Automated Processing and Direct Migration  
of Ambient Seismic Data

Aaron Joseph Girard  
B.Sc. Geophysical Engineering from Colorado School of Mines  
M.Sc Applied Geophysics from IDEA-League (TU Delft; ETH Zürich; RWTH Aachen)

Centre for Energy Geoscience  
School of Earth Sciences

Supervisors: A/Prof. Jeffrey Shragge (Principal)  
Prof. Michael Dentith (Coordinating)  
Prof. David Lumley

This thesis is presented for the degree of  
Doctor of Philosophy of The University of Western Australia
Thesis Declaration

I, Aaron Joseph Girard, certify that:

This thesis has been substantially accomplished during enrolment in the degree.

This thesis does not contain material which has been accepted for the award of any other degree or diploma in my name, in any university or other tertiary institution.

No part of this work will, in the future, be used in a submission in my name, for any other degree or diploma in any university or other tertiary institution without the prior approval of The University of Western Australia and where applicable, any partner institution responsible for the joint-award of this degree.

This thesis does not contain any material previously published or written by another person, except where due reference has been made in the text.

The work(s) are not in any way a violation or infringement of any copyright, trademark, patent, or other rights whatsoever of any person.

The work described in this thesis was funded by an Australian Government Research Training Program Scholarship, an Australian Postgraduate Award (APA), a UWA Top-up Scholarship and an Australian Society of Exploration Research Foundation Grant.

This thesis contains published work and/or work prepared for publication, some of which has been co-authored.

Signature: Date: May 31, 2019
Abstract

Active-source seismic imaging using well-defined man-made energy sources is a standard Earth exploration methodology for identifying and imaging subsurface structure associated with hydrocarbon and minerals deposits. In situations where it is not feasible to utilise high-energy man-made sources for Earth exploration, though, ambient seismic energy naturally propagating in the Earth can be used as an alternative source to active-source body-wave seismic investigations. However, long recording times and significant data signal processing are required to mitigate surface-wave energy and coherent noise sources such that sufficient reflected body-wave energy can be correlated to create a stable image. Even for these scenarios, identifying and validating imaged body-wave reflection events remains challenging. I address these technical challenges in this thesis as a set of three interrelated investigations.

Study 1 develops an ambient 2D direct migration approach using a novel ambient extended imaging condition (EIC) to demonstrate the velocity sensitivity of ambient seismic recordings in the migration image domain. Tests with varying global velocity perturbations show a characteristic reflector moveout in EIC gathers that can serve as a diagnostic of reflected ambient body-wave energy. This imaging formalism, under idealised circumstances, yields comparable results to conventional seismic methods, which effectively extends the use of EIC gather-based image validation to ambient scenarios.

In real ambient seismic recordings, though, extracting body-wave arrivals remains a challenging task largely because ambient records are generally dominated by surface-wave energy and coherent noise. While the conventional cross-correlation plus stack (CC+S) methodology (i.e., interferometry) is effective, it squares the magnitude of unwanted coherent noise events (e.g., surface waves, burst-like or strong monochromatic energy) that commonly overpower ambient body-wave signal and thereby increase the challenges associated with ambient data processing. To address these issues, Study 2 develops and applies a data processing workflow to uncorrelated ambient seismic data as an alternate, though more computationally expensive method to post-CC+S processing strategies for mitigating coherent noise. The novel automated processing workflow automatically identifies and mitigates coherent noise events without severely degrading the remaining waveforms. After each processing step a number of quality control (QC) measures are applied to monitor the convergence rate of CC+S waveforms - and for evidence of emerging body-wave reflection events. The processing flow is applied to a 3D ambient seismic data set acquired at a mine site near Lalor Lake, Manitoba, Canada. The QC analyses suggest that automated
preprocessing of uncorrelated ambient seismic recordings successfully mitigates unwanted impulsive and monochromatic coherent noise events, ideally resulting in a data set sufficient for identifying reflectors with a 3D direct migration imaging approach.

While the direct migration approach applied to synthetic data in Study 1 yields desirable results, there are additional challenges associated with expanding the 2D imaging methodology to 3D field data scenarios. Study 3 begins by examining the negative effects of limited and irregular array geometry on imaging aperture and illumination. Identifying regions sub-optimally illuminated by the array reduces the risk of misidentifying or overinterpreting spuriously imaged artefacts. Next, I address challenges associated with imbalanced data spectral content in direct migration investigations by introducing a deconvolution (extended) imaging condition designed to balance wavefield spectra contributing to the final image. I then quantify whether imaged correlations represent subsurface reflections or spurious noise correlations using multi-dimensional EIC gathers in both horizontal lags and global velocity perturbations, which assist with discriminating between correlations with and without expected reflector moveout characteristics. I illustrate an ability to image subsurface correlations using the fully processed Lalor Lake data set and the 3D ambient deconvolution extended imaging condition approach. Results from 95 recording hours are consistent over hourly time intervals and reveal that many ambient imaged reflections that are spatially coincident in numerous well-illuminated locations with expected subsurface reflectors. Ambient images are compared to existing post-stack time migration results from a spatially coincident active-source 3D seismic survey, and highlight NE-dipping geologic structure associated with the Lalor Lake mineral deposit. These observations help to corroborate previously existing seismic and geological interpretations. Overall, the processing and 2D/3D imaging workflows developed in this thesis can be utilised in other ambient seismic scenarios to improve the reliability and interpretability of reflection images from ambient seismic data.
Acknowledgments

“No man is an island, entire of itself; every man is a piece of the continent.”

John Donne

This research was supported by an Australian Government Research Training Program (RTP) Scholarship. I thank The University of Western Australia (UWA) Top-up scholarship program, as well as the UWA Centre for Energy Geoscience and David Lumley. The Lalor Lake seismic data sets used in this thesis were graciously provided by Saeid Cheraghi and Gilles Bellefleur from the Geological Survey of Canada. Data processing was made possible through the reproducible open-source package Madagascar (www.reproducibility.org). Compute resources were generously provided by the Center for Wave Phenomenon at the Colorado School of Mines and the UWA Centre for Energy Geoscience.

My fellow students at the Centre for Energy Geoscience - Ben Witten, Rafael Medeiros de Souza, Thomas Hoskin and Lisa Gavin - all deserve a special mention for the valuable input and assistance and support they have given while I worked on this project. The staff at the Centre for Energy Geoscience - Toby Potter, Rie Kamei, Nader Issa, and Taka Miyoshi - were always available for advice and direction, as well as assistance with computers, equipment and simply listening to my problems, and deserve my thanks.

Staff at the Colorado School of Mines (CSM) Geophysics Department and particularly the Center for Wave Phenomenon essentially adopted me as one of their own for the last part of this project, and I am greatly appreciative for all of their help and assistance. Students at CSM - particularly Thomas Rapstine, Joseph Capriotti, Elizabeth Maag, Iga Pawelec, Vladimir Li, Colton Kohne, Tuğrul Konuk, Oscar Jarillo Michel, Can Oren, Jihyun Yang, Qifan Liu, Patipan Saengduean, Sagar Singh and Aaron Prunty - were sources of assistance and inspiration while finishing my UWA thesis physically at CSM.

My principal supervisor Jeff Shragge also deserves a special mention for assisting me with the logistical and physical aspects of moving to the other side of the planet (twice) during this project. He has also been a deep well of advice, an inspiration, and on top of that, a friend. My coordinating supervisor Mike Dentith has my thanks for picking up my project half-way through and supporting me from a world away during my thesis completion.

I also thank my parents, Joe and Audrey, for all of the assistance and support they have provided to make this thesis a reality. My brother, Mark, was also invaluable in sharing his experiences and helping me know that every PhD student meets seemingly insurmountable challenges and makes mistakes, and that that is the point: to learn how to overcome them.
Most importantly, I thank my wife, Ionela, for her monumental and unwavering support during this endeavour, and for walking with me through the many challenges this project has brought.
Authorship Declaration: Co-authored Publications

This thesis contains work that has been submitted, peer reviewed and accepted for publication.

First published manuscript

Details of the work: Manuscript 1 develops the theory for direct wavefield migration of ambient seismic recordings and applies the methodology to the Lalor Lake ambient seismic data set. It has been accepted for publication as Girard, A.J. and J. Shragge, 2019, Direct Migration of Ambient Seismic Data: Geophysical Prospecting (submitted August 2018; accepted April 2019).

Location in thesis: This appears as Chapter 2 and parts of Chapter 4 in this work.

Student contribution to work: The original idea for the manuscript was devised by myself and A/Prof. Shragge in the first months of my studies. I developed the imaging workflows and experiments and modified some existing code to generate ambient synthetic data. I processed the Lalor Lake ambient seismic data set and created images using the methodology developed for the synthetic experiments. I also wrote the manuscript which was reviewed, edited and improved upon through numerous iterations with A/Prof. Shragge.

Co-author Signature: Date: May 30, 2019

Second published manuscript

Details of the work: Manuscript 2 develops the methodology and processing workflow required to extract (reflected) body-wave energy from ambient seismic recordings. It has been accepted for publication as Girard, A.J. and J. Shragge, 2019, Automated Processing Strategies for Ambient Seismic Data: Geophysical Prospecting (submitted November 2018; accepted April 2019).

Location in thesis: This appears as Chapter 3 in this work.
Student contribution to work: The idea for this work developed over many months of discussion with A/Prof. Shragge while developing an ambient imaging framework. Through much numerical testing I found that the data had statistical properties that could be optimised and automatically utilised by the processing workflow. I developed the filters with assistance from A/Prof. Shragge and wrote the manuscript, which was again reviewed, edited and improved upon through numerous iterations with A/Prof. Shragge.

Co-author Signature: Date: May 30, 2019

I, Michael Dentith, certify that the student statements regarding their contribution to each of the works listed above are correct.

Coordinating supervisor signature: Date: June 3, 2019
# Contents

Thesis Declaration iii  
Abstract v  
Acknowledgments vii  
Authorship Declaration ix  
Contents xi  
List of Figures xv

## 1 Introduction

1.1 Styles of Seismic Data Acquisition ........................................ 3  
1.2 Acoustic Wave Propagation ............................................... 4  
1.2.1 One-way Wave-equation Approximations ............................ 6  
1.3 Active Source Imaging ..................................................... 6  
1.3.1 Imaging Condition .................................................... 8  
1.3.2 Extended Imaging Condition ......................................... 9  
1.4 Ambient Imaging ............................................................ 9  
1.4.1 Interferometry .......................................................... 11  
1.4.2 Direct Ambient Migration ............................................. 12  
1.5 Ambient Data Reality ...................................................... 14  
1.6 Research Challenges and Thesis Contributions ......................... 15  
1.6.1 Study 1: Imaging, Identifying and Validating Imaged P-wave Reflectivity ......................................................... 16  
1.6.2 Study 2: Ambient Data Conditioning Through Processing ........ 17  
1.6.3 Study 3: Application to Field Data Set in 3D ...................... 18

## 2 Direct Migration of Ambient Seismic Data: Theory 19

2.1 Introduction ................................................................. 19  
2.2 Active-Source Imaging Framework .................................... 22  
2.2.1 Acoustic Wave Propagation ......................................... 22  
2.2.2 Active-source Imaging Conditions ................................. 24
## Contents

2.3 Ambient Source Imaging Framework ........................................ 24
2.4 Image Convergence ......................................................... 26
  2.4.1 Image Variance ...................................................... 27
2.5 Experiment I: Layered Half-space Model .................................. 28
  2.5.1 2D Elastic Model .................................................... 28
  2.5.2 Data Generation ..................................................... 29
  2.5.3 Zero-offset Imaging .................................................. 30
  2.5.4 EIC Gathers ......................................................... 32
  2.5.5 Image Convergence .................................................. 37
2.6 Experiment II: Overthrust Model .......................................... 38
2.7 Discussion ................................................................. 44
2.8 Conclusions ............................................................... 45

3 Automated Processing Strategies for Ambient Seismic Data ................. 47
  3.1 Introduction ............................................................. 47
  3.2 Lalor Lake Data Set ..................................................... 51
  3.3 Extracting Body-wave Signal From Ambient Data ....................... 52
  3.4 Processing Methodology ................................................ 54
    3.4.1 Processing workflow ............................................. 54
    3.4.2 Quality Control Measures ...................................... 56
  3.5 Processing the Lalor Lake Ambient Data set ............................ 62
    3.5.1 Steps 1 and 2: Masking and Data Selection ..................... 63
    3.5.2 Step 3: Time Debursting ....................................... 65
    3.5.3 Step 4: Frequency Debursting .................................. 66
    3.5.4 Multi-hour stack ................................................. 68
  3.6 Discussion ............................................................... 70
    3.6.1 Value of Quality Control Measures .............................. 70
    3.6.2 Value of “informed” stacking ................................... 71
    3.6.3 Caveats .......................................................... 71
  3.7 Conclusions ............................................................. 72

4 Direct Migration of Ambient Seismic Data: Field Test Study ................. 75
  4.1 Introduction ............................................................. 75
  4.2 Ambient Direct Migration ............................................... 76
    4.2.1 3D Ambient Imaging Condition ................................ 76
    4.2.2 3D Ambient Extended Imaging Condition ...................... 77
    4.2.3 Ambient Deconvolution Imaging Condition .................... 77
    4.2.4 Ambient Deconvolution Extended Imaging Condition .......... 77
  4.3 Lalor Lake Dataset Postprocessing .................................... 78
  4.4 3D Field Data Imaging Challenges .................................... 79
    4.4.1 Array Geometry Effects ....................................... 80
    4.4.2 Cross-correlation versus Deconvolution Imaging Conditions . 82
CONTENTS

4.4.3 Imaging Raw Data versus Processed Data .................................. 84
4.5 Imaging Results ........................................................................... 85
  4.5.1 Imaging Single Hours of Fully Processed Data ......................... 85
  4.5.2 Image Convergence ............................................................... 90
  4.5.3 Extended Images ................................................................. 91
  4.5.4 Global Velocity Model Perturbations ..................................... 94
4.6 Imaging Multi-hour Recordings ...................................................... 98
  4.6.1 Comparing Ambient and Active-Source Images ....................... 98
4.7 Discussion ................................................................................ 105
4.8 Conclusions .............................................................................. 107

5 Conclusions .................................................................................. 109
  5.1 Summary of Conclusions .......................................................... 109
    5.1.1 Study 1: Direct Migration of Ambient Seismic Data .............. 110
    5.1.2 Study 2: Automated Processing Strategies for Ambient Seismic Data 110
    5.1.3 Study 3: Direct Migration of Ambient Seismic Field Data .......... 111
  5.2 Lessons learned ......................................................................... 112
  5.3 Future work ............................................................................... 113
    5.3.1 Application to Other 3D Ambient Seismic Data Sets ............ 113
    5.3.2 Modelling Coherent Noise ............................................... 113
    5.3.3 Multi-Channel Filtering .................................................... 114
    5.3.4 Azimuthal Weighting and/or Balancing ............................... 114
    5.3.5 Computational Runtime Improvements ............................... 115

Bibliography .................................................................................... 116
List of Figures

1.1 Cartoon showing types of seismic energy that generally propagate in the Earth. The red lines represent active-source seismic energy (here denoted as a vibroseis truck). The orange line shows an earthquake source, arriving from far outside the receiver area. Yellow stars and lines show microseismic reservoir fracturing source energy coming from a well. Green “wavefields” show man-made sources from the surface (here traffic). Blue “wavefields” show natural energy sources from the surface (here ocean waves). All sources create energy that propagates through the Earth and can be recorded by the receivers (inverted yellow triangles).

1.2 Cartoon showing the forward modelling operation for active-source experiments: a source excited at $f_s$ propagates as source wavefield $U_S$ downward into the Earth, where it diffracts off a subsurface scatterer, $p$, and propagates to the surface as receiver wavefield $U_R$ that is recorded at $f_r$ as data $d$.

1.3 Cartoon showing source and receiver imaging operations for active-source experiments. (a) The forward modelling operation injects and causally propagates the source information from $f_s$ into the subsurface creating source wavefield $U_S$, equivalent to the left hand side of Figure 1.2. (b) The single-scattered adjoint receiver wavefield, $U_R^*$, is reconstructed from data, which are propagated in reverse time into the subsurface equivalent to the time-reversed right hand side of Figure 1.2. When using the correct velocity model for propagation, both wavefields will correlate at the same time at the correct scatterer location $p$ and can be stacked to create an image through a correlation-based (or deconvolution-based) imaging condition.

1.4 Extended image gathers at the same image point over lag $\lambda_x$ for a simple active-source experiment illustrating migration velocity models that are: (a) too slow and showing a characteristic “frown”; (b) correct; and (c) too fast and showing a characteristic “smile”. Interestingly, a poorly focused S-wave centred at 2.4-2.6 s in panels (a)-(c) changes shape less but depth more than the P-wave energy as velocity is perturbed. This is expected since the P-wave velocity used for acoustic imaging of elastically forward modelled data is always too high for correctly focusing S-wave energy.
1.5 Cartoon showing the forward modelling operation for ambient seismic experiments. An upgoing ambient incident plane wave reflects off the free surface and is recorded at $f'_a$ as data $d$, before propagating as wavefield $R_+$ down to the scatterer, $p$, that generates an upgoing scattered wavefield $R_-$, which travels to the surface and is measured at $f_a$ as data $d$. 

1.6 Cartoon showing the causal and anti-causal imaging operations for ambient seismic experiments. (a) Causally propagating data recorded at $f'_a$ with the correct migration velocity creates wavefield $R_+$, which correctly models the “source-side” contribution of the ambient wavefield, but incorrectly models the “receiver-side” contributions ($R_-$) from $f_a$. (b) Time-reverse propagation of data recorded at $f_a$ reconstructs wavefield $R^\dagger_-$, which correctly models the “receiver-side” contributions, but inaccurately reconstructs the “source-side” contributions ($R^\dagger_+$) from $f'_a$. Correlating the two wavefields results in an image of the scatterer, along with the erroneous spurious correlations involving wavefields ($R_-$) and ($R^\dagger_+$) illustrated by the dashed lines [equation (1.13)].

1.7 Example traces from Lalor Lake ambient data set with different types of large-magnitude coherent noise events. (a) Stationary in time represented by a strong broadband event related to a dynamite blast (i.e., nearly simultaneously across an array). (b) Stationary in space, likely related to road/mine site activity coming from a single direction. (c) Stationary in frequency, likely related to electro-mechanical vibratory noise from on-site electrical generators. Note that the amplitude scale changes in these panels.
2.1 Cartoon showing forward modelling and adjoint imaging operations for active-source and ambient experiments. (a) Forward modelling of an injected source $S(f_s)$ at location $f_s$ creates source wavefield $U_S$ that diffracts off a subsurface scatterer at $p$, reflected toward the surface as wavefield $U_R$ and recorded as data $d(f_r)$ at location $f_r$. (b) An adjoint imaging operation causally propagates the known source information into the subsurface at $f_s$, creating source wavefield $U_S$. (c) The estimated singly scattered adjoint (receiver) wavefield, $U_R^\dagger$, created by injecting recorded data processed to be only primary reflections $d_p(f_r)$ at $f_r$, propagates in reverse time into the subsurface. When using the correct velocity, the wavefields in (b) and (c) will correlate at the scatterer $p$ to create an image. (d) An upgoing ambient incident plane wave reflects off the surface and is recorded as data, $d(f'_a)$, before propagating down to the scatterer $p$ as wavefield $R_+$, which diffracts wavefield information upward as $R_-$, to be measured as data $d(f_a)$. (e) Causally propagating $d_p(f'_a)$ creates wavefield $R_+$, which correctly models the “source-side” contribution of the ambient wavefield, but incorrectly models the “receiver-side” contributions $d_p(f_a)$, denoted as $(R_-)$. (f) Adjoint propagation of data $d_p(f_a)$ reconstructs wavefield $R_+^\dagger$, which correctly models the “receiver-side” contributions from $f_a$, but inaccurately reconstructs the “source-side” contributions $d_p(f'_a)$, denoted as $(R_+^\dagger)$. Correlating the wavefields $R_+$ in (e) and $R_+^\dagger$ in (f) results in an image of the scatterer at $p$, along with the erroneous crosstalk through correlations with wavefields $(R_-)$ and $(R_+^\dagger)$.

2.2 Two-layer half-space model with a sinusoidal density interface centred about 2 km depth. The P- and S-wave velocities are a constant $V_P=2.0$ km/s and $V_S=V_P/\sqrt{3}$ km/s, respectively, and upper and lower densities of $\rho=1.0$ g/cm$^3$ and $\rho=2.0$ g/cm$^3$, respectively. The common receivers are denoted in red at 10 m intervals, which is identical to the model spatial sampling. The sources used for the active-source experiment are denoted as blue Xs, at 40 m intervals. The 50 passive sources are denoted as yellow Xs in the lower half-space. The white lines denote the two locations (anticline, left, and syncline, right) where I extract EIC gathers.

2.3 Ambient seismic modelling results. (a) Propagated ambient wavefield snapshot with (b) corresponding $k_x$-$k_z$ spectrum. (c) Surface-recorded ambient data for the model presented in Figure 2.2 with (d) corresponding $k_x$-$\omega$ spectrum.

2.4 (a-c) Images created with active-source data and equation (2.4). (d-f) Images created with passive data and equation (2.6). (g-i) Images created with ambient data and equation (2.6). The left, centre and right columns show each experiment calculated with the P-wave velocity error ($dV$) of -5%, 0% and +5%, respectively.
2.5 2D EIC gathers for the active-source experiment extracted at the noted anticline (a-c) and syncline (d-f) locations in Figure 2.2 for dV of -5% (a and d), 0% (b and e) and +5% (c and f). ................................. 33

2.6 3D EIC gathers at the anticline location in Figure 2.2 for the (a) active-source and (b) ambient experiments. Each plane on the third axis represents a different dV ranging between ±5% error in 1% increments. ............... 34

2.7 2D EIC gathers for the ambient experiment extracted at the noted anticline (a-c) and syncline (d-f) locations in Figure 2.2 for dV of -5% (a and d), 0% (b and e) and +5% (c and f). ................................. 35

2.8 Comparison of the total stacked image using ambient recordings and the passive imaging condition after stacking: (a) four windows, (b) 16 windows, (c) 64 windows and (d) 128 windows. ................................. 36

2.9 (a) Normalised convergence curve for the ambient image with the correct velocity from Figure 2.4h. (b) Absolute value of the gradient of the normalised convergence curve. Each of the panels in Figures 2.8a–d is marked as a vertical line, showing that the RMS decreases as the image stability increases. ......................................................... 37

2.10 (a) Overthrust density model showing the receivers at every location (red dots), active source locations (blue Xs) and the two locations for EIC gather extraction (white lines), where the left and right lines are the thrust-fault and 1D geology locations, respectively. (b) Active-source and (c) ambient Overthrust zero-offset images using correct migration velocity. .......... 39

2.11 Overthrust 2D EIC gathers for the active-source experiment in the (a-c) thrust-fault [left line in Figure 2.10a] and (d-f) the 1D geology [right line in Figure 2.10a] locations for dV of -5% (a and d), 0% (b and e) and +5% (c and f). ................................. 40

2.12 Overthrust 2D EIC gathers for the ambient experiment for the (a-c) thrust-fault [left line in Figure 2.10a] and (d-f) 1D geology [right line in Figure 2.10a] locations for dV of -5% (a and d), 0% (b and e) and +5% (c and f). ....... 41

2.13 Overthrust 3D EIC gathers as in Figure 2.6. (a) Active-source and (b) ambient examples for the thrust-fault location shown in left line of Figure 2.10a. (c) Active-source and (d) ambient examples for the 1D geology location shown in right line of Figure 2.10a. ................................. 43

2.14 (a) Normalised convergence curve for the Overthrust ambient image with dV=0%. (b) Absolute value of the gradient of the normalised Overthrust convergence curve. ......................................................... 44
3.1 Map showing the ambient seismic survey area and its location in Manitoba, Canada (inset). Filled circles show geophone locations with those in yellow forming the 2D sub-array for CC+S processing described herein. The red star shows the location used to calculate power spectral density (PSD) plots. The mine site (left side) access road (south) and connecting main road (lower right) are visible on the map. ................................. 52

3.2 Workflow showing the processing sequence steps and computed QC outputs: (1) global PSD plots; (2) hourly and global CC+S panels; (3) hourly RMS convergence plots; and (4) hourly beamsteering plots. ................................. 54

3.3 PSD plots of the entire recording time for a single receiver (red star in Figure 3.1) using roughly 30 s windows (4096 samples at 8 ms sample rate) to calculate individual spectra. The plots show (a) the entire raw data set and (b) the result of applying the data selection and masking processing steps. Blanked-out intervals represent windows excluded based on the criterion of RMS energy exceeding the established threshold value. ................................. 56

3.4 PSD plots of the entire recording time without the blanked-out windows (solid blue in Figure 3.3b) for a single trace (red star in Figure 3.1) using roughly 30 s (4096 samples) windows to calculate the spectrum. (a) Raw data. (b) After data masking and selection. (c) After time debursting. (d) After frequency debursting. ................................. 57

3.5 CC+S panels from a 2D line of receivers (yellow circles in Figure 3.1) correlated with a single master trace for two different hours at each processing stage. The upper and lower rows represent quiet and noisy hours, respectively. (a and e) Raw data. (b and f) After data masking and selection. (c and g) After time debursting. (d and h) After frequency debursting. Red points denote samples clipped (high-energy) for visualisation purposes, which are included to highlight improvements from applying the processing workflow. The emergence of a prominent reflection event can be observed in (d) at ±0.15 s lags. ................................. 59
3.6 (a) RMS convergence plots for the quiet hour of CC+S panels shown in Figures 3.5a–d. As more CC+S windows are included, the RMS energy effectively decreases as inverse square of accumulated time (i.e., windows). The coloured lines denote different processing steps. Green: Raw data. Blue: After window masking and selection. Red: Time debursted. Pink: Frequency debursted. (b) Semi-log of the absolute value of the normalised convergence curve gradient, showing that the convergence rate decreases asymptotically (though with some jitter). (c) RMS convergence plots for the noisy hour of CC+S panels shown in Figures 3.5e–h. The same trend as observed in (a) is evident, though some increases occur at particular windows. More windows have been removed in the masking process, as indicated by the shorter length of the three processed curves with respect to the green line (raw data). (d) Semi-log of the absolute value of the normalised convergence curve gradient in (c), showing that the convergence rate is similar to the quieter hour, but with increased jitter. 60

3.7 Beamsteering plots of PWDS for the quiet (upper row) and noisy (lower row) hour data showing the apparent velocity (from 1.5 to 5.5 km/s) and directivity (360° of azimuth) of wavefield arrivals. (a and e) Raw data. (b and f) After masking and data selection. (c and g) After time debursting. (d and h) After frequency debursting. Red colours are higher amplitudes. Evidence for coherent high-velocity P-wave arrivals is observed in (d) and to a lesser extent in (h). 62

3.8 Illustrating the masking procedure. (a) Result of stacking envelopes (blue curve) computed for each raw data trace to calculate the 70% masking threshold (red curve). Both curves show scaled amplitudes ranging from zero (left) to one (right). (b) Result of applying the masking from where the masking operator (pink line, also valued from zero to one) calculated from the red curve in (a). 64

3.9 Frequency-domain debursting example. (a) Time-domain example trace before (blue) and after (red) debursting. (b) Partial spectrum before (blue) and after (red) debursting. While the representative trace in (a) is somewhat altered, the debursted spectrum in (b) is well matched at frequencies where magnitudes are below the user-selected threshold. 66

3.10 Averaged CC+S panels for 95 selected hours, largely falling between 1-10 hours after blasting. (a) Raw. (b) After selection and masking. (c) After time- and frequency-debursting. 68
3.11 A single extracted CC+S trace at a virtual source-receiver offset of 1.0 km for the 95 selected hours. (a) Raw data. (b) After masking and data selection. (c) After time and frequency debursting. While there is consistency across entire data set, the final processing results show greater stationarity and less ringing over full time axis. (d) 227 hours of raw data, which highlight periods of significant variability when compared to panel (a).......

4.1 Map showing the ambient seismic survey area and its location in Manitoba, Canada (inset). Filled-in circles show recording station locations. Magenta circles indicate the geophones used for imaging. Yellow circles represent geophones removed for imaging because of their distance from the denser array. Red letters and blue stars show locations selected for extended image locations shown in subsequent figures. The cyan rectangle shows the maximum migration aperture. Black stars indicate recording stations of Line 133, used for interpretation.  

4.2 Acquisition footprint effects of the full Lalor Lake acquisition geometry on the horizontal plane-wave test image. (a) 0.5 km, (b) 1.0 km. (c) 1.5 km. (d) 2.0 km. (e) 2.5 km. (f) 3.0 km. Each plane is visible in the cross-sections (g) $[E,N]=[1-4, 2.5]$ km and (h) $[E,N]=[2.5, 1-4]$ km.  

4.3 Acquisition footprint effects of the restricted Lalor Lake acquisition geometry on the horizontal plane-wave test image. (a) 0.5 km, (b) 1.0 km. (c) 1.5 km. (d) 2.0 km. (e) 2.5 km. (f) 3.0 km. Each plane is visible in the cross-sections (g) $[E,N]=[1-4, 2.5]$ km and (h) $[E,N]=[2.5, 1-4]$ km.  

4.4 Comparison of 3D direct migration image volumes of the fully processed quiet hour data with the deconvolution imaging condition from equation (4.3) for different $\epsilon^2$ values: (a) 0.1; (b) 1.0; (c) 3.7; and (d) 10.0.  

4.5 Comparison of 3D direct migration results using a cross-correlation imaging condition for both raw and fully processed quiet- and noisy-hour data. (a) Raw data image from the noisy hour. (b) Raw data image from the quiet hour. (c) Processed data image from the noisy hour. (d) Processed data image from the quiet hour.  

4.6 Comparison of 3D images with the deconvolution operator from both raw and fully processed hours from the Lalor Lake data set. (a) Raw data image from a noisy hour. (b) Raw data image from a quiet hour. (c) Processed data image from a noisy hour. (d) Processed data image from a quiet hour.  

4.7 3D images from the fully processed quiet-hour data set at four locations migrated using only the selected receivers (magenta) noted in Figure 4.1: (a) $[E,N] = [2.36, 2.54]$ km, (b) $[E,N] = [3.06, 2.66]$ km, (c) $[E,N] = [2.54, 2.54]$ km, and (d) $[E,N] = [2.28, 2.22]$ km.  

