicists, 587–588.

Gazdag, J., 1978, Wave equation migration with the phase-shift method: Geophysics, **43**, 1342–1351.

Gazdag, J., and P. Sguazzero, 1984, Migration of seismic data by phase shift plus interpolation: Geophysics, **49**, 124–131.

Grandi, A., A. Mazzotti, and E. Stucchi, 2007, Multicomponent velocity analysis with quaternions: Geophysical Prospecting, **55**, 761–777.

Gülünay, N., 2003, Seismic trace interpolation in the fourier transform domain: Geophysics, **68**, 355–369.

Guo, Q., and T. Alkhalifah, 2016, Elastic reflection based waveform inversion in isotropic media: Presented at the 78th EAGE Conference and Exhibition, EAGE.

Hamilton, W., 1866, Elements of quaternions: Longmans Green.

Hennenfent, G., and F. J. Herrmann, 2006, Seismic denoising with nonuniformly sampled curvelets: Computing in Science & Engineering, **8**, 16–25.

———, 2008, Simply denoise: Wavefield reconstruction via jittered undersampling: Geophysics, **73**, V19–V28.

Hestenes, M. R., and E. Stiefel, 1952, Methods of conjugate gradients for solving linear systems: NBS, **49**.

Higginbotham, J., M. Brown, and C. Macesanu, 2010, Depth migration velocity model building with wave equation imaging: First Break, **28**, 27–33.

Hindriks, K., and A. J. Duijndam, 1999, Reconstruction of 3-d seismic signals irregularly sampled along two spatial coordinates: Geophysics, **65**, 253–263.

Hoffe, B. H., G. F. Margrave, R. R. Stewart, D. S. Foltinek, H. C. Bland, and P. M. Manning, 2002, Analyzing the effectiveness of receiver arrays for multicomponent seismic exploration: Geophysics, **67**, 1853–1868.

Hu, J., and G. T. Schuster, 1998, Migration deconvolution: SPIE's International Symposium on Optical Science, Engineering, and Instrumentation, International Society for Optics and Photonics, SPIE, 118–124.

Ji, J., 2009, An exact adjoint operation pair in time extrapolation and its application in least-squares reverse-time migration: Geophysics, **74**, H27–H33.

Jin, S., 2010, 5d seismic data regularization by a damped least-norm fourier inversion: Geophysics, **75**, WB103–WB111.

Kaplan, S. T., P. S. Routh, and M. D. Sacchi, 2010, Derivation of forward and adjoint operators for least-squares shot-profile split-step migration: Geophysics, **75**, S225–S235.

Kendall, R., 2006, Advances in land multicomponent seismic: Acquisition, processing, and interpretation: CSEG Recorder, **31**, 65–70.

Knopoff, L., and A. F. Gangi, 1959, Seismic reciprocity: Geophysics, **24**, 681–691.

Kreimer, N., and M. D. Sacchi, 2011, A tensor higher-order singular value decomposition (hosvd) for pre-stack simultaneous noise-reduction and interpolation: SEG Technical Program Expanded Abstracts, Society of Exploration Geophysicists, 3069–3074.

Kühl, H., and M. Sacchi, 2003, Least-squares wave-equation migration for AVP/AVA inversion: Geophysics, **68**, 262–273.

Kuo, J., and T. Dai, 1984, Kirchhoff elastic wave migration for the case of noncoincident source and receiver: Geophysics, **49**, 1223–1238.

Lambaré, G., J. Virieux, R. Madariaga, and S. Jin, 1992, Iterative asymptotic inversion in the acoustic approximation: Geophysics, **57**, 1138–1154.

Li, X., 1998, Fracture detection using PP and PS waves in multicomponent sea-floor data: Presented at the SEG Technical Program Expanded Abstracts, Society of Exploration Geophysicists.

Li, X., P. Cary, and R. Couzens, 2011, The deconvolution of multicomponent trace vectors: Presented at the CSEG conference and exhibition, CSEG.

Li, X., and J. Yuan, 1999, Geophone orientation and coupling in three-component sea-floor data: a case study: Geophysical Prospecting, **47**, 995–1013.

Liu, B., and M. D. Sacchi, 2004, Minimum weighted norm interpolation of seismic records: Geophysics, **69**, 1560–1568.

MacLeod, M., M. Hadley, K. Reynolds, A. Tura, et al., 1999, Multicomponent analysis of obc data: Presented at the Offshore Technology Conference, Offshore Technology Conference.

Menanno, G. M., and A. Mazzotti, 2012, Deconvolution of multicomponent seismic data by means of quaternions: Theory and preliminary results: Geophysical Prospecting, **60**, 217–238.

Miller, G., and H. Pursey, 1955, On the partition of energy between elastic waves in a semi-infinite solid: Proceedings of the Royal Society of London a: mathematical, physical and engineering sciences, The Royal Society, 55–69.

Naghizadeh, M., 2012, Seismic data interpolation and denoising in the frequency-wavenumber domain: Geophysics, **77**, V71–V80.

———, 2015a, Double-weave 3d seismic acquisitionpart 1: Sampling and sparse fourier reconstruction: Geophysics, **80**, WD143–WD162.

———, 2015b, Double-weave 3d seismic acquisitionpart 2: Seismic modeling and subsurface fold analyses: Geophysics, **80**, WD163–WD173.

Naghizadeh, M., and M. Sacchi, 2012, Multicomponent f-x seismic random noise attenuation via vector autoregressive operators: Geophysics, **77**, V91–V99.

Naghizadeh, M., and M. D. Sacchi, 2007a, Multistep autoregressive reconstruction of seismic records: Geophysics, **72**, V111–V118.

———, 2007b, Multistep autoregressive reconstruction of seismic records: Geophysics, **72**, V111–V118.

———, 2010, Seismic data reconstruction using multidimensional prediction filters: Geophysical Prospecting, **58**, 157–173.

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

Øzbek, A., A. K. Øzdemir, and M. Vassallo, 2009, Interpolation by matching pursuit: SEG Technical Program Expanded Abstracts, Society of Exploration Geophysicists, 3254–3258.

Rickett, J. E., 2003, Illumination-based normalization for wave-equation depth migration: Geophysics, **68**, 1371–1379.

Rickett, J. E., and P. C. Sava, 2002, Offset and angle-domain common image-point gathers for shot-profile migration: Geophysics, **67**, 883–889.

Rodriguez, I. V., D. Bonar, and M. Sacchi, 2012, Microseismic data denoising using a 3c group sparsity constrained time-frequency transform: Geophysics, **77**, V21–V29.

Ronen, J., 1987, Wave-equation trace interpolation: Geophysics, **52**, 973–984.

Rosales, D. A., S. Fomel, B. L. Biondi, and P. C. Sava, 2007, Wave-equation angle-domain common-image gathers for converted waves: Geophysics, **73**, S17–S26.

Sacchi, M. D., T. J. Ulrych, and C. J. Walker, 1998, Interpolation and extrapolation using a high-resolution discrete fourier transform: IEEE Transactions on Signal Processing, **46**, 31–38.

Sacchi, M. D., J. Wang, and H. Kuehl, 2007, Estimation of the diagonal of the migration blurring kernel through a stochastic approximation: SEG Technical Program Expanded Abstracts, Society of Exploration Geophysicists, 2437–2441.

Sangwine, S. J., 4-7 Oct 1998, The problem of defining the Fourier transform of a colour image: Image Processing, 1998. ICIP 98. Proceedings. 1998 International Conference on, **1**, 171–175 vol.1.

Sangwine, S. J., and T. A. Ell, 1999, Hypercomplex auto- and cross-correlation of color images: Presented at the International Conference on Image Processing, IEEE.

———, 2000, The discrete Fourier transform of a colour image: Proc. Image Processing II Mathematical Methods, Algorithms and Applications, Chichester, UK, 430–441.

Scales, J., 1987, Tomographic inversion via the conjugate gradient method: Geophysics, **52**, 179–185.

Schonewille, M., A. Klaedtke, and A. Vigner, 2009, Anti-alias anti-leakage fourier transform: SEG Technical Program Expanded Abstracts, Society of Exploration Geophysicists, 3249–3253.

Schwarz, U., 1978, Mathematical-statistical description of the iterative beam removing technique (method clean): Astronomy and Astrophysics, **65**, 345.

Shan, G., and B. Biondi, 2008, Plane-wave migration in tilted coordinates: Geophysics, **73**, S185–S194.

