ENHANCING LAND SEISMIC DATA WITH COMPRESSIVE SENSING AND PROCESSING

by

Iga K. Pawelec

© Copyright by Iga K. Pawelec, 2023

All Rights Reserved

A thesis submitted to the Faculty and the Board of Trustees of the Colorado School of Mines in partial fulfillment of the requirements for the degree of Doctor of Philosophy (Geophysics).

Golden, Colorado

Date _____

Signed: _____

Iga K. Pawelec

Signed: _____

Dr. Paul C. Sava
Thesis Advisor

Golden, Colorado

Date _____

Signed: _____

Dr. Paul C. Sava
Professor and Department Head
Department of Geophysics

ABSTRACT

Active-source seismic methods can provide a wealth of information about the structural and stratigraphic makeup of the subsurface as well as about physical and reservoir properties of rocks. However, high-resolution imaging subsurface techniques are challenging to apply for land seismic data. Unlike in marine acquisition where the source signal propagates through a near-homogeneous acoustic water layer, land acquisition places sources and receivers on a free-surface boundary between air and often poorly consolidated sediments. This near-surface layer traps most of the source-generated energy and gives rise to surface and guided waves and scattering noise that can propagate with extremely slow velocities. Wavefields triggered in the near surface have short wavelengths, and thus are difficult to acquire non-aliased which is the main source of discrepancy between marine and land seismic data quality. In this thesis, I research compressive sensing approaches to reduce the number of sensors required for non-aliased recordings of land wavefields and methods to improve regularly sampled aliased data. I consider a multi-channel extension of compressive sensing using both signal and its spatial derivative with a common sparse support constraint but conclude that due to noise and sub-optimal recovery algorithm such an approach yields negligible benefits over a single-channel variant. I establish that complex wavelet domain is an optimal choice for sparsifying highly non-stationary land wavefields for single-channel compressive sensing and develop thresholding techniques that can be used for a sparsity-promoting data reconstruction and for interpolation beyond aliasing. To demystify the complex wavelet domain, I represent complex wavelet coefficients on their idealized Fourier domain support with frequency-wavenumber octave bands representing different scales and orientations. Such approach provides the direct link between complex wavelet scales and orientations and phase velocities, enabling straightforward definitions of velocity filters that are localized in time, space, frequency and wavenumber and can therefore yield better signal and noise separation than the traditional approaches based solely on the Fourier domain.

TABLE OF CONTENTS

ABSTRACT iii

LIST OF FIGURES vii

LIST OF ABBREVIATIONS xii

ACKNOWLEDGMENTS xiii

DEDICATION xv

CHAPTER 1 INTRODUCTION 1

CHAPTER 2 MISSING TRACE RECONSTRUCTION FOR 2D LAND SEISMIC DATA WITH RANDOMIZED SPARSE SAMPLING 6

2.1 Introduction 6

2.2 Challenges in Data Reconstruction 10

2.2.1 Aliasing 10

2.2.2 Gap Pattern 13

2.2.3 Dynamic Range and Fourier Domain Representation 14

2.3 Theory 18

2.3.1 Sparse Signal Recovery 18

2.3.2 Fourier Transform 19

2.3.3 Discrete Wavelet Transform 20

2.3.4 Complex Wavelet Transform 22

2.3.5 Curvelet Transform 25

2.4 Data Reconstruction Examples 25

2.5 Discussion 29

2.6 Conclusions 31

2.7 Acknowledgements 32

CHAPTER 3 HIGH DYNAMIC RANGE LAND WAVEFIELD RECONSTRUCTION FROM RANDOMIZED ACQUISITION 36

3.1 Introduction 37

3.2 Complex Wavelet Transform 42

3.3 Compressive Sensing 46

3.4 Iterative Soft Thresholding 47

3.5 Data Reconstruction Examples 49

3.6 Conclusions 52

CHAPTER 4 TOWARDS 'GOOD' SEISMIC DATA: BEYOND ALIAS INTERPOLATION AND FILTERING IN THE COMPLEX WAVELET DOMAIN 56

4.1 Introduction 57

4.2 Complex Wavelet Reconstruction and Filtering 60

4.2.1 Complex Wavelet Transform 61

4.2.2 Rotational Variance 63

4.2.3 Multiscale $\mathbb{C}$WT Reconstruction 65

4.2.4 $\mathbb{C}$WT Velocity Filter Design 67

4.3 Examples 70

4.3.1 Regular Data Interpolation Beyond Aliasing 71

4.3.2 Infilling Large Gaps 74

4.3.3 Airblast Removal 76

4.4 Discussion 78

4.5 Conclusions 79

CHAPTER 5 MULTI-CHANNEL COMPRESSIVE SENSING FOR SEISMIC DATA RECONSTRUCTION USING JOINT SPARSITY 80

5.1 Introduction 81

5.2 Wavefield Derivatives 84

5.3 Compressive Sensing 86

5.4 Data Reconstruction Examples 91

5.4.1 1D Sparse Sinusoids 91

5.4.2 Synthetic 3D Test 94

5.4.3 2D Field Data 99

5.5 Discussion and Conclusions 108

CHAPTER 6 CONCLUSION 110

REFERENCES 114

APPENDIX A DERIVING 2D GSE 127

APPENDIX B COPYRIGHT PERMISSIONS 130

B.1 Society of Exploration Geophysicists 130

LIST OF FIGURES

Figure 2.1 (a-c) Land data sampled at 1.25 m, 5 m and 10 m trace interval, and (d-f) the corresponding frequency spectra. Data amplitudes are gained for display. Note that aliasing occurs even at 5 m sampling interval due to the slow surface waves. Data sampled at 10 m are difficult to interpret. 12

Figure 2.2 Dynamic range of the land seismic data. Note that significant amplitudes span several orders of magnitude and that most of the shot-generated energy is trapped in the near-source region. 13

Figure 2.3 Plane wave (a) of constant amplitude and (b) with amplitude decay proportional to $1/r$. (c and d): Fourier domain representations of (a and b). Note that the Fourier representation of (b) is much less sparse than that of (a). . 17

Figure 2.4 The nonlinear approximation error of a seismic field record as a function of sparsity ratio $\rho = K/N$, where N represents the total number of coefficients in the transform domain and K is the number of top magnitude coefficients used in the nonlinear approximation. The Fourier transform performs the worst, whereas the curvelets and wavelets are closely matched. 20

Figure 2.5 Idealized Fourier domain support for 2D real wavelet transform. (a) One level decomposition with white box in the middle representing low-frequency signal approximation. The blue, red and black represent vertical, horizontal and diagonal details, respectively. Level two wavelet decomposition in (b) results from further decomposing the white box in (a). 21

Figure 2.6 Idealized Fourier domain support for 2D dual-tree $\mathbb{C}$WT . (a) One level decomposition with white box in the middle representing low-frequency signal approximation. The colors denote support of distinct detail subspaces. Level two wavelet decomposition in (b) results from further decomposing the white box in (a). 22

Figure 2.7 (a and b) Magnitudes and (c and d) phases of the two-level complex wavelet decomposition of a circle. Note the selective directional sensitivity and phase shift between the two trees. 24

Figure 2.8 (a) Field data and (b) their spectrum used in reconstruction experiments. Note that the field data are critically sampled. 27

Figure 2.9 Summary of the reconstruction results for field data in Figure 2.8. Solid lines indicate the average performance of each domain. 28

Figure 2.10 From left to right: the worst, average and best case complex wavelet-based reconstruction for $\delta = 0.5$. (d-e) show the data difference and (g-i) show the $f - k$ spectrum of the reconstructed data. 33

Figure 2.11 From left to right: the worst, average and best case curvelet-based reconstruction for $\delta = 0.5$. (d-f) show the data difference and (g-i) show the $f - k$ spectrum of the reconstructed data. 34

Figure 2.12 From left to right: the worst, average and best case real wavelet-based reconstruction for $\delta = 0.5$. (d-f) show the data difference and (g-i) show the $f - k$ spectrum of the reconstructed data. 35

Figure 3.1 (a)-(c): Seismic records used for reconstructions experiments and (d)-(f) the corresponding $f - k$ spectra. (a) Land record with 1.25 m receiver spacing, (b) land record with 2.0 m receiver spacing, and (c) synthetic data with 10.0 m receiver spacing. All records are normalized such that their peak amplitude is unity. We use symmetric logarithmic scale to emphasize the large data dynamic range. $f - k$ spectra are displayed with logarithmic color mapping. 42

Figure 3.2 Magnitudes of complex wavelet coefficients for wavefield in Figure 3.1(b): (a) before and (b) after random trace subsampling. (c) and (d) are zoom on the $\mathbb{C}$WT coefficient magnitudes from yellow boxes in (a) and (b). Random trace subsampling introduces stripe-like pattern mimicking missing trace geometry while the coefficients representing the remaining traces are strengthened. 45

Figure 3.3 Reconstructions of the wavefields in Figure 3.1: (a)-(c) basis pursuit with SPGL1 solver, (d)-(f) IST, and (g)-(i) POCS. 54

Figure 3.4 Data differences between fully sampled wavefields from Figure 3.1 and their reconstructions from Figure 3.3: (a)-(c) basis pursuit with SPGL1 solver, (d)-(f) IST, and (g)-(i) POCS. 55

Figure 4.1 Gap orientation sensitivity test. We remove a piece of an arch from the circularly symmetric object, run complex wavelet reconstruction, and repeat the test at multiple angular orientations. The top row shows reconstructions for 0°, 45°, and 90°, respectively, and the middle row shows difference between original and reconstructed object magnified 10x. Each reconstruction is quantified by NRMSE and plotted against gap azimuth (bottom). Dashed lines show primary angular sensitivity directions for a 2D $\mathbb{C}$WT . The most favorable reconstructions are indicated by green stars while the least favorable are marked with red crosses. 64

Figure 4.2 (a) Magnitude and (b) phase of complex wavelet coefficients for a land seismic record sampled at 2 m and 4 ms. The largest coefficients correspond to the high energy events and to the features supported by given scale and orientation (e.g., that fast first arrivals have strong coefficients at 62.5 - 125 Hz, -0.1 - 0.1 1/m, which follows the geophysical intuition). The phase spectrum reveals that high frequency, high wavenumber features do not form many coherent patterns, indicating that these $\mathbb{C}$WT bands primarily support noise. 68

Figure 4.3 Mapping of different phase velocities to $\mathbb{C}$WT orientations assuming spatial sampling of 2 m and temporal sampling of (a) 1 ms and (b) 4 ms. Note that changing the temporal sampling alters the distribution of phase velocities among $\mathbb{C}$WT scales and orientations which has implications for noise suppression and signal separation. 69

Figure 4.4 (a) Fully sampled wavefield and (b) wavefield after spatial subsampling by a factor of 4. (c) and (d) are the corresponding FK spectra. 72

Figure 4.5 Reconstructions of aliased wavefield from Figure 4.4(b) on a fine grid. (a) Reconstruction by basis pursuit, and (b) and (c) are obtained with MIST using subsampled data or reconstructed data as an input. (d)-(e) show the corresponding data differences, while (f)-(h) show the FK spectra. 73

Figure 4.6 (a) Reference data with a 160m (16 traces) gap marked with the black box and the reconstructed wavefields using (b) basis pursuit and (c) MIST. (d)-(f) show the gap interior. MIST has better amplitude accuracy, but is sometimes unable to accurately reproduce event moveouts, for example the diffraction tail around 2.47 s and 1200 m visible in (d). 75

Figure 4.7 Complex wavelet domain filter masks for airblast suppression. Note that we only target the scales and orientations that are the most affected by the airblast energy, leaving others unchanged. 77

Figure 4.8 Airblast removal in the complex wavelet domain. (a) raw data, (b) after removing the airblast, and (c) data difference. Data are gained and clipped for display. Note that the majority of the airblast energy is removed without affecting other events. 77

Figure 5.1 The recovery probability, average mean-squared error and its standard deviation for (a) single channel and (b) multi-channel recovery of a 1D signal sparse in the Fourier domain. K/N is the fraction of non-zero coefficients needed to describe the signal, while M/N is the fraction of available samples from the fully sampled signal. The transition between successful and failed recovery is clear for both scenarios, but reconstructions are more consistent for the multi-channel case. 83

Figure 5.2 1D signal composed of two sinusoids recorded with a multi-channel sampling system: by taking discrete samples of a signal (top) and its first derivative (bottom). Large light circles on both panels indicate Nyquist sampling locations while small dark circles show one realization of compressive sampling. Note that compressive samples consist of only 22% of the Nyquist samples. 85

Figure 5.3 Fourier domain representation of the signal shown in Figure 5.2. Note that severe subsampling in compressive sampling makes it challenging to accurately estimate the true support of the underlying signal (8 and 24 Hz in this example). 88

Figure 5.4 Weights applied to the derivative channel (top) and the effect of combining compressive signal samples with weighed compressive derivative samples. Note that in this instance, the combined signal spectrum yields highest energy coefficients at 8 and 24 Hz, the same as the Nyquist sampled signal. However, this property may not hold for other realizations of signal sampling. 90

Figure 5.5 Reconstruction of a 1D signal shown at the top panel in Figure 5.2 using single- and multi-channel IHT. Note that the multi-channel approach allows to recover the Nyquist samples up to the numerical precision. 91

Figure 5.6 Quality metrics for single- and multi-channel compressive sensing reconstruction (top and bottom rows, respectively). (a) and (c) are the probabilities of recovering the correct Fourier domain support while (b) and (d) show the quality of the reconstructed signal quantified by SNR. 93

Figure 5.7 Reconstruction of coarsely sampled seismic data. (a), (b) are the slices from densely sampled 3D volume and (c), (d) are the compressive samples after 80% subsampling. 95

Figure 5.8 Single-channel reconstruction for the wavefield depicted in Figure 5.7 and the corresponding differences. The SNR for the recovered volume is 11.60 dB. 96

Figure 5.9 Multi-channel reconstruction for the wavefield depicted in Figure 5.7 and the corresponding data differences. The SNR for the recovered volume is 12.69 dB. 97

Figure 5.10 Simultaneous sparse approximation of 4 channels: (a) data samples, (b) time derivative, (c) x derivative, and (d) y derivative. Enforcing the common sparse support constraint helps to preserve the physical relationship between the channels. 98

Figure 5.11 (a) Land seismic gather and (b) its inline derivative. The symmetric logarithmic scale is used to showcase the range of significant amplitudes. . . . 100

Figure 5.12 Normalized root-mean-square error computed from (a) only the reconstructed missing traces and (b) full reconstructed wavefield including the originally present traces. Amplitudes were not altered before computing the NRMSE. The green star, yellow diamond, and red cross indicate the best, average, and worst multi-channel reconstruction from (a). Note that the best and the worst missing trace reconstructions do not not have smallest and largest full wavefield NRMSE. 103

Figure 5.13 Normalized root-mean-square error computed from (a) only the reconstructed missing traces and (b) full reconstructed wavefield including the originally present traces. Amplitudes were range compressed before computing the NRMSE. Green star, yellow diamond, and red cross indicate the best, average, and the worst multi-channel reconstruction from Figure 5.12(a). Note that the range of NRMSE after range compression is smaller than for unaltered wavefields. 103

Figure 5.14 (a) Single- and (b) multi-channel reconstructions for a realization with the smallest missing trace NRMSE (Figure 5.12(a)). (c) and (d) are overlays of the respective reconstructions on the data from Figure 5.11(a). Both reconstructions are comparable, with the multi-channel approach offering a slight improvement at the near offsets. 104

Figure 5.15 (a) Single- and (b) multi-channel reconstructions for a realization with the largest missing trace NRMSE (Figure 5.12(a)). (c) and (d) are overlays of the respective reconstructions on the data from Figure 5.11(a). Despite the overall poor NRMSE, both reconstructions are mostly accurate in terms of phase, with the most significant errors within large continuous gaps. 105

Figure 5.16 (a) Single- and (b) multi-channel reconstructions for a realization with the smallest range compressed missing trace NRMSE (Figure 5.13(a)). (c) and (d) are overlays of the respective reconstructions on the data from Figure 5.11(a). Range compression partially mitigates the NRMSE bias towards high amplitudes and reveals that both approaches perform well at farther offsets and struggle with accurate amplitude recovery at near offsets. . . 106

Figure 5.17 (a) Single- and (b) multi-channel reconstructions for a realization with the largest range compressed missing trace NRMSE (Figure 5.13(a)). (c) and (d) are overlays of the respective reconstructions on the data from Figure 5.11(a). Both reconstructions struggle with the significant number of missing traces at near offsets and early times, introducing notable phase and amplitude errors. . 107

Figure B.1 SEG's permission to reproduce published work in a thesis 130

LIST OF ABBREVIATIONS

Center for Wave Phenomena . . . . . CWP

Colorado School of Mines . . . . . CSM

complex wavelet transform . . . . . $\mathbb{C}$WT

compressive sensing . . . . . CS

discrete wavelet transform . . . . . DWT

full waveform inversion . . . . . FWI

iterative hard thresholding . . . . . IHT

iterative soft thresholding . . . . . IST

multiscale iterative soft thresholding . . . . . MIST

normal moveout . . . . . NMO

normalized root mean squared error . . . . . NRMSE

prediction error filter . . . . . PEF

projection onto convex sets . . . . . POCS

short time Fourier transform . . . . . STFT

signal to noise ratio . . . . . SNR

ACKNOWLEDGMENTS

First, I would like to thank my advisor and mentor, Paul Sava, for encouraging me to think outside of the box, to question, and to be relentless in search of answers. I appreciate the trust and the freedom I enjoyed while working on this thesis and am thankful for the many iTeam support sessions we had during the trying pandemic times.

I would like to thank Mike Wakin - for sparking my interest in the compressive sensing by doing a great job as a teacher, and for the many discussions we had on various aspects of signal processing. Our interactions opened my perspective on research outside of geophysics and taught me how to communicate science in a more approachable manner.

I also thank my other committee members: Whitney Trainor-Guitton, Jeffrey Shragge, and Luis Tenorio for making time to chat about research challenges and providing valuable perspectives.

I want to thank all the CWP sponsors who provide funds for students like me to conduct their research. I am grateful for their presence and continued interest in our research, expressed by regular attendance at our annual consortium meetings.

There are many people in the Department of Geophysics (GP) and Center for Wave Phenomena (CWP) who helped me in various ways throughout my time here. I thank my fellow CWPers who were there to help in and outside of school: Daniel Rocha, Ivan Lim, Thomas Rapstine, Odette Aragao, Khalid Almuteri, Werter Silva, Nicholas Dorogy, Jihyun Yang, and Aaron Girard. I am grateful to GP and CWP staff; Michelle Szobody for invaluable administrative assistance and fabulous coffee chats; Joana Perez, Dawn Umpleby, Debra Marufo, Gina Schwieger, Emilia Clayton, Noelle Vance, and Lynn Lundebrek who keep the department and consortium running like a well-oiled machine.

Special thanks to Diane Witters, who introduced me to the art of technical writing and helped to fit in with the Mines community. I value the lessons learned and the friendship that blossomed.

I am very grateful for the continued support of my family in pursuing my dreams. My parents - Joanna and Piotr, and my brother Filip, were always there for me just a video call away. Though I sometimes struggle to express it, the fact that I can always count on them means a lot to me.

My grandparents have also played an important role in this PhD journey. My grandpa Adam, for whom this thesis is dedicated, has always encouraged me in all of my pursuits, in sports and academia alike. His unwavering faith that I can do anything I set my mind to has helped me through some difficult times. Though He is no longer here to see the journey's end, I like to believe that he would have been proud of what I have achieved.

Finally, I thank Agnieszka and Jim Folcik for adopting me as an unofficial family member and showing me that family does not have to consist of blood relations only.

In the loving memory of grandpa Adam, who always believed in me.

CHAPTER 1

INTRODUCTION

Out of many geophysical methods, active-source seismic surveys gained the most popularity as a source of large-scale subsurface information. Enabled by state-of-the-art imaging technology such as full waveform inversion (FWI) and the ever-increasing computational power, seismic data provide detailed subsurface imagery, delineating structures, stratigraphy, reservoir properties, and time-lapse changes in these properties. Although the imaging technology itself is an active research area that sees constant improvement by implementing more realistic physics of wave propagation or by adding reasonable model constraints, the biggest challenge to high-resolution imaging is seismic data quality.

Land seismic data are somewhat notorious for being significantly more challenging than their marine counterparts. First, land seismic data are sometimes acquired in areas with elevated background noise, e.g., due to local traffic or industrial activities. Atmospheric conditions such as wind and precipitation combined with sub-optimal geophone coupling also contribute to the seismic background noise (Barajas-Olalde and Jeffreys, 2014; Dybing et al., 2019; Nørmark, 2011). The commonly used land seismic source - vibroseis - generates signals outside of its prescribed sweep such as sweep harmonics, acoustic noise, and engine-related mechanical noise (Bagaini et al., 2014; Denisov et al., 2019; Polom, 1997). For example, in Pawelec and Sava (2020), I show that continuous time records over an ultra-dense array enable characterization of the source-generated noise, in turn allowing for its successful removal.

The biggest source of data quality discrepancy between marine and land data is the geologic environment in which seismic sources and receivers are deployed. In an offshore setting with towed streamer acquisition, sources and receivers are situated in an acoustic water layer with small spatial variations in wave propagation velocities. Recently gaining in popularity, ocean bottom nodes can be placed on a seafloor, but the source signal is still generated in the

near-homogeneous water layer. Conversely, in an onshore setting the sources and receivers are commonly placed on a free surface boundary between air and the often poorly consolidated near surface. These conditions pose many problems to land seismic acquisition (Keho and Kelamis, 2012). The majority of the source-generated energy is trapped in the near-surface layer, propagating as surface or guided waves and scattering on subsurface heterogeneities. Stork (2020) describes the mechanism of near-surface noise generation and remarks that land seismic data suffer from 10-100 times more shot-generated noise than marine data. Handling noise that is markedly stronger than signal can be problematic, especially when that noise is aliased. A subpar noise suppression lowers data quality, which in turn adversely affects seismic imaging.

It has been a long-standing assertion that the key to overcoming most problems associated with land seismic data is to increase trace density. Jack (2003) recognized that the then current way of acquiring seismic data may not be the most optimal and that hardware changes might be necessary to achieve higher data quality. Keho and Kelamis (2012) remark that both signal and noise need to be sampled properly in long-azimuth, long-offset geometries which would necessitate channel counts in hundreds of thousands or even millions. Ourabah et al. (2014) study the sampling effect on noise attenuation, surface consistent processing, velocities, the final imaging and attributes, and conclude that dense sampling improves resolution and yields better signal-to-noise content in the final image. Manning et al. (2018) further the case for nodal technology that unlocks unprecedented channel counts, arguing that such technology not only helps to acquire superior quality data, but also lowers health, safety, and environment risks, increases operational efficiency, and reduces acquisition costs.

Wireless nodal technology is now available (Manning et al., 2019), and the increase in trace density is starting to reverse the curse of low-quality land seismic data. Ourabah et al. (2019) compare wireless single sensor recordings to cabled systems in a desert environment and show that nodes yield high-quality data that are almost indistinguishable from those acquired on a wired system. Yanchak et al. (2018) present a case study comparing the processing outcomes for a legacy data and a new, ultra-high density survey with 72 million traces/km^2. Although the same

state-of-the-art technology is used for processing both surveys, the ultra-dense dataset yields a step change in imaging quality, allowing for the detailed delineation of previously unresolvable structural features. An important part in this success story is the ability to record surface waves without aliasing. Surface waves, often referred to as ground roll, can yield detailed near surface models that are useful for improving static corrections and velocity model resolution for migration (Socco et al., 2010b; Strobbia et al., 2011). Non-aliased surface waves are also easier to remove from reflection records once they fulfilled their purpose (Alkatheeri et al., 2020; Manning et al., 2019).

It may seem that ultra-high density regular acquisition is a panacea for the land data quality issues; however, there are two main problems with that solution: 1) sampling intervals needed for recording non-aliased wavefields can be on the order of 2 m or less, which remains prohibitively expensive for large acquisition projects, and 2) terrain obstacles and other access restrictions prevent the placement of sources and/or receivers in certain locations, resulting in irregular acquisitions with spatial sampling gaps.

The first challenge can be addressed within a compressive sensing (CS) framework (Candès et al., 2006b; Candès and Tao, 2006; Donoho, 2006). CS is a sampling/sensing paradigm that facilitates capturing signals from reduced irregular measurements compared to the regular sampling requirements (Shannon, 1948). Instead of using large number of regularly placed sensors, it is possible to deliberately randomize sampling and take advantage of a sparse signal representation in some domain with a sparsity-promoting reconstruction algorithm to recover a fully sampled wavefield on a regular grid. Therefore, finding an appropriate sparse domain and sparse recovery strategy are key to the success of the CS approach. The challenge of gaps and regular aliased data requires developing technology that can use the available information in clever ways to interpolate beyond aliasing and to infill the gaps.

In this thesis, I develop compressive sensing and processing solutions for overcoming the land seismic data acquisition-related challenges. Each chapter is a self-contained paper that examines some aspect of data acquisition and reconstruction problem and proposes practical solutions.

In Chapter 2, which is a paper entitled "Missing trace reconstruction for 2D land seismic data with randomized sparse sampling" published in *Geophysics*, I discuss several possible choices for sparsifying land seismic wavefields. I describe how the non-stationary characteristics of land wavefields make the Fourier domain an infeasible choice, leaving curvelets (Candès et al., 2006a) and discrete complex wavelets (Kingsbury, 2001; Selesnick et al., 2005) as the main contestants due to their localized character and directional selectivity. My experiments with multiple realizations of random trace geometry at different trace decimation ratios of a densely sampled land record reveal that complex wavelets consistently yield the best reconstructions in terms of signal-to-noise ratio, which establishes them as the optimal sparse domain for land seismic data reconstruction problems.

In Chapter 3, entitled "High dynamic range land wavefield reconstruction from randomized acquisition" (submitted to *Geophysics*), I develop a sparsity-promoting reconstruction that is optimized for complex wavelet domain. By estimating scale- and orientation-dependent thresholds and using them within an iterative soft thresholding algorithm (Beck and Teboulle, 2009; Daubechies et al., 2004; Donoho, 1995), I obtain high-quality data reconstructions that are superior to those obtained following a basis pursuit approach with a spectral projected gradient solver (SPGL1) (van den Berg and Friedlander, 2009, 2011) and from hard thresholding in the projection onto convex sets (POCS) framework. My technique reduces the number of required sampling points or obtains higher signal-to-noise ratio reconstructions from the same trace number.

In Chapter 4, entitled "Towards 'good' seismic data: beyond alias interpolation and filtering in the complex wavelet domain" (submitted to *Geophysics*), I address the problem of regularly undersampled data and large data gaps. Modifying the algorithm from Chapter 3 to leverage the multiresolution aspect of the complex wavelet transform ($\mathbb{C}$WT), I demonstrate its ability for interpolation beyond aliasing. I also take this opportunity to showcase the $\mathbb{C}$WT-enabled filtering capabilities that enhance signal and noise separation with a minimal adverse impact on the signal.

Chapter 5, entitled “Multi-channel compressive sensing for seismic data reconstruction using joint sparsity”, is based on the work I presented at 2022 IMAGE meeting. I use it to explore the possibility of further increasing the average spacing between sampling points or improving reconstruction quality by incorporating derivative information into the compressive sensing framework. I use joint sparse support as a constraint for jointly reconstructing wavefield and its derivatives. Though initial tests on synthetic data look promising, multi-channel reconstruction does not yield appreciable uplift in reconstruction quality when applied to field data, suggesting that other ways of combining different data channels should be considered.

I conclude this thesis with a discussion of key insights from each chapter and my thoughts on how these insights might shape future research in the quest for ‘good’ land seismic data.

CHAPTER 2

MISSING TRACE RECONSTRUCTION FOR 2D LAND SEISMIC DATA WITH RANDOMIZED SPARSE SAMPLING

A paper published in *Geophysics*

Iga Pawelec[1,2], Michael Wakin[2], Paul Sava[2]

Acquisition of high-quality land seismic data requires (expensive) dense source and receiver geometries to avoid aliasing- related problems. Alternatively, acquisition using the concept of compressive sensing (CS) allows for similarly high-quality land seismic data using fewer measurements provided that the designed geometry and sparse recovery strategy are well matched. We have developed a complex wavelet-based sparsity-promoting wavefield reconstruction strategy to overcome challenges in land seismic data interpolation using the CS framework. Despite having lower angular sensitivity than curvelets, complex wavelets improve the reconstruction of sparsely acquired land data while being faster and requiring less storage. Unlike the Fourier transform, the complex wavelet transform localizes aliasing-related artifacts likely to be present in field data and yields reconstructions with fewer artifacts and higher signal-to- noise ratios. We determine that the data recovery success depends on the number and the geometry of the missing traces as revealed by analyzing reconstructions from multiple realizations of trace geometry and data decimation ratios. Using half the number of traces required by the regular sampling rules and thus reducing the acquisition costs, we find that data are appropriately reconstructed provided that there are no large gaps in the strategic places.

2.1 Introduction

Land seismic data are notorious for being particularly challenging to handle. The challenges can be separated into two categories: acquisition and processing. On the acquisition side, the

[1]Primary researcher and author.
[2]Colorado School of Mines, Golden, Colorado, United States of America.

main challenge is that poor coupling between receivers and the medium (sand, soil or ice) significantly reduces the signal-to-noise ratio (S/N) of the recorded signal. There are also legal and operational access restrictions that can result in regions with no data coverage. On the processing side, challenges are associated with various types of noise (though in some applications, what we term here as noise can be used to derive information about the subsurface). The ambient noise background is typically higher on land than in a marine environment due to the proximity to different cultural sources of noise (e.g., industry, traffic). Even remote areas can present unique noise challenges such as ice breaking in the Arctic (Li et al., 2017). Furthermore, the commonly used vibroseis source generates both the desired sweep and other signals (harmonics of the sweep, sound generated by the engine) (Denisov et al., 2019; Polom, 1997), which need to be attenuated either during acquisition or processing. Finally, one of the most important causes of land data quality degradation is the highly heterogeneous and generally poorly consolidated near surface. This portion of the subsurface, called the low-velocity layer (Sheriff, 2002), gives rise to strong and dispersive surface waves and scattering noise.

Due to these challenges, land seismic data quality tends to be low, which can adversely impact reservoir characterization. In the case of unconventional reservoirs, the value of acquiring seismic data in the first place is sometimes questioned. However, even in these scenarios, seismic data can provide a wealth of information useful for characterizing unconventional plays, such as anisotropy maps, fracture orientation maps, and pressure distribution. The land seismic acquisition goal is to increase the reliability of seismic-derived products, which depend critically on having dense and high-S/N seismic data as an input.

Although we are on the verge of a revolution in terms of the number of geophones deployable in a cost-effective manner (Manning et al., 2018), some sites with particularly slow surface waves require ultrafine sampling (on the order of 1 m), and certain areas are inaccessible to both sources and receivers. To overcome these economic and logistical obstacles, one can explore alternative ways of acquiring data that could yield substantial savings on the acquisition efforts. Mosher et al. (2017) show that a novel approach, using ideas from compressive sensing (CS), can significantly

speed-up acquisition without jeopardizing the data quality. The underlying idea is to randomize receiver placement and shot timing according to compressive sampling methodologies and to subsequently solve a large-scale regularization problem, recovering nonaliased data from sparse measurements. Most conventional 3D acquisition geometries have regular, but coarse, sampling in at least one direction (Trad, 2009) and sometimes large data gaps due to access restrictions. In contrast, compressive sensing surveys have deliberate irregular sampling and often simultaneous shooting, account for geometry restrictions, and are tuned for the data recovery strategy.

Commonly, data recovery strategies rely on the presence of known, repetitive patterns in the data and/or data sparsity in some representation. Current techniques for infilling missing data can be divided in several categories, including prediction error filters (PEFs) (Claerbout and Fomel, 2014; Crawley et al., 1999; Fomel and Claerbout, 2003; Naghizadeh and Sacchi, 2009; Spitz, 1991), matrix or tensor completion (Kreimer and Sacchi, 2012; Kreimer et al., 2013; Kumar et al., 2015; Ma, 2013), rank reduction (Chen et al., 2016b; Gao et al., 2013a; Trickett et al., 2010) and machine learning (Jia and Ma, 2017; Jia et al., 2018; Pilikos and Faul, 2017; Wang et al., 2019). However, the most widely used interpolation techniques are transform-based approaches that can take advantage of known data characteristics in the transform domain. Such methods are well-studied in the context of data aliasing and irregular sampling, and rely on data representation in a known transform domain to recover missing information. Although different transforms have been used, including Radon transform (Hollander and Yilmaz, 2019; Kabir and Verschuur, 1995; Wang et al., 2010; Yu et al., 2007) and wavelet or seislet transforms (Gan et al., 2015; Liu et al., 2017; Yu et al., 2007), the Fourier transform remains the most popular choice because it is easy to implement and fast to compute, as long as data are regularly sampled. Liu and Sacchi (2004) develop a framework for data recovery based on L_2 norm minimization, using spectral weights bootstrapped from $f - k$ representation of data that can be extended to five dimensions (Trad, 2009). The Fourier domain is also used in the projection onto convex sets (POCS) method described by Abma and Kabir (2006). Duijndam et al. (1999b) tackle the problem of arbitrarily irregular sampling and leverage a weighting scheme based on adjacent sample distances to

reconstruct data with one varying spatial coordinate, whereas Xu et al. (2010, 2005) propose an antileakage version of the Fourier transform, which can handle irregular geometry and mitigate aliasing problems. One downside of Fourier-based approaches is that data have to be windowed for non-stationarity; as a consequence, only local information can be used for interpolation. Another attractive transform for seismic data interpolation gaining significant popularity is the curvelet transform (Hennenfent et al., 2010; Herrmann, 2010; Naghizadeh and Sacchi, 2010a). Curvelets provide an optimally sparse representation of seismic wavefields (Candès and Demanet, 2005), but their redundancy implies that, for a dataset of size N, as many as $7 \times N$ curvelet coefficients have to be computed, depending on the chosen number of scales and angles, which can be prohibitively expensive for large 3D datasets. Furthermore, the choice of number of scales and angles in the curvelet transform is nonintuitive and strongly data-dependent.

In this paper, we discuss the challenges inherent in seismic data reconstruction - such as significant temporal or spatial changes in signal amplitude and frequency content or data aliasing - and we propose how to overcome them by exploiting the complex wavelet domain (Kingsbury, 2001; Selesnick et al., 2005). We demonstrate that complex wavelets outperform the curvelets in the challenging field data reconstruction examples while being faster to compute and requiring less memory. We also emphasize the key role that gap pattern plays for successful data recovery.

The paper is organized as follows. First, we explore the specific challenges inherent in seismic data reconstruction with an emphasis on land seismic data. We then discuss the conditions for sparse data recovery and review several transforms available for sparse representation of seismic data. This review is followed by a description of our data recovery approach and field data examples to compare the performance of a wavelet transform (real and complex) with that of a curvelet transform. Finally, we share key insights from our analysis, including computational advantages of using the complex wavelet transform ($\mathbb{C}$WT) for seismic data reconstruction, the relevance of data decimation ratio with respect to the Nyquist wavenumber, and the impact of gap patterns on successful data recovery.

2.2 Challenges in Data Reconstruction

Every seismic survey is different, but there are some common characteristics of land seismic data, which make them particularly difficult to reconstruct: the presence of data aliasing, the pattern of missing traces, the size of data gaps, and the large dynamic range. Land seismic data tend to be more problematic than marine seismic data due to the highly complex heterogeneous near surface, which traps most of the energy released by the seismic source and produces slowly propagating surface waves (Keho and Kelamis, 2012). In the following, we discuss the challenging features of land seismic data in more detail and explain how different transforms handle them.

2.2.1 Aliasing

In land seismic data, aliasing of surface waves can be especially severe due to the much lower velocities of the surface waves com- pared with the body waves. Figure 2.1 shows the same land data record sampled at different trace intervals (coarse sampling results from discarding a portion of the full data) and their frequency-wavenumber spectra. Note that the energy corresponding to the surface waves is present between 10 and 60 Hz, and the phase velocity is approximately 200 m/s. The spatial sampling needed to acquire nonaliased data would be

$$\Delta x \leq \frac{v_{\min}}{2f_{\max}}, \tag{2.1}$$

where $v_{\min}$ is the slowest surface-wave velocity and $f_{\max}$ is the maximum frequency in the data. We can conclude that the needed sampling interval for recording the data shown in Figure 2.1 nonaliased is $\Delta x \leq 1.66$ m, which is much finer than would typically be chosen for a large-scale exploration project.

High-fidelity recording of surface waves is beneficial for near-surface characterization (Foti et al., 2014). For example, one could use surface-wave analysis to build a velocity model for a low-velocity layer (Socco et al., 2010b) or to compute static corrections (Papadopoulou et al., 2020), both of which can lead to an enhanced image of the subsurface. Furthermore, nonaliased surface waves are much easier to remove from shot record for subsequent reflection processing

(Manning et al., 2018). Although the surface-wave velocity is site dependent and the case presented in Figure 2.1 has unusually low velocities, it clearly demonstrates that, to obtain nonaliased representation of surface waves, one needs to either drastically increase the sampling on the receiver side or to develop a strategy to reconstruct the nonaliased wavefield from reduced measurements. Because multiple sensors are generally cheaper to deploy on land than multiple sources, most efforts to acquire high-density data are focused on increasing receiver coverage, with distributed acoustic sensing (Bakulin et al., 2020; Parker et al., 2014) and affordable point receivers (Manning et al., 2018) as prime examples. The wavefield reconstruction effort, on the other hand, would ideally incorporate the prior knowledge of the spatial wavenumbers expected from data with known acquisition limitations (access restrictions, available channels) to come up with a survey design allowing for reconstruction of a nonaliased wavefield in the early processing.

Reconstruction of an aliased wavefield poses a serious challenge. One possible approach uses (PEFs) (Spitz, 1991) by exploiting the idea that filter coefficients derived from low, nonaliased frequencies can be used to interpolate aliased data components (Naghizadeh and Sacchi, 2009). A similar idea is used by Gülünay (2003) in $f - k$ trace interpolation and by Zwartjes and Sacchi (2007) whose Fourier reconstruction with sparse inversion can handle irregular data sampling and aliasing. Naghizadeh and Sacchi (2010a) use nonaliased scales in the curvelet domain for reconstructing aliased data, and Gan et al. (2015) take advantage of low-pass-filtered data to interpolate using seislets (Fomel and Liu, 2010). Another popular data reconstruction strategy, the minimum weighted norm interpolation (Liu and Sacchi, 2004), requires adjustments to spectral weights to handle aliased data because additional energy is present for aliased components. Despite these advances, the degree of aliasing present in some land seismic data can be severe, suggesting that altering the data reconstruction approach would be a better solution (Baraniuk and Steeghs, 2017).

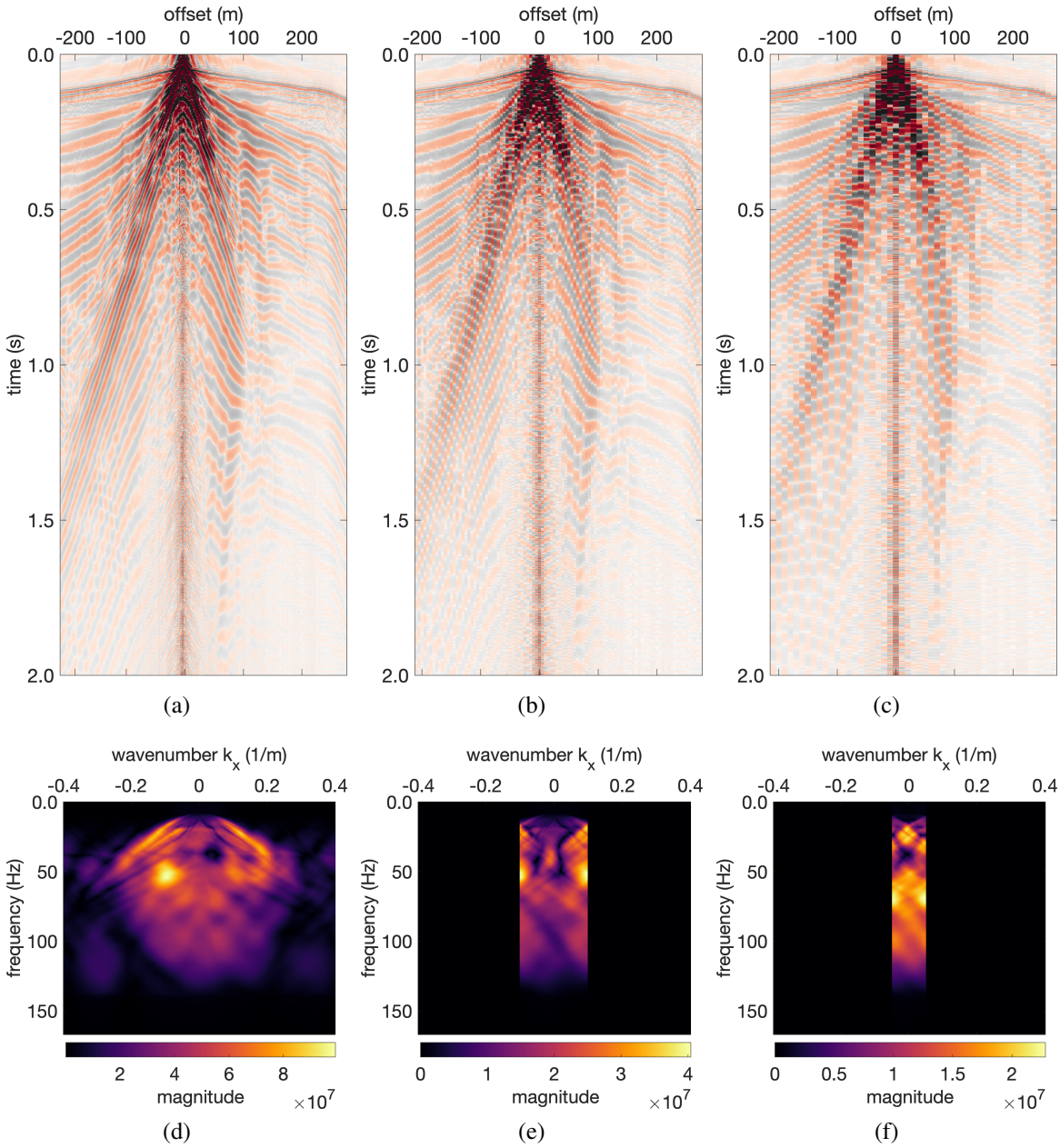


Figure 2.1 (a-c) Land data sampled at 1.25 m, 5 m and 10 m trace interval, and (d-f) the corresponding frequency spectra. Data amplitudes are gained for display. Note that aliasing occurs even at 5 m sampling interval due to the slow surface waves. Data sampled at 10 m are difficult to interpret.