4.8 3D images from the fully processed noisy-hour data set at the same four locations as in Figure 4.7: (a) $[E,N] = [2.36, 2.54]$ km, (b) $[E,N] = [3.06, 2.66]$ km, (c) $[E,N] = [2.54, 2.54]$ km, and (d) $[E,N] = [2.28, 2.22]$ km.
4.9 Convergence curves for the deconvolution imaging condition. (a) Convergence curve for the noisy (blue) and quiet (magenta) hours. (b) Gradient of the convergence curves for the noisy (blue) and quiet (magenta) hours. . . . 90

4.10 EIC gathers created from the fully processed quiet-hour data set at locations of interest denoted as red letters in Figure 4.1. Each EIC gather has both horizontal lags $\lambda_x$ and $\lambda_y$ and shows image focusing centred at zero lag. . . 92

4.11 EIC gathers created from the fully processed noisy-hour data set at locations of interest denoted as red letters in Figure 4.1. Each EIC gather has both horizontal lags $\lambda_x$ and $\lambda_y$ and shows image focusing centred at zero lag. . . 93

4.12 EIC gathers created from the fully processed quiet hour at locations outside of the dense receiver array areas, shown as blue stars in Figure 4.1. These EIC gathers show most energy focused at non-zero lags. . . . . . . . . . . . 94

4.13 Images at the same location as Figure 4.7a using fully processed quiet-hour data and the reference velocity with (a) $-5\%$ and (b) $+5\%$ perturbations. . 95

4.14 Set of $\lambda_x$-$dV$ panels created from the fully processed quiet-hour data set at locations of interest denoted as red letters in Figure 4.1. EIC gather has both horizontal lags $\lambda_x$ and global velocity perturbation $dV$ to show how image focusing changes at zero lag as migration velocity changes. . . . . . 96

4.15 Set of $\lambda_x$-$dV$ panels created from the fully processed noisy-hour data set at locations of interest denoted as red letters in Figure 4.1. EIC gather has both horizontal lags $\lambda_x$ and global velocity perturbation $dV$ to show how image focusing changes at zero lag as migration velocity changes. . . . . . 97

4.16 3D images at location A from (a) 95 imaged and stacked hours and (b) the active-source PoSTM volume, and location B from (c) 95 imaged and stacked hours and (d) the active-source PoSTM volume. Ambient images are migrated using only the selected receivers (magenta) as noted in Figure 4.1. 99

4.17 3D images at location C from (a) 95 imaged and stacked hours and (b) the active-source PoSTM volume, and location D from (c) 95 imaged and stacked hours and (d) the active-source PoSTM volume. Ambient images are migrated using only the selected receivers (magenta) as noted in Figure 4.1. 100

4.18 Stability panel over the 95 imaged hours extracted at $[E,N] = [2.36, 2.54] \text{ km}$. This shows the continuity of the imaged location after the automated processing and a deconvolution imaging workflow. (a) Unfiltered result showing moderate differences over calendar time and (b) stacked trace of all 95 traces. (c) Filtered result to highlight continuity of highest amplitudes and remove jitter and (b) stacked trace of all 95 traces. The difference in the stack trace is negligible between the two, and the filtered version is much easier to interpret. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 101

4.19 2D annotated images extracted at $E=2.54 \text{ km}$ for (a) ambient and (b) active-source images. Annotations indicate locations of prominent reflectors. 103
4.20 2D annotated images extracted at E=2.54 km for (a) ambient and (b) active-source images. Annotations indicate locations of prominent reflectors. 103
4.21 2D annotated images extracted at N=2.78 km for (a) ambient and (b) active-source images. Annotations indicate locations of prominent reflectors. 104
4.22 2D annotated images extracted at Line 133 from the original survey grid points for the (a) ambient and (b) active-source data. (c) A geological model of Line 133 from Roots et al. (2017) for comparison. . . . . . . . . . . . . . 106
LIST OF FIGURES
Chapter 1

Introduction

“I don’t know where I am going, but I am on my way.”

Voltaire

Geophysical exploration for minerals and hydrocarbons has been undertaken for over a century. Of all the geophysical techniques available, active-source three-dimensional (3D) seismic imaging has become the method of choice due to the high-resolution capabilities, ease and speed of exploring large areas, and relative low cost of surveying per information gained. Yilmaz (2001) and references therein detail the currently recognised practice of using 3D active-source (land and marine) seismic data to image subsurface geological structure based on wave-propagation physics, data signal processing, and an ability to estimate model parameters (e.g., velocity).

There are numerous situations, though, where 3D active-source seismic acquisition becomes impractical. Reasons for this include, but are not limited to:

- **geographical**: survey locations may be extremely difficult or impossible to reach with the required equipment (e.g., mountains, swamps, arctic);

- **cultural**: landowners do not permit seismic acquisition vehicles or active dynamite sources on their premises (e.g., archaeological sites, cropland, native title land) or do not want their land adversely affected by surveying; and

- **financial**: it is prohibitively expensive to acquire 3D seismic data because of the high barrier to entry introduced by the cost of transporting the equipment and personnel required to run a large-scale survey (i.e., active-source seismic 3D surveys typically cost millions of dollars).

Based on these limitations and constraints, there is significant industry interest in pursuing other methods of seismic exploration that do not employ active sources or dense recording geometries. For these scenarios, many researchers are investigating alternative approaches for acquiring, imaging and interpreting seismic data from “non-active” sources.
Figure 1.1: Cartoon showing types of seismic energy that generally propagate in the Earth. The red lines represent active-source seismic energy (here denoted as a vibroseis truck). The orange line shows an earthquake source, arriving from far outside the receiver area. Yellow stars and lines show microseismic reservoir fracturing source energy coming from a well. Green “wavefields” show man-made sources from the surface (here traffic). Blue “wavefields” show natural energy sources from the surface (here ocean waves). All sources create energy that propagates through the Earth and can be recorded by the receivers (inverted yellow triangles).

There are a large number of mechanisms that can act as potential sources of wave energy for imaging purposes (e.g., vibroseis, dynamite, earthquakes, microseisms, ocean waves, road traffic). Figure 1.1 illustrates how some representative sources can excite energy that propagates throughout a 1D subsurface model. To the left of the figure an active source (here a vibroseis truck) excites the Earth’s surface, sending energy into the subsurface with directions represented by rays. This energy interacts with geologic discontinuities, scatters upward, and is measured at the surface by geophones (inverted yellow triangles). Similar sources also may be excited within the subsurface due to natural (earthquake, orange ray) or man-made/induced (reservoir fracing, yellow stars and rays) sources. Other sources with less easily definable characteristics are shown as natural (blue ocean waves) or man-made (green urban traffic energy) wavefields propagating as elliptical wavefronts.

Terrestrial-based sources generate two categories of propagating energy - surface waves and body waves. Surface waves (\( R \), Rayleigh and \( L \), Love waves) travel along the interface between the layered Earth and the atmosphere (i.e., at the Earth’s free surface), interact with near-surface structure (relative to their wavelength), and decay exponentially with depth. Body waves, emitted as either P-waves (\( P \), primary or pressure) or S-waves (\( S \), secondary or shear), travel through the Earth, interact with structure, and re-emerge at the surface. Thus, it is possible to consider an idealised composite recorded wavefield \( U \) to be the superposition of these wavetypes:

\[
U \simeq P + S + R + L + N,
\]  

where \( N \) is the noise component that neither necessarily obeys a wave equation (e.g.,
1.1 Styles of Seismic Data Acquisition

For seismic exploration purposes, there are numerous methods of acquiring 3D seismic data as well as a variety of imaging and inversion methodologies that utilise the recorded wavefield $U$ in different ways. Broadly speaking, these approaches can be divided into one of three classes of investigation:

- **Active Seismic:** acquiring data from a known source usually generated by an explosion (e.g., dynamite), vibroseis truck or compressed air gun. This source excites waves that propagate through a medium, scatter off of subsurface structure, and travel back up to the surface where they are recorded ideally unaliased by a detector array. This scenario is represented in Figure 1.1 as red rays emanating from a vibroseis source. The early-arrival wavefield components are dominated by P waves, which are used for conventional seismic reflection imaging through a process termed migration (Sheriff and Geldart, 1995). The excited S- and surface-wave components (for terrestrial surveys), though useful in some imaging and inversion contexts, contaminate the recorded P-waves as coherent noise (i.e., ground roll) and must be mitigated through signal processing. Noise $N$ is relatively generally quite low, except at low frequencies (i.e., $<3$ Hz) where it is difficult to generate active source energy. Active-source imaging is the most common approach in exploration and crustal seismology and uses P waves acquired in many independent and short time windows (e.g., 8-12 s).

- **Passive Seismic:** acquiring data containing wavefield information from non-surface-related natural or induced sources such as (teleseismic) earthquakes or micro-seismicity (e.g., reservoir fracturing). There is interest in measuring or estimating the source location, mechanism and signature because this information helps to better characterise subsurface dynamics (e.g., geomechanics, stress state, fracture propagation) than just structural imaging alone. This scenario is represented in Figure 1.1 as orange (earthquake source) and yellow (fracing source) rays, emanating from a number of subsurface locations, that are measured on the same receivers as the active-source acquisition scenario. Passive signals are dominated by high-energy surface-wave components; however, there are often short windows containing relatively strong body-wave energy (e.g., teleseismic or microseismic arrivals) that, when isolated, are useful for migration-style imaging and inversion of Earth properties with minor modification of standard active-source imaging techniques. Noise $N$ is generally at
low-to-moderate relative levels, and distinct source signals usually are identifiable in the data. This strategy is most common in crustal and lithospheric studies, academic seismology, as well as industry-oriented microseismic and induced seismicity investigations. These studies utilise short time windows (several to 10s of seconds) where $P$ and $S$ energy levels are high relative to $R$, $L$ and $N$, that are extracted from long-duration continuous recordings (hours or days).

- **Ambient Seismic**: acquiring data containing energy from unknown sources without requiring explicit knowledge of the source itself. These signals are usually the background Earth motion generated from the same types of mechanisms used in passive seismic methods; however, the source origin, mechanism, frequency structure, and amplitude generally are not known or necessarily determinable. This energy is characterised by high surface-wave ($R$ and $L$) and low body-wave ($P$ and $S$) energy. The ambient scenario is represented in Figure 1.1 as man-made (green, due to traffic) and natural (blue, due to ocean waves) “wavefields” propagating with elliptical wavefronts through the subsurface. Relative to passive scenarios, the interest is not on isolating strong P- or S-wave energy bursts; rather, it is on measuring body-wave arrivals from a well-sampled range of azimuths and apparent velocities. The $P$ and $S$ energy may be of relatively similar strength to each other and to $N$; however, $R$ and $L$ are generally larger in magnitude. Accordingly, the ambient strategy requires recording, processing and stacking of long-duration continuous records (e.g., hours or days) to develop sufficient signal-to-noise ratios (S/N) of stacked body-wave energy.

This thesis focuses on investigations that use the ambient seismic data scenario; in particular, it focuses on developing and identifying novel data processing and imaging analyses that are required to use ambient P-wave reflected energy to generate and validate 3D images of subsurface reflectivity. To motivate these novel developments, I first discuss some of the key concepts behind the governing (isotropic acoustic) wave propagation and the use of propagated wavefields for active-source seismic imaging purposes. After defining these concepts, I describe how to extend these approaches and develop new tools for ambient seismic imaging based on an extension of the direct wavefield migration approach (Artman, 2006).

### 1.2 Acoustic Wave Propagation

In active-source seismic experiments, a man-made and often well-characterised (or measured) seismic source is excited at or near the Earth’s surface at a known time and location. The associated source wavefield energy propagates through the Earth, interacts with contrasts in material properties, and is scattered upwards toward the surface. The response from this propagating wavefield is then measured by receivers (geophones), generally located at the surface. The resulting surface-recorded data are a superposition of direct arrivals from the injected source and the Earth’s response from (discontinuous) structure, and can include all of the different wave-mode types (for terrestrial surveys) discussed above.
1.2. ACOUSTIC WAVE PROPAGATION

Figure 1.2 illustrates this seismic forward modelling operation. The explosive source excited at location $f_s$ propagates as “source wavefield” $U_S$ to a scatterer $p$, representing a discontinuity in material properties. A scattered wavefield generated at point $p$ then propagates back toward the surface as singly scattered “receiver wavefield” $U_R$, where it is measured by a receiver at location $f_r$ as data $d$.

This synthetic representation of active-source seismic data can be simulated numerically by applying an acoustic forward modelling operator (e.g., two-way or one-way finite-difference wave-equation approximation), an injected source-time function (e.g., a Ricker wavelet), and an Earth model. In this thesis, I assume that it is justifiable to use an isotropic acoustic wave equation and Earth model (i.e., rather than an elastic wave equation) and to ignore any wave-type conversion or anisotropy effects\(^1\). While these effects can significantly influence wave propagation and are important for advanced active-source modelling and imaging investigations, these extensions are beyond both the purpose and scope of this thesis.

The time-domain acoustic wave equation for the scalar pressure field, the solution of which approximately represents the vertical motion of wavefield $U$, at location $x$ and depth $z$ from a source function $S(t)$ located at $[x, z] = [x_s, 0]$ m, may be defined as:

$$\nabla^2 U = \frac{1}{c^2} \frac{\partial^2 U}{\partial t^2} \quad \text{subject to} \quad \begin{cases} U(x, z, t) = 0, & t < 0, \\ U(x, z = 0, t) = \delta(x - x_s)\delta(t) * S(t), & 0 \leq t \leq T, \end{cases}$$

where $\nabla^2$ is the Laplacian operator, $c$ is the P-wave velocity of the medium, $\delta(\cdot)$ is a Dirac delta function, $*$ indicates temporal convolution, and $T$ is the record length. The top constraint of equation (1.2) is the initial condition stating that there is no wavefield disturbance for $t < 0$ s (i.e., quiescent before source excitation). The bottom constraint is the free-surface boundary condition\(^2\) where the pressure wavefield must be equal to the forcing term at the source location and zero elsewhere. This forcing term can be represented as $\delta(x - x_s)\delta(t) * S(t)$. The total wavefield modelled by implementing equation (1.2) is represented as

$$U = U_S + U_R,$$

which is the superposition of the non-scattered source wavefield $U_S$ and the scattered receiver

\(^1\)This approximation is often used in exploration seismology to reduce the overall computational requirements of the various imaging and inversion activities.

\(^2\)Formally, boundary conditions are required on all boundaries. Here, it is assumed that Helmholtz radiation boundaries exist such that the wavefields vanish at all non-free surface boundaries. In practice, these can be implemented with absorptive or perfectly matched layer (PML) numerical boundary conditions.
wavefield $U_R$. Therefore, one must use signal processing approaches to approximately separate these two contributions before applying the types of imaging operations described below.

### 1.2.1 One-way Wave-equation Approximations

There are computational benefits from solving the acoustic wave equation using one-way approximations in the frequency domain (Claerbout, 1985), including lower memory requirements and faster, though lower accuracy, numerical solvers. In many seismic imaging scenarios, such as the acoustic imaging problem in non-complex media investigated herein, one-way frequency-domain wave propagation represents an acceptable trade-off between numerical accuracy and computational complexity (i.e., runtime). Throughout this thesis I use the one-way frequency-domain solvers described by Shragge and Shan (2010) to generate wavefield solutions. In the frequency domain, the forward-modelled wavefield $U(x, z, \omega)$ can be represented as:

$$ L U = 0 \quad \text{subject to} \quad U(x, z = 0, \omega) = \delta(z) \delta(x - x_s) S(\omega), \quad (1.4) $$

where $\omega$ is angular frequency and $L$ is a partial-differential operator represented by

$$ L = \nabla^2 + \frac{\omega^2}{c^2}. \quad (1.5) $$

For 1D media, it is possible to develop a dispersion relationship between velocity $c$, angular frequency $\omega$, and vertical and lateral wavenumbers $k_z$ and $k_x$: $k_z^2 + k_x^2 = \frac{\omega^2}{c^2}$. This wavenumber expression is used to extrapolate a causal downward one-way source wavefield $U_S$ (which is not equal to the total wavefield $U$ because of the one-way propagation approximation and a lack of upward-propagating scattered wavefield energy) from depth level $z$ to $z + \Delta z$ according to:

$$ U_S(k_x, z + \Delta z, \omega) = U_S(k_x, z, \omega) e^{-ik_z \Delta z} \quad \text{subject to} \quad U_S(x, z = 0, \omega) = \delta(x - x_s) S(\omega), \quad (1.6) $$

where $i$ is the imaginary unit, and $\Delta z$ is the extrapolation depth step increment. One can model $U_S(k_x, z, \omega)$ by extrapolating from the surface to the maximum depth $z_{max}$ over number of depth steps $N_z$ (i.e., from $z = 0$ to $z = N_z \Delta z = z_{max}$ in $N_z$ steps). Scenarios involving more complex velocity models, $v = v(x, y)$, require using higher-order one-way propagator approximations like those discussed in Shragge and Shan (2010).

### 1.3 Active Source Imaging

The goal of active-source seismic imaging is to estimate the Earth’s reflectivity response. Figure 1.3 shows the imaging operation requiring two distinct wavefields: [1] the known source function injected at the source location and propagated causally (i.e., forward in time) to create source wavefield $U_S$ (Figure 1.3a); and [2] the singly scattered contributions
1.3. ACTIVE SOURCE IMAGING

Figure 1.3: Cartoon showing source and receiver imaging operations for active-source experiments. (a) The forward modelling operation injects and causally propagates the source information from $f_s$ into the subsurface creating source wavefield $U_S$, equivalent to the left hand side of Figure 1.2. (b) The single-scattered adjoint receiver wavefield, $U_R^\dagger$, is reconstructed from data, which are propagated in reverse time into the subsurface equivalent to the time-reversed right hand side of Figure 1.2. When using the correct velocity model for propagation, both wavefields will correlate at the same time at the correct scatterer location $p$ and can be stacked to create an image through a correlation-based (or deconvolution-based) imaging condition.

estimated from the measured data injected at the geophone locations and propagated anti-causally (i.e., in reverse time) to reconstruct the adjoint receiver wavefield $U_R^\dagger$, where $^\dagger$ indicates adjoint time reversal (Figure 1.3b). Given the correct velocity model, the two propagated wavefields will be co-spatially and co-temporally located at the scatterer $p$ and can be cross-correlated to create an image through an imaging condition.

The reverse-time propagation of an adjoint wavefield $U^\dagger$ can be done by modifying equation (1.2). This operation uses the numerical adjoint of the forward modelling operator, and is thus called the adjoint modelling operator. The governing partial differential equation (PDE) for modelling an adjoint wavefield $U^\dagger$ (from data recorded on the free surface) is given by:

\[
\nabla^2 U^\dagger = \frac{1}{c^2} \frac{\partial^2 U^\dagger}{\partial t^2} \quad \text{subject to} \quad \begin{cases} 
U^\dagger(x, z, t) = 0, & t > T, \\
U^\dagger(x, z = 0, t) = d_p(x, t), & T \geq t \geq 0,
\end{cases}
\]

where $d_p(x, t)$ represents data recorded at $f_r$ in Figure 1.2 that has been pre-processed to contain only primary reflections. The bottom constraint is the free-surface boundary condition where the adjoint wavefield must be equal to the data at the receiver locations at the associated time step. In this representation, anti-causal propagation is apparent by the reversed time ordering from $t = T$ back to $t = 0$ (i.e., the source excitation time). This requires an initial condition that $U^\dagger$ be quiescent for $t > T$.

It is straightforward to reconstruct an adjoint wavefield $U^\dagger(x, z, \omega)$ in the frequency domain using a frequency-domain modelling operator:

\[
\mathcal{L}U^\dagger = 0 \quad \text{subject to} \quad U^\dagger(x, z = 0, \omega) = \delta(z)d_p(x, \omega),
\]

\[7\]
where $L$ is defined as in equation (1.5). Similarly, one can propagate an anti-causal downgoing one-way adjoint receiver wavefield $U_R^\dagger$ using the adjoint one-way propagation operator (the complex conjugate to the forward one-way propagation operator):

$$U_R^\dagger(k_x, z + \Delta z, \omega) = U_R^\dagger(k_x, z, \omega)e^{+ik_z\Delta z} \text{ subject to } U_R^\dagger(x, z = 0, \omega) = d_p(x, \omega)$$  \hspace{1cm} (1.9)

where $k_z$ is defined by equation (1.6). Again, propagation from $z = 0$ to $z = N_z\Delta z = z_{max}$ in $N_z$ steps reconstructs an approximate one-way adjoint receiver wavefield $U_R^\dagger(k_x, z, \omega)$, which is desirable because the downward propagating one-way modelling procedure does not generate any upgoing wavefield energy.

1.3.1 Imaging Condition

Having simulated the causal source wavefield $U_S(x, z, \omega)$ in equation (1.6) (Figure 1.3a) and the anti-causal adjoint receiver wavefield $U_R^\dagger(x, z, \omega)$ in equation (1.9) (Figure 1.3b), it is possible to create an image of the scatterer at point $p$ by cross-correlating the two wavefields through an imaging condition (Claerbout, 1971). When repeating this process over many sources, the resulting image stack will highlight subsurface reflector locations (Claerbout, 1968). While there are numerous migration methods defined in the literature\(^3\), in this thesis I use the one-way wave-equation migration (WEM) style of active-source imaging (Claerbout, 1970), which employs forward and adjoint one-way wave-equation solutions of the acoustic wave equation (Claerbout, 1985) described above as the propagation kernel.

Mathematically, the imaging condition can be represented in the frequency domain as

$$I(x) = \sum_{\omega,e} \Re\left[U_S(x, \omega, e)U_R^\dagger(x, \omega, e)\right],$$  \hspace{1cm} (1.10)

where image $I(x)$ is obtained by cross-correlating the complex conjugate of source wavefield $U_S$ (identified with a overline as $\overline{U_S}$) with adjoint receiver wavefield $U_R^\dagger$ and summing the real component ($\Re$) over each contributing frequency $\omega$ and source location $e$.

When the WEM operation is computed with the correct velocity model for a full aperture of shots, each with sufficient S/N and processed to have only primary reflections distributed over a broad aperture, a clear subsurface migration image will emerge. However, using incorrect velocity model parameters will cause wavefields $U_S$ and $U_R^\dagger$ to propagate to incorrect subsurface locations at the wrong times leading to images with incorrectly placed reflectivity with poorer focus than if the correct velocity model were used. Because velocity models must be estimated from the data themselves, they are often inaccurate; thus, it is helpful to have additional quality control (QC) measures that can be used to examine the correctness of the migration velocity model.

---

\(^3\)Migration methods are commonly named after the method used for wave propagation: Kirchhoff (hyperbola summation) (French, 1974; Schneider, 1978), Fourier-domain (Stolt, 1978; Gazdag, 1978) and reverse-time migration (RTM) (Baysal et al., 1983).
1.4. AMBIENT IMAGING

1.3.2 Extended Imaging Condition

One way to examine velocity model correctness is by examining images created through an extended imaging condition (EIC) (de Bruin et al., 1990; Prucha et al., 1999; Sava and Fomel, 2006; Sava and Vasconcelos, 2011). To formulate an EIC, the conventional imaging condition is extended with additional lag parameter dimensions, which allows practitioners to examine whether the image focus could be improved by shifting wavefields in opposing directions prior to correlation. The EIC lags can be included in equation (1.10) as spatial-shift vector $\mathbf{\lambda} = [\lambda_x, \lambda_y, \lambda_z]$:

$$ I(x, \mathbf{\lambda}) = \sum_{\omega, e} \Re \left[ U_S(x + \mathbf{\lambda}, \omega, e) U_R^\dagger(x - \mathbf{\lambda}, \omega, e) \right]. $$  \hspace{1cm} (1.11)

The resulting extended image gather has as many additional dimensions as lag terms included.

Examining $I(x, \mathbf{\lambda})$ gathers can be quite useful in the context of migration velocity analysis (MVA). In an MVA approach, if a well-focused $I(x, \mathbf{\lambda})$ gather has a maximum at zero lag ($|\mathbf{\lambda}| = 0$ m), then the velocity model is considered to be correct. Figure 1.4b illustrates this scenario using a single lag in the horizontal direction, $\lambda_x$. However, when the focus of the extended image is at non-zero lag, as illustrated by the characteristic migration “frown” in Figure 1.4a and “smile” in Figure 1.4c, the implication is that the subsurface model used to generate wavefields $U_S$ and $U_R^\dagger$ is incorrect. Even though imaged energy in Figures 1.4a and 1.4c is centred about $|\mathbf{\lambda}| = 0$ m, much of the energy is poorly focused, especially when compared to the well-focused image in Figure 1.4b. Extended images may also be used to discriminate between different wave-type energy (i.e., S-wave energy propagated with the correct or close to correct P-wave velocity will not focus whereas P-wave energy will, shown in Figure 1.4), and they routinely serve as the basis for image-domain tomography schemes used to update velocity models using focusing-based MVA criterion (Liu and Bleistein, 1995).

1.4 Ambient Imaging

The physics and conceptual framework of ambient body-wave imaging are closely related to active-source imaging techniques. The same wave-equation physics applies to ambient energy propagating through the Earth as for active-source energy. Therefore, any measured ambient wavefield energy can be propagated through a subsurface model with the same acoustic forward and adjoint modelling operators described in the sections above. However, there are two main differences in ambient scenarios that need to be addressed: [1] there is neither an explicit or characterisable source wavefield to model $U_S$ nor a way to isolate the singly scattered contributions to form $U_R^\dagger$, and [2] the signal-to-noise ratio (S/N) of $P$ relative to other wavefield components is much lower since $R$ and $L$ usually dominate.

To illustrate the imaging consequences of these differences, Figure 1.5 depicts an unknown representative ‘source’ function arriving as a plane wave at surface location $f'_a$. 


Figure 1.4: Extended image gathers at the same image point over lag $\lambda_x$ for a simple active-source experiment illustrating migration velocity models that are: (a) too slow and showing a characteristic “frown”; (b) correct; and (c) too fast and showing a characteristic “smile”. Interestingly, a poorly focused S-wave centred at 2.4-2.6 s in panels (a)-(c) changes shape less but depth more than the P-wave energy as velocity is perturbed. This is expected since the P-wave velocity used for acoustic imaging of elastically forward modelled data is always too high for correctly focusing S-wave energy.

This wave reflects off of the free surface while simultaneously being recorded as data $d$ at $f_a'$. Downward-reflected energy propagates into the subsurface as wavefield $R_+$, eventually interacting with the subsurface discontinuity at point $p$. The scattered energy from $p$ propagates upward as wavefield $R_-$ to the surface, where it is again recorded at $f_a$ as data $d$. The $+$ and $-$ subscripts indicate down- and up-going energy with increasing time (i.e., causal propagation). Note that the rays depict the same travel paths as the source energy in the active-source example, save for the initial incoming wave path. This observation suggests an underlying connection with active-source imaging.

The original conjecture regarding the possibility of utilising background ambient signal to image the subsurface was introduced by Claerbout (1968), which showed that the autocorrelation of an earthquake seismogram can be compared to a seismic source from active-source experiments. The autocorrelation in this case is important because the response from the earthquake source will optimally correlate at zero lag.
Recent work has extended Claerbout’s autocorrelation conjecture to cross-correlation of spatially offset receivers. The concept of time-distance correlation was developed in Duvall et al. (1993) with applications for helioseismology to mimic impulsive sources on the surface of the sun. Rickett and Claerbout (1999) used the time-distance correlations from solar data and compared the results to active-source seismic signals acquired for petroleum exploration purposes in the Gulf of Mexico, and termed this style of investigation acoustic daylight imaging. This concept was expanded in Cole (1995), which reported the result of cross-correlating ambient seismic data with the goal of imaging subsurface reflectivity.

To illustrate these scenarios, Figure 1.6 depicts the cross-correlation of energy arriving at $f_a$ with that at $f'_a$, where the temporal lag between the two arrivals is indicative of the two-way travel time from $f'_a \rightarrow p \rightarrow f_a$. The cross-correlation operation essentially resets the “absolute clock” to zero, and disregards the initial incoming ray path effects and the absolute time of excitation.

There are two classes of methods derived from Claerbout’s original conjecture that use ambient data for investigating the subsurface: [1] data-domain interferometry employing cross-correlate and stack (CC+S) processing of recorded ambient energy (Wapenaar et al., 2004b); and [2] image-domain direct ambient migration involving wave propagation and applying a CC+S imaging condition (Artman, 2006). I describe these two approaches below.

### 1.4.1 Interferometry

Interferometry represents a multidimensional expansion of the imaging principle theory introduced by Claerbout (1968). As described in Schuster et al. (2004), interferometric (or daylight) imaging can be implemented between two receivers as long as both measure energy corresponding to the same source event. Interferometry aims to extract Green’s functions from passive seismic recordings (Wapenaar, 2004), which represent the Earth’s response to an impulsive point source excited at one receiver location as measured at a second. The frequency-domain Green’s function $G_{12}(x_1, x_2, \omega)$ between receivers located at $x_1$ and $x_2$ may be approximately reconstructed through a CC+S process as:

$$G_{12}(x_1, x_2, \omega) \simeq \frac{1}{N_w} \sum_{n=1}^{N_w} v_{1n}(x_1, \omega)v_{2n}(x_2, \omega),$$  \hspace{1cm} (1.12)

where $v_{in}$ is measured vertical particle velocity at receiver $i$ in time window $n$. The entire operation is stacked over the total number of windows, $N_w$. Given a Green’s function
in equation (1.12), one can ideally recover the active-source data that would have been produced by an impulse source located at \(x_1\) or \(x_2\). When extended to a wide range of virtual source and receiver locations, this process creates interferometric or virtual shot gathers, which can be used in the same way as conventional shot gathers in active-source imaging experiments (Wapenaar et al., 2004a; Wapenaar and Fokkema, 2006). Imaging using virtual source data created via interferometric techniques has proven successful at numerous locations and environmental conditions, including in the Saudi and Libyan deserts (Draganov et al., 2006, 2009), with the Long Beach array in greater Los Angeles, California (Nakata et al., 2011, 2015), at Lalor Lake mine site in Manitoba, Canada (Cheraghi et al., 2015; Roots et al., 2017) and at Kylylahti Mine Area, Finland (Chamarzuk et al., 2018; Väkevä et al., 2018) among others.