Silver, P. G., and W. W. Chan, 1991, Shear wave splitting and subcontinental mantle deformation: Journal of Geophysical Research: Solid Earth, **96**, 16429–16454.

Simmons, J. L., 2009, Converted-wave splitting estimation and compensation: Geophysics, **74**, D37–D48.

Spitz, S., 1991, Seismic trace interpolation in the fx domain: Geophysics, **56**, 785–794.

Stanton, A., N. Kazemi, and M. D. Sacchi, 2013, Processing seismic data in the presence of residual statics: SEG Technical Program Expanded Abstracts, Society of Exploration Geophysicists, 1838–1842.

Stanton, A., N. Kreimer, D. Bonar, M. Naghizadeh, and M. Sacchi, 2012, A comparison of 5d reconstruction methods: SEG Technical Program Expanded Abstracts, Society of Exploration Geophysicists, 1–5.

Stanton, A., and M. Sacchi, 2011, Multicomponent seismic data reconstruction using the quaternion fourier transform and POCS: SEG Technical Program Expanded Abstracts, Society of Exploration Geophysicists, 1267–1272.

———, 2012, Five-dimensional vector data interpolation: Presented at the 74th EAGE Conference and Exhibition incorporating EUROPEC 2012, EAGE.

———, 2015, Least squares wave equation migration of elastic data: Presented at the 77th EAGE Conference and Exhibition 2015, EAGE.

Stanton, A., and M. D. Sacchi, 2013, Vector reconstruction of multicomponent seismic data: Geophysics, **78**, V131–V145.

———, 2017, Elastic least-squares one-way wave equation migration: Geophysics, **82**.

Stanton, A., M. D. Sacchi, R. Abma, and J. A. Stein, 2015, Mitigating artifacts in projection onto convex sets interpolation: SEG Technical Program Expanded Abstracts, Society of Exploration Geophysicists, 3779–3783.

Stanton, A., M. D. Sacchi, and N. Kazemi, 2016, Seismic data analysis in julia: Presented at the 78th EAGE Conference and Exhibition 2016-Workshops, EAGE.

Starck, J.-L., E. J. Candès, and D. L. Donoho, 2002, The curvelet transform for image denoising: IEEE Transactions on image processing, **11**, 670–684.

Stein, J. A., S. Boyer, K. Hellman, and J. Weigant, 2010, Application of POCS interpolation to exploration: SEG Technical Program Expanded Abstracts, Society of Exploration Geophysicists, 3565–3568.

Stewart, R. R., 2009, The measure of full motion: Multicomponent seismic exploration and its value: SEG Technical Program Expanded Abstracts, Society of Exploration Geophysicists, 4243–4248.

Stewart, R. R., J. E. Gaiser, R. J. Brown, and D. C. Lawton, 2002, Converted-wave seismic exploration: Methods: Geophysics, **67**, 1348–1363.

———, 2003, Converted-wave seismic exploration: Applications: Geophysics, **68**, 40–57.

Stoffa, P. L., J. T. Fokkema, R. M. de Luna Freire, and W. P. Kessinger, 1990, Split-step fourier migration: Geophysics, **55**, 410–421.

Stolt, R. H., and A. B. Weglein, 2012, Seismic imaging and inversion: Volume 1: Application of linear inverse theory: Cambridge University Press.

Sun, R., and G. A. McMechan, 1986, Pre-stack reverse-time migration for elastic waves with application to synthetic offset vertical seismic profiles: Proceedings of the IEEE, **74**, 457–465.

———, 2001, Scalar reverse-time depth migration of prestack elastic seismic data: Geophysics, **66**, 1519–1527.

Sun, R., G. A. McMechan, and H.-H. Chuang, 2011, Amplitude balancing in separating P- and S-waves in 2D and 3D elastic seismic data: Geophysics, **76**, S103–S113.

Tarantola, A., 1986, A strategy for nonlinear elastic inversion of seismic reflection data: Geophysics, **51**, 1893–1903.

———, 1987, Inverse problem theory. Methods for data fitting and model parameter estimation: Elsevier.

Trad, D., 2009, Five-dimensional interpolation: Recovering from acquisition constraints: Geophysics, **74**, V123–V132.

Trickett, S., L. Burroughs, A. Milton, L. Walton, and R. Dack, 2010, Rank-reduction-based trace interpolation: SEG Technical Program Expanded Abstracts, Society of Exploration Geophysicists, 3829–3833.

Tsvankin, I., 2012, Seismic signatures and analysis of reflection data in anisotropic media, 3 ed.: Society of Exploration Geophysicists.

van Der Baan, M., 2006, Pp/ps wavefield separation by independent component analysis: Geophysical Journal International, **166**, 339–348.

Vassallo, M., A. Özbek, K. Özdemir, and K. Eggenberger, 2010, Crossline wavefield reconstruction from multicomponent streamer data: Part 1multichannel interpolation by matching pursuit (mimap) using pressure and its crossline gradient: Geophysics, **75**, WB53–WB67.

Wang, J., 2005, Least-squares wave-equation AVP/AVA migration of common azimuth data: PhD thesis, University of Alberta.

Wang, J., H. Kuehl, and M. Sacchi, 2005, High-resolution wave-equation ava imaging: Algorithm and tests with a data set from the western canadian sedimentary basin: Geophysics, **70**, S91–S99.

Wang, J., M. Ng, and M. Perz, 2010a, Seismic data interpolation by greedy local radon transform: Geophysics, **75**, WB225–WB234.

Wang, J., and M. D. Sacchi, 2007, High-resolution wave-equation amplitude-variation-withray-parameter (avp) imaging with sparseness constraints: Geophysics, 11–18.

———, 2009, Structure constrained least-squares migration: SEG Technical Program Expanded Abstracts, Society of Exploration Geophysicists, 2763–2767.

Wang, P., A. Gomes, Z. Zhang, and M. Wang, 2016, Least-squares rtm: Reality and possibilities for subsalt imaging: SEG Technical Program Expanded Abstracts, Society of Exploration Geophysicists, 4204–4209.

Wang, S.-Q., X. Gao, and Z.-X. Yao, 2010b, Accelerating pocs interpolation of 3d irregular seismic data with graphics processing units: Computers & Geosciences, **36**, 1292–1300.

———, 2010c, Accelerating POCS interpolation of 3D irregular seismic data with graphics processing units: Comput. Geosci., **36**, no. 10, 1292–1300.

Wapenaar, C., H. Cox, and A. Berkhout, 1992, Elastic redatuming of multicomponent seismic data1: Geophysical prospecting, **40**, 465–482.

Wikel, K., R. Kendall, R. Bale, J. Grossman, and K. DeMeersman, 2012, 4d-3c geomechanical study of in-situ bitumen recovery in nw canada using toe-to-heel air injection: first break, **30**, 55–65.

Witten, B., and J. Shragge, 2006, Quaternion-based signal processing: SEG Technical Program Expanded Abstracts, Society of Exploration Geophysicists, 2862–2866.

Wong, M., B. L. Biondi, and S. Ronen, 2015, Imaging with primaries and free-surface multiples by joint least-squares reverse time migration: Geophysics, **80**, S223–S235.

Xie, X., and R. Wu, 2005, Multicomponent prestack depth migration using the elastic screen method: Geophysics, **70**, S30–S37.

Xu, S., Y. Zhang, D. Pham, and G. Lambaré, 2005, Antileakage fourier transform for seismic data regularization: Geophysics, **70**, V87–V95.

Yan, J., and W. Ross, 2013, Improving the stability of angle gather computation using poynting vectors: Presented at the SEG Technical Program Expanded Abstracts, Society of Exploration Geophysicists.

Yan, J., and P. Sava, 2008a, Isotropic angle-domain elastic reverse-time migration: Geophysics, **73**, S229–S239.

———, 2008b, Isotropic angle-domain elastic reverse-time migration: Geophysics, **73**, S229–S239.

Yoon, K., and K. J. Marfurt, 2006, Reverse-time migration using the poynting vector: Exploration Geophysics, **37**, 102–107.

Zhang, Q., H. Zhou, H. Chen, and J. Wang, 2016, Least-squares reverse time migration with and without source wavelet estimation: Journal of Applied Geophysics, **134**, 1–10.

Zhang, Y., L. Duan, and Y. Xie, 2015, A stable and practical implementation of least-squares reverse time migration: Geophysics, **80**, V23–V31.

