UNIVERSITY OF CALGARY

Event Detection and Classification Using Distributed Acoustic Sensors

by

Heather Hardeman

A THESIS

SUBMITTED TO THE FACULTY OF GRADUATE STUDIES

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE

DEGREE OF DOCTOR OF PHILOSOPHY

GRADUATE PROGRAM IN MATHEMATICS AND STATISTICS

CALGARY, ALBERTA

NOVEMBER, 2019

Abstract

This thesis is on the mathematics of seismic data acquisition and processing, with a particular focus on distributed acoustic sensing. Distributed acoustic sensing (DAS) is a relatively new means of seismic acquisition. It utilizes strain on a fibre-optic cable to deduce information about nearby events. DAS systems possess the potential for a wide variety of applications beyond seismic acquisition including security and monitoring. In order to pursue these applications, developing techniques for detecting and identifying events in DAS-acquired data is essential.

In this thesis, we explore various methods for locating and classifying events in data acquired using a distributed acoustic sensor. We begin with an investigation of reflection and transmission coefficients as they define where events and anomalies occur in data. We find exact solutions for these coefficients and use them to provide insight into the success of numerical methods in modelling seismic data. Then, we consider seismic processing techniques such as wavelets and time-frequency analysis. We also develop a wavelet transform: the inverted wavelet transform.

An explanation of distributed acoustic sensing and how it works is provided. Afterwards, we produce models of DAS-acquired data and use these models to offer insight into the amplitude response of a DAS system. It also enables the employment of a homotopy to compare different formations of fibre.

Applications to distributed acoustic sensors fill the final chapters of the work. The first example involves the use of DAS for acquiring vertical seismic profiles at the Containment and Monitoring Institute's Field Research Station in Newell County, AB. We then employ Gaussian mixture models and independent component analysis to detect a vehicle signal in data acquired using a DAS system. To address classification of events, we utilize a convolutional neural network to identify events in microseismic data as well as in data monitoring someone walking and digging next to a distributed acoustic sensor. This investigation leads to a discussion of feature-based image registration with regards to distributed acoustic sensing acquired data. Finally, we establish the Hyperbola Method to determine the distance between an event and the DAS system from the data.

Preface

This thesis is an original work by the author. No part of this thesis has been previously published.

Acknowledgments

Thank you to all my friends, family, and teachers, near and far, who encouraged and supported me as I pursued mathematics. Your support has meant the world to me. I would not have made it this far without you.

To my committee, thank you taking the time to go through this rather large thesis and learn what I have accomplished during my time at the University of Calgary.

To Ms. Monosky and Dr. Martin, thank you for encouraging me to pursue Mathematics instead of English. Sometimes you do not realize what is best for you, and you both helped me find what was best for me.

To Melissa, thank you for emboldening me to use my weird, but concrete ideas for analytical arguments in my literature essays. It made me a better applied mathematician, because no idea is too weird not to try as long as it is well-founded.

To Dr. Pigott, thank you for introducing me to the University of Calgary for my PhD.

To Dr. Robinson, thank you for deciding I wanted to do a masters thesis before I did and helping me see how much fun mathematical research is. Thank you for keeping in contact and your continual encouragement over the course of my time at the University of Calgary.

To Dr. Hope, thank you for introducing me to a physics research program that showed me how much fun research in general really is. The experience made me unafraid to look in areas outside my own for answers.

To Joel, thank you for always answering my random LaTeXquestions. Seriously, you were a lifesaver.

To Dr. Tyler and Dr. Martin, thank you for introducing me to board games and tabletop RPGs which have been fundamental to building my support network as I moved to Winston-Salem and then to Calgary. Friends are essential to success and you gave me the building blocks to make those friends.

To Fotech, thank you for allowing me to use your data to run several experiments included in this body of work.

To CREWES and the CREWES sponsors, thank you for the support over the course of my degree.

To Matt, thank you for your guidance these last three years. Not only did you teach me what it was like to work in industry, but you also showed me how to succeed in graduate school. Thank you for everything.

To Dr. Rios, thank you for your help over the last few years and agreeing to be my co-supervisor.

To Dr. Lamoureux, thank you for your guidance these last five years. Thank you for seeing my grad school application and deciding that you wanted to work with me. Thank you for your patience and guidance as I learned seismic imaging and geophysics. It has truly been an honor working with you during my time at the University of Calgary, and I look forward to future collaborations. I truly appreciate all your guidance and support. Finally, thank you for always bearing with my comparisons to *Jurassic Park*. You only have to deal with it one more time, I promise.

Finally, to Geoff and Rachel, thank you for all of your encouragement throughout my degree and especially as I reached the finish line. Thank you for always believing in me, even when I couldn't or didn't. I love you both more than you know! Now, it's y'all's turn!

To Geoff, Rachel, and Petunia

Table of Contents

List of Figures and Illustrations

List of Tables

List of Symbols, Abbreviations and Nomenclature

Symbol or abbreviation	Definition
BP	Basis Pursuit
CaMI	Containment and Monitoring Institute
CNN	convolutional neural network
CWT	Continuous Wavelet Transform
DAS	distributed acoustic sensing
Fibre Distance	displacement along the length of the fibre
FRS	Field Research Station
GMM	Gaussian mixture model
ICA	Independent Component Analysis
IWT	inverted wavelet tree
Path of Fibre	displacement along the path of the fibre
PRF	pulse repetition frequency
SST	Synchrosqueezing Transform
TSWT	tree-structured wavelet transform
VSP	vertical seismic profiles

Epigraph

There is nothing like looking, if you want to find something. You certainly usually find something, if you look, but it is not always quite the something you were after.

- J.R.R. Tolkien, *The Hobbit*

Chapter 1

Introduction

When introducing the paleontologists Alan Grant and Elli Sattler in the 1993 film *Jurassic Park*, Steven Spielberg provides an excellent example of the method for acquiring seismic data. In the scene, the team at the dig site shoot a shotgun blast into the ground to create sound waves. These sound waves collide with the velociraptor's bones which, in turn, reflect a portion of the wave back. Sensors pick up the reflections and the computer uses the data collected by the sensors to create an image of the dinosaur's skeleton. Seismic imaging utilizes sound waves similarly. When collecting seismic data, explosives, a vibrator truck, or even a blunt object hitting a circular metallic disk on the ground creates sound waves which travel deep into the Earth's surface and ricochet off anomalies, such as dinosaur bones or fossil fuel reservoirs. Sensors on the Earth's surface, often geophones, read these reflections and create a data set which can be used to determine what lies in the ground.