Figure 1.6: Cartoon showing the causal and anti-causal imaging operations for ambient seismic experiments. (a) Causally propagating data recorded at \(f_a'\) with the correct migration velocity creates wavefield \(R_+\), which correctly models the “source-side” contribution of the ambient wavefield, but incorrectly models the “receiver-side” contributions \((R_-)\) from \(f_a\). (b) Time-reverse propagation of data recorded at \(f_a\) reconstructs wavefield \(R_+^\dagger\), which correctly models the “receiver-side” contributions, but inaccurately reconstructs the “source-side” contributions \((R_-^\dagger)\) from \(f_a'\). Correlating the two wavefields results in an image of the scatterer, along with the erroneous spurious correlations involving wavefields \((R_-)\) and \((R_-^\dagger)\) illustrated by the dashed lines [equation (1.13)].

One of the challenges in accurately applying interferometric methods stems from the stacking operation itself: once energy passes through a CC+S process it is impossible to reverse the stacking action. Thus, if any strong coherent or otherwise non-Gaussian noise is present in the pre-correlated data, this information can strongly bias the resulting CC+S gather. Unfortunately, because most coherent noise sources recorded in ambient data sets are of much greater magnitude than the desired P-wave components, the former commonly dominate the resulting CC+S gathers and any resulting image that uses these as input is adversely compromised. Thus, developing and applying a robust data processing workflow to condition ambient data prior to applying the interferometric CC+S process would represent a desirable but as yet untested approach.

1.4.2 Direct Ambient Migration

The second form of ambient imaging is direct migration, which applies active-source seismic migration principles directly on propagated uncorrelated ambient seismic data to generate a subsurface image (Draganov et al., 2004; Artman, 2006, 2007). Using a migration approach
implies that imaging and identifying subsurface reflection events can be achieved in the image domain through a CC+S process. Note the contrast between image-domain CC+S (direct migration) and data-domain CC+S (interferometry).

As with active-source imaging, the migration operation requires two wavefields: one propagating causally into the subsurface, the other propagating anti-causally. In direct migration, causal wavefield $R_+$ is created from ambient data using forward modelling operators that propagate the recorded ambient data recorded at $f_a$ throughout the model. However, because the ‘source’ information is inherently measured in the data at $f_a$, this propagation operation also creates wavefield components with the wrong causality, $(R_-)$. Figure 1.6a illustrates both of these wavefield components. The second required wavefield, anti-causal $R_-^\dagger$, is created using the adjoint modelling operator by reverse-time propagating the data recorded at $f_a$. Like the forward modelling operation, the non-separability of the ‘source’ information from the singly scattered information creates another wavefield with the wrong causality ($R_+^\dagger$) by reverse-time propagation of energy recorded at $f_a$ (in Figure 1.6b). Cross-correlation of the forward and adjoint wavefields using an ambient imaging condition creates an image at point $p$. Ideally, this image will be dominated by the desired correlation between $R_+$ and $R_-^\dagger$; however, due to the non-separability of the pre- and post-scattered wavefield information, one should expect that non-zero correlations involving wavefield $(R_-)$ and $(R_+^\dagger)$ will cause spurious image events. Over sufficiently long recording time (large $N_w$), though, those spurious correlations ideally will be inconsistent, incoherent, and thus tend to zero while the coherent correlations of $R_+$ and $R_-^\dagger$ will increasingly contribute to the image stack.

An ambient imaging condition (Artman, 2006) may be developed in a similar fashion as the active-source imaging condition in equation (1.10):

$$I_A(x, N_w) = \frac{1}{N_w} \sum_{n=1}^{N_w} \sum_{\omega} \Re \left[ (R_+ + R_-)(R_+^\dagger + R_-^\dagger) \right] \simeq \frac{1}{N_w} \sum_{n=1}^{N_w} \sum_{\omega} \Re \left[ R_+ R_-^\dagger \right], \quad (1.13)$$

where $I_A$ is the ambient image and the sum is over a (large) number of windows, $N_w$, instead of the active-source index $e$. The approximate equality assumes convergence of the direct migration CC+S process.

A key novel extension reported in this thesis is that the ambient imaging condition in equation (1.13) can be extended to define an EIC in the same fashion as equation (1.11) for the active-source imaging condition. This extension is achieved by introducing opposing spatial lags to the $R_+$ and $R_-^\dagger$ wavefields:

$$I_A(x, \lambda, N_w) \simeq \frac{1}{N_w} \sum_{n=1}^{N_w} \sum_{\omega} \Re \left[ R_+ (x + \lambda, \omega, n) R_-^\dagger (x - \lambda, \omega, n) \right]. \quad (1.14)$$

The resulting ambient EIC is based on the same wave-equation physics as its active-source counterpart, and thus exhibits similar characteristics and shares the same potential uses. For example, if the subsurface velocity model used to create wavefields $R_+$ and $R_-^\dagger$ were
incorrect, the resulting ambient image should neither fully focus nor have energy imaged at the correct location (Girard and Shragge, 2015). One of the main motivations of this thesis is to demonstrate that the ambient EIC can be used in the ambient imaging process and provide researchers with a new tool to help evaluate the resulting ambient images.

1.5 Ambient Data Reality

The ambient imaging theory discussed in either of the interferometry or direct migration scenarios described above both implicitly require satisfying key assumptions: that the recorded data volume largely represents signal usable for P-wave reflection analysis and has wavefield arrivals appropriately distributed in frequency content, illumination direction and apparent velocity. In practice, though, these assumptions are never realized owing to the realities of field ambient data, particularly due to the presence of large-magnitude coherent noise events. In ambient field data, there are commonly three types of problematic coherent large-magnitude noise events: [1] stationary in time (i.e., simultaneously coherent across an array); [2] stationary in space (i.e., continuously coming from the same location but at different times); and [3] stationary in frequency (i.e., persistent monochromatic emissions). Each of these is detrimental to satisfying the underlying ambient assumptions.

Figure 1.7 shows an example of each type of noise using the Lalor Lake field data set analysed in this thesis. Figure 1.7a presents a signal from a dynamite blast that represents a broadband, large-magnitude coherent noise event [1] that is considered noise because the associated arrivals are strongly apparent in all traces in the array and will dominate any CC+S process. Figure 1.7b shows an interpreted road traffic noise event representing relatively high-amplitude events that can repeatedly occur at the same location [2], but at inconsistent time intervals. These events tend to bias the ambient imaging analysis toward the response of a narrow range of azimuths and apparent velocities. Figure 1.7c shows that even during recording periods exhibiting much lower amplitude recordings, there can be significant monochromatic noise energy generated at particular frequencies [3]. These events will have high correlation coefficients at spectral energy peaks and cause ringing in CC+S results.

These three types of large-magnitude coherent noise events present challenges to ambient imaging experiments because they can forestall achieving the underlying goal of imaging and identifying reflected P-wave energy. Because ambient P-wave reflections are much lower magnitude than surface-wave energy and the described coherent noise events, this leads to an imaging scenario where $S/N \ll 1$. While one could mitigate surface-wave energy prior to correlation by removing low frequencies through high-pass filtering, this approach would remove any overlapping lower-frequency P-wave content required for imaging purposes.

A second approach for overcoming low $S/N$ in CC+S processing is to use long-time ambient recordings. Ideally, coherent P-wave reflection event correlations will be stationary and coherently stack, over long times, while incoherent correlations with background (Gaussian) noise will be non-stationary and stack away. However, when the background noise is non-Gaussian or otherwise coherent, as in the examples shown above, the stacking
process may no longer converge. Thus, it is important to investigate novel approaches for mitigating or eliminating sources of coherent noise prior to applying any CC+S process, be they interferometry or direction migration.

1.6 Research Challenges and Thesis Contributions

The aim of this thesis is to develop novel approaches to ambient seismic data processing, direct migration, and ambient migration image analysis and interpretation. Specifically, the investigations reported in this thesis address the following three research questions:

[1] How to adapt concepts and techniques from active-source seismic imaging to help directly migrate, identify and interpret imaged ambient P-wave reflection energy?

![Figure 1.7: Example traces from Lalor Lake ambient data set with different types of large-magnitude coherent noise events. (a) Stationary in time represented by a strong broadband event related to a dynamite blast (i.e., nearly simultaneously across an array). (b) Stationary in space, likely related to road/mine site activity coming from a single direction. (c) Stationary in frequency, likely related to electro-mechanical vibratory noise from on-site electrical generators. Note that the amplitude scale changes in these panels.](image-url)
CHAPTER 1. INTRODUCTION

[2] How to develop signal processing workflows that will (better) condition ambient data to enable the types of ambient direct migration approaches developed in [1]?

[3] Given successful solutions to both [1] and [2], is it possible to make these concepts, techniques and workflows work on 3D ambient field data?

The research presented in this thesis addresses these three key challenges as a set of three integrated studies. The novel research contributions of each may be summarised as follows:

[1] To identify reflected body-wave signals in direct migrated ambient seismic data, I develop a novel ambient extended imaging condition used for direct migration of recorded ambient seismic data. Applying this process results in a set of 3D extended image gathers that show how the moveout characteristics of imaged wavefield energy change based on lag shifts and velocity perturbations. These moveout characteristics may then be used as a tool to identify energy types and help validate imaged P-wave reflectivity.

[2] To mitigate coherent noise and enhance the S/N of P-wave energy relative to other components, I develop a new automated ambient seismic processing workflow that aims to mitigate the types of coherent noise described above in the pre-correlation data domain. I also apply a suite of QC measures to both visualise and quantify improvements in the conditioned P-wave signal throughout the processing workflow.

[3] To show the value of these procedures outlined in [1] and [2], I implement both the automated processing and extended imaging direct migration workflows on a 3D ambient data set from Lalor Lake mine site in Manitoba, Canada. I also introduce a deconvolution (extended) imaging condition to balance the spectrum of the final images. The results illustrate the procedures’ ability to remove coherent noise and highlight reflection images in ambient field data that correspond to active-source images.

In addition to the Introduction and Conclusions and Future Work discussed in Chapters 1 and 5, respectively, this thesis presents these three technical studies as Chapters 2-4, abstracts of which are outlined below.

1.6.1 Study 1: Imaging, Identifying and Validating Imaged P-wave Reflectivity

Utilising ambient seismic energy naturally propagating in the Earth as an alternative approach to active body-wave seismic investigations has been a topic of interest for a number of decades. However, because ambient surface-wave arrivals typically are of much greater amplitude than ambient body-wave energy, significant data signal processing and long recording times are required to mitigate this and other coherent noise sources, and to correlate sufficient reflected body-wave energy to converge to a stable image. Even for
these scenarios, identifying and validating imaged body-wave reflection events remains challenging. In active-source investigations, extended imaging condition (EIC) gathers are used to examine velocity (in)accuracy. Herein, I develop an ambient direct migration approach that uses a novel ambient EIC. I simulate synthetic ambient seismic data for two different models and conduct a series of numerical experiments to demonstrate the velocity sensitivity of ambient seismic recordings in the migration image domain. Tests with varying global velocity perturbations show a characteristic reflector moveout in EIC gathers that can serve as a diagnostic of reflected ambient body-wave energy. I illustrate that this imaging formalism, under idealised circumstances, gives comparable results to conventional seismic methods, which extends the use of EIC gather-based image validation to ambient scenarios. I assert that this may be a valuable tool for the validation of ambient migration techniques that to date have yielded largely inconclusive results.

The work in this chapter is expanded from Girard and Shragge (2015). The expanded work has been submitted to the journal Geophysical Prospecting in a manuscript entitled “Direct Migration of Ambient Seismic Data,” with myself as the lead author and co-authored by my PhD supervisor Dr. Jeffrey Shragge. It includes results from Chapter 4 of this thesis. The current manuscript status is “accepted for publication.”

1.6.2 Study 2: Ambient Data Conditioning Through Processing

Extracting body-wave arrivals from ambient seismic recordings remains a challenging task, largely because ambient records are usually dominated by surface-wave energy. Most ambient seismic data processing strategies aimed at enhancing body-wave energy focus on a cross-correlation plus stack (CC+S) methodology. While this approach is useful for interferometric investigations, it effectively squares the magnitude of unwanted coherent noise events (e.g., surface waves, burst-like or strong monochromatic energy) that commonly overpower ambient body-wave signal. Accordingly, coherent noise events are usually treated with a binary accept-or-reject decision of individual data windows based on root-mean-squared energy considerations.

Applying a data processing workflow to uncorrelated ambient seismic data represents an alternate strategy for mitigating coherent noise. However, this prestack methodology requires significant computational effort due to the often Terabyte-sized data volumes. To make this approach feasible, I outline an automated processing workflow to automatically identify and mitigate coherent noise events that otherwise does not severely degrade the remaining waveforms. After each processing step I apply a number of quality control (QC) measures to monitor the convergence rate of CC+S waveforms - and for evidence of emerging body-wave reflection events. I apply the processing flow to an ambient seismic data set acquired on a large-N array at a mine site near Lalor Lake, Manitoba, Canada. My QC analyses suggest that automated preprocessing of uncorrelated ambient seismic recordings successfully mitigates unwanted impulsive and monochromatic coherent noise events. Accordingly, I assert that applying an automated data-processing approach would be beneficial for body-wave and other imaging and inversion analyses applied to ambient
seismic recordings.

The work in this chapter is expanded from Girard and Shragge (2016, 2017). The expanded work has been submitted to the journal Geophysical Prospecting in a manuscript entitled “Automated Processing Strategies for Ambient Seismic Data,” also with myself as the lead author and co-authored by my PhD supervisor Dr. Jeffrey Shragge. The current manuscript status is “accepted for publication.”

1.6.3 Study 3: Application to Field Data Set in 3D

After pre-processing ambient seismic recordings to enhance reflected body-wave energy using an automated data processing workflow, the next step is to use this ambient data volume to generate images of the subsurface reflectors. One way to do this is through the direct migration procedure discussed in Chapter 2. However, there are additional challenges associated with expanding this imaging methodology to 3D scenarios including: irregular array geometry; imbalanced frequency, illumination azimuth, and apparent velocity content; and uncertainty of whether imaged events represent true subsurface reflectors or spurious noise correlations.

The first challenge to address is how limited and irregular array geometry affects the imaging aperture and illumination. I examine this issue for the processed Lalor Lake data set by using a synthetic horizontal plane-wave test to explore the acquisition footprint of imaged planar reflectors located over a range of model depths. Identifying poorly or non-illuminated regions is helpful for not misidentifying or overinterpreting spuriously imaged artefacts. The second challenge associated with imbalanced Lalor Lake ambient data spectral content largely can be addressed by introducing a deconvolution (extended) imaging condition designed to balance the spectra of wavefields contributing to the final image. The third challenge of quantifying whether imaged correlations represent subsurface reflections or spurious noise correlations is addressed using the concepts of extended imaging condition (EIC) gathers introduced in Chapter 2. Building multi-dimensional EIC gathers in both horizontal lags and global velocity perturbations assists with discriminating between correlations with and without expected reflector moveout characteristics.

I conduct a number of imaging experiments that illustrate an ability to image subsurface correlations using the fully processed Lalor Lake data set and the 3D ambient deconvolution extended imaging condition approach. The repeatability of the 3D ambient imaging experiment over many hours is illustrated through stationarity plots to track changes in stacked images over a multi-hour recording time. Finally, the ambient imaging results show that many ambient imaged reflections are spatially coincident in numerous well-illuminated locations with expected subsurface reflectors when compared to post-stack time migration results from an active-source seismic survey.

Results from this chapter have been incorporated into the accepted manuscript entitled “Direct Migration of Ambient Seismic Data” that includes content from Chapter 2.
Chapter 2

Direct Migration of Ambient Seismic Data: Theory

“It’s not what you look at that matters, it’s what you see.”

Henry David Thoreau

2.1 Introduction

Over the past few decades there has been much interest in body-wave imaging in the absence of active seismic sources; however, overall community success has been somewhat limited due to a lack of validatable field examples. Previous investigations using passive and ambient seismic recordings include interferometric methods to develop pseudo-source gathers (Wapenaar et al., 2004a; Draganov et al., 2006; Forghani and Snieder, 2010; Draganov et al., 2013; de Ridder, 2014; Cheraghi et al., 2015); microseismic energy to image subsurface fractures (Reshetnikov et al., 2009) as well as for elastic velocity inversion (Witten and Shragge, 2017); and direct ambient wavefield migration involving wave-equation propagation and passive imaging conditions to investigate subsurface geology (Artman and Shragge, 2003; Draganov et al., 2004; Artman, 2006). A common theme in these ambient examples is the challenge of isolating, conditioning and validating the body-wave arrivals while removing high-energy coherent noise events (e.g., blasts, traffic, resource production activities) detrimental to the ambient imaging and inversion processes (Draganov et al., 2013; Nakata et al., 2015; Girard and Shragge, 2016).

There are a number of investigations that explore ambient body-wave energy in experiments progressing from 1D wavefield correlation/interferometry to full-wavefield 3D ambient subsurface imaging. Cole (1995), which presents an early example of extracting subsurface information from ambient recordings, shows that cross-correlation of two traces can recover a reflectivity estimate, assuming flat layer geology and plane-wave arrivals. However, this approach did not hold for more complex field data largely because of not
honouring the requirements of 1D geology, sufficient signal-to-noise ratio (S/N) and planar arrivals. Artman (2006) reports the first major investigation into using ambient recordings for seismic exploration at the Valhall field in the Norwegian North Sea. This work employs a technique for isolating body waves and generating P-wave reflection images from autocorrelated ambient wavefield records; however, this approach was hampered by largely unidirectional coherent noise sources originating from nearby production facilities, which effectively dominated any (other) reflected body-wave signal present in the recordings. At the same location, de Ridder et al. (2014) demonstrates that time-lapse tomography may be implemented with ambient seismic surface-wave energy to determine velocity changes over calendar time. In another series of examples, Draganov et al. (2007, 2009, 2013) retrieve P-wave reflection images from ambient seismic recordings at exploration scales in the Saudi and Libyan deserts. More recently, Issa et al. (2017) images geological reflectivity (∼2 km) using co-located active and ambient seismic recordings at a CO₂ sequestration site near the coast in Otway, Victoria, Australia. These examples demonstrate the possibility of undertaking body-wave imaging with ambient seismic energy and/or validating images against active-source results.

There are three key challenges that are more significant in ambient seismic imaging than in active-source methods: (1) identifying and removing coherent noise; (2) handling (often severely) non-uniform illumination; and (3) identifying and validating imaged reflectivity. Coherent noise is demonstrably detrimental to ambient seismic imaging because it commonly dominates any existing reflected body-wave event with S/N often ≪ 1 (Artman, 2006). This contrasts with conventional reflection seismic data where S/N of 5 to 10 or greater are common. Even if the ambient data were not severely contaminated by coherent noise, illumination issues can lead to recorded responses being dominated by arrivals from only a limited range of incidence angles and back-azimuths, which generates spatially band-limited wavefields for imaging and therefore leaves subsurface features poorly or un-illuminated. This contrasts with active-source scenarios where the distribution of source illumination and receiver coverage largely can be controlled. Lastly, directly identifying and validating imaged body-wave energy in ambient seismic recordings remain extremely challenging tasks because non-body-wave energy can be spuriously imaged and appear as signal, leading to misinterpretations.

Without being able to directly identify coherent body-wave events in raw ambient data, as is done in conventional seismic experiments, one may consider examining alternate domains where these signals potentially are more prominent. The most common alternative is the cross-correlation-plus-stack (CC+S) domain. A CC+S approach using an interferometric technique to build pseudo-source gathers (Claerbout, 1968; Wapenaar, 2004; Bakulin and Calvert, 2006) enables the use of conventional seismic processing methods. While the CC+S approach has proven successful in extracting refracted body-wave arrivals in surface-wave-dominated ambient recordings (Nakata et al., 2011), it has shown less promise when applied to highlight reflected ambient body-wave energy (Nakata et al., 2015). Importantly, because CC+S is an irreversible operation, any high-energy coherent noise
events contributing to the CC+S gather will overpower weaker reflected body-wave energy, a scenario to which I attribute many of the challenges in ambient body-wave imaging.

Because of these coherent noise issues in CC+S approaches, herein I perform body-wave reflection analysis directly in the migrated image domain. I use a “direct migration” approach (Artman, 2007) to build up a migrated image volume by propagating ambient wavefield energy causally and anti-causally through a subsurface velocity model and correlating the resulting wavefields using a novel imaging condition. Using the migrated image domain inherently will include wavefield correlations that image P-wave reflections and, ideally, minimise the influence of other wave modes through the combined effects of wave propagation and stacking over many event windows.

Since wave propagation and imaging are dependent on the underlying velocity model, it is possible to explore the sensitivity of migration results to the P-wave velocity model through use of an extended imaging condition (EIC) (Sava and Fomel, 2003; Sava and Biondi, 2004; Biondi, 2006; Vasconcelos et al., 2010; Sava and Vasconcelos, 2011). An EIC introduces a spatial lag term, $\lambda$, which represents an opposing symmetric vector shift between two wavefields about the image point prior to CC+S. While a temporal lag, $\tau$, can also be used (Vasconcelos et al., 2010), herein I consider only spatial lags, which leads to the so-called subsurface offset-domain common-image gathers (Biondi, 2006). The main utility of EIC gathers is that when the migration velocity is correct, imaged reflected energy should focus to a (band-limited) point about zero lag in the EIC gather volume. However, when the migration velocity is incorrect, imaged reflected energy will not focus to a point and will appear as a “smile” at shallower depths or a “frown” at greater depths depending on whether the migration velocity model is too fast or slow, respectively. This quantifiable reflector moveout signature is a diagnostic of the imaging velocity correctness and is used in image-domain adjoint-state tomography algorithms to update velocity models based on minimising EIC gather mis-focussing (Biondi and Symes, 2004; Shen and Symes, 2008; Symes, 2008; Yang and Sava, 2015; Witten and Shragge, 2017).

Images created from ambient seismic data are likewise sensitive to the migration velocity model. Thus, applying an EIC through direct migration of long-recording ambient data should image events (if present) with moveout signatures characteristic of body-wave reflections. Moreover, by completing several migrations with different scaled global migration velocity models, it may be possible to identify and validate imaged events as ambient P-wave reflections based on their characteristic spatial-lag and velocity moveout signature in 3D EIC gathers. To illustrate these conjectures, I begin by outlining current active-source migration practice using sources at known times and locations. I then recast this direct migration framework for use with (passive and) ambient wavefield recordings, and determine the stability of incrementally stacked ambient images by quantifying a measure of image convergence (Issa et al., 2017). After exploring the sensitivity of EIC gathers to velocity change, I undertake a number of synthetic experiments that show the similarity in reflector moveout signature of EIC gathers in active-source, passive and ambient scenarios. Results from these experiments demonstrate the applicability of ambient
EIC gathers to exploration-scale investigations, provided that the recorded data and final image satisfy key requirements (i.e., not dominated by coherent noise sources or surface waves, exhibit well-balanced illumination, and have converged to a stable migration result).

2.2 Active-Source Imaging Framework

Conventional active-source acoustic seismic imaging is based on the concept of single scattering, as I illustrate for the idealised scenario shown in Figure 2.1. The forward modelling problem assumes that a known seismic source, $S(f_s)$, originating at a specific location $f_s$ and time (assumed to be $t = 0$ s), propagates as a source wavefield $U_S$ into the subsurface (left arrow in Figure 2.1a). This source wavefield interacts with discontinuities in material properties at $p$, returns to the surface as a singly scattered wavefield $U_R$ (right arrow in Figure 2.1a), and is measured as seismic data, $d(f_r)$ at receiver location $f_r$.

For the adjoint (migration) imaging problem, one begins by forward modelling a source wavefield, $U_S$, from an assumed source function, $S(f_s)$, throughout the medium (Figure 2.1b). Similarly, a separate receiver wavefield, $U_R^\dagger$, is reconstructed from recorded data $d_p(f_r)$ (processed to be only the singly scattered contributions, or “primaries”) through propagation in reverse time throughout the subsurface model (Figure 2.1c). This reverse-time operation is denoted by superscript $\dagger$ on the reconstructed wavefield term. If the velocity model is correct, the two wavefields will coincide spatially and temporally at the locations of discontinuities (i.e., at $p$). Correlating wavefields $U_S$ and $U_R^\dagger$ through an imaging condition thus generates an image of the scatterer (Claerbout, 1968; Berkhout, 1982). Therefore, to fully develop the active-source imaging framework described above two operations must be specified: (1) a numerical method for acoustic wave propagation; and (2) an imaging condition used to generate an image.

2.2.1 Acoustic Wave Propagation

Numerical acoustic wave propagation may be undertaken in either the time or frequency domain. Here, I restrict the discussion to frequency-domain propagators for reasons that will be discussed below. I define a frequency-domain wave-equation modelling operator, $\mathcal{L}$, as

$$\mathcal{L} = \nabla^2 + \omega^2 s^2,$$

where $\nabla^2$ is the Laplacian operator, $\omega$ is the angular frequency, and $s$ is slowness or reciprocal of velocity (i.e., $s = \frac{1}{v}$). These one-way downward-continuation operators can be implemented in a causal ($\mathcal{L}_+$) or anti-causal ($\mathcal{L}_-$) sense and effectively introduce the necessary phase shift as one progresses the wavefield between successive depth levels (Claerbout, 2010). Conventional active-source imaging of surface-recorded data requires forward (causal) and adjoint (anti-causal) propagation of the estimated source wavefield energy, $S(f_s) = S(x_s, z = 0, t)$, and receiver data processed to be only primary reflections,
\[ d_p(f_r) = d_p(x, z = 0, t), \] Figures 2.1b and 2.1c respectively, according to:

\[
\begin{bmatrix}
L_+ & 0 \\
0 & L_-
\end{bmatrix}
\begin{bmatrix}
U_S \\
U_R^\dagger
\end{bmatrix}
= 
\begin{bmatrix}
S(f_s) \\
d_p(f_r)
\end{bmatrix},
\tag{2.2}
\]

where \( x \) and \( y \) are the lateral coordinates, \( z \) is depth, and \( U_S = U_S(x, \omega) \) and \( U_R^\dagger = U_R^\dagger(x, \omega) \) are the reconstructed source and receiver wavefields extrapolated throughout subsurface volume \( x = (x, y, z) \). Reconstructing wavefields \( U_S \) and \( U_R^\dagger \) requires using a slowness model because of the functional dependence of the propagator, \( L = L(s) \), in equation (2.1).

**Figure 2.1:** Cartoon showing forward modelling and adjoint imaging operations for active-source and ambient experiments. (a) Forward modelling of an injected source \( S(f_s) \) at location \( f_s \) creates source wavefield \( U_S \) that diffracts off a subsurface scatterer at \( p \), reflected toward the surface as wavefield \( U_R \) and recorded as data \( d(f_r) \) at location \( f_r \). (b) An adjoint imaging operation causally propagates the known source information into the subsurface at \( f_s \), creating source wavefield \( U_S \). (c) The estimated singly scattered adjoint (receiver) wavefield, \( U_R^\dagger \), created by injecting recorded data processed to be only primary reflections \( d_p(f_r) \) at \( f_r \), propagates in reverse time into the subsurface. When using the correct velocity, the wavefields in (b) and (c) will correlate at the scatterer \( p \) to create an image. (d) An upgoing ambient incident plane wave reflects off the surface and is recorded as data, \( d(f'_a) \), before propagating down to the scatterer \( p \) as wavefield \( R_+ \), which diffracts wavefield information upward as \( R_- \), to be measured as data \( d(f_a) \). (e) Causally propagating \( d_p(f'_a) \) creates wavefield \( R_+ \), which correctly models the “source-side” contribution of the ambient wavefield, but incorrectly models the “receiver-side” contributions \( d_p(f_a) \), denoted as \( (R_-) \). (f) Adjoint propagation of data \( d_p(f_a) \) reconstructs wavefield \( R_+^\dagger \), which correctly models the “receiver-side” contributions from \( f_a \), but inaccurately reconstructs the “source-side” contributions \( d_p(f'_a) \), denoted as \( (R_+^\dagger) \). Correlating the wavefields \( R_+ \) in (e) and \( R_+^\dagger \) in (f) results in an image of the scatterer at \( p \), along with the erroneous crosstalk through correlations with wavefields \( (R_-) \) and \( (R_+^\dagger) \).
Herein, I apply high-angle, one-way, frequency-domain modelling operators (Shan and Biondi, 2005; Shragge and Shan, 2010) for $L_+$ and $L_-$ because, in my examples, there are neither strong lateral velocity variations, nor do I expect significant overturning events. I note that these operators are more computationally and memory efficient (though of lower accuracy) than the two-way operators used in reverse-time migration (RTM), and also stack over the full waveform unlike Kirchoff migration operators. Thus, they represent a judicious accuracy and computational trade-off for the types of direct ambient migration scenarios and models discussed herein.

2.2.2 Active-source Imaging Conditions

Using the definitions of wavefields $U_S$ and $U_R^\dagger$ from equation (2.2), one can pose the CC+S process as an imaging condition (Claerbout, 1971) throughout subsurface $x$:

$$I(x) = \sum_{\omega,e} \Re \left[ U_S(x,\omega,e)U_R^\dagger(x,\omega,e) \right], \quad (2.3)$$

where image $I(x)$ is obtained by summing over each contributing frequency $\omega$ and source location $e$. The overline ($\overline{U}$) denotes complex conjugation and $\Re$ the real part of the calculation.

To make an EIC, a vector spatial lag term, $\lambda$, can be introduced into equation (2.3) (Sava and Vasconcelos, 2011). The corresponding EIC gather is then given by:

$$I(x,\lambda) = \sum_{\omega,e} \Re \left[ U_S(x+\lambda,\omega,e)U_R^\dagger(x-\lambda,\omega,e) \right]. \quad (2.4)$$

Note that equation (2.4) reduces to equation (2.3) where the lag term is zero ($|\lambda|=0$), which is also known as a "zero-offset" imaging condition.

2.3 Ambient Source Imaging Framework

The active-source imaging framework described above can be adapted to ambient seismic scenarios because all propagating seismic energy must follow the same physical laws, regardless of the source (assuming a linear elastic model). Like active-source scenarios, ambient direct migration requires propagating acquired data through a subsurface velocity model using the same wavefield propagators ($L_-$ and $L_+$) as the active-source case, and correlating the resulting wavefields to image discontinuous subsurface structure. An EIC for ambient seismic analysis can be developed that is similar to that in equation (2.4); however, unlike in active-source scenarios, it must be defined such that it does not require explicitly separating source wavefield energy (i.e., $U_S$) from non-source contributions (i.e., $U_R$), for reasons discussed below.