Zwartjes, P., and A. Gisolf, 2007, Fourier reconstruction with sparse inversion: Geophysical Prospecting, **55**, 199–221.

APPENDIX A

The non-orthonormality of the composition/ decomposition operators

The composition operator for elastic wave propagation gets its 3 columns from the eigenvectors of the Christoffel equation. As such, we expect this matrix to be orthonormal, with the inverse (decomposition operator) being simply the adjoint of the composition operator. Numerically this computation is complicated by the fact that the vertical wavenumber must be computed from the material parameters corresponding to each wave mode and the horizontal wavenumbers via the dispersion relation (Bale, 2006). This requires the columns of the decomposition operator to be individually normalized, making the composition operator non-orthogonal, and requiring the inverse (decomposition operator) to be computed[1].

Here I derive the normalized decomposition operator taking into account the non-orthogonality of the columns of the composition matrix. Starting with the normalized composition operator

$$\mathbf{\Gamma} = \begin{bmatrix} \boldsymbol{\gamma}_1 & \boldsymbol{\gamma}_2 & \boldsymbol{\gamma}_3 \end{bmatrix} = \begin{bmatrix} \frac{k_x}{a} & \frac{-k_y}{b} & \frac{-k_x k_z(s)}{c} \\ \frac{k_y}{a} & \frac{k_x}{b} & \frac{-k_y k_z(s)}{c} \\ \frac{k_z(p)}{a} & 0 & \frac{b^2}{c} \end{bmatrix}, \tag{A.1}$$

where $a^2 = k_x^2 + k_y^2 + k_z(p)^2$, $b^2 = k_x^2 + k_y^2$, and $c^2 = k_x^2 k_z(s)^2 + k_y^2 k_z(s)^2 + b^4$. The inner product $< \boldsymbol{\gamma}_1, \boldsymbol{\gamma}_3 >= \frac{k_z(p)b^2}{ac} - \frac{k_z(s)b^2}{ac} = 0 \iff v_p = v_s$. Clearly the columns of $\mathbf{\Gamma}$ are linearly dependent making the matrix non-orthonormal. Indeed, the determinant

---

[1] This is not an issue when decomposing wavefields during time extrapolation (such as in anisotropic elastic Reverse Time Migration). In this case, within each time step, the spatial wavenumbers are computed independently and the orthonormality of the composition/decomposition operators remains valid. This is perhaps another argument in favor of extrapolation stepping in time rather than in depth.

$|\mathbf{\Gamma}| = \frac{bd^2}{ac} \neq 1$, where $d^2 = k_x^2 + k_y^2 + k_z(p)k_z(s)$. Rather than using the adjoint, the decomposition operator is found by

$$\mathbf{\Gamma}^{-1} = \frac{1}{|\mathbf{\Gamma}|} adj(\mathbf{\Gamma}) \tag{A.2}$$

which is found to be

$$\mathbf{\Gamma}^{-1} = \frac{1}{d^2} \begin{bmatrix} k_x a & k_y a & k_z(s)a \\ \frac{-k_y d^2}{b} & \frac{k_x d^2}{b} & 0 \\ \frac{-k_x k_z(p)c}{b^2} & \frac{-k_y k_z(p)c}{b^2} & c \end{bmatrix}. \tag{A.3}$$

The adjoint of the elastic scattering operator requires the adjoints of $\mathbf{\Gamma}$ and $\mathbf{\Gamma}^{-1}$. The adjoint of the composition operator (used to decompose the receiver side wavefield for a given depth step in elastic migration) is given by

$$\mathbf{\Gamma}^{\dagger} = \begin{bmatrix} \frac{k_x}{a} & \frac{k_y}{a} & \frac{k_z(p)}{a} \\ \frac{-k_y}{a} & \frac{k_x}{b} & 0 \\ \frac{-k_x k_z(s)}{c} & \frac{-k_y k_z(s)}{c} & \frac{b^2}{c} \end{bmatrix}, \tag{A.4}$$

while the adjoint of the decomposition operator (used to compose the receiver side wavefield for a given depth step in elastic migration) is given by

$$(\mathbf{\Gamma}^{-1})^{\dagger} = \frac{1}{d^2} \begin{bmatrix} k_x a & \frac{-k_y d^2}{b} & \frac{-k_x k_z(p)c}{b^2} \\ k_y a & \frac{k_x d^2}{b} & \frac{-k_y k_z(p)c}{b^2} \\ k_z(s)a & 0 & c \end{bmatrix}. \tag{A.5}$$

APPENDIX B

CG for multicomponent model and data vectors

Conjugate Gradients (CG) is a numerical method to estimate a solution vector, $\boldsymbol{x}$ from the forward problem $\boldsymbol{Ax} + \boldsymbol{noise} = \boldsymbol{b}$. Supposing $\boldsymbol{A}$ is a mxn positive-definite matrix, $\boldsymbol{x}$ is a model vector of length n, and $\boldsymbol{b}$ is a data vector of length m, we seek the minimum norm solution via the following cost function:

$$J = ||\boldsymbol{Ax} - \boldsymbol{b}||_2^2 + \mu ||\boldsymbol{x}||_2^2, \tag{B.1}$$

Where μ is a tradeoff parameter controlling the level of data fitting versus the influence of the norm of the model. Algebraically the solution to this problem is determined by taking the derivative of the cost function and setting it to zero giving

$$\boldsymbol{x} = (\boldsymbol{A}^T\boldsymbol{A} + \mu \boldsymbol{I})^{-1}\boldsymbol{A}^T\boldsymbol{b}. \tag{B.2}$$

Rather than solving the problem algebraically, it is often more practical to solve the problem using an iterative method such as CG outlined in algorithm 1 (Hestenes and Stiefel, 1952; Scales, 1987).

It should be noted that the linear operator, $\boldsymbol{A}$, and its adjoint $\boldsymbol{A}^T$, are generally not matrices, although it is theoretically possible for them to be. In the many cases it is ill advised to express a linear operator as a matrix because it is either A) enormous in size or B) filled with redundant entries or zeros. Typically we take a linear operator $\boldsymbol{A}$, and its adjoint $\boldsymbol{A}^T$ to mean an efficient pair of computer programs that have been carefully designed to pass an adjointness, or *dot-product*, test that passes within machine precision.

Algorithm 1 Conjugate Gradients

```
$\boldsymbol{d} \leftarrow$ observed data vector
$\mu \leftarrow$ tradeoff parameter
$N_{iter} \leftarrow$ number of iterations
$tol \leftarrow$ tolerance to exit program early
$\boldsymbol{r} = \boldsymbol{d}$
$\boldsymbol{s} = \boldsymbol{g} = \boldsymbol{A}^T\boldsymbol{r}$
$\boldsymbol{x} = \boldsymbol{0}$
$\gamma = \boldsymbol{g}^T\boldsymbol{g}$
$\gamma_0 = \gamma$
$cost_0 = \boldsymbol{r}^T\boldsymbol{r}$
for $j = 1 : N_{iter}$ do
  $\boldsymbol{t} = \boldsymbol{A}\boldsymbol{s}$
  $\Delta = \boldsymbol{t}^T\boldsymbol{t} + \mu\boldsymbol{s}^T\boldsymbol{s}$
  if $\Delta <= tol$ then
    BREAK
  end if
  $\alpha = \gamma/\Delta$
  $\boldsymbol{x} = \boldsymbol{x} + \alpha\boldsymbol{s}$
  $\boldsymbol{r} = \boldsymbol{r} - \alpha\boldsymbol{t}$
  $\boldsymbol{g} = \boldsymbol{A}^T\boldsymbol{r}$
  $\boldsymbol{g} = \boldsymbol{g} - \mu\boldsymbol{x}$
  $\gamma_0 = \gamma$
  $\gamma = \boldsymbol{g}^T\boldsymbol{g}$
  $cost = \boldsymbol{r}^T\boldsymbol{r} + \mu\boldsymbol{x}^T\boldsymbol{x}$
  $\beta = \gamma/\gamma_0$
  $\boldsymbol{s} = \beta\boldsymbol{s} + \boldsymbol{g}$
  if $\sqrt{\gamma} <= \sqrt{\gamma_0}tol$ then
    BREAK
  end if
end for
return $\boldsymbol{x}$
```

Notice the inner products in algorithm 1 are collapsing vectors into scalar quantities (effectively a measure of length). To extend this algorithm to the case of multicomponent model and data vectors we require the linear operators $\boldsymbol{A}$ and $\boldsymbol{A}^T$ to act on multicomponent-valued model and data vectors respectively. We also need the inner product calculations to collapse over individual components. Finally, to control the influence of individual vector norms it is useful to incorporate a multicomponent-valued trade-off parameter. Algorithm 2 outlines the changes necessary for CG to operate on multicomponent-valued model and data vectors. This *multicomponent conjugate gradients* algorithm could find many useful applications in multicomponent seismic data processing. In chapter 3 of this thesis I use it to interpolate multicomponent seismic records. In chapters 4 and 5 of this thesis I use it to perform elastic least squares migration. More generally, this algorithm could be used in any linear inversion in which the input data and output models are multicomponent or multiparameter in nature.