While a good outline of the general process behind seismic imaging, the film is not completely correct. One of the paleontologists is able to deduce how the dinosaur died using the seismic image they generated. Although the film was released over a quarter of a century ago, today's technology can only resolve large geological features that are tens to hundreds of meters in size, lying hundreds to thousands of meters below the Earth's surface after some processing has occurred. The scientists acquired and generated the image of the skeleton in a matter of seconds using a single shotgun blast into the ground. Even today, typical seismic acquisition requires a couple of days for a complete study of the area. For example, a seismic study in Australia recently lasted 36 days (APPEA, 2019). Another issue is that multiple sources are necessary for acquisition instead of the single shotgun blast used in the film in order to obtain any reasonable degree of accuracy. Complex seismic acquisitions necessitate more time and more sources. After acquisition, the data must be processed which often involves correcting the geometry based on the placement of the sources and receivers (Aminzadeh et al., 2003). Several weeks or even months can pass between acquisition and the completion of data processing. A good deal of movie magic occurs in *Jurassic Park*'s example of seismic imaging; however, it still provides a basic idea of the fundamentals of seismic acquisition.

Given the film's release in 1993, the paleontologists in *Jurassic Park* likely utilized geophones as the receiver when they acquired the seismic data of the dinosaur's skeleton. While image quality has not progressed to the point of the film, there exist many new ways to collect seismic data. One novel method of interest is called Distributed Acoustic Sensing (DAS). It utilizes a laser interferometer and a fibre-optic cable as the receiver. Imperfections in the fibre reflect the laser when strain is placed on the cable by sound waves in the ground. The laser is inside a box called an interrogator which collects the information from the reflections caused by the strain on the fibre to create a seismic data set. One benefit to DAS is that it is sensitive; it can even detect footsteps near a fibre-optic cable. Thus, this form of acquisition may bring us a step closer to the film's results.

Distributed acoustic sensing in seismic acquisition is the main focus of this dissertation. This paper compiles various studies conducted over the last few years focusing on aspects of and applications to distributed acoustic sensing acquired data; however, it begins with an exploration of seismic processing techniques in general seismic data and culminates with an emphasis on applications to distributed acoustic sensing.

We break this thesis into three parts. The first part, consisting of Chapters 2 to 4, explores seismic experiments. In Chapter 2, we discuss the reflection problem in seismic imaging. Solving for reflection coefficients are a recurring problem in geophysical experiments, especially seismic modeling. Knowledge of these coefficients helps create a more accurate model. Thomson first described the reflectivity method in Thomson (1950) for finding these reflections. Kind computed reflection coefficients for layered media in Kind (1976). Müeller provides a tutorial of the method in Mueller (1985). A great deal of work has been done into solving for reflection coefficients. More work on calculating reflection coefficients in a variety of different ways can be found in Ursenbach and Hasse (2004), Ma et al. (2004), Sharma and Ferguson (2010), Daley et al. (2010), Lira et al. (2011), Lamoureux et al. (2012), Lamoureux et al. (2013), and Jabbari and Innanen (2015). Chapter 2 presents an extension of the work done in Lamoureux et al. (2012) and Lamoureux et al. (2013). Using plane waves, we solve for reflection and transmission coefficients when a velocity jump and a velocity ramp is present in different dimensions. We then compare the results of the exact solution of the wave equation in multiple dimensions using the analytic equations to the reflection and transmission coefficients to numerical solutions of the wave equation produced using finite difference methods.

Chapter 3 provides a description of wavelets and wavelet transforms. Wavelets have been around since 1910 when Haar developed what is known as the Haar wavelet transform today (Mallat, 2009). Their use in geophysical applications began in the 1980s with the work of Morlet et al. (1982a), Morlet et al. (1982b), Grossmann and Morlet (1984), and Goupillaud et al. (1984). Within the field of geophysics, wavelets were of particular interest given their nature as local transforms, unlike the Fourier transform which affects the entire data set (Fleet, 2008). From 1989 to 1995, the works of Mallat, Daubechies, Meyer, Chui, Wornell, and Holschneider provided further insight into wavelets; cf. Mallat (1989b), Mallat (1989a), Daubechies (1988), Daubechies (1992), Meyer (1992), Chui (1992a), Wornell (1995), and Holschneider (1995), respectively. More information about wavelets can be found in Meyer (1993), Benedetto and Frazier (1993), Chui (1992b), Ruskai et al. (1992), Farge et al. (1993), Beylkin et al. (1991), and Meyer and Roques (1993). For a textbook on wavelets, see Mallat (2009). The author provides an overview of discrete wavelet transforms in Fleet (2008). The authors describe the geophysical applications of wavelets in Kumar and Foufoula-Georgiou (1997). One particular development in wavelet analysis of interest to us is the tree-structured wavelet transform (Chang and Kuo, 1993). In Chapter 3, we explain how the tree-structured wavelet transform is employed to analyze data. The wavelets in a tree-structure wavelet transform correspond to a specific frequency band and scale the original data to smaller sizes. With this in mind, we then introduce the inverted wavelet tree which allows for scale invariance in image-recognition techniques, especially in Chapter 10.