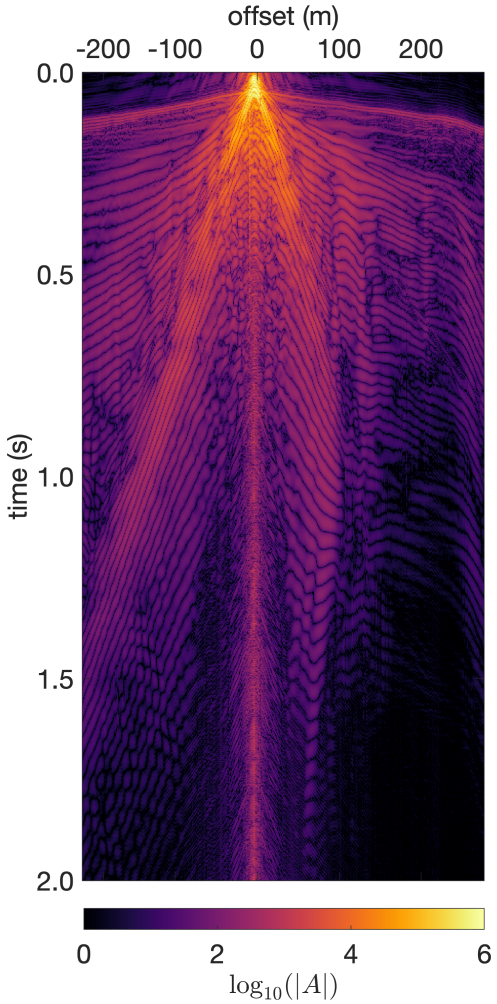


Figure 2.2 Dynamic range of the land seismic data. Note that significant amplitudes span several orders of magnitude and that most of the shot-generated energy is trapped in the near-source region.

2.2.2 Gap Pattern

Historically, seismic data have been acquired on a regular grid or are regularized after acquisition - a pragmatic choice because many processing and imaging algorithms require regular spacing. However, such acquisition is limited by the Nyquist - Shannon sampling theorem (Candès, 2006) that dictates a sampling rate of more than two points per wavelength for successful recovery of a nonaliased signal. The number of sensors needed to record good-quality, nonaliased land data on a regular grid is exceedingly high for slowly propagating waves. The advent of CS (Candès et al., 2006b) has opened up new exciting possibilities for signal

reconstruction from incomplete information. Hennenfent and Herrmann (2008) and Herrmann (2010) examine randomized acquisition using much fewer sensors than a regular-grid survey and achieve data density and quality comparable to regularly sampled, dense grid acquisition. Mosher et al. (2017) demonstrate that CS can be successfully applied to field seismic acquisition. The success of CS depends on finding the favorable gap pattern combined with sparse signal representation in a known transform domain.

2.2.3 Dynamic Range and Fourier Domain Representation

Although visually seismic data do not look more complex than many natural images (i.e., photographs of real objects), the key differences lie in the dynamic range and representation of seismic data in the Fourier domain. Dynamic range can be defined as the ratio between the largest and smallest values that a given signal can assume. Figure 2.2 shows a land shot record with the absolute values of amplitudes displayed on a logarithmic scale. Note that the coherent events that could be reliably labeled as seismic signals easily span five orders of magnitude. With the exception of high- dynamic-range (HDR) images that are stored as floating point numbers (i.e., 32 bits per color channel), natural images tend to have a low dynamic range with a fixed number of possible values per color channel: 256 for 8 bit images and 65,536 for 16 bit images. The bigger the dynamic range, the more challenging it is to find a sparse signal representation that does not compromise low data amplitudes. Thus, the many advances in image compression and reconstruction are not immediately applicable to seismic data. One can partially bypass that problem by sorting seismic data into common-offset gathers (smaller lateral amplitude variation) or by applying reversible gain functions (to preserve the relative amplitudes). Both operations may reduce the dynamic range, enhance sparsity in some transform domains, and yield a better reconstruction of missing data. However, as we will demonstrate, sparsity alone does not guarantee a successful signal recovery.

Another challenge in seismic data reconstruction is the Fourier-domain data representation. Natural images have comparable sampling rates in all image dimensions. Seismic data are

typically oversampled along the time axis and undersampled along the spatial axes. The differences in sampling lead to differences in the Fourier-domain representation of these two types of data and consequently differences in how transforms relying on sampling of the Fourier space see them. For instance, wavelets and curvelets are sensitive to angular sampling (curvelets more so than wavelets). However, while downsampling the seismic data along the time axis can usually be done without the risk of information loss, the same cannot be said about the spatial dimension. As a result, the curvelet and wavelet transforms may behave in nonintuitive ways, with large energy transform coefficients concentrated at unexpected scales and angles. Thus, the conventional wisdom about angles represented by the transform applies to data with similar sampling along all dimensions. Because seismic data typically do not meet this requirement, it is beneficial to adjust the sampling such that spatial and temporal frequencies occupy a similar portion of the Fourier space in all dimensions. This allows one to use insights derived from the application of curvelet- and wavelet-based natural image reconstruction to improve the process of seismic data interpolation.

The curvelet domain is near-optimal for representing wave phenomena (Candès and Demanet, 2005). The curvelet transform divides the frequency plane into dyadic bands, which are then split into overlapping angular wedges doubling in every other dyadic scale. The curvelet transform is highly redundant: there is no unique representation of a signal in the curvelet domain, and the number of curvelet coefficients is much larger than the number of data points. This feature of the curvelet transform is favorable for denoising and finding sparse signal representations, but it comes at the expense of increased storage requirements, which makes curve- lets a memory-expensive choice for large data sets.

$\mathbb{C}$WT , on the other hand, offers a good middle ground between the Fourier and curvelet domains. In a sense, complex wavelets can be viewed as an isotropic version of curvelets (Douma and de Hoop, 2007), although wavelets have more limited directional sensitivity. Complex wavelets provide a multiresolution approximation and, in two dimensions, are sensitive to six directions: $\pm 15^\circ, \pm 45^\circ$ and $\pm 75^\circ$. Another advantage of the wavelet transform is its linear

computational complexity and only 2^D redundancy for $D-$dimensional signals, thus making wavelets suitable for analysis of large datasets.

Due to their multiscale nature, wavelets are also well-suited for handling nonstationary signals. The large dynamic range of seismic data is particularly difficult to handle by data reconstruction algorithms, so windowing or data gaining is often used to avoid dealing with the full data range. Consider the way humans would interpolate missing data: we would look at the available portion of data to find patterns and then infill the gaps assuming that observed trends are also present in gaps. However, given a raw land seismic record, such a task becomes nearly impossible; because unless gain or trace balancing is applied, only a small range of offsets and early times are visible to the eye. We would be unable to interpolate something we cannot see.

Numerical interpolation algorithms struggle in the same way. Many approaches can only be applied to small data windows or to amplitude-processed data because the transform-domain representation they use is strongly affected by the dynamic range. Consider, for example, a plane wave of constant amplitude (Figure 2.3(a)). The Fourier representation of this object is also a line with a few nonzero coefficients (Figure 2.3(c)). However, if one introduces an offset-dependent amplitude decay on the order of $1/r$, where r is the offset, the spectral representation changes: a large region of non- zero coefficients surrounds the previously sparse line (Figure 2.3(b) and Figure 2.3(d)). Plane waves with decaying amplitude do not have sparse frequency-domain representations, causing attempts at signal recovery to fail if the algorithm relies on sparsity. In the case of local transforms such as the wavelet or curvelet transforms, large transform coefficients correspond to strong events, enabling much better recovery of signals with decaying amplitudes. We compare wavelet and curvelet domain data recovery schemes to overcome the dynamic range problem without the necessity of amplitude preprocessing. This approach enables interpolation of raw land seismic data and typically aliased surface waves, which in turn has the potential to solve some of the key near-surface challenges (Keho and Kelamis, 2012) and improve reservoir characterization.

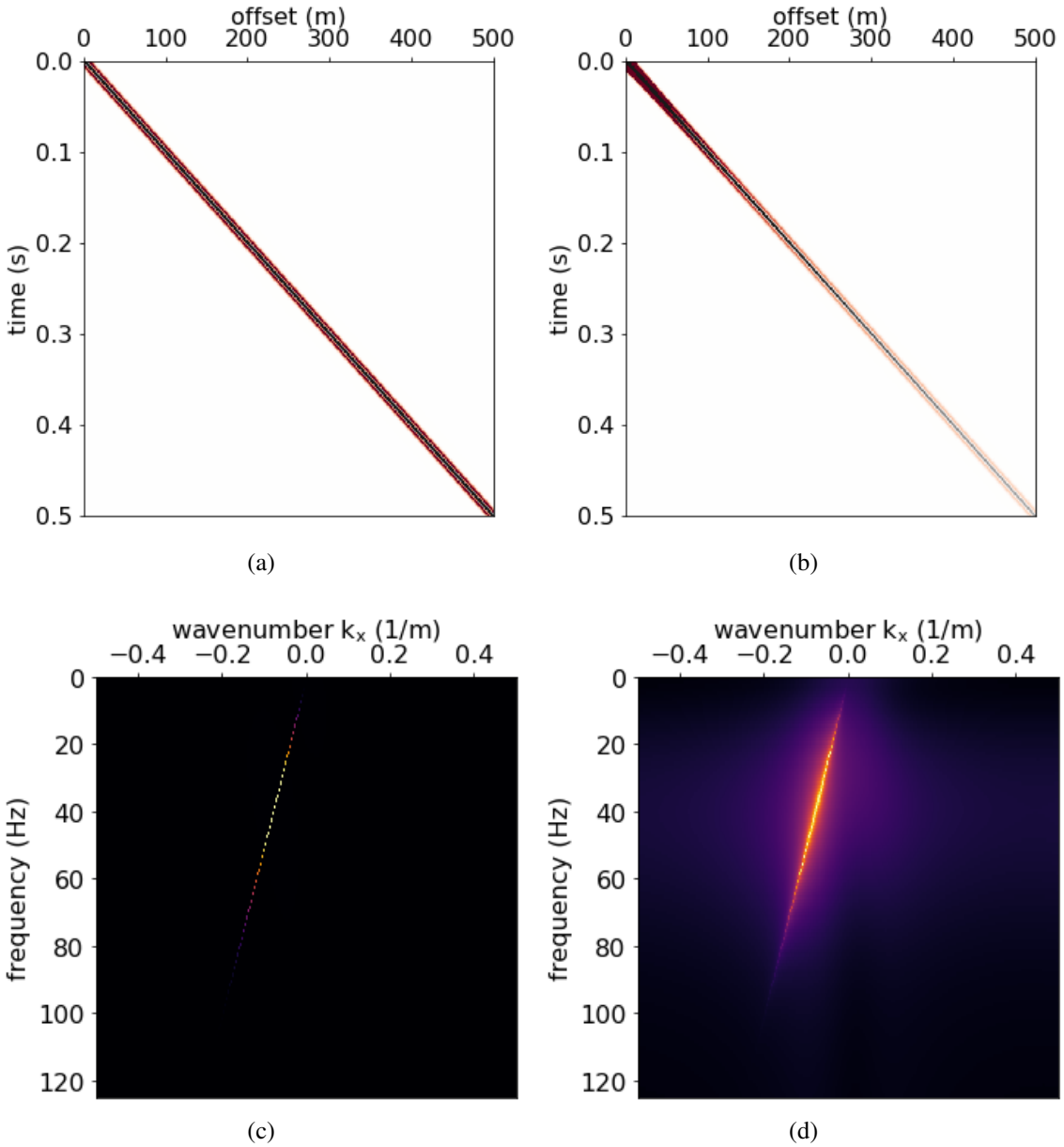


Figure 2.3 Plane wave (a) of constant amplitude and (b) with amplitude decay proportional to $1/r$. (c and d): Fourier domain representations of (a and b). Note that the Fourier representation of (b) is much less sparse than that of (a).

2.3 Theory

2.3.1 Sparse Signal Recovery

Consider an $N-$length signal $\mathbf{m}$ that can be represented as a vector of coefficients $\boldsymbol{\alpha}$ in some basis or dictionary expansion: $\mathbf{m} = \boldsymbol{\Phi}^T\boldsymbol{\alpha}$. $\mathbf{m}$ is said to be sparse if only $K \ll P$, where $P \geq N$ is the number of the dictionary coefficients in $\boldsymbol{\alpha}$, is non-zero. More commonly for seismic data, $\mathbf{m}$ is compressible when sorted coefficients $\boldsymbol{\alpha}$ decay rapidly enough to zero, so that $\boldsymbol{\alpha}$ can be well-approximated as sparse using a small number of large magnitude coefficients (Baraniuk et al., 2010).

In seismic data acquisition, we acquire $\mathbf{d} = \mathbf{Sm}$, where $\mathbf{d}$ is the recorded data, $\mathbf{S}$ is a sampling matrix, and $\mathbf{m}$ is the full data volume needed for processing and inversion. The matrix $\mathbf{S}$ in this instance represents the layout of sources and receivers, and, in the case of simultaneous acquisition, shot timing as well. Using the techniques from CS, it is possible to recover the full data volume from the sparse acquisition under certain assumptions.

Successful recovery of $K-$sparse or compressible signal depends on three key components: the sampling strategy, the sparsifying transform, and the sparsity-promoting recovery algorithm. Results from CS suggest that sparse signals can be recovered without loss of information if the sampling matrix satisfies the restricted isometry property (RIP) (Baraniuk, 2007). RIP is satisfied with high probability for Gaussian matrices (each entry is independent and follows a normal distribution) and random Bernoulli matrices (entries are ± 1 with equal probability) or when sampling nonuniformly Fourier-sparse signals. Depending on the choice of the sampling matrix, the number of measurements to recover a K-sparse signal is at $M = \mathcal{O}(K \log(N/K))$, where N is signal length and M is the number of measurements. However, this result may not hold for nonuniform sampling in other domains (e.g., for nonuniform sampling of a wavefield in conjunction with wavelet or curvelet transform).

Seismic data can be represented in the sparse domain $\boldsymbol{\alpha}$ as $\mathbf{d} = \mathbf{S}\boldsymbol{\Phi}^T\boldsymbol{\alpha}$. The matrix that ideally satisfies the RIP in this case is $\mathbf{S}\boldsymbol{\Phi}^T$. If $\mathbf{S}$ has sample locations chosen uniformly at random (meaning that each combination of the given number of sample locations is equally

probable) with a sufficient number of measurements ($M = \mathcal{O}(K\text{polylog}(N/K))$) and $\mathbf{\Phi}$ is a Fourier transform, the RIP is satisfied with high probability and the sparsity promoting recovery can be achieved by solving the following ℓ_1 optimization problem:

$$\tilde{\boldsymbol{\alpha}} = \arg\min_{\boldsymbol{\alpha}} \|\boldsymbol{\alpha}\|_1 \text{ subject to } \mathbf{d} = \mathbf{S}\mathbf{\Phi}^T\boldsymbol{\alpha}. \tag{2.2}$$

However, for other transforms (such as the wavelet transform or the curvelet transform), there is no practical algorithm to compute RIP constants (Herrmann, 2010) and provide similar recovery guarantees.

Finding a good sparsifying transform for seismic data is the key for applying techniques from compressive sensing to infill the missing traces. The sparsifying capabilities of a transform can be quantified by approximating a target signal with the K largest transform coefficients and computing the approximation error. In the following, we review a couple of popular choices for a sparse domain and highlight their benefits and pitfalls.

2.3.2 Fourier Transform

The classic D-dimensional Fourier transform is an excellent tool to examine spatiotemporal frequencies present in seismic data. The transform is orthogonal, and its computational complexity is $\mathcal{O}(N \log N)$. One advantage of the Fourier-domain representation of seismic data is that it is straightforward to interpret which coefficients represent the majority of coherent seismic energy. However, Fourier coefficients are not localized in time and space, making it challenging to identify what features of data are represented by specific coefficients. Furthermore, the typically high number of nonzero Fourier coefficients necessary to represent the data depends on sampling in time and space. In other words, the Fourier representation of seismic data is rarely sparse. Figure 2.4 shows the non-linear approximation of a 2D field seismic record using $\rho = K/P$ fraction of the largest transform coefficients (P represents the number of transform coefficients) in different domains. The approximation error is quantified by S/N that reflects the power of the original signal to the power of non-linear approximation error. The Fourier domain is the least sparse of all tested, as discussed next.

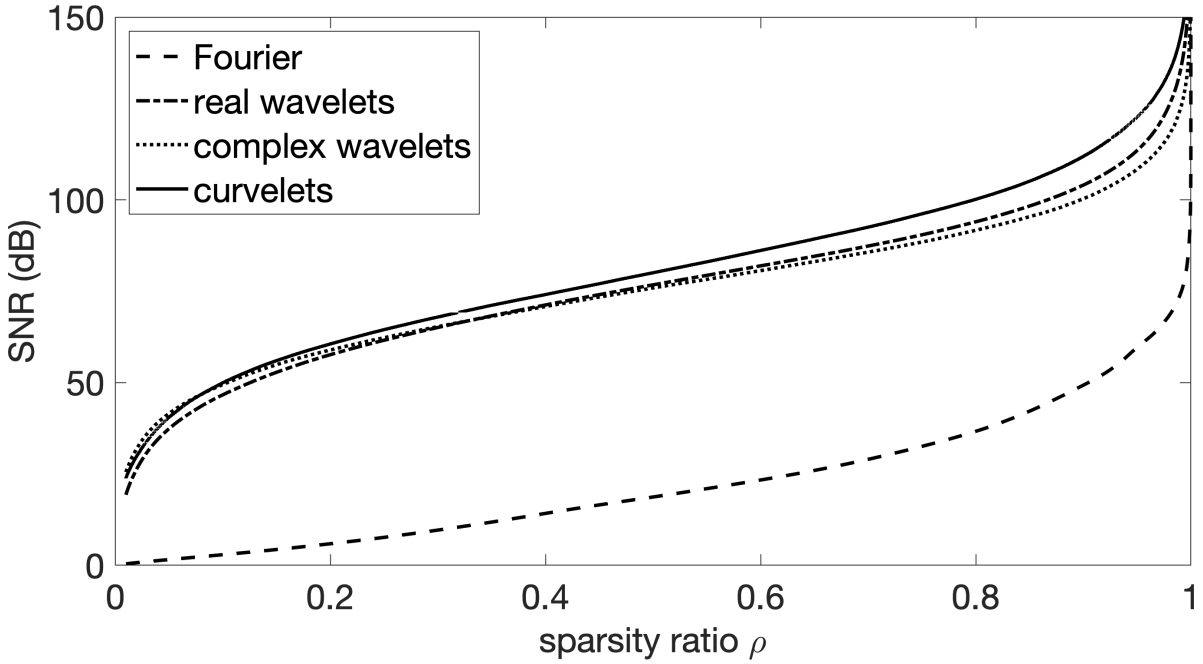


Figure 2.4 The nonlinear approximation error of a seismic field record as a function of sparsity ratio $\rho = K/N$, where N represents the total number of coefficients in the transform domain and K is the number of top magnitude coefficients used in the nonlinear approximation. The Fourier transform performs the worst, whereas the curvelets and wavelets are closely matched.

2.3.3 Discrete Wavelet Transform

The introduction of wavelets in geophysics (Morlet et al., 1982a,b) has been motivated by the desire to examine the characteristics of seismic reflection signals (amplitude, shape, frequency and phase) with time. Although it is possible to use the windowed Fourier transform for this purpose, the achieved time-frequency resolution is often insufficient to identify subtle signal characteristics due to the fixed window size. The wavelet transform (Mallat, 1989a) achieves improved resolution and time-frequency characteristics of nonstationary signals by varying the window size based on the scale. The continuous wavelet transform can be sampled to obtain the discrete wavelet transform (DWT) which forms an orthonormal basis for a large class of wavelets (e.g., Daubechies, symlets or coiflets). This feature makes wavelets an attractive choice for multiresolution approximations of signals (Mallat, 1989b), allowing us to analyze signal characteristics in different frequency bands with high accuracy and to localize them in space-time

domain at the same time.

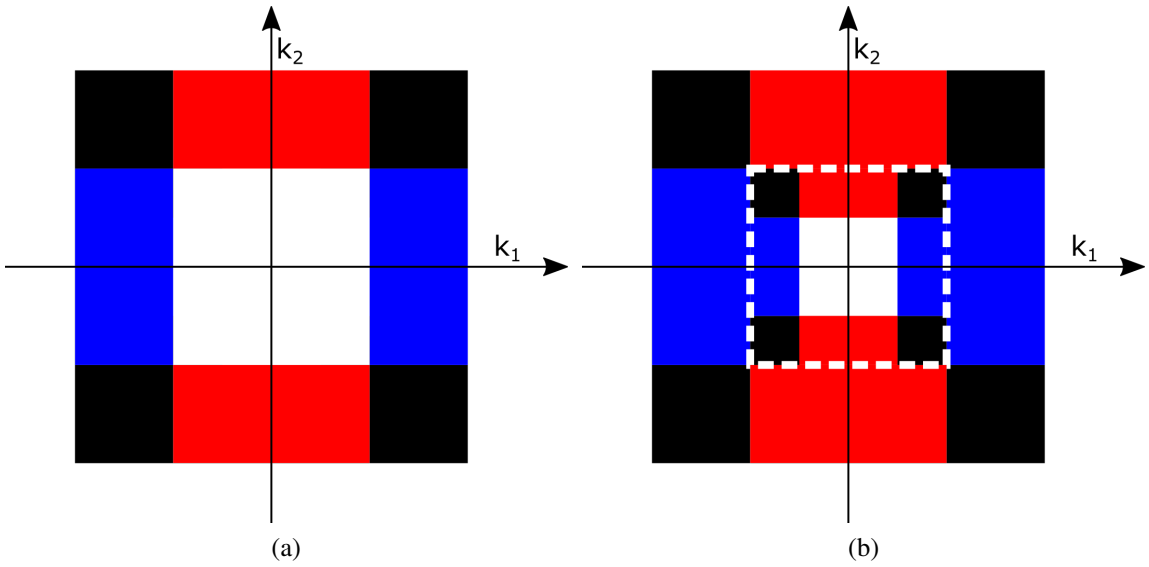


Figure 2.5 Idealized Fourier domain support for 2D real wavelet transform. (a) One level decomposition with white box in the middle representing low-frequency signal approximation. The blue, red and black represent vertical, horizontal and diagonal details, respectively. Level two wavelet decomposition in (b) results from further decomposing the white box in (a).

The multiresolution character of the wavelet transform can be understood by examining the idealized Fourier domain support of wavelet coefficients. Figure 2.5 shows that support for a 2D signal, but in higher dimensions similar reasoning applies. A first-level wavelet decomposition splits a signal into two parts along each dimension: the low- part and the high-frequency parts. The 2^D colors in Figure 2.5(a) represent different bandwidths along signal dimensions. Thus, in two dimensions, the white box corresponds to the low frequency along both axes, the black boxes correspond to high frequency along both axes, and the red and blue represent low-high and high-low frequency portions of the signal, respectively. The black, blue and red boxes can also be linked to directional sensitivity of the wavelet transform. Because real-valued signals have symmetric amplitude spectra, the real DWT cannot distinguish between events dipping to the left or the right. Thus, for a 2D case, only three directions can be distinguished: $0°, \pm 45°$, and $90°$. More generally, in D dimensions the sensitivity is along $2^D - 1$ directions. Figure 2.5(b) depicts the second level wavelet decomposition of a 2D signal. Note that by decomposing further, we

increase the number of bands in which directional details can be observed. Thus, the real DWT allows examination of the signals in multiple bands, which results in a more sparse representation of HDR data such as wavefields when compared to the Fourier transform (see Figure 2.4).

2.3.4 Complex Wavelet Transform

ℂWT is an enhancement to the real DWT offering attractive additional properties: near shift invariance and directional selectivity in two dimensions and higher dimensions (Selesnick et al., 2005).

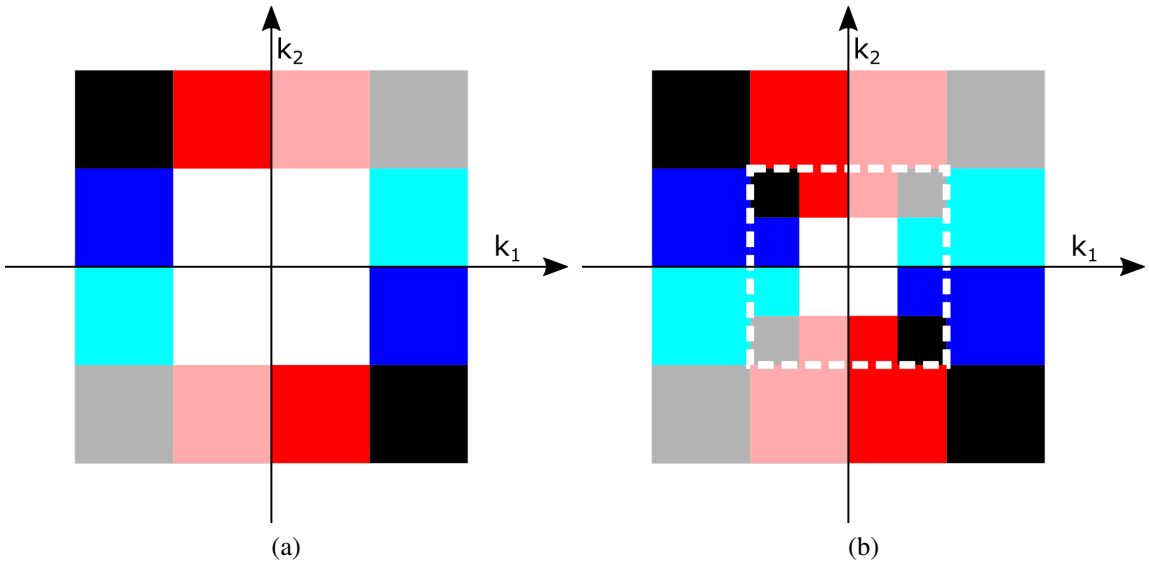


Figure 2.6 Idealized Fourier domain support for 2D dual-tree ℂWT . (a) One level decomposition with white box in the middle representing low-frequency signal approximation. The colors denote support of distinct detail subspaces. Level two wavelet decomposition in (b) results from further decomposing the white box in (a).

In a classic DWT, a small shift of a signal greatly perturbs the wavelet coefficients around signal singularities such as zero traces - an undesirable property while using overlapping spatial windows for seismic data reconstruction. Furthermore, the limited directional selectivity of DWT does not distinguish between left- and right-dipping events on a seismic record. The ℂWT is able to overcome the shortcomings of DWT by replacing the classical real wavelet with a complex,

approximately analytic wavelet:

$$\psi_C(t) = \psi_r(t) + \mathrm{i}\psi_i(t), \tag{2.3}$$

where real functions $\psi_r(t)$ and $\psi_i(t)$ are even and odd, respectively. Similarly, $\mathbb{C}$WT uses a complex scaling function:

$$\phi_C(t) = \phi_r(t) + \mathrm{i}\phi_i(t), \tag{2.4}$$

with properties similar to those of the complex wavelet. The scaling function acts as a low-pass filter, while the wavelet function is a band-pass filter. Because the analytic signals have one-sided amplitude support in the Fourier domain, the $\mathbb{C}$WT is able to distinguish between events of opposing dip. Figure 2.6 shows the idealized Fourier domain support of complex wavelet coefficients in two dimensions (similar principles apply in higher dimensions). Note that the number of distinct detail subspaces increased from three for DWT to six for $\mathbb{C}$WT.

One way of achieving an approximately analytic wavelet is by forming a slightly redundant frame such that both $\psi_r(t)$ and $\psi_i(t)$ form orthonormal or biorthogonal bases (Selesnick et al., 2005). An implementation following this approach is the dual-tree $\mathbb{C}$WT (Kingsbury, 2001) based on filter banks. In fact, the dual-tree $\mathbb{C}$WT can be computed using the infrastructure used for classical DWT, though with specially designed filters (Selesnick et al., 2005). This makes the complex wavelet transform fast to compute ($\mathcal{O}(N)$ complexity) and a 2^D times redundant tight frame (independent of the decomposition level) for $D-$dimensional signals with the added benefit of enhanced angular sensitivity to $\pm 15^\circ$, $\pm 45^\circ$ and $\pm 75^\circ$.

To understand how the dual-tree $\mathbb{C}$WT works, consider the complex wavelet decomposition of a binary image of a circle shown in Figure 2.6. A circle is a simple object with all angles equally represented; thus it is optimal for assessing the directional selectivity of the $\mathbb{C}$WT. Because the transform yields complex-valued coefficients, we can interpret the decompositions in terms of magnitude (Figure 2.7(a) and Figure 2.7(b)) and phase (Figure 2.7(c) and Figure 2.7(d)). Note that despite the uniform distribution of angles, the energy of complex wavelet coefficients is not distributed evenly between detail subspaces. This implies that the $\mathbb{C}$WT is not equally sensitive to

all angles in the data. The study of phase plots suggests that there is a phase shift of about $90°$ between the two trees except for the low frequency approximation. The shift ensures the enhanced directional selectivity of the $\mathbb{C}$WT.

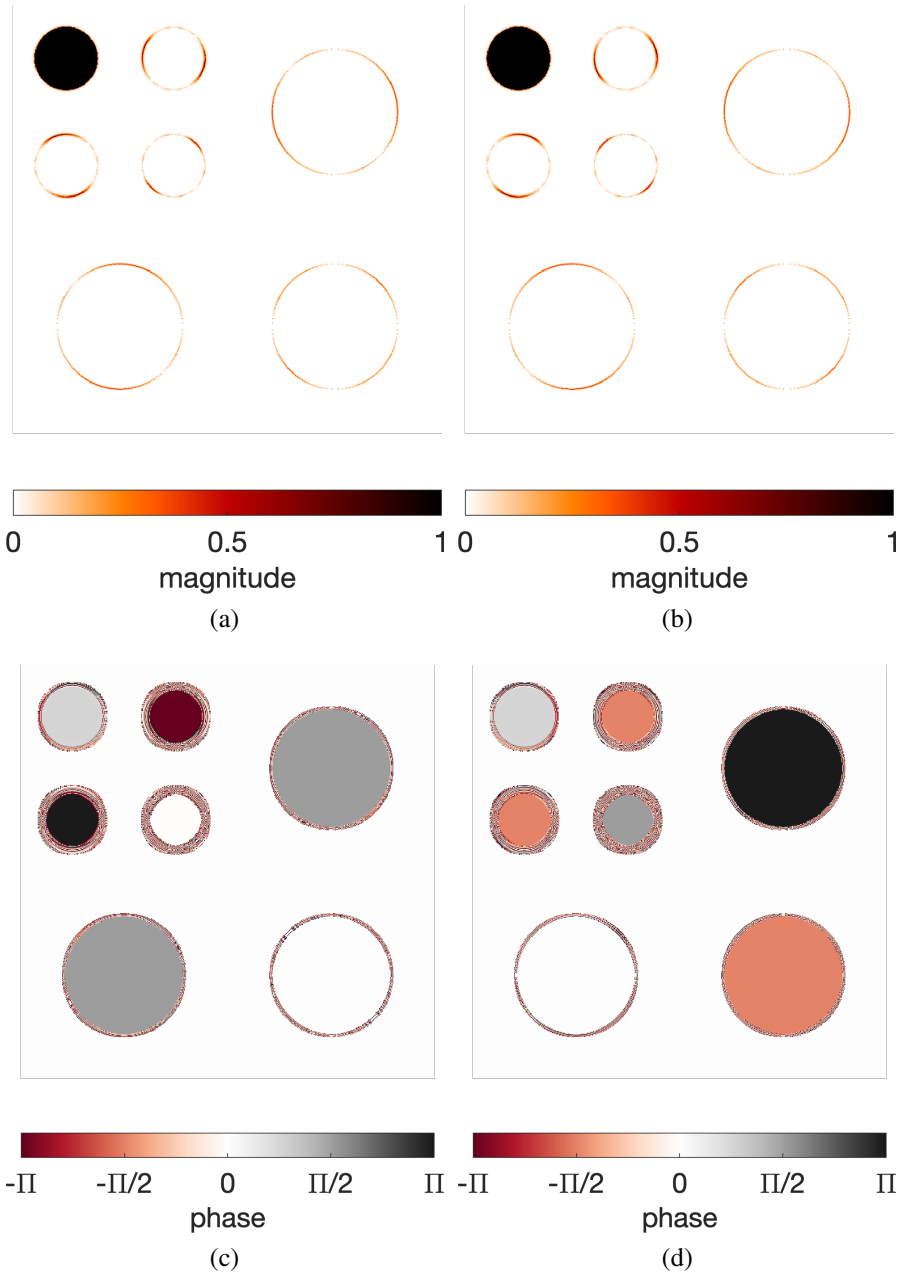


Figure 2.7 (a and b) Magnitudes and (c and d) phases of the two-level complex wavelet decomposition of a circle. Note the selective directional sensitivity and phase shift between the two trees.

The straightforward interpretation of complex wavelet coefficients combined with enhanced directional selectivity and limited redundancy of the $\mathbb{C}$WT makes it an attractive domain for sparsely representing seismic data. Note that in Figure 2.4, $\mathbb{C}$WT outperforms DWT when $\rho < 0.35$, making $\mathbb{C}$WT a good transform for sparsity-promoting recovery of seismic data.

2.3.5 Curvelet Transform

The curvelet transform can be thought of as a localized oriented Fourier transform or an anisotropic generalization of $\mathbb{C}$WT . Because curvelets are specified by scale, angle, and position, the transform contains location, orientation, and frequency information. These features come at an expense of a highly redundant representation. Redundancy depends on the number of scales and angles, but is much higher than that of $\mathbb{C}$WT. For the example presented in Figure 2.4, there are 7.2 times more curvelet coefficients than data samples.

Although the curvelet transform is said to provide an optimally sparse representation of wave propagators (Candès and Demanet, 2005), the gains may not be as significant as expected when applied to field seismic data. One pitfall of field data is the often prominent issue of aliased energy. Because the Fourier transform is at the core of the digital implementation of a curvelet transform (with a similar computational complexity of $\mathcal{O}(N \log N)$) (Candès et al., 2006a), it may spread aliasing artifacts across the entire domain instead of keeping them localized (Yu et al., 2017). Note that in Figure 2.4, the curvelets outperform real and complex wavelets for $\rho > 0.1$, but not as much as one would expect from a similar experiment with synthetic data.

2.4 Data Reconstruction Examples

We demonstrate data reconstruction with field data using real wavelets, complex wavelets and curvelets. We design our experiments for data reconstruction using compressive seismic acquisition, which follows the uniform random downsampling strategy, similar to Herrmann (2010). Let $\delta = \frac{n}{N}$ denote data decimation ratio, where N is the number of traces in the undecimated data and n is the number of traces remaining after random data decimation. For every value of δ, we generate 100 realizations of uniform random sampling, with different

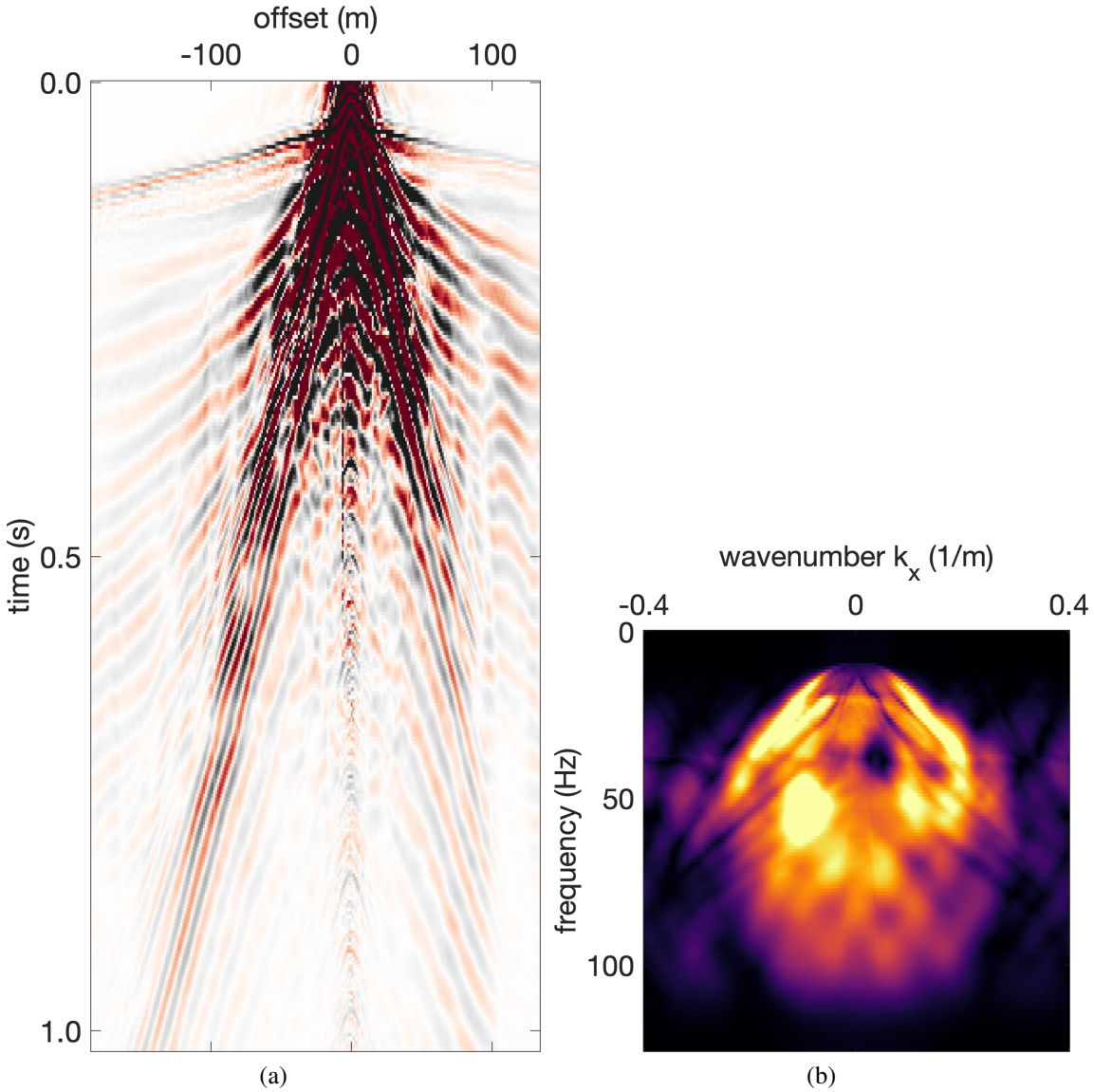
geometries of missing traces. We then run the reconstruction for each realization to evaluate the impact of gap pattern on the reconstructed data.

We formulate the data reconstruction as sparsity promoting ℓ_1 optimization problem in the tested domains. Because field data inevitably contain noise, instead of solving equation 2.2, we change the constraints such that $\|\mathbf{d} - \mathbf{S}\mathbf{\Phi}^T\boldsymbol{\alpha}\|_2^2 \leq \epsilon$, where ϵ is a noise level inferred from log-amplitude plot similar to Figure 2.2. This is to ensure that there is no data over-fitting which could introduce high frequency artifacts to the reconstructed wavefields. The optimization is solved using the spectral projected gradient solver (van den Berg and Friedlander, 2009). To quantify the quality of the reconstruction, we use the S/N ratio, defined as:

$$S/N = 20\log_{10}\Big(\frac{\|x\|}{\|x - \hat{x}\|}\Big), \tag{2.5}$$

where x represents the original full data and $\hat{x}$ represents the reconstructed data. We also examine the individual reconstructions, data difference and the FK spectrum of the reconstructed data to understand the domain-specific reconstruction artifacts and the effect of geometry on the reconstructed wavefields.

The undecimated shot is shown in Figure 2.8(a) and the corresponding FK spectrum in Figure 2.8(b). The data come from a mountainous region. The dominant lithologies are shale and sandstone with occasional intrusions of igneous rocks. Note that this field record showcases the described challenges with land data: slow and dispersive surface waves, source-generated noise, and large dynamic range. Furthermore, field data are critically sampled, i.e. they occupy the entire FK space. The real benefit of compressive sensing is when one can decimate such data (which would introduce aliasing if done regularly) and still recover the nonaliased wavefield without the loss of information.



(a) (b)

Figure 2.8 (a) Field data and (b) their spectrum used in reconstruction experiments. Note that the field data are critically sampled.