Algorithm 2 Conjugate Gradients for multicomponent model and data vectors

```
$\vec{\boldsymbol{d}} \leftarrow$ observed multicomponent-valued data vector
$\mu_1, \mu_2, \mu_3 \leftarrow$ tradeoff parameters for components 1, 2 and 3
$N_{iter} \leftarrow$ number of iterations
$tol \leftarrow$ tolerance to exit program early
$\vec{\boldsymbol{r}} = \vec{\boldsymbol{d}}$
$\vec{\boldsymbol{s}} = \vec{\boldsymbol{g}} = \boldsymbol{A}^T \vec{\boldsymbol{r}}$
$\vec{\boldsymbol{x}} = \vec{\boldsymbol{0}}$
$\gamma = \boldsymbol{g}_1^T \boldsymbol{g}_1 + \boldsymbol{g}_2^T \boldsymbol{g}_2 + \boldsymbol{g}_3^T \boldsymbol{g}_3$
$\gamma_0 = \gamma$
$cost_0 = \boldsymbol{r}_1^T \boldsymbol{r}_1 + \boldsymbol{r}_2^T \boldsymbol{r}_2 + \boldsymbol{r}_3^T \boldsymbol{r}_3$
for $j = 1 : N_{iter}$ do
  $\vec{\boldsymbol{t}} = \boldsymbol{A} \vec{\boldsymbol{s}}$
  $\Delta = \boldsymbol{t}_1^T \boldsymbol{t}_1 + \boldsymbol{t}_2^T \boldsymbol{t}_2 + \boldsymbol{t}_3^T \boldsymbol{t}_3 + \mu_1 \boldsymbol{s}_1^T \boldsymbol{s}_1 + \mu_2 \boldsymbol{s}_2^T \boldsymbol{s}_2 + \mu_3 \boldsymbol{s}_3^T \boldsymbol{s}_3$
  if $\Delta <= tol$ then
    BREAK
  end if
  $\alpha = \gamma / \Delta$
  $\vec{\boldsymbol{x}} = \vec{\boldsymbol{x}} + \alpha \vec{\boldsymbol{s}}$
  $\vec{\boldsymbol{r}} = \vec{\boldsymbol{r}} - \alpha \vec{\boldsymbol{t}}$
  $\vec{\boldsymbol{g}} = \boldsymbol{A}^T \vec{\boldsymbol{r}}$
  $\vec{\boldsymbol{g}} = \vec{\boldsymbol{g}} - \vec{\mu} \cdot \vec{\boldsymbol{x}}$
  $\gamma_0 = \gamma$
  $\gamma = \boldsymbol{g}_1^T \boldsymbol{g}_1 + \boldsymbol{g}_2^T \boldsymbol{g}_2 + \boldsymbol{g}_3^T \boldsymbol{g}_3$
  $cost = \boldsymbol{r}_1^T \boldsymbol{r}_1 + \boldsymbol{r}_2^T \boldsymbol{r}_2 + \boldsymbol{r}_3^T \boldsymbol{r}_3 + \mu_1 \boldsymbol{x}_1^T \boldsymbol{x}_1 + \mu_2 \boldsymbol{x}_2^T \boldsymbol{x}_2 + \mu_3 \boldsymbol{x}_3^T \boldsymbol{x}_3$
  $\beta = \gamma / \gamma_0$
  $\vec{\boldsymbol{s}} = \beta \vec{\boldsymbol{s}} + \vec{\boldsymbol{g}}$
  if $\sqrt{\gamma} <= \sqrt{\gamma_0} tol$ then
    BREAK
  end if
end for
return $\vec{\boldsymbol{x}}$
```

APPENDIX C

Angle gathers in wave equation migration

Angle gathers find many useful applications in seismic imaging. Migration velocity analysis is often carried out in an iterative layer-stripping where angle gathers are progressively flattened starting from shallow events and working downward. After final migration, gathers can be muted to mitigate the effects of imaging artifacts or to produce angle stacks for interpretation. In least-squares migration angle gathers provide an extended axis to regularize the inversion and improve convergence.

A popular technique to calculate angle gathers is to measure the direction of energy flow for the propagating source and receiver-side wavefields, followed by a calculation of the opening angle between these vectors (Yoon and Marfurt, 2006). This is known as the Poynting vector approach. While this approach is efficient and provides high-resolution angle estimates, it often suffers from artefacts that can interfere with convergence in least-squares migration.

An alternative approach employs an extended imaging condition (in time, subsurface offset, depth, or a combination of axes) followed by slant stack to convert to angle (Rickett and Sava, 2002). This technique is expensive because it requires an additional ”for loop” over spatial lags in the very inner loop of the imaging algorithm, but provides a smooth artefact-free extended that is effective for use in least-squares migration.

This appendix outlines the implementation of these two strategies.

C.1 Poynting vector to angle formulation

Poynting vectors are calculated via

$$S_i = -\tau_{ij}\dot{u}_j \tag{C.1}$$

where $\boldsymbol{u}$ is the particle displacement and τ_{ij} is the stress tensor (Dickens and Winbow, 2011), which can be approximated by

$$\boldsymbol{S} \approx \triangledown P \frac{\partial P}{\partial t} P \tag{C.2}$$

where P is the pressure (Yoon and Marfurt, 2006). The steps to evaluate the derivatives in equation 2 for shot profile wave equation migration are as follows (Higginbotham et al., 2010):

1. calculate gradient components for the source side wavefield: $U^s_x(w, k_x, k_z) = k_x U^s(w, k_x, k_z)$
2. inverse Fourier transform over the spatial axes (x and z)
3. obtain $U^s_x(x, z)$ that corresponds to the time of reflection by calculating the zero-lag cross correlation with the receiver wavefield: $U^s_x(x, z) = \int U^s_x(w, x, z) U^{g^*}(w, x, z) d\omega$.
4. normalize the elements of $U^g(x, z) = U^g_x(x, z)\hat{\boldsymbol{x}} + U^g_z(x, z)\hat{\boldsymbol{z}}$
5. repeat steps 1-4 for the z-component of the source side wavefield to obtain $U^g_z(x, z)$

These steps are repeated for the x and z components of the receiver side wavefield to obtain the unit vectors: $\hat{\boldsymbol{U}}^s(x, z)$, and $\hat{\boldsymbol{U}}^g(x, z)$. In practice we find that we must multiply $U^s_x(x, z)$ by -1 to obtain unit vectors following the convention shown in figure C.1.

Using these vectors we may calculate the angle of incidence with respect to the vertical:

$$\alpha(x, z) = tan^{-1} \left(\frac{\hat{U}^s_x(x, z)}{\hat{U}^s_z(x, z)} \right) \tag{C.3}$$

or with respect to the reflector normal:

$$\theta(x, z) = \frac{1}{2} cos^{-1} \left(\hat{\boldsymbol{U}}^s(x, z) \cdot \hat{\boldsymbol{U}}^g(x, z) \right). \tag{C.4}$$

For 2D imaging the signed incidence angle can be found by multiplying by the sign of the cross product:

$$\theta(x, z) = sgn \left(\hat{\boldsymbol{U}}^s(x, z) \times \hat{\boldsymbol{U}}^g(x, z) \right) \frac{1}{2} cos^{-1} \left(\hat{\boldsymbol{U}}^s(x, z) \cdot \hat{\boldsymbol{U}}^g(x, z) \right) \tag{C.5}$$

while for 3D imaging, we can compute the reflection azimuth (Dickens and Winbow, 2011).