The final chapter of this part, Chapter 4, contains an exploration of time-frequency analysis methods. Many of the techniques in this chapter relay on wavelet transforms such as the Continuous Wavelet Transform and Synchrosqueezing Transform. The Continuous Wavelet Transform was developed by Morlet and Grossman (Grossmann and Morlet, 1984). Thakur and Wu employed the works of Daubechies, Lu, and Wu on Synchrosqueezing theory to develop the Synchrosqueezing Transform which retrieves the instantaneous frequency information from a signal (Thakur and Wu, 2011). Basis Pursuit is a inverse problem which uses a wavelet dictionary to decompose a signal based on a true global optimization (Chen et al., 1998). Time-frequency analysis was first developed in the late 1990s (Chopra and Marfurt, 2006). Its introduction enables the consideration of a variety of seismic attributes such as the amplitude attribute and phase attribute. Seismic attributes provide more information about seismic data. In Han et al. (2015), the authors consider the amplitude and phase attributes produced once a time-frequency analysis method is applied to seismic data. They also developed a new seismic attribute: the derivative of corrected phase attribute which highlights the edges of anomalies in seismic data. The authors of Han et al. (2015) focus on the application of their new attribute to one seismic data set processed using Basis Pursuit. In Chapter 4, we expand the application of their new seismic attribute to different time-frequency analysis methods as well as more seismic data.

We include Chapters 2 and 4 for a complete record of the work done over the course of the author's doctoral program. While these two chapters focus on data that is not collected using a DAS system, they both deal with one of the overarching themes of this thesis: event detection. Anomalies in data tell the investigator where reflections occur in the physical process. Using time-frequency analysis, we consider a method which detects the edges of anomalies in seismic data, thus providing an estimate of the event's location. The following two parts of the thesis, encompassing Chapters 5 to 11, focus on event detection using distributed acoustic sensors.

The second part of the thesis comprises Chapters 5 to 7 and focuses on Distributed Acoustic Sensing (DAS) and fibre-optic cables. In Chapter 5, we introduce the acquisition technique called Distributed Acoustic Sensing. Fibre-optic sensing was invented over 40 years ago (Culshaw and Kersey, 2008). Curtis D. Kissinger patented the technology in 1967 (Kimbell, 2013). The invention of single mode optical fibre a decade later led to ideas of combing the fibre with interferometers. While originally employed for communication technologies, the use of fibre-optic cables and interferometry finds applications in a variety of areas including monitoring and acquisition. The authors provide more information about fibre optic sensing in Grattan and Meggitt (2000). In 2017, Hartog published a work explaining various fibre optic sensing methods including Distributed Vibration Sensing, or Distributed Acoustic Sensing, as well as various experiments of these sensing techniques (Hartog, 2017). Chapter 5 contains and a detailed explanation how the distributed acoustic sensor uses interferometry and strain on a fibre-optic cable to collect information.

Creating numerical models of seismic data is often the first step in any seismic experiment. Since the rise of distributed acoustic sensing, modeling such data has become necessary. Innanen introduces a geometrical model of a DAS fibre response in Innanen (2016). In Eaid et al. (2017), the authors create a 3D elastic finite difference wave model to model the response of arbitrarily shaped fibres. Innanen and Eaid went on to develop a model for a variety of fibre formations following the CREWES fibre geometry and sensing model

(Eaid and Innanen, 2018). Martins et al. addressed the question of modeling DAS systems attached to fibre-optic cables used in telecommunications and provide a tutorial for its development (Martins et al., 2017). In Masoudi and Newson (2017), the authors provided a numerical model the building blocks of the distributed acoustic sensor which allowed for an investigation of each of these parts on their own. Combining these modeled building blocks produce a model of the sensor's response. The authors of Sherman et al. (2019) developed numerical models of the DAS response and employed them to identify methods for finding event locations and geometry. They then build on this model by accounting for different physical aspects such as rock physics, fluid flow, and elastic wave propagation. In Chapter 6, we utilize analytic methods to produce models of DAS data with an elementary comparison of the effect different parameters cause in the data. We model the amount of stretching at a point s along the fibre-optic cable of a DAS sensor by using the strain tensor ϵ and the path of the fibre.

Chapter 7 investigates questions regarding the structure of the fibre-optic cable in DAS acquisition projects and provides bounds for the amplitude response of the fibre; see Theorems 7.6 and 7.7. We find that the bounds depended on the eigenvalues of the strain tensor. The amplitude response thus depends on the least and greatest fractional change that occurs along the entire fibre in the DAS system. The structure of the fibre-optic cable plays an important part in the response of the DAS sensor to a source. Two formations of particular interest are straight laid fibre and helically wound fibre; see Kuvshinov (2015), Hartog (2017), and Boer et al. (2013). Figure 1.1 provides an example of what we mean by straight fibre (top) versus helically wound fibre (bottom). One issue with straight fibre is that it is less sensitivity to waves that hit the fibre perpendicularly (Wuestefeld and Wilks, 2019). Development of the helically wound fibre for DAS systems occurred to increase broadside sensitivity; cf. Wuestefeld and Wilks (2019), Hornman et al. (2013), Boer et al. (2013), Mateeva et al. (2014), Yavuz et al. (2016), and Hornman (2016). The introduction of helically wound fibre brought the question of the fibre's sensitivity in a DAS system to the