Figures 2.1d–f present the ambient migration scenario to parallel the active-source migration scheme discussed in Figures 2.1a–c. Figure 2.1d shows an upgoing ambient body-wave event (here symbolised as a plane wave) recorded at the free surface as $d(f_a')$. 

24
2.3. AMBIENT SOURCE IMAGING FRAMEWORK

The free-surface-reflected, down-going wavefield component then follows the same wave physics as an active source wavefield \((U_S)\) as it propagates into the subsurface, scatters off of discontinuities, and is subsequently recorded as \(d(f_a)\). Thus, these data are related to the active-source data, \(d(f_r)\), and contain reflected ambient wavefield arrivals; however, there is one key difference.

Ideally, I could record separable \(d(f'_a)\) and \(d(f_a)\), but because I know nothing about the ambient source field (i.e., arrival time, spatial extent or origin frequency content), I am left with processing \(d(f_a)\) as a composite of source and receiver wavefield information. Analogously to the active-source scenarios, Figure 2.1e shows the recorded ambient “source” \(d(f'_a)\) [i.e., contained within \(d(f_a)\)] forward-propagated as wavefield \(R_+\), denoted as a solid line. Similarly, Figure 2.1f shows the ambient data \(d(f_a)\) [i.e., containing \(d(f'_a)\)] that are time-reverse propagated as wavefield \(R_+^\dagger\), again denoted as a solid line. Wavefields \(R_+\) and \(R_+^\dagger\) will correlate and form an image at the scatterer \(p\) if the velocity model is correct (Artman, 2006). However, unlike in the active-source scenarios, in addition to the desired cross-correlation components (i.e., the two solid lines) there are now unwanted crosstalk components \((R_-)\) and \((R_+^\dagger+)\) that will correlate at incorrect locations in the migrated image domain (i.e., correlations involving a dashed line). These crosstalk components exist because it is not possible to separate the equivalent of source-and-receiver-side contributions in the ambient recordings. Over long recording times, though, crosstalk components ideally will stack out leaving only singly scattered contributions in the ambient migration image. Thus, by modifying the active-source imaging condition to account for these differences, one can define a direct migration approach where the equivalent of the active-source reflectivity response can be approximated from ambient data (Artman, 2006), provided one uses a “sufficient” number of event windows not dominated by coherent noise or suffering from narrow illumination issues. Related approaches have been discussed in Claerbout (1968) and Wapenaar et al. (2004b).

The aforementioned ambient direct migration approach can be implemented in a way parallel to the active-source migration [equations (2.2) and (2.3)]. First, I propagate recorded ambient data according to:

\[
\begin{bmatrix}
  L_+ & 0 \\
  0 & L_-
\end{bmatrix}
\begin{bmatrix}
  R_+ \\
  R_+^\dagger
\end{bmatrix} =
\begin{bmatrix}
  d(f_a) \\
  d(f_a)
\end{bmatrix}.
\]

(2.5)

However, instead of having source and receiver data, \(S(f_s)\) and \(d(f_r)\), as in equation (2.2), I now have only ambient data \(d(f_a)\) containing the unseparated “source” and “receiver” wavefield contributions. To generate the two reconstructed wavefields, I use the previously described one-way frequency-domain modelling operators, \(L_+\) and \(L_-\). The resulting causally and anti-causally propagated ambient wavefields, \(R_+\) and \(R_+^\dagger\), respectively, replace wavefields \(U_S\) and \(U_R^\dagger\) from equation (2.2). Having reconstructed the two wavefields, I can then define an ambient EIC similar to equation (2.4), which results in the following
expression for an ambient EIC gather:

$$I_A(x, \lambda, N_w) = \frac{1}{N_w} \sum_{n=1}^{N_w} \sum_{\omega} \Re \left[ R_+^\dagger (x + \lambda, \omega, n) R_- (x - \lambda, \omega, n) \right], \quad (2.6)$$

where index $n$ refers to the corresponding ambient data record in the $n^{\text{th}}$ of $N_w$ total windows [i.e., in place of the active shot index $e$ in equation (2.3)]. $N_w$ is included as a normalisation factor.

In the linear system in equation (2.5), operator $L_+$ correctly models the source wavefield contributions of $d(f_a)$, but not the primary reflections (i.e., downgoing legs represented by the dashed line ($R_-$) in Figure 2.1e). Similarly, operator $L_-$ correctly propagates the primary reflections of $d(f_a)$, but not the source wavefield components (i.e., downgoing leg represented by the dashed line ($R_+^\dagger$) in Figure 2.1f). Assuming a variety of illumination angles, this crosstalk largely will not correlate in the same location for different windows, meaning that one may stack over many correlation windows to constructively build up the primary reflection contributions while stacking away crosstalk components. This is also true for coherent noise not following the single-scattering assumption. The rate at which the ambient EIC gather converges to a stable result [i.e., $\lim_{N_w \to \infty} I_A(x, \lambda, N_w)$ in equation (2.6)] is a measure of “image convergence.”

### 2.4 Image Convergence

Because coherent noise will not correlate, and crosstalk will not appear, at the same subsurface locations for each imaged window, it is required to quantify when these unwanted contributions have been stacked out and a stable image has formed. To do this, I apply the concept of image convergence introduced by Issa et al. (2017). This procedure requires calculating the root mean square (RMS) energy in the cumulative stack after adding each window, and then normalising the result by the window fold to generate a measure of how much “different” energy is introduced by adding each successive window [hence, the dependence on $N_w$ in equation (2.6)].

Assuming that a recorded ambient wavefield consists of signal $S$ (i.e., body-wave components) and noise $N$ (i.e., all other contributions) such that $R_+(x + \lambda, \omega, n) = S_+(x + \lambda, \omega, n) + N_+(x + \lambda, \omega, n)$ [with complement $R_- (x - \lambda, \omega, n) = S_- (x - \lambda, \omega, n) + N_- (x - \lambda, \omega, n)$], I can redefine the EIC gather $I_A(x, \lambda, N_w)$ from equation (2.6) as:

$$I_A(x, \lambda, N_w) = \frac{1}{N_w} \sum_{n=1}^{N_w} \sum_{\omega} \left( S_+ + N_+ \right) \left( S_- + N_- \right). \quad (2.7)$$

For low numbers of $N_w$ and energetic $N_+$ and $N_-$ terms, the $\overline{N_+N_-} + \overline{S_+N_-} + \overline{N_+S_-}$ contributions will dominate $\overline{S_+S_-}$. However, when using a relatively large number of windows, the individually weaker signal components with stationary paths ideally will begin to cumulatively dominate the stack because the noise components are assumed to be
2.4. IMAGE CONVERGENCE

non-stationary. Thus, I represent the expanded right-hand-side of equation (2.7) as two terms:

\[ I_A = \left[ \frac{1}{N_w} \sum_{n=1}^{N_w} \sum_{\omega} S_+ S_- \right] + \left[ \frac{1}{N_w} \sum_{n=1}^{N_w} \sum_{\omega} \left( N_+ N_- + S_+ N_- + N_- S_- \right) \right] \equiv I_{\text{stat}} + I_{\text{noise}}, \]

(2.8)

where \( I_{\text{stat}} \) is the desired stationary (i.e., fully converged) image stack to be found, and \( I_{\text{noise}} \) is comprised of unwanted noise correlations to be minimised by stacking over a “sufficient” \( N_w \).

2.4.1 Image Variance

I am interested in assessing the asymptotic behaviour of \( I_A \) as \( N_w \) increases, which I define as image convergence. An image is considered to be converged when a stack of ambient data windows no longer changes (within tolerance) when adding more windows. I accomplish this by examining image variance as a function of the total number of contributing windows. Following Issa et al. (2017), I take the variance of equation (2.8), which yields an image variance panel \( I_{\text{var}} \):

\[ I_{\text{var}}(x, N_w) \equiv \text{Var}(I_A(x, N_w)) = \text{Var}(I_{\text{stat}}(x)) + \frac{1}{N_w} \left[ \sum_{n=1}^{N_w} \sum_{\omega} \left( N_+ N_- + S_+ N_- + N_- S_- \right) \right]^2, \]

(2.9)

where the variance in the overall image stack, \( \text{Var}(I_A) \), approaches that of the ideal stationary image panel, \( \text{Var}(I_{\text{stat}}) \), because the variance of the \( I_{\text{noise}} \) term ideally will tend to zero as \( n \) increases.

To visualise the overall rate of convergence, one may compute how the cumulative global RMS image energy changes as a function of \( N_w \) panels in equation (2.9). To do this, I compute the global RMS image energy in image \( I_{\text{var}} \) after the addition of each new window:

\[ I_{\text{var}}(n) = \frac{1}{n^2} \sum_{x} I_A^2(x, n), \quad n = 1, N_w. \]

(2.10)

Plotting \( I_{\text{var}} \) as a function of contributing windows generates a convergence curve that quantifies the variance in total energy added by each successive window. Ideally, \( I_{\text{var}} \) should converge to \( \text{Var}(I_{\text{stat}}) \) if \( N_+ \) and \( N_- \) have favourable statistical properties (i.e., Gaussian distribution of coherent noise events), but will take an increased number of \( N_w \) to converge - if at all - if the noise components have characteristics unfavourable for ambient imaging (e.g., a strongly biased non-Gaussian noise events).

Creating a stable and interpretable ambient image requires optimising the S/N of stacked energy. Achieving this goal requires: (1) ensuring that each window has favourable statistical properties (i.e., nearly uniform illumination direction, broad-band spectrum, and minimal coherent noise events) ideally by applying ambient signal processing workflow prior to imaging [see, e.g., Girard and Shragge (2016)]; and (2) stacking over a sufficient number of windows such that the image volume achieves convergence. I demonstrate these
CHAPTER 2. DIRECT MIGRATION OF SYNTHETIC AMBIENT DATA

below in a series of numerical experiments.

2.5 Experiment I: Layered Half-space Model

I conduct three numerical experiments on an idealised 2D elastic half-space model to test the EIC gather and image convergence theory described above. I simulate synthetic data volumes and subsequently perform imaging experiments using three different energy models: active (i.e., conventional), passive (i.e., microseismic) and ambient sources. I choose these three source types to illustrate that ambient direct migration and imaging conditions, under idealised conditions, can yield results comparable to those where the sources are known.

2.5.1 2D Elastic Model

I perform tests on a synthetic two-layer half-space model (Figure 2.2) with constant P- and S-wave velocities ($V_P=2.0$ km/s and $V_S=V_P/\sqrt{3}$ km/s, respectively), and upper and lower layer densities of $\rho=1.0$ g/cm$^3$ and $\rho=2.0$ g/cm$^3$, respectively. Figure 2.2 also shows the sinusoidal interface forming the density contrast as well as the acquisition geometry used for the active, passive and ambient experiments. The model is of dimension $[nx, nz] = [3072, 4096]$ with a 2.5 m grid spacing in both dimensions. Although simplistic, this model is useful for illustrating how different subsurface dips influence image focussing, particularly when using an incorrect migration velocity model. The white grid lines represent...
the locations where I extract and visualise EIC gathers in the ensuing figures. The left and right lines are termed the “anticline” and “syncline” locations below, respectively.

2.5.2 Data Generation

I generate active-source, passive and ambient data sets using a 2D elastic finite-difference time-domain modelling code (Weiss and Shragge, 2013). The active-source experiment is used as a control to compare the passive and ambient migration methods and to highlight any similarities and differences that may exist. I model shots at 151 locations with a 40 m source spacing (blue Xs in Figure 2.2), using a custom broad-band wavelet with energy ranging between 2-50 Hz. Receivers are located at 10 m intervals and span a 10 km recording aperture (red dots in Figure 2.2).

For the passive seismic experiment I model 50 microseismic-like sources at random locations below the undulating reflector using the same receiver locations as in the active-source experiment. The wavefields arrive as up-going events that are recorded at the free surface [i.e., \( d(f_a') \)]. After being reflected downward, the wavefield scatters off the discontinuity before progressing upward to be recorded as passive data, \( d(f_a) \). Given these propagation paths, the goal of this experiment is two fold: (1) to illustrate the applicability of the (ambient) direct migration approach with unknown source locations and mechanisms; and (2) to demonstrate that the ambient EIC gathers exhibit interpretable results even when the source information is not explicitly defined in the imaging formalism.

I model ambient wavefield data using a modified version of the 2D elastic modelling code referred to above. Instead of using an active-source wavelet, though, I seed two random (i.e., Gaussian) wavefield states for the first two time steps \( t_0 \) and \( t_{-1} \). This results in energy propagating not from a single source location, but roughly equiprobably from any direction at all subsurface locations. I filter these idealised wavefield states in \( k_x-k_z \) space to generate data within the typical ambient experimental range (0.1 to 30 Hz). Using different random seeds, I model 128 windows with 13-second recording times to simulate a \( \sim 30 \)-minute ambient data set. In this idealised scenario there is no dominant energy source, nor unique wavelet, which could represent a field recording after careful filtering and balancing of ambient records through applying a specialised processing workflow (Girard and Shragge, 2016). I note that surface-wave energy is not specifically modelled in the data volume, other than that naturally arising through elastic wave propagation. Again, I emphasise that this is an idealised representation of ambient recordings.

Figure 2.3a shows a representative forward-modelled ambient wavefield snapshot that appears incoherent with a similarly incoherent \( k_x-k_z \) spectrum (Figure 2.3b). Figure 2.3c shows an ambient data record, while Figure 2.3d presents the corresponding \( \omega-k_x \) spectrum. Though Figures 2.3a and 2.3c do not look like standard seismic data, the recording does indeed carry information about the subsurface, as revealed through direct migration experiments below.
CHAPTER 2. DIRECT MIGRATION OF SYNTHETIC AMBIENT DATA

Figure 2.3: Ambient seismic modelling results. (a) Propagated ambient wavefield snapshot with (b) corresponding $k_x$-$k_z$ spectrum. (c) Surface-recorded ambient data for the model presented in Figure 2.2 with (d) corresponding $k_x$-$\omega$ spectrum.

2.5.3 Zero-offset Imaging

I applied the active-source imaging formalism of equations (2.2) and (2.3) to the 151 modelled active-shot gathers. Figures 2.4a–c present the migration results for three different velocity models using the conventional zero-offset imaging condition [equation (2.3)]. Figure 2.4b shows the image when using the correct acoustic migration velocity ($V$). Figures 2.4a and 2.4c present the images when the migration velocity is globally perturbed by $dV=-5\%$ and $dV=+5\%$, respectively. The most prominent effect of using the different migration velocities is reflector mislocation in depth. There is also modest image defocussing, which is difficult to observe in zero-offset images. In all panels, the artefacts visible below the reflector are due to PS- and SS-wave correlations, which generate the observed crosstalk events.

While an active-source imaging formalism of passive data can be applied to processed and separated source and receiver microseismic wavefields [i.e., first arriving source $S$
Figure 2.4: (a-c) Images created with active-source data and equation (2.4). (d-f) Images created with passive data and equation (2.6). (g-i) Images created with ambient data and equation (2.6). The left, centre and right columns show each experiment calculated with the P-wave velocity error (dV) of -5%, 0% and +5%, respectively.
and scattered wavefield coda $d$, here I apply the ambient direct migration strategy [i.e., equations (2.5) and (2.6)] that does not require their explicit separation. Figure 2.4e shows the result of applying the ambient zero-offset imaging condition with the correct migration velocity ($dV=0\%$) and stacking over all 50 passive seismic records. In this case, the reconstructed $R_+$ and $R_+$† wavefields coincide and correctly image the reflector. However, there are spurious events imaged above the reflector, which again are due to crosstalk between the different wave modes. Figures 2.4d and 2.4f show the passive imaging results using a $dV$ of -5% and +5%, respectively. The characteristics of the passive images in Figures 2.4d–f are similar to the active-source images (Figures 2.4a–c). The major difference is that the S-wave crosstalk now occurs in the near surface, which is expected because no processing was applied to remove S-wave information from the upgoing wavefield.

I now apply the same ambient direct migration technique and zero-offset imaging condition to the 128 modelled ambient data windows. Figure 2.4h shows the image created using $dV=0\%$. Evidently, ambient P-wave arrivals are able to correctly image the undulating discontinuity. Figures 2.4g and 2.4i show the ambient images when using $dV$ of -5% and +5%, respectively. These images show both the mislocation and misfocussing of the undulating discontinuity when compared to Figure 2.4e. A key observation is that the ambient image has the same character as the active- and passive-source images. There is visible crosstalk, though at lower amplitudes compared to those of the imaged reflector. The most prominent difference between the ambient and the other image types is the lack of the source character in the former. This is because the active and passive experiments use a realistic broad-band wavelet, while the ambient images are generated from the aforementioned filtered random wavefield states. Overall, Figure 2.4 demonstrates that the active-source and ambient direct migration algorithms (under idealised conditions and using the same velocity model) are capable of generating similar images. However, it is difficult to know whether the images are kinematically accurate without examining their characteristic reflector moveout in EIC gathers.

### 2.5.4 EIC Gathers

An EIC in either equation (2.4) or equation (2.6) can be applied with any combination of lags ($\lambda$) in any spatial axis ($\lambda_x$ or $\lambda_z$) or as a time lag ($\tau$) in the time axis; however, for demonstration purposes, I show only the horizontal lag, $\lambda_x$. In active-source seismic experiments, reflector moveout over $\lambda_x$ in EIC gathers is well understood (Yang and Sava, 2010) and is commonly used in migration velocity analysis (MVA). Figure 2.5 presents active-source EIC gathers extracted at the anticline (Figures 2.5a–c) and syncline (Figures 2.5d–f) locations (see Figure 2.2). Figures 2.5b and 2.5e show EIC gathers from the two locations when using $dV=0\%$. Note that the energy largely focusses at the correct depth and at zero horizontal lag ($\lambda_x=0$). EIC gathers created with a $dV$ of -5% (Figures 2.5a and 2.5d) show the expected “frown” moveout, with a non-trivial amount of energy focused at non-zero $\lambda_x$. The EIC gathers created with a $dV$ of +5% (Figures 2.5c and 2.5f) show the expected “smile” moveout, similarly at non-zero lags. Comparing the asymmetry of imaged reflectors
Figure 2.5: 2D EIC gathers for the active-source experiment extracted at the noted anticline (a-c) and syncline (d-f) locations in Figure 2.2 for dV of -5% (a and d), 0% (b and e) and +5% (c and f).
Figure 2.6: 3D EIC gathers at the anticline location in Figure 2.2 for the (a) active-source and (b) ambient experiments. Each plane on the third axis represents a different $dV$ ranging between ±5% error in 1% increments.

in different panels, the reflection moveout character is evidently influenced by geologic structure.

To better illustrate the characteristic moveout of fast and slow velocities, I generate a set of EIC gathers at a sampling density finer than ±5%. Figure 2.6a shows a 3D gather where the third axis is $dV$ (in percent), after migrating the active-source data 11 times with $dV$ ranging from -5% to +5% in 1% increments. This visualisation shows that image location and moveout are both altered with changing $dV$. The resulting reflection moveout surface in Figure 2.6a can be used to identify velocity inaccuracy (Sava and Fomel, 2003; Tisserant and Biondi, 2005; Yang and Sava, 2010; Mulder, 2014). The deeper S-wave artifacts also relocate with different $dV$, but do not change the “tightness” of focus, implying that the migration velocity is far from correctly imaging S-wave modes.

Figures 2.7a–c show ambient EIC gathers at the anticline location for $dV$ of -5%, 0% and +5%, respectively. The reflection moveout characteristics exhibited at the anticline location under the fast and slow perturbations (Figures 2.7a and 2.7c) are similar to those in the active-source example. The same is true at the syncline location (Figures 2.7d–f). While some noise correlations persist in the ambient EIC gathers after stacking 128
Figure 2.7: 2D EIC gathers for the ambient experiment extracted at the noted anticline (a-c) and syncline (d-f) locations in Figure 2.2 for dV of -5% (a and d), 0% (b and e) and +5% (c and f).
windows, the imaged undulating reflector is the dominant feature in every EIC gather. S-wave correlations are not as dominant in the ambient EIC gathers as they are in the active-source example, further enhancing the imaged ambient P-wave reflection.

Figure 2.6b presents a 3D ambient EIC gather to parallel the active-source EIC gather shown in Figure 2.6a. The reflection moveout character in $\lambda_x$-$dV$ space is a “cross” focussed at $[\lambda_x, dV] = [0 \text{ m}, 0\%]$, which is very similar to the active-source result. Thus, the resulting moveout surface in Figure 2.6b could be used to identify velocity errors in the same way as that in Figure 2.6a. The S-wave correlations are apparent in both of these 3D EIC gathers, which further demonstrates the similarity between the active-source and ambient EICs for these idealised scenarios.

![Figure 2.6b](image_url)

**Figure 2.6b:** Comparison of the total stacked image using ambient recordings and the passive imaging condition after stacking: (a) four windows, (b) 16 windows, (c) 64 windows and (d) 128 windows.
2.5.5 Image Convergence

With ambient migrated images it is important to quantify how many data windows ($N_w$) are required to generate a stable result. Intuitively, the reflector evident in Figure 2.4h (a stack of 128 windows migrated with the correct velocity) would be less prominent if fewer windows were used for imaging. However, an important question remains: how low can this number be pushed? Figure 2.8 presents a test of a number of stacked windows to illustrate the concept of image convergence. Figure 2.8a shows the result when migrating only four ambient data windows and illustrates that there is an insufficient S/N to correctly interpret the peak correlation at the discontinuity. With 16 windows, (Figure 2.8b), the reflector becomes more apparent, but spurious correlations remain that could lead to misinterpretation. Quadrupling the number of windows to 64 (Figure 2.8c) generates an image closer to the expected result that is qualitatively converging. Doubling the amount of windows again to 128 (Figure 2.8d) shows little difference from Figure 2.8c, implying

![Convergence Curve](image1.png)

![Convergence Curve Gradient](image2.png)

**Figure 2.9:** (a) Normalised convergence curve for the ambient image with the correct velocity from Figure 2.4h. (b) Absolute value of the gradient of the normalised convergence curve. Each of the panels in Figures 2.8a–d is marked as a vertical line, showing that the RMS decreases as the image stability increases.
that the image has largely converged.

I now use equation (2.8) to provide a quantitative measure of image convergence. I note that with the introduction of each successive window, the semilog RMS curve of $I_{var}$ (Figure 2.9a) appears to be asymptoting toward a stable result. As expected based on the images in Figures 2.8c–d, little difference is noted between the RMS values between 64 or 128 windows. The convergence curve gradient, plotted as a semilog graph in Figure 2.9b, shows that the general rate of convergence decreases asymptotically (though with some jitter). In general, this means that when the coherent noise is sufficiently small, the crosstalk terms [Figures 2.1e and 2.1f and equation (2.8)] will destructively interfere when including additional windows, thereby improving the image until reaching convergence. Once the image has converged, no matter how many more windows are added, the image will not be significantly modified - unless the underlying noise characteristics are significantly altered.

### 2.6 Experiment II: Overthrust Model

While the previous experiment presents a proof of concept that the ambient direct migration method is sufficient for imaging a single idealised reflector with results comparable to active-source images, it is important to show how well the methodology works when applied to a more complex model. Accordingly, I employ the previously discussed active and ambient modelling and imaging methods on a 2D subset of the SEG/EAEG Overthrust model (Aminzadeh et al., 1994, 1995). The chosen 2D subset has dimensions $[nx, nz] = [800, 176]$ with 25 m grid spacing in both dimensions. P-wave velocities in this section range between $V_P = 2.4 - 6.0$ km/s, with $V_S = V_P / \sqrt{3}$. Density ranges between $\rho = 1.15 - 2.65$ g/cm$^3$.

Figure 2.10a shows the 2D density model and indicates the active shots (blue Xs), receivers (red dots) and EIC gather locations (white lines).

To generate an active-source seismic survey, I model 181 shot locations with 100 m spacing using the previously described 2D elastic modelling code. Using these data and the zero-offset imaging condition in equation (2.4), I construct an active-source image using the correct migration velocity $dV = 0\%$ (Figure 2.10b). Similarly, I model 128 ambient windows of 4 s duration using similar initial wavefield state conditions as described in the previous experiment. Using the ambient data and equation (2.6), I construct an ambient acoustic image again using $dV = 0\%$ (Figure 2.10c). Compared to the sinusoidal model, the ambient image is less clear with lower interpretability, which is perhaps expected owing to the more complex wave paths due to significant structural heterogeneities. However, the overall structure and reflectivity patterns are evident in the ambient image and remain interpretable to a certain degree.

To gain a better understanding of the image, I look at EIC gathers extracted at two model locations: one through the thrust-fault region (left), and one through the more straightforward 1D geology (right). The goal is to identify the characteristic reflection moveout in the same way as in the previous experiment. Figure 2.11 shows the ambient EIC gathers for the active-source experiment in Figure 2.10b at the two locations (shown in Figure 2.10a) for $dV = -5\%$, $dV = 0\%$ and $dV = +5\%$. Figures 2.11b and 2.11e show the EIC
2.6. EXPERIMENT II: OVERTHRUST MODEL

Figure 2.10: (a) Overthrust density model showing the receivers at every location (red dots), active source locations (blue Xs) and the two locations for EIC gather extraction (white lines), where the left and right lines are the thrust-fault and 1D geology locations, respectively. (b) Active-source and (c) ambient Overthrust zero-offset images using correct migration velocity.
Figure 2.11: Overthrust 2D EIC gathers for the active-source experiment in the (a-c) thrust-fault [left line in Figure 2.10a] and (d-f) the 1D geology [right line in Figure 2.10a] locations for $dV$ of -5% (a and d), 0% (b and e) and +5% (c and f).
2.6. EXPERIMENT II: OVERTHRUST MODEL

Figure 2.12: Overthrust 2D EIC gathers for the ambient experiment for the (a-c) thrust-fault [left line in Figure 2.10a] and (d-f) 1D geology [right line in Figure 2.10a] locations for dV of -5% (a and d), 0% (b and e) and +5% (c and f).
gathers generated using $dV=0\%$, which are well focussed at correct discontinuity depths at $\lambda_x=0$. The EIC gathers generated with $dV=-5\%$ (Figures 2.11a and 2.11d) show the characteristic “frown” moveout. The same is true for Figures 2.11c and 2.11f, where the “smiles” are clearly evident in both panels. Overall, Figure 2.11 shows the expected results when applying the active-source direct migration approach to active-source data.

Figure 2.12 presents the ambient EIC gathers extracted at the same two thrust-fault (Figures 2.12a–c) and 1D geology (Figures 2.12d–f) locations as in Figure 2.11. Figures 2.12b and 2.12e show EIC gathers generated using $dV=0\%$, and are fairly well focussed at $\lambda_x=0$. The reflection character at shallower depths in these ambient EIC gathers mimic the reflections observed in Figures 2.7b and 2.7e. This indicates that even in the more complex model, this characteristic focussing can be used to identify reflections in ambient EIC gathers. Figure 2.12a shows less obvious characteristic “frown” moveout than I would expect when using $dV=-5\%$, especially in deeper regions. This is perhaps not surprising considering the complex structure at the thrust-fault EIC gather location. However, in Figure 2.12d the characteristic “frowns” are more evident, suggesting that in less complex model regions it is easier to identify characteristic reflector moveout in ambient EIC gathers. The relative interpretability is lower for Figure 2.12c, where the “smiles” are not as clearly visible in the thrust-fault location, but are more evident in Figure 2.12f where the geology is closer to 1D.

As previously, I generate a set of 3D EIC gathers with a denser $dV$ sampling to better illustrate the characteristic $\lambda_x-dV$ reflector moveout. Figure 2.13a shows a 3D EIC gather where the axes are the same as in Figure 2.6. This panel shows that image location and reflector moveout are both altered with the changing migration velocity. Effects from S-wave correlations are not seen as in the previous model, which is likely due to the more complex Overthrust model structure. Figure 2.13b presents a 3D ambient EIC gather showing that the reflector moveout in $\lambda_x-dV$ space retains the “cross” shape centred about $[\lambda_x, dV] = [0 \text{ m}, 0\%]$, but not as strongly correlated as in the previous experiment. The correlation is still evident, though weaker, at deeper reflection points, highlighting some of the challenges associated with interpreting ambient images. It is interesting to note that the focal point in the active-source EIC gathers (Figures 2.11b and 2.13a) is at the zero-crossing at $[\lambda_x, dV] = [0 \text{ m}, 0\%]$, while in the ambient EIC gathers (Figures 2.12b and 2.13b) the focal point is at the peak energy. This suggests the importance of accounting for differences in the source signatures for various image types. The 3D EIC gathers computed in the 1D geology (Figures 2.13c and 2.13d) exhibit a characteristic reflector moveout and zero-lag correlations similar to, but more prominent than, those from the thrust-fault location. Finally, I note that interpreting 3D EIC gathers is significantly easier than interpreting only 2D slices, as illustrated by comparing Figure 2.13b with Figures 2.12a–c and Figure 2.13d with Figures 2.12d–f. Thus, I recommend migrating with a range of velocities and generating 3D EIC gather panels to help one scan the resulting volume for moveout characteristic of body-wave reflections.

Though the ambient images and their corresponding EIC gathers are not as easily
Figure 2.13: Overthrust 3D EIC gathers as in Figure 2.6. (a) Active-source and (b) ambient examples for the thrust-fault location shown in left line of Figure 2.10a. (c) Active-source and (d) ambient examples for the 1D geology location shown in right line of Figure 2.10a.
Figure 2.14: (a) Normalised convergence curve for the Overthrust ambient image with \( dV = 0\% \). (b) Absolute value of the gradient of the normalised Overthrust convergence curve.

interpreted as those from the simpler sinusoidal model, it is interesting to examine whether the image convergence in this experiment has similar results to the previous model. The ambient image convergence curve, calculated using equation (2.9) and shown in Figure 2.14a, shows similar asymptotic behaviour as the previous experiment. The convergence curve gradient, shown in Figure 2.14b, also exhibits similar character. This indicates that even in more complex models, it should be possible to use convergence curves and their gradients to identify an optimal number of windows. This implies that image convergence could be a valuable tool for identifying when an ambient image is unlikely to improve further through inclusion of additional windows.