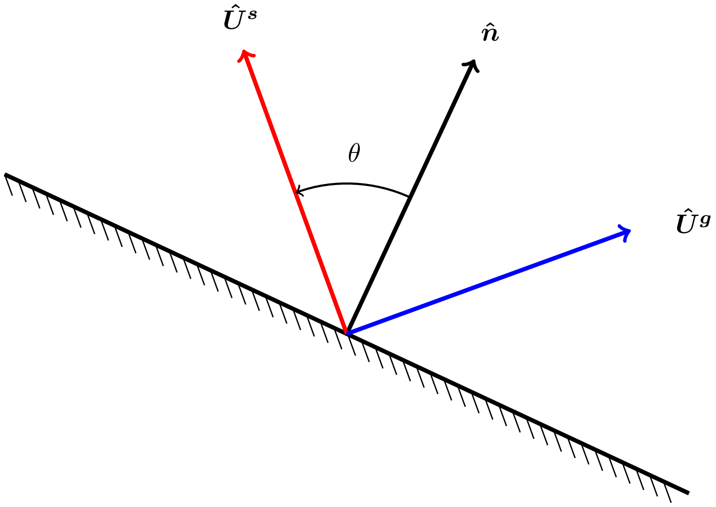


Figure C.1: Source and receiver wavefield Poynting vector conventions used when compute opening angle, θ, with resect to reflector normal, $\vec{\boldsymbol{n}}$.

These calculations often lead to noisy angles because of the normalization to unit vectors and the complicated receiver side wavefield. Yan and Ross (2013) propose a smooth solution by calculating an angle that minimizes a cost function for a small window. A straightforward alternative is to avoid the use of opening angle, and to define the angle solely using source wavefield gradients. This provides smooth source side incidence angles defined with respect to vertical.

After angles have been computed for a migrated shot gather, the shot is linearly interpolated onto a regular grid of angle bins in the final image

$$m(\boldsymbol{x})(\lfloor\frac{\theta(\boldsymbol{x})-\theta_o}{\Delta\theta}\rfloor])+=(1-\alpha(\boldsymbol{x}))m_{1shot}(\boldsymbol{x}) \tag{C.6}$$

$$m(\boldsymbol{x})(\lfloor\frac{\theta(\boldsymbol{x})-\theta_o}{\Delta\theta}\rfloor+1)+=\alpha(\boldsymbol{x})m_{1shot}(\boldsymbol{x}) \tag{C.7}$$

where $\alpha(\boldsymbol{x})=(\theta(\boldsymbol{x})-\lfloor\frac{\theta(\boldsymbol{x})-\theta_o}{\Delta\theta}\rfloor\Delta\theta-\theta_o)/\Delta\theta$. Conversely, for the forward operator the migrated image is sprayed to each migrated shot (prior to demigration) using

$$m_{1shot}(\boldsymbol{x})=(1-\alpha(\boldsymbol{x}))m(\boldsymbol{x})(\lfloor\frac{\theta(\boldsymbol{x})-\theta_o}{\Delta\theta}\rfloor])+\alpha(\boldsymbol{x})m(\boldsymbol{x})(\lfloor\frac{\theta(\boldsymbol{x})-\theta_o}{\Delta\theta}\rfloor+1) \tag{C.8}$$

Equations 10 and 11 form an adjoint/forward operator pair that allows for least-squares migration using Poynting vectors.

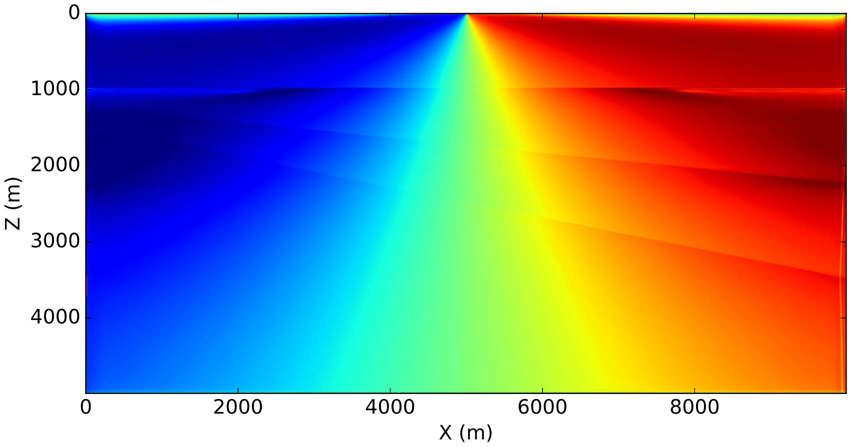


Figure C.2: Propagation angles for a source at X=5000m, computed using the Poynting vector method.

Using the velocity model from figure 2.3, and data shown in figure 2.4 as input, the Poynting vector method of computing angle gathers begins with the computation of propagation angles by downward continuing the source wavelet and computing its spatial gradients. The result of this process is shown in figure C.2.

Next, the angles are used to bin individually migrated shot gathers into angle bins via the linear interpolation given in equation C.7. The results of this process are shown in figure C.3, where source side incidence angles between -10.0 and 10.0 degrees from vertical have been stacked to generate a near angle image.

An angle gather at $X = 5000m$, directly underlying the source position, is shown in figure C.4

From this figure it is apparent that while the Poynting vector method creates high-resolution angle gathers, there is a tradeoff with numerical instability, creating high amplitude artefacts. The next section outlines an alternative method to compute angle gathers that produces fewer artefacts.

C.2 Subsurface offset to angle formulation

The subsurface offset to angle method uses the Fourier transform to map from subsurface offset to the half opening angle, γ, via

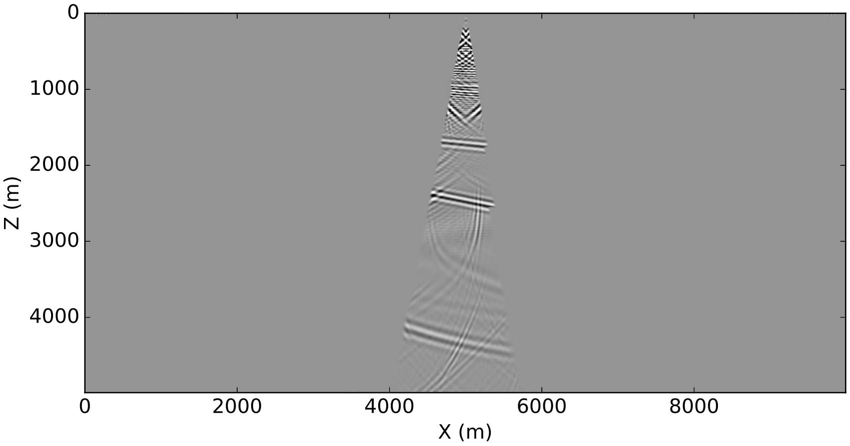


Figure C.3: Near angle stack from the Poynting vector method using source side incidence angles between -10.0 and 10.0 degrees from vertical.

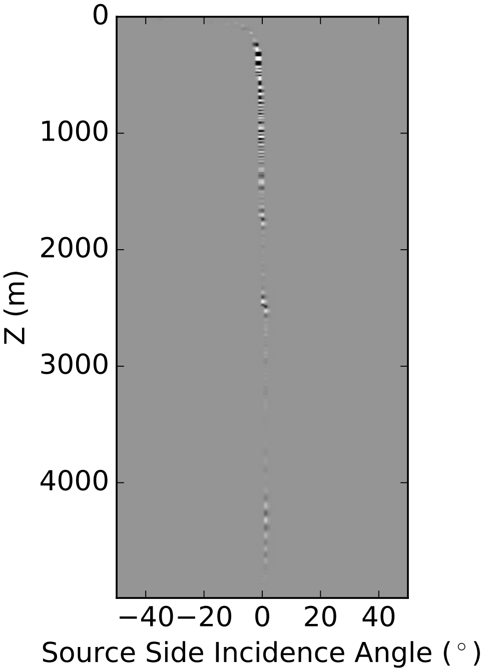


Figure C.4: Angle gather at $X = 5000m$ generated for a single shot via the subsurface offset method. While the resolution of the angles and depths is slightly lower than for the equivalent Poynting vector angle gather (figure), the artefacts are fewer.

$$tan\,\gamma = \frac{k_h}{k_z} \tag{C.9}$$

where k_h, and k_z are the subsurface offset and depth wavenumbers respectively.