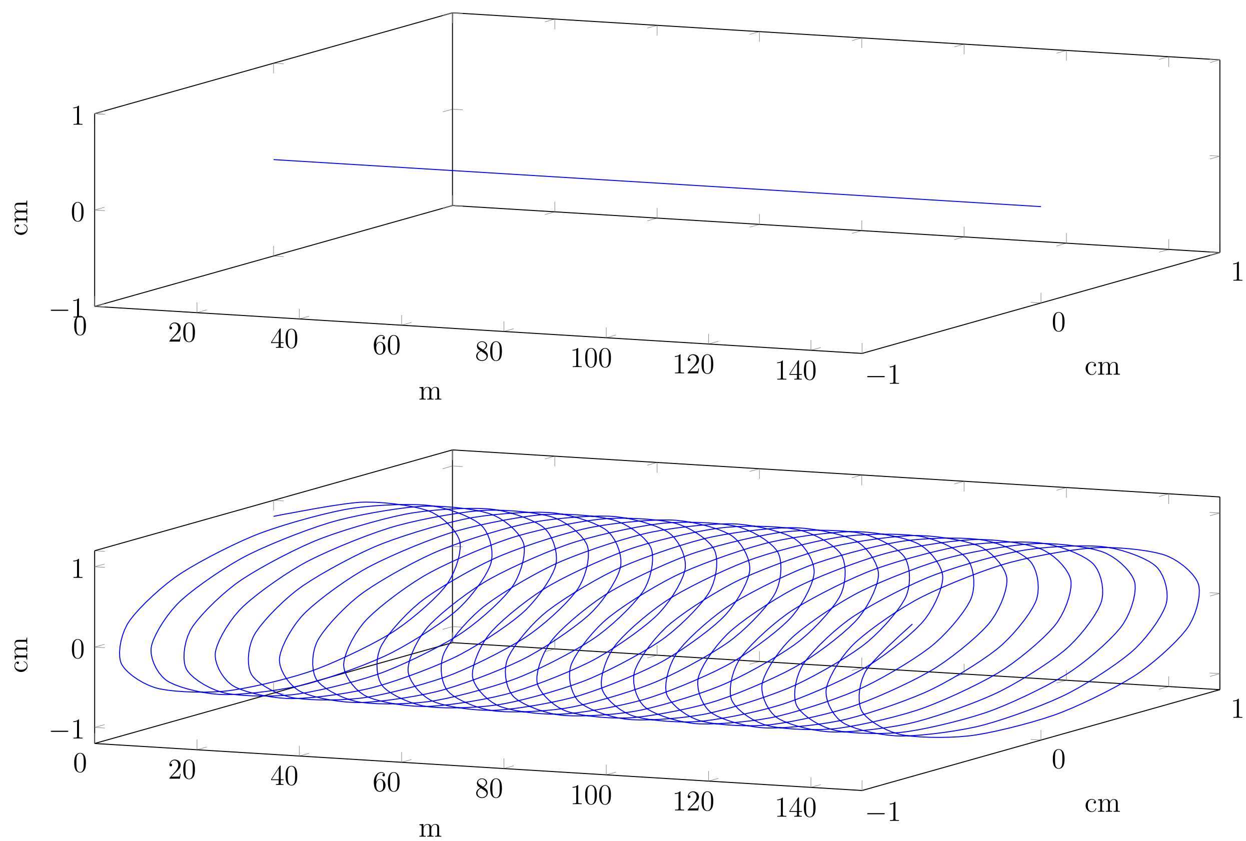


Figure 1.1: (Top) An example of a straight fibre optic cable in a DAS sensor. (Bottom) An example of a helically wound fibre-optic cable in a DAS sensor.

forefront. Gabai et al. developed a new parameter for defining DAS sensitivity; see Gabai and Eyal (2016). This new parameter allows the comparison of different distributed acoustic sensors. In Gabai and Eyal (2017), the authors generalized the parameter to define DAS sensitivity in a more intuitive manner. They also show how the static backscatter profile and DAS sensitivity are related. In Eaid and Innanen (2018), the authors provide models of the sensitivity of DAS-acquired data given different geometric formations of the fibre. In Chapter 7, we provide further insight into the investigation of DAS sensitivity. Using the classical homotopy theory of topological spaces, we also produce a homotopy which deforms the helically wound fibre to the straight fibre; cf. Theorems 7.13 and 7.14. The deformation allows us to offer comparisons of the two different formations of fibre-optic cable in a DAS sensor by studying the data over the homotopy as well as the strain matrices at each point in the data over the homotopy.

The final part consists of Chapters 8 – 11, and provides several studies on applications to DAS acquired data. In Chapter 8, we use DAS and fibre-optic cables to acquire vertical seismic profiles (VSP) of the two observation wells at the Containment and Monitoring Institute's Field Research Station in Newell County, AB. Mestayer et al. introduced the use of DAS to acquire walk-away vertical seismic profiles (Mestayer et al., 2011). They concluded that installing fibre permanently in wells would allow for low-cost nonintrusive geophysical monitoring due to the DAS results for the walk-away VSP providing nearly equivalent images to conventional borehole geophone methods. Mateeva et al. shows the results of similar field trails (Mateeva et al., 2012). They also investigated its employment in VSP applications such as checkshots, imaging, and time-lapse monitoring, and concluded the the use of DAS would make VSP a viable experiment in a variety of wells including ones which were not accessible to geophones. Nizkous et al. considered the configuration of the fibre which provided similar results to a 3D VSP survey with conventional geophones (Nizkous et al., 2015). At the CaMI FRS, Gordon investigates two VSP surveys focused on the use of DAS (Ferrebus, 2019). One survey tests the use of DAS to acquire zero offset VSPs and another survey compares the results of helically wound and straight fibre optic cables from a walk away VSP acquisition. Chapter 8 contains the results of two different data collections at the site collected using an interrogator from Fotech Solutions. We show the results of the straight laid fibre in both wells and use standard VSP processing techniques to identify the caprock which will be employed to contain the CO^2 in future experiments. The caprock is identifiable in both studies which verifies the use of DAS as a viable option for monitoring and the acquisition of vertical seismic profiles.

Chapters 9 and 10 contain event detection experiments involving machine learning. Given a distributed acoustic sensor can generate terabytes worth of data over the course of an acquisition, machine learning is an natural step towards locating points of interest in large data. The authors of Aktaş et al. (2017) apply a deep convolutional neural network to data collected using a phase-optical time domain reflectometry (OTDR) sensing system to identify events in the data. Prior to application of their CNN, they propose a new preprocessing denoising method which they apply to the data before having the CNN identify six types of events. The authors found that their proposed denoising method increased the accuracy of their CNN in classifying events. Huot and Biondi used unsupervised learning to identify recurring events which might interfere with locating the signals of useful events; see Huot and Biondi (2018). They taught a CNN to denoise the data and combined a CNN with a Markov decision process to map the acquired data to its geometry. An general adversarial neural network (GAN) is utilized in Shiloh et al. (2018) to transform synthetic data of a DAS response to mimic real data in order to build a training set. Using these GAN generated images for the training set, they trained a neural network to identify between walking and noise in a DAS-acquired data set. The authors continue to develop this work in Shiloh et al. (2019). Bublin compares the results of using classical deep learning techniques and an approaching combing image processing and deep learning to detect events in DAS data (Bublin, 2019). Their work shows that the combination of image processing and deep learning holds the most promise for event detection. In Shi et al. (2019), the authors use only the time-space data from a Φ-OTDR as the CNN's input in order to classify five types of events in the data.