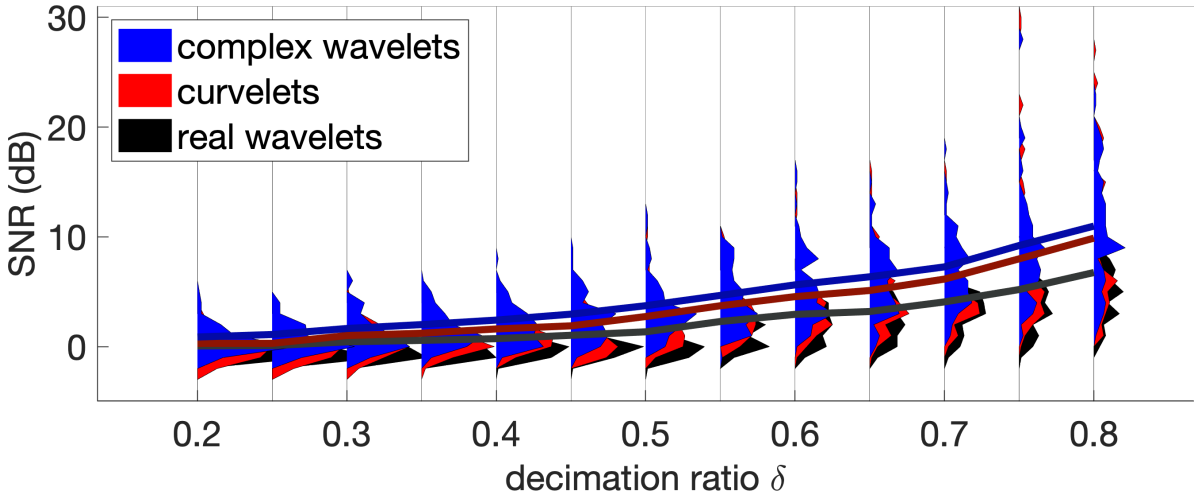


Figure 2.9 Summary of the reconstruction results for field data in Figure 2.8. Solid lines indicate the average performance of each domain.

The reconstructions for real wavelets, curvelets and complex wavelets are summarized in Figure 2.9. On average, the complex wavelets perform the best in terms of S/N, followed by curvelets and real wavelets. The reconstruction performance is consistently poor when only a small fraction of original traces are kept ($\delta < 0.3$). If more that 70% of original traces are kept ($\delta > 0.7$), there is a big spread in the reconstruction quality in all domains, with the best and the worst reconstruction differing by as much as 30dB. This suggests that the geometry of missing traces has a big impact on the results.

Because we use the uniform random sampling, there is no control over the size of the gaps. Scenarios in which gaps are bigger and contain unique information not present in the remaining traces yield poor reconstructions compared to the scenarios with more but smaller gaps.

Figure 2.10, Figure 2.11, and Figure 2.12 (pages 33-35) show the worst, average, and the best reconstructions in terms of S/N, the data difference and the FK spectrum of the reconstructed data for $\delta = 0.5$ using complex wavelets, curvelets and real wavelets, respectively. Note that although the reconstructed data may look similar within the same reconstruction domain, the S/N is related to the FK spectrum of the reconstructed data: high values of S/N imply higher fidelity of spectral representation and vice versa. Accurate frequency content, especially the low temporal

frequencies, is essential in many applications for reservoir characterization, for example the acoustic impedance inversion or full waveform inversion. Poorly reconstructed data may hinder instead of help the subsequent processing and inversions.

Comparing the reconstructions from different domains we note that complex wavelets and curvelets perform similarly, while the real wavelets struggle the most to reconstruct the data with the accurate spectrum. The reconstructions utilizing DWT suffer from aliased coefficients - a known shortcoming of a critically sampled wavelet transform (Selesnick et al., 2005). Thus, the DWT is not a good candidate for sparse recovery with field data. The main difference between the complex wavelets and curvelets is the reconstruction artifacts. Complex wavelets are unable to reconstruct the first arrivals, resulting in significantly lower amplitudes where traces were missing. This behavior stems from the fact that first arrivals contain high frequencies which tend to be represented by smaller transform coefficients than the low frequencies. Furthermore, as shown in Figure 2.7, the $\mathbb{C}$WT is not equally sensitive to all angles in the data so without providing additional structural information in the recovery process, the angles represented by small transform coefficients may be lost. Curvelets have better angular coverage and do not struggle with the first arrival, but instead exhibit strong "wrap-around" artifacts along the time axis, likely due to their implementation using wrapped Fourier transforms. Given the consistently better reconstruction quality and reduced artifacts in the reconstructed wavefields, complex wavelets are the best choice for field data reconstruction out of all tested options.

2.5 Discussion

The geometry of the missing traces plays a key role in the successful recovery of nonaliased wavefield. All local transforms struggle with gaps comparable to the transform support. Furthermore, the reconstruction algorithm cannot create new information. Suppose that a diffraction is present in the region of missing receivers. Unless we use data from additional shots which register that diffraction from a different angle, we cannot recover the diffraction inside a gap as it is missing from the registered data. Thus, the fidelity of the reconstructed data improves

as more information is included (e.g., multiple shots). Furthermore, if the seismic data are viewed as a multidimensional cube, one can resort the data in a way which makes gaps appear smaller. Changing the sorting may also have a benefit of decreasing the lateral variability of data (e.g., in common offset gathers), which in turn enhances the sparsity in many transform domains and has the potential to improve the reconstruction quality.

We show in Figure 2.4 that curvelets and wavelets can sparsely represent seismic data, but as evidenced by the presented examples, sparsity is not sufficient to ensure high-quality reconstruction. Additional considerations include the implementation details of the transform as described before and the coherence between sampling scheme and data representation in the transform domain. Ideally, the sampling scheme and the data sparsifying transform should be incoherent, such as with uniform random sampling and the Fourier domain. However, the Fourier transform cannot sufficiently sparsify the challenging land seismic data and other transforms have some degree of coherence with randomly missing traces. Thus, randomizing along more than one dimension (e.g. by randomizing shot timing during simultaneous acquisition) is beneficial, as shown by Herrmann (2010).

The benefits of compressive sensing come from acquiring less data than required by Nyquist theorem for regularly sampled data. When discussing the data decimation ratio, it is important to keep the Nyquist wavenumbers in mind. Suppose we decimate oversampled data uniformly by half. We would be able to infill the decimated traces with little effort using one of the many techniques for interpolation of nonaliased, regular data. With half of the traces missing at random, the challenge is driven by the gap size, but reconstruction tends to be successful. However, when sampling below Nyquist requirements, the reconstructions can range from poor to good (see Figure 2.9), strongly depending on the geometry of the missing traces. This has important implications for acquisition design. To acquire data at a lower cost with fewer receivers than needed by the Nyquist theorem, one needs to develop a reconstruction strategy which provides recovery guarantees with high probability. Within the framework presented here, ensuring good quality reconstruction still requires a high channel count due to the very slow surface waves.

One promising future direction for complex wavelet-based data reconstruction is an extension to 5D. Because seismic data do not vary in space as rapidly as in time, 5D $\mathbb{C}$WT of seismic data is likely to be sparser than a 2D $\mathbb{C}$WT, requiring fewer measurements for good reconstruction. 5D interpolation is demonstrated to work well with windows of data in the Fourier domain. Using wavelets would facilitate the analysis of larger subsets, limited only by the available computer memory. An additional benefit of extending the framework to higher dimensions is enhanced directional selectivity: the number of directions is $2^{2D-1} - 2^{D-1}$, which yields 28 directions in 3D and 496 directions in 5D. This would further help to sparsely represent the seismic data.

2.6 Conclusions

We demonstrate that ℓ_1 sparsity-promoting optimization utilizing the $\mathbb{C}$WT is an attractive alternative to curvelet domain reconstruction for challenging land seismic data. $\mathbb{C}$WT is faster and less redundant than the curvelet transform; in addition, it is straightforward to extend to four or more dimensions. Furthermore, when applied to critically sampled or aliased field data, $\mathbb{C}$WT keeps the aliasing artifacts local instead of spreading them throughout the entire domain, thus limiting artifacts in the reconstructed data. Despite superior angular sensitivity, curvelets are not as successful on field data used in this chapter due to the Fourier-domain based implementation spreading aliasing effects and resulting in strong artifacts in the reconstructed data.

Random sampling in conjunction with the complex wavelet-domain sparsity-promoting algorithm allows us to reduce the number of channels needed to acquire nonaliased wavefields below the limit imposed by the Nyquist theorem, but the exact decimation ratio strongly depends on the specific receiver geometry. The successful recovery is characterized by high S/N values, which are good indicators of Fourier domain fidelity of the reconstructed data. A sparsity promoting data preprocessing or extension to higher dimensions is needed to increase the geometry robustness of the data reconstruction.

2.7 Acknowledgements

We would like to acknowledge Dawson and Sercel for providing equipment and technical support at the CSM Geophysical Field Camp. We extend our thanks to GTI for allowing us to test wireless nodes and conduct dense sampling experiments. SEG provided partial financial support of the Field Camp. We also acknowledge the CSM undergraduate students who conducted the field data acquisition.

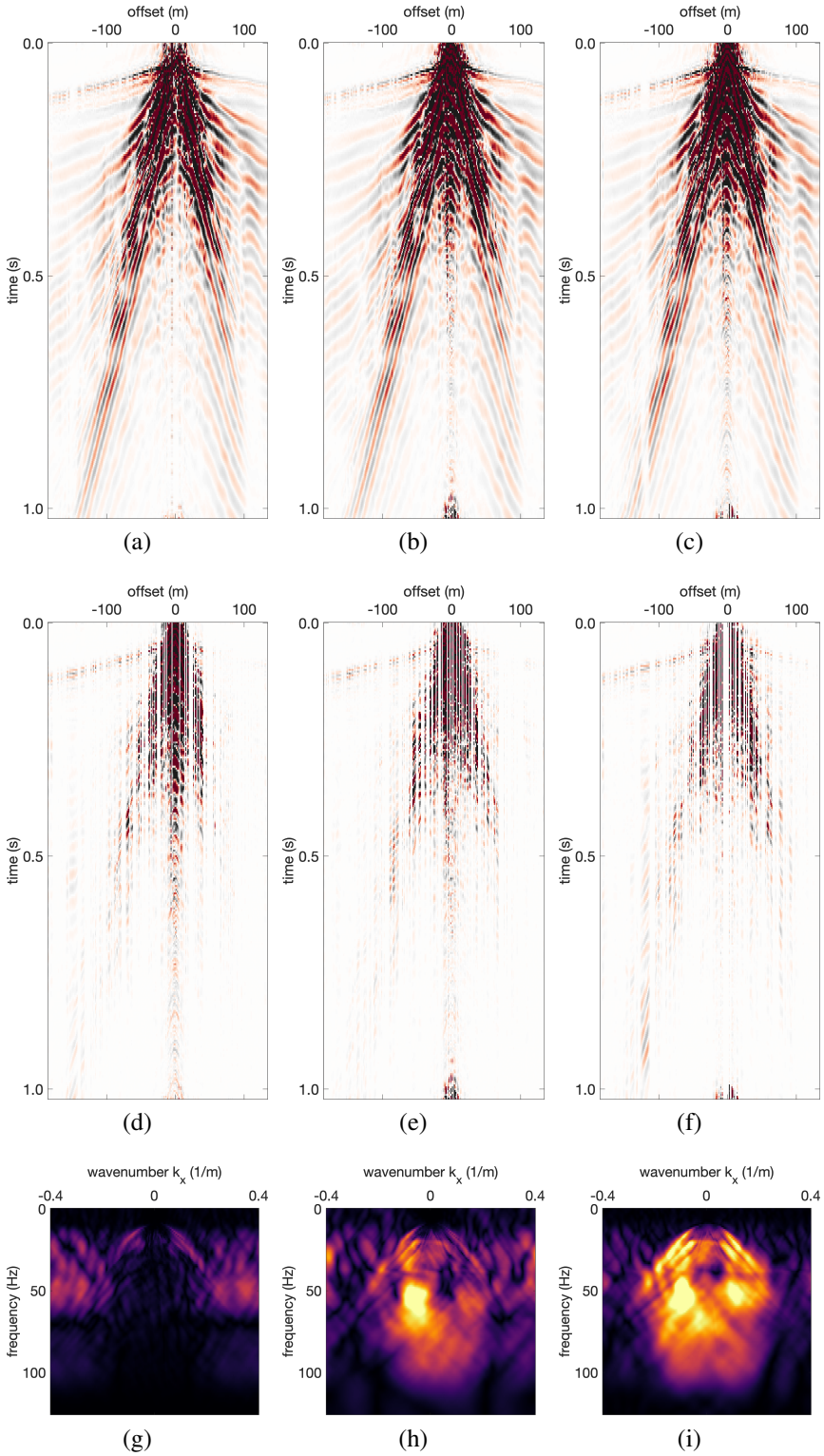


Figure 2.10 From left to right: the worst, average and best case complex wavelet-based reconstruction for $\delta = 0.5$. (d-e) show the data difference and (g-i) show the $f - k$ spectrum of the reconstructed data.

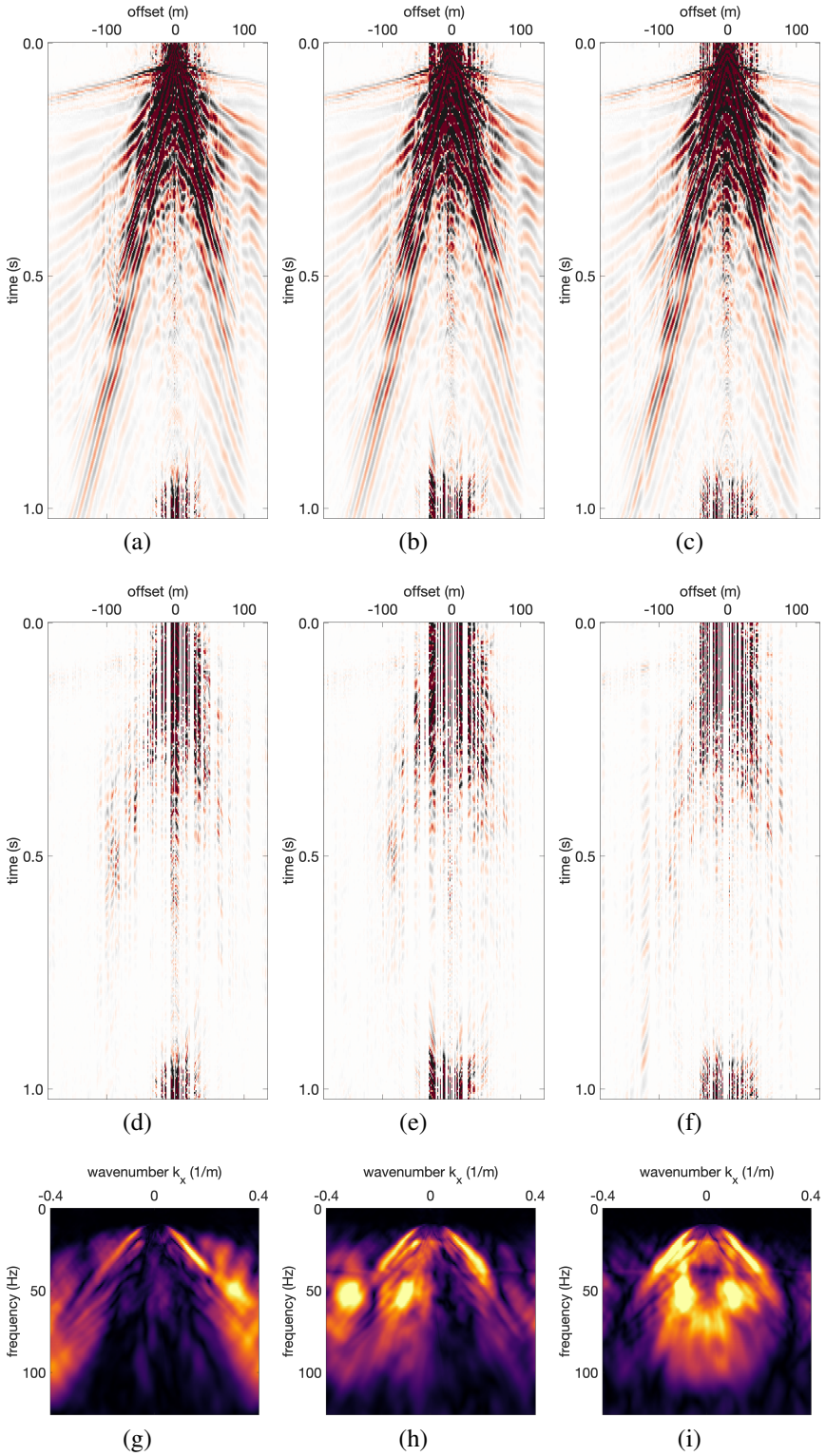


Figure 2.11 From left to right: the worst, average and best case **curvelet**-based reconstruction for $\delta = 0.5$. (d-f) show the data difference and (g-i) show the $f - k$ spectrum of the reconstructed data.

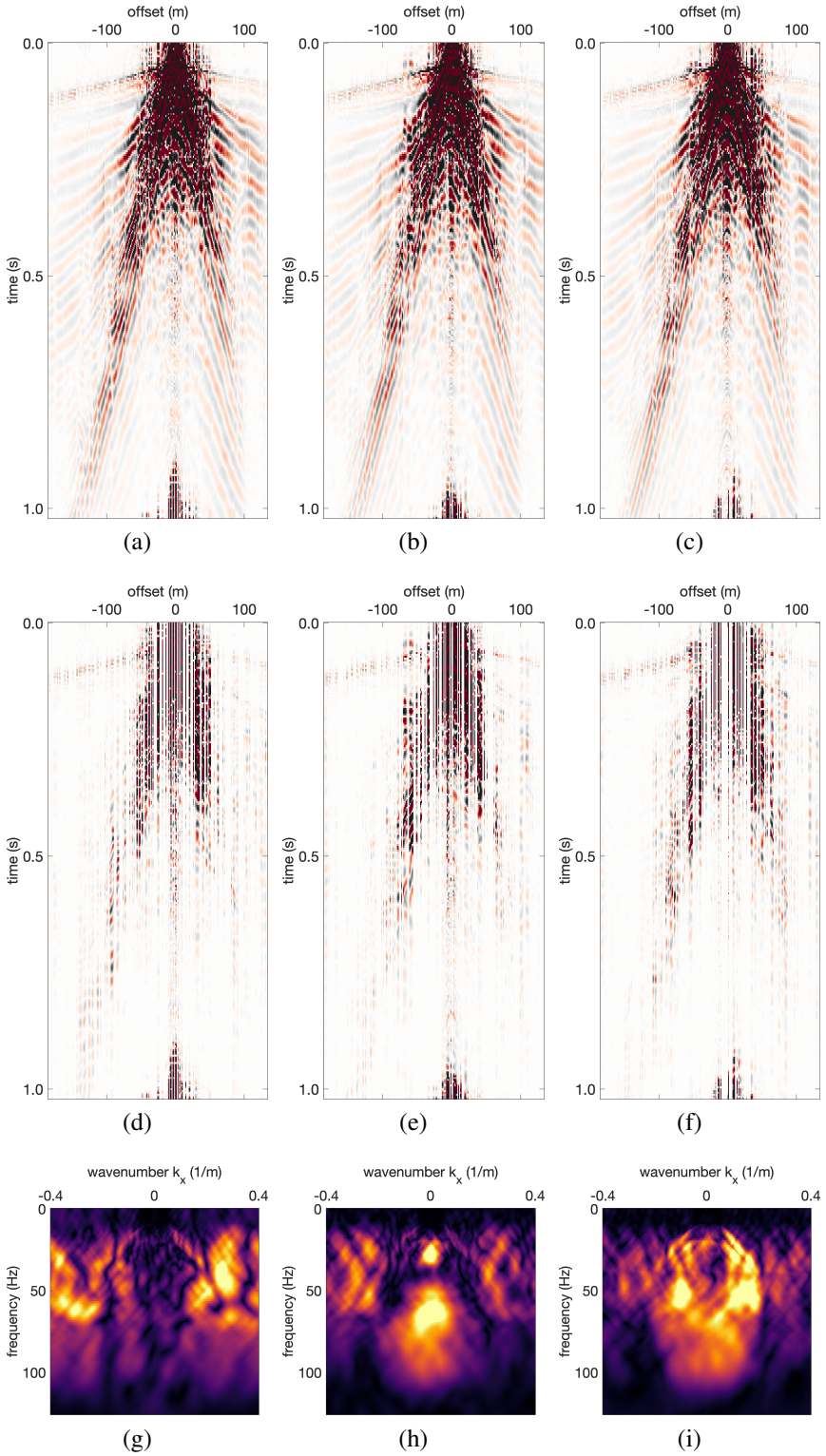


Figure 2.12 From left to right: the worst, average and best case **real wavelet**-based reconstruction for $\delta = 0.5$. (d-f) show the data difference and (g-i) show the $f - k$ spectrum of the reconstructed data.

CHAPTER 3

HIGH DYNAMIC RANGE LAND WAVEFIELD RECONSTRUCTION FROM RANDOMIZED ACQUISITION

A paper submitted to *Geophysics*

Iga Pawelec[1,2], Paul Sava[2]

Compressive sensing (CS) is an alternative to regular Shannon-Nyquist sampling capable of capturing the same amount of information about seismic wavefields from much reduced measurements. It relies on randomized sampling patterns and a sparse data representation in a domain of interest to reconstruct the regularly sampled object. This technology applied to land seismic acquisition design promises to deliver the same data quality as a regular acquisition with fewer samples or superior data quality with the same number of samples. Compressive sensing is a strong contender for helping to overcome some of the challenges associated with land seismic data, providing the ability to affordably record the intricate wavefields propagating in the near surface without aliasing. Recording high-quality nonaliased wavefields can lead to significant improvements in the understanding of the near-surface geology and in suppressing noise. However, near-surface wave phenomena tend to propagate with low velocities and contain the majority of the source-generated energy, leaving the much weaker reflection signals barely discernible. This results in data rich in high wavenumber content and with amplitudes of interest that span several orders of magnitude. When handling high dynamic range non-stationary data, the Fourier domain is not optimal for providing sparse representation - a necessary condition for successful application of compressive sensing. In contrast, a discrete complex wavelet transform can localize high energy features, has good directional selectivity, and is near-shift invariant. Combined, these properties allow complex wavelets to represent detail-rich wavefields with high dynamic range in a compact form which is essential for a sparsity-promoting reconstruction.

[1]Primary researcher and author.
[2]Colorado School of Mines, Golden, Colorado, United States of America.

Another important ingredient of a successful CS data reconstruction is the choice of an appropriate sparsity-promoting algorithm. Since complex wavelets provide a multiscale signal representation, global approaches that simply rely on the distribution of high energy coefficients in the transform domain are not the best choice. To account for differences in coefficient energy distributions at different decomposition levels of the complex wavelet transform, we develop a scale- and orientation- dependent iterative soft thresholding scheme (IST) for reconstructing land wavefields. Our approach requires little parametrization, is easy to implement, and is robust to reconstructed wavefield sampling grid and dynamic range. We test the developed soft thresholding scheme on different wavefields with randomly missing traces, and compare the data reconstructions to the spectral projected gradient solver and projection onto convex sets. We quantify the reconstructions by a direct comparison of Fourier coefficients between fully sampled and reconstructed wavefields. Taking log10 of Fourier coefficients prior to computing the quality metric de-emphasizes the importance of magnitude match while highlighting Fourier coefficient support accuracy which usually translates into good structural fidelity of reconstructed data. We find that our IST performs consistently among all examples, yielding high structural fidelity while performing gentle denoising.

3.1 Introduction

Land seismic data suffer from a number of quality issues, such as low signal-to-noise ratio, phase distortions, or environmental noise. The majority of these problems stem from the near-surface noise (Regone, 1997) and can be addressed by acquiring nonaliased wavefields. However, achieving this goal following the classical sampling theorem (Jerri, 1977; Shannon, 1948) is prohibitively expensive due to the required trace density. The near-surface materials often consist of loose soil and poorly consolidated sediments, giving rise to strong and slow surface waves, guided waves, and scattering noise, especially in presence of lateral heterogeneities (Stork, 2020). The waves traveling in the near surface are only sensitive to the upper tens or hundreds of meters, provide no information on the deeper structure, and are

considered noise for most imaging applications. However, it is recognized that high-fidelity recording of the surface waves is valuable in building near-surface models and computing reliable static corrections (Monk, 2020; Socco et al., 2010a; Strobbia et al., 2011). The need to record nonaliased surface waves is thus twofold: (1) to constrain the near surface model, and (2) to remove them from seismic records without compromising the significantly weaker reflections.

The sampling needed for regular nonaliased recording is dictated by the slowest phase velocity expected in the seismic record and the highest frequency present in that signal. Figure 3.1(a) and Figure 3.1(b) show examples of densely sampled land records with their corresponding $f - k$ spectra in Figure 3.1(d) and Figure 3.1(e). Note that the strong surface-wave energy has phase velocities below 200 m/s and reaches frequencies of 50-60 Hz. In these scenarios, the regular sampling interval would have to be less than 2 m which is not feasible for large scale surveys: terrain obstacles and access restrictions often prevent placement of sources and/or receivers at desired locations. Stork (2023) argues that some locations for sources or receivers can be predicted as noisy and problematic prior to the survey, furthering the case for deliberate irregular acquisition.

Enabled by recent advances in nodal acquisition technology (Freed, 2008; Manning et al., 2018), compressive sensing (CS) emerges as a feasible methodology for acquiring data on an irregular grid and with smaller channel count than required for a regular survey, allowing for recovery of nonaliased wavefield on a regular grid as a post-acquisition step. Developed in the signal processing community, CS is one possible solution to the problem of retaining the needed information about the signal of interest from reduced measurements. Early works by Candès et al. (2006b) and Donoho (2006) establish the CS theoretical framework. There are two main elements to the CS approach: selection of sampling scheme that is as incoherent as possible with respect to the domain in which data are considered sparse (Candès and Romberg, 2007; Candès and Wakin, 2008), and the appropriate choice of reconstruction strategy.

Different sampling schemes have been considered for geophysical applications. Hennenfent and Herrmann (2008) show that for localized transforms such as curvelet transform it may be

beneficial to constrain randomized sampling with the maximum gap size, ensuring that there are no regions containing only zeros within the entire operator support. Naghizadeh and Sacchi (2010b) investigate multidimensional sampling operators in the Fourier domain and conclude that randomly sampled multidimensional data should only suffer from a minor spectral distortion. Jiang et al. (2018) design a CS field trial defining their source and receiver sampling locations on regularly indexed grid accounting for access restrictions and implementing gap size control while minimizing aliased energy in the sparse domain. Though the benefits of CS for seismic data acquisition have been demonstrated (Blymyer et al., 2021; Jiang et al., 2019; Mosher et al., 2017), finding the optimal reconstruction strategy for specific sampling scheme and sparse domain remains an open research area.

One of the most popular choices for a sparse domain in seismic data acquisition and reconstruction is the Fourier domain. Liu and Sacchi (2004) introduce minimum weighted norm interpolation with spectral weights derived directly from data and adjusted iteratively. Trad (2009) extend the reconstruction approach to five dimensions. Though memory-consuming, such an extension is particularly favorable for seismic data sparse recovery: 5D objects tend to be more sparse than 2D or even 3D objects. Because wavefields usually vary the fastest along time, and less rapidly over space, adding spatial dimensions (sources and receivers) introduces reciprocal information, contributing to the sparsity of the entire object. For the same reason, successful reconstruction of 2D gathers tends to be more challenging, particularly on land, where amplitudes can vary by several orders of magnitude over short distances. The gathers shown in Figure 3.1(a) and Figure 3.1(b) are examples of 2D land wavefields showcasing that particular feature which we highlight by using a semi-logarithmic color normalization available in *matplotlib* (Hunter, 2007). The strongest energy is concentrated in the source region and can be four orders of magnitude higher than the energy 200 m away. For contrast, we also show a window from a synthetic gather in Figure 3.1(c). Note that, though the wavefield is intricate, the amplitude variability over distance is not significant. High dynamic range objects such as raw land seismic wavefields containing near offsets are not well-approximated as sparse in the Fourier domain (Pawelec et al.,

2021). The $f-k$ spectra in Figure 3.1(d) and Figure 3.1(e) illustrate why: rapid energy decay introduces large number of non-zero Fourier coefficients, and setting these coefficients to zero leads to substantial signal distortions. As a result, Fourier-based reconstructions require data windowing, range compression, or working on a single frequency slice at a time.

The discrete complex wavelet transform ($\mathbb{C}$WT) does not suffer from this problem. Complex wavelets are localized in time and space and have directional sensitivity (six orientations in 2D and twenty eight orientations in 3D). With linear computation time and only 2^D redundancy for $D-$dimensional signals, complex wavelets are particularly attractive for representing and analyzing non-stationary signals. In particular, $\mathbb{C}$WT found applications in denoising (Julayusefi et al., 2012; Peng et al., 2014; Yu et al., 2017, 2022) and discontinuity detection (Liang et al., 2019), but its use for data reconstruction is not well-studied despite indications that it shows promise, outperforming some popular contestants such as curvelets (Li et al., 2014; Pawelec et al., 2021).

Because complex wavelets can sparsely represent high dynamic range wavefields, they allow for reconstructing detail-rich land records without windowing or amplitude processing (i.e., gain or normalization). However, reconstruction success strongly depends on the reconstruction approach. The basis pursuit with projected spectral gradient (SPGL1) method (van den Berg and Friedlander, 2009, 2011) is a common choice for solving sparsity-promoting regularization problems because its numerical implementations are readily available and can handle complex numbers. In seismic interpolation problems, a popular data reconstruction strategy is projection onto convex sets (POCS) due to its ease of implementation and relatively straightforward parametrization (Abma and Kabir, 2006; Gao et al., 2013b). We show that yet another approach, an iterative soft thresholding (IST) (Daubechies et al., 2004; Donoho, 1995), can match or better the reconstructions obtained from SPGL1 and POCS, as demonstrated by Fourier domain SNR. Though iterative soft thresholding with wavelets, often referred to as 'wavelet shrinkage' (Donoho et al., 1995), has been successfully used in image restoration problems, we are not aware of its application to seismic data reconstruction. Our main contribution is developing a strategy

for scale- and orientation-dependent thresholding that is inspired by the link of the mixed norm minimization problem with ℓ_1 constraint to the maximum *a posteriori* estimate with a Laplacian prior (Alliney and Ruzinsky, 1994; Figueiredo et al., 2007; Tibshirani, 1996). The only parameter required for the initial threshold estimation is the noise standard deviation σ_n. We show that this parameter can be estimated directly from data, yielding threshold values that converge on a good reconstruction. As a result, IST is easy to use and implement and yields high-quality reconstructions. Furthermore, if one adjusts the stopping criterion appropriately, IST can also be used as a denoising tool.

This manuscript begins with a brief review of the basic properties of the complex wavelet transform and introduces the intuition behind correspondence of scales and orientations and Fourier frequency-wavenumber bands. Next, we discuss the compressive sensing framework and the iterative thresholding approach. We then show reconstructions for randomly subsampled wavefields with 50% missing traces which highlight the favorable qualities of soft thresholding and POCS. Finally, we summarize our findings and conclude that iterative soft thresholding is an attractive complex-wavelet domain reconstruction strategy, easy to implement, and capable of reconstructing structurally-consistent wavefields.

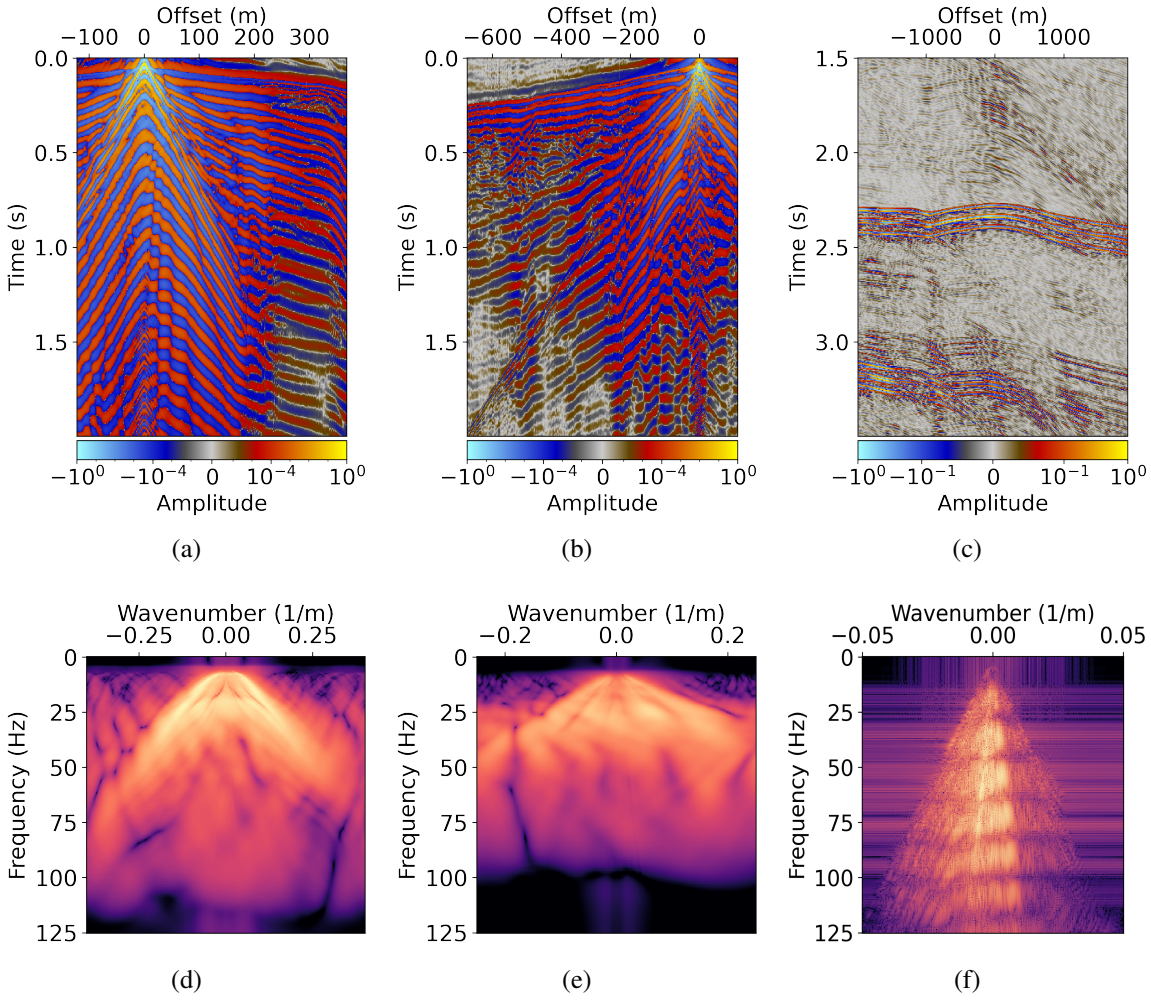


Figure 3.1 (a)-(c): Seismic records used for reconstructions experiments and (d)-(f) the corresponding $f - k$ spectra. (a) Land record with 1.25 m receiver spacing, (b) land record with 2.0 m receiver spacing, and (c) synthetic data with 10.0 m receiver spacing. All records are normalized such that their peak amplitude is unity. We use symmetric logarithmic scale to emphasize the large data dynamic range. $f - k$ spectra are displayed with logarithmic color mapping.

3.2 Complex Wavelet Transform

The discrete complex wavelet transform ($\mathbb{C}$WT) extends the real wavelet transform by introducing several desirable qualities absent from a critically sampled real wavelet transform (Selesnick et al., 2005). First, unlike real wavelets that exhibit oscillatory behavior around singularities, the magnitude of $\mathbb{C}$WT is consistently high around such features, allowing for more

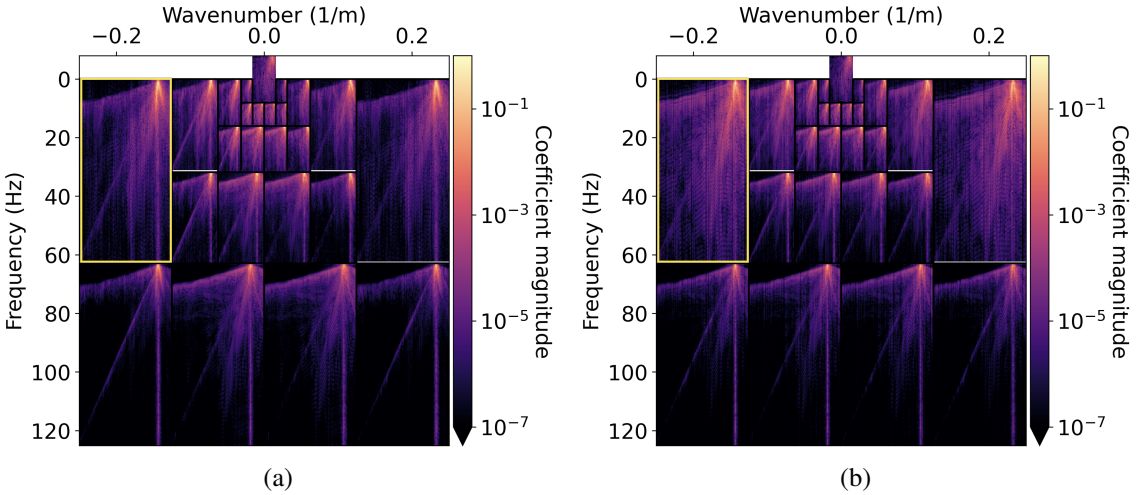
intuitive coefficient interpretation and feature identification. Second, $\mathbb{C}$WT is near-shift invariant:, i.e., small signal shift causes only a minor change in coefficient magnitude with a near-linear phase encoding of the shift amount. Thus, this property limits the number of possible coefficient patterns introduced for example by shifting the sampling matrix, yielding more predictable reconstruction results. Third, complex wavelets in two and more dimensions offer enhanced directional selectivity, avoiding the 'mixing' of directions characteristic of the real transform (e.g., its inability to distinguish between $\pm 45°$). Finally, $\mathbb{C}$WT is less sensitive to coefficient aliasing that results from transform implementation via an imperfect filter bank (filters do not completely attenuate signal's spectral content outside of filter's pass band).

While $\mathbb{C}$WT is a redundant transform, the redundancy factor does not depend on scale and is 2^m for m-D signals (Kingsbury, 2001; Selesnick et al., 2005). Similarly to the Fourier transform, when $\mathbb{C}$WT is used for analyzing real-valued signals, coefficient symmetries can be exploited, allowing for memory savings with the right transform implementation. We use $\mathbb{C}$WT implementation via dual-tree filter bank of real wavelet filters, with the two tree branches providing real and imaginary parts of complex coefficients. The filters used for $\mathbb{C}$WT must meet stringent design criteria including perfect reconstruction and jointly yielding an approximately analytic transform. Kingsbury (2001) and Selesnick (2002) provide detailed requirements.

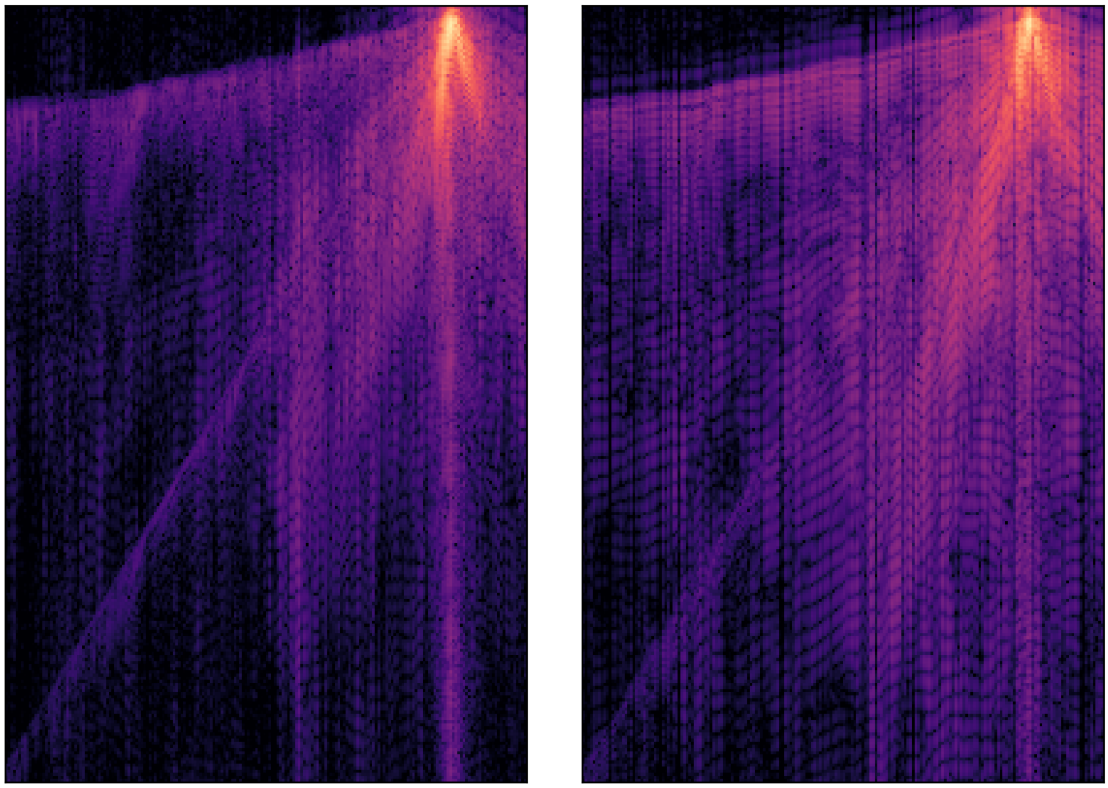
The orientations in the complex wavelet transform are related to the idealized Fourier domain support of each wavelet. The common interpretation of these orientations in 2D is related to features oriented at six specific angles: $\pm 15°$, $\pm 45°$, and $\pm 75°$. However, while such interpretation may be helpful for natural images, it gives little insight for seismic data with spatial and temporal dimensions. Thus, it is helpful to represent complex wavelet coefficients on their idealized Fourier domain support. While the coefficients at specific scale and orientation are supposed to represent signal features from specific octave band, they can also contain some frequency and wavenumber information outside of that band because wavelets are not perfectly analytical and the filters implementing the transform imperfect. Despite that, showing $\mathbb{C}$WT coefficients in the Fourier space offers valuable insight. Figure 3.2 visualizes coefficient

magnitudes for a transform of the wavefield from Figure 3.1(b) before and after random trace subsampling. The correspondence of $\mathbb{C}$WT's angles to more natural phase velocities is not simply that steeply dipping events correspond to high angles. Instead, the highest concentration of complex wavelet coefficients corresponding to a particular dipping event depends on spatial and temporal sampling. Because wavefields are frequently oversampled in time, one may want to resample seismic data such that all sampling dimensions (space and time) have usable information up to their respective Nyquist wavenumber or frequency, allowing for improved representation of specific wavefield features in the complex wavelet domain.