Incorporating angle dependent reflectivity into the adjoint operator, $\boldsymbol{L}^{\dagger}$, involves

1. computing the imaging condition for discrete spatial lags between the source and receiver wavefields
2. taking the Fourier transform of the subsurface offset image along the depth and offset axes
3. performing the mapping from subsurface offset wavenumber to angle via equation C.9
4. taking the inverse Fourier transform along the depth axis

while the forward operator, $\boldsymbol{L}$, involves

1. taking the Fourier transform of the image along the depth axis
2. performing the adjoint mapping from angle to subsurface offset wavenumber via equation C.9
3. taking the inverse Fourier transform of the subsurface offset image along the depth and offset axes
4. computing the receiver wavefields at depth for discrete spatial lags between the source and the image point

The above method parameterizes the images as a function of the half opening angle between the source and receiver wavefields. For PP reflections the equivalence of source and receiver side velocity means the opening angle is bisected by the reflector normal vector, making the half opening angle equivalent to the source side incidence angle with respect to reflector normal. For PS reflections there is no such equivalence, and the half opening angle does not adequately describe the zero incidence angle where polarity reversal is expected (except for the special case of a flat-lying reflector). While a post-processing step can be used to correctly position the polarity reversal for PS reflections prior to stacking (Rosales et al., 2007), we observe that simply using the half opening angle provides an adequate extended imaging condition for elastic least-squares migration. A positive feature of this method is that it provides a smooth artefact free extended axis. Considering that angle gathers should be flat if an accurate migration velocity is used, this axis is effective to precondition least-squares migration by removing dipping energy and noise.

Using the velocity model from figure 2.3, and data shown in figure 2.4 as input, the subsurface offset method generates laterally lagged images. Three such images are shown in figure C.5. These lagged images are then transformed to angle via equation C.9.

A near angle stack generated via the subsurface offset method is shown in figure C.6, while an angle gather at $X = 5000m$ is shown in figure C.7.

Comparing the near angles section created using the Poynting vector method (figure C.3) with the section created using the subsurface offset method (figure C.6), the subsurface method results in a section with fewer artefacts. The half opening angle and the source side propagation angle with respect to vertical are equivalent for flat reflectors. Because of this, the events in figures C.3 and C.6 align for the first (flat) event, however, later dipping events do not align due to the difference in the angle definitions.

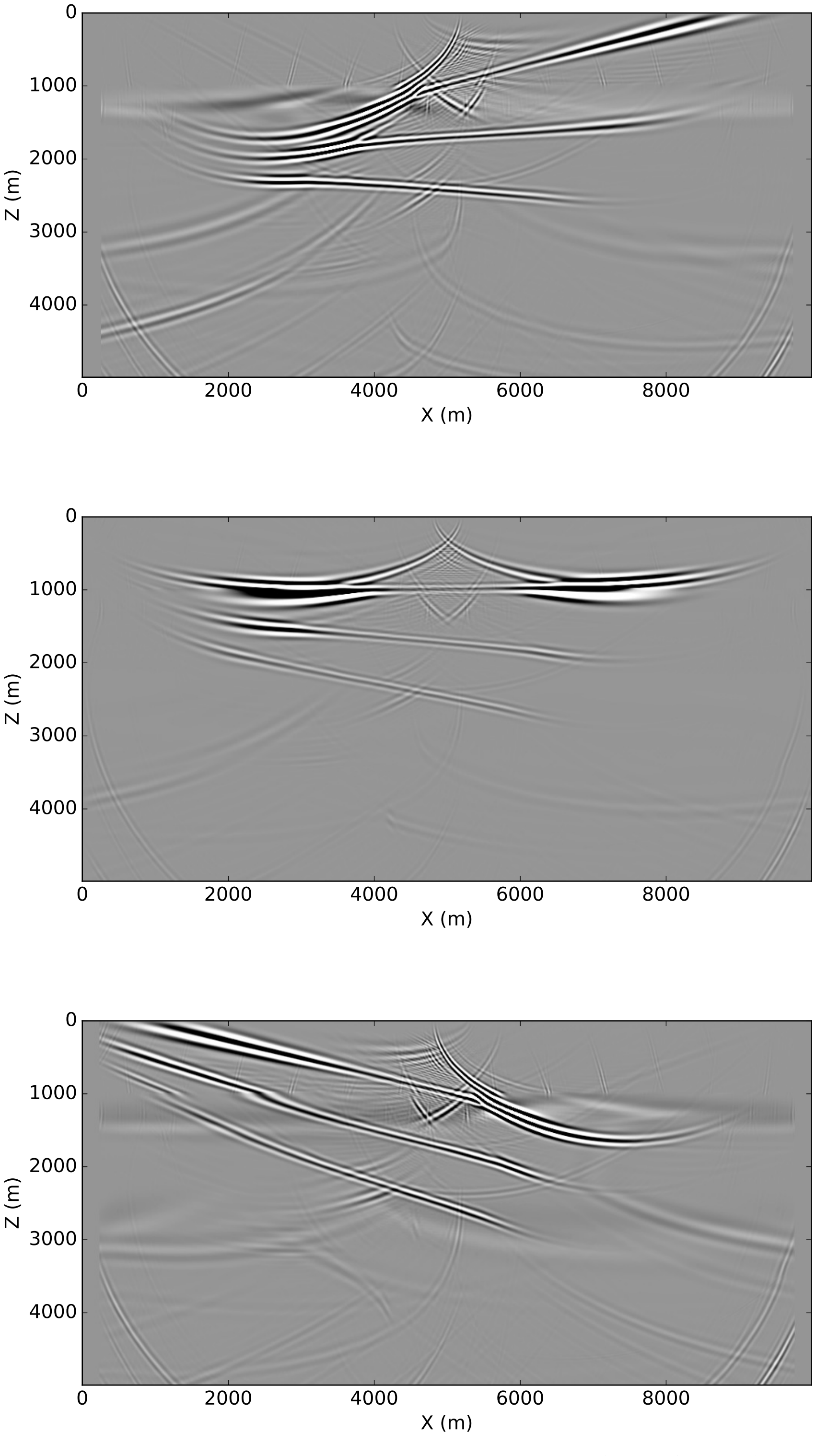


Figure C.5: Image for a subsurface offsets of -250m (top), 0m (middle), and +250m (bottom).

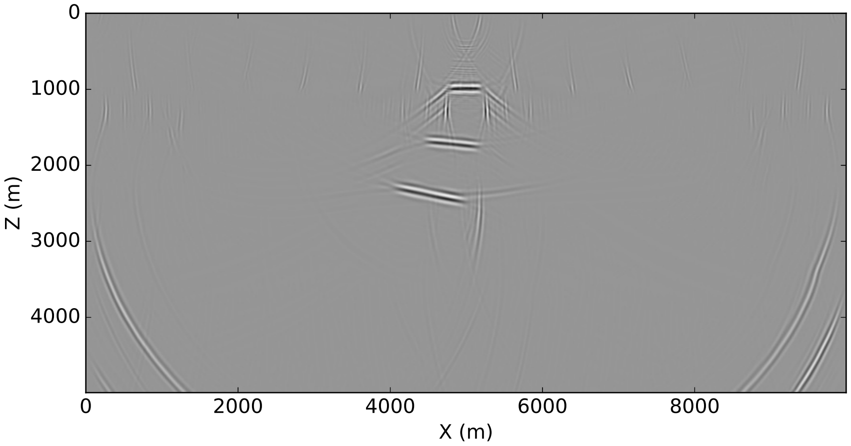


Figure C.6: Near angle stack from the subsurface offset method using half-opening angles from γ = -10.0 to 10.0 degrees.

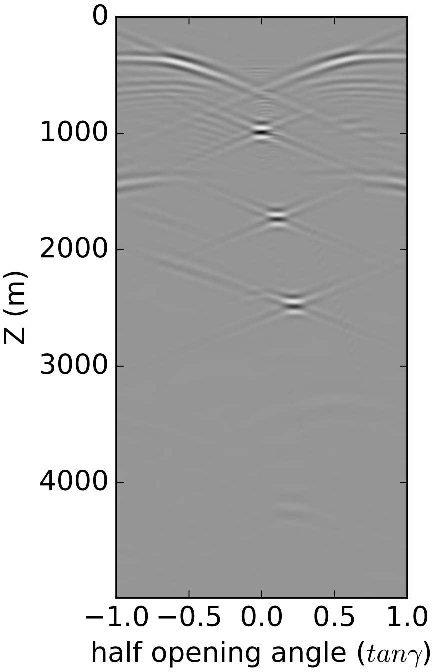


Figure C.7: Angle gather at $X = 5000m$ generated for a single shot via the subsurface offset method. While the resolution of the angles and depths is slightly lower than for the equivalent Poynting vector angle gather (figure), the artefacts are of lower amplitude.