Many of the current practices for event detection using machine learning involve neural networks for image recognition. In Chapter 9, we utilize a cluster analysis technique in Chapter 9 called Gaussian Mixture Models (GMM) to cluster the signal of a vehicle near the fibre in a DAS system. Shamsa and Paydayesh employed independent component analysis and Gaussian mixture models to detect a microseismic signal (Shamsa and Paydayesh, 2019). Their goal was to separate the fracture signal from the background noise in the data. Inspired by Shamsa and Paydayesh (2019), we use a GMM for foreground detection of the vehicle signal with independent component analysis to highlight the vehicle's signal. We apply this method in an attempt to make it easier to detect events in DAS data. While increasing the precision of the detection, the independent component analysis does not produce a better detection when there is only one event to consider. Future investigations for this technique should application to data containing more than one type of signal to offer a better judgment of how independent component analysis would improve results. With only the vehicle's signal, the GMM often provided the best detection of the signal though less precise than when independent component analysis was employed.

Chapter 10 provides an exploration of employing convolutional neural networks for image-recognition purposes in DAS data. We use the UFLDL Deep Learning tutorial to build a convolutional neural network that finds and identifies events in DAS-acquired data (Ng et al., 2013). In the first example, we train a convolutional neural network to identify footsteps near the fibre-optic cable. The convolutional neural network is then utilized to identify microseismic events in microseismic data. In the second example, a convolutional neural network learns to identify walking events, digging events, and noise events near the fibre. We test the convolutional neural network on two types of data acquired using a DAS sensor. Our results suggest that there are several degrees of freedom to investigate with regards to image registration when using image-recognition techniques to locate events in DAS-acquired data. Not only should the velocity, shift, and scale of an event be considered, but also the different variables used when acquiring data with a DAS sensor such as the pulse-repetition-frequency and the gauge length.

Chapter 11 deals with the concept of finding the exact location of events in the real world based on how they appear in the data. Matt McDonald proposed a similar problem with respect to events in microseismic data acquired using a DAS system at the 2015 PIMS Industrial Problem Solving Workshop (McDonald, 2015). He suggested minimizing the residual between the arrival times of the wavefronts in the data to those computed using the Eikonal equations for the model. In Webster et al. (2016), the authors discuss how a deviated well, or a well which is not entirely straight, provides the equivalent of a second and third component sensor when using a distributed acoustic sensor. The deviation allows for the direction of the event in the data to be determined. Using this technique, they compared the results of locating microseismic events with a DAS system to pinpointing the location of events using three component geophones. The authors found that the DAS method was often within 1 to 2m from the event unlike the geophone method which was between 20 and 50m from the event. The authors of Williams et al. (2017) also investigated this question of locating microseismic events in DAS-acquired data. They began by modeling the one-component DAS data and experimenting with choosing P-wave arrivals as well as P-wave and S-wave arrivals given different geometries of the DAS system. As with Webster et al. (2016), they found that deviating the fibre in the sensor improved locating the events. In Chapter 11, we introduce the Hyperbola Method which uses the hyperbolic nature of signals in DAS data to calculate the distance between the signal and the fibre-optic cable. We assume that the fibre used to create the data is not deviated. The Hyperbola Method provides more knowledge about where events are with respect to the fibre by using the distance between the P-wave and S-wave responses. This method allows for the potential to locate fibre-optic cables which are used in DAS acquisition projects.

Finally, we conclude in Chapter 12 with a discussion of future work which can be explored from the work presented in these chapters and draw our final conclusions about the work contained in this dissertation. For now, we begin this study by discussing a fundamental concept of seismic imaging: the reflection coefficient.

Chapter 2

Exact Solutions of Reflection Coefficients for the Acoustic Wave Equation

Reflection coefficients are fundamental to detecting geological structures below the Earth's surface. These coefficients describe the portion of the seismic wave which is reflected by the anomaly in the subsurface. Given the complicated nature of the Earth's subsurface, an exact solution of the wave equation which accurately describes its structure cannot always be found. As such, reflection coefficients are computed numerically so that we can use this information to discover where these objects reside underground. Unfortunately, numerical results do not offer the same amount of precision in detection as an exact solution.

In order to judge how successful numerical methods estimate reflections in seismic models, we solve for exact solutions of reflection coefficients in this chapter for plane waves giving varying velocities. We examine the case when plane waves are present, as it is possible to solve the wave equation exactly in this case. Therefore, the exact solution of the reflection coefficient for these equations is obtainable.

We focus on two specific velocity wavefields for our models: a velocity jump and a velocity ramp. The velocity jump models a quick jump from a low velocity to a higher velocity while the velocity ramp provides a model of the speed of a seismic wave through a transition zone. For the velocity ramp, we find that the reflection coefficients are dependent on the frequency in each dimension. For the second and third dimension, the frequency depends on the angle at which the wave hits the velocity ramp. We show the reflection coefficient results for different angles in two-dimensions for one solution to the acoustic wave equation and see that the reflection coefficient decreases as the frequency increases at each of the five angles we choose.

We consider a one-dimensional numerical model and a two-dimensional numerical model at the end of this chapter. While some error is expected, given that we are using numerical methods, the size of that error should be relatively small if we are to consider it an adequate model of the wave equation for those cases. We used finite difference methods to create both models. Between the two models, the error of the numerical solution to the wave equation increases by a factor of 7. The one-dimensional model differs from the exact solution to the wave equation produced using the reflection and transmission coefficients by about 0.03. The approximated two-dimensional solution varies from the exact solution by approximately 0.2. The error is less than 1% for the one-dimensional case when considering the difference of the norms of the exact and approximated solutions divided by the norm of the exact solution and 1.64% for the two dimensional case.

In the following section, we introduce the reflection problem for various velocity fields. In the first section, we outline the reflection problem. Then, we offer a summary of work done in Lamoureux et al. (2012) and Lamoureux et al. (2013) for the one-dimensional case and extend the results for general velocity jumps and ramps in one dimension. In the following two sections, we extend the velocity fields to two and three dimensions before solving for their respective reflection coefficients. In the final section, we compare the exact solution using the reflection coefficient equations and the approximated solution produced using finite difference methods.