2.7 Discussion

The synthetic ambient EIC gather examples presented here demonstrate that the \( \lambda_x - dV \) reflector moveout characteristics are similar to those noted in active-source EIC gathers.
Because of the similarity in the images in all experiments, ambient EIC gathers may be considered comparable to their active-source counterparts (under idealised scenarios) and potentially useful for similar analyses, such as migration velocity analysis. Ambient EIC gathers show that the characteristic reflector moveout is evident even in more complex structure (e.g., Overthrust). Thus, exploring a well-sampled $\lambda_x$-$dV$ space in a fully converged ambient image (i.e., Figures 2.6b, 2.13b and 2.13d) would provide a stringent test, as well as a tool, for validating any imaged ambient reflectivity. Events lacking the characteristic P-wave reflector moveout are more likely related to coherent noise sources. Unless these coherent noise sources are removed through careful processing, they will still be imaged and likely appear as spurious reflectivity, potentially causing misinterpretation. Thus, I assert that applying a 3D EIC gather test through an ambient direct migration approach could improve the reliability and interpretability of ambient images.

Finally, in these idealised scenarios I recognise that no dominant coherent noise sources (e.g., blasts, traffic, resource production activities) were added, which would be expected in ambient recordings. When coherent noise is prevalent in ambient recordings and not removed through processing, there are two possible outcomes: (1) the image will not converge due to the inclusion of many stronger coherent noise sources, or (2) the image will converge, but be dominated by non-body-wave energy that either images the location of coherent noise generation or is misinterpreted as a dominant crosstalk reflector. In either scenario, it would be judicious to process the ambient records prior to applying the direct ambient migration and 3D EIC gather framework (Girard and Shragge, 2016, 2019b).

### 2.8 Conclusions

I demonstrate that a direct migration method modified from active-source seismology represents a viable approach for ambient imaging. When the coherent noise component in ambient seismic data is sufficiently low, the direct migration approach images body-wave reflections in the similar manner as active-source imaging, assuming a converged image and sufficiently balanced illumination. The ambient direct-migrated extended imaging condition (EIC) gathers developed herein have similar character to those created during active-source migration, even for more complex geological scenarios. Therefore, as with their active-source counterparts, ambient EIC gathers can be used as a tool for interpreting and validating imaged body-wave reflections.

The advantages of this ambient EIC approach over other ambient imaging approaches are threefold. First, there is no need to identify or separate source information or to use interferometric methods to explicitly synthesise a source wavefield prior to imaging. Second, if the coherent noise component is sufficiently small, when using the correct migration velocity a converged ambient EIC gather will show reflectivity focused at the correct locations and exhibit characteristic reflector moveout while non-reflection correlations (i.e., spurious and coherent noise) will not. Third, scanning the $\lambda_x$-$dV$ dimension of 3D EIC gathers represents a tool for interpreting and validating of any imaged ambient body-wave reflectivity.
The major caveat not addressed herein is the importance of improving P-wave signal-to-noise ratio (S/N) and balancing illumination prior to migration for scenarios where strong surface waves and other coherent noise events are present. In these experiments, there were only weak surface waves and no other added sources of coherent noise to cause spurious correlations. Thus, before applying this method to field data, it would be important to implement a careful processing workflow to mitigate these issues.

Acknowledgements

I thank the Australian Government support of this project through an “Australian Government Research Training Program Scholarship”, The University of Western Australia (UWA) Top-up scholarship program, as well as the UWA Centre for Energy Geoscience and David Lumley for their continued support of this project. Data processing was made possible through the reproducible open-source package Madagascar (www.reproducibility.org). Compute resources were generously provided by the Center for Wave Phenomenon at the Colorado School of Mines and the UWA Centre for Energy Geoscience.
Chapter 3

Automated Processing Strategies for Ambient Seismic Data

“...when you have eliminated the impossible, whatever remains, however improbable, must be the truth.”

Sherlock Holmes in *The Sign of the Four* by Sir Arthur Conan Doyle

3.1 Introduction

Acquiring long-time continuous recordings on fixed seismic arrays is an increasingly common strategy for measuring the data required for understanding Earth structure and physical properties. Historically, fixed arrays have been used more in academic seismology experiments aimed at imaging larger-scale crustal and/or lithospheric structure and inverting for deep Earth properties using passive earthquake [e.g., teleseismic (Bostock et al., 2001; Abe et al., 2007)] or ambient (Moschetti et al., 2007; Bensen et al., 2007) data. However, with greater channel counts and denser acquisition geometries (i.e., so-called large-N recordings), fixed-array or nodal data acquisition increasingly is being used for imaging and inversion purposes at finer length scales. Examples range from imaging oil and gas reservoirs (Draganov et al., 2009, 2013) and monitoring carbon sequestration (Xu et al., 2012; Ugalde et al., 2013) to assisting with mineral exploration (Cheraghi et al., 2015; Olivier et al., 2015) and improving the understanding of seismic hazards (Czarny et al., 2016; Olivier and Brenguier, 2016).

The success of imaging and inversion methods based on long-time recordings is strongly contingent upon the ability to extract or enhance the desired signal content while removing or mitigating coherent noise events detrimental to imaging and inversion processes. However, differentiating between what is signal and what is noise in long-time recordings depends on the particular imaging or inversion problem at hand. For example, signal in passive
seismic problems can be noise in ambient seismic problems (and vice versa); thus, I must
define each term in relation to the survey objectives.

Herein, I use the term ‘passive’ to denote signals from high-energy events [relative to
background root-mean-square (RMS) energy] that have a reasonably determinable spatial
and temporal origin and occur within a predictable volume with an expected magnitude
range and frequency content. Passive seismic events are measured by most, if not all,
stations in a large-N array, and can be separated into individual windows usually with a
relatively high signal-to-noise ratio (S/N). Windowed passive events commonly are used
to determine elastic properties and source parameters. Examples include microseismic
monitoring of fracing operations (Eisner et al., 2006; Maxwell, 2014; Witten and Shragge,
2017), deployments of broad-band arrays for recording earthquakes for crustal lithospheric
structure and hazard studies (Rondenay et al., 2001; Schmandt and Clayton, 2013), and
arrays to monitor for subsurface changes in volcanic environments (Sens-Schönfelder and
Wegler, 2006; Brenguier et al., 2008).

Conversely, I use the term ‘ambient’ to denote deterministic signals contained in the
background wavefield energy that neither have a predictable nor necessarily determinable
source location, signature, magnitude nor frequency content. Terrestrial long-period
ambient seismic recordings include all wave types (i.e., P-, S-, Rayleigh and Love) with
energy incident upon a recording array potentially from any direction - though rarely
evenly distributed over solid angle. While ambient recordings likely include energy from
specific deterministic sources (e.g., microseismicity), they are not dominated by passive
events over the entire recording period. Passive events in long-time recordings are useful
for passive investigations; however, in the context of ambient analysis, I consider them
to be coherent noise sources to be mitigated or ideally eliminated from the ambient data
volume.

Many published ambient seismic experiments involving long-time recordings focus on
surface waves acquired in high-energy environments. These experiments generally employ
interferometric methods (Shapiro et al., 2005; Gerstoft et al., 2006; de Ridder, 2014) based
on ‘cross-correlation plus stack’ (CC+S) processing of ambient waveforms measured on
receiver pairs. Examples of investigations employing a CC+S strategy include using long-
term seafloor installations to determine the near-surface S-wave velocity ($V_S$) structure
(Landèş et al., 2009), employing permanent land arrays in areas with heavy urban vehicle
traffic to invert for $V_S$ structure of the top 300 m (Nakata et al., 2011), and using shear- and
surface-wave-dominated earthquake coda measured on temporary arrays to characterise
the underlying lithosphere (Ruigrok et al., 2011).

Body-wave reconstruction through a CC+S approach is significantly more difficult than
its surface-wave counterpart (Roux et al., 2005; Forghani and Snieder, 2010). Ambient
imaging experiments aiming to recover weaker body-wave energy generally use interfero-
metric methods (Nakata et al., 2011; Xu et al., 2012; Draganov et al., 2013; Nakata et al.,
2015). While the results are promising in areas exhibiting persistent and relatively high-
amplitude body-wave energy (i.e., urban environments), these high-energy environments
3.1. INTRODUCTION

do not always coincide with the chosen area of investigation. After recovering body-wave energy, the virtual source method (Bakulin and Calvert, 2006) can be used to synthesise the equivalent of active-source seismic data. To a lesser extent, some investigations apply various body-wave analyses to higher-frequency ambient data (>40 Hz) with reasonable results (Draganov et al., 2013; Nakata et al., 2015; Edne and Halliday, 2016).

The goal of body-wave CC+S processing is to extract the Green’s function and fully reconstruct the reflection response of the medium as measured between ‘master’ and ‘auxiliary’ receiver pairs (Wapenaar, 2004). Ideally, the Green’s function encapsulates all wavefield scattering in the medium underlying the two receivers (as if the master receiver were an active source) that would arrive within the trace record length. Thus, CC+S panels serve as proxies for Green’s functions and are useful for producing images of Earth structure and inverting for material properties without specific knowledge of the location or timing of source energy generation (Lobkis and Weaver, 2001; Campillo and Paul, 2003). This approach works optimally when body-wave signal arrives at an array equally from all directions (i.e., equipartitioned) and there are no coherent noise events to dominate the overall stack (Artman, 2007). However, in practice these assumptions are never realised since wavefield directivity is more commonly biased toward a few incidence angles and backazimuths related to persistent energy sources. When ambient body-wave S/N is low, summation over long time periods enhances the coherent signal via the power of stack while down-weighting non-stationary noise events; however, this process by itself does not address directionality issues.

Applying CC+S methodology to raw ambient seismic data to enhance body-wave energy has (at least) one major drawback: once the CC+S operation combines energy from different windows, it is difficult or impossible to remove large-amplitude coherent noise events because stacking is an irreversible operation. A common strategy for eliminating these noise sources is to employ a window-selection process that identifies and removes (or down-weights) noisy windows prior to CC+S processing. This can be achieved, e.g., by manually or automatically eliminating windows in the time domain (Nakata et al., 2015) or through a sign-bit normalisation (Oren and Nowack, 2017). While the former approach excludes windows with undesirable energy from the stack, this strategy has at least two potential drawbacks: [1] each removed window may have useful body-wave signal; and [2] removing too many windows may leave insufficient energy for the CC+S process to converge to a stable result.

An alternative strategy would be to automatically process ambient data to filter out coherent noise in the uncorrelated domain prior to applying CC+S. This approach offers three key benefits. First, processing ambient data directly in the uncorrelated prestack domain is more effective than post-processing of CC+S pseudo-shot gathers at mitigating broadband (e.g., drilling, fracturing, blasting) and narrowband (e.g., electrical generators) coherent noise events. A second benefit is partially addressing wavefield directivity issues measured in ambient records. Whether or not strong coherent noise events are isolated in time, they often arrive from a limited range of incidence angles and backazimuths.
Removing or down-weighting directionally dominant coherent noise events significantly helps to generate a data volume with more balanced wavefield energy directivity. Third, applying a fully automated processing and quality control (QC) workflow is important considering that the human and computational effort required to handle the growing ambient data volumes. While automated processing and QC techniques have been explored previously in the context of microseismic event picking (Tselentis et al., 2012) and velocity tomography (Leontarakis et al., 2015), surface-wave tomography (Orfanos et al., 2016), and shear-wave velocity structure delineation (Giannopoulos et al., 2017), no examples are to be found within the literature with respect to automated processing aimed at enhancing ambient body-wave arrivals.

Herein, I investigate whether automated data processing of ambient seismic recordings in the prestack domain by removing stronger coherent noise sources can enhance weaker body-wave energy. To devise an effective ambient data processing strategy, I first classify large-magnitude coherent noise events in one of three categories: [1] localised in time where high-energy burst-like events occur within a single time window; [2] localised in space where energy persistently originates from a specific location and contributes significantly to CC+S panels, but with spatially localised wavefield illumination; and [3] localised in frequency where persistent monochromatic energy, usually from one or more narrowband sources, causes ringing in time-domain CC+S panels. To mitigate these three types of coherent noise I develop and apply an automated four-step processing methodology. I first automatically mask residual high-amplitude events that are coherent across the array. As a second step, I use a window selection process, similar to Nakata et al. (2011), to remove windows with RMS energy greater than a chosen threshold. However, I am less aggressive in my window selection because my first step already masked out the highest amplitude events. The final two steps involve removing residual burst-like energy on a trace-by-trace basis in the time domain (Claerbout, 2014), and applying a related frequency-domain debursting operator to mitigate strong stationary monochromatic noise. By applying this prestack processing workflow I demonstrate convergence improvements in processed data and assert that the resulting ambient data volume is better conditioned for ensuing imaging and migration activities.

I evaluate my automated ambient data processing workflow on an ambient data set recorded in March 2013 by the Geological Survey of Canada at the Lalor Lake mine site in northern Manitoba, Canada. I begin by describing the Lalor Lake survey site and the acquired ambient data set. Based on the data characteristics I then develop my prestack processing workflow used to mitigate coherent noise events and help balance wavefield directivity. Because each processing step modifies the data set, I also describe and apply a sequence of quality control (QC) measures that help us quantitatively assess whether the automated workflow procedure is converging toward a stable result. The manuscript concludes with discussion on the value of applying a comprehensive ambient data processing workflow, computing a range of QC metrics, and measuring the convergence rates of CC+S processes.
3.2 Lalor Lake Data Set

The Lalor Lake survey site is situated in northern Manitoba, Canada on the Canadian Shield (see inset, Figure 3.1). This stable cratonic setting is characterised by very thin (<5 m) regolith overlying granitic basement rocks with a very fast bulk P-wave velocity ($V_P=6.0$ km/s) known from a co-located active-source seismic experiment (Bellefleur et al., 2015). Mining interests in the area focus on volcanogenic massive sulphide (VMS) deposits containing economic quantities of gold, silver, copper and zinc. The rock units hosting such deposits can exhibit significant contrasts in acoustic impedance, which often generate prominent reflectivity commonly observed in hard-rock active-source seismic experiments (Dentith and Mudge, 2014). For a more detailed geological background of Lalor Lake and the surrounding region, the reader is referred to Bellefleur et al. (2015), Cheraghi et al. (2015) and Miah et al. (2015).

The Lalor Lake mine site is located far from major population centres, with the nearest small town (Snow Lake, pop. 700) situated 8 km north-east of the survey area and the largest regional centre (Flin Flon, pop. 5,000) located by road 215 km to the west. The only nearby transportation facilities are a single gravel road designated for mine site traffic (south of the array in Figure 3.1), rarely used otherwise, and a seldom-used railroad track 6.5 km to the south-east. No other industrial activity occurs within an 8 km radius of the mine site, with the area being largely untouched boreal forest and lakes. Thus, there are unlikely to be any significant anthropogenic sources of seismic energy other than the mine site itself.

The Lalor Lake ambient seismic survey was acquired by the Geological Survey of Canada (GSC) in early March 2013 during the Canadian winter when surficial water sources were frozen to a significant depth. Ambient seismic data were acquired on an array of 336 vertical-component 10 Hz OYO GSR Geospace geophones, deployed in a $2 \times 2$ km$^2$ area with 100 m inline and 360 m to 400 m cross-line station spacing (directions A and B, respectively, in Figure 3.1). The total recording time was 227 hours at a 2 ms sampling rate.

The seismic array, situated largely over the active Lalor Lake mine site, was actively recording during shaft construction, which occurred at about 0.8 km depth. There were at least two underground detonations per day, as well as transportation activity from heavy vehicles removing extracted material via the main access road. Significant coherent monochromatic noise is also evident in the seismic recordings, likely due to generators providing on-site electricity and electro-mechanical vibrations of mining-related infrastructure. Other mining-related coherent noise events are present throughout the data volume (e.g., blasting, drilling, extracted material transport, and mine-induced microseisms).

The Lalor Lake data set proved to be a judicious choice for ambient body-wave research for a number of reasons: [1] the recording array is situated on the Canadian Shield (with thin regolith) with a tectonic history favourable for observing seismic reflectivity; [2] a remote location with low anthropogenic noise levels other than mine-site-related energy;
CHAPTER 3. AUTOMATED PROCESSING

[Image 73x506 to 466x785]

Figure 3.1: Map showing the ambient seismic survey area and its location in Manitoba, Canada (inset). Filled circles show geophone locations with those in yellow forming the 2D sub-array for CC+S processing described herein. The red star shows the location used to calculate power spectral density (PSD) plots. The mine site (left side) access road (south) and connecting main road (lower right) are visible on the map.

[3] a large-N array with an appropriate receiver density (336 channels within a 2x2 km$^2$ area) and sufficient aperture to capture reflected ambient wavefield energy at a range of incidence angles and azimuths; [4] hibernal data acquisition that precludes local marine wave action as a source of coherent noise energy and leads to excellent coupling of receivers to the frozen earth; and [5] the availability of both an active-source 3D seismic survey (Bellefleur et al., 2015) and a virtual-source interferometry study (Cheraghi et al., 2015) for comparison purposes.

3.3 Extracting Body-wave Signal From Ambient Data

To illustrate the challenges of recovering body-wave signals in the presence of coherent noise events like those observed in the Lalor Lake ambient data set, it is instructive to develop a data model. I consider a model where the observed vertical particle velocity $v_{in} = v_{in}(x_i, n, t)$ for a trace at location $x_i$ in window $n$ with relative time $t$ is the superposition of a supposed body-wave signal, $S_{in} = S_{in}(x_i, n, t)$, and unwanted noise events, $N_{in} = N_{in}(x_i, n, t)$,

$$v_{in} = S_{in} + N_{in}, \quad i = 1, N_r \quad \text{and} \quad n = 1, N_w,$$

(3.1)
3.3. EXTRACTING BODY-WAVE SIGNAL FROM AMBIENT DATA

where \( N_r \) is the number of receivers, and \( N_w \) is the total number of windows. I further refine my noise model by splitting \( N_{in} \) into a combination of coherent noise, \( N^{C}_{in} \), highly correlated between two stations, and random/incoherent noise, \( N^{R}_{in} \), assumed to represent a Gaussian process ideally uncorrelated between traces:

\[
v_{in} = S_{in} + N^{C}_{in} + N^{R}_{in}.
\]  

(3.2)

An approximate frequency-domain (\( \omega \)) Green’s function, \( G_{ij}(x_i, x_j, \omega) \), between receiver locations \( x_i \) and \( x_j \) may be extracted by cross-correlating frequency-domain wavefield traces \( v_{in} \) and \( v_{jn} \) and summing over all windows (Wapenaar, 2004). For my model, this is given by:

\[
G_{ij}(x_i, x_j, \omega) \approx \frac{1}{N_w} \sum_{n=1}^{N_w} v_{in} v_{jn} = \frac{1}{N_w} \sum_{n=1}^{N_w} (S_{in} + N^{C}_{in} + N^{R}_{in})(S_{jn} + N^{C}_{jn} + N^{R}_{jn}),
\]  

(3.3)

where \( \overline{v} \) represents the complex conjugate of \( v \). When \( N_w \) is large, CC+S processing generally sees the random/incoherent noise contribution \( (\sum_{n=1}^{N_w} N^{R}_{in} N^{R}_{jn}) \) tend to zero. However, if the stacked coherent noise correlations are large in magnitude relative to the stacked correlated body-wave signal, the former will dominate the overall stack:

\[
\lim_{N_w \to \infty} G_{ij}(x_i, x_j, \omega) \approx \frac{1}{N_w} \sum_{n=1}^{N_w} N^{C}_{in} N^{C}_{jn}.
\]  

(3.4)

Conversely, where the stacked coherent noise correlations do not dominate persistent stacked correlated signal components, either naturally or through data processing, the CC+S process works to enhance the latter:

\[
\lim_{N_w \to \infty} G_{ij}(x_i, x_j, \omega) \approx \frac{1}{N_w} \sum_{n=1}^{N_w} S_{in} S_{jn}.
\]  

(3.5)

Even for CC+S scenarios where equation (3.5) holds and one has recovered a suite of interferometric shot gathers, introducing a single high-energy coherent noise event in the \( p^{th} \) window recorded on both receivers (i.e., \( N^{C}_{ip} \) and \( N^{C}_{jp} \)) into the overall stack could skew the results:

\[
\lim_{N_w \to \infty} G_{ij}(x_i, x_j, \omega) \approx \frac{N^{C}_{ip} N^{C}_{jp}}{N_w} + \frac{1}{N_w} \sum_{n=1}^{N_w} S_{in} S_{jn} \approx \frac{N^{C}_{ip} N^{C}_{jp}}{N_w}.
\]  

(3.6)

Thus, care must be taken to forestall introduction or minimise coherent noise events to meet S/N requirements. I aim to achieve this by developing and applying a data processing workflow to uncorrelated ambient prestack data.
3.4 Processing Methodology

This section develops an automated processing workflow applicable to ambient seismic recordings in the uncorrelated domain. Analogous to active-source seismic data processing, my multi-stage workflow involves testing and validating a sequence of filtering steps and associated QC measures. Each processing tool represents a modification of a standard processing technique that is generally neither new nor purpose built. However, due to the large data volumes acquired in ambient surveys (here >700 Gb), the presence of time windows with strong mine-related activity, and the lack of impulsive shot records to QC data directly, I deemed that an automated data-processing approach was necessary to meet the processing objectives within a reasonable time frame.

3.4.1 Processing workflow

My automated processing workflow aims to automatically identify and mitigate large-amplitude short-time events (e.g., mine blasts), other lower-amplitude burst-like energy (e.g., mine traffic), as well as any strong monochromatic energy (e.g., electro-mechanical noise at fundamental and harmonic frequencies) that may be present in the ambient recordings. Figure 3.2 depicts the processing workflow including the computed QC outputs;

![Workflow Diagram](image)

**Figure 3.2:** Workflow showing the processing sequence steps and computed QC outputs: (1) global PSD plots; (2) hourly and global CC+S panels; (3) hourly RMS convergence plots; and (4) hourly beamsteering plots.
I defer providing specific implementation details to the Processing section below. The first processing step is designed to eliminate high-energy events considered part of the coherent noise ($N_{co}$ in equation (3.2)). For such windows, I define a masking operation that locates statistically high-energy events representing consistent arrivals across the array. I apply a sliding window operator to smoothly mask out these events. Even though this step removes some recorded data, it prevents the resulting windows from being dominated by high-energy seismic events and retains a significant percentage of unmasked data that would be otherwise rejected.

As a second step, I remove windows with residual high RMS energy amplitudes not handled by the masking operation. To do so, I compute short- versus long-term averages (McEvilly and Majer, 1982) to prioritise high-energy windows for removal. Based on this information, I define a global magnitude threshold to eliminate windows with abnormally large RMS energy values. This might mean that some hours have no remaining windows, which is different than introducing a local threshold that results in each hour maintaining a consistent number of windows. I do this because of diurnal variations in environmental conditions (i.e., mine work shifts, weather) that cause some hours to exhibit relatively high noise levels.

After mitigating high-energy coherent noise events, the third step addresses burst-like energy that, while not contributing greatly to a window’s RMS value, would otherwise lead to spurious high-amplitude correlations. I minimise high-energy spikes using a time-domain deburring method (Claerbout, 2014) that lowers large-amplitude events within individual traces to a prescribed level. I note that this is a judicious alternative to hard clipping, which can introduce frequency-domain artefacts via discontinuous particle accelerations.

The fourth and final step mitigates strong monochromatic energy, which manifests as ringing in time-domain CC+S panels. I modify the time-domain deburring method of Claerbout (2014) to operate on Fourier magnitude spectra (leaving the original phase untouched). This filter down-weights monochromatic energy of significantly greater magnitude than the background spectrum to a user-specified level. While notch filtering is commonly applied to remove various types of monochromatic noise in active-source seismic experiments (Linville, 1994), designing non-stationary notch filters for each window and characteristic frequency makes automation impractical. In addition, notch filters can introduce artefacts by affecting phase information. My approach differs from notch filtering in that it neither requires *a priori* knowledge of the frequency structure nor designs a suite of filters to remove energy at specific peak frequencies. Debursts leaves filtered spectral magnitudes at levels commensurate with those of nearby Fourier components, which is not necessarily true with notch filtering. Removing strong localised frequency-domain energy helps to minimise coherent monochromatic noise in processed ambient data, while (largely) preserving the phase component important for reconstructing the Green’s function kinematics through CC+S processing.
3.4.2 Quality Control Measures

To provide information on the individual and cumulative effects of applying each processing step to the Lalor Lake ambient data set, I also compute and show the results of applying a suite of QC measures: power spectral density (PSD) plots, CC+S panels, RMS convergence curves, and beamsteering diagrams.

Power Spectral Density

PSD plots are useful for visualising variations in the frequency content of a large ambient data set over calendar time (Draganov et al., 2013). I use PSD plots as the primary data reconnaissance tool in this study. Figure 3.3a shows a representative PSD for unprocessed ambient data acquired at a single station (denoted by the red star in Figure 3.1) for the 227 hours of recording time. Each vertical trace represents an individual spectrum calculated from a consecutive 30 s tapered window with zero overlap. Figure 3.3b shows the PSD from Figure 3.3a after applying the masking and window selection steps, while Figure 3.4a

![Figure 3.3](image)

**Figure 3.3:** PSD plots of the entire recording time for a single receiver (red star in Figure 3.1) using roughly 30 s windows (4096 samples at 8 ms sample rate) to calculate individual spectra. The plots show (a) the entire raw data set and (b) the result of applying the data selection and masking processing steps. Blanked-out intervals represent windows excluded based on the criterion of RMS energy exceeding the established threshold value.
3.4. PROCESSING METHODOLOGY

Figure 3.4: PSD plots of the entire recording time without the blanked-out windows (solid blue in Figure 3.3b) for a single trace (red star in Figure 3.1) using roughly 30 s (4096 samples) windows to calculate the spectrum. (a) Raw data. (b) After data masking and selection. (c) After time debursting. (d) After frequency debursted.
shows the PSD after removing the excluded parts of the PSD in Figure 3.3b to assist with interpretability.

The raw PSD plots in Figures 3.3a and 3.4a contain prominent vertical and horizontal striping. Vertical lines represent high-energy events concentrated over short time intervals (e.g., mine blasts), while horizontal lines indicate persistent monochromatic energy likely attributable to the aforementioned electro-mechanical vibrations. Since both energy types violate the equipartition requirements for Green’s function extraction, they need to be mitigated without significantly distorting the underlying data characteristics.

Cross-Correlation Plus Stack Panels

CC+S panels are an effective QC tool for determining the degree to which coherent noise events are mitigated via the above processing workflow. These gathers are similarly useful when monitoring for the emergence of reflected body-wave arrivals with moveouts consistent with the known bulk P-wave velocity structure ($V_P \approx 6.0 \text{ km/s}$). Figure 3.5a shows an hour-long CC+S panel computed using equation (3.3) for 40 stations situated on the longest survey line (Figure 3.1, yellow receivers) with a master trace. The chosen hour exhibits strong burst-like energy as well as significant monochromatic ringing, likely originating from the mine site. Figure 3.5e shows a similar CC+S panel, but for a noisy hour. While the background noise correlations are more dominant than Figure 3.5a, surface-wave correlations remain the more prominent features. Thus, this corresponds with the scenario described by equation (3.4), where coherent noise dominates any ambient body-wave signal that may be present. Figure 3.5a also shows events with linear (and perhaps dispersive) moveouts with apparent velocities ranging between 2.0-2.5 km/s. This observation implies that much energy arrives in the form of shear and surface waves.

Convergence Curves

Examining the overall RMS energy in CC+S gathers after introducing each successive window helps to determine both whether the CC+S process is converging and how many windows are required to achieve convergence. To examine the rate of convergence I follow Issa et al. (2017), which defines an image variance measure that is directly comparable to the CC+S panels considered herein. The rate of convergence of a CC+S panel for master and auxiliary traces $x_i$ and $x_j$ at correlation lag $\tau$ over total accumulated time $T$ can be defined as:

$$C(x_i, x_j, \tau, T) = C_{stat}(x_i, x_j, \tau) + \frac{1}{T} \int_0^T N^2(x_i, x_j, \tau, t') dt',$$  \hspace{1cm} (3.7)

where $C_{stat}$ is the stationary (i.e., fully converged) CC+S panel that would exist in an idealised noise-free environment [i.e., equation (3.5)], noise component $N(x_s, x_r, \tau, t')$ is similar to the noise contribution in equation (3.3) (i.e., $N^C$ and $N^R$ components), and $t'$ is a dummy integration variable. Issa et al. (2017) further shows that taking the variance of
3.4. PROCESSING METHODOLOGY

Figure 3.5: CC+S panels from a 2D line of receivers (yellow circles in Figure 3.1) correlated with a single master trace for two different hours at each processing stage. The upper and lower rows represent quiet and noisy hours, respectively. (a and e) Raw data. (b and f) After data masking and selection. (c and g) After time debursting. (d and h) After frequency debursting. Red points denote samples clipped (high-energy) for visualisation purposes, which are included to highlight improvements from applying the processing workflow. The emergence of a prominent reflection event can be observed in (d) at $\pm 0.15$ s lags.
Figure 3.6: (a) RMS convergence plots for the quiet hour of CC+S panels shown in Figures 3.5a–d. As more CC+S windows are included, the RMS energy effectively decreases as inverse square of accumulated time (i.e., windows). The coloured lines denote different processing steps. Green: Raw data. Blue: After window masking and selection. Red: Time debursted. Pink: Frequency debursted. (b) Semi-log of the absolute value of the normalised convergence curve gradient, showing that the convergence rate decreases asymptotically (though with some jitter). (c) RMS convergence plots for the noisy hour of CC+S panels shown in Figures 3.5e–h. The same trend as observed in (a) is evident, though some increases occur at particular windows. More windows have been removed in the masking process, as indicated by the shorter length of the three processed curves with respect to the green line (raw data). (d) Semi-log of the absolute value of the normalised convergence curve gradient in (c), showing that the convergence rate is similar to the quieter hour, but with increased jitter.

equation (3.7) and recasting the problem to allow for \( N_w \) discrete windows yields:

\[
C\text{var}(x_i, x_j, \tau, N_w) \equiv \text{Var} (C_{\text{stat}}(x_i, x_j, \tau)) + \frac{1}{N_w^2} \sum_{n=1}^{N_w} \text{Var}(x_i, x_j, \tau, n) \text{,} \tag{3.8}
\]

where \( n \) is the window index, and \( \text{Var}(x_i, x_j, \tau, n) \) is the variance of all noise components expanded from equation (3.3) (i.e., \( \overline{N_{jn} N_{jn}} + \overline{N_{in} S_{jn}} + \overline{N_{in} N_{jn}} \)), with \( N \) comprised of both coherent and non-coherent noise components, \( N^C + N^R \). The variance calculation is normalised by the inverse square of the number of windows used to compute the cumulative RMS energy window stack. In the overall CC+S gather, \( C\text{var} \) approaches that of the ideal stationary CC+S panel, \( \text{Var} (C_{\text{stat}}) \), because when \( N \) has favourable (Gaussian) statistical properties the variance of the noise term is ideally zero in the limit \( N_w \to \infty \). However, introducing coherent non-Gaussian noise events will delay or altogether prevent CC+S panels from converging toward \( \text{Var} (C_{\text{stat}}) \).
To establish an easily interpretable QC tool, I use an hourly RMS “convergence curve” formed by stacking over the $x_i$, $x_j$ and $\tau$ variables in equation 3.8:

$$C(N_W) = \sqrt{\frac{\sum_{i,j=1}^{N_r} \sum_{\tau=\tau_{\text{min}}}^{\tau_{\text{max}}} C_{\text{var}}^2(x_i, x_j, \tau, N_W)}}.$$  

(3.9)

where $\tau_{\text{min}}$ and $\tau_{\text{max}}$ are the minimum and maximum computed correlation lags, respectively.