APPENDIX D

Seismic.jl: seismic data analysis in Julia[1].

Geophysical programming typically falls into one of two categories: small-scale prototyping for teaching or research using high-level languages (e.g. Python, Matlab, or Mathematica) or large-scale implementations for production applications or large numerical experiments using statically compiled languages (e.g. C, C++, Fortran, or Java). Julia offers a bridge between these two categories. In Julia we have the ability to quickly create prototypes that can also solve larger problems efficiently. It does this by using a sophisticated type system and multiple dispatch (specified function behavior for various combinations of argument types). The language design allows Julia's just-in-time (JIT) compiler to execute programs at impressive speeds. Below is a table provided by the creators of Julia that benchmarks the performance of several popular languages for various algorithms.

	Fortran	Julia	Python	R	Matlab	Octave	Mathematica	JavaScript	Go	LuaJIT	Java
fib	0.70	2.11	77.76	533.52	26.89	9324.35	118.53	3.36	1.86	1.71	1.21
parse_int	5.05	1.45	17.02	45.73	802.52	9581.44	15.02	6.06	1.20	5.77	3.35
quicksort	1.31	1.15	32.89	264.54	4.92	1866.01	43.23	2.70	1.29	2.03	2.60
mandel	0.81	0.79	15.32	53.16	7.58	451.81	5.13	0.66	1.11	0.67	1.35
pi_sum	1.00	1.00	21.99	9.56	1.00	299.31	1.69	1.01	1.00	1.00	1.00
rand_mat_stat	1.45	1.66	17.93	14.56	14.52	30.93	5.95	2.30	2.96	3.27	3.92
rand_mat_mul	3.48	1.02	1.14	1.57	1.12	1.12	1.30	15.07	1.42	1.16	2.36

Table D.1: Benchmark times relative to C (smaller is better, C performance = 1.0). Adapted from http://julialang.org. Please refer to the original for technical specifics.

Some features of the language that we find particularly useful for geophysical applications are its ability to execute multi-level *for loops* without the performance penalty of other high-level languages (eg. Python, and Matlab). The language also allows for C and Python libraries to be called directly with no overhead, and has functionality for large-scale parallel computing.

---

[1] Parts of this appendix have been published in Stanton et al. (2016)

Method

The Signal Analysis and Imaging Group have created a module named *Seismic.jl* that can be installed from the Julia command line by typing *Pkg.add("Seismic")*. The package contains utilities for reading and writing SEG-Y, SU (Seismic Unix), and RSF (Madagascar) data formats, and uses a simple internal format that emulates Madagascar's RSF format to store data where headers and data are kept in separate binary files. By storing data and headers separately we can efficiently compute header statistics and manipulate data based on header values. The type system in Julia allowed us to create a custom Header type that contains many useful indices for 5D processing and imaging.

Seismic.jl contains modules for many conventional data manipulations including windowing, sorting, 5D geometry calculation, 5D binning, as well as wrapper modules for processing groups or patches of data in parallel. The processing functionality of the package is a work in progress, but already includes semblance, NMO, band-pass and FK filtering, FX-deconvolution, Radon demultiple, 5D interpolation, stacking, and 3D shot-profile acoustic and elastic wave equation migration, de-migration and least-squares migration.

Examples

To try these examples yourself we ask the interested reader to visit `http://juliabox.org` where a Linux virtual machine can be accessed in the browser for no cost. Below we show a simple seismic data interpolation exercise using the Projection Onto Convex Sets (POCS) algorithm (Abma and Kabir, 2006). The example can be reproduced by following `https://goo.gl/gYtWZs`. To create a simple synthetic dataset to test the algorithm we use the *SeisLinearEvents* program and decimate 50% of the traces randomly. Figure D.1 shows the decimated input data and its FK-spectrum. The data are reconstructed using 100 iterations of POCS with results shown in figure D.2.

As a second example, we compute reflector dip for a velocity model using Plane Wave Destruction (PWD) (Claerbout, 1992). In figure D.3 the velocity model was used as an input to the program SeisPWD, with computed reflector normal directions plotted with arrows throughout the model. This example can be reproduced by following `https://goo.gl/zu40wg`.

For our last example, we use the Teapot Dome dataset available on the SEG's open data website. In this simple example, we convert from SEGY to SEIS format and calculate geometry information including binned midpoint X and Y coordinates. Extracting coordinates from the header file, we map shot and receiver coordinates and an image of the fold in figures D.4 and D.5 respectively. To follow this example visit `https://goo.gl/kwvFqQ`.

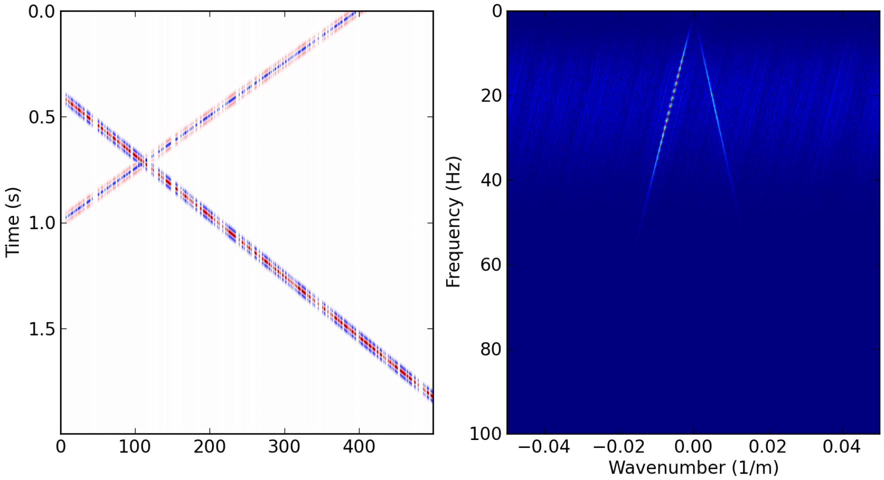


Figure D.1: decimated seismic data used as input to the SeisPOCS program. The data have 50% missing traces.

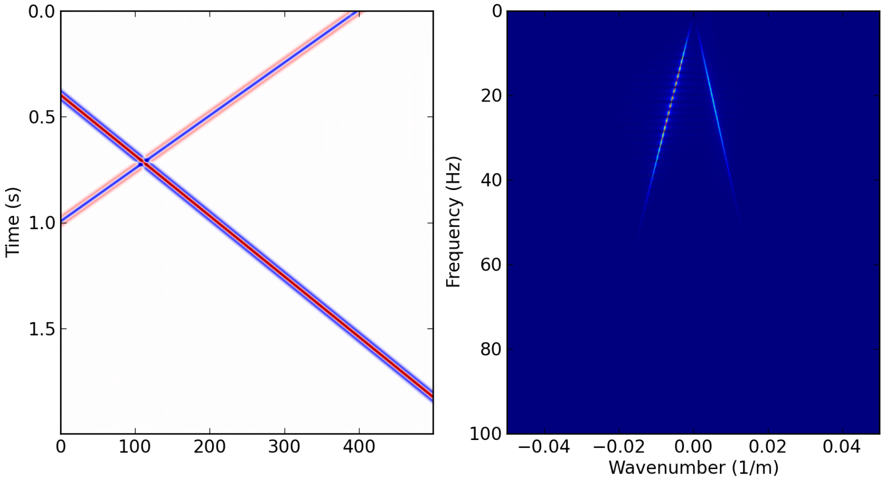


Figure D.2: Seismic data interpolated using the SeisPOCS program.

Conclusions

Julia offers the unique ability to write simple prototypes for geophysical research and teaching that can also scale to solve large problems efficiently. We have introduced a set of utilities for writing and manipulating seismic data in the Julia language.