2.1 The reflection problem

With regards to the velocity field, we focus on a velocity field with two specific traits. We are interested in velocities varying only in the z-direction. We restrict the velocity to one dimension, the z-direction, as that is the traditional downward direction in geophysics, in order to limit the complexity of the problem. As the dimension of the model increases, the more complicated the problem becomes; limiting to one dimension makes the problem more manageable for computational purposes. In the geophysics literature, such a restriction is called the 1.5D or 2.5D version of the problem.

The second trait of interest is when a velocity field has a transition zone. As discussed previously, a velocity $c(z)$ has a transition zone if a portion of $c(z)$ is non-constant. For instance, in Hardeman and Lamoureux (2016), we considered a velocity field

$$c(z) = \begin{cases} c_1 & z < z_0; \\ f(z) & z_0 \leq z \leq z_0 + L; \\ c_2 & z_0 + L < z; \end{cases} \tag{2.1}$$

where c_1, c_2, z_0 are constants, $f(z)$ is an arbitrary continuous function of z, and L is the length of the transition zone. The velocity varies for some function $f(z)$ for $z_0 \leq z \leq z_0 + L$. We denote this region as our transition zone. For $z < z_0$, we denote this as the top region as it is above the transition zone; for $z_0 + L < z$, we have the bottom region because it is below the transition zone. As an illustration, we will consider a velocity jump in each dimension.

Consider what occurs in a general reflection problem (Mueller, 1985). We model this problem using regional solutions. In the top region, there is a plane wave and reflection from the start of the transition region that we write as

$$u_{\text{top}} = u_{\text{inc}} + u_{\text{ref}}. \tag{2.2}$$

The incident wave, defined as $u_{\text{inc}} = u(\mathbf{x}, z, t)$, for some solution u of the acoustic wave equation

$$\rho(z)\frac{\partial^2 u}{\partial t^2} = \nabla(K(z) \cdot \nabla u) \tag{2.3}$$

where $\rho(z) = c^{\alpha-2}(z)$ and $K(z) = c^{\alpha}(z)$, for some parameter $\alpha \in \mathbb{N}$. The function u represents the displacement field. The term $\mathbf{x}$ depends on the dimension we are discussing. For instance, if we are focusing on the one-dimensional case, then $u(\mathbf{x}, z, t) = u(z, t)$. In two dimensions, $u(\mathbf{x}, z, t) = u(x, z, t)$, etc. Generally, $u(\mathbf{x}, z, t) = u(x_1, x_2, ..., x_{m-1}, z, t)$. The incident wave u_{inc} is positive in the z-direction, as that is conventionally chosen as the downward direction in geophysical problems. The reflection is defined as

$$u_{\text{ref}} = Ru(\mathbf{x}, -z, t), \tag{2.4}$$

where R is the reflection coefficient. As opposed to u_{inc}, the reflected portion of the wave u_{ref} will be negative in the z-direction. Hence,

$$u_{\text{top}}(\mathbf{x}, z, t) = u(\mathbf{x}, z, t) + Ru(\mathbf{x}, -z, t). \tag{2.5}$$

We designate what occurs in the transition region by

$$u_{\text{trans}}(\mathbf{x}, z, t) = u(\mathbf{x}, z, t). \tag{2.6}$$

We employ Separation of Variables to solve the acoustic wave equation which separates the PDE into ODEs for each variable. Typically, the solutions to the z-ODE would be of the form $Z(z) = AZ_1(z) + BZ_2(z)$ for arbitrary A, B and functions $Z_1(z)$ and $Z_2(z)$. In the bottom region, we have the portion of the wave which is transmitted through the transition zone

$$u_{\text{bottom}}(\mathbf{x}, z, t) = Tu(\mathbf{x}, z, t) \tag{2.7}$$

where T is the transmission coefficient. The solution u in each of these regional solutions is the portion of the exact solution u of the acoustic wave equation corresponding to that particular region.

Given that we are considering piecewise functions, we must enforce some continuity conditions. We wish to have displacement continuity:

$$u_{\text{top}} = u_{\text{trans}} \quad \text{at } z = z_0; \tag{2.8}$$

$$u_{\text{trans}} = u_{\text{bottom}} \quad \text{at } z = z_0 + L; \tag{2.9}$$

and continuity of force:

$$K_{\text{top}} \frac{\partial}{\partial z} \left(u_{\text{top}} \right) = K_{\text{trans}} \frac{\partial}{\partial z} \left(u_{\text{trans}} \right) \quad \text{at } z = z_0; \tag{2.10}$$

$$K_{\text{trans}} \frac{\partial}{\partial z} \left(u_{\text{trans}} \right) = K_{\text{bottom}} \frac{\partial}{\partial z} \left(u_{\text{bottom}} \right) \quad \text{at } z = z_0 + L. \tag{2.11}$$

After applying these continuity conditions to the regional solutions, we get a system of equations which we use to solve for the reflection coefficient R.

For simplification, we write

$$\frac{\partial}{\partial x_i} f = f_{x_i}; \tag{2.12}$$

$$\frac{\partial^2}{\partial x_i^2} = f_{x_i x_i}; \tag{2.13}$$

for variables x_i and functions f.

2.2 One-Dimensional Reflection Coefficients

In one dimension, we denote the solutions to the one-dimensional acoustic wave equation,

$$\rho(z) u_{tt} = \left(K(z) u_z \right)_z, \tag{2.14}$$

as $u(z,t)$.