For a fully sampled wavefield without missing traces, one would expect that the high-energy wavelet coefficients correspond to high-energy portions of wavefields. That intuition holds true in Figure 3.2, with the complex wavelet coefficients displayed on logarithmic scale. Trace subsampling introduces artifacts, but rather than being spread randomly throughout the whole domain, these effects are fairly localized. At the finest scale, we see high-energy coefficients introduced for high-wavenumber, low-frequency quadrant (top left corner marked by a yellow box). Figure 3.2(c) and Figure 3.2(d) are a zoom-in at the scale and orientation marked by the box. Coefficients are plotted on the same color scale, showing that after random trace subsampling, missing trace magnitudes are weakened, while the magnitudes representing the remaining traces are strengthened. The strengthening is likely because gaps in the wavefield are perceived as singularities. This furthers the case for soft thresholding: we aim to rebuild coefficient magnitudes inside gaps while weakening them for the available data. A sparsity promoting algorithm that simply seeks to preserve the high-energy coefficients in the transform domain while disregarding weak coefficients as noise is likely to reinforce data discontinuities. By driving these weak-energy coefficients corresponding to gap locations to zero we may ultimately prevent the recovery of the missing information. One way to counteract that is to introduce a low-resolution initial estimate (Gorodnitsky and Rao, 1997).



(a) (b)



(c) (d)

Figure 3.2 Magnitudes of complex wavelet coefficients for wavefield in Figure 3.1(b): (a) before and (b) after random trace subsampling. (c) and (d) are zoom on the $\mathbb{C}$WT coefficient magnitudes from yellow boxes in (a) and (b). Random trace subsampling introduces stripe-like pattern mimicking missing trace geometry while the coefficients representing the remaining traces are strengthened.

3.3 Compressive Sensing

Compressive sensing is a sampling / sensing paradigm allowing for recovery of signals from incomplete measurements (Candès et al., 2006b; Candès and Tao, 2006). The key assumption is that the target signal is or can be approximated as sparse in some representation, and that it is judiciously sampled. Sparsity means that the signal's information content can be expressed concisely, for example by encoding locations and values of large wavelet coefficients in image compression problems (Taubman, 2002). The concept of 'judicious sampling' is related to the relationship between the sensing basis - 'the data domain' in the case of seismic acquisition - and the representation basis - data representation in the sparse domain. Formally, that relationship is described by coherence μ

$$\mu(\mathbf{\Phi}, \mathbf{\Psi}) = \sqrt{n} \cdot \max_{1 \leq k,j \leq n} |\langle \phi_k, \psi_j \rangle|, \tag{3.1}$$

where ϕ and ψ represent basis functions from sensing and representation basis, respectively. When the two bases have correlated elements, the coherence is large. In general, it is desirable to find the pairing of bases for which the coherence is as small as possible, i.e., the bases that do not have correlated elements. The combination of spikes in the physical domain and the Fourier basis has maximal incoherence, hence one of the reasons for its popularity. Wavelet transforms are local; when used as a sparse basis for signals acquired with random spikes, the sampling artifacts should not spread through the whole domain. It leads to an interesting conundrum for compressive sensing of seismic signals. From a practical point of view, the non-uniform sampling (i.e., either placing the receiver and/or source on the grid or skipping the sampling location altogether) is the easiest to achieve for seismic acquisition. Additional constraints can be placed on the geometry, for example to limit the gap size for the benefit of reconstruction strategies or to account for natural obstacles. For this type of sampling, Fourier representation seems like a natural choice. However, if the goal is to recover raw high dynamic range records, the Fourier basis cannot sufficiently sparsify such signals. Complex wavelets are less ideal from the coherency point of view, but provide significantly sparser wavefield representation, making them

a viable alternative to the Fourier basis.

After selecting the sensing and sparse bases, the underlying signal can be recovered by a sparsity-promoting algorithm. There are many possible choices that fall into at least two categories: greedy pursuits that approximate signals looking for locally optimal representations one coefficient at a time (Needell and Tropp, 2009; Tropp, 2004; Tropp and Gilbert, 2007) and convex relaxation methods that solve a convex optimization problem that approximates the target signal (Candès et al., 2006c; Daubechies et al., 2004; van den Berg and Friedlander, 2009, 2011). The convex optimization usually uses the ℓ_1 norm to enforce the sparsity constraint and is formulated as

$$\min \|\boldsymbol{\alpha}\|_1 \text{ s. t. } \|\mathbf{d} - \mathbf{S}\boldsymbol{\Psi}^H\boldsymbol{\alpha}\|_2^2 \leq \epsilon, \tag{3.2}$$

where $\boldsymbol{\alpha}$ is a coefficient vector in the sparse domain, $\mathbf{S}$ is the sampling matrix, $\boldsymbol{\Psi}$ is the sparsifying transform, and $\mathbf{d}$ is the recorded data. Basis pursuit and iterative soft thresholding fall into this latter family of algorithms. Projection onto convex sets (POCS) falls into the greedy pursuit family.

3.4 Iterative Soft Thresholding

Wavelet thresholding, sometimes referred to as 'shrinkage' (Chambolle et al., 1998; Donoho et al., 1995) is a well-established technique used for image restoration (Bioucas-Dias and Figueiredo, 2007; Donoho, 1995; Raj and Venkateswarlu, 2012) and aims to recover a high-quality image from degraded samples, with degradation that can be purely due to noise, or blurring, or both. Iterative soft thresholding also arises naturally as a way of solving a mixed norm minimization problem:

$$\min_{\boldsymbol{\alpha}} \mathcal{J} = \|\mathbf{d} - \mathbf{G}\boldsymbol{\alpha}\|_2^2 + \lambda\|\boldsymbol{\alpha}\|_1. \tag{3.3}$$

In the context of complex wavelet seismic data reconstruction, $\mathbf{G} = \mathbf{S}\boldsymbol{\Psi}^H$, with sampling matrix $\mathbf{S}$, and inverse complex wavelet transform $\boldsymbol{\Psi}^H$, while $\boldsymbol{\alpha}$ represents the complex wavelet coefficients of fully sampled wavefield. λ is the regularization parameter that in the IST fulfills the role of threshold. The algorithm iterations involve matrix-vector multiplications with $\mathbf{G}$ and

$\mathbf{G}^H$ followed by the shrinkage / soft threshold step (Beck and Teboulle, 2009; Chambolle et al., 1998; Figueiredo and Nowak, 2003):

$$\boldsymbol{\alpha}_{k+1} = \mathcal{T}_{\lambda t}\Big(\boldsymbol{\alpha}_k - 2t\mathbf{G}^H(\mathbf{G}\boldsymbol{\alpha}_k - \mathbf{d})\Big), \tag{3.4}$$

where t is the step size and the soft thresholding operator is defined as

$$\mathcal{T}_\lambda(\boldsymbol{\alpha})_i = (|\alpha_i| - \lambda)_+ \mathrm{sgn}(\alpha_i) = \begin{cases} \alpha_i + \lambda & \text{for } \alpha_i \leq -\lambda \\ 0 & \text{for } |\alpha_i| < \lambda \\ \alpha_i - \lambda & \text{for } \alpha_i \geq \lambda \end{cases}, \tag{3.5}$$

with the threshold λ. Note that this operator is defined for real-valued functions, however, the extension to complex numbers is straightforward. Let $\alpha_i = re^{j\theta}$. Then, the complex thresholding operator is simply $\mathcal{T}_\lambda(re^{j\theta}) = \mathcal{T}_\lambda(r)e^{j\theta}$ (Daubechies et al., 2004).

Determining thresholds appropriate for a specific problem can be a challenge. The most common approach is to find a global threshold, or in other words, a single threshold value that is applied to all coefficients. While performance of such an approach is reasonable if the noise is stationary, the results can quickly deteriorate when the noise power varies from sample to sample (Lo and Selesnick, 2006). For instance, signal-to-noise ratio for seismic data depends on offset, and the disturbance introduced by irregular sampling does not affect all complex wavelet decomposition coefficients equally (Figure 3.2(b)). Therefore, introducing scale- and orientation-dependent thresholds has merit because it accounts for different coefficient energy levels among transform scales and orientations. The question remains: how does one determine the appropriate thresholds? The answer lies in the link of the mixed norm optimization problem from equation 3.3 to the probabilistic framework (Alliney and Ruzinsky, 1994; Figueiredo et al., 2007; Tibshirani, 1996). From the Bayesian perspective, one can view equation 3.3 as a maximum *a posteriori* solution to finding $\boldsymbol{\alpha}$ from noisy observations

$$\mathbf{d} = \mathbf{G}\boldsymbol{\alpha} + \mathbf{n}, \tag{3.6}$$

where the noise $\mathbf{n}$ is assumed to be Gaussian and white with variance σ_n, and $\boldsymbol{\alpha}$ belongs to a Laplace distribution. Under these assumptions, Lo and Selesnick (2006) suggest the following

threshold definition

$$\lambda = \frac{\sqrt{2}\sigma_n^2}{\sigma_w}, \tag{3.7}$$

where σ_n is the standard deviation of the noise and σ_w is scale- and orientation-dependent standard deviation for the wavelets coefficients with a zero-mean Laplace distribution (Selesnick, 2009). Chang et al. (2000) advocate for a similar threshold but without $\sqrt{2}$ factor, providing full derivation and proving its near-optimality yielding denoising results within 5% mean squared error of the best soft thresholding benchmark. Although originally intended for natural image denoising, $\lambda = \frac{\sigma_n^2}{\sigma_w}$ is well-suited for seismic data reconstruction within iterative soft thresholding framework, provided that parameters σ_w and σ_n are selected appropriately. As suggested by Selesnick (2009), σ_w can be estimated as $\widehat{\sigma}_w = \sqrt{\max(\mathbb{E}[|\alpha|^2] - \sigma_n^2), 0}$, with α representing noisy wavelet coefficients at specific scale and orientation. The noise level σ_n can be estimated as median absolute deviation (MAD):

$$\widehat{\sigma}_n = k \cdot 1.4826 \cdot \text{median}(|\alpha_i - \tilde{\alpha}|), \tag{3.8}$$

where $\tilde{\alpha}$ is the median of wavelet coefficients at specific scale and k is the empirically selected integer scaling factor. MAD is a robust statistical estimator and a popular choice when the noise is unknown (Donoho et al., 1995). A numerical constant of 1.4826 is related to the Gaussian distribution assumed for the noise. Because the imprint of sampling in the complex wavelet domain is the smallest for the coarsest scale, we recommend using the coarsest available scale with all orientations to obtain the value of $\widehat{\sigma}_n$, as in examples shown here. The choice of k depends on the level of data subsampling and the amplitude of events surrounding the gap. We find that values of $k > 10$ are required when there are substantial gaps in high amplitude features.

3.5 Data Reconstruction Examples

To assess the performance of complex wavelet domain reconstruction, we randomly remove 50% of the traces for wavefields shown in Figure 3.1. The three wavefields are used to avoid drawing conclusions from potentially biased results. We opt against introducing gap size control

in this instance to better understand what the reconstruction limitations are. Note that the Fourier spectra of the tested wavefields indicate that they are critically sampled or suffer from minor aliasing. This is an important point because we might suffer potential information loss, particularly if significant gaps happen to fall in the near-offset region where the energy concentration is high and events steeply dipping.

Each wavefield has the same number of samples in time and space and therefore the gap location indices are also the same in each scenario. The peak amplitude is normalized to unity in all cases, but the minimum non-zero amplitude differs and is $1.48 \cdot 10^{-10}$, $4.86 \cdot 10^{-11}$, and $1.53 \cdot 10^{-8}$, respectively, clearly showing that wavefields have different dynamic ranges which impacts the reconstruction. Recall that highest energy wavelet coefficients correspond to high energy concentrations in data - however, there is also sampling noise resulting in relatively high coefficients in some places and near zero coefficients in others. The reconstructions should, ideally, start with the strong events and progressively stretch towards the low amplitude events while reducing sampling imprint on data. The bigger range of data to be reconstructed, the more iterations with high initial thresholds the reconstruction may need to restore events of interest.

The large dynamic range of the wavefields is also the reason to opt for Fourier domain rather than data domain quality assessment. Any single number describing pixel-by-pixel differences between two objects can be misleading and counter-intuitive to human observer's perception, especially when large distortions or gaps are introduced (Eskicioglu and Fisher, 1995). That assessment is further compromised when both high- and low-amplitude features of an object are important. To partially avoid these pitfalls, we compute signal-to-noise ratios (SNR) in the Fourier domain for the original Fourier coefficients and for logarithmic transformation of these coefficients. First, the features of any specific trace are spread throughout Fourier domain, avoiding bias introduced by near-zero reconstruction difference where data are available and the large but often high-frequency, high-wavenumber differences where traces were missing. Second, evaluating the logarithm of Fourier coefficient magnitudes before computing SNR helps to emphasize the coefficient structure, which is important for accurate representation of slopes in

data.

Figure 3.3 (page 54) shows reconstructions with basis pursuit, iterative soft thresholding, and projection onto convex sets in rows 1, 2, and 3, with their respective Fourier SNRs. The most visually striking difference is for the synthetic wavefield: the basis pursuit approach leaves many gaps unfilled or partially unfilled while both soft thresholding and POCS overcome most of these gaps. The evaluation of data differences between original fully-sampled wavefields and their reconstructions in Figure 3.4 (page 55) shows that POCS is the most successful in restoring amplitude, while phase is consistent and reliable in both reconstructions. Interestingly, one gap remaining in the data is partially filled, depending on amplitude, wavelength and orientation of the events surrounding it. This behavior is caused by uneven coefficient energy distribution among orientations: for a circularly symmetric object, coefficients encoding some angles are weaker than coefficients encoding other angles, leading to an unfortunate reconstruction artifacts (partially unfilled gap) when wavefield features happen to align with orientations represented by weaker energy coefficients.

Land wavefield reconstruction differences are more difficult to spot, partially due to the logarithmic scale spanning six orders of magnitude necessary to make all events visible. The wavefield in Figure 3.1(b) is missing near-offset information from the subsampled data, causing the reconstruction in that region to be relatively difficult, particularly within the first 0.3 s. The basis pursuit additionally suffers from discontinuities in the reconstructed first arrival that thresholding methods manage to avoid.

One important advantage of the soft thresholding approach over POCS is that there is no requirement to perfectly match the original data to achieve the reconstruction: since both original traces and missing traces are reconstructed at the same time, one can decide on what level of original data fitting is desirable - the same is also true for basis pursuit with denoising. Here, we show iterative soft thresholding with a minor denoising that can be depicted on data differences in Figure 3.4. Denoising operations primarily target both coherent and incoherent high wavenumber events. Specifically for POCS, the original traces are re-injected at each iteration. While

re-injection might be desirable when data are high quality, such an operation can be problematic in more noisy scenarios, resulting in reconstructing noise rather than signal inside gaps.

3.6 Conclusions

Complex wavelet domain thresholding techniques are successful in restoring land seismic wavefields with randomly subsampled traces. Owing to the localized nature of the transform, both high- and low-amplitude events can be captured in a relatively sparse representation, allowing for accurate reconstructions of these features without the need for data windowing or amplitude processing. Both iterative soft thresholding and projection onto convex sets with support-based thresholding functions yield superior results to the basis pursuit with projected spectral gradient.

The success of soft thresholding relies on the scale- and orientation-based threshold definitions. These thresholds are estimated directly from the subsampled data, with the noise parameter computed using directional coefficients in the coarsest scale and multiplied by an integer constant. The choice of constant depends on the degree of data subsampling, especially if gaps are located around strong events. We find that $k = 5$ is a good choice for data with less amplitude variation, or when gaps are not concentrated around near offsets, but $k = 40$ is needed to try and restore missing near offset information, particularly since 98.7% of the recorded shot energy is contained in $(-20, 20)$ m offset range. Soft thresholding has a natural termination point when the desired data residual is reached, making such an operation an attractive approach for denoising while reconstructing. Though POCS yields improved amplitude fidelity in the presented synthetic case, IST has consistently the highest SNR for log-compressed Fourier coefficients which indicates a better overall structural fidelity of the IST-reconstructed wavefields. Thanks to the provided recipe for finding σ_n and σ_w, our IST approach requires little user input and testing, making it fast and easy to use.

Projection onto convex sets also yields satisfactory reconstructions but needs more customization to select optimal bounds for starting and ending thresholding points as well as for the thresholds progression. For land wavefields, exponential thresholding scheme proves the best

while the synthetic example benefits from linear step progression with a degree of Gaussian smoothing in the first few iterations. POCS is a good choice for reconstructing high quality data, but because the original unaltered traces are re-injected at each POCS iteration, the denoising effect cannot be achieved, affecting the reconstruction.

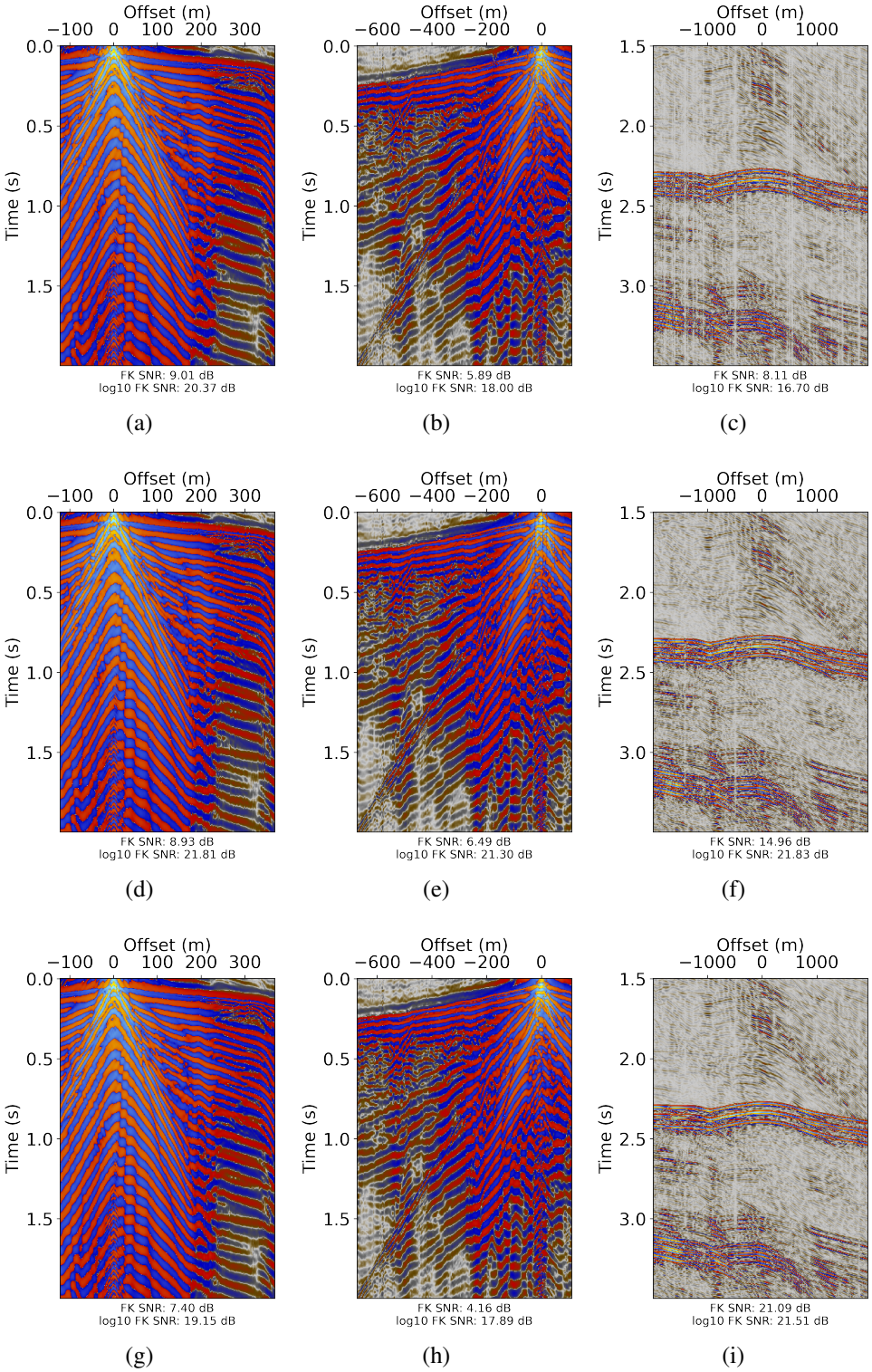


Figure 3.3 Reconstructions of the wavefields in Figure 3.1: (a)-(c) basis pursuit with SPGL1 solver, (d)-(f) IST, and (g)-(i) POCS.

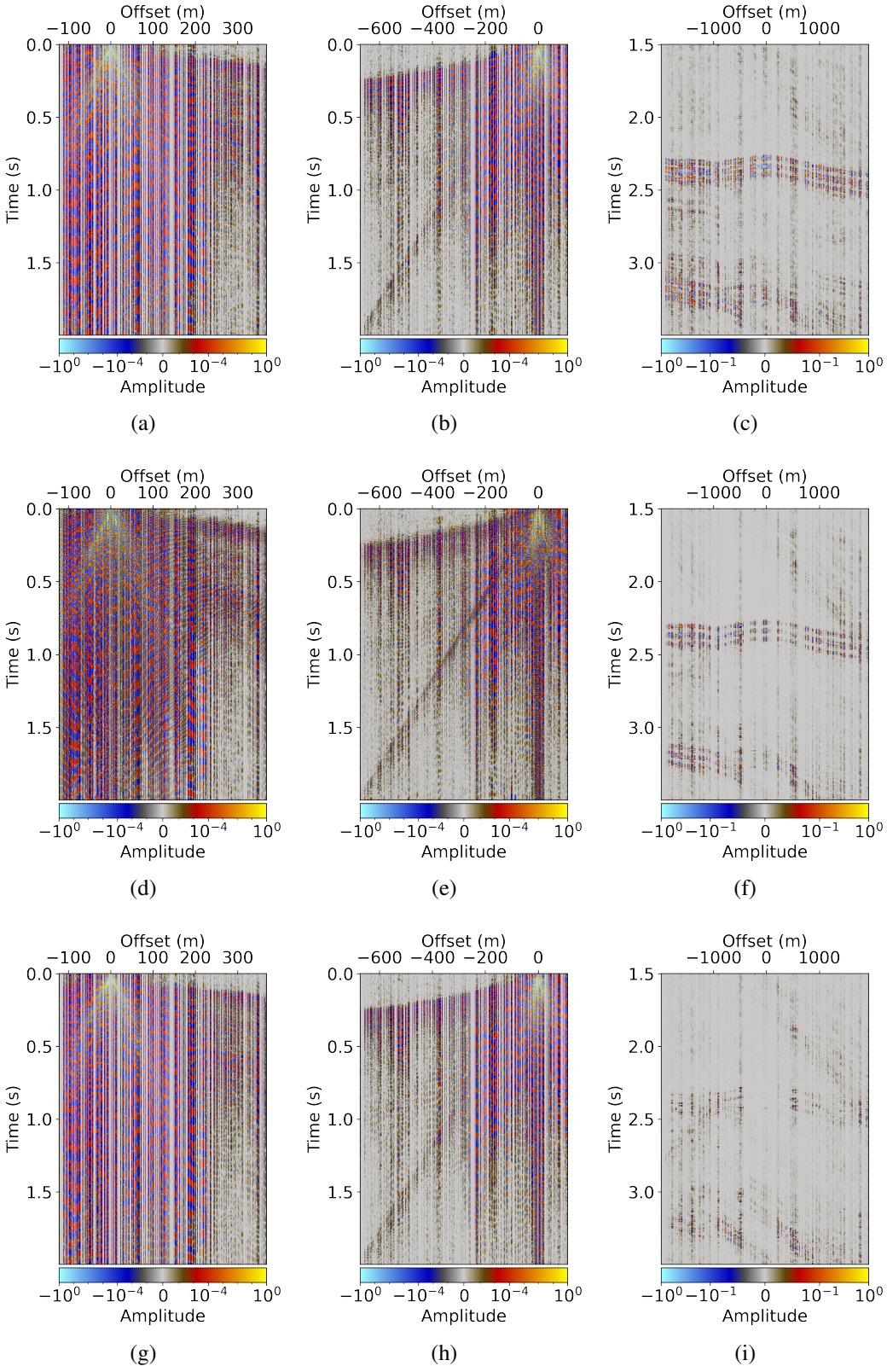


Figure 3.4 Data differences between fully sampled wavefields from Figure 3.1 and their reconstructions from Figure 3.3: (a)-(c) basis pursuit with SPGL1 solver, (d)-(f) IST, and (g)-(i) POCS.

CHAPTER 4

TOWARDS 'GOOD' SEISMIC DATA: BEYOND ALIAS INTERPOLATION AND FILTERING IN THE COMPLEX WAVELET DOMAIN

A paper submitted to *Geophysics*

Iga Pawelec[1,2], Paul Sava[2]

Seismic data can provide a wealth of information about subsurface. The process of obtaining interpretable volumes is involved and intricate, but it starts with the raw data acquired in the field. 'Good' raw data should enable the use of state-of-the-art noise suppression tools and thus allow for high-resolution subsurface imaging. The key to good data is the ability to capture wavefields without gaps and aliasing which in most situations is impossible or prohibitively expensive. As a result, data are often regularly sampled, but contain gaps in areas of restricted access and suffer from aliasing artifacts. Both of these challenges can be tackled after transforming data into a domain in which the their key characteristics can be easily identified and leveraged for guiding wavefield reconstruction. We find that a discrete complex wavelet transform is an optimal domain for representing intricate land seismic wavefields. Taking advantage of the multiscale decomposition and seismic data sparsity in the complex wavelet domain, we develop a multiscale soft thresholding algorithm (MIST) that uses low-frequency, low-wavenumber information to guide the reconstruction of higher frequencies and wavenumbers. Our approach is successful in interpolation beyond aliasing and in infilling large gaps, but its effectiveness is sampling-dependent, requiring access to nonaliased wavelet coefficients at the deepest possible decomposition level. We also demonstrate the filtering outcomes possible to achieve for nonaliased data. Exploiting the spatio-temporal and frequency-wavenumber localization capability of complex wavelet transform, we develop a velocity filter that allows to target noise of

[1]Primary researcher and author.
[2]Colorado School of Mines, Golden, Colorado, United States of America.

particular phase velocity or range of velocities that is also restricted in time and space. We demonstrate the filter by removing airblast energy from a land seismic record. Our filter preserves features underlying noise and does not affect any signals outside of the predicted airblast trajectory.

4.1 Introduction

Active seismic data are acquired to provide detailed subsurface information such as structural or stratigraphic images, depositional sequences, material and reservoir properties (e.g., wave propagation velocities, anisotropy, density, porosity, fluid saturation), or time-lapse differences in these properties. The confidence in inferences made from the acquired seismic volumes depends on the data quality and uncertainties associated with processing and inversion. But what does it mean to have 'good data'? The answer may not be obvious. An important ingredient of data quality is trace density (Alkatheeri et al., 2020; Ourabah et al., 2015; Yanchak et al., 2018). Dense acquisitions are characterized by high fold which is directly related to the random incoherent noise suppression by stacking (Schimmel and Paulssen, 1997). However, for most onshore acquisitions it is the coherent, slowly propagating waves that contaminate the desired data and degrade the migrated volumes (Manning et al., 2019; Ourabah et al., 2019; Regone et al., 2015; Stork, 2020). Examples of such noise include surface and guided waves, airblast, and scattering noise. These types of coherent noise are best addressed when recorded nonaliased.

nonaliased recordings of complex land wavefields can be difficult to achieve in practice. If one follows the traditional regular sampling theorem (Jerri, 1977; Shannon, 1948) or its m-D extension (Petersen and Middleton, 1962), the sampling interval required is dictated by the slowest apparent velocity of the observed wavefield and the highest frequency present in the slow moveout features. For traditional field land surveying, this criterion may call for regular sampling intervals of less than 2 m, which is prohibitively expensive for most acquisition campaigns. Furthermore, land acquisitions frequently have to be planned around existing infrastructure and natural obstacles, leading to gaps in the acquired seismic volumes.

A possible solution to these problems is compressive sensing (CS) acquisition that uses randomized sampling and sparse data representation in some domain to reconstruct the nonaliased wavefield as a post-acquisition step (Jiang et al., 2019, 2018; Li et al., 2012; Mosher et al., 2017). However, that solution is not always a viable option. One problem is that the compressive sensing approach cannot be applied to the already acquired regular data, as most seismic data are collected on regular or approximately regular grids. When aliasing occurs, such sampling solutions do not lend themselves to data reconstruction following compressive sensing rules because aliases masquerade as real signals, and make aliased data appear sparse. Any sizable gaps in data are also a problem regardless of sampling scheme (Trad et al., 2005). Large concentrations of missing traces introduce illumination gaps and ultimately may prevent one from achieving a desired imaging objective. A 5D interpolation (Trad, 2009) is helpful when reciprocal information about the gap interior is available (i.e., skipped shot locations infilled from receiver data when possible). However, such information is not always present or of sufficient quality, and the often high dynamic range of land seismic data introduces further complications.

Though at the first glance it may not seem like the problems of regular aliased data and significant data gaps are related, simple solutions can help to address them both. Spitz (1991) shows that nonaliased low frequency information can be used to guide the interpolation of higher frequency aliased events by designing appropriate prediction error filters (PEFs) in the F-X domain. Gülünay (2003) demonstrates that FK domain interpolation operator designed to manipulate low nonaliased frequencies can guide the unraveling of aliased data. However, a common problem with these and other Fourier domain approaches (Duijndam et al., 1999b; Gao et al., 2013b; Naghizadeh, 2012; Naghizadeh and Innanen, 2011; Zwartjes and Sacchi, 2007) is that they assume stationary signals. When the wavefield considerably varies in amplitude and frequency as a function of time and space, methods assuming signal stationarity struggle. To overcome this problem, Liu and Fomel (2011) introduce adaptive PEFs, obtaining the nonstationary coefficients by solving global regularized least squares problem. Guitton and Claerbout (2010) develop a pyramid transform to estimate a single PEF that accounts for

non-stationarity. Another solution is to take advantage of local transforms, e.g., the approach described by Naghizadeh and Sacchi (2010a) that uses nonaliased curvelet scales for interpolation. The curvelet transform (Candès et al., 2006a) is an oriented and localized generalization of the Fourier transform. While it provides sparse representations of wavefields (Candès and Demanet, 2005), it is a highly redundant transformation, with a scale-dependent redundancy level. A less redundant alternative to curvelets is the discrete complex wavelet transform which we elect to use for data reconstruction and filtering.

The discrete complex wavelet transform ($\mathbb{C}$WT) was developed to overcome certain shortcomings of its real-valued counterpart that made wavelets sub-optimal for image processing. Unlike the real wavelet transform, $\mathbb{C}$WT offers near shift invariance, the ability to distinguish between dip directions, and an intuitive interpretation for detecting signal singularities (Selesnick et al., 2005). While the transform implementation necessitates redundancy, its degree depends only on the dimensionality of decomposed objects and is 2^m for m-dimensional signals, meaning that the number of $\mathbb{C}$WT coefficients representing data is $N \times 2^m$. Furthermore, the transform has an efficient implementation via filter bank with the computational complexity of $\mathcal{O}(N)$ which is better than the $\mathcal{O}(N \log N)$ complexity of the Fast Fourier transform. Combined, these features make the $\mathbb{C}$WT an attractive domain for performing data interpolation.

$\mathbb{C}$WT naturally decomposes signals into different octave bands from the Fourier domain, therefore it can easily adapt the philosophy of using nonaliased (low-frequency, low-wavenumber) scales to guide the data reconstruction in large gaps and for regularly missing aliased traces. We describe how to adapt an iterative soft thresholding algorithm (Daubechies et al., 2004) for a multiscale reconstruction that follows $\mathbb{C}$WT decomposition levels and provide examples that illustrate its effectiveness. In particular, we show that multiscale iterative soft thresholding (MIST) succeeds in interpolation beyond aliasing and in infilling large gaps, yielding reconstructed wavefields that can be subsequently processed without suffering the consequences of inadequate sampling.

We also demonstrate how the $\mathbb{C}$WT can be used to accomplish the next task in the quest for the good quality data, namely, wavefield denoising. Nonaliased wavefields make it easier to separate signal and noise by their characteristics in the transform domain. $\mathbb{C}$WT is localized in time, space, frequency, and wavenumber, enabling straightforward definition of velocity filters. Different types of noise such as surface waves, airblast, guided waves, or first arrivals that have near-linear moveout or range of moveouts can be suppressed with a well-defined velocity filter. Such filters are commonly applied in the Fourier domain but because of the stationarity assumption and the lack of time-space localization, an FK filter targeting specific slope erases all instances of events with similar slopes, thus risking the loss of valuable diffractions or reflection tails. We can avoid this shortcoming by defining a velocity filter in the $\mathbb{C}$WT domain. We explain how to determine which scales and orientations of the $\mathbb{C}$WT should be used for filtering and how to incorporate spatial moveout information for best filtering outcomes.

4.2 Complex Wavelet Reconstruction and Filtering

Geophysicists frequently deal with highly non-stationary signals. It was the need to study the minute changes in amplitude, phase, frequency and shape of seismic reflection signals as a function of propagation time that motivated Morlet et al. (1982b) to use wavelets as an analysis tool to characterize wave scattering regimes for different wavelengths and media types. Although non-stationary signals can be analyzed with the short time Fourier transform (STFT), wavelets are an attractive alternative because they provide a multiresolution decomposition of signals. In STFT, an analysis window is selected at the start and kept the same for all times and frequencies. The time-frequency resolution, which is the ability to distinguish between two discrete spikes in time domain and two pure frequency sinusoids in the frequency domain, is therefore fixed and bounded by the Heisenberg uncertainty (Gabor, 1946):

$$\Delta t \Delta f \geq \frac{1}{2}, \tag{4.1}$$

with Δt and Δf indicating time and frequency resolution, respectively. Keeping the resolution fixed over the entire time-frequency plane imposes limitations on how well specific signal features

can be localized in time or in frequency. The idea behind the wavelet transforms is to vary Δt and Δf such that the time resolution becomes arbitrarily good at high frequencies while the frequency resolution becomes arbitrarily good at low frequencies (Rioul and Vetterli, 1991), thus obtaining the multiresolution representations of signals. This is achieved by keeping the relative bandwidths constant, most commonly using 2:1 scaling that yields octave bandwidths (Kingsbury, 1999). Naturally, wavelet decompositions are still subject to the uncertainty principle from equation 4.1, but by selecting long analysis windows for low frequency and short windows for high frequency content, our ability to distinguish time-frequency signal characteristics is significantly improved.

The multiresolution signal analysis is not restricted to 1D cases. A discrete 2D wavelet transform is used in image processing problems such as denoising, restoration, compression, and segmentation. However, using a critically sampled real wavelet transform in image processing can be challenging because of its strong shift dependence and the lack of sufficient directional selectivity. Critical sampling refers to the transform implementation via filter bank with analysis (decomposition) filters whose number equals the downsampling factor. Shift variance implies that the energy concentration of wavelet coefficients at each decomposition scale depends on time or space shifts of the input signal, which can lead to erratic and inconsistent processing outcomes. The limited directional sensitivity means that it is not possible to distinguish the direction of dipping features, for instance, in 2D it is not possible to tell the difference between $\pm 45^{\circ}$. These shortcomings led to the development of discrete complex wavelet transform ($\mathbb{C}$WT) that is near shift invariant and capable of distinguishing dip directions (Kingsbury, 1999, 2001; Selesnick et al., 2005).

4.2.1 Complex Wavelet Transform

By introducing approximately analytic complex wavelets (whose support is finite, preventing them from being exactly analytic), the $\mathbb{C}$WT is able to mimic desirable properties of the Fourier transform, providing complex coefficients whose phases depend on signal shifts almost linearly while magnitudes are largely shift invariant (Kingsbury, 2001; Selesnick et al., 2005). The

complex wavelet is defined as

$$\psi_C(t) = \psi_r(t) + \mathrm{i}\psi_i(t), \tag{4.2}$$

with real-valued functions $\psi_r(t)$ and $\psi_i(t)$ that are odd and even, respectively. Wavelets act as highpass filters. To achieve a lowpass at the lowest decomposition level, a complex scaling function is also introduced:

$$\phi_C(t) = \phi_r(t) + \mathrm{i}\phi_i(t), \tag{4.3}$$

with similar properties to the complex wavelet. The transform is implemented via a dual-tree filter bank, as described by Kingsbury (2001) and Selesnick et al. (2005). The m-D wavelet can be obtained by taking a tensor product of wavelets along all dimensions. Directional selectivity is achieved because $\mathbb{C}$WT uses analytic wavelets, allowing to distinguish between positive and negative wavenumbers. If we use a seismic convention for Fourier spectrum representation of real-valued signals, the frequency axis contains only positive frequencies, while wavenumbers can be either positive or negative. A specific scale and orientation in $\mathbb{C}$WT corresponds to an octave band representing frequency and either positive or negative wavenumbes. Such support allows for six distinct orientations in 2D and 28 orientations in 3D transform.

Due to the favorable properties described above, complex wavelets found many interesting applications such as volume registration where one seeks to detect and compensate for motion between images or volumes (Chen and Kingsbury, 2012), video denoising (Selesnick and Li, 2003) and a wide array of image processing problems, where $\mathbb{C}$WT outperforms their real-valued transform alternative (Chitchian et al., 2009; Kingsbury, 1999; Miller and Kingsbury, 2008; Raj and Venkateswarlu, 2012).

Although we often display geophysical data as if they are images, we have to remember that they represent physical quantities as a function of time, or space, or both, rather than the scaled light intensity or color coding for pixels. As a consequence, the range of possible coefficient values for transformed geophysical data differs from that of image data. Moreover, data sampling rate and the number of samples available in each sampling direction impose inherent limitations

on what processing outcomes can be achieved with the aid of $\mathbb{C}$WT because they control the number of possible decomposition levels and central frequencies and wavenumbers corresponding to $\mathbb{C}$WT scales and orientations. Understanding how geophysical data, and seismic data in particular, are represented in the complex wavelet domain is the key to optimizing data reconstruction and to designing effective filters.

4.2.2 Rotational Variance

One potential problem with $\mathbb{C}$WT is that certain orientations correspond to central frequencies and wavenumbers that are further from the origin than others. The orientations formed by combining lowpass and highpass filters have central frequencies that are approximately $\sqrt{1.8}$ closer to the origin than orientations formed by combining highpass filters (Kingsbury, 2006). This has interesting implications for data reconstruction, as illustrated next.

Consider a circularly symmetric object such as in Figure 4.1. For such an image, the transform orientations should represent angles and because a circle has the same amount of all angles, the coefficient energy distribution should be uniform among all orientations. However, due to center frequency shift, that is not the case. To understand how this affects reconstruction and assess the effect of rotational variance on the reconstruction quality, we conduct a numerical experiment. First, we remove a portion of the data in a shape of circular sector with a 6° angle. Starting with a gap centered at 0° azimuth, we then run a sparsity-promoting reconstruction. We repeat that process for gaps oriented along all azimuths with a 1° increment. Each reconstruction is quantified by the normalized root mean squared error (NRMSE) and plotted against gap azimuth. The higher the NRMSE, the less favorable a particular gap orientation is from the reconstruction point of view and vice-versa. Example reconstructions for $0^\circ, 45^\circ$ and 90° are shown in the top row of Figure 4.1, and the middle row shows the difference (magnified by 10) between fully sampled object and its reconstruction. Note that reconstruction error is the biggest for 45° and related to amplitude, while phase is reconstructed accurately. The bottom plot in Figure 4.1 highlights the good orientations with green stars and poor orientations with red

crosses. As expected, $45°, 135°, 225°$ and $315°$ are poor because they represent gaps aligned with orientations for which two high-pass filters were used to achieve the decomposition. Interestingly, when the gaps are nearly aligned with the sampling axes (in this case, x or y), reconstructions quickly shift from good to bad or the other way around. We are unsure about the cause of such behavior but suspect that it might be related to representing a circular sector with square pixels with limited number of samples.