Figures 3.6a and 3.6c show the RMS convergence curves [equation (3.9)] on a semi-log plot [i.e., $\log_{10} C(N_W)$] for the low- and high-noise CC+S panels presented in Figures 3.5a–d and Figures 3.5e–h, respectively. The green lines denote the RMS energy of the cumulative CC+S panel as a function of $N_W$ panel windows. I note that the green lines do not decrease monotonically when adding more windows, indicating the introduction of correlations inconsistent with previous cumulative panels. These are more evident in the noisy-hour convergence curve (Figure 3.6c) where variations in the raw data result (green) show correlations inconsistent with the rest of the windows.

The convergence curve gradient [i.e., $\log_{10}|\nabla C(N_W)|$] is an effective measure of the convergence rate of the CC+S process. The observed convergence curve rates in Figures 3.6b and 3.6d decrease largely, but not exclusively, monotonically since some windows increase the $\log_{10}|\nabla C(N_W)|$ measure and presumably degrade the overall S/N of the CC+S stack. A higher amount of jitter in the green line is also evident in Figure 3.6b relative to Figure 3.6a, though this is exacerbated by the log$_{10}$ plot. Conversely, the noisy-hour convergence curve gradient (green line in Figure 3.6d) shows significantly more jitter compared to the quiet-hour convergence curve gradient (green line in Figure 3.6b).

**Beamsteering**

To find evidence of ambient body-wave arrivals and examine wavefield directivity, it is helpful to use beamsteering methods to investigate the data’s plane-wave dip spectrum (PWDS). I use a beamsteering approach (Cole, 1995) to calculate data coherency in each window when stacking over a range of plane-wave slowness components in the x- and y-directions ($p = [p_x, p_y]$). The PWDS, computed relative to a reference trace situated in the middle of the recording array, are then converted to apparent velocity and backazimuth.

The cumulative PWDS stack for all windows in one hour of recording time shows the time-averaged directivity and apparent velocity of arriving wavefield components.

Figure 3.7 shows beamsteering plots for the same data hours used in Figure 3.5. Figure 3.7a, which presents the quiet-hour raw data results, highlights energy arrivals from backazimuths corresponding to the mine site and access road locations with apparent velocities ranging between 2.0-2.5 km/s. These arrivals likely represent a combination of shear- and surface-wave contributions. I also note the presence of energy with faster apparent velocities in the stacked beamsteering plots; however, little stands out as coherent signal. Figures 3.7b–d show beamsteering plots for the same quiet hour after applying
Figure 3.7: Beamsteering plots of PWDS for the quiet (upper row) and noisy (lower row) hour data showing the apparent velocity (from 1.5 to 5.5 km/s) and directivity (360° of azimuth) of wavefield arrivals. (a and e) Raw data. (b and f) After masking and data selection. (c and g) After time debursting. (d and h) After frequency debursting. Red colours are higher amplitudes. Evidence for coherent high-velocity P-wave arrivals is observed in (d) and to a lesser extent in (h).

the masking plus selection, time- and frequency-debursting steps, respectively. The fully processed beamsteering plot in Figure 3.7d shows more focused energy at faster apparent velocities than equivalent plots at earlier processing stages. Figures 3.7e–h show a set of beamsteering plots after applying the same progression of processing steps as the overlying panels, but for the noisy hour presented in Figures 3.5e–h. These panels show the effects of heavier road traffic and documented mine construction blasts, which result in more concentrated coherent noise than in the previous set of panels (Figures 3.7a–d). I note that there is more overall energy with the strongest energy contributions again interpreted to be shear and surface waves.

3.5 Processing the Lalor Lake Ambient Data set

I now apply the above data processing workflow to the Lalor Lake ambient data set to examine how well it mitigates both coherent and incoherent noise events while otherwise preserving the quality of the underlying ambient waveforms. Each step is applied on a window-by-window basis with parameters selected to suit this particular data set. I choose windows of roughly 30 s duration (4096 samples) for two reasons: [1] to form an optimal record length for using the cuFFT package on a graphical processing unit (GPU) infrastructure (Step 4); and [2] to have sufficiently long windows to effectively minimise
event truncation. I illustrate the effects of applying each step using a sequence of figures below.

3.5.1 Steps 1 and 2: Masking and Data Selection

Due to their high-energy nature, mine blasts would dominate the CC+S process in terms of wavefield amplitude and directivity and thus must be removed. While it may be possible to use known detonation times from the mine-site blasting records to manually remove associated energy, this approach would suffer from two main drawbacks: [1] it is not automated; [2] any blast events not in the mine-site records or high-energy events unrelated to blasting would dominate the data. The automated masking workflow described above addresses these issues.

I first scan for large-magnitude events generating consistent arrivals across the array by computing and stacking the wavefield envelope of each trace in the window and using the large magnitudes contained therein to define wavefield mask filters (blue trace in Figure 3.8a). The RMS energy of the stacked envelope trace defines the time (or times) with the highest energy levels, which likely include events that would dominate any ambient body-wave signal present in the data. I subtract this trace from the stacked enveloped trace and apply a cosine taper to the first and last 150 samples to forestall introduction of edge artefacts. The 20\textsuperscript{th} and 30\textsuperscript{th} percentiles of the RMS energy of the scaled enveloped stack (blue curve in Figure 3.8b) are used to define the clip function (red curve in Figure 3.8b), which is subtracted from the scaled smoothed enveloped stack trace. This procedure results in a masking trace (pink curve in Figure 3.8b) where the top 20% of energy is set to zero, the bottom 70% is left untouched, and the intervening 10% has a cosine taper applied. The resulting mask is then multiplicatively applied to each trace in the window. The masking percent is easily modified based on data characteristics, either from examining the QC results or the peak PSD values.

The masking process removes not only blast energy, but that from other significant mine-related events, thereby reducing coherent noise contributions in the ambient data volume. Figures 3.3a–b illustrate these differences, where the blue regions in Figure 3.3b indicate windows eliminated by the threshold operation. While the effects of applying the masking operator are subtle, the observed differences between Figures 3.4a–b (PSD plots of the raw and masked data with masked times removed, respectively) are entirely due to this step. Because impulsive events are broad band, removing them tends to narrow the frequency distribution and result in more coloured spectra. I note that the remaining data still exhibit windows with much higher RMS energy levels than the overall average. This could be for a number of reasons, including mine traffic, vibrations from electrical generators, and even microseismic activity related to mining activity and post-blast relaxation. Therefore, I establish a threshold for retaining or rejecting these high-energy windows. My preferred method is to calculate the RMS energy of each window and determine a global threshold percentage (here 60%, meaning I discard 40% of the data). I then remove all windows with greater RMS values since the associated energy will likely dominate CC+S panels.
At the Lalor Lake site I have found that the threshold chosen on masked data for any one hour is consistent with the threshold requirements for the remaining hours in the data set. I assert this to be generally true while the background seismic energy regime remains broadly stationary (e.g., no earthquakes, diurnal variations or changing environmental factors like wind or rain storms). However, significant changes in the background energy levels likely would require establishing time-varying threshold parameters.

The most obvious effects of applying the data masking and windowing processes are illustrated by the differences between the panels in Figures 3.3a–b and Figures 3.5a–b. The removal of noisy windows from Figure 3.3a to obtain Figure 3.3b (indicated by the blue regions) illustrates the degree to which the window selection process affects the result. By inspecting Figure 3.3a I observe that the eliminated energy is mostly related to high-energy coherent noise events. Similarly, the differences between Figures 3.5a–b (cumulative CC+S panels of the raw and masked data, respectively) show less coherent noise energy focused at non-zero correlation lags, suggesting a reduction in the associated mine site activity captured in the acquired waveforms. In terms of RMS convergence (Figure 3.6a), though, the data masking and windowing operations do not necessarily improve this metric. For example, the raw data curve with all windows (longer green line) approaches convergence; however, after applying filtering Steps 1 and 2 (blue line) the volume with fewer windows

![Image of Figures 3.8](image.png)

**Figure 3.8:** Illustrating the masking procedure. (a) Result of stacking envelopes (blue curve) computed for each raw data trace to calculate the 70% masking threshold (red curve). Both curves show scaled amplitudes ranging from zero (left) to one (right). (b) Result of applying the masking from where the masking operator (pink line, also valued from zero to one) calculated from the red curve in (a).
may not converge to a lower RMS energy value. The convergence curve gradient in Figure 3.6b, though, shows that the masked data (blue line) converge slightly faster than the raw data (green line), as evidenced by the lower values of the masked data gradient. The beamsteering plot in Figure 3.7b shows the result of applying Steps 1 and 2 on the directivity of the arriving energy. In comparison with Figure 3.7a, this panel now somewhat highlights arrivals from backazimuths other than that of the mine site. Arriving wavefield energy is less attributable to the mine site itself, because the highest energy windows and coherent events have been removed. A similar observation is noted in the beamsteering plot in Figure 3.7f, which still exhibits a significant amount of mine-related surface-wave energy, but at lower levels than that observed in Figure 3.7e. Overall, I observe that these processing steps facilitate CC+S convergence in some, but not all, hourly windows.

3.5.2 Step 3: Time Debursting

I remove burst-like data from individual seismic traces using an $L_1$ iteratively reduced least-squares (IRLS) debursting approach (Claerbout, 2014). This debursting operation addresses residual high-energy events remaining after performing the data masking and window selection steps. Instead of clipping samples with high amplitudes, it reshapes the trace to conform to the original shape while reducing the highest amplitudes. To apply this filter, I calculate the envelope for each individual trace and choose a preservation threshold (here 70%) based on the window RMS energy and the highest residual spike amplitude remaining after Step 2. Waveforms with magnitudes greater than this threshold are reduced to the selected value without clipping, while those with lower amplitudes are unaffected by the debursting process.

The effects of applying the $L_1$ debursting operator are illustrated by observing the differences between the pre- and post-time-debursted PSD plots (Figures 3.4b–c), where energy at particularly dominant frequencies is minimised over the entire recording time. The high-energy background observed in the CC+S panels in Figure 3.5b is reduced after time debursting (CC+S in Figure 3.5c), as evidenced by a greater number of red pixels clipped at an overall lower RMS energy level. However, when examining the noisy hour CC+S panels, the overall RMS energy reduction due to the time-debursting step is evident but less obvious than in the PSD plots due to the ringing nature of both pre- and post-debursted CC+S panels. While the differences between Figures 3.5b–c are somewhat difficult to visualise, the most significant influence of the time-debursting step is illustrated by the convergence plots in Figures 3.6a and 3.6c. The red line shows that the CC+S panel converges to a somewhat lower RMS energy value than the raw data result (green) and is generally smoother than the masked curve (blue). This implies that Step 3 has improved the CC+S panel because the variance of the RMS energy is lower than that of the input. The beamsteering plots in Figure 3.7c show that the influence of the time-debursting step on the wavefield directivity is minimal. However, there is relatively more focused energy with faster apparent velocity (light blue) in Figure 3.7c than in Figure 3.7b, meaning that most of the energy removed during this step is likely shear- and surface-wave contributions.
The time-debursting step does not greatly influence the beamsteering result in Figure 3.7g with respect to energy arriving at faster apparent velocities, but it does reduce the energy of apparent shear- and surface-wave arrivals from the mine site backazimuth at apparent velocities between 2.0-2.6 km/s. Overall, the above QC panels show that time-domain debursting is a helpful processing tool for highlighting signals and reducing higher-energy shear- and surface-wave arrivals not removed during the previous processing steps.

### 3.5.3 Step 4: Frequency Debursting

While the time-debursting step mitigates much of the localised spatially coherent noise, some spectra are still dominated by monochromatic energy, which is observable as a horizontal line in the PSD (Figures 3.4b–c) and as coherent ringing events in CC+S panels (Figures 3.5b–c). In particular, strong persistent monochromatic coherent noise is observed at 6.5, 11.0 and 17.0 Hz in the PSD.

To mitigate these effects I adapt the $L_1$ debursting IRLS approach of Claerbout (2014) for frequency-domain use. I first Fourier Transform ($\mathcal{F}$) particle velocity trace $v_{in}$ and

![Figure 3.9: Frequency-domain debursting example. (a) Time-domain example trace before (blue) and after (red) debursting. (b) Partial spectrum before (blue) and after (red) debursting. While the representative trace in (a) is somewhat altered, the debursted spectrum in (b) is well matched at frequencies where magnitudes are below the user-selected threshold.](image-url)
calculate the resulting magnitude $M$

$$M = \sqrt{\Re[\mathcal{F}(v_{in})]^2 + \Im[\mathcal{F}(v_{in})]^2},$$  \hfill (3.10)

and phase $\Phi$

$$\Phi = \arctan\left(\frac{\Im[\mathcal{F}(v_{in})]}{\Re[\mathcal{F}(v_{in})]}\right),$$  \hfill (3.11)

where $\Re$ and $\Im$ are the real and imaginary parts, respectively. I then apply the L$_1$ deburring algorithm to the Fourier magnitude spectrum [equation (3.10)] to generate a deburred spectral magnitude, $M_D$, which I use to reconstruct the frequency-deburred signal according to

$$v_{in}^D = \Re\left[\mathcal{F}^{-1}(M_{D}e^{i\Phi})\right],$$  \hfill (3.12)

where $\mathcal{F}^{-1}$ is the inverse Fourier transform operator. Note that the frequency-domain deburring operation maintains the original phase. Figure 3.9a shows a representative trace segment in the time domain before (blue) and after (red) applying the frequency-deburring procedure. While the representative trace appears somewhat altered, the frequency-domain deburred spectrum (red line in Figure 3.9b) is only modified for frequencies with spectral components falling above the user-selected threshold (blue).

When applied to the Lalor Lake data set, this operation slightly reduces the overall trace energy (see PSD in Figure 3.4d) and removes dominant ringing frequencies in the CC+S panels (see Figures 3.5d and 3.5h). The RMS energy variance of the CC+S panels after frequency deburring (magenta curves in Figures 3.6a–b) decreases monotonically and converges toward a lower level than any other curve. This improvement is even more evident in the noisy-hour RMS convergence curves, where the frequency-deburred result (magenta curves in Figures 3.6c–d) has lower RMS energy meaning that the benefits of this processing step are likely greater when applied to noisier data. This implies that the variance of each window is lower after applying the frequency-domain deburring step, which likely leads to more coherent CC+S panels. The RMS convergence curve in this noisy hour (Figure 3.6c) has not yet asymptoted and would likely improve further when including additional windows.

The beamsteering plots after applying frequency deburring (Figure 3.7d) show significantly reduced wavefield energy at slower apparent velocities ($\sim 2.0-2.6$ km/s), while arrivals at faster apparent velocities ($\sim 3.5-4.7$ km/s) are now more prominent. While most energy is coming from the directions of the road and mine site, focused body-wave events are now emerging from other directions (e.g., 45° azimuth and $\sim 5.0$ km/s in Figure 3.7d). Overall, the L$_1$ frequency-deburring approach is very useful for reducing dominant frequency components without requiring a priori knowledge of the underlying frequency structure or damaging the phase (i.e., Figure 3.9b).
3.5.4 Multi-hour stack

Because my overall motivation of applying the data processing workflow is to create a stable subsurface image using ambient body-wave reflection energy, it is pertinent to stack together multiple hours of CC+S panels both to potentially enhance S/N and to visualise hour-to-hour variations. Figure 3.10a shows a CC+S panel for 95 hours of stacked raw data selected based on S/N characteristics. These 95 hours generally correspond to between ~1-10 hours after blasting (02:00 am - 12:00 pm). Interestingly, these panels show little variation from the single-hour CC+S panel in Figure 3.5a. While this is not entirely unexpected because the RMS convergence curve in Figure 3.6a and those like it for other hours have already converged, a key question is whether different hours converge to a similar result. Figure 3.10b shows a CC+S panel for the masked and selected data for the same 95 hours. This result is similar to Figure 3.10a, as one might expect due to the minimal visual difference between the two panels. Figure 3.10c shows a CC+S panel from the same 95 hours after applying the full data processing workflow. The results are also

![Figure 3.10](image-url)

**Figure 3.10:** Averaged CC+S panels for 95 selected hours, largely falling between 1-10 hours after blasting. (a) Raw. (b) After selection and masking. (c) After time- and frequency-debursting.
3.5. PROCESSING THE LALOR LAKE AMBIENT DATA SET

similar to Figure 3.5d, except for the sharper focus at the large-magnitude event at about ±0.15 s two-way travel time. In addition, even fewer spiked and clipped points are observed relative to Figure 3.5d, indicating a reduction in spurious noise through the multi-hour stacking process.

Another interesting way to visualise the stationarity of the processed data over the entire recording time is to extract the same trace (i.e., the same virtual source and receiver locations, $x_s$ and $x_r$) from each hourly CC+S panel. Figure 3.11 shows a composite of single traces ($x_s-x_r$ offset of 1 km) extracted from the 95 selected CC+S gathers for raw (Figure 3.11a), masked/selected (Figure 3.11b), and fully processed (Figure 3.11c) data. While there is consistency between the raw-data hours shown in Figure 3.11a, there are variations over the full recording time that may prevent convergence of the fully stacked CC+S gather. After masking and selection there are still variations along the time dimension, such as the persistent ringing energy throughout Figure 3.11b. The fully processed panel (Figure 3.11c) shows not only greater consistency, but reduced ringing, throughout the entire recording time. For comparison purposes, Figure 3.11d presents CC+S results for the same $x_s-x_r$ pair for the raw data for all recording hours. There are blocks of hours that clearly exhibit quite different non-stationary noise characteristics; importantly, these hours have largely been removed through the window masking and

![Figure 3.11: A single extracted CC+S trace at a virtual source-receiver offset of 1.0 km for the 95 selected hours. (a) Raw data. (b) After masking and data selection. (c) After time and frequency debursting. While there is consistency across entire data set, the final processing results show greater stationarity and less ringing over full time axis. (d) 227 hours of raw data, which highlight periods of significant variability when compared to panel (a).]
selection process in Steps 1 and 2 (see Figure 3.3).

The most important improvement due to the processing workflow is the enhancement of potential body-wave reflections visible in the processed CC+S panels in Figures 3.10c and 3.11c. This enhancement is not visible or, at best, is very difficult to interpret in Figures 3.10a and 3.11a, respectively. While the most prominent stationary feature is the possible body-wave reflection at ±0.15 s, there is evidence for a number of weaker events (e.g., ±0.4 s and ±0.6 s) in Figure 3.10c with moveouts consistent with those of P-wave reflections. However, with such a large a P-wave velocity, moveout is nearly flat in this instance. This observation is important in the context of Lalor Lake ambient P-wave migration investigation reported in Girard and Shragge (2019b).

3.6 Discussion

Due to very low S/N of P-wave energy in ambient recordings, successfully applying a data processing workflow aimed at isolating P-wave reflections remains a challenging task. In this section I highlight three issues based on my experience of processing the Lalor Lake ambient data set: [1] the value of including quality control (QC) measures in the processing workflow; [2] the value of ‘informed’ data stacking; and [3] the associated caveats and potential drawbacks of applying an automated data processing approach.

3.6.1 Value of Quality Control Measures

My initial attempts at brute-force processing (and imaging) of the Lalor Lake data set met with limited success, largely because of the lack of interpretable coherent ambient body-wave signal that could serve as a processing QC guide. To address this issue, I developed and applied a number of standard and novel QC measures that allowed us to observe the degree to which specific data components were removed or enhanced by each data processing step. By comparing a sequence of power spectral density (PSD) plots, I was able to observe whether coherent (and often stationary) noise components (i.e., blasts, ringing from electro-mechanical noise, subsurface- and surface-based mining activity) were successfully removed. The main value of PSD plots as a data reconnaissance tool was in simultaneously viewing ambient waveforms across the entire survey recording time. These panels can be examined and interpreted for the presence of coherent monochromatic and impulsive noise sources. Combining CC+S panels and RMS convergence plots proved helpful for determining which windows added ‘different’ energy than those previously stacked as well as how many windows were sufficient to generate converged gathers. Moreover, examining a composite of single CC+S traces from the selected 95 hours demonstrated not just that CC+S panels generally converged, but did so with a remarkable degree of consistency. This observation helped to boost confidence in my overall interpretations. Finally, the beamsteering plots show both the dominant directivity and apparent velocity of wavefield arrivals, which provided important clues regarding potential source locations and wave types. Overall, I view these QC measures as important contributors to the overall...
3.6. DISCUSSION

3.6.2 Value of “informed” stacking

One of the first questions raised during my processing of the Lalor Lake ambient seismic data set was what is the value of the masking plus window selection versus doing a straightforward window selection? I view the inclusion of the masking process as a judicious addition for two reasons. First, an entire window of ‘good’ ambient data may be contaminated by only a single large-magnitude short-duration coherent noise event; discarding the entire trace may mean omitting usable body-wave energy from the final processed data set. Second, removing only the highest energy windows does not necessarily improve convergence because relatively high-energy events may remain in some otherwise relatively low-energy windows. I attribute the masking operation as the main reason why other window-selection-only processing workflows generally remove more windows than I have done here. Interestingly, the combined masking plus window selection operation does not guarantee a faster convergence rate than when using raw data (blue and green lines of Figure 3.6a, respectively), but it does slightly improve the CC+S result (Figure 3.5b). This non-intuitive observation persists throughout the data set and provides justification for why a processing scheme more comprehensive than either window selection or window selection plus masking is able to provide a satisfactory processed ambient data set.

A second important question was how many windows are required to achieve convergence of, e.g., the CC+S process? I address this question using the concept of image variance as expressed through hourly RMS convergence curves like those shown in Figure 3.6. Importantly, curves from all 95 selected hours have either converged or nearly converged. Moreover, as demonstrated by Figure 3.11c, the same virtual source-receiver CC+S trace extracted over the 95 selected hours shows remarkable consistency and provides persistent evidence of body-wave reflection events (e.g., ±0.15 s) that would otherwise be difficult to interpret in non-debursted CC+S panels (Figures 3.11a and 3.11b). While one hour is largely sufficient to achieve convergence for this particular location and data set, I recognise that convergence properties will depend on the particular acquired data set and its underlying signal and coherent noise characteristics.

3.6.3 Caveats

Finally, I note that introducing additional steps in the ambient data processing workflow leads to a greater overall computational cost than most CC+S processing methodologies. The results generated herein were computed using CPU- and GPU-parallelised algorithms executed on a combination of high-end GPU workstations and a limited number (<10) of cluster computing nodes. Once the codes were developed, automated processing for each hour of the Lalor Lake data set took roughly 25 minutes. Fortunately, the high degree of data parallelism means that this processing workflow is scalable to larger computational facilities if they were made available for this type of project.

The work described herein represents one example of a workflow that may be used to

robustness of the automated ambient data processing approach.
process an ambient data set to enhance P-wave signals. I recognise that there likely are other processing approaches or tools better suited for processing the Lalor Lake (or other) ambient data sets. Similarly, I stress that other processing tools and flows may be more optimal or appropriate for ambient recordings acquired in different noise environments. However, my intention in communicating my experience in processing the Lalor Lake data set is to emphasise the net benefit of developing an automated ambient processing workflow capable of reducing or eliminating coherent noise events and enhancing the underlying ambient body-wave signals, thereby facilitating body-wave imaging and inversion analyses.

3.7 Conclusions

This chapter explores an automated processing approach for denoising ambient seismic recordings with the aim of enhancing P-wave energy for ensuing ambient imaging and inversion investigations. I develop and apply a four-step processing workflow. The first step uses a masking function to mitigate dominant coherent noise events measured on all (or most) receivers within the window. The second step identifies and removes windows retaining a global high energy content after Step 1. The third step uses time-domain debursting to automatically despike localised impulsive events not necessarily coherent across the recording array. The final step applies an adapted frequency-domain debursting to the Fourier magnitude spectrum to remove dominant monochromatic coherent noise contributions. While each step requires a priori information about the data set and one or two global threshold parameters, these can be determined from initial data reconnaissance. I also implement several quality control (QC) methods [power spectral density (PSD) plots, cross-correlation plus stack (CC+S) panels, root-mean-square (RMS) energy convergence curves and beamsteering diagrams] to investigate the improvements introduced by each processing step and to monitor for the emergence of P-wave reflections. I view these as important tools for ensuring that the processing converges toward the desired result.

I used the Lalor Lake ambient data set to illustrate the utility of my automated ambient data processing workflow. Using the outlined QC measures, I demonstrated that the processing workflow down-weights coherent noise and reduces S- and surface-wave energy while enhancing P-wave energy which, though weaker, can be identified in CC+S panels based on characteristic moveout. RMS energy curves largely converged within the one-hour observation windows and were more likely to do so monotonically when using the fully processed rather than the raw CC+S panels. I observed that the CC+S waveforms for the same virtual source-receiver pair were largely stationary over 95 selected hours, further indicating successful convergence of the CC+S process. The beamsteering analysis proved to be an invaluable tool for determining the directivity and apparent velocity of ambient wavefield arrivals associated with mining related activity. These panels also showed enhancement of suspected P-wave reflections at the expected range of apparent velocities, especially for fully processed results. Based on these observations I assert that the processed Lalor Lake data set is better conditioned for any ensuing ambient body-wave imaging or inversion experiments. Finally, while I recognise that following this approach leads to
3.7. CONCLUSIONS

increased computational complexity compared to standard ambient data processing, I view this cost as a justifiable trade off for recovering a better-conditioned ambient data set.

Acknowledgements

I acknowledge project support from the Australian Government through an “Australian Government Research Training Program Scholarship”, The University of Western Australia (UWA) Top-up scholarship program, and members of the UWA Centre for Energy Geoscience. Drs. Saeid Cheraghi and Gilles Bellefleur from the Geological Survey of Canada graciously provided the Lalor Lake ambient data set. Data processing was conducted using the reproducible open-source package Madagascar (www.reproducibility.org). Computing resources were generously provided by the Center for Wave Phenomenon at the Colorado School of Mines and the UWA Centre for Energy Geoscience.
Chapter 4

Direct Migration of Ambient Seismic Data: Field Test Study

“Look deep into nature, and then you will understand everything better.”

Albert Einstein

4.1 Introduction

Ambient direct migration involves propagating wavefields from data recorded at the surface into the subsurface and applying some form of a cross-correlation plus stack (CC+S) imaging condition to construct an image. Chapter 2 outlines the 2D wave propagation and imaging condition methodology used in this thesis research for both active-source and ambient imaging scenarios. In that chapter, I also formulate a novel ambient extended imaging condition (EIC) based on extensions of active-source EIC theory. Ambient EICs are demonstrably helpful for identifying (un)focused ambient wavefield energy, and can also be used as a discriminant between ‘true’ (i.e., reflected P-wave) and ‘false’ (i.e., coherent noise) reflectivity. I have demonstrated that the direct wavefield migration and the use of an EIC work well for 2D imaging scenarios with synthetic data sets that are largely free of coherent noise; however, these are idealised cases due to limited approximations of true field conditions.

Chapter 3 develops a novel automated data processing workflow aimed at minimising coherent noise contributions in ambient data sets. I applied this automated processing workflow to the Lalor Lake 3D ambient field survey to evaluate its benefits and those of the introduced quality control (QC) tools. The processing methodology also demonstrated convergence of CC+S processes, improved stationarity of results over 95 selected hours, and up-weighting of energy consistent with reflected P-wave arrivals in an interferometric setting. However, Chapter 3 did not demonstrate whether the signal processing methodology significantly improved the result of the ambient direct migration procedure outlined in Chapter 2.
CHAPTER 4. DIRECT MIGRATION OF AMBIENT FIELD DATA

This chapter looks to demonstrate the benefits of combining the two processing and imaging procedures outlined in Chapters 2 and 3 to improve the quality and interpretability of 3D ambient direct migration images. To do so, I extend the 2D ambient direct migration and EIC theory presented in Chapter 2 to 3D scenarios and evaluate the added benefit of using a deconvolution (extended) imaging condition. I then address the imaging challenges associated with 3D array geometries and imbalanced illumination caused by irregular array geometry and aperture limitation. I use the conditioned Lalor Lake data set from Chapter 3 to demonstrate an ability to image subsurface reflectors known to exist based on the correlation of ambient seismic imaging results with that from results of an active-source seismic investigation (Bellefleur et al., 2015). I conclude with a discussion on the stationarity of interpreted locations over many hours of migrated recorded data and the relationship to previous active-source migration results.

4.2 Ambient Direct Migration

Direct migration of ambient seismic data requires two key procedures: a wave propagation operator to generate wavefields through a model domain, and an imaging condition to correlate wavefields and form an image. The one-way ambient 2D propagation outlined in Chapter 2 extrapolates wavefields between successive depth steps, starting with the surface-recorded data as the initial condition. This procedure is used to create two one-way ambient wavefields, one propagated forward in time, the other (adjoint) in reverse time. The imaging condition takes these two frequency-domain wavefields and correlates them to generate an image. An EIC is created by correlating the same wavefields after introducing a spatial lag to shift the wavefields in opposing directions prior to CC+S. The usefulness of these gathers is that they show whether or not wavefield information is optimally focussed at zero lag and thus representing a well-focused image constructed with the correct velocity model. Below I discuss the four types of imaging conditions that will be applied in this chapter.