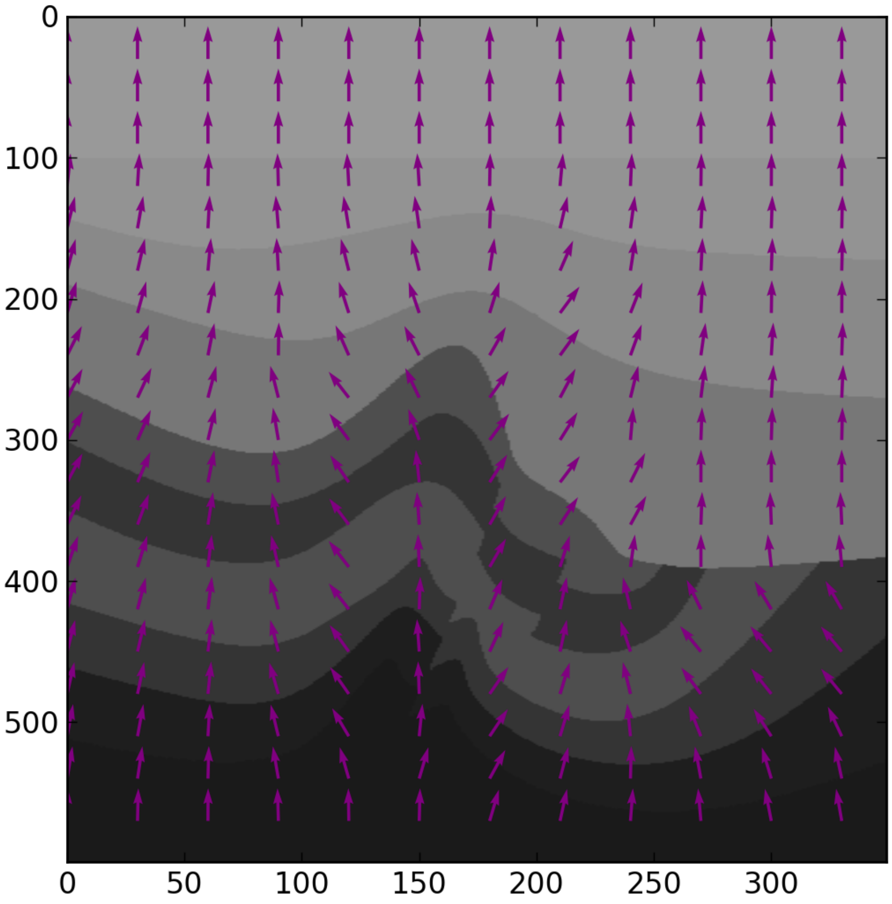


Figure D.3: A velocity model used to demonstrate the program SeisPWD. Computed normal vectors are plotted as purple arrows.

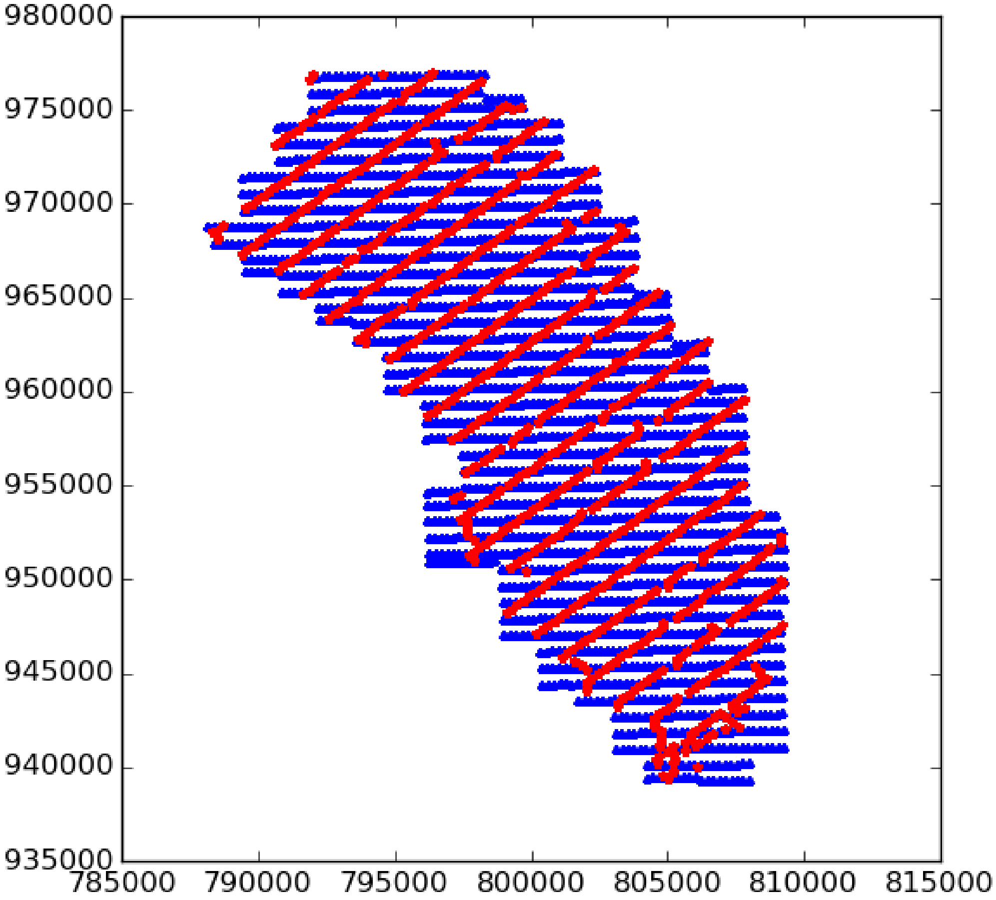


Figure D.4: Shot and receiver coordinates extracted from the Teapot Dome dataset.

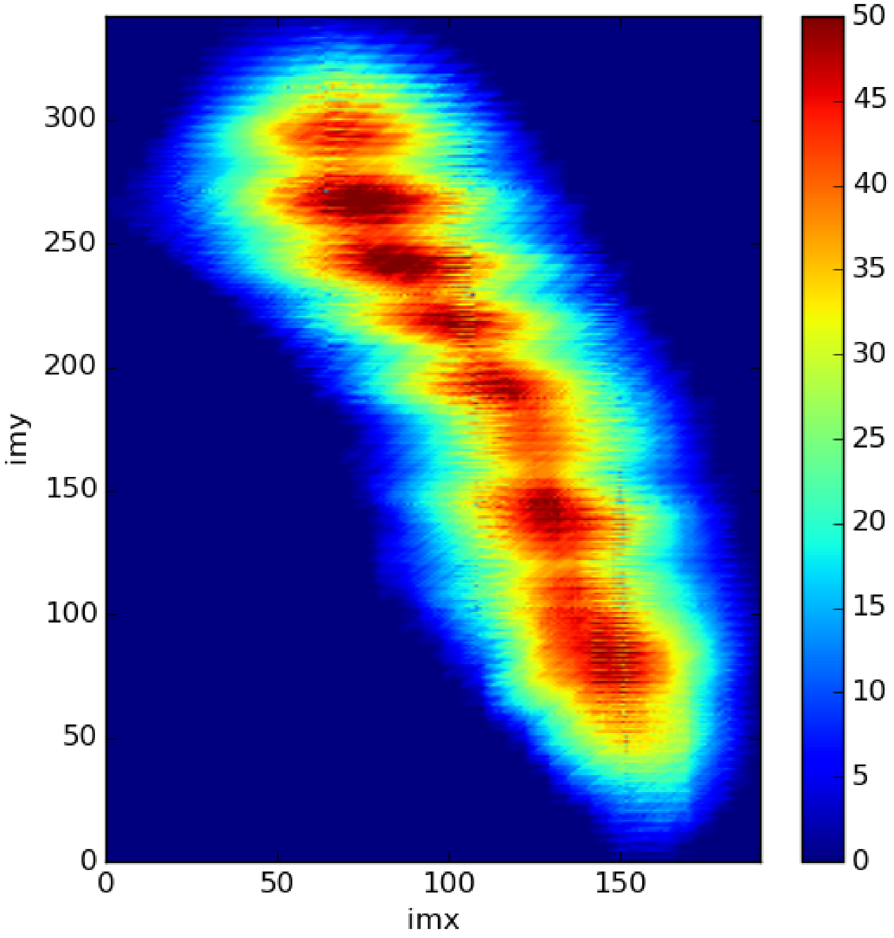


Figure D.5: Fold map computed from the Teapot Dome dataset.