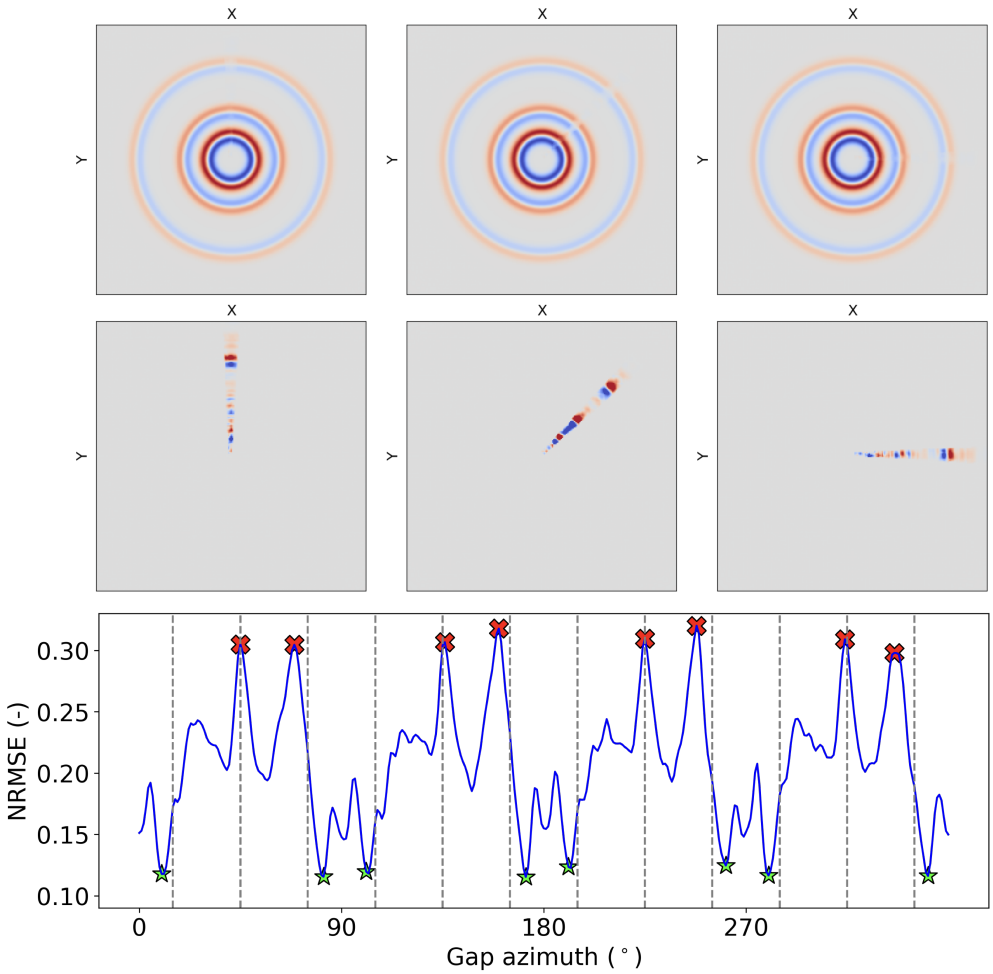


Figure 4.1 Gap orientation sensitivity test. We remove a piece of an arch from the circularly symmetric object, run complex wavelet reconstruction, and repeat the test at multiple angular orientations. The top row shows reconstructions for $0°$, $45°$, and $90°$, respectively, and the middle row shows difference between original and reconstructed object magnified 10x. Each reconstruction is quantified by NRMSE and plotted against gap azimuth (bottom). Dashed lines show primary angular sensitivity directions for a 2D $\mathbb{C}$WT . The most favorable reconstructions are indicated by green stars while the least favorable are marked with red crosses.

4.2.3 Multiscale $\mathbb{C}$WT Reconstruction

As a consequence of rotational variance and unequal coefficient energy distribution among orientations, we can expect certain wavefield features to be more challenging to reconstruct than others, simply because they are represented by weaker coefficients than other features of a similar amplitude but different slope in the data domain. One way to address that problem in the reconstruction algorithm is to minimize the effect of unequal coefficient distribution among scales and orientations by introducing scale- and orientation-dependent thresholds within an iterative soft thresholding algorithm. In Chapter 3, we describe the threshold estimation process that takes advantage of the available data to obtain parameters σ_n and $\sigma_w(s,\gamma)$ that are then used to estimate the thresholds as $\lambda_{s,\gamma} = \frac{\sigma_n^2}{\sigma_w(s,\gamma)}$. We refer to that chapter for further details.

Let $\mathbf{d}$ be the acquired data, $\mathbf{S}$ the sampling operator projecting the acquired data on the desired regular fine grid, $\mathbf{\Psi}$ the forward $\mathbb{C}$WT, and $\boldsymbol{\alpha}$ the complex wavelet domain representation of the fully sampled wavefield. The data recovery problem can then be formulated as a mixed norm optimization problem:

$$\min_{\boldsymbol{\alpha}} \mathcal{J} = \|\mathbf{d} - \mathbf{\Psi}^H \mathbf{S}^H \boldsymbol{\alpha}\|_2^2 + \lambda \|\boldsymbol{\alpha}\|_1, \tag{4.4}$$

with the regularization parameter λ. The solution can be found by iterative soft thresholding algorithm (Beck and Teboulle, 2009; Chambolle et al., 1998; Figueiredo and Nowak, 2003) whose general step is

$$\boldsymbol{\alpha}_{k+1} = \mathcal{T}_{\lambda t}\Big(\boldsymbol{\alpha}_k - 2t\mathbf{S}\mathbf{\Psi}(\mathbf{\Psi}^H\mathbf{S}^H\boldsymbol{\alpha}_k - \mathbf{d})\Big), \tag{4.5}$$

where t denotes the step size. We define the complex thresholding operator for scale- and orientation- dependent thresholds $\lambda_{s,\gamma}$. Let $\alpha_k = r_k e^{j\theta_k}$. Then

$$\mathcal{T}_{\lambda_{s,\gamma}}(\alpha_k) = \mathcal{T}_{\lambda_{s,\gamma}}(r_k)e^{j\theta_k} = \begin{cases} (r_k + \lambda_{s,\gamma})e^{j\theta_k} & \text{for } r_k \geq \lambda_{s,\gamma} \\ 0 & \text{for } r_k < \lambda_{s,\gamma} \end{cases}, \tag{4.6}$$

Rather than threshold all scales simultaneously, we modify the algorithm by introducing a multiscale approach. A pseudo code summarizing our strategy is described in Algorithm 1. We start by estimating the initial thresholds $\lambda_{s,\gamma}$ from data projected onto the target grid, $\mathbf{Sd}$. The

initial complex wavelet domain approximation of the fully sampled wavefield is obtained as $\boldsymbol{\alpha}_0 = \boldsymbol{\Psi}\mathbf{P}\mathbf{S}\mathbf{d}$, where $\mathbf{P}$ denotes a smoothing operator along the spatial dimensions. The purpose of this low resolution initial guess is to mitigate potential influence of data gaps at the deepest complex wavelet decomposition level. Starting with these deep level coefficients, we apply thresholding at the active scale and set all of the higher frequency, higher wavenumber scales to 0. When no further improvement can be achieved from a given scale, we move to a finer level and recompute the thresholds from an object that has original data traces where available and the lower scale approximation elsewhere. By doing so, we minimize the impact that the gaps presence has on the distribution and energy level of complex wavelet coefficients. The final reconstruction is achieved either when we reach the desired relative data residual or the maximum number of iterations.

A threshold re-computation step is important for reconstructing regularly missing traces. When the $\mathbb{C}$WT is computed, each data decomposition level means a data subsampling by a factor of 2 along all axes. For example, the level 1 decomposition coefficients are computed from data downsampled by a factor of 2, level 2 means downsampling by a factor of 4, level 3 by a factor of 8, etc. If data have enough spatial and temporal samples to allow for complex wavelet decomposition with at least one scale not suffering from aliasing effects, we can use it for beyond aliasing data reconstructions. However, for scales corrupted by aliasing effects, the original thresholds are sub-optimal and likely to reinforce a form of periodicity inherent in sampling, ultimately leading to poor reconstructions. In such situations, one may try to apply a constant velocity normal moveout (NMO) correction to data prior attempting the reconstruction so as to lessen the aliasing impact by making dipping events less steep (Yu et al., 2017).

A similar reasoning can be applied when one is concerned with reconstructing data within large gaps. The deepest decomposition level is the least affected by the gap presence, allowing us to gradually rebuild the missing information. When the gap is large enough to impact wavelet coefficients at all scales, the gap effect on the coarsest scale can be lessened by smoothing data along the spatial dimension. The degree of smoothing should be chosen based on the gap size and

maximum decomposition level so as to eliminate wavelet coefficient discontinuity that gap introduces.

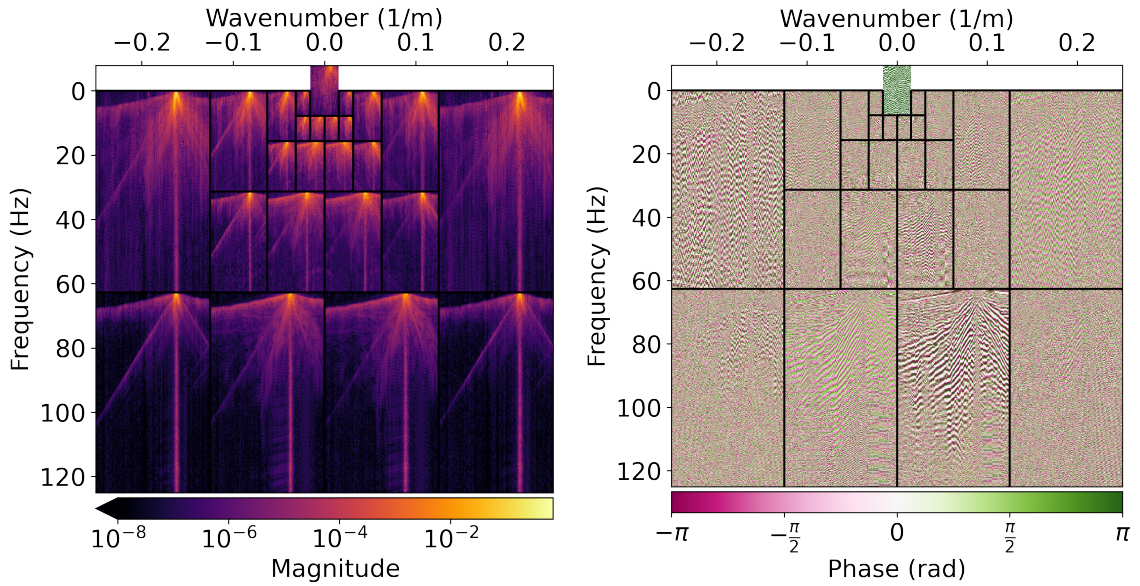
Algorithm 1 MIST

Inputs:
$\mathbf{d}, \mathbf{S}, \sigma_n, \text{maxlevel}, \text{maxit}, \epsilon, tol$
Initialize:
$\lambda_{s,\gamma} \leftarrow f(\mathbf{Sd}, \sigma_n)$ # Thresholds from available data
$\boldsymbol{\alpha}_0 \leftarrow \boldsymbol{\Psi}\mathbf{PSd}$ # Low resolution estimate
$s_k \leftarrow \text{maxlevel}$ # Maximum CWT decomposition level
$rez_0 \leftarrow \|\mathbf{d}\|_2^2$ # Initial data residual
$k \leftarrow 0$
while $\frac{rez_k}{rez_0} > \epsilon$ **and** $k < \text{maxit}$ **do**
  if $s \geq s_k$ **then**
    $\boldsymbol{\alpha}_{k+1}(s,\gamma) \leftarrow \mathcal{T}_{\lambda_{s,\gamma}}\Big(\boldsymbol{\alpha}_k - \mathbf{S}\boldsymbol{\Psi}(\boldsymbol{\Psi}^H\mathbf{S}^H\boldsymbol{\alpha}_k - \mathbf{d})\Big)$
  else
    $\boldsymbol{\alpha}_{k+1}(s,\gamma) \leftarrow 0$
  end if
  $rez_k \leftarrow \|\mathbf{d} - \boldsymbol{\Psi}^H\mathbf{S}^H\boldsymbol{\alpha}_k\|_2^2$
  if $\frac{rez_k - rez_{k-1}}{rez_k} \leq \text{tol}$ **then**
    continue
  else
    $s_k \leftarrow s_k - 1$
    $\lambda_{s,\gamma} \leftarrow f(\mathbf{Sd}, \sigma_n, \boldsymbol{\alpha_k})$ # Use current data approximation inside gaps and original data where available to get new thresholds
  end if
  $k \leftarrow k + 1$
end while
Output:
$\mathbf{m} \leftarrow \boldsymbol{\Phi}^H\boldsymbol{\alpha_k}$

4.2.4 $\mathbb{C}$WT Velocity Filter Design

$\mathbb{C}$WT could also be used for data filtering, which is not a new idea. Yu and Whitcombe (2008) and Yu et al. (2017) discuss the use of $\mathbb{C}$WT as a dip filter, pointing out that the success of filtering depends on the ability to separate the signal and noise in the transform domain and suggesting the use of constant velocity NMO correction in order to improve that separation. The filter is then defined to target specific scales and orientations in the transform domain and

removes the undesirable energy by setting all targeted coefficients to 0. However, knowing which scales and orientations to target can be difficult and non-intuitive, particularly when the commonly assumed image processing jargon uses angles to describe orientations.

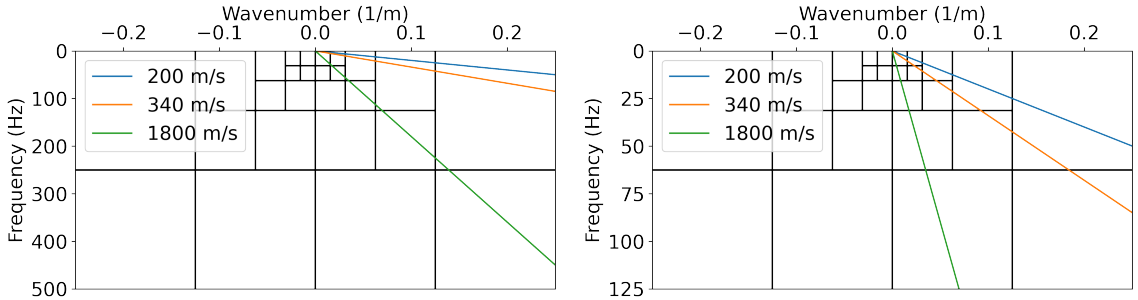


(a) (b)

Figure 4.2 (a) Magnitude and (b) phase of complex wavelet coefficients for a land seismic record sampled at 2 m and 4 ms. The largest coefficients correspond to the high energy events and to the features supported by given scale and orientation (e.g., that fast first arrivals have strong coefficients at 62.5 - 125 Hz, -0.1 - 0.1 1/m, which follows the geophysical intuition). The phase spectrum reveals that high frequency, high wavenumber features do not form many coherent patterns, indicating that these $\mathbb{C}$WT bands primarily support noise.

To help with geophysical intuition behind complex wavelet scales and orientations, consider displaying wavelet coefficients on their idealized Fourier support. As explained earlier, scales and orientations represent Fourier octave bands whose central frequencies depend on data spatial and temporal sampling. Figure 4.2 shows the type of display we are advocating for. The horizontal axis represents wavenumbers while the vertical axis represents frequencies. Solid black lines mark the boundaries between scales and orientations. By identifying the box location in the $f - k$ coordinates, we are able to assign frequencies and wavenumbers to a particular scale and orientation.

Assigning the $f-k$ bands to $\mathbb{C}$WT scales and orientations means that we can also map phase velocities. Figure 4.3 shows how such mapping can be done. Note that $x-t$ sampling defines spatial and temporal Nyquist, directly affecting whether specific phase velocities are mapped to the same or different $\mathbb{C}$WT orientation. Both panels are generated assuming the spatial sampling of 2 m, but temporal sampling is selected as 1 ms in Figure 4.3(a) and 4 ms in Figure 4.3(b). For a 1 ms case, the velocity of 1800 m/s falls into the orientation that is typically associated with $45°$, but for a 4 ms sampling it is assigned to $15°$ orientation. Therefore, trying to associate seismic events with $\mathbb{C}$WT orientations based purely on the angular interpretation of orientations may be misleading. Plots like the ones shown in Figure 4.3 are easy to generate for arbitrary samplings, decomposition levels, and velocities and are a valuable tool to aid in selecting scales and orientations for filtering or in adjusting data sampling that would achieve orientation separation between specific events.



(a) (b)

Figure 4.3 Mapping of different phase velocities to $\mathbb{C}$WT orientations assuming spatial sampling of 2 m and temporal sampling of (a) 1 ms and (b) 4 ms. Note that changing the temporal sampling alters the distribution of phase velocities among $\mathbb{C}$WT scales and orientations which has implications for noise suppression and signal separation.

Before we move on to defining a complex wavelet domain filter, we need to discuss one other aspect of $\mathbb{C}$WT : the spatio-temporal localization. Coefficient magnitudes and phases in Figure 4.2 form patterns that distinctly resemble seismic data. This is because $\mathbb{C}$WT is also localized in time and space. To find $t-x$ coordinates corresponding to a particular coefficient, one only needs to know the original data coordinates and the decomposition level. Spatial and

temporal sampling at scale s is $d\mathbf{x}_s = d\mathbf{x}_0/2^s$ and $dt_s = dt_0/2^s$, with $d\mathbf{x}_0$ and dt_0 representing the data sampling. Therefore, to map a $t - x$ slope between decomposition levels, one only needs to appropriately scale the coordinates.

3D seismic data $\mathbf{d}(x, y, t)$ can be represented by their 3D $\mathbb{C}$WT :

$$\boldsymbol{\alpha}(\mathbf{x}, \mathbf{y}, \mathbf{t}, \boldsymbol{\gamma}, \mathbf{s}) = \boldsymbol{\Phi}\mathbf{d}(\mathbf{x}, \mathbf{y}, \mathbf{t}), \tag{4.7}$$

with orientation γ and scale s. One can define a binary filter in the transform domain as (Yu et al., 2017):

$$\mathbf{Q}(\mathbf{x}, \mathbf{y}, \mathbf{t}, \boldsymbol{\gamma}, \mathbf{s}) = \begin{cases} 1, & \text{if Q} \in \text{selected } (\mathbf{x}', \mathbf{y}', \mathbf{t}', \boldsymbol{\gamma}', \mathbf{s}') \\ 0, & \text{otherwise} \end{cases} \tag{4.8}$$

Therefore, filtering can be viewed as an application of binary masks to complex wavelet coefficients. Yu et al. (2017) suggest to define filter coefficients over scales and orientations only. However, such an approach disregards spatio-temporal localization of complex wavelets, hindering filtering outcomes for events that are not sufficiently distinguished from useful signals in scales and orientations only. We propose to define a filtering mask that is time- and space-dependent for each scale and orientation:

$$\mathbf{Q}(\boldsymbol{\gamma}, \mathbf{s}) = f(\mathbf{x}, \mathbf{y}, \mathbf{t}), \tag{4.9}$$

with the mask values that can take any values from $[0, 1]$ interval. This allows for introducing moveout-based filtering masks that have a more gradual transition between pass and reject bands, mitigating edge effects. We illustrate this idea in the following section.

4.3 Examples

In this section, we demonstrate the practical uses for our multiscale reconstruction and complex wavelet domain velocity filtering. We start with interpolation beyond aliasing example showcasing wavefield recovery from regularly missing traces. Next, we show how the multiscale approach handles infilling large gaps and what can be done to improve the reconstruction outcome. Finally, we present the $\mathbb{C}$WT velocity filter design to suppress the airblast from a land seismic record.

4.3.1 Regular Data Interpolation Beyond Aliasing

Regular coarse sampling that results in spatially aliased data is a common problem. Figure 4.4(a) shows a fully sampled synthetic wavefield with a wide range of data amplitudes and event orientations. We use symmetric logarithmic color normalization from *matplotlib* (Hunter, 2007) for wavefield display to better highlight weaker signals. To test the capability of MIST for reconstructing regularly missing aliased data, we only keep every fourth trace from Figure 4.4(a) leading to coarsely sampled wavefield in Figure 4.4(b). FK spectra for both wavefields are shown in Figure 4.4(c) and Figure 4.4(d).

To evaluate the performance of MIST, we compare it with a sparsity promoting basis pursuit. The goal of this comparison is to judge whether the sparsity assumption alone would be sufficient to recover the missing data. Figure 4.5 summarizes the reconstruction outcomes for basis pursuit and MIST. Since the initial threshold estimation plays an important role for MIST, we show a reconstruction for thresholds obtained from the available coarsely sampled data and another reconstruction where the initial thresholds are estimated based on the wavefield recovered from the previous MIST reconstruction.

Figure 4.5(a) shows that sparsity alone is not sufficient to infill gaps in regularly subsampled data. The wavefield reconstructed following the basis pursuit suffers from sampling artifacts and has severe problems in recovering accurate amplitudes. MIST does markedly better, maintaining structural consistency in most places. Data differences obtained by subtracting reconstructed data from a fully sampled wavefield in Figure 4.4(a) reveal that the reconstruction error is the most significant for steeply dipping events as well as some near-horizontal features and becomes smaller if we re-run the reconstruction with thresholds recomputed from a previous reconstruction. Note that the spectrum from Figure 4.5(h) still has a remnant sampling imprint that dominates the basis pursuit reconstruction. However, that imprint is much smaller than for the original data, allowing the re-computed thresholds to handle these spectral artifacts.

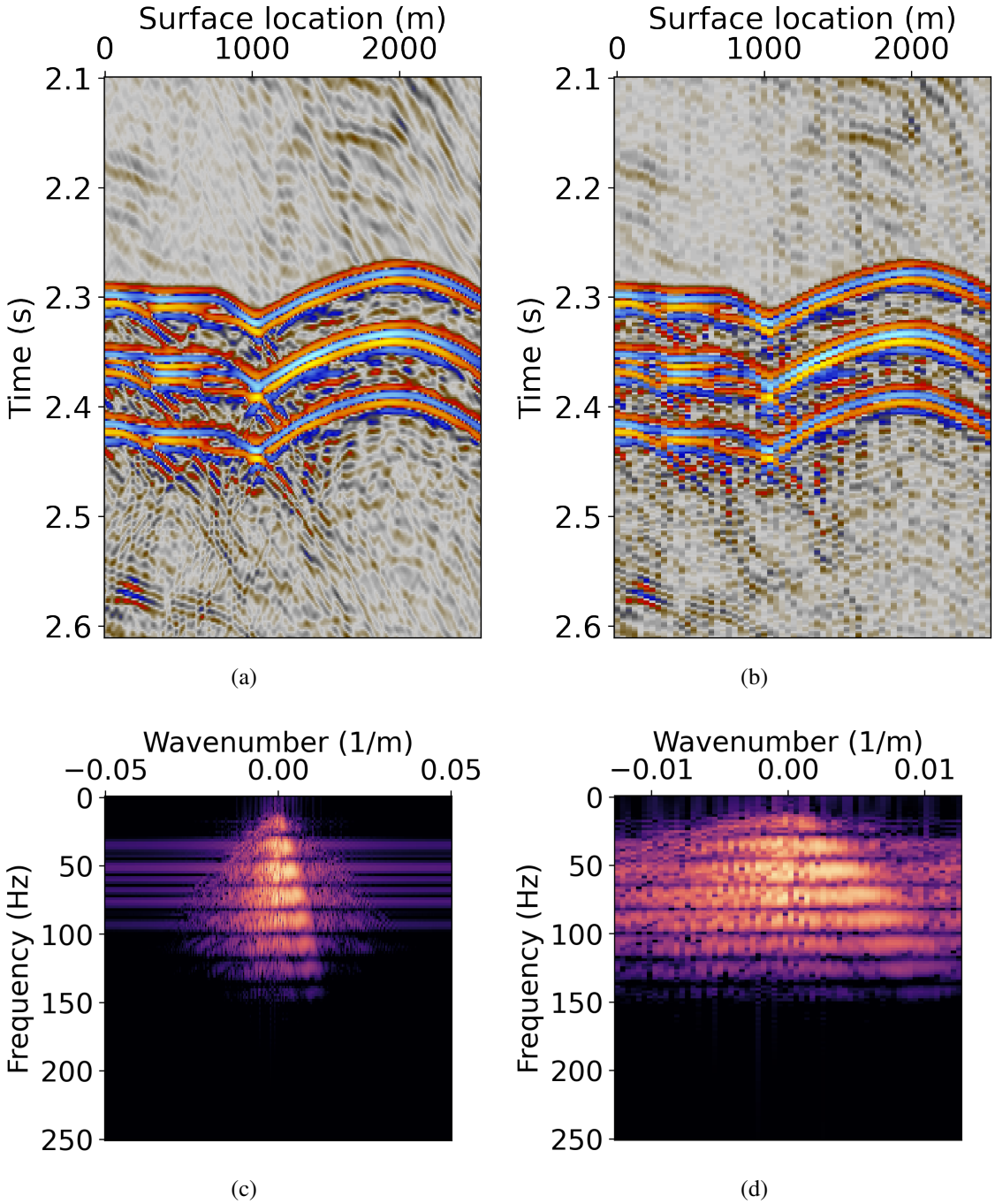


Figure 4.4 (a) Fully sampled wavefield and (b) wavefield after spatial subsampling by a factor of 4. (c) and (d) are the corresponding FK spectra.

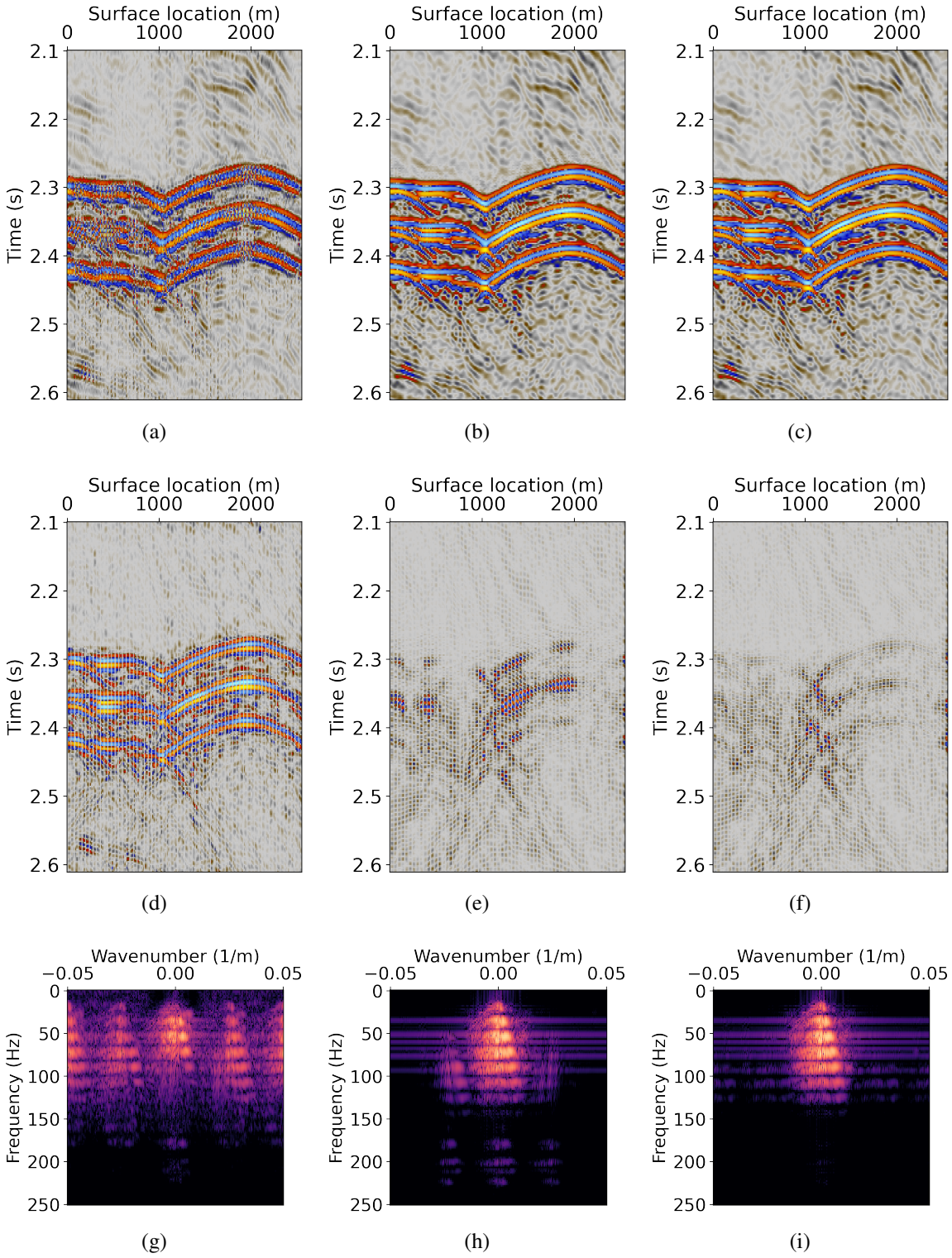


Figure 4.5 Reconstructions of aliased wavefield from Figure 4.4(b) on a fine grid. (a) Reconstruction by basis pursuit, and (b) and (c) are obtained with MIST using subsampled data or reconstructed data as an input. (d)-(e) show the corresponding data differences, while (f)-(h) show the FK spectra.

4.3.2 Infilling Large Gaps

In land acquisition, we frequently face access restrictions preventing the placement of sources and/or receivers at desired locations. To simulate such a scenario and evaluate how data reconstruction handles gaps, we introduce a 16-traces gap (marked with a black box in Figure 4.6(a)). Its size in the physical domain is 160 m, and the dominant wavelength corresponding to the strongest event is approximately 450 m.

Similarly to the previous experiment, we compare MIST reconstruction against the basis pursuit solution. Figure 4.6(b) and Figure 4.6(c) show the reconstructions in the context of the entire wavefield, while Figure 4.6(e) and Figure 4.6(f) zoom-in on details inside the gap. Basis pursuit reconstruction struggles to recover reflection amplitudes and moveouts, with fidelity getting progressively worse going from gap edges to its middle. The coherent but weak energy before reflections is barely represented. In contrast, MIST reconstruction recovers reflections with accurate moveout but amplitudes are slightly weaker than those of the original data. Weak energy before and after reflections is also reconstructed, but it is the moveouts of those events that are of particular interest. Comparing the zoomed-in reconstruction from Figure 4.6(f) to the reference data in Figure 4.6(d), we note that the reconstruction created events that do not exist in reference data (for example, a near-infinite velocity feature around 2.47 s). This happens because interpolation is incapable of creating new information - we can only use what we have in clever ways to fill in the missing data. By removing a portion of data, we risk permanent information loss. Data surrounding the gap do not provide sufficient information about some events - especially those that are not continuous through the gap. Therefore, one has to be careful about spurious events that might be created through interpolation.

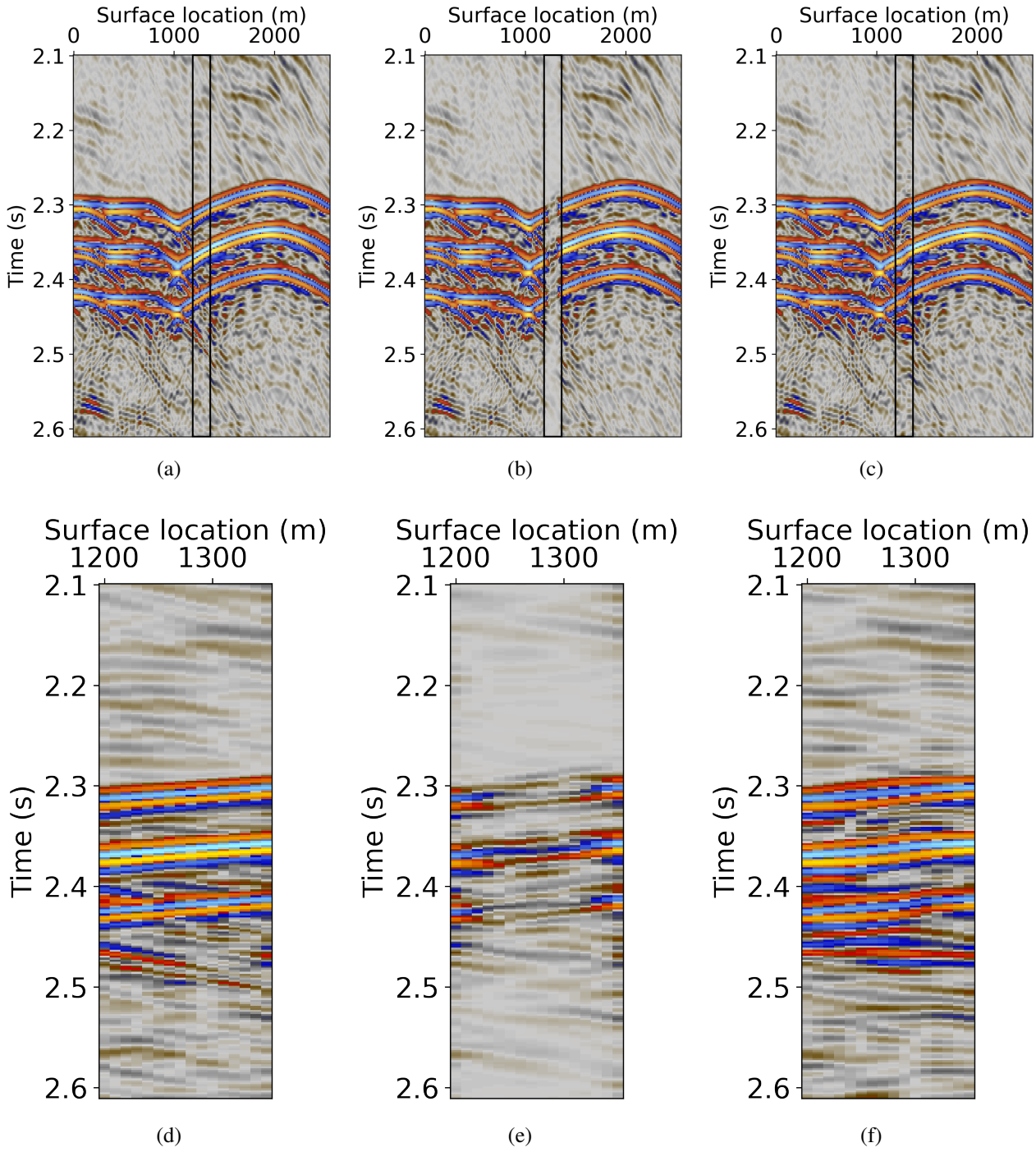


Figure 4.6 (a) Reference data with a 160m (16 traces) gap marked with the black box and the reconstructed wavefields using (b) basis pursuit and (c) MIST. (d)-(f) show the gap interior. MIST has better amplitude accuracy, but is sometimes unable to accurately reproduce event moveouts, for example the diffraction tail around 2.47 s and 1200 m visible in (d).

4.3.3 Airblast Removal

The airblast or air wave is frequently observed on land seismic records. Traveling with the speed of sound in the air (usually around 330 - 340 m/s), the airblast is a broadband signal that obscures valuable reflection data and provides no information about subsurface. It is desirable to have this signal removed with as little effect on the underlying data as possible. Rather than using a surgical mute, we demonstrate the use of a $t-x-f-k$-localized complex wavelet domain velocity filter to accomplish this task.

The filtering is done by multiplying complex wavelet coefficients with the corresponding filter masks whose values can vary continuously between 0 and 1. Because the moveout of an airblast can be predicted with a known source location and air wave speed, we are able to define appropriate masks as shown in Figure 4.7. We use a chart from Figure 4.3(a) to pinpoint $\mathbb{C}$WT orientations that should contain the majority of the energy we seek to remove. Figure 4.8 shows an ariblast contaminated land record, its filtered version, and the difference between raw and filtered data. To avoid signal distortions at near offsets (-90,90) m, we can exclude them from the filtering either normalizing the corresponding $\mathbb{C}$WT filter coefficients to 1 or by re-injecting near offset traces after filtering. Note that the filter is effective: the energy corresponding to the airblast is largely removed and the underlying signals are preserved. Due to the spatio-temporal filter localization, none of the steeply dipping events that do not follow the predicted airblast moveout are affected. This result is promising for a wide breadth of filtering applications for noise that can be separated in $t-x-f-k$ space, such as surface and guided waves, first breaks, or vibroseis-generated chimney noise (Bagaini et al., 2014).

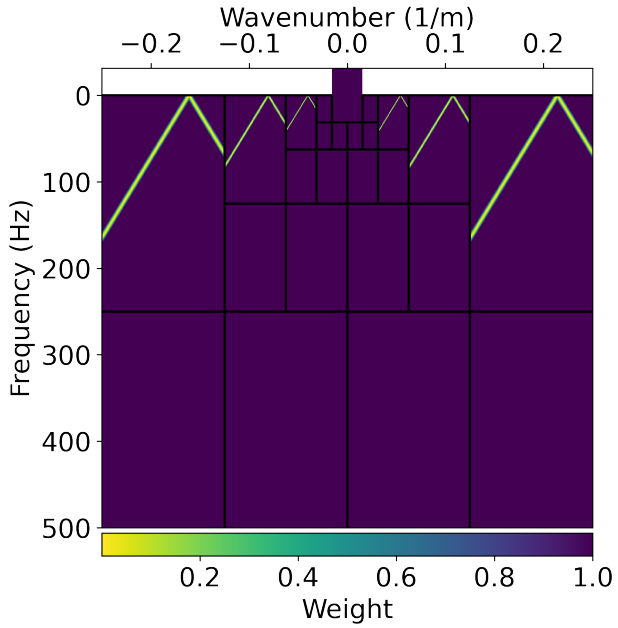


Figure 4.7 Complex wavelet domain filter masks for airblast suppression. Note that we only target the scales and orientations that are the most affected by the airblast energy, leaving others unchanged.

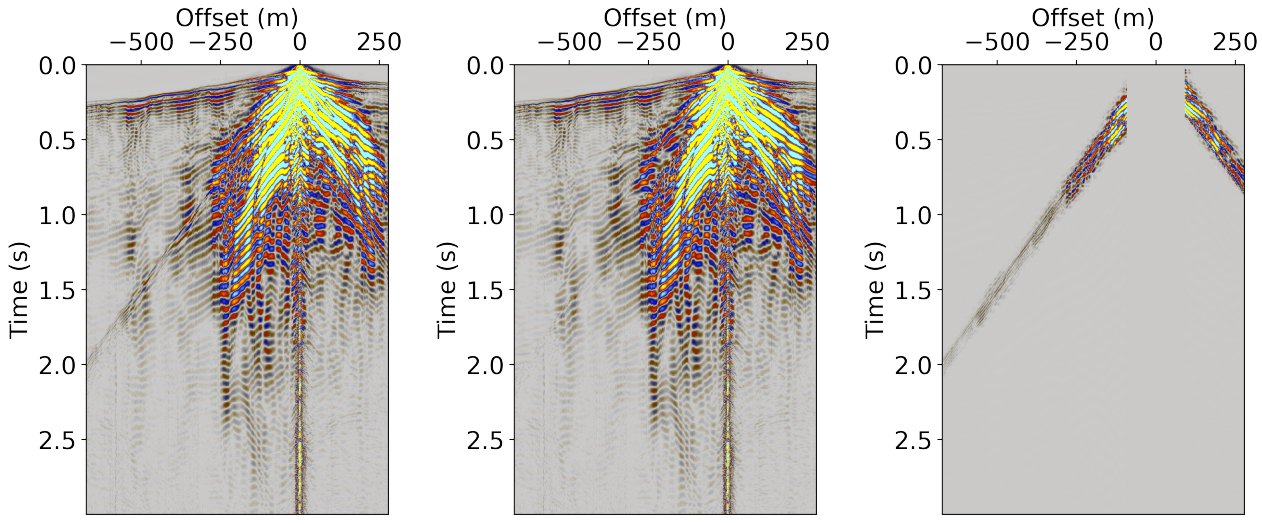


Figure 4.8 Airblast removal in the complex wavelet domain. (a) raw data, (b) after removing the airblast, and (c) data difference. Data are gained and clipped for display. Note that the majority of the airblast energy is removed without affecting other events.

4.4 Discussion

We demonstrate that complex wavelets can be successfully applied to solve a variety of geophysical problems involving data reconstruction and filtering. The key to successful processing outcomes is in the recording parameters of signals: array length and recording time affect the number of possible decomposition levels of $\mathbb{C}$WT and effectively control transform resolution, while spatial and temporal sampling intervals define Nyquist wavenumber and frequency. Because $\mathbb{C}$WT decomposes signals into octave bands, data sampling affects feature separability.

We find that a multiscale iterative soft thresholding is effective in restoring aliased signals. However, this technique assumes that there is a scale for which the aliasing effects are minimal or non-existent. When aliasing has severe imprint on the coefficients at the deepest possible decomposition level, the reconstruction is likely to fail. A partial solution for overcoming this limitation is to use a constant velocity NMO prior to data reconstruction to lessen the degree of aliasing (Yu et al., 2017).

Another important factor to consider is the data dimensionality. In this chapter, we consider 2D examples for reconstruction and filtering, but our techniques can be extended to higher dimensions. Working in higher dimensions, in particular when including reciprocal information from shot and receiver gathers, can have implications for infilling large gaps. Although it is still true that no data reconstruction technique can create new information, access to a volume with multiple illuminations increases chances that the features which should be present inside a gap are captured somewhere in the data volume. Thus, we expect higher-dimensional reconstructions to yield better gap infills.

For the reasons mentioned above, $\mathbb{C}$WT reconstruction and filtering outcomes are data dependent. It is important to consider data sampling to establish what can be reasonably achieved. With favorable data parameters, $\mathbb{C}$WT velocity filters are effective at suppressing coherent noise that can be separated in $t - x - f - k$ space because we can define the spatial and temporal origin of signals to be suppressed as well as their moveout or range of moveouts. Examples include

surface waves, guided waves, and first breaks. One can also define a narrow time-space cone to address chimney noise generated by vibroseis at near offsets (Bagaini et al., 2014).

4.5 Conclusions

Seismic data processing, imaging, and inversion outcomes are the best when we have access to nonaliased data that do not have significant gaps. The complex wavelet domain can help in achieving that if the acquired data do not already have these desirable attributes. The multiresolution character of wavelets naturally predisposes them to address sampling challenges. We show that by leveraging nonaliased information from the deepest decomposition level, our multiresolution iterative soft thresholding is successful both in reconstruction beyond aliasing and in large gap reconstruction.

One of the advantages of nonaliased data is the significantly improved ability for noise suppression. Separating signal and noise is contingent upon achieving the best possible noise separation in the transform domain. Because $\mathbb{C}$WT is localized in $t - x - f - k$ space, it allows one to design filters based on velocities as well as expected $t - x$ moveouts. The advantage of such design is demonstrated with airblast removal: we show that $\mathbb{C}$WT velocity filter is effective in removing airblast energy and does not affect any signals outside of expected noise trajectory.