4.2.1 3D Ambient Imaging Condition

Extending the conceptual framework of 2D ambient direct migration discussed in Chapter 2 to 3D field imaging scenarios requires handling the challenges associated with irregular and limited array geometries. For the 3D field investigations I use the one-way frequency-domain 3D wavefield propagators defined in Shragge and Shan (2010). Once the required ambient causal and anti-causal wavefields have been interpolated and propagated, the resulting wavefields may be correlated using a 3D version of the imaging condition defined in equation (2.3):

$$I_A(x, N_w) = \frac{1}{N_w} \sum_{n=1}^{N_w} \sum_\omega \Re \left[ R_+ (x, \omega, n) R_-^* (x, \omega, n) \right] ,$$  \hspace{1cm} (4.1)
4.2 AMBIENT DIRECT MIGRATION

where $\mathbf{x} = (x, y, z)$ represents a 3D rather than a 2D model, and all other symbols are as defined in Chapter 2.

4.2.2 3D Ambient Extended Imaging Condition

Evaluating a 3D ambient EIC requires applying wavefield correlations not just over the three spatial dimensions in equation (4.1), but over horizontal lags in the $x$- and $y$-dimensions as well. These lags are introduced as symmetric opposing vector wavefield shifts when evaluating the 3D EIC. The 3D ambient EIC is expressed similarly to that in equation (2.6):

$$I_A(\mathbf{x}, \mathbf{\lambda}, N_w) = \frac{1}{N_w} \sum_{n=1}^{N_w} \sum_{\omega} \Re \left[ R_+(\mathbf{x} + \mathbf{\lambda}, \omega, n) R_+^\dagger (\mathbf{x} - \mathbf{\lambda}, \omega, n) \right], \quad (4.2)$$

but with the vector lag term $\mathbf{\lambda} = (\lambda_x, \lambda_y)$ modified for the 3D case to include $\lambda_y$ and the 3D domain $\mathbf{x}$ defined in equation (4.1). This allows image focusing to be assessed over a 3D $\lambda_x$-$\lambda_y$-$z$ volume. In this experiment I examine 3D EIC gathers, but when identifying moveout over the EIC gather through velocity (d$V$, as defined in Chapter 2), I will utilise only a single horizontal lag $\lambda_x$ for visualisation purposes.

4.2.3 Ambient Deconvolution Imaging Condition

One imaging challenge discussed in Chapter 2 is that frequency spectra of field data are usually poorly balanced. In active-source imaging this can be addressed by modifying the source acquisition strategy or through a seismic data spectral whitening operation; however, in ambient seismic scenarios these procedures are challenging to apply due to both the lack of defined sources and continuity of events across traces. Therefore, in addition to the frequency debursting operation in Chapter 3, I also use a deconvolution operation to help balance wavefield spectra during the imaging condition:

$$I_{A \text{decon}}(\mathbf{x}, N_w) = \frac{1}{N_w} \sum_{n=1}^{N_w} \sum_{\omega} \Re \left[ \frac{R_+(\mathbf{x}, \omega, n) R_+^\dagger (\mathbf{x}, \omega, n)}{|R_+(\mathbf{x}, \omega, n)|^2 + \epsilon^2} \right], \quad (4.3)$$

where the wavefield magnitude is given by $|R| = \sqrt{\Re^2(R) + \Im^2(R)}$, and $\epsilon^2$ is a small positive constant introduced as a stabilisation factor. Deconvolution operations are known to help whiten spectra when there are uneven and irregularly distributed components (including coherent noise events stationary in frequency). The deconvolution operator is introduced in Claerbout (1971), and has been previously applied to active-source shot-profile migration (Valenciano and Biondi, 2005; Guitton et al., 2007), and in interferometry (Wapenaar et al., 2011) investigations.

4.2.4 Ambient Deconvolution Extended Imaging Condition

Extended images have the same challenges associated with unbalanced spectra as conventional ambient images. Therefore, this thesis also tests the benefits of introducing a
deconvolution operator to the formation of EIC gathers. Extending the deconvolution ambient 3D imaging condition in equation (4.3) requires including a set of spatial lags \( \lambda = (\lambda_x, \lambda_y) \), which lead to the 3D ambient deconvolution EIC:

\[
I_{A \text{decon}}(x, \lambda, N_w) = \frac{1}{N_w} \sum_{n=1}^{N_w} \sum_{\omega} \Re \left[ \begin{array}{c}
R_+^{\dagger}(x + \lambda, \omega, n)R_+^{\dagger}(x - \lambda, \omega, n) \\
\left| R_+(x + \lambda, \omega, n) \right| \left| R_-^{\dagger}(x - \lambda, \omega, n) \right| + \epsilon^2
\end{array} \right].
\] (4.4)

Introducing these \( \lambda \) lags is done in a similar fashion as to the 3D ambient cross-correlation imaging condition in equation (4.2). The resulting 5D EIC gather has three spatial coordinates, \( x = (x, y, z) \), as well as two horizontal lags \( (\lambda_x, \lambda_y) \). Because densely sampled 5D volumes would require significant amounts of input/output (I/O) and disk space, herein I compute 5D EICs only in a few locations to “diagnose” which correlations are the most likely candidates of imaged P-wave reflected energy.

### 4.3 Lalor Lake Dataset Postprocessing

Chapter 3 discusses the processing of the 3D ambient Lalor Lake data set and the selection of 95 highly consistent hours from the full 227 hours of recording. Each selected data hour has between 40 and 109 event windows of \( \sim 30 \) s duration with high-energy events masked out and localised spikes in the time and frequency domains, respectively removed through time and frequency debursting filtering.

Preparing the processed data for direct migration required applying several post-processing steps. First, I adjusted the wavefield migration aperture by removing stations from the ends of long lines falling outside of the dense 2D receiver grid. Figure 4.1 shows the removed stations as yellow circles. This step was viewed as necessary because there is little-to-no crossline aperture in these locations, and the migrated data will generate severe elliptical imaging artefacts that easily could be mistaken for real structure (see tests below). Second, I applied a bandpass filter with the high-cut at half of the final image Nyquist frequency (31.25 Hz for a sampling rate of 16 ms) to ensure that no spurious correlations are introduced by image-domain aliasing (Artman et al., 2005). Finally, I interpolated the data from acquisition coordinates to the propagation grid using 2D spatial linear operators within the 3D migration code to match the regularly sampled propagation grid defined at \( \Delta x = \Delta y = 20 \) m intervals. Wavefields were extrapolated from the assumed flat surface by \( \Delta z = 20 \) m depth increments such that a discrete wavefield was defined at regular \( \Delta x = \Delta y = \Delta z = 20 \) m intervals. The final volume dimensions are \( N_x \times N_y \times N_z = 256 \times 256 \times 256 \) for a \( 5.1 \times 5.1 \times 5.1 \) km\(^3\) total volume, shown as a cyan box in Figure 4.1, which are cut back to \( 150 \times 150 \times 150 \) in a \( 3.0 \times 3.0 \times 3.0 \) km\(^3\) volume for visualisation of more well constrained areas in the figures below. The box does not appear square in this figure due to the projection used in mapping at the 54.88°N latitude.

The data were migrated in the frequency domain, using data Fourier transformed to 337 frequencies ranging from 2.0-22.5 Hz at a 0.06 Hz interval, and using a constant propagation velocity of \( V_P = 6.0 \) km/s. The choice of frequencies comes from both the

78
increased coherent noise at higher frequencies and the increased cost of migrating those high frequencies, while the propagation velocity was chosen due to both moveout velocity analyses in Chapter 3 and regional knowledge in Bellefleur et al. (2015). This is close to the Roots et al. (2017) approach that applied a linearly increasing $v(z)$ profile between 6.0-7.0 km/s over a $\sim$4.0 km depth interval.

### 4.4 3D Field Data Imaging Challenges

Expanding the imaging condition to 3D includes handling challenges associated with irregular array geometry, imbalanced spectra of input wavefields, and correctly identifying true reflectivity versus coherent noise. Thus, before applying an imaging operation to ambient field data and interpreting the results, it is judicious to first address these issues. I examine the array geometry effects by testing its influence (i.e., acquisition footprint) on a set of synthetic horizontal planar reflectors. Next, I test the benefits of whitening wavefield spectra during imaging by comparing the results of the 3D deconvolution ambient imaging condition to its cross-correlation counterpart. Finally, I explore the difference in

![Map showing the ambient seismic survey area and its location in Manitoba, Canada (inset). Filled-in circles show recording station locations. Magenta circles indicate the geophones used for imaging. Yellow circles represent geophones removed for imaging because of their distance from the denser array. Red letters and blue stars show locations selected for extended image locations shown in subsequent figures. The cyan rectangle shows the maximum migration aperture. Black stars indicate recording stations of Line 133, used for interpretation.](image)

**Figure 4.1:** Map showing the ambient seismic survey area and its location in Manitoba, Canada (inset). Filled-in circles show recording station locations. Magenta circles indicate the geophones used for imaging. Yellow circles represent geophones removed for imaging because of their distance from the denser array. Red letters and blue stars show locations selected for extended image locations shown in subsequent figures. The cyan rectangle shows the maximum migration aperture. Black stars indicate recording stations of Line 133, used for interpretation.
imaging results from raw and fully processed data to illustrate the value of applying the pre-migration ambient processing workflow outlined in Chapter 3. I explore each of these effects on the Lalor Lake “noisy” and “quiet” hours that were used to QC the processing workflow in Chapter 3.

4.4.1 Array Geometry Effects

Before imaging the fully processed Lalor Lake ambient seismic data, it is important to understand the influence of acquisition geometry on the migration results. To examine this, I undertake a synthetic experiment to show the array footprint at different depths from

![Figure 4.2: Acquisition footprint effects of the full Lalor Lake acquisition geometry on the horizontal plane-wave test image. (a) 0.5 km, (b) 1.0 km. (c) 1.5 km. (d) 2.0 km. (e) 2.5 km. (f) 3.0 km. Each plane is visible in the cross-sections (g) [E,N]=[1-4, 2.5] km and (h) [E,N]=[2.5, 1-4] km.](image-url)
plane-wave sources initially covering the full migration aperture (Figure 4.2). I initiate the experiment with five horizontal plane-wave impulse responses [i.e., the exploding reflector model of Claerbout (1985)] excited at known subsurface depths. I then propagate these sources to the surface using a representative constant velocity model and measure their responses recorded at the true Lalor Lake geophone locations. I then apply the direct migration methodology to the synthetic data recorded with the correct acquisition footprint to image the plane-wave sources. Thus, any deviations in the image from a horizontal planar response are due to the irregularity and limited aperture of the recording array.

I visualise the resulting planar image response in depth slices at the exploding reflector locations in Figures 4.2a–f, and in inline and crossline sections extracted from the centre.

**Figure 4.3:** Acquisition footprint effects of the restricted Lalor Lake acquisition geometry on the horizontal plane-wave test image. (a) 0.5 km, (b) 1.0 km. (c) 1.5 km. (d) 2.0 km. (e) 2.5 km. (f) 3.0 km. Each plane is visible in the cross-sections (g) \([E,N]=[1-4,2.5]\) km and (h) \([E,N]=[2.5,1-4]\) km.
of the array in Figures 4.2g–h. At depths <750 m there is a more significant acquisition footprint effect on the imaged horizontal reflectors; at depths >750 m the reflectors are better imaged, though the depth slices show that the horizontal reflectors have not fully healed; however, illumination would improve when including a range of plane-wave sources. Poor illumination is also observed at the perimeter of the array, including at the ends of the longer lines of receivers (denoted in yellow in Figure 4.1).

Because of acquisition footprint effects on the 3D direct migration operation, it is pertinent to remove isolated receivers that are not part of the dense 3D receiver array. In particular, one inline and three crossline segments extending well beyond the denser 2D main array grid were removed prior to imaging. These 44 receivers are denoted as yellow circles in Figure 4.1. To examine the imaging difference between the full and reduced arrays, I recreate the synthetic plane-wave test using only the magenta receivers in Figure 4.1. Figure 4.3 shows the results of the synthetic test with the 44 outlying receivers removed. The resulting image is similar to the result of the full array, though some differences do exist particularly at shallower depths. Relative to the same depth slice using all receivers (Figure 4.2a), these differences are manifest as reduced cross-hatching in the plane-wave imaging results at 0.5 and 1.0 km depth slices (e.g., at [E,N]=[1.5, 3.5] km in Figures 4.3a–b. This observation repeats, though less prominently, for deeper events, and is consistent with this being the area where most receivers were removed. The effect is also observed by comparing the cross-sections of Figure 4.2h with Figure 4.3h, where fewer spurious correlations are present on the right hand side between horizontal reflectors at 1.0 and 1.5 km depth. Reducing spurious correlations caused by a limited number of receivers will assist with future interpretations at those depths. Based on this plane-wave test I expect interpretable migration results as shallow as 0.5 km in the ambient seismic data recorded on the Lalor Lake ambient array geometry.

4.4.2 Cross-correlation versus Deconvolution Imaging Conditions

Introducing the deconvolution operator to the imaging condition is known to help balance spectral amplitudes; however, choosing the values of $\epsilon^2$ in equations (4.3) and (4.4) can significantly alter the resulting image. Choosing too high of an $\epsilon^2$ value does nothing to counter the spuriously correlated ringing energy and illumination issues in the image (i.e., it effectively reduces to a weighted cross-correlation imaging condition); using too low of an $\epsilon^2$ value can overwhiten the spectrum by overemphasising all correlations in the image (regardless of frequency content of the data), and may cause instability when there is division by a near-zero value. Therefore, to highlight these image differences and to help chose an appropriate $\epsilon^2$ value, I ran a suite of tests for the fully processed quiet-hour data using a wide range of $\epsilon^2$ values. I step up the $\epsilon^2$ value by orders of magnitude from 0.1 to 10 for experiments shown in Figure 4.4. Increasing $\epsilon^2$ tends to decrease the magnitude of image correlations, particularly at depth, and allows for easier interpretation of the remaining correlations. An $\epsilon^2 = 0.1$ value (Figure 4.4a) was deemed too low because many of the higher-energy noise correlations are boosted, particularly at the edges of the
Figure 4.4: Comparison of 3D direct migration image volumes of the fully processed quiet hour data with the deconvolution imaging condition from equation (4.3) for different $\epsilon^2$ values: (a) 0.1; (b) 1.0; (c) 3.7; and (d) 10.0.
migration array, and may be incorrectly interpreted as coming from true structure. An $\epsilon^2 = 1$ value (Figure 4.4b) produced more reasonable results, but high-energy correlations still remain at unlikely reflection locations at >2 km depth. An $\epsilon^2 = 10$ value (Figure 4.4d) was deemed too high because too much energy was removed from the image. Thus, the value for $\epsilon^2$ was chosen at 3.7, or 10% of the RMS energy of the input data volume, resulting in an image with fairly well-balanced amplitudes (Figure 4.4c).

A good way to visualise the benefits of balancing spectral amplitudes on imaging is to directly compare imaged data from the Lalor Lake data set using cross-correlation and deconvolution imaging conditions in equations (4.1) and (4.3) for both the raw and fully processed data. Figures 4.5a–b show the raw data from the noisy and quiet hours imaged with the cross-correlation ambient imaging condition, respectively. The most prominent imaged events are the correlations ringing with increasing depth. Clear illumination issues caused by coherent noise coming from a single location are observed, as evidenced by the high-amplitude events in the depth slices and the North-South cross-sections of Figure 4.5 at 3 km Northing, as well as at [E,N]=[2.0, 2.5] km. Cross-correlation ambient images using fully processed data (Figures 4.5c–d) exhibit fewer ringing events and coherent noise events, and show more reasonably focused energy. However, there is still evidence of ringing and directed coherent noise, particularly in the fully processed noisy-hour data (Figure 4.5c) at 3.0 km Northing.

To show the benefit of applying the deconvolution imaging condition, I use the same quiet- and noisy-hour data to generate deconvolution migration images (Figure 4.6). The improvement is evident for raw data for both the noisy (Figure 4.6a) and quiet (Figure 4.6b) hours, in particular the removal of almost all ringing energy. Numerous coherent events remain, and are more likely to be related to true reflectivity. The images in Figures 4.6c–d combine the processing workflow and the deconvolution imaging condition, leading to migration results with the least amount of obvious spurious correlations, with the most balanced amplitudes, and that are most likely to result in a successful interpretation.

### 4.4.3 Imaging Raw Data versus Processed Data

To highlight the benefits of both the automated processing workflow and the data selection routine in Chapter 3, Figure 4.6 presents images for the raw and fully processed data sets for the quiet and noisy hours. Figure 4.6a presents an image from the raw noisy hour that shows high-amplitude correlations at 0.5-0.7 km depth spread across the majority of both presented cross-sections. The depth slices exhibit better focusing directly beneath the cross-section intersection point. While this may look like an interpretable image, many spurious correlations no doubt exist, particularly around the mine site and at depths consistent with mine shaft construction (0.835 km). Comparably, the image created from the raw quiet hour in Figure 4.6b shows many of the same correlations as Figure 4.5a. The depth slice, showing lower amplitude correlations to the north of the cross-section intersection, likely highlights reflected energy.

Comparing images from raw and fully processed data illustrates the benefits of applying
the automated processing workflow. In the fully processed noisy-hour image (Figure 4.6c),
the largest effect is a reduction in amplitudes to the west of the image intersection in
both the depth slice and cross-section panels. Because this is the direction of the mine
shaft, this is likely where most coherent noise originates. The lower amplitude correlations
in this region suggest that processing has reduced the coherent noise. However, there
are still numerous spurious correlations at potential reflector locations in deeper regions
(e.g., [E,N]=[1.5-1.8,2.5-3.0] km in Figure 4.6c). The fully processed quiet-hour image
(Figure 4.6d) fewer prominent correlations at depth, as may be expected due to the limited
2×2 km² array. Most importantly, the areal extent of imaged reflectivity between 0.5-0.8 km
depth is small and more reasonable considering the reflector from the estimated ore body
(Bellefleur et al., 2015).

This result illustrates that the automated processing workflow helps to mitigate coherent
noise events that lead to spurious image correlations. However, it also provides motivation
for a post-migration analysis to identify the ‘best’ images with minimal spurious correlations
(i.e., in the noisy-hour processed data image in Figure 4.6c) so as to not greatly bias final
image interpretations.

4.5 Imaging Results

Having examined the consequences of applying the automatic processing workflow and
deconvolution imaging operator to improve direct wavefield migration of ambient seismic
data recorded on sparse irregular arrays, I now compare the imaging results from the 3D
direct migration operation.

4.5.1 Imaging Single Hours of Fully Processed Data

Based on the results above, I perform single-hour imaging tests involving the fully processed
quiet- and noisy-hour data sets. Images from each hour are formed by stacking numerous
sub-images from each contributing window (108 and 80 for the quiet and noisy hours).
Figures 4.7 and 4.8 show the imaging results for the quiet and noisy hours, respectively, with
the four locations of the crosshair intersections indicated by the red letters in Figure 4.1.
Each panel examines a different image location and depth slice to identify image correlations
most likely associated with true reflectivity. The image from the first location (Figure 4.7a),
which shows energy focused at the crosshair location at 0.74 km depth with relatively
large areal extent, suggests a flat or low-angle NE-dipping reflector. The second panel
(Figure 4.7b) at 0.84 km depth shows focusing in two locations (one at the crosshair and
second at 1.0 km farther to the west), each with a fairly small areal extent. The third panel
(Figure 4.7c) shows energy focused at 2.0 km depth; however, the restricted 2 × 2 km²
array aperture is not optimal for imaging reflectors at this depth as it is near the expected
maximum depth of investigation. The fourth panel (Figure 4.7d) shows higher amplitude
image correlations, which are likely reflection energy since there is little coherent noise
expected at this depth location.
Figure 4.5: Comparison of 3D direct migration results using a cross-correlation imaging condition for both raw and fully processed quiet- and noisy-hour data. (a) Raw data image from the noisy hour. (b) Raw data image from the quiet hour. (c) Processed data image from the noisy hour. (d) Processed data image from the quiet hour.
Figure 4.6: Comparison of 3D images with the deconvolution operator from both raw and fully processed hours from the Lalor Lake data set. (a) Raw data image from a noisy hour. (b) Raw data image from a quiet hour. (c) Processed data image from a noisy hour. (d) Processed data image from a quiet hour.
Figure 4.7: 3D images from the fully processed quiet-hour data set at four locations migrated using only the selected receivers (magenta) noted in Figure 4.1: (a) [E,N] = [2.36, 2.54] km, (b) [E,N] = [3.06, 2.66] km, (c) [E,N] = [2.54, 2.54] km, and (d) [E,N] = [2.28, 2.22] km.
Figure 4.8: 3D images from the fully processed noisy-hour data set at the same four locations as in Figure 4.7: (a) \([E,N] = [2.36, 2.54] \) km, (b) \([E,N] = [3.06, 2.66] \) km, (c) \([E,N] = [2.54, 2.54] \) km, and (d) \([E,N] = [2.28, 2.22] \) km.
In the first noisy-hour image panel (Figure 4.8a) the areal extent of the reflector is similar to that of the quiet hour, but there are additional spurious correlations associated with ringing energy at \([E,N]=[1.0,2.5]\) km. The second panel (Figure 4.8b) shows focusing at the same locations as in the corresponding quiet-hour panel. The third panel (Figure 4.8c) does not show the expected reflector at 2.0 km depth, which is more evident in the corresponding quiet-hour panel (Figure 4.7c). The fourth panel (Figure 4.8d) has several large-amplitude focal points at the 0.5 km depth between Northing 2.8-3.2 km. Those may be related to coherent noise sources since the quiet-hour results did not show energy focused well at these locations. Overall, the main difference is that the noisy-hour imaging results have lower amplitudes at the examined subsurface locations. Interestingly, the differences between the quiet- and noisy-hour images are not immediately obvious, suggesting that the processing workflow likely sufficiently mitigated coherent noise sources and the deconvolution imaging condition has better balanced wavefield spectra.

### 4.5.2 Image Convergence

Calculating image convergence is important for understanding the stability of a stacked ambient image. Image convergence can be shown as a curve of the total RMS energy in the stacked image after adding each successive window. Thus, this provides a measure of how similar the energy recently added to the stack is to that already present. This is similar to the convergence concept presented for CC+S panels in Chapter 3, but now includes the deconvolution migration operator and wavefield propagation influences.

I calculated convergence curves using the methodology outlined in equation (2.10), which is similar to that presented in equation (3.9). Image convergence curves calculated from the sub-image stack for the quiet- and the noisy-hour data (Figure 4.9a) show that both are progressing toward convergence, though have not yet fully converged. This observation is somewhat expected due to the similarity with the RMS convergence curves created as processing QC panels in Chapter 3. The curves exhibit some jitter, which is not as evident in the RMS curves created from the CC+S operation; however, these may be due to an incorrect migration velocity model. The gradient curves (Figure 4.9b) show

![Figure 4.9](image.png)

**Figure 4.9:** Convergence curves for the deconvolution imaging condition. (a) Convergence curve for the noisy (blue) and quiet (magenta) hours. (b) Gradient of the convergence curves for the noisy (blue) and quiet (magenta) hours.
that jitter does not greatly affect convergence. It is interesting to note the similarity of this curve with those calculated from the CC+S panels in Chapter 2 (Girard and Shragge, 2019a): even though they are calculated using different imaging methods, the convergence and convergence rates are still similar. This observation suggests that RMS convergence curves are a good measure of image stability that is independent of the imaging method used.

4.5.3 Extended Images

The EIC gather locations were chosen after generating conventional ambient images and identifying image points with relatively high-amplitude correlations at locations consistent with those of expected reflections from active-source images (Bellefleur et al., 2015) and reflectivity observed in the deconvolution migration results for the fully processed quiet hour. Figures 4.10 and 4.11 show the quiet- and noisy-hour EIC gathers, respectively. In both figures, much of the energy focuses at non-zero lags in both directions ($\lambda_x$ and $\lambda_y$). The best examples in each panel are for depths 0.74 km in Figure 4.10a, 0.84 km in Figure 4.10b, 0.74 km in Figure 4.10c, and 0.5 km in Figure 4.10d. In Figure 4.10c the depth slice is shown at 1.98 km rather than the main focusing location because the EIC gather shows reflectivity at the depth expected from the active-source survey results.

There are some areas of significant non-zero lag focusing, particularly at negative lag distances in Figures 4.10a and 4.10c between 0.5-0.8 km depth. This misfocusing may be due to array geometry, the imperfect migration velocity model, multiple arrivals, or coherent noise sources. While all explanations are plausible, the most likely causes are coherent noise or acquisition footprint, since this is both a near-surface and near-mine-site location with minimal array coverage.

The EIC gathers from the noisy-hour data (Figure 4.11) show minor variations, as may be expected from the limited variations between the conventional images at the same locations. However, there is more spuriously correlated energy at non-zero lags than in the EIC gathers in Figure 4.10, particularly in Figure 4.11a at $[\lambda_x, z] = [-0.2, 0.8]$ km, Figure 4.11b at $[\lambda_x, z] = [-0.2, 1.6]$ km, Figure 4.11c at $[\lambda_x, z] = [+0.2, 1.4]$ km, and Figure 4.11d at $[\lambda_x, z] = [+0.2, 0.8]$ km. This observation suggests that multi-hour imaging results may be improved by stacking fewer quiet hours rather than maximising the number of hours and thereby possibly introducing spurious correlations at non-zero lags into the image stack.

For comparison, I also extract EIC gathers at array perimeter locations to show that without appropriate receiver coverage, there are unlikely to be focused image correlations at a reflection location. Figure 4.12a shows the EIC gather extracted at the location of the western blue star in Figure 4.1, and illustrates that most energy focuses at non-zero lags in the $\lambda_x$-$\lambda_y$ space. However, there is correlation at $\lambda_x = 0$ for the illustrated depth (0.74 km), which may be the expected reflector, but with so much unfocused energy around, this is not a viable interpretation. Figure 4.12b shows the EIC gather extracted at the eastern edge of the array in Figure 4.1, and shows even poorer correlation at zero-lag than
Figure 4.10: EIC gathers created from the fully processed quiet-hour data set at locations of interest denoted as red letters in Figure 4.1. Each EIC gather has both horizontal lags $\lambda_x$ and $\lambda_y$ and shows image focusing centred at zero lag.
Figure 4.11: EIC gathers created from the fully processed noisy-hour data set at locations of interest denoted as red letters in Figure 4.1. Each EIC gather has both horizontal lags $\lambda_x$ and $\lambda_y$ and shows image focusing centred at zero lag.
CHAPTER 4. DIRECT MIGRATION OF AMBIENT FIELD DATA

Figure 4.12: EIC gathers created from the fully processed quiet hour at locations outside of the dense receiver array areas, shown as blue stars in Figure 4.1. These EIC gathers show most energy focused at non-zero lags.

Figure 4.12a. Because most energy is not correlated well at zero lag, it is reasonable to conclude that both EIC gathers and conventional images near the edges of the migration aperture are not representative of the subsurface.

4.5.4 Global Velocity Model Perturbations

As identified in Chapter 2, one way to explore the effects of velocity errors in the migration and imaging process is to perturb the velocity model globally and identify the differences. Figure 4.13a shows an image at the same location as Figure 4.7a but migrated with a velocity that is 5% slower, which can be stated as a global velocity perturbation $dV = -5\%$. Figure 4.13b shows the image when $dV = +5\%$, which is faster. The differences in these images show various focusing changes, particularly an improved focusing of the dipping reflector at $[E,N,z]=[2.36,1.0,1.0]$ km in the image with $dV = +5\%$ in Figure 4.13b and better focusing at $[E,N,z]=[2.6,2.54,2.1]$ km in the image with $dV = -5\%$ in Figure 4.13a.

To more clearly identify the effects of modifying the velocity, I use a more dense sampling than $dV = \pm 5\%$ and create a combined EIC gather with $\lambda_x$ and $dV$ as the extended axes, as outlined in Chapter 2. I create an EIC gather set by migrating the data at each velocity perturbation at 1% increments for $-5\% \leq dV \leq +5\%$ and looking at the changes in $\lambda_x$ over $dV$. The four $\lambda_x$-$dV$ plots for the quiet hour are shown in Figure 4.14 at the same locations as the conventional EIC gathers. In Figure 4.14a the focusing improves as $dV$ increases, particularly at a 1.5 km depth. Conversely, the focusing increases with
decreased velocity, as evidenced at \([dV, z] = [3\%, 2.3 \text{ km}]\). Both results suggest that the velocity model used for migration could be improved to better focus energy. Similarly, at the next location shown in Figure 4.14b, there is evidence of focusing improvement as velocity increases at \([dV, z] = [3-5\%, 1.5 \text{ km}]\) and decreases at \([dV, z] = [-3-1\%, 2.1 \text{ km}]\). The same changes in focus at zero lag as velocity changes are seen in the \(\lambda_x-dV\) panels locations in Figures 4.14c–d. This is a consistent but more interpretable result than simply looking at zero-offset images with different migration velocities in Figure 4.13.

When examining the \(\lambda_x-dV\) panels for the same locations for images created with the fully processed noisy-hour data (Figure 4.15) similar focus-velocity changes are observed. These changes in focus are evidence that these \(\lambda_x-dV\) panels can be used to identify migration velocity model errors as well as imaged events that have moveout characteristic of reflected body-wave energy. There are, though, some differences between the noisy and quiet hours. The most prominent is between Figures 4.14d and 4.15d, where the reflector at 2.8 km depth in the former is absent from the latter results. At this depth, there is fairly limited illumination aperture for reflectors, so it is not surprising to see greater differences at depths \(>2.0 \text{ km}\), however, such a difference is likely related to coherent noise in the data that is imaged at this location.

Overall, the value of \(\lambda_x-dV\) panels is shown by identifying reflectors that focus better when the migration velocity is globally perturbed. This suggests that images created through direct wavefield migration of processed ambient seismic data can be used for future inversion operations.

\textbf{Figure 4.13:} Images at the same location as Figure 4.7a using fully processed quiet-hour data and the reference velocity with (a) \(-5\%\) and (b) \(+5\%\) perturbations.
Figure 4.14: Set of $\lambda_x$-$dV$ panels created from the fully processed quiet-hour data set at locations of interest denoted as red letters in Figure 4.1. EIC gather has both horizontal lags $\lambda_x$ and global velocity perturbation $dV$ to show how image focusing changes at zero lag as migration velocity changes.
Figure 4.15: Set of $\lambda_x$-$dV$ panels created from the fully processed noisy-hour data set at locations of interest denoted as red letters in Figure 4.1. EIC gather has both horizontal lags $\lambda_x$ and global velocity perturbation $dV$ to show how image focusing changes at zero lag as migration velocity changes.
4.6 Imaging Multi-hour Recordings

Thus far, I have only explored the imaging results from single hours of data that exhibit considerable similarities. When exploring the images for all hours, it is important to examine the variability between images and identify those more optimal for multi-hour stacking. In Chapter 3 I identified 95 fully processed recording hours that show desirable stability characteristics following CC+S operations. Here, I explore the benefits of stacking images created from all optimal 95 fully processed hours.