In Lamoureux et al. (2012), the authors found reflection coefficients of solutions to the one-dimensional acoustic wave equation

$$\rho(x)u_{tt} = \left(K(x)u_x\right)_x, \tag{2.15}$$

where they established a relationship between the density $\rho(x)$ and the bulk modulus $K(x)$ using some parameter $\alpha \in \mathbb{N}$. They defined this relation as $\rho(x) = c^{\alpha-2}(x)$ and $K(x) = c^{\alpha}(x)$, where $c(x)$ is the velocity. The relation preserves the ratio

$$\frac{K(x)}{\rho(x)} = c^2(x). \tag{2.16}$$

The authors examined two different cases: varying density and varying modulus. Respectively, these are the cases when $\alpha = 0$ and $\alpha = 2$. They studied these cases for two different velocity ramps in the one-dimensional case and found the reflection coefficients for each case. In particular, they were interested in the velocity with a jump discontinuity

$$c(x) = \begin{cases} 1 & x < 0; \\ 2 & 0 < x; \end{cases} \tag{2.17}$$

and the velocity ramp

$$c(x) = \begin{cases} 1 & x < 1; \\ x & 1 < x < 2; \\ 2 & 2 < x. \end{cases} \tag{2.18}$$

After studying the case when $\alpha = 0, 2$ for the velocity ramp in Equation (2.18), the authors solved for an exact solution for the reflection coefficient with regards to the one-dimensional velocity ramp for general α:

$$R(\omega) = \frac{e^{2i\omega}(2^{n_1} - 2^{n_2})(n_1 + n_2)}{2^{n_2}(2i\omega + 2\sqrt{1/4 - \omega^2}) + 2^{n_1}(-2i\omega + 2\sqrt{1/4 - \omega^2})}, \tag{2.19}$$

where $n_{1,2} = (1 - \alpha) \pm \sqrt{(1-\alpha)^2/4 - \omega^2}$ for frequency ω.

The following year, the authors extended their work to find exact solutions for reflection coefficients given more general velocities in Lamoureux et al. (2013). In particular, they considered a velocity jump

$$c(x) = \begin{cases} c_1 & x < 0; \\ c_2 & 0 < x; \end{cases} \tag{2.20}$$

and a velocity ramp

$$c(x) = \begin{cases} c_1 & x < 0; \\ c_1 + mx & 0 \leq x \leq L; \\ c_2 & L < x; \end{cases} \tag{2.21}$$

where $m = (c_2 - c_1)/L$.

For the velocity ramp in Equation (2.21) when only the modulus varies ($\alpha = 0$), they found the reflection coefficient

$$R(\omega) = \frac{(r_2 - r_1)(n_1 + n_2)}{2(r_1 - r_2)(i\omega/m) + 2(r_1 + r_2)\sqrt{1/4 - (\omega/m)^2}}, \tag{2.22}$$

where $r_1 = (c_2/c_1)^{n_1}$, $r_2 = (c_2/c_1)^{n_2}$. The parameters n_1 and n_2 are defined as before but with $\alpha = 0$, i.e. $n_1 = -1/4 + \sqrt{1/4 - (\omega/m)^2}$ and $n_2 = -1/4 - \sqrt{1/4 - (\omega/m)^2}$ where ω is frequency.

2.2.1 One-dimensional velocity jump

In Lamoureux et al. (2013), the authors found the reflection coefficients when there is a velocity jump

$$c(z) = \begin{cases} c_1 & z < z_0; \\ \\ c_2 & z > z_0; \end{cases} \tag{2.23}$$

where $z_0 = 0$. In this section, we explore the case for any $z_0 \in \mathbb{R}$. Solving the one-dimensional acoustic wave equation with this velocity jump gives the solution

$$u(z,t) = \begin{cases} e^{i\omega(\pm z/c_1 - t)} & z < z_0; \\ \\ e^{i\omega(\pm z/c_2 - t)} & z > z_0. \end{cases} \tag{2.24}$$

Since there is no transition zone, we define the regional solutions to be

$$u_{\text{top}}(z,t) = e^{i\omega(z/c_1 - t)} + Re^{i\omega(-z/c_1 - t)}; \tag{2.25}$$

$$u_{\text{bottom}}(z,t) = Te^{i\omega(z/c_2 - t)}. \tag{2.26}$$

The continuity conditions for this problem are

$$u_{\text{top}} = u_{\text{bottom}} \quad \text{at } z = z_0; \tag{2.27}$$

$$K_{\text{top}} \left(u_{\text{top}} \right)_z = K_{\text{bottom}} \left(u_{\text{bottom}} \right)_z \quad \text{at } z = z_0; \tag{2.28}$$

thus the system of equations reduces to

$$e^{i\omega z_0/c_1} + Re^{-i\omega z_0/c_1} = Te^{i\omega z_0/c_2}; \tag{2.29}$$

$$c_1^{\alpha-1} \left(e^{i\omega z_0/c_1} - Re^{-i\omega z_0/c_1} \right) = c_2^{\alpha-1} Te^{i\omega z_0/c_2}. \tag{2.30}$$

Using these equations, we solve to find

$$R(\alpha,\omega) = -e^{i2z_0\omega/c_1}\frac{(c_1c_2^\alpha - c_1^\alpha c_2)}{(c_1c_2^\alpha + c_1^\alpha c_2)}; \tag{2.31}$$

$$T(\alpha,\omega) = 2e^{iz_0\omega/c_1}e^{-iz_0\omega/c_2}\frac{c_1^\alpha c_2}{c_1c_2^\alpha + c_1^\alpha c_2}. \tag{2.32}$$

We will compare this to the results we get in higher dimensions.

Impedance at $\alpha = 1$

Given the relation defined for the modulus $K(z)$ and the density $\rho(z)$, Equation (2.31) provides the opportunity to compare the effect of varying the modulus and density on the reflection which occurs when the wave crosses a jump discontinuity. As described earlier in Lamoureux et al. (2012), the authors took particular interest in the cases when $\alpha = 0$ and $\alpha = 2$. They noted that a consequence in the different choice of α shows a flip in polarity once the wave is reflected.

Consider the case when $\alpha = 1$. Using Equation (2.31), we find

$$R(1,\omega) = -e^{i2z_0\omega/c_1}\frac{(c_1c_2^1 - c_1^1c_2)}{(c_1c_2^1 + c_1^1c_2)} = 0. \tag{2.33}$$

The fact that impedance remains constant in the case when $\alpha = 1$ provides a good explanation for why the reflection coefficient $R(1,\omega) = 0$. Recall that the impedance is given by $I = \rho(z)c(z)$. For $\alpha = 1$, we get

$$I = \rho(z)c(z) = c^{1-2}(z)c(z) = \frac{1}{c(z)}c(z) = 1. \tag{2.34}$$

Reflections are dependent on a change in the impedance between layers in the Earth. If the impedance is constant, then there should not be a reflection. The case $\alpha = 1$ shows this fact.