CHAPTER 5

MULTI-CHANNEL COMPRESSIVE SENSING FOR SEISMIC DATA RECONSTRUCTION USING JOINT SPARSITY

Based on a paper presented at Second International Meeting for Applied Geoscience & Energy in August 2022

Iga Pawelec[1,2], Michael Wakin[2], Paul Sava[2]

Despite the many advances in data reconstruction technology and significant increase in the number of channels available for recording the particle motion, acquiring non-aliased seismic data at high signal-to-noise ratio remains a challenge. Areas with restricted access, difficult terrain, and slow near-surface velocities can prevent the acquisition of properly sampled wavefields, which in turn significantly complicates the suppression of the near surface related noise. There are two popular strategies available for reducing the sampling requirement without loss of information. Recording wavefield and its derivatives yields multiple pieces of information at each sampling point, creating a multi-channel signal and allowing for an increased distance between samples. Compressive sensing (CS) is an alternative way of data acquisition relying on randomized sampling and known data patterns in some domain to reconstruct a fully sampled seismic data volume from reduced measurements. These two approaches can be combined by recording multi-channel information on a randomized grid. In this chapter, we modify the iterative hard thresholding algorithm (IHT) to allow for combining different data inputs (channels) and thus estimating joint sparse signal support. First, we use Fourier-sparse 1D signals to build intuition about the probabilistic nature of compressive sensing and its multi-channel extension. Then, we apply complex wavelet domain IHT on multiple realizations of randomly subsampled 2D land seismic field data with available spatial derivative channel. Our intuition-building 1D experiment

[1]Primary researcher and author.
[2]Colorado School of Mines, Golden, Colorado, United States of America.

demonstrates that the multi-channel approach yields higher success rate for signal recovery and that intuition seems to hold for a 3D synthetic data. However, in the field data application the definition of success is less clear. Depending on which traces are used to compute the normalized root-mean-square error (NRMSE) and how they are scaled, the reconstruction quality evaluation for a specific realization can be misleading. We show that using range compression before computing NRMSE yields results that better align with the intuition of a human observer. Comparing ‘good‘ and ‘bad‘ reconstructions reveals that the wavefields recovered for single- and multi-channel approach are fairly similar, with the multi-channel approach showing slightly better event continuity for some realizations. The biggest reconstruction challenge is posed by the near-offset region, with large concentration of high-amplitude, short-wavelength features. We demonstrate that, even for realizations with high NRMSE, phase of the wavefield tends to be recovered accurately, while the amplitude error can be significant.

5.1 Introduction

Obtaining non-aliased seismic data with high signal-to-noise ratio (SNR) can be a challenge, particularly in difficult terrain or in areas with access restrictions. The conventional approach to seismic data acquisition relies on recording the particle motion on a regular grid, with spacing between stations dictated by the maximum frequency of the signal and the slowest velocity (water velocity for marine acquisition, or near-surface velocities for land acquisition). However, due to terrain obstacles and slow near-surface velocities (sometimes < 200 m/s), dense regular sampling on land can be economically infeasible or even impossible.

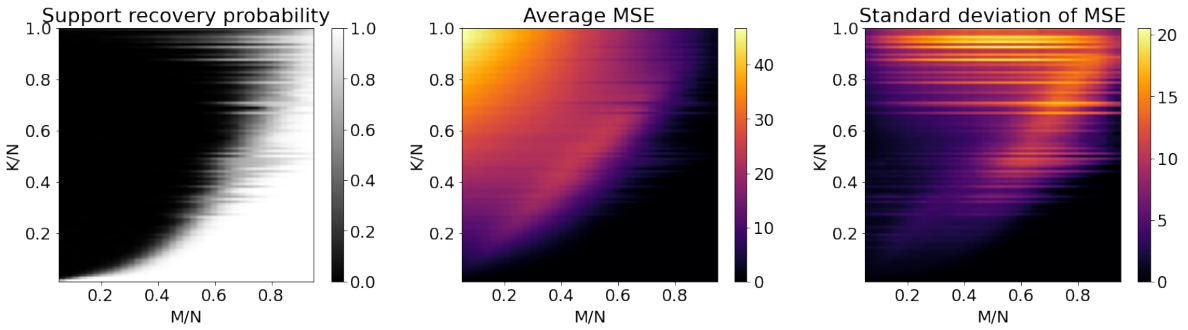
In light of the challenges associated with regularly sampled data, there are many available techniques to correct for deviations from regular grids and to mitigate the aliasing. A large number of these techniques use the Fourier representation of seismic data. For instance, Liu and Sacchi (2004) develop a framework for data recovery based on L_2 norm minimization, using spectral weights bootstrapped from frequency-wavenumber (FK) representation of data. Similar strategy extended to five dimensions (Trad, 2009) is even more successful because data in higher

dimensional spaces tend to have more compact representations (they achieve higher sparsity level) and thus are easier to reconstruct. Duijndam et al. (1999a) tackle the problem of arbitrarily irregular sampling and leverage a weighting scheme based on adjacent sample distances to reconstruct data with one varying spatial coordinate. Xu et al. (2010, 2005) propose an antileakage version of the Fourier transform that can both handle irregular geometry and mitigate aliasing problems.

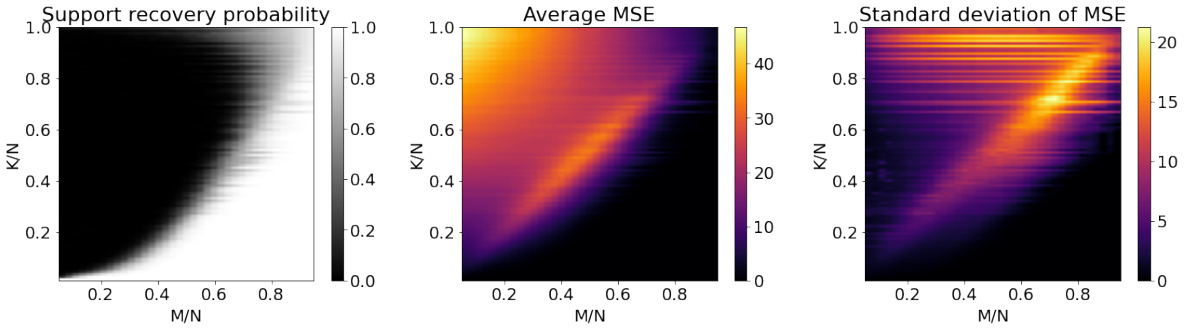
The advent of wireless nodal acquisition shifts the way we think about land data acquisition. An increasingly popular alternative to regular sampling is deliberately randomized acquisition which exploits the data patterns in some domain (Allegar et al., 2017; Jiang et al., 2019; Mosher et al., 2017). Assuming that these patterns can be approximated as sparse, compressive sensing (CS) theory offers many strategies for successful recovery of fully sampled signals on a regular grid from sub-Nyquist number of sampling points. The challenge for CS land seismic data acquisition is to find a data representation that captures all features of raw records (including amplitudes spanning several orders of magnitude and instantaneous phase) in what can be considered a sparse form. As shown in Figure 5.1, there is a trade-off between data sparsity level and the likelihood of successful reconstruction. The less sparse the pattern, the more measurements are needed to ensure success. Pawelec et al. (2021) demonstrate CS recovery for a complex raw land seismic record that reiterates the intuition that the gap pattern and the number of measurement points is critical: the success of recovery depends on a specific realization of random sampling geometry.

We propose to use spatial derivatives of particle motion in addition to the particle motion signal within the CS framework to further reduce the number of measurement points or improve the SNR of the recovered data. Following the terminology from the signal processing community, we use the term ‘multi-channel’ to refer to multiple measurements available at the same location, i.e. the particle motion signal and its spatial derivative(s). We develop a framework for simultaneous sparse approximation of wavefields and their derivatives using the iterative hard thresholding algorithm (Blumensath and Davies, 2008). Expanding the Fourier domain

implementation reported in (Pawelec et al., 2022) to discrete complex wavelet transform, we present synthetic and field data examples showing that the multi-channel approach yields superior reconstructions in terms of average SNR per reconstructed missing trace when compared to single-channel reconstructions in majority of tested scenarios. However, due to the complex interactions between specific realizations of a sampling pattern and signals' representation in the domain of interest (Fourier or complex wavelet), there are also cases when the additional derivative information has an adverse effect on the wavefield reconstruction quality.



(a)



(b)

Figure 5.1 The recovery probability, average mean-squared error and its standard deviation for (a) single channel and (b) multi-channel recovery of a 1D signal sparse in the Fourier domain. K/N is the fraction of non-zero coefficients needed to describe the signal, while M/N is the fraction of available samples from the fully sampled signal. The transition between successful and failed recovery is clear for both scenarios, but reconstructions are more consistent for the multi-channel case.

5.2 Wavefield Derivatives

Following the traditional sampling theorem (Jerri, 1977; Shannon, 1948) and its extension to multidimensional signals (Petersen and Middleton, 1962) we can recover signals without the loss of information provided that all measurement points on the sampling grid are available. This approach to sampling can be extended for multi-channel signals. Linden (1959) provides a reconstruction formula for simultaneous sampling of a band-limited function and its derivative. Papoulis (1977) generalizes that result showing that a band-limited function $f(t)$ can be uniquely described by samples of m linear systems with input $f(t)$ sampled at $\frac{1}{m}$ the Nyquist rate. Cheung (1993) extends the theory to multidimensional signals. In the Appendix A, we show the derivation of a 2D generalized sampling expansion (GSE) for the orthogonal sampling matrix. Note that to increase the distance between the samples by a factor of two, four channels are required.

An interesting application of generalized sampling expansion is presented by Robertsson et al. (2008). They use it to mitigate the issue of coarse crossline sampling in towed streamer data. The pressure gradient is computed from three-component measurements of particle velocity and then combined with independently recorded pressure, providing a two-channel measurement at each sampling location. Paired with some simple data processing, such an approach allows one to recover data up to three times the Nyquist wavenumber compared to the single channel measurements.

A natural question arises: is it possible to push the GSE method for seismic data reconstruction even further? Vassallo et al. (2010) show examples of unraveling multiply aliased data relying on the known velocity of the water and predicting aliasing patterns for pressure and its derivative. Their approach, however, cannot be readily applied in land acquisition because the near-surface velocities are usually unknown a priori, are rarely constant, and vary as a function of space.

Working with onshore data and with access to wavefield spatial derivatives, Muyzert et al. (2019) recover land records with the ground roll aliased up to three times, similar to the results shown by Robertsson et al. (2008). Pushing beyond that result on land may require a different

approach. Because we cannot rely on known aliasing patterns and obtaining higher-order derivatives can be challenging due to noise considerations, a possible solution is to explore multi-channel recordings in a compressive sensing framework.

Compressive sensing is known for being able to recover sparse signals from significantly reduced measurements, provided that the sampling strategy and sparse domain are carefully selected. Supplementing the compressive sensing framework with the derivative information should, in principle, push the limits of what is possible with one channel only - or with regularly sampled multi-channel signals. Because differentiation is a linear operation, wavefield derivatives have the same bandwidth as the original data, providing an important commonality between different channels, i.e. a shared Fourier domain support. This allows for the use of derivatives in the Fourier-based multi-channel CS framework by exploiting the idea of joint sparsity.

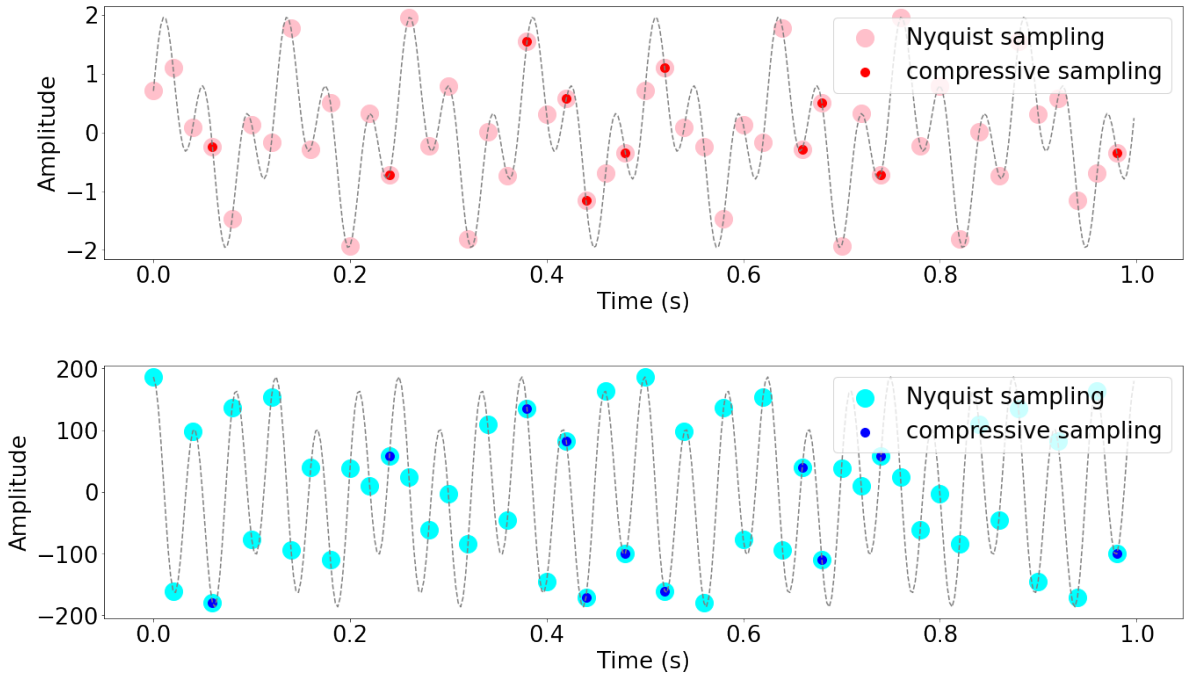


Figure 5.2 1D signal composed of two sinusoids recorded with a multi-channel sampling system: by taking discrete samples of a signal (top) and its first derivative (bottom). Large light circles on both panels indicate Nyquist sampling locations while small dark circles show one realization of compressive sampling. Note that compressive samples consist of only 22% of the Nyquist samples.

5.3 Compressive Sensing

Compressive sensing is a sampling paradigm allowing for recovery of signals from incomplete measurements (Candès et al., 2006b). The key assumption is that the target signal is or can be approximated as sparse in some representation. In particular, an $N-$length $K-$sparse signal can be expressed using only K non-zero coefficients, usually with $K \ll N$. Recovering that signal means finding the signal support – locations of non-zero coefficients – and values of the coefficients at these locations.

Successful recovery of $K-$sparse or compressible signal depends on three key components: the sampling strategy, the data sparsifying transform, and the sparsity-promoting recovery algorithm. Results from compressive sensing suggest that sparse signals can be recovered without loss of information if the sampling matrix satisfies the restricted isometry property (RIP) (Baraniuk, 2007). RIP is satisfied with high probability for Gaussian matrices (each entry is independent and follows a normal distribution) and random Bernoulli matrices (entries are ± 1 with equal probability) or when sampling non-uniformly Fourier-sparse signals. Depending on the choice of the sampling matrix, the number of measurements to recover a K-sparse signal is $M = \mathcal{O}(K \log(N/K))$.

Let the data in sparse domain be represented as

$$\mathbf{d} = \mathbf{S}\mathbf{\Phi}^H \boldsymbol{\alpha}, \tag{5.1}$$

where $\mathbf{d}$ is the recorded wavefield, $\mathbf{S}$ is the sampling matrix, $\mathbf{\Phi}$ is the sparsifying transform, and $\boldsymbol{\alpha}$ are the signal coefficients in the sparse domain. From a practical point of view, only the non-uniform sampling (i.e., either placing the receiver and/or source on the grid or skipping the sampling location altogether) can be achieved for seismic acquisition. Although the Fourier domain is not optimal for sparsely representing seismic data, it is possible to obtain a data window that can be approximated as Fourier-sparse with a clever combination of sorting, pre-processing, and windowing. Furthermore, the Fourier domain is natural to consider for extending to the multi-channel case given the relationship between signal and its derivative:

$f'(t) = \mathcal{F}^{-1}(i\omega F(\omega))$.

In the joint sparse recovery problem, also known as simultaneous sparse approximation or multiple measurement vector problem, rather than dealing with one sparse signal, we are attempting to recover an ensemble of p signals:

$$\mathbf{d}_i = \mathbf{S}\mathbf{\Phi}^H \boldsymbol{\alpha}_i \; i = 1, \ldots, p. \tag{5.2}$$

Figure 5.2 shows an example of a 1D signal composed of two sinusoids and recorded with two channels: by recording samples of signal and its derivative. Each signal is individually sparse in the Fourier domain (Figure 5.3), but there is also a relationship between the channels. That relationship can be theoretical, as is the case here, or statistical. Formally, the interrelations between channels are described by a joint sparsity model (Duarte et al., 2005). The model we consider in this chapter is the common sparse support, where each individual signal coefficient vector has the same support, but the coefficient values can differ.

Consider the matrix $\mathbf{A} = [\boldsymbol{\alpha}_1 \cdots \boldsymbol{\alpha}_p]$. If all $\boldsymbol{\alpha}_i$ share the same support and are K-sparse, then $\mathbf{A}$ contains $N - K$ zero rows. Many algorithms are available for recovering this type of signals. They include simultaneous orthogonal matching pursuit (Tropp et al., 2006), convex relaxation (Tropp, 2006), subspace-based methods (Lee et al., 2012), deep learning (Palangi et al., 2016) and Bayesian approaches (Chen et al., 2016a; Wipf and Rao, 2007). Due to the ease of implementation we consider a modification of the iterative hard thresholding (IHT) (Blumensath and Davies, 2008), as described next.

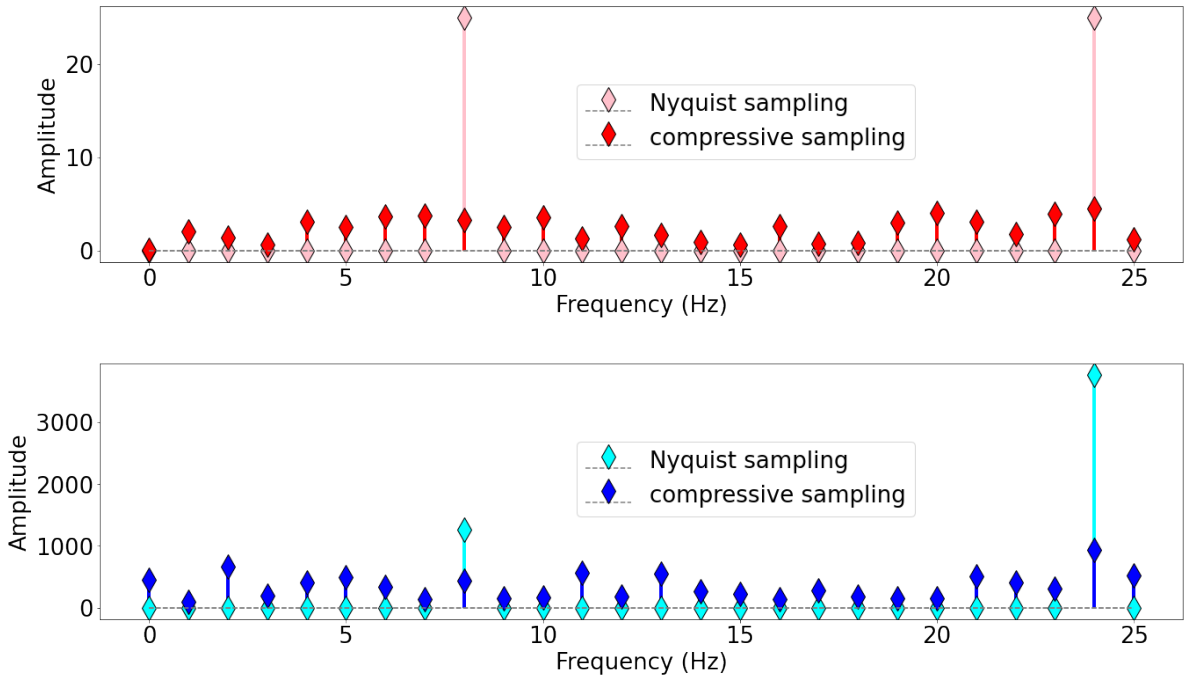


Figure 5.3 Fourier domain representation of the signal shown in Figure 5.2. Note that severe subsampling in compressive sampling makes it challenging to accurately estimate the true support of the underlying signal (8 and 24 Hz in this example).

In the case of single-channel reconstruction, we are interested in solving the following K-sparse optimization problem:

$$\min_{\boldsymbol{\alpha}} \mathcal{J} = \|\mathbf{d} - \mathbf{S}\boldsymbol{\Phi}^H\boldsymbol{\alpha}\|_2^2 \text{ subject to } \|\boldsymbol{\alpha}\|_0 \leq K, \tag{5.3}$$

where $\|\cdot\|_0$ refers to the number of non-zero entries. This problem can be solved with the following iterative algorithm:

$$\boldsymbol{\alpha}^{n+1} = T_K\Big(\boldsymbol{\alpha}^n + \boldsymbol{\Phi}(\mathbf{d} - \mathbf{S}\boldsymbol{\Phi}^H\boldsymbol{\alpha}^n)\Big), \tag{5.4}$$

with the non-linear thresholding operator T_K retaining only K coefficients with the largest magnitude

$$T_K(\alpha_i) = \begin{cases} 0 \text{ if } |\alpha_i| < \lambda_K(\boldsymbol{\alpha}), \\ \alpha_i \text{ if } |\alpha_i| \geq \lambda_K(\boldsymbol{\alpha}). \end{cases} \tag{5.5}$$

The threshold $\lambda_K(\boldsymbol{\alpha})$ is set to K-th largest absolute value of $\boldsymbol{\alpha}^n + \boldsymbol{\Phi}(\mathbf{d} - \mathbf{S}\boldsymbol{\Phi}^H\boldsymbol{\alpha}^n)$. For jointly sparse signals with different physical units, we reformulate the problem from equation 5.3 into

$$\min_{\boldsymbol{\alpha_1}\cdots\boldsymbol{\alpha_p}} \sum_{i=1}^{p} \|\mathbf{d}_i - \mathbf{S}\boldsymbol{\Phi}^H\boldsymbol{\alpha}_i\|_2^2 \text{ subject to } \|\sum_{i=1}^{p} \mathbf{W}^i\boldsymbol{\alpha}_i\|_0 \leq K. \tag{5.6}$$

In this formulation, we aim to honor the acquired signal samples while promoting $K-$sparse joint support. To estimate this support, the Fourier or complex wavelet domain representations of the respective channels are combined into one. This is achieved by weighing the respective spectra such that all spectral coefficients corresponding to the first channel (signal samples) remain unchanged ($\mathbf{W}^1$ is an identity matrix). For the remaining channels, $\mathbf{W}$ is a diagonal matrix with entries defined as

$$\mathbf{W}_i^n = s_i \frac{1}{j\omega_i + \epsilon_n},\ \epsilon_n = \frac{\|\boldsymbol{\Phi}\mathbf{d_n}\|_2}{\|\boldsymbol{\Phi}\mathbf{d_1}\|_2}, \tag{5.7}$$

where ω_i represents cycles in appropriate physical units (for example, $\omega_i = 2\pi f_i$ for temporal signals), and s_i denoting a scalar regulating the contribution of the $i-th$ channel (by default set to 1) for the Fourier domain or simply as

$$\mathbf{W}_i^n = s_i \frac{\|\boldsymbol{\Phi}\mathbf{d_1}\|_2}{\|\boldsymbol{\Phi}\mathbf{d_n}\|_2}, \tag{5.8}$$

for the complex wavelet domain. The purpose of the weighting term is thus either to partially 'undo' the derivative operation, using ϵ as a damping factor that also balances the relative contribution of derivative channels with respect to data channel, or to normalize derivative channel contributions with respect to the data channel. An example of the Fourier domain weighting function for a 1D time signal is depicted in the top panel of Figure 5.4. With the weight applied to the derivative channel (dark blue diamonds), its spectral coefficients have similar magnitudes to the coefficients of the signal channel (red diamonds). Adding these two together (black diamonds) yields a joint spectrum whose top two highest energy coefficients are at the same frequencies as for the underlying sparse signal. When correct sparse support can be estimated, signal recovery tends to be good, as discussed in the Examples section.

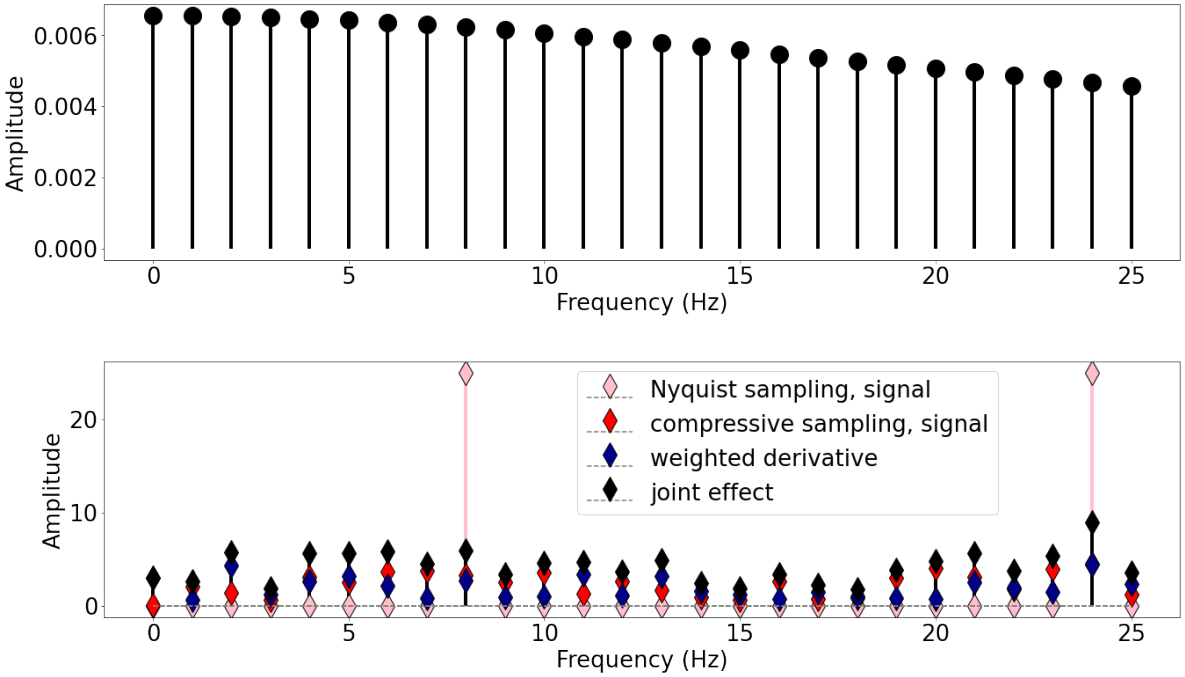


Figure 5.4 Weights applied to the derivative channel (top) and the effect of combining compressive signal samples with weighed compressive derivative samples. Note that in this instance, the combined signal spectrum yields highest energy coefficients at 8 and 24 Hz, the same as the Nyquist sampled signal. However, this property may not hold for other realizations of signal sampling.

5.4 Data Reconstruction Examples

5.4.1 1D Sparse Sinusoids

A simple, proof-of-concept experiment demonstrating the joint sparse recovery for Fourier-sparse signals is reconstructing the superposition of sinusoids. With compressive samples shown in Figure 5.2 and weights from Figure 5.4, we use the iterative hard thresholding setting $k = 2$. Recovered Nyquist samples are shown in Figure 5.5. Supplying the derivative information helps to estimate the correct signal support, leading the reconstruction towards the correct solution. Single channel on its own is insufficient to correctly estimate signal support in this case, leading to erroneous estimate for one of the coefficients and thus incorrect solution.

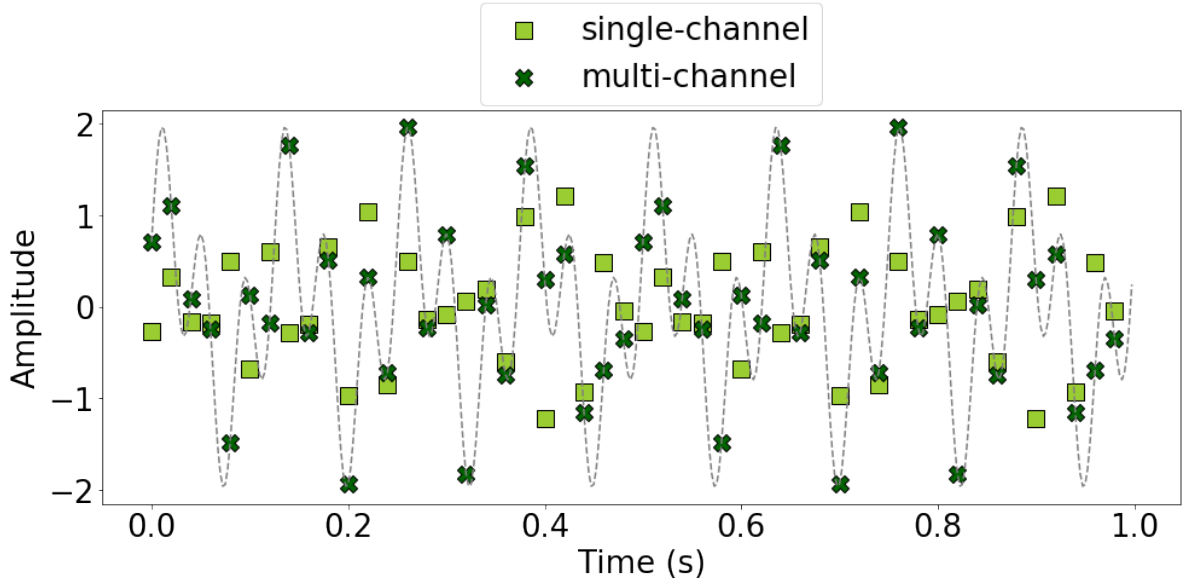


Figure 5.5 Reconstruction of a 1D signal shown at the top panel in Figure 5.2 using single- and multi-channel IHT. Note that the multi-channel approach allows to recover the Nyquist samples up to the numerical precision.

A similar experiment is repeated for different sparsity ratios K/N (number of frequency coefficients needed to represent the signal divided by the total number of frequency coefficients for a given signal length) and data subsampling ratios M/N, with 200 realizations for each $(K/N, M/N)$ combination. In our test, we assume that all sparse signal components have equal strength (i.e., the amplitude of the sinusoid is always the same). The compressive sampling of

both channels is done by selecting a fixed number of samples M from a Nyquist-sampled $N-$length signal (like in Figure 5.2) uniformly at random.

Figure 5.6 shows the success rate of signal reconstruction between different signal sparsity levels, numbers of kept samples, and realizations. The success is defined as recovering the correct Fourier domain support. Note that the successful recoveries also have very high values of SNR, often as high as 300 dB, indicating signal recovery within the numerical precision. That is only possible when signals are exactly $K-$sparse and are noise-free. The volumes also highlight the probabilistic nature of compressive sensing. There is a clear and sharp transition between success and failure, that is best captured by summary statistics derived from all realizations. Figure 5.1 is an example of such: the probability of successful recovery is averaged over the number of realizations, and the mean-squared-error is used to assess the accuracy of each reconstruction. We find that the multi-channel approach is successful with fewer samples than the single-channel approach, but behaves more erratically in the transition zone between 100% success and 100% failure rate.

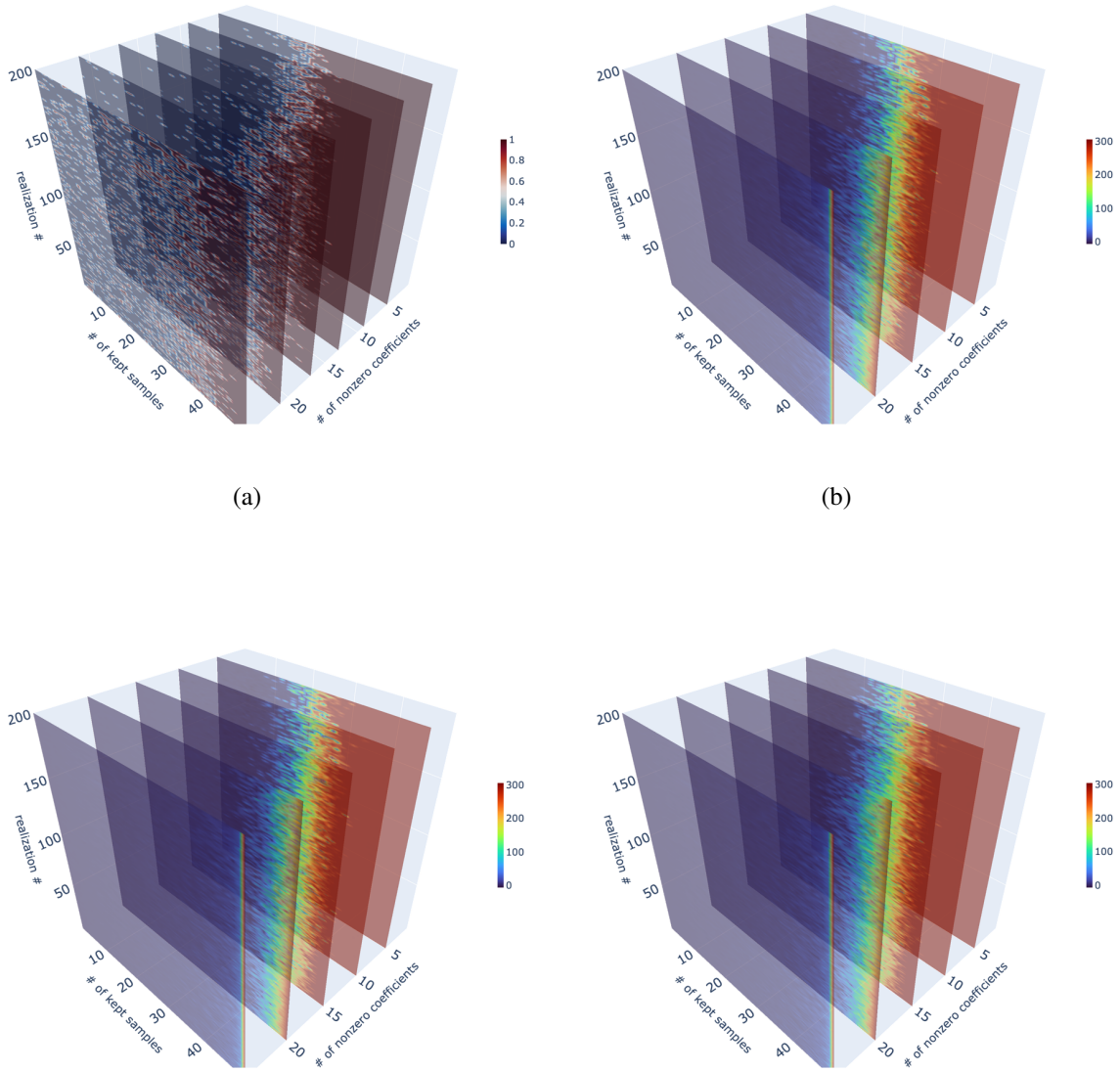


Figure 5.6 Quality metrics for single- and multi-channel compressive sensing reconstruction (top and bottom rows, respectively). (a) and (c) are the probabilities of recovering the correct Fourier domain support while (b) and (d) show the quality of the reconstructed signal quantified by SNR.

5.4.2 Synthetic 3D Test

We test the applicability of multi-channel compressive sensing on a simulated seismic shot. Unlike in the previous experiment, synthetic seismic data are not perfectly sparse in the Fourier domain. However, if the amplitude variability is not large or if we consider only a small data window, we can approximate the data as sparse, thus making them $K-$compressible. Figure 5.7 shows two slices through a densely sampled 3D seismic volume with all available traces at the top and compressively sampled data at the bottom. The subsampling ratio for this experiment is 80%, and due to the computational cost involved, we only consider one realization of the missing trace geometry. Figure 5.8 and Figure 5.9 show the reconstruction results for single- and multi-channel approaches, respectively. The multi-channel reconstruction yields results with higher SNR (12.69 dB compared to 11.60 dB) and better event continuity. The near-offset reconstruction artifacts stemming from the difficulty of approximating fast amplitude decay as sparse are also reduced for the multi-channel case. Additionally, the multi-channel approach reconstructs all channels while enforcing the common sparse support constraint (Figure 5.10). Maintaining this consistency between the channels is particularly valuable when exploring the use of wavefield derivatives for other applications, such as denoising or mode separation.

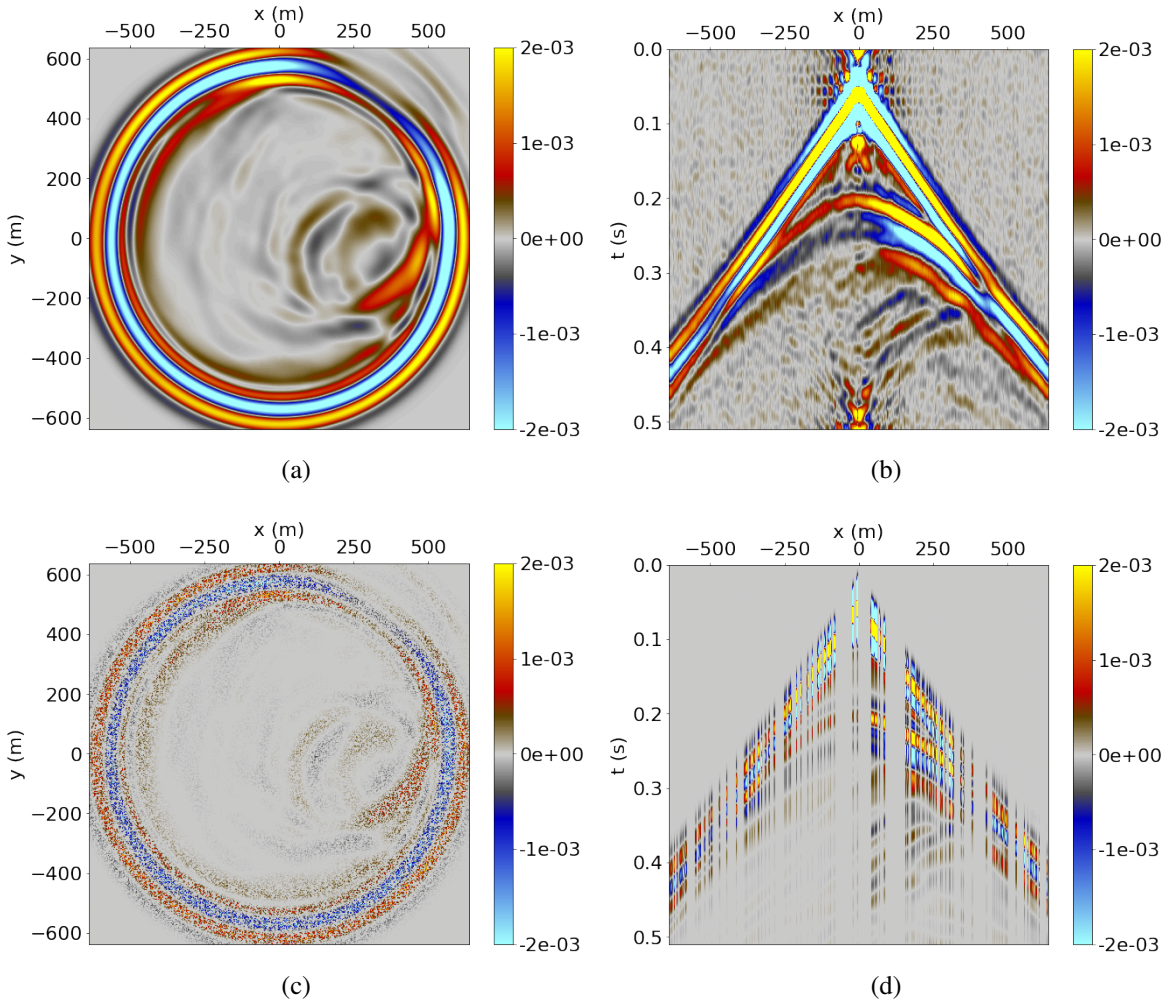


Figure 5.7 Reconstruction of coarsely sampled seismic data. (a), (b) are the slices from densely sampled 3D volume and (c), (d) are the compressive samples after 80% subsampling.

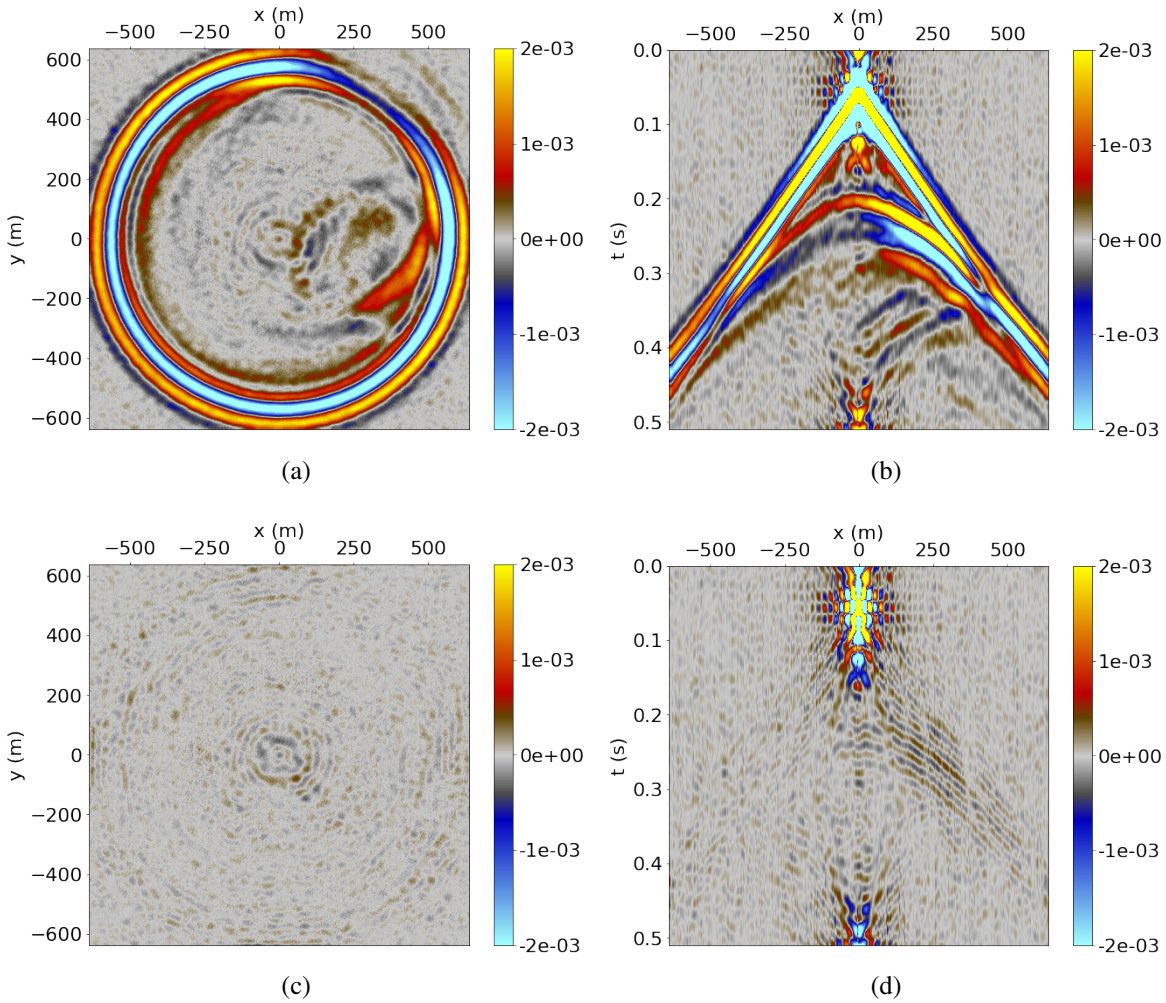


Figure 5.8 Single-channel reconstruction for the wavefield depicted in Figure 5.7 and the corresponding differences. The SNR for the recovered volume is 11.60 dB.