The multi-hour images in Figures 4.16 and 4.17 are presented at the same image cross-section locations as the single-hour images for comparison purposes. There seem to be only minor image differences from the single-hour results, which is likely due to the fact that the individual images have largely converged. While stacking additional consistent hours is unlikely to change the image, observing consistently focused image energy over many hours reduces interpretation uncertainty.

One way to visualise the stability of the stacked image over multiple time windows is to extract the same single trace from each hour’s image. In this case, at the first image location for all previous examples, [E,N] = [2.36, 2.54] km, I plot the same image location from every hour to create a stability plot (Figure 4.18) in the similar fashion as that introduced in Chapter 3. Figure 4.18a shows the stability plot extracted from each imaged hour at the same location. It shows considerable stationarity in the images at all depths over the entire 95 hours of imaged data, with few changes over the hourly time intervals; however, there is some jitter, implying differences in the higher frequency components. Stacking the traces over recording time (Figure 4.18b) shows the image at that location, and can identify reflections. The addition of a high-cut filter (Figure 4.18c) reduces some of this jitter and presents a more consistent result, which further highlights the consistency of the image correlations. Stacking the traces over recording time (Figure 4.18d) shows that the filtering had a negligible effect on the stacked image, and therefore justifies the use of a high-cut filter for interpretation purposes.

4.6.1 Comparing Ambient and Active-Source Images

With a quantifiably stable image created from many imaged hours of data, it is important to identify whether or not the image created from the ambient data is functional for exploration purposes. In this instance, I am greatly benefiting from the results from an active-source experiment carried out in Bellefleur et al. (2015), which resulted in a post-stack time migrated (PoSTM) image volume. The location of the sources and receivers in this volume cover a larger area than the receivers for the Lalor Lake ambient data set examined in this thesis (16 km² versus 4 km² area), encapsulating the entire ambient area of investigation. The active-source data set also had more receivers than the ambient data set (2685 versus 336), and therefore has a wider aperture and better resolution due to the broadband source frequency content (Bellefleur et al., 2015). However, the ambient survey is entirely coincident with the exploration area, as shown in Cheraghi et al. (2015) and
Figure 4.16: 3D images at location A from (a) 95 imaged and stacked hours and (b) the active-source PoSTM volume, and location B from (c) 95 imaged and stacked hours and (d) the active-source PoSTM volume. Ambient images are migrated using only the selected receivers (magenta) as noted in Figure 4.1.
Figure 4.17: 3D images at location C from (a) 95 imaged and stacked hours and (b) the active-source PoSTM volume, and location D from (c) 95 imaged and stacked hours and (d) the active-source PoSTM volume. Ambient images are migrated using only the selected receivers (magenta) as noted in Figure 4.1.
4.6. IMAGING MULTI-HOUR RECORDINGS

Figure 4.18: Stability panel over the 95 imaged hours extracted at [E,N] = [2.36, 2.54] km. This shows the continuity of the imaged location after the automated processing and a deconvolution imaging workflow. (a) Unfiltered result showing moderate differences over calendar time and (b) stacked trace of all 95 traces. (c) Filtered result to highlight continuity of highest amplitudes and remove jitter and (b) stacked trace of all 95 traces. The difference in the stack trace is negligible between the two, and the filtered version is much easier to interpret.
Roots et al. (2017), so image points are therefore directly comparable.

Because the active-source PoSTM image volume was organised in an inline-crossline format at an angle oblique to the ambient migration geometry used herein, I modified the geometry using 2D linear interpolation to UTM coordinates to facilitate comparison with the ambient-seismic image volume presented above. Figures 4.16 and 4.17 show image panels for both ambient and active-source images extracted at the same geographic image locations as in previous figures for comparison. In Figure 4.16d there is a reflector dipping to the north from an expected ore body at about \([\text{E}, \text{N}, z]=[2.36, 2.54, 0.8]\) km, which is where the intersection occurs in the bottom right part of the panel. Figure 4.16c a reflector at the same location with the same dipping characteristic, suggesting that the processing and imaging workflow applied herein is highlighting reflected body-wave energy. The active-source images have a higher frequency content and resolution than the ambient images due to the wider and denser receiver coverage. However, there are clear similarities between the active-source PoSTM images and the fully processed ambient direct migrated images. Particularly, in the active-source image in Figure 4.17b, there is a reflection similar to that interpreted in Bellefleur et al. (2015) to arise from an ore bearing formation at (in this image) \([\text{E}, \text{N}, z]=[2.54, 1.0-3.0, 1.0-1.5]\) km. This reflector is also evident in the ambient image in Figure 4.17a at the same location, except for the range \([\text{E}, \text{N}, z]=[2.54, 1.0-1.5, 1.0-1.5]\) km, where the image quality is reduced due to the ambient array geometry. Due to the numerous differences between the images (frequency content, aperture, illumination, imaging condition) it is difficult to directly compare images in this format.

It is possible, however, to bandpass the active-source PoSTM image to simulate a frequency content that is similar to that of the ambient image. The result (Figure 4.19b) shows a different energy at each reflector than observed in the broad-band active-source image in Figure 4.16d. The bandpassed active-source image is consistent with the ambient image (Figure 4.19a, which is the same as Figure 4.16c). The most prominent effect of the bandpass operation on the PoSTM image is the reduction of amplitude of the deeper reflector, indicated by the arrow at 2.4 km depth. This is a good indication that the correlated energy in the ambient image are likely reflected energy. However, differences still arise due to differences in migration velocity and illumination. This is particularly obvious on the left arrow of both images in Figure 4.19, where the reflector in the ambient image appears misplaced and misfocused, likely due to migration aperture issues at the margins.

An easier way to facilitate comparison between the active-source PoSTM and direct migrated ambient images is to extract 2D image cross-sections from each at the same location. Figure 4.20 shows a North-South cross-section extracted from each image volume at \(\text{E} = 2.54\) km, which goes nearly through the centre of the ambient grid. The dipping reflector at R1 matches well between the ambient (Figure 4.20a) and the active-source (Figure 4.20b) images, and is consistent with the Bellefleur et al. (2015) interpretation of being related to to be the footwall of the ore body of interest in the active-source experiment. Locations R2 and R3 are prominent reflectors in the active-source image, but
Figure 4.19: 2D annotated images extracted at E=2.54 km for (a) ambient and (b) active-source images. Annotations indicate locations of prominent reflectors.

have a less-dominant signature in the ambient image, possibly due to tuning effects of the lower-frequency ambient image. It is also possible that this is more poorly resolved due to the low receiver coverage in this area of the array. A similar reflector is visible, however, in both images at the deeper location R4, and consistent with the Bellefleur et al. (2015) interpretation of being related to volcanic intrusions to the north of the mine site. Location R5 shows some correlation between the images at a somewhat flat reflector that is more prominent in the active source image at 2.3 km depth. The same in the active-source image is evident at location R6, which is also more prominent in the ambient image. The illumination of this reflector at such a depth is less certain due to the depth of investigation.

Figure 4.20: 2D annotated images extracted at E=2.54 km for (a) ambient and (b) active-source images. Annotations indicate locations of prominent reflectors.
of the ambient image (∼2.0 km) but also because it may not be within the illumination aperture of either experiment. The location denoted R7 shows a reflector that could be interpreted in the active-source image, but is more prominent in the ambient image. This could be due to a multiple or ringing energy associated with the reflector at location R1.

Figure 4.21 shows an East-West cross-section extracted from each image volume at N=2.78 km, which again goes nearly through the centre of the ambient grid. The dipping reflector at depth, between R1 and R2, is prominent in the active-source image (Figure 4.21b), visible in the ambient image (Figure 4.21a), and is consistent with the Bellefleur et al. (2015) interpretation as the lower footwall of a deeper ore body. A similarly dipping shallower reflector at R3 is clearly identified in both images, and is consistent with the previous interpretations of being related to the upper ore body footwall which is of interest to the mining company. The shallowest reflector in this image is denoted as R4, and is interpreted to have a larger areal extend than denoted by the annotation arrow. The reflection is apparent in both images, and may be related to the ore body located in this vicinity at 0.8 km depth. The reflection at D1 is a different dipping reflector than R3, but is not as visible in the ambient image, perhaps due to the lower frequency content. However, the ambient reflector does show a localized discontinuity, indicating that some signal from D1 may be present in the image. Location F1 is the most interesting part of this comparison, because it appears more prominently than the surrounding geology in the ambient image than in the active-source image. However, it is uncertain why this region is seemingly better illuminated in the ambient image. Thus, the overall interpretation of the ambient image is broadly consistent with that in Bellefleur et al. (2015).

A more direct comparison can be made by converting the ambient image coordinate system to the original active-source coordinate system and comparing with a geologic model. Schetselaar et al. (2017) provides a detailed 3-D geological model based on over 220 exploration and delineation boreholes plus in-mine underground drill holes which have been drilled in and around the Lalor deposit (borehole information was unavailable for this

Figure 4.21: 2D annotated images extracted at N=2.78 km for (a) ambient and (b) active-source images. Annotations indicate locations of prominent reflectors.
4.7 DISCUSSION

Roots et al. (2017) extract Line 133 (receivers of which are indicated in Figure 4.1 as black stars) from this model, which is reproduced here as Figure 4.22c with the ambient (Figure 4.22a) and active-source (Figure 4.22b) images along the same line. While the ambient image is of slightly lower quality than in earlier representations due to the use of a sinc interpolation to convert coordinate systems to the active-source inline and crossline geometry, there are still reflections evident that correlate to both the active-source image and the geologic model. Location H1 is interpreted as the hanging wall of an ore deposit, and is evident in the active-source image, but is too shallow in the ambient image to be fully imaged. A second hanging wall at location H2 is well recovered in the ambient image, but at a slightly different depth, likely due to the constant migration velocity being slightly too slow. The footwall denoted at F1 is also well recovered in the ambient image, though with less certainty in location due to both the lower frequency and velocity error associated with propagation. Thus, the overall interpretation of the ambient image, though still somewhat uncertain, is broadly consistent with the geological model interpretation presented in Roots et al. (2017). There are also reflections that match between the ambient and active-source images - labelled R1, R2, R3 and R4 - that are not in the geological model, but offer more evidence that the ambient image is largely coincident with the active-source image.

4.7 Discussion

The overall imaging experiment undertaken in this chapter combines two key factors: automated ambient data processing and direct migration. While these had been examined separately in previous chapters, they had not yet been examined together on field data. One key result from this is the finding that field data with unbalanced spectra caused spurious and ringing events that made final image interpretation challenging. To address this, I introduce a deconvolution (extended) imaging condition to help whiten the spectrum of the final image correlations. The deconvolution operator also significantly improves images from unprocessed data over images created from the same data with a standard cross-correlation imaging condition; however, combining fully processed data with the deconvolution imaging condition resulted in the most interpretable images.

Another important interpretability factor is the stationarity of the resulting image stack. While image convergence has proven to be important when identifying reflected energy in ambient seismic data, convergence properties are dependent on the overall recording time used to generate the results. By computing a stability plot from a single location (or multiple plots from multiple image locations) over many hours of images prior to stack shows the degree to which an imaged location remains stationary over calendar time. The plot provides evidence of both convergence and stationarity, and thereby increases confidence that the correlated energy is indeed related to an imaged subsurface reflector.

Justifying interpretations in ambient seismic data is challenging due to the overall lower frequency and lower resolution that conventional active-source images. In this experiment a direct comparison is made to an active-source image volume from the same location. There are similarities in the image volumes that relate to reflections interpreted from
Figure 4.22: 2D annotated images extracted at Line 133 from the original survey grid points for the (a) ambient and (b) active-source data. (c) A geological model of Line 133 from Roots et al. (2017) for comparison.
the active-source image volume. The most convincing reflector identified at location R1 in Figures 4.20a–b and location R3 in Figures 4.21a–b is dipping to the northeast from 0.5-1.5 km depth, which leads to a similar interpretation as from the active-source image by Bellefleur et al. (2015) and Roots et al. (2017).

4.8 Conclusions

Direct wavefield migration of 3D ambient seismic field data requires addressing challenges associated with irregular array geometries, imbalanced spectra of input data and identifying true reflected energy at reflector locations. I examine the irregular array geometry in the Lalor Lake data set by undertaking a plane-wave migration test that highlights a number of receivers that do not greatly contribute to the overall illumination, thereby increasing the opportunities for misinterpretation of resulting imaged reflectivity. Imbalanced spectra of input data are addressed by including a deconvolution (extended) imaging condition to whiten the spectra of propagating wavefields, which helps to reduce ringing and other spurious correlations in the image results. The image volume is interpreted with the assistance of extended image gathers that identify moveout and focusing in the spatial lag domain $\lambda$, as well as over global velocity perturbations $dV$. Each of these image extensions allows identification of reflectors and reflector focusing by changing the migration and imaging parameters to observe moveout of possible reflectors in the images.

Identifying reflectors and reflector moveout in a single hour of recorded ambient seismic data is valuable only if the reflectors are stationary over calendar time. To identify the consistency of the images, 95 selected hours were processed and migrated to form individual hour images. A stationarity panel built from the same image trace from each migrated hour shows how consistent the images are over time, increasing the reliability of any interpretations. The value of ambient seismic direct wavefield migration is finally shown by comparing the results to a co-located active-source seismic data set and observing comparable reflectors in both images that generally support pre-existing geological interpretations of the Lalor Lake mineral deposit system.
Chapter 5

Conclusions

“Begin at the beginning,” the King said, gravely, “and go on till you come to the end; then stop.”

Lewis Carroll, Alice in Wonderland

5.1 Summary of Conclusions

The use of ambient seismic recordings in the exploration for Earth resources has increased in recent years, particularly in the mining and oil and gas sectors. The reasons for this largely stem from the increase in data storage capacity and computational resources, the need for non-destructive methods for subsurface investigation, and a desire for continuous monitoring of resource-producing areas. With the development of interferometric and direct migration cross-correlation plus stack (CC+S) methods, there are now several imaging approaches to utilise ambient seismic recordings for exploration purposes.

While there have been numerous successful examples using these two ambient imaging approaches, significant challenges with the current methodologies remain. In particular, the presence of coherent noise events of significantly larger magnitude than ambient reflected P-wave energy leads to scenarios where CC+S methods generate uninterpretable and uncertain ambient imaging results - regardless of whether the analysis is conducted by data-domain interferometry or image-domain direct migration. The key issue is that noise-mitigation processing is performed after applying CC+S operations, which makes it difficult or impossible to remove the effects of high-magnitude coherent noise to reveal weaker imaged reflected-wave P-wave energy. Even for situations where input data are largely free of coherent noise, quantitatively discriminating imaged ambient P-wave reflection energy from other contributing wave modes remains challenging.

To address these issues, Chapters 2 and 3 develop a workflow to identify and mitigate coherent noise in ambient seismic data prior to CC+S procedures, as well as an extended-image-based direct migration methodology that provides a necessary test for discriminating imaged reflected body-wave energy from spurious imaged noise contributions. Chapter 4 demonstrates that these approaches are applicable to the Lalor Lake field data, which
CHAPTER 5. CONCLUSIONS

is characterised by sufficient acquisition aperture and spatial sampling density for wave-equation-based imaging purposes. The key findings of each study are outlined below.

5.1.1 Study 1: Direct Migration of Ambient Seismic Data

Chapter 2 addresses the research question of how to adapt concepts and techniques from active-source seismic imaging to help directly migrate, identify and interpret imaged ambient P-wave reflection energy. When coherent noise energy in ambient seismic data is sufficiently low, the direct migration approach images P-wave reflections in the similar manner as active-source imaging. A key requirement is that the image must have converged to a stable result, a criterion that can be demonstrated using an image convergence measure developed herein. The ambient direct-migrated extended imaging condition (EIC) gathers developed herein have similar character to those created during active-source migration, even for more complex geological scenarios. Therefore, ambient EIC gathers can be used as a tool for interpreting and validating imaged body-wave reflections as they are currently used in active-source seismic investigations.

The advantages of this direct migration plus EIC gather approach over other reported ambient imaging approaches are threefold. First, there is no need to identify or separate source information or to use interferometric methods to explicitly synthesise source and receiver wavefields prior to imaging. Second, for scenarios involving low coherent noise energy and correct migration velocity models, a converged ambient EIC gather will show P-wave reflection energy focused at the correct subsurface locations. Spurious correlations unrelated to reflection events (i.e., coherent noise) largely will not focus at zero lag. Thus, this combined methodology presents a way to discriminate between unwanted wave modes and imaged P-wave signal. Third, remigrating the volume with a range of globally perturbed velocity models will generate an EIC gather with an added dimension of velocity-error $\Delta V$. Scanning the lag-versus-velocity-error ($\lambda - \Delta V$) coordinate space for characteristic moveouts over the $\Delta V$ dimension represents a tool for interpreting and validating imaged ambient body-wave reflectivity. This is because imaged P-wave reflected events have different sensitivity to velocity changes than other noise wavemodes. These novel contributions should improve the overall ambient imaging practice by allowing for quantifiable interpretations of ambient seismic reflection images.

5.1.2 Study 2: Automated Processing Strategies for Ambient Seismic Data

Chapter 3 develops a novel signal processing workflow to condition ambient seismic data and thereby extends the direct migration approach developed in Chapter 2 to scenarios typified by contamination with large-magnitude coherent noise events. This automated workflow aims to identify and mitigate three types of large-magnitude coherent noise that create challenges for imaging operations: stationary in time (i.e., simultaneously coherent across an array), stationary in space (i.e., continuously coming from the same location but at different times), and stationary in frequency (i.e., persistent monochromatic emissions).
The automated processing workflow has four steps. The first step uses a masking function to mitigate dominant coherent noise events (i.e., relative to the ambient RMS energy background) measured on all (or most) receivers within a relatively short time window. The second step identifies and removes windows retaining a global high energy content after applying the masking procedure of Step 1. The third step uses a time-domain debursting algorithm to automatically despike localised impulsive events not necessarily coherent across the recording array. The final step applies an adapted frequency-domain debursting algorithm to the Fourier magnitude spectrum to remove dominant monochromatic coherent noise contributions. While each step requires a priori information about the data set and one or two global threshold parameters, these are easily determined from an initial data reconnaissance.

To investigate the improvements introduced by each processing step and to monitor for the emergence of P-wave reflections, I also implement several quality control (QC) methods [power spectral density (PSD) plots, cross-correlation plus stack (CC+S) panels, root-mean-square (RMS) energy convergence curves and beamsteering diagrams]. These tools are important for identifying coherent noise prior to processing as well as ensuring that the four-step workflow converges toward the desired result. More importantly, these QC tools, along with stationarity panels, suggest that the migration results converge to a consistent subsurface image. The processing results provide evidence that the processed Lalor Lake data set is better conditioned for any ensuing ambient body-wave imaging or inversion experiments. Beyond the Lalor Lake data set, these processing results show that defining an automated workflow for pre-CC+S ambient seismic data is not only feasible, but should adequately prepare other ambient seismic data sets for imaging and potentially, inversion purposes.

5.1.3 Study 3: Direct Migration of Ambient Seismic Field Data

Chapter 4 expands the 2D imaging workflow to 3D scenarios and applies it to an ambient seismic data set acquired at Lalor Lake mine site in Manitoba, Canada. This requires addressing a number of imaging challenges. First, I compensate for the unbalanced frequency spectra generally observed in ambient seismic data by introducing a deconvolution operation to the ambient (extended) imaging condition. This operation helps to whiten the spectrum when there are uneven and irregularly distributed spectral components. Second, I examine the effects of irregular and limited receiver distribution in the array by propagating a series of horizontal plane waves and imaging the results using the Lalor Lake array geometry. The test shows that removal of isolated receivers should improve interpretability of resulting images by identifying regions lacking receiver coverage or with a strong acquisition footprint. Furthermore, the results indicate that the subsurface is well illuminated by the array at depths as shallow as 500 m. Overall, these tests show that using the ambient deconvolution imaging condition and fully processed data generates interpretable direct migration results with reflectivity located at places well illuminated by the array.

After addressing these challenges, I show that applying the ambient deconvolution
extended imaging condition to build 3D EIC gathers helps users to examine focusing in both horizontal spatial lag dimensions. I present stability images calculated for 95 selected hours using a stationarity plot, which shows that image points do not vary significantly between recording hours. Finally, I show that the automated ambient processing and ambient deconvolution imaging operation applied to the Lalor Lake seismic data set produces similar results to post-stack time migrated (PoSTM) images from the co-located active-source seismic experiment. In particular, there are strong similarities between the two images in locations with NE-dipping reflectors appearing at locations coincident with those interpreted from the ore body. Accordingly, these observations are consistent with interpretations from previous investigations. Interpretation challenges remain in deeper regions of the ambient images due to aperture limitations, which is not true for the active-source PoSTM. Thus, I expect that ambient image interpretation would be improved with denser receiver arrays.

5.2 Lessons learned

Imaging subsurface reflectors using ambient recorded seismic data comes with several other challenges that were not outlined in the previous chapters: (1) limited array aperture; (2) low array acquisition density; and (3) presence of surface-wave arrivals.

Increasing the number of sensors in the array increases the overall imaging aperture would allow for farther offsets and therefore allow for more accurate imaging in deeper regions. Similar to active-source seismic investigations where a good “rule of thumb” is that, in an image, reflectors can be resolved up to a depth of the longest offsets in the receiver array. Thus, a broader array would help to resolve deeper reflectors when using ambient recorded data.

Ambient seismic imaging experiments would benefit from denser receiver arrays, and a corresponding spatial aliasing in the recorded domain. Most ambient seismic imaging experiments to date are modelled after regional or tectonic seismology experiments, where receiver offset is more important design criterion than acquisition density. Accordingly, the receiver spacing tends to be much wider than in active-source experiments - in this case, 100 m by 400 m spacing in the ambient experiment compared to 50 m spacing in the active-source image. With a denser receiver array, greater spatial resolution and improved wavefield healing is achieved during migration (resulting in fewer issues with migration swings, which can be challenging during interpretation, as seen in the Lalor Lake example). This combination would improve ambient seismic migration results. With the decreasing cost and increased use of nodal acquisition systems, denser and larger arrays would not only be beneficial to the direct migration of ambient seismic data, but is becoming increasingly practical.

In this thesis, I assumed that the surface wave energy was of lower importance than other processing issues. While the imaging results seem to support this, I recognise that this may not be the case, either here or in other experiments. To address this issue, additional multi-channel noise elimination techniques and likely elastic imaging may be important to
5.3 Future work

While the work detailed in Chapters 2-4 highlights how automated data processing and direct migration approaches can be applied to ambient synthetic and field data, there are a number of interesting challenges and opportunities that have yet to be fully addressed. I detail some of these potential extensions below.

5.3.1 Application to Other 3D Ambient Seismic Data Sets

This thesis successfully applies the workflow to a single data set in a unique ambient environmental setting; however, the effectiveness of this developed processing and imaging workflow likely will be different when applied to other data sets, especially for those acquired in significantly different noise environments. Thus, it is likely that the automated processing workflow will need to be tailored to suit the specific environment to handle the types of stationary and coherent noise present in ambient data recordings. For example, the Long Beach array experiments in the urban environment of Los Angeles, California contained significant variations in diurnal effects from traffic (i.e., rush hours), industrial noise from the port, wave- and weather-dominated influences from the nearby coastline, as well as possible seismicity over a wide range of magnitudes. Each noise type may need to be handled in a different fashion than discussed herein, potentially with alternative filters and processing steps. However, regardless of the specific filtering workflow selected, I assert that the overall strategy and usefulness of applying an automated ambient data processing should remain valid.

5.3.2 Modelling Coherent Noise

Because handling large-magnitude coherent noise events is of utmost importance for generating accurate and interpretable synthetic and field ambient images, an important question is how does one devise realistic 3D controlled ambient imaging experiments to test the robustness of data processing and interferometry/direct migration workflows and their sensitivity to coherent noise? One approach commonly used in active-source seismology is to build 3D Earth models and simulate pre-stack seismic data volumes. While this could be done for an ambient data set from a deterministic physics and computing perspective, an interesting question is how would one model realistic coherent noise sources for the types (and potentially others) identified in thesis and thereby generate a realistic synthetic ambient data set for testing purposes?

Based on the experience in this thesis, I assert that modelling a realistic noisy ambient seismic data set would help to identify problems associated with existing data processing and 3D imaging workflows. When statistically relevant noise is modelled and added to synthetic data sets, new and different processing tools can be assessed for different types of coherent noise. It may even be beneficial to use such 3D ambient modelling tools to build
an “industry standard” synthetic ambient data set useful for future community processing and imaging benchmarking experiments in the same way that seismic data sets modelled through the SEAM program have done for active-source marine (Fehler and Keliher, 2011) and land (Oristaglio, 2012; Regone et al., 2017) investigations. This would allow for a partnership between academia and industry to explore the problems with ambient seismic exploration, develop novel and cutting-edge processing and imaging solutions, and could lead to tangible improvements in ambient seismic exploration practice.

5.3.3 Multi-Channel Filtering

In this thesis, I outline the types of coherent noise present in the ambient seismic data set from Lalor Lake, Manitoba. Most of the filtering steps applied to address these noise types are single-channel operations; however, there are advantages of multi-channel processing - especially in removing energy with dip-based filtering of apparent velocities corresponding to those of shear- and surface-wave arrivals. Introducing different and additional filtering steps may be useful for better conditioning data prior to applying direct migration. For example, Fourier-based 3D dip-filtering ($f-k$) methods commonly used in conventional seismic data processing are difficult to apply in ambient scenarios because of irregular acquisition geometries and the theoretical and computational challenges of performing accurate wavefield interpolation of ambient waveforms. However, sparse 3D Radon filtering methods do not suffer from these limitations (Trad et al., 2003) and are commonly used in active-source land seismic data processing to remove noise associated with direct and surface-wave arrivals. Ideally, the same could be done in ambient seismic recordings to reduce the influence of these direct arrivals. The overall impact on the imaging process would be to reduce the energy propagated into the model, eliminate many of the spurious CC+S correlations of these wave modes, and thereby lead to more interpretable images and higher interpretation confidence.

5.3.4 Azimuthal Weighting and/or Balancing

The beamsteering analysis in Chapter 3 proved to be an invaluable tool for determining the directivity and apparent velocity of ambient wavefield arrivals associated with mining-related activity. These panels also show enhancement between raw and processed results of suspected P-wave reflections at the expected range of apparent velocities. However, the question remains as to whether an invertible $\tau$-azimuth-apparent-velocity transform would be as useful for balancing wavefield arrivals to better satisfy equipartition assumptions of CC+S processing?

Based on the Lalor Lake data set processing experience, I assert that developing and applying an azimuthal weighting scheme prior to imaging would allow for more equally distributed illuminating wavefields, which would be beneficial for the overall imaging operation. This azimuthal weighting could be implemented through 3D Radon filtering in the $\tau$-azimuth-apparent-slowness domain to more evenly balance wavefield illumination of subsurface reflectors. This operation could also be used to filter out or down-weight
specific wave types (i.e., surface-wave energy) based on the apparent moveout velocity, thereby preventing this energy from entering migration images. More uniform illumination of subsurface reflectors should improve the correlation of the ambient imaging condition at those reflectors, and improve the overall image interpretability.

5.3.5 Computational Runtime Improvements

Finally, another workflow improvement would be to parallelise all processing steps for advanced computing infrastructure platforms (e.g., GPU). In these experiments I completed the data processing component of the overall workflow on a CPU-based high-performance computing system with runtimes averaging 40-50% of the total recording time when processing data down-sampled to a 62.5 Hz Nyquist frequency. When including quality control (QC) measures, the computational efforts increased by several fold (about 3x), depending on number of windows removed. Direct migration of processed data using a GPU-parallelised code on a high-end GPU workstation required 80% of the total recording time (e.g., imaging one hour of data took 48 minutes on a single high-end P100 GPU card). Direct wavefield migration and applying the extended imaging condition on the same workstation required 120% of the total recording time, and including the velocity perturbation EIC gathers increased the runtime another tenfold. This means that the combined processing, QC and imaging runtimes were longer than the actual recording time.

In ambient seismic monitoring scenarios, it is advisable to reduce the analysis runtime to be equal to or less than the recording time. Fully GPU-parallelising all processing and imaging workflow steps and exploiting data parallelism to expand the analysis to multiple GPU nodes would address this concern and help to achieve the goal of real-time ambient data processing and imaging analysis. Solving these challenges could mean that these workflows could become a standard tool for 3D and time-lapse applications of ambient seismic exploration and monitoring.
Bibliography


Bellefleur, G., E. Schetselaar, D. White, K. Miah, and P. Dueck, 2015, 3D seismic imaging of the Lalor volcanogenic massive sulphide deposit, Manitoba, Canada: Geophysical Prospecting, 63, no. 4, 813–832.


Cheraghi, S., J. A. Craven, and G. Bellefleur, 2015, Feasibility of virtual source reflection seismology using interferometry for mineral exploration: A test study in the Lalor Lake volcanogenic massive sulphide mining area, Manitoba, Canada: Geophysical Prospecting, 63, no. 4, 833–848.


——–, 1971, Toward a unified theory of reflector mapping: Geophysics, 36, no. 3, 467–481.


——–, 2010, Basic Earth Imaging: Stanford Exploration Project.

——–, 2014, Geophysical Image Estimation by Example: LULU Press.


———, 2019a, Automated processing strategies for ambient seismic data: accepted to Geophysical Prospecting.

———, 2019b, Direct migration of ambient seismic data: accepted to Geophysical Prospecting.


Miah, K. H., G. Bellefleur, E. Schetselaar, and D. K. Potter, 2015, Seismic properties and effects of hydrothermal alteration on Volcanogenic Massive Sulfide (VMS) deposits at the Lalor Lake in Manitoba, Canada: Journal of Applied Geophysics, 123, 141 – 152.

Moschetti, M., M. Ritzwoller, and N. Shapiro, 2007, Surface wave tomography of the western United States from ambient seismic noise: Rayleigh wave group velocity maps: Geochemistry, Geophysics, Geosystems, 8, no. 8, (Q08010).


Olivier, G., and F. Brenguier, 2016, Interpreting seismic velocity changes observed with ambient seismic noise correlations: Interpretation, 4, no. 3, SJ77–SJ85.


———, 2015, Image-domain wavefield tomography with extended common-image-point gathers: Geophysical Prospecting, 63, no. 5, 1086–1096.