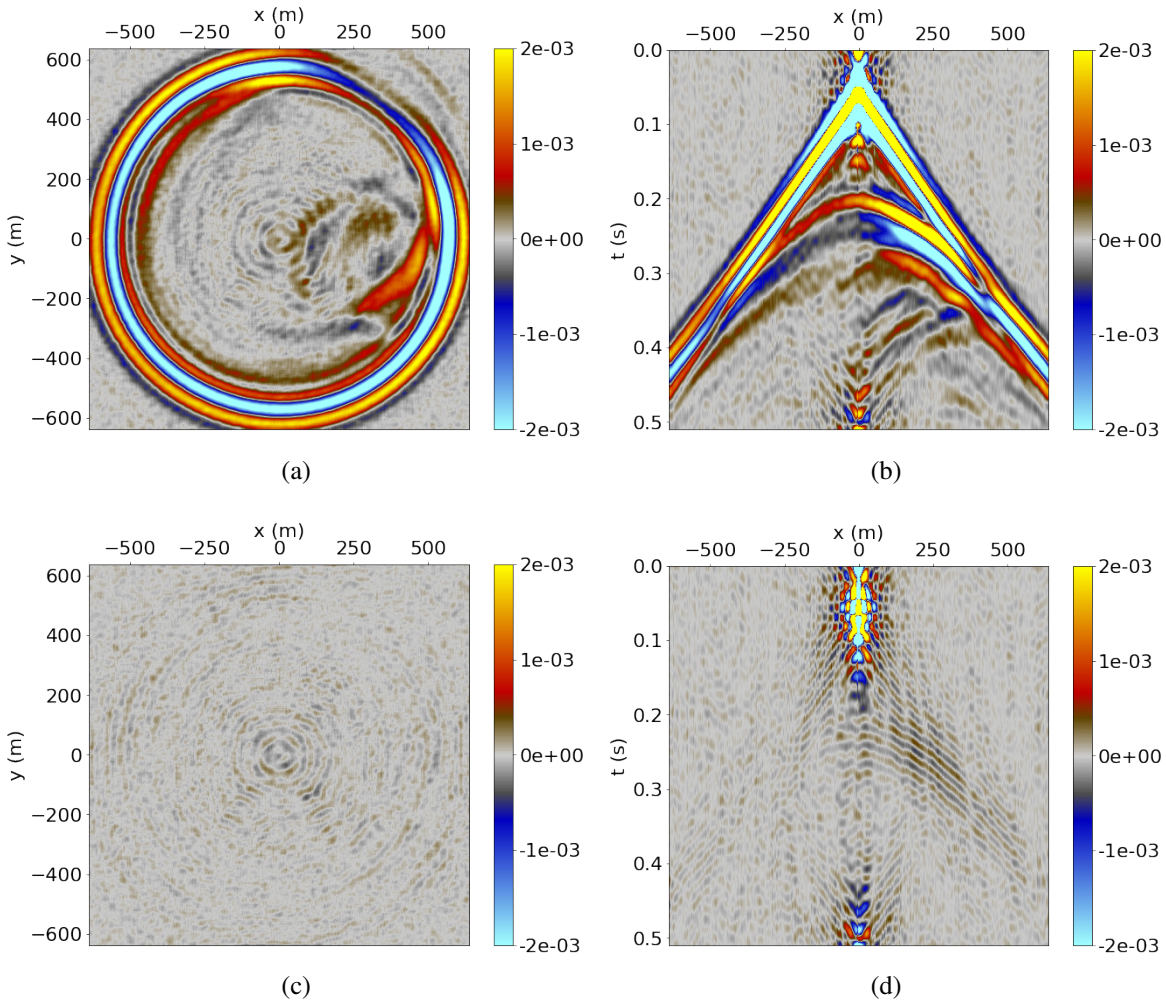


Figure 5.9 Multi-channel reconstruction for the wavefield depicted in Figure 5.7 and the corresponding data differences. The SNR for the recovered volume is 12.69 dB.

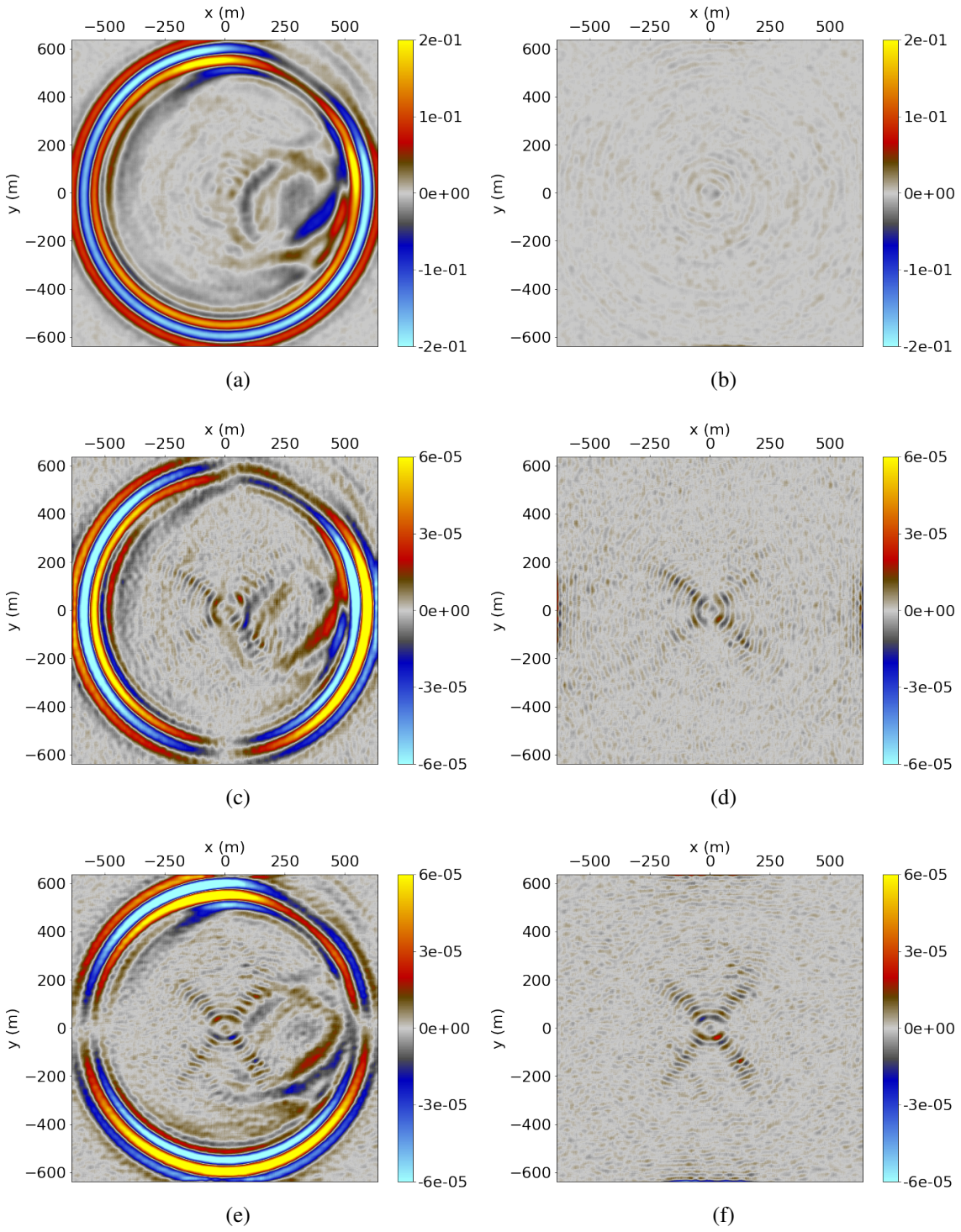
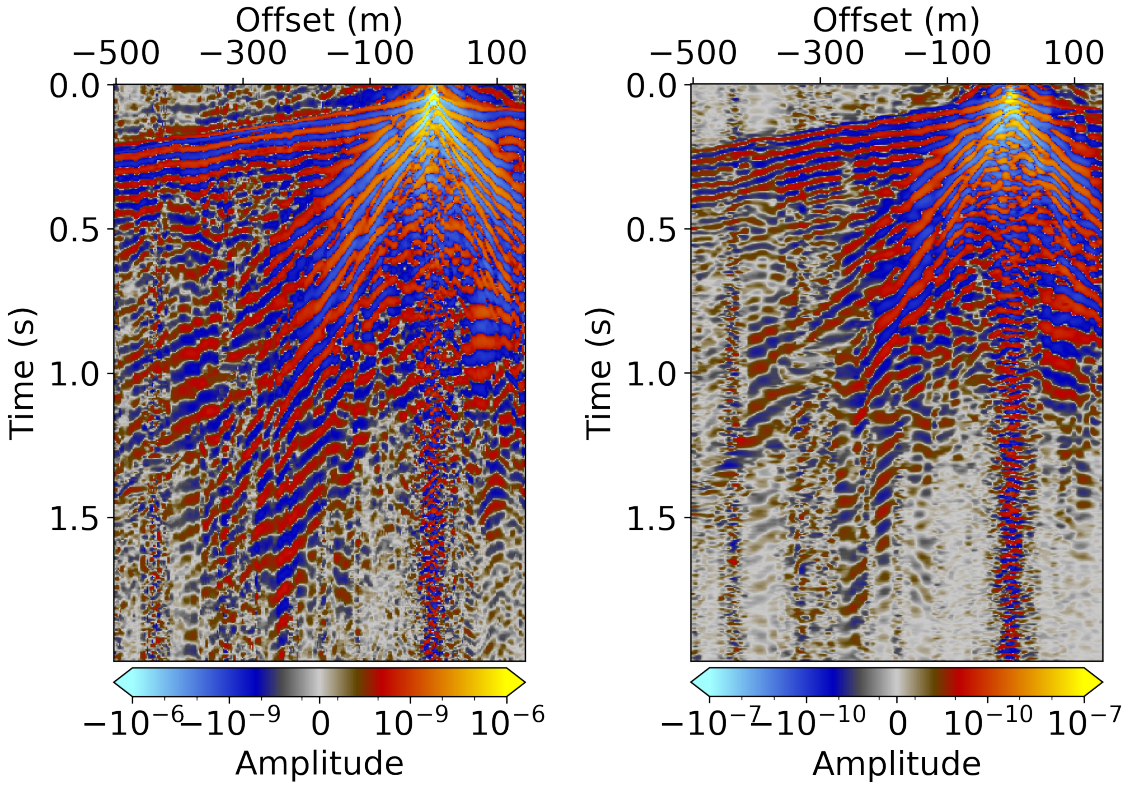


Figure 5.10 Simultaneous sparse approximation of 4 channels: (a) data samples, (b) time derivative, (c) x derivative, and (d) y derivative. Enforcing the common sparse support constraint helps to preserve the physical relationship between the channels.

5.4.3 2D Field Data

To see whether the proposed approach has merit in a real seismic acquisition scenario, we devise reconstruction experiments using field data acquired as a part of Colorado School of Mines Geophysical Field Camp. Data were collected in Golden, Colorado on the foothills of Mount Zion with 6-120 Hz, 3 dB/oct vibroseis sweep as a source signal. Source and receiver spacing is 2 m, and the inline derivative information is obtained by taking a centered finite-difference approximation from the mini gradient array formed by three receivers placed approximately 0.25 m apart.

Figure 5.11 shows gathers selected for reconstruction experiments. Note that the inline derivative is more noisy than data themselves, with noise likely affecting derivative estimation around -400 m offset. We use a symmetric logarithmic scale for data display to emphasize large dynamic range for significant events, with one order of magnitude difference between data and derivative scale. The same scale is later used for displaying reconstructed wavefields in Figure 5.12 through Figure 5.17.



(a) (b)

Figure 5.11 (a) Land seismic gather and (b) its inline derivative. The symmetric logarithmic scale is used to showcase the range of significant amplitudes.

In this experiment, we subsample data from Figure 5.11 following the uniform random distribution, retaining 60% of the original traces. Mimicking the intuition-building 1D example, we compute 200 realizations of random sampling geometry to obtain an unbiased assessment for the performance of the multi-channel reconstruction. Unlike the previous example, we implement single- and multi-channel IHT in the complex wavelet domain because complex wavelets provide a better, sparser representation for land seismic data than Fourier coefficients (Pawelec et al., 2021).

Quantifying reconstruction success for a complex wavefield object is much more challenging than for a signal with known sparse support. First, wavefields are compressible rather than sparse in any chosen analysis domain. Second, complex wavelets form a redundant dictionary, meaning

that more than one support is possible and can yield comparable data reconstruction. This prevents us from comparing the support of fully sampled data to that of recovered data. Third, when handling field data, noise is always present in the records but its level and character can vary as a function of time and space. While reconstructing data, a choice has to be made about the level of fitting for the original data, and depending on that choice, one can achieve simultaneous reconstruction and denoising. Therefore, any pixel-by-pixel comparison of the original and reconstructed data can misrepresent the geophysical value of the reconstructed wavefield. Finally, unlike in image reconstruction problems, seismic amplitudes vary over several orders of magnitude, with relatively weak energy features being just as, if not more important as the strong energy features. Although there are many metrics available for image quantification (Eskicioglu and Fisher, 1995), most rely on sample-by-sample comparisons between the original and distorted - or in our case reconstructed - object. Such comparisons are inevitably biased towards high energy features and often fail to convey the level of distortion in the reconstructed wavefield in a geophysically meaningful way. And yet, some sort of metric is required to judge the quality of an individual reconstruction as well as to compare different realizations.

To explore the reconstruction metric problem, we use the normalized root-mean-square error (NRMSE) in four ways. We compute NRMSE for data without any amplitude alteration using either only the missing traces (Figure 5.12(a)) or the entire wavefield (Figure 5.12(b)). We then repeat the same process but for data that are first range compressed with a logarithmic transformation (Figure 5.13). Note that range compression is only for NRMSE computations, all reconstructions are done on data with their native range.

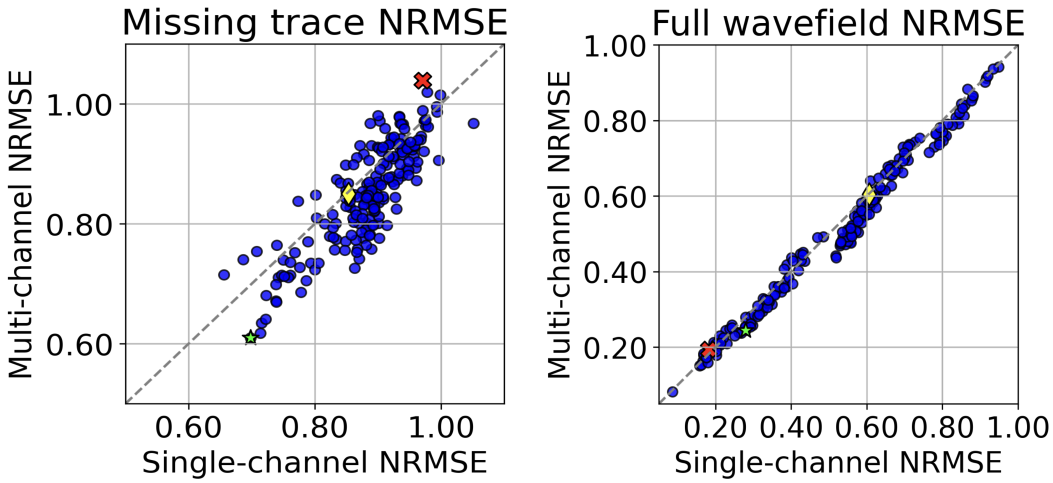
Figure 5.12 and Figure 5.13 summarize performance of single- and two-channel reconstruction approach. The best, average, and the worst realizations for the two-channel reconstructions from Figure 5.12(a) are marked with a star, a diamond, and a cross, respectively and are tracked throughout all other assessments. Note that the missing trace geometry with the highest two-channel NRMSE has one of the smallest errors when the full wavefield is considered in computations. This happens because the traces that are not missing are reconstructed almost up

to the numerical precision: the fewer near-offset traces are missing, the better the reconstruction is judged to be regardless of how well the weak energy features at further offsets are preserved. Looking at the NRMSE computed for range-compressed data, we can see that the green star representing the most favorable two-channel geometry from Figure 5.12(a) is also one of the better ones, though not the best. However, the least favorable two-channel geometry is assessed to be somewhere in the middle of the distribution, with plenty realizations that are judged as worse.

Figure 5.14 through Figure 5.17 show examples of specific single- and multi-channel reconstructions for different sampling realizations that are quantified as good or bad according to different metrics. The differences between original and reconstructed data can be difficult to assess on a difference plot, so as an alternative, we show an overlay of the reconstructed data on the original wavefield. Areas that appear brighter are where the largest differences are concentrated.

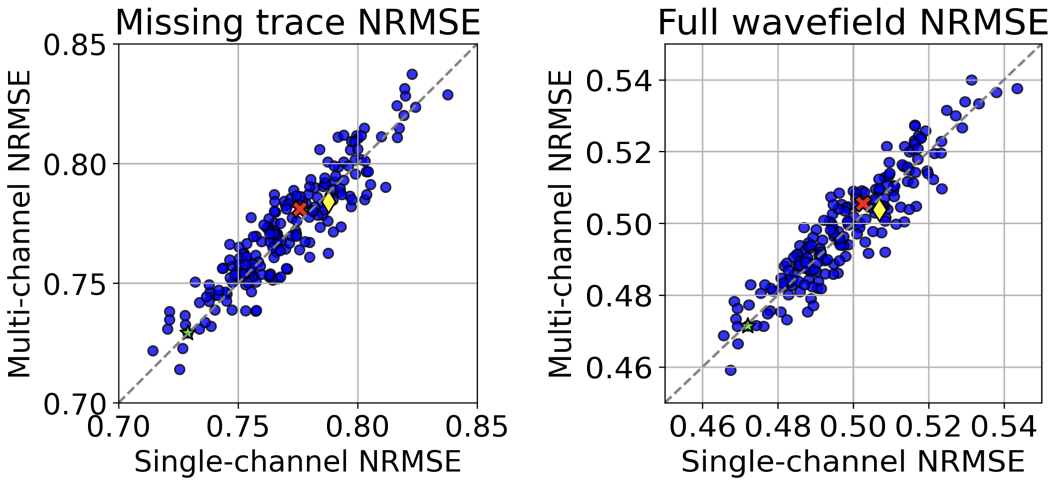
A common theme for all reconstructions is the struggle to accurately recover the high energy concentration at near offsets. Depending on the severity of data subsampling in that area, the error might be primarily in amplitude (Figure 5.15 and Figure 5.16) or both amplitude and phase (Figure 5.17). This is likely because near the source, there is a large concentration of short wavelength features whose amplitude changes rapidly. As a result, even a minor data subsampling in that area can cause information loss.

Comparing the single- and multi-channel reconstructions for the same realizations, we note that the recovered wavefields are quite similar. In the near offset areas where the differences are the most apparent, the addition of a spatial derivative channel tends to help with event continuity (Figure 5.14(b) and Figure 5.15(b)). However, no matter which of the reconstruction metrics is selected, there are also realizations for which NRMSE of the multi-channel approach is higher than for a single-channel. This is consistent with a synthetic 1D example which demonstrated that while the multi-channel approach is, on average, more successful in data reconstruction at specific data subsampling ratio, there can be specific cases where single-channel approach proves superior.



(a) (b)

Figure 5.12 Normalized root-mean-square error computed from (a) only the reconstructed missing traces and (b) full reconstructed wavefield including the originally present traces. Amplitudes were not altered before computing the NRMSE. The green star, yellow diamond, and red cross indicate the best, average, and worst multi-channel reconstruction from (a). Note that the best and the worst missing trace reconstructions do not not have smallest and largest full wavefield NRMSE.



(a) (b)

Figure 5.13 Normalized root-mean-square error computed from (a) only the reconstructed missing traces and (b) full reconstructed wavefield including the originally present traces. Amplitudes were range compressed before computing the NRMSE. Green star, yellow diamond, and red cross indicate the best, average, and the worst multi-channel reconstruction from Figure 5.12(a). Note that the range of NRMSE after range compression is smaller than for unaltered wavefields.

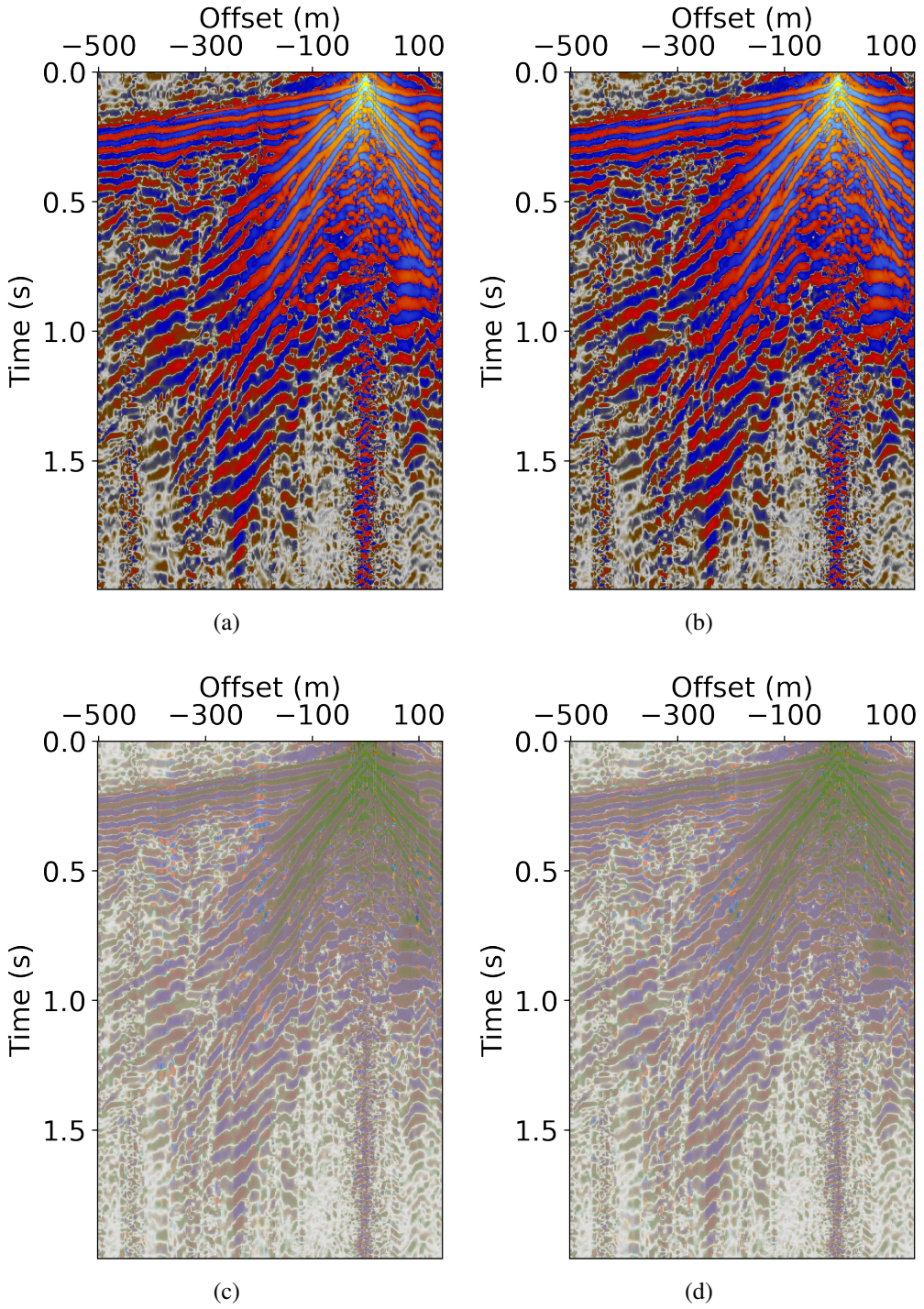


Figure 5.14 (a) Single- and (b) multi-channel reconstructions for a realization with the smallest missing trace NRMSE (Figure 5.12(a)). (c) and (d) are overlays of the respective reconstructions on the data from Figure 5.11(a). Both reconstructions are comparable, with the multi-channel approach offering a slight improvement at the near offsets.

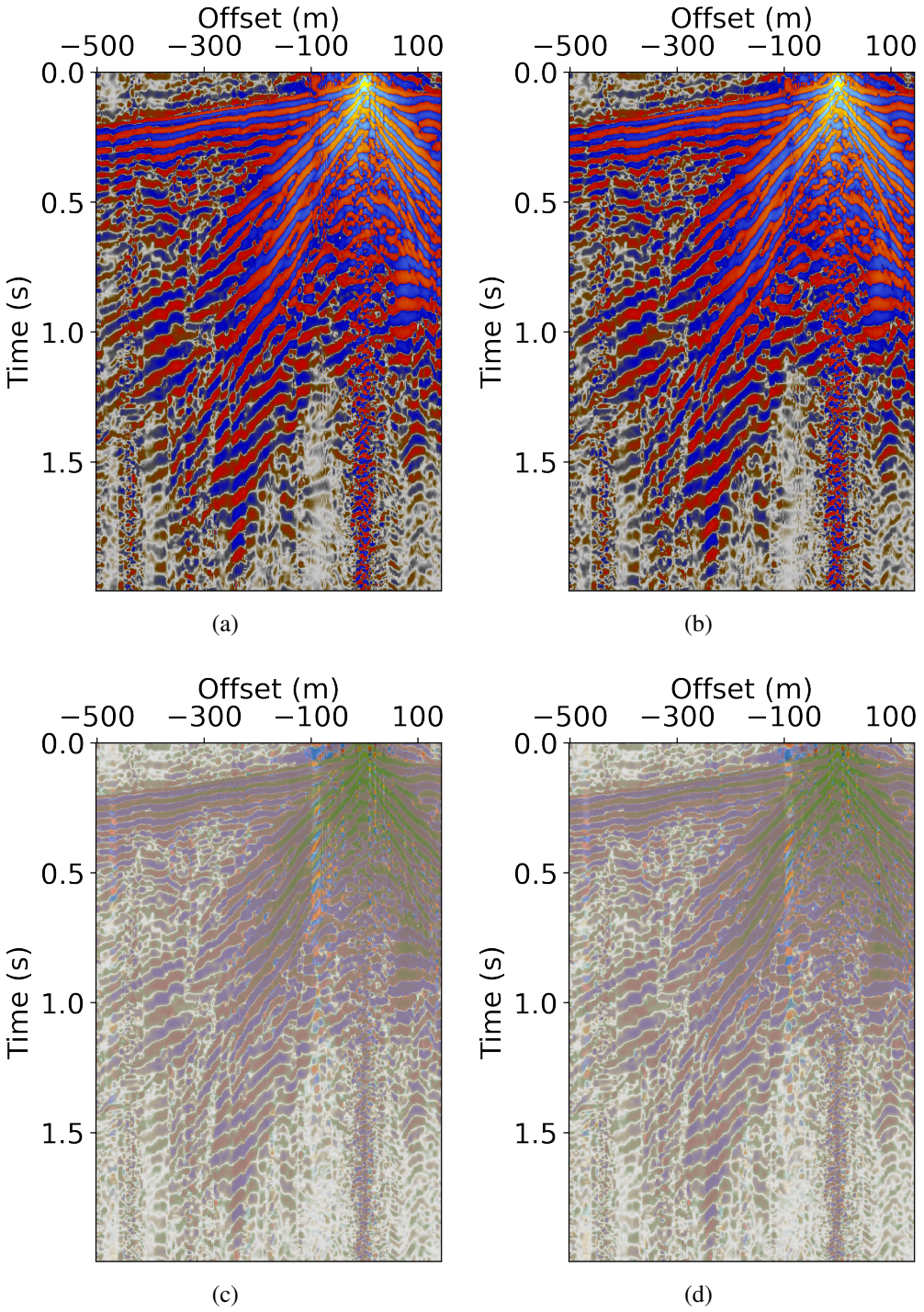


Figure 5.15 (a) Single- and (b) multi-channel reconstructions for a realization with the largest missing trace NRMSE (Figure 5.12(a)). (c) and (d) are overlays of the respective reconstructions on the data from Figure 5.11(a). Despite the overall poor NRMSE, both reconstructions are mostly accurate in terms of phase, with the most significant errors within large continuous gaps.

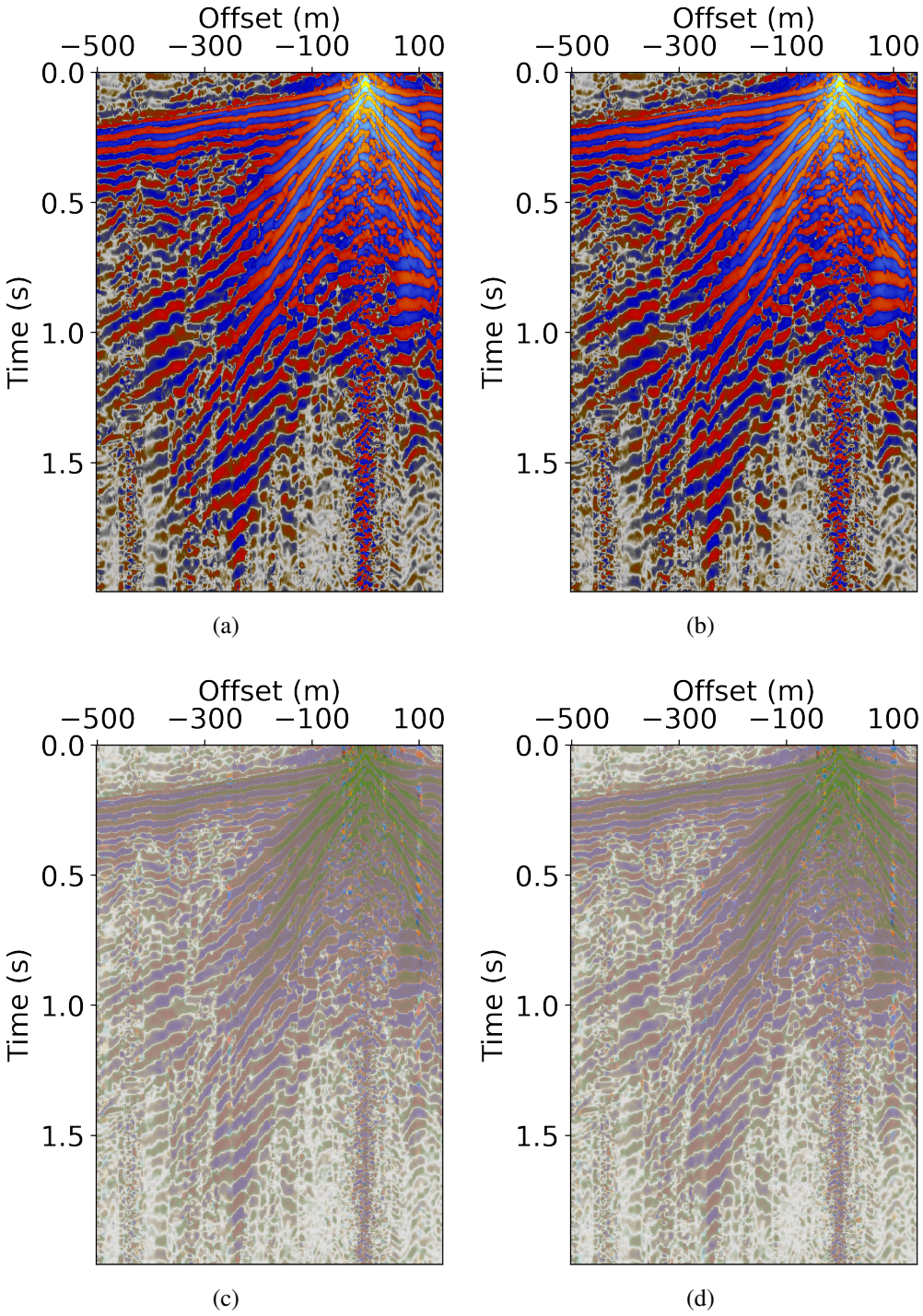


Figure 5.16 (a) Single- and (b) multi-channel reconstructions for a realization with the smallest range compressed missing trace NRMSE (Figure 5.13(a)). (c) and (d) are overlays of the respective reconstructions on the data from Figure 5.11(a). Range compression partially mitigates the NRMSE bias towards high amplitudes and reveals that both approaches perform well at farther offsets and struggle with accurate amplitude recovery at near offsets.

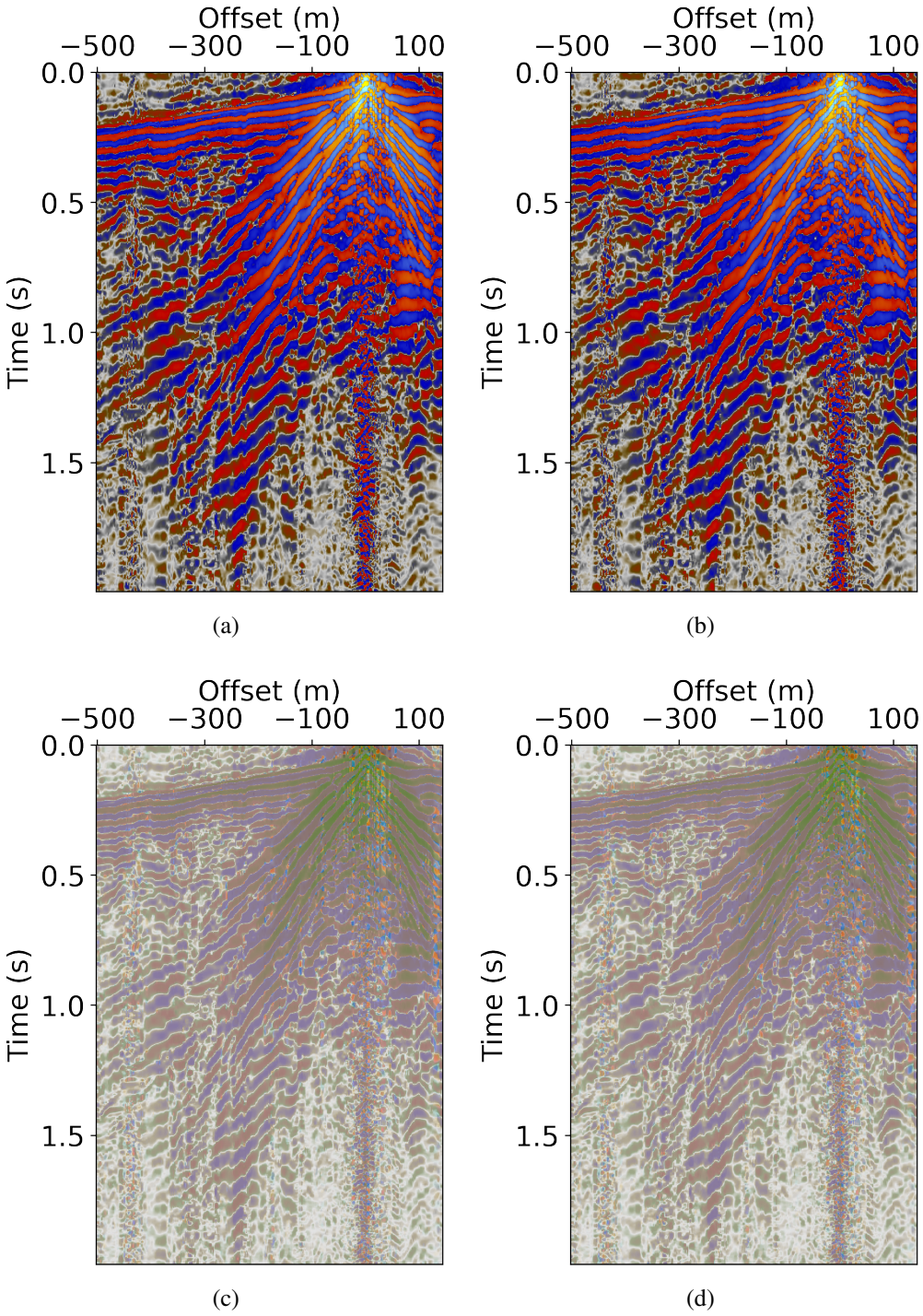


Figure 5.17 (a) Single- and (b) multi-channel reconstructions for a realization with the largest range compressed missing trace NRMSE (Figure 5.13(a)). (c) and (d) are overlays of the respective reconstructions on the data from Figure 5.11(a). Both reconstructions struggle with the significant number of missing traces at near offsets and early times, introducing notable phase and amplitude errors.

5.5 Discussion and Conclusions

We show that using derivative information in the multi-channel compressive sensing with joint sparse support constraint can help to recover a signal for which single-channel reconstruction struggles. Our synthetic 1D experiment reveals that, on average, the addition of extra bit of information helps. Perhaps somewhat against the intuition, there are also cases where the additional channel steers the reconstruction in the wrong direction, ultimately leading to its failure or poor performance. Though one might be tempted to explain this phenomenon on the grounds of stronger noise present in the field data derivative channel, the same is the case for the ideal synthetic signals. The underlying cause is the probabilistic nature of the compressive sensing experiment with random sampling following uniform distribution. As shown in Figure 5.1, there is the transition zone between the area of certain success and certain failure. For data of a given sparsity or compressibility level, the success is only certain for some subsampling ratios but once it enters the transition zone, it is the specific realization of sampling geometry that decides the outcome.

Although the 3D synthetic reconstruction seems promising, field data application requires the use of complex wavelets as a sparse domain, making the joint sparse support assumption a bit weaker. Adding the spatial derivative channel to aid the reconstruction does not yield appreciable uplift in the reconstruction quality. This might be attributed to noise in data and in derivative, or a by-product of our assumption about shared signal support and the use of IHT to enforce it.

Conventional image quality metrics are not an optimal choice for assessing the quality of land seismic data reconstruction. This is because land wavefields have high dynamic range, with amplitudes varying over several orders of magnitude. Inevitably, the metrics are biased towards high energy concentrations, and as a result, near offsets and noisy channels with spike-like, high energy noise are driving the reconstruction assessment. While near offsets are important in some applications, the primary focus for imaging is the reflections that are much weaker and typically require longer offsets to resolve. It is therefore important to be able to recover the amplitude and phase of the weak energy just as well as the amplitude of the high energy events. Arguably,

because the high energy tends to be concentrated in the surface waves that are only sensitive to the near-surface material properties, the weaker reflection energy is more important for depth imaging. When the reconstruction metric is strongly biased to high energy, the reconstruction that utterly fails to recover any of the weak energy can still be assessed as one of the best. At the same time, if there are amplitude and/or phase errors in near offsets, the reconstructions can be judged as poor even if the far offsets are reconstructed with high fidelity. Including only the missing traces and range compressing the data prior to metric computation can mitigate the bias to some extent, giving a more realistic view of the reconstruction.

A potential value of the multi-channel CS is that it simultaneously reconstructs both wavefield and its derivatives on the full grid which may provide interesting opportunities for noise suppression and wave mode separation. With more research to establish performance limits and best practices in field data application, the multi-channel CS using the joint sparsity is a promising technique that could bring us closer to solving the bad land data challenge.

CHAPTER 6

CONCLUSION

In this thesis, I look for practical solutions to data challenges associated with proper sampling of land seismic wavefields characterized by high dynamic range and an abundance of short wavelength features. My research led me to the discrete complex wavelet transform as a tool for sparsifying such wavefields and for targeted filtering. In Chapter 2, I establish complex wavelets as a sparsifying transform, making it possible to acquire land seismic wavefields from fewer measurements than dictated by regular sampling rules. Despite the fact that complex wavelets are less redundant and directionally selective than curvelets, their implementation via linear filter banks prevents the spread of aliasing artifacts over the entire domain, contributing to the consistently higher reconstruction SNR than for curvelet-based reconstructions.

In Chapter 3, I develop an iterative soft thresholding approach with scale- and orientation-dependent thresholds for complex wavelet domain data reconstruction to allow for reduction in the required channel count or an improvement in reconstruction quality for a given channel number. Compared to other algorithms used for seismic data reconstruction with sparsity constraints, the proposed method yields the most structurally consistent reconstructions. One of the important lessons from this chapter is understanding the effect of high dynamic range on the popular image quality metrics: a small error in reconstruction of near offsets can skew the metric such that it implies that the entire wavefield is reconstructed poorly which is at odds with human data interpretation. This perceptive dissonance is due to the fact that over 98% of a source-generated energy is contained to within 20 m of the source, and so the reconstructions at far offsets have negligible contribution towards the quality metric value. I find that computing the reconstruction quality metrics in a Fourier domain yields values that better align with a human intuition, and to improve it further while minimizing the effect of high dynamic range, one can take a logarithm of the Fourier coefficient magnitudes before computing the metric. The

log-Fourier SNR shows itself to be a good indicator of structural consistency in reconstructed wavefields.

In Chapter 4 I extend the reconstruction approach from Chapter 3 a step further by leveraging the multiscale character of the complex wavelet transform and thus allowing for interpolation beyond aliasing of regularly undersampled data, and for bridging large data gaps. This reconstruction algorithm called MIST is the most robust tool for approaching sampling-related data problems in the complex wavelet domain. With the sparsity assumption at its core, MIST can also be used in random sampling scenarios. Moreover, one can select the degree to which the available data samples are to be matched, which gives an opportunity for denoising during data reconstruction.

In Chapter 5, I combined ideas from generalized sampling expansion and compressive sensing under the umbrella of joint sparse recovery with a common sparse support constraint. Despite initial promising indications that the derivative channel can improve data reconstructions, field data application proved disappointing. However, despite negative results, this chapter highlights an important, but often overlooked aspect of compressive sensing: namely, its probabilistic nature. Drawing conclusions from a single realization can be misleading: even for a simple 1D signal made of sinusoids, the addition of derivative channel helps to recover the target signal on average, but there are instances when multi-channel approach fails while single-channel succeeds. This also applies to multiple sampling geometries within the same trace subsampling ratio. The locations of data gaps play an important role because land seismic data are non-stationary and because any transform we select to sparsify them will have certain limitations. For instance, complex wavelets struggle with reconstructing certain dipping events because coefficient energy is not equally distributed among all orientations.

In all presented examples, the success of complex wavelet reconstruction depends on the data sampling rate which affects spatial and temporal Nyquist, and on the number of available samples which affects the possible decomposition levels and thus the transform resolution. I observed that reconstructions tend to be better when data have similar normalized bandwidth along space and

time, i.e., when all data dimensions contain signals up to a similar fraction of their respective Nyquist. On the other hand, for the velocity filtering technique proposed in Chapter 4, the role of sampling is a little bit different. As long as signals are not aliased and follow trajectories specified by particular moveouts, it is possible to define an effective filter for their suppression. However, data sampling rates affect which scales and orientations of $\mathbb{C}$WT contain the combination of frequencies and wavenumbers that we seek to suppress. I find that with the knowledge of that mapping the filter design is straightforward and the filtering outcomes favorable.

An interesting area for future research is extending the complex wavelet transform to five dimensions and evaluating seismic data reconstruction and filtering in that extended domain. Wai Lam Chan et al. (2004) provide a recipe for constructing $m-$dimensional directional wavelets within a generalization of $\mathbb{C}$WT called hypercomplex wavelet transform. We can use that recipe to test different wavelet designs to select the one is best suited for representing land wavefields and thus achieve better processing outcomes than the current state-of-the-art 5D framework in the Fourier domain (Trad, 2009).

Complex wavelets could also prove valuable in tracking time-lapse changes. Because coefficient phase is an almost linear function of signal shifts, one can use that information for estimating motion between two images or between grids, a problem known in signal processing community as image registration. I think there is a potential to use wavelets for monitoring the movement of CO_2 through formations and reservoir depletion.

And what does the future hold for the continued improvement of land seismic data quality? The growing availability of high channel count systems opens up opportunities for experimentation - what is the smallest wavelength within the bandwidth of the active source-generated signal that could be used for subsurface characterization? The ultra-dense acquisition experiments like the one Colorado School of Mines sometimes conducts at geophysical field camps (pushing sampling to as low as 1.25 m for vibroseis data) can further our understanding of wave phenomena at the scale that was previously not possible. I believe that insights from such experiments will shape how we design future acquisitions, in which I expect

CS to play a significant part. Because no modeling is capable of reproducing the level of intricacy we are observing in the ultra-dense recordings, experimental field data are invaluable source of insight. In fact, that inability to reproduce accurate physics of wave propagation and create sufficiently accurate starting models is another reason why land data do not provide the same kind of high-resolution imagery that we came to expect from marine settings. It will be a joint effort of fine-tuning the acquisition and developing the imaging tool that will eventually let us bridge the gap between marine and land seismic imaging.

REFERENCES

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

Alkatheeri, T., G. Cambois, M. Mahgoub, G. Nyein, P. Vasilyev, I. Mihaljevic, and M. Doheim, 2020, Processing ultra–high density 3D seismic surveys: an Abu Dhabi case study: Abu Dhabi International Petroleum Exhibition & Conference, SPE, 1–10.

Allegar, N., F. J. Herrmann, and C. C. Mosher, 2017, Introduction to this special section: Impact of compressive sensing on seismic data acquisition and processing: Leading Edge, **36**, 640.

Alliney, S., and S. Ruzinsky, 1994, An algorithm for the minimization of mixed l1 and l2 norms with application to Bayesian estimation: IEEE Transactions on Signal Processing, **42**, 618–627.

Bagaini, C., M. Laycock, C. Readman, E. Coste, and C. Anderson, 2014, Seismo-acoustic characterization of a seismic vibrator: Society of Exploration Geophysicists International Exposition and 84th Annual Meeting SEG 2014, 3877–3881.

Bakulin, A., I. Silvestrov, and R. Pevzner, 2020, Surface seismics with DAS: An emerging alternative to modern point-sensor acquisition: The Leading Edge, **39**, 808–818.

Barajas-Olalde, C., and A. Jeffreys, 2014, Seismic wind noise experiments using a portable wind tunnel: 76th EAGE Conference and Exhibition 2014, EAGE, 16–19.

Baraniuk, R. G., 2007, Compressive Sensing [Lecture Notes]: IEEE Signal Processing Magazine, **24**, 118–121.

Baraniuk, R. G., V. Cevher, M. F. Duarte, and C. Hegde, 2010, Model-based compressive sensing: IEEE Transactions on Information Theory, **56**, 1982–2001.

Baraniuk, R. G., and P. Steeghs, 2017, Compressive sensing: A new approach to seismic data acquisition: The Leading Edge, **36**, 642–645.

Beck, A., and M. Teboulle, 2009, A fast iterative shrinkage-thresholding algorithm for linear inverse problems: SIAM Journal on Imaging Sciences, **2**, 183–202.

Bioucas-Dias, J., and M. Figueiredo, 2007, A new TwIST: two-step iterative shrinkage/thresholding algorithms for image restoration: IEEE Transactions on Image Processing, **16**, 2992–3004.

Blumensath, T., and M. E. Davies, 2008, Iterative Thresholding for Sparse Approximations: Journal of Fourier Analysis and Applications, **14**, 629–654.

Blymyer, D., K. Koster, and G. Warren, 2021, A 108km2 Compressive Sensing Processing Trial: First International Meeting for Applied Geoscience & Energy Expanded Abstracts, Society of Exploration Geophysicists, 2630–2634.

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

Candès, E., and J. Romberg, 2007, Sparsity and incoherence in compressive sampling: Inverse Problems, **23**, 969–985.

Candès, E. J., 2006, Compressive sampling: Proceedings of the International Congress of Mathematicians, Madrid, Spain, 1433–1452.

Candès, E. J., and L. Demanet, 2005, The curvelet representation of wave propagators is optimally sparse: Communications on Pure and Applied Mathematics, **58**, 1472–1528.

Candès, E. J., J. Romberg, and T. Tao, 2006b, Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information: IEEE Transactions on Information Theory, **52**, 489–509.

Candès, E. J., J. K. Romberg, and T. Tao, 2006c, Stable signal recovery from incomplete and inaccurate measurements: Communications on Pure and Applied Mathematics, **59**, 1207–1223.

Candès, E. J., and T. Tao, 2006, Near-optimal signal recovery from random projections: universal encoding strategies?: IEEE Transactions on Information Theory, **52**, 5406–5425.

Candès, E. J., and M. B. Wakin, 2008, An Introduction To Compressive Sampling: IEEE Signal Processing Magazine, **25**, 21–30.

Chambolle, A., R. De Vore, Nam-Yong Lee, and B. Lucier, 1998, Nonlinear wavelet image processing: variational problems, compression, and noise removal through wavelet shrinkage: IEEE Transactions on Image Processing, **7**, 319–335.

Chang, S., Bin Yu, and M. Vetterli, 2000, Adaptive wavelet thresholding for image denoising and compression: IEEE Transactions on Image Processing, **9**, 1532–1546.

Chen, H., and N. Kingsbury, 2012, Efficient registration of nonrigid 3-D bodies: IEEE Transactions on Image Processing, **21**, 262–272.

Chen, W., D. Wipf, Y. Wang, Y. Liu, and I. J. Wassell, 2016a, Simultaneous Bayesian sparse approximation with structured sparse models: IEEE Transactions on Signal Processing, **64**, 6145–6159.

Chen, Y., D. Zhang, S. Zu, S. Gan, W. Huang, X. Chen, and Z. Jin, 2016b, Simultaneous denoising and reconstruction of 5-D seismic data via damped rank-reduction method: Geophysical Journal International, **206**, 1695–1717.

Cheung, K. F., 1993, A multidimensional extension of Papoulis' generalized sampling expansion with the application in minimum density sampling: Advanced topics in Shannon sampling and interpolation theory, 85–119.

Chitchian, S., M. A. Fiddy, and N. M. Fried, 2009, Denoising during optical coherence tomography of the prostate nerves via wavelet shrinkage using dual-tree complex wavelet transform: Journal of Biomedical Optics, **14**, 014031.

Claerbout, J., and S. Fomel, 2014, Geophysical image estimation by example: Lulu.com.

Crawley, S., R. Clapp, and J. Claerbout, 1999, Interpolation with smoothly nonstationary prediction-error filters, *in* SEG Technical Program Expanded Abstracts 1999: Society of Exploration Geophysicists, 1154–1157.

Daubechies, I., M. Defrise, and C. De Mol, 2004, An iterative thresholding algorithm for linear inverse problems with a sparsity constraint: Communications on Pure and Applied Mathematics, **57**, 1413–1457.

Denisov, M. S., A. Egorov, E. A. Kurin, and M. B. Shneerson, 2019, Vibroseis Harmonic Noise Elimination Based on Optimized Recursive Filtering: 81st EAGE Conference and Exhibition 2019, European Association of Geoscientists & Engineers, 1–5.

Donoho, D., 1995, De-noising by soft-thresholding: IEEE Transactions on Information Theory, **41**, 613–627.

Donoho, D. L., 2006, Compressed sensing: IEEE Transactions on Information Theory, **52**, 1289–1306.

Donoho, D. L., I. M. Johnstone, G. Kerkyacharian, and D. Picard, 1995, Wavelet Shrinkage: Asymptopia?: Journal of the Royal Statistical Society: Series B (Methodological), **57**, 301–337.

Douma, H., and M. V. de Hoop, 2007, Leading-order seismic imaging using curvelets: Geophysics, **72**, S231–S248.

Duarte, M. F., S. Sarvotham, M. B. Wakin, D. Baron, and R. G. Baraniuk, 2005, Joint sparsity models for distributed compressed sensing: Proceedings of the Workshop on Signal Processing with Adaptative Sparse Structured Representations, 1–4.

Duijndam, A. J. W., M. A. Schonewille, and C. O. H. Hindriks, 1999a, Reconstruction of band-limited signals, irregularly sampled along one spatial direction: Geophysics, **64**, 524–538.

———, 1999b, Reconstruction of band-limited signals, irregularly sampled along one spatial direction: Geophysics, **64**, 524–538.

Dybing, S. N., A. T. Ringler, D. C. Wilson, and R. E. Anthony, 2019, Characteristics and spatial variability of wind noise on near-surface broadband seismometers: Bulletin of the Seismological Society of America, **109**, 1082–1098.

Eskicioglu, A. M., and P. S. Fisher, 1995, Image quality measures and their performance: IEEE Transactions on Communications, **43**, 2959–2965.

Figueiredo, M., and R. Nowak, 2003, An EM algorithm for wavelet-based image restoration: IEEE Transactions on Image Processing, **12**, 906–916.

Figueiredo, M. A., R. D. Nowak, and S. J. Wright, 2007, Gradient projection for sparse reconstruction: application to compressed sensing and other inverse problems: IEEE Journal of selected topics in signal processing, **1**, 586–597.

Fomel, S., and J. F. Claerbout, 2003, Multidimensional recursive filter preconditioning in geophysical estimation problems: Geophysics, **68**, 577–588.

Fomel, S., and Y. Liu, 2010, Seislet transform and seislet frame: Geophysics, **75**, V25–V38.

Foti, S., C. Lai, G. J. Rix, and C. Strobbia, 2014, Surface wave methods for near-surface site characterization: CRC Press.

Freed, D., 2008, Cable-free nodes: The next generation land seismic system: The Leading Edge, **27**, 878–881.

Gabor, D., 1946, A theory of communication. Part 1: The analysis of information.: Journal of the Institution of Electrical Engineers-part III: radio and communication engineering, **93**, 429–441.

Gan, S., S. Wang, Y. Chen, Y. Zhang, and Z. Jin, 2015, Dealiased seismic data interpolation using seislet transform with low-frequency constraint: IEEE Geoscience and Remote Sensing Letters, **12**, 2150–2154.

Gao, J., M. D. Sacchi, and X. Chen, 2013a, A fast reduced-rank interpolation method for prestack seismic volumes that depend on four spatial dimensions: Geophysics, **78**, V21–V30.

Gao, J., A. Stanton, M. Naghizadeh, M. D. Sacchi, and X. Chen, 2013b, Convergence improvement and noise attenuation considerations for beyond alias projection onto convex sets reconstruction: Geophysical Prospecting, **61**, 138–151.

Gorodnitsky, I. F., and B. D. Rao, 1997, Sparse signal reconstruction from limited data using FOCUSS: A re-weighted minimum norm algorithm: IEEE Transactions on Signal Processing, **45**, 600–616.

Guitton, A., and J. Claerbout, 2010, An algorithm for interpolation in the pyramid domain: Geophysical Prospecting, **58**, 965–976.

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

Hennenfent, G., L. Fenelon, and F. J. Herrmann, 2010, Nonequispaced curvelet transform for seismic data reconstruction: A sparsity-promoting approach: Geophysics, **75**, WB203–WB210.

Hennenfent, G., and F. J. Herrmann, 2008, Simply denoise: Wavefield reconstruction via jittered undersampling: Geophysics, **73**, V19–V28.

Herrmann, F. J., 2010, Randomized sampling and sparsity: Getting more information from fewer samples: Geophysics, **75**, WB173–WB187.

Hollander, Y., and O. Yilmaz, 2019, An acceleration method for the anti-leakage parabolic Radon transform for seismic data interpolation, *in* SEG Technical Program Expanded Abstracts 2019: Society of Exploration Geophysicists, 4480–4484.

Hunter, J. D., 2007, Matplotlib: A 2D Graphics Environment: Computing in Science & Engineering, **9**, 90–95.

Jack, I., 2003, Land seismic technology: Where do we go from here?: First Break, **21**, 33–36.

Jerri, A., 1977, The Shannon sampling theorem—Its various extensions and applications: A tutorial review: Proceedings of the IEEE, **65**, 1565–1596.

Jia, Y., and J. Ma, 2017, What can machine learning do for seismic data processing? An interpolation application: Geophysics, **82**, V163–V177.

Jia, Y., S. Yu, and J. Ma, 2018, Intelligent interpolation by Monte Carlo machine learning: Geophysics, **83**, V83–V97.

Jiang, T., P. Eick, Y. Jiang, T. Li, H. Hao, W. Chu, R. Holt, D. Blymyer, K. Koster, and D. Enns, 2019, Compressive sensing seismic processing tests on a high density blended land data set: SEG Technical Program Expanded Abstracts 2019, Society of Exploration Geophysicists, 4505–4509.

Jiang, T., Y. Jiang, D. Clark, R. Gray, and D. Brost, 2018, A compressive seismic field trial and reconstruction test using regular indexing: SEG Technical Program Expanded Abstracts, Society of Exploration Geophysicists, 86–90.

Julayusefi, M., A. R. Goudarzi, R. H. Bovanloo, and M. Shamounzadeh, 2012, Application of the 2-D dual-tree CWT as an image processing technique to attenuate remnant random noise of GPR signals.: International Geophysical Conference and Oil & Gas Exhibition, Istanbul, Turkey, 17-19 September 2012, Society of Exploration Geophysicists and The Chamber of Geophysical Engineers of Turkey, 1–4.

Kabir, M. M. N., and D. J. Verschuur, 1995, Restoration of missing offsets by parabolic Radon transform: Geophysical Prospecting, **43**, 347–368.

Keho, T. H., and P. G. Kelamis, 2012, Focus on land seismic technology: The near-surface challenge: The Leading Edge, **31**, 62–68.

Kingsbury, N., 1999, Image processing with complex wavelets: Philosophical Transactions of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences, **357**, 2543–2560.

———, 2001, Complex wavelets for shift invariant analysis and filtering of signals: Applied and Computational Harmonic Analysis, **10**, 234–253.

———, 2006, Rotation-invariant local feature matching with complex wavelets: European Signal Processing Conference.

Kreimer, N., and M. D. Sacchi, 2012, A tensor higher-order singular value decomposition for prestack seismic data noise reduction and interpolation: Geophysics, **77**, V113–V122.

Kreimer, N., A. Stanton, and M. D. Sacchi, 2013, Tensor completion based on nuclear norm minimization for 5D seismic data reconstruction: Geophysics, **78**, V273–V284.

Kumar, R., C. Da Silva, O. Akalin, A. Y. Aravkin, H. Mansour, B. Recht, and F. J. Herrmann, 2015, Efficient matrix completion for seismic data reconstruction: Geophysics, **80**, V97–V114.

Lee, K., Y. Bresler, and M. Junge, 2012, Subspace methods for joint sparse recovery: IEEE Transactions on Information Theory, **58**, 3613–3641.

Li, C., C. Mosher, Y. Tao, and Y. Zhang, 2017, A low-rank approximation method for ice-break noise attenuation of simultaneous vibroseis data, *in* SEG Technical Program Expanded Abstracts 2017: Society of Exploration Geophysicists, 4930–4934.

Li, C., C. C. Mosher, and S. T. Kaplan, 2012, Interpolated compressive sensing for seismic data reconstruction: SEG Technical Program Expanded Abstracts 2012, Society of Exploration Geophysicists, 1–6.

Li, H., W. Yang, and X. Yong, 2014, Sparse reconstruction of 3D stochastic trace missing seismic data: Beijing 2014 International Geophysical Conference & Exposition, Beijing, China, 21-24 April 2014, Society of Exploration Geophysicists and Chinese Petroleum Society, 288–291.

Liang, Y., Z. Wang, X. Li, Z. Gao, Y. Xu, S. Zhang, and Y. Cai, 2019, Seismic discontinuous detection using dual-tree complex wavelet transform: 81st EAGE Conference and Exhibition 2019, European Association of Geoscientists & Engineers, 1–5.

Linden, D. A., 1959, A Discussion of Sampling Theorems: Proceedings of the IRE, **47**, 1219–1226.

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

Liu, Y., and S. Fomel, 2011, Seismic data interpolation beyond aliasing using regularized nonstationary autoregression: Geophysics, **76**, V69–V77.

Liu, Y., P. Zhang, and C. Liu, 2017, Seismic data interpolation using generalised velocity-dependent seislet transform: Geophysical Prospecting, **65**, 82–93.

Lo, W. Y., and I. W. Selesnick, 2006, Wavelet-domain soft-thresholding for non-stationary noise: 2006 International Conference on Image Processing, IEEE, 1441–1444.

Ma, J., 2013, Three-dimensional irregular seismic data reconstruction via low-rank matrix completion: Geophysics, **78**, V181–V192.

Mallat, S. G., 1989a, A theory for multiresolution signal decomposition: the wavelet representation: IEEE Transactions on Pattern Analysis and Machine Intelligence, **11**, 674–693.

———, 1989b, Multiresolution approximations and wavelet orthonormal bases of Lˆ2(R): Transactions of the American Mathematical Society, **315**, 69–87.

Manning, T., D. Ablyazina, and J. Quigley, 2019, The nimble node - million-channel land recording systems have arrived: Leading Edge, **38**, 706–714.

Manning, T., C. Brooks, A. Ourabah, M. Popham, D. Ablyazina, V. Zhuzhel, E. Holst, and N. Goujon, 2018, The case for a nimble node, towards a new land seismic receiver system with unlimited channels: SEG Technical Program Expanded Abstracts 2018, Society of Exploration Geophysicists, 21–25.

Miller, M., and N. Kingsbury, 2008, Image denoising using derotated complex wavelet coefficients: IEEE Transactions on Image Processing, **17**, 1500–1511.

Monk, D. J., 2020, Survey design and seismic acquisition for land, marine, and in-between in light of new technology and techniques: Society of Exploration Geophysicists.

Morlet, J., G. Arens, E. Fourgeau, and D. Giard, 1982a, Wave propagation and sampling theory—Part II: Sampling theory and complex waves: Geophysics, **47**, 222–236.

Morlet, J., G. Arens, E. Fourgeau, and D. Glard, 1982b, Wave propagation and sampling theory—Part I: Complex signal and scattering in multilayered media: Geophysics, **47**, 203–221.

Mosher, C. C., C. Li, F. D. Janiszewski, L. S. Williams, T. C. Carey, and Y. Ji, 2017, Operational deployment of compressive sensing systems for seismic data acquisition: The Leading Edge, **36**, 661–669.

Muyzert, E., N. Allouche, P. Edme, and N. Goujon, 2019, A five component land seismic sensor for measuring lateral gradients of the wavefield: Geophysical Prospecting, **67**, 97–113.

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

Naghizadeh, M., and K. A. Innanen, 2011, Seismic data interpolation using a fast generalized Fourier transform: Geophysics, **76**, V1–V10.

Naghizadeh, M., and M. D. Sacchi, 2009, f-x adaptive seismic-trace interpolation: Geophysics, **74**, V9–V16.

———, 2010a, Beyond alias hierarchical scale curvelet interpolation of regularly and irregularly sampled seismic data: Geophysics, **75**, WB189–WB202.

———, 2010b, On sampling functions and Fourier reconstruction methods: Geophysics, **75**, WB137–WB151.

Needell, D., and J. A. Tropp, 2009, CoSaMP: Iterative signal recovery from incomplete and inaccurate samples: Applied and Computational Harmonic Analysis, **26**, 301–321.

Nørmark, E., 2011, Wind and rain induced noise on reflection seismic data: Near Surface 2011 - 17th European Meeting of Environmental and Engineering Geophysics, 12–14.

Ourabah, A., A. Crosby, C. Brooks, E. Manning, K. Lythgoe, D. Ablyazina, V. Zhuzhel, E. Holst, and T. Knutsen, 2019, A comparative field trial of a new nimble node and cabled systems in a desert environment: 81st EAGE Conference and Exhibition 2019, European Association of Geoscientists & Engineers, 1–5.

Ourabah, A., M. Grimshaw, J. Keggin, M. Kowalczyk-Kedzierska, J. Stone, E. Murray, S. Cooper, and L. Shaw, 2014, Acquiring and Imaging Ultra High Density Land Seismic Data - Practical Challenges and the Impact of Spatial Sampling: 76th European Association of Geoscientists and Engineers Conference and Exhibition 2014: Experience the Energy - Incorporating SPE EUROPEC 2014, 4060–4064.

Ourabah, A., J. Keggin, C. Brooks, D. Ellis, and J. Etgen*, 2015, Seismic acquisition, what really matters?: SEG Technical Program Expanded Abstracts 2015, Society of Exploration Geophysicists, 6–11.

Palangi, H., R. Ward, and L. Deng, 2016, Distributed compressive sensing: a deep learning approach: IEEE Transactions on Signal Processing, **64**, 4504–4518.

Papadopoulou, M., F. Da Col, B. Mi, E. Bäckström, P. Marsden, B. Brodic, A. Malehmir, and L. V. Socco, 2020, Surface-wave analysis for static corrections in mineral exploration: a case study from central Sweden: Geophysical Prospecting, **68**, 214–231.

Papoulis, A., 1977, Generalized Sampling Expansion: IEEE Transactions on Circuits and Systems, **24**, 652–654.

Parker, T., S. Shatalin, and M. Farhadiroushan, 2014, Distributed Acoustic Sensing – a new tool for seismic applications: First Break, **32**.

Pawelec, I., and P. Sava, 2020, Multiscale analysis and denoising of near offset vibroseis data: SEG Technical Program Expanded Abstracts 2020, Society of Exploration Geophysicists, 3264–3268.

Pawelec, I., M. Wakin, and P. Sava, 2021, Missing trace reconstruction for 2D land seismic data with randomized sparse sampling: Geophysics, **86**, P25–P36.

———, 2022, Multichannel compressive sensing for seismic data reconstruction using joint sparsity: Second International Meeting for Applied Geoscience & Energy, Society of Exploration Geophysicists and American Association of Petroleum Geologists, 2636–2640.

Peng, C., H. Jin, and P. Wang, 2014, Noise attenuation for multi-sensor streamer data via cooperative de-noising: Society of Exploration Geophysicists International Exposition and 84th Annual Meeting SEG 2014, 1878–1882.

Petersen, D. P., and D. Middleton, 1962, Sampling and reconstruction of wave-number-limited functions in N-dimensional euclidean spaces: Information and Control, **5**, 279–323.

Pilikos, G., and A. C. Faul, 2017, Bayesian feature learning for seismic compressive sensing and denoising: Geophysics, **82**, O91–O104.

Polom, U., 1997, Elimination of source-generated noise from correlated vibroseis data (the 'ghost-sweep' problem): Geophysical Prospecting, **45**, 571–591.

Raj, V. N. P., and T. Venkateswarlu, 2012, Denoising of medical images using dual tree complex wavelet transform: Procedia Technology, **4**, 238–244.

Regone, C., M. Fry, and J. Etgen, 2015, Dense sources vs. dense receivers in the presence of coherent noise: A land modeling study: SEG Technical Program Expanded Abstracts, **34**, 12–16.

Regone, C. J., 1997, Measurement and identification of 3-D coherent noise generated from irregular surface, *in* Carbonate seismology: Society of Exploration Geophysicists, 11, 281–306.

Rioul, O., and M. Vetterli, 1991, Wavelets and signal processing: IEEE Signal Processing Magazine, **8**, 14–38.

Robertsson, J. O. A., I. Moore, M. Vassallo, K. Özdemir, D.-j. V. Manen, and A. Özbek, 2008, On the use of multicomponent streamer recordings for reconstruction of pressure wavefields in the crossline direction: Geophysics, **73**, A45–A49.

Schimmel, M., and H. Paulssen, 1997, Noise reduction and detection of weak, coherent signals through phase-weighted stacks: Geophysical Journal International, **130**, 497–505.

Selesnick, I., 2009, A Derivation of the Soft-Thresholding Function: Polytechnic Institute of New York University (lecture notes), 1–9.

Selesnick, I., R. Baraniuk, and N. Kingsbury, 2005, The dual-tree complex wavelet transform: IEEE Signal Processing Magazine, **22**, 123–151.

Selesnick, I. W., 2002, The design of approximate Hilbert transform pairs of wavelet bases: IEEE Transactions on Signal Processing, **50**, 1144–1152.

Selesnick, I. W., and K. Y. Li, 2003, Video denoising using 2D and 3D dual-tree complex wavelet transforms: Wavelets: Applications in Signal and Image Processing X, 607.

Shannon, C. E., 1948, A mathematical theory of communication: Bell System Technical Journal, **27**, 379–423.

Sheriff, R. E., 2002, Encyclopedic dictionary of applied geophysics: Society of Exploration Geophysicists.

Socco, L. V., D. Boiero, S. Foti, M. Maraschini, C. Piatti, P. Bergamo, F. Garofalo, M. Pastori, and G. D. Molino, 2010a, Surface wave analysis for S-wave static correction computation: Society of Exploration Geophysicists International Exposition and 80th Annual Meeting 2010, SEG 2010, 1929–1933.

Socco, L. V., S. Foti, and D. Boiero, 2010b, Surface-wave analysis for building near-surface velocity models—Established approaches and new perspectives: Geophysics, **75**, 83–75.

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

Stork, C., 2020, How does the thin near surface of the earth produce 10–100 times more noise on land seismic data than on marine data?: First Break, **38**, 67–75.

———, 2023, Using 'Noise-Corrected OVT Fold' to optimise land seismic acquisition risk, costs, and quality: First Break, **41**, 45–49.

Strobbia, C., L. Andreas, P. Vermeer, and A. Glushchenko, 2011, Surface waves: use them then lose them. Surface-wave analysis, inversion and attenuation in land reflection seismic surveying: Near Surface Geophysics, **9**, 503–513.

Taubman, D. S., 2002, JPEG2000: Image compression fundamentals, standards and practice: Journal of Electronic Imaging, **11**, 286.

Tibshirani, R., 1996, Regression shrinkage and selection via the lasso: Journal of the Royal Statistical Society: Series B (Methodological), **58**, 267–288.

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

Trad, D., J. Deere, and S. Cheadle, 2005, Challenges for land data interpolation: CSEG Annual Convention, Expanded Abstracts, 309–311.

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

Tropp, J. A., 2004, Greed is good: algorithmic results for sparse approximation: IEEE Transactions on Information Theory, **50**, 2231–2242.

———, 2006, Algorithms for simultaneous sparse approximation. Part II: Convex relaxation: Signal Processing, **86**, 589–602.

Tropp, J. A., and A. C. Gilbert, 2007, Signal recovery from random measurements via orthogonal matching pursuit: IEEE Transactions on Information Theory, **53**, 4655–4666.

Tropp, J. A., A. C. Gilbert, and M. J. Strauss, 2006, Algorithms for simultaneous sparse approximation. Part I: Greedy pursuit: Signal Processing, **86**, 572–588.

van den Berg, E., and M. P. Friedlander, 2009, Probing the Pareto frontier for basis pursuit solutions: SIAM Journal on Scientific Computing, **31**, 890–912.

———, 2011, Sparse Optimization with Least-Squares Constraints: SIAM Journal on Optimization, **21**, 1201–1229.

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

Wai Lam Chan, Hyeokho Choi, and R. Baraniuk, 2004, Directional hypercomplex wavelets for multidimensional signal analysis and processing: 2004 IEEE International Conference on Acoustics, Speech, and Signal Processing, IEEE, 996–9.

Wang, B., N. Zhang, W. Lu, and J. Wang, 2019, Deep-learning-based seismic data interpolation: a preliminary result: Geophysics, **84**, V11–V20.

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

Wipf, D. P., and B. D. Rao, 2007, An empirical Bayesian strategy for solving the simultaneous sparse approximation problem: IEEE Transactions on Signal Processing, **55**, 3704–3716.

Xu, S., Y. Zhang, and G. Lambaré, 2010, Antileakage Fourier transform for seismic data regularization in higher dimensions: Geophysics, **75**, WB113–WB120.

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

Yanchak, D., D. Monk, and J. Versfelt, 2018, Egypt West Kalabsha 3D broadband ultrahigh density seismic survey: SEG Technical Program Expanded Abstracts 2018, Society of Exploration Geophysicists, 4080–4084.

Yu, Z., R. Abma, J. Etgen, and C. Sullivan, 2017, Attenuation of noise and simultaneous source interference using wavelet denoising: Geophysics, **82**, V179–V190.

Yu, Z., J. Ferguson, G. McMechan, and P. Anno, 2007, Wavelet-Radon domain dealiasing and interpolation of seismic data: Geophysics, **72**, V41–V49.

Yu, Z., Z. Liu, G. Zhan, P. Cheng, B. Wang, and W. Fei, 2022, DAS noise attenuation using wavelet stack: Second International Meeting for Applied Geoscience & Energy, Society of Exploration Geophysicists and American Association of Petroleum Geologists, 592–596.

Yu, Z., and D. Whitcombe, 2008, Seismic noise attenuation using 2D complex wavelet transform: 70th EAGE Conference and Exhibition incorporating SPE EUROPEC 2008, European Association of Geoscientists & Engineers, 2050–2054.

Zwartjes, P. M., and M. D. Sacchi, 2007, Fourier reconstruction of nonuniformly sampled, aliased seismic data: Geophysics, **72**, V21–V32.

APPENDIX A

DERIVING 2D GSE

N-D signal can be represented by samples of itself and its filtered versions. The interpolation formula is

$$f(\mathbf{x}) = \sum_{i=0}^{L-1} \sum_{\mathbf{n}} g_i(\mathbf{V}_g \mathbf{n}) y_i(\mathbf{x} - \mathbf{V}_g \mathbf{n}). \tag{A.1}$$

L is the number of linear systems. To reduce the sampling of N-dimensional system to $\frac{1}{m}$ Nyquist, $L = m^N$ linear filters are needed. g_i are sample values. y_i are the interpolation functions corresponding to the respective g_i. $\mathbf{V}$ is the ND sampling martix whose columns correspond to sampling vectors in i-th direction. $\|\mathbf{v}_i\|$ is the sampling interval in the i-th direction.

$$\langle \mathbf{v}_j, \mathbf{u}_k \rangle = 2\pi\delta_{jk} \tag{A.2}$$

Let us define a 2D signal in its native domain, with sampling on rectangular grid such that $\Delta x = \frac{1}{2k_x}$ and $\Delta y = \frac{1}{2k_y}$. Then we have

$$\mathbf{V} = \begin{bmatrix} \frac{1}{2k_x} & 0 \\ 0 & \frac{1}{2k_y} \end{bmatrix} \text{ and } \mathbf{U} = \begin{bmatrix} 4\pi k_x & 0 \\ 0 & 4\pi k_y \end{bmatrix}. \tag{A.3}$$

Set $m = 2$. Then the sampling and periodicity matrices become:

$$\mathbf{V_g} = \begin{bmatrix} \frac{1}{k_x} & 0 \\ 0 & \frac{1}{k_y} \end{bmatrix} \text{ and } \mathbf{U_g} = \begin{bmatrix} 2\pi k_x & 0 \\ 0 & 2\pi k_y \end{bmatrix}. \tag{A.4}$$

Now, we need $L = 4$ linear systems to generate 4 sample sets. Let them be:

$$H_0(2\pi v_x, 2\pi v_y) = 1 \tag{A.5}$$
$$H_1(2\pi v_x, 2\pi v_y) = j2\pi v_x \tag{A.6}$$
$$H_2(2\pi v_x, 2\pi v_y) = j2\pi v_y \tag{A.7}$$
$$H_3(2\pi v_x, 2\pi v_y) = -2\pi v_x 2\pi v_y. \tag{A.8}$$

Sampling density D_g of each sample set is $det(\mathbf{U_g}) = 4\pi^2 k_x k_y$. To obtain interpolation functions y_i, we need to solve the following system:

$$\mathbf{H}^T\mathbf{y} = \mathbf{e},$$

where $H_{k,i} = H_k(\boldsymbol{\omega} + \mathbf{U_g}\mathbf{q}_i)$, with $\mathbf{q}_i$ being $k-$ary representation of integer i, and with the carrier vector entries $e_i = \mathrm{e}^{j\mathbf{q}_i^T\mathbf{U_g}^T\mathbf{x}}$. The interpolation functions are then obtained by

$$y_i(\mathbf{x}) = \frac{1}{D_g}\int_{\mathcal{C}_{g0}} Y_i(\boldsymbol{\omega}, \mathbf{x})\mathrm{e}^{j\boldsymbol{\omega}^T\mathbf{x}}d\boldsymbol{\omega}. \tag{A.9}$$

$$(\mathbf{H}^T)^{-1} = \frac{1}{4\pi^2 k_x k_y}\begin{bmatrix} 4\pi^2(v_x+k_x)(v_y+k_y) & -4\pi^2 v_x(k_y+v_y) & -4\pi^2 v_y(k_x+v_x) & 4\pi^2 v_x v_y \\ j2\pi(k_y+v_y) & -j2\pi(k_y+v_y) & -j2\pi v_y & j2\pi v_y \\ j2\pi(k_x+v_x) & -j2\pi v_x & -j2\pi(k_x+v_x) & j2\pi v_x \\ -1 & 1 & 1 & -1 \end{bmatrix}$$

The carrier vector in this instance is

$$\mathbf{e} = \begin{bmatrix} 1 & e^{j2\pi k_x x} & e^{j2\pi k_y y} & e^{j2\pi(k_x x + k_y y)} \end{bmatrix}^T \tag{A.10}$$

Thus, we can compute Fourier domain representation of interpolation functions $\mathbf{y}$:

$$\begin{aligned}
Y_0 &= \frac{1}{4\pi^2 k_x k_y}\Big(4\pi^2(v_x+k_x)(v_y+k_y) - 4\pi^2 e^{j2\pi k_x x}v_x(k_y+v_y) - 4\pi^2 e^{j2\pi k_y y}v_y(k_x+v_x) \\
&\quad +4\pi^2 e^{j2\pi(k_x x+k_y y)}v_x v_y\Big) \\
Y_1 &= \frac{j}{4\pi^2 k_x k_y}\Big(2\pi(k_y+v_y) - 2\pi(k_y+v_y)e^{j2\pi k_x x} - 2\pi v_y e^{j2\pi k_y y} + 2\pi v_y e^{j2\pi(k_x x+k_y y)}\Big) \\
Y_2 &= \frac{j}{4\pi^2 k_x k_y}\Big(2\pi(k_x+v_x) - 2\pi v_x e^{j2\pi k_x x} - 2\pi(k_x+v_x)e^{j2\pi k_y 2\pi v_y} + 2\pi v_x e^{j2\pi(k_x x+k_y y)}\Big) \\
Y_3 &= \frac{1}{4\pi^2 k_x k_y}\Big(-1 + e^{j2\pi k_x x} + e^{j2\pi k_y y} - e^{j2\pi(k_x x+k_y y)}\Big)
\end{aligned}$$

To obtain interpolators in the native domain, we can use equation A.9, with $\mathcal{C}_{g0}$ defined by a $[\pi k_x, \pi k_y]$ rectangle with a vertex at $[-\pi k_x, -\pi k_y]$. Since integration region is rectangular, we can use Fubini's theorem and split double integral into two cascading integrals. Following that approach, we derive the following 2D interpolation functions:

$$y_0 = \frac{\sin^2(k_x\pi x)\sin^2(k_y\pi y)}{k_x^2k_y^2\pi^4x^2y^2} = \mathrm{sinc}^2(k_xx)\mathrm{sinc}^2(k_yy) \tag{A.11}$$

$$y_1 = \frac{\sin^2(k_x\pi x)\sin^2(k_y\pi y)}{k_x^2k_y^2\pi^4x^2y^2} = \mathrm{sinc}^2(k_xx)\mathrm{sinc}^2(k_yy)x \tag{A.12}$$

$$y_2 = \frac{\sin^2(k_x\pi x)\sin^2(k_y\pi y)}{k_x^2k_y^2\pi^4x^2y^2} = \mathrm{sinc}^2(k_xx)\mathrm{sinc}^2(k_yy)y \tag{A.13}$$

$$y_3 = \frac{\sin^2(k_x\pi x)\sin^2(k_y\pi y)}{k_x^2k_y^2\pi^4x^2y^2} = \mathrm{sinc}^2(k_xx)\mathrm{sinc}^2(k_yy)xy. \tag{A.14}$$

APPENDIX B

COPYRIGHT PERMISSIONS

B.1 Society of Exploration Geophysicists

Here is the written affirmation of SEG's policy on allowing authors to reproduce their work as a part of a thesis (Figure B.1).

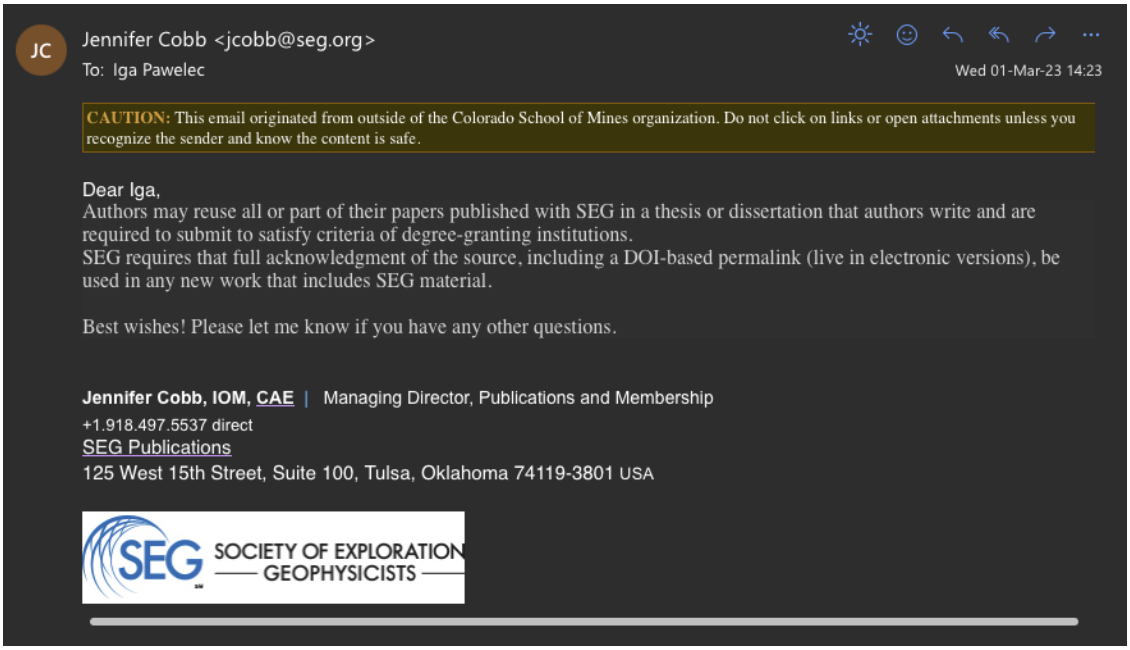


Figure B.1 SEG's permission to reproduce published work in a thesis