2.2.2 One-dimensional velocity piecewise ramp

In Lamoureux et al. (2013), the authors also found the reflection coefficient for the one-dimensional acoustic wave equation when a velocity ramp is present. We will consider general one-dimensional velocity ramps:

$$c(z) = \begin{cases} c_1 & z < z_0; \\ c_{\text{trans}}(z) & z_0 \leq z \leq z_0 + L; \\ c_2 & z_0 + L < z; \end{cases} \tag{2.35}$$

where L is the length of the transition zone. In the 1D case, there is only the normal incidence case. We find that solutions to the 1D acoustic wave equation are of the form:

$$u(x,t) = \begin{cases} e^{i\omega(\pm z/c_1 - t)} & z < z_0; \\ Z(z)e^{-i\omega t} & z_0 \leq z \leq z_0 + L; \\ e^{i\omega(\pm z/c_2 - t)} & z_0 + L < z; \end{cases} \tag{2.36}$$

where $Z(z) = AZ_1(z) + BZ_2(z)$ for constants A, B. Given the presence of a transition zone, the regional solutions are

$$u_{\text{top}}(z,t) = e^{i\omega(z/c_1 - t)} + Re^{i\omega(-z/c_1 - t)}; \tag{2.37}$$

$$u_{\text{trans}}(z,t) = AZ_1(z)e^{-i\omega t} + BZ_2(z)e^{-i\omega t}; \tag{2.38}$$

$$u_{\text{bottom}}(z,t) = Te^{i\omega(z/c_2 - t)}; \tag{2.39}$$

which gives the following system of equations when we apply the continuity conditions from Equations (2.8) to (2.10)

$$e^{i\omega(z_0/c_1-t)} + Re^{i\omega(-z_0/c_1-t)} = [AZ_1(z_0) + BZ_2(z_0)]e^{-i\omega t}; \tag{2.40}$$

$$[AZ_1(z_0+L) + BZ_2(z_0+L)]e^{-i\omega t} = Te^{i\omega((z_0+L)/c_2-t)}; \tag{2.41}$$

$$\frac{i\omega}{c_1}e^{i\omega(z_0/c_1-t)} - \frac{i\omega}{c_1}Re^{i\omega(-z_0/c_1-t)} = [AZ_1'(z_0) + BZ_2'(z_0)]e^{-i\omega t}; \tag{2.42}$$

$$[AZ_1'(z_0+L) + BZ_2'(z_0+L)]e^{-i\omega t} = \frac{i\omega(z_0+L)}{c_2}Te^{i\omega((z_0+L)/c_2-t)}. \tag{2.43}$$

We reduce this system of equations to give

$$e^{i\omega z_0/c_1} + Re^{-i\omega z_0/c_1} = AZ_1(z_0) + BZ_2(z_0), \tag{2.44}$$

$$AZ_1(z_0+L) + BZ_2(z_0+L) = Te^{i\omega(z_0+L)/c_2}, \tag{2.45}$$

$$\frac{i\omega}{c_1}e^{i\omega z_0/c_1} - \frac{i\omega}{c_1}Re^{-i\omega z_0/c_1} = AZ_1'(z_0) + BZ_2'(z_0), \text{ and} \tag{2.46}$$

$$AZ_1'(z_0+L) + BZ_2'(z_0+L) = \frac{i\omega}{c_2}Te^{i\omega(z_0+L)/c_2}. \tag{2.47}$$

We turn this into the matrix problem

$$\begin{bmatrix} Z_1(z_0) & Z_2(z_0) & -e^{-i\omega z_0/c_1} & 0 \\ Z_1(z_0+L) & Z_2(z_0+L) & 0 & -e^{i\omega(z_0+L)/c_2} \\ Z_1'(z_0) & Z_2'(z_0) & \frac{i\omega}{c_1}e^{-i\omega z_0/c_1} & 0 \\ Z_1'(z_0+L) & Z_2'(z_0+L) & 0 & -\frac{i\omega}{c_2}e^{i\omega(z_0+L)/c_2} \end{bmatrix} \begin{bmatrix} A \\ B \\ R \\ T \end{bmatrix} = \begin{bmatrix} e^{i\omega z_0/c_1} \\ 0 \\ \frac{i\omega}{c_1}e^{i\omega z_0/c_1} \\ 0 \end{bmatrix}. \tag{2.48}$$

Solving this equation, we get

$$R(\alpha,\omega) = e^{i2z_0\omega/c_1}\frac{\omega^2 N_1 - c_1c_2N_2 + i\omega(c_1N_3 + c_2N_4))}{\omega^2 N_1 + c_1c_2N_2 + i\omega(c_2N_4 - c_1N_3)}; \tag{2.49}$$

$$T(\alpha,\omega) = -e^{i\omega(z_0+L)/c_2}e^{i\omega z_0/c_1}\frac{i2c_2\omega N_5}{\omega^2 + c_1c_2N_2 + i\omega(c_2N_4 - c_1N_3)}; \tag{2.50}$$

where

$$N_1 = Z_1(z_0 + L)Z_2(z_0) - Z_2(z_0 + L)Z_1(z_0); \tag{2.51}$$

$$N_2 = Z_1'(z_0 + L)Z_2'(z_0) - Z_2'(z_0 + L)Z_1'(z_0); \tag{2.52}$$

$$N_3 = Z_1(z_0 + L)Z_2'(z_0) - Z_2(z_0 + L)Z_1'(z_0); \tag{2.53}$$

$$N_4 = Z_1'(z_0 + L)Z_2(z_0) - Z_2'(z_0 + L)Z_1(z_0); \tag{2.54}$$

$$N_5 = Z_1(z_0 + L)Z_2'(z_0 + L) - Z_2(z_0 + L)Z_1'(z_0 + L). \tag{2.55}$$

2.3 Two-dimensional reflection coefficients

For the two-dimensional case, we denote the solutions to the two-dimensional acoustic wave equation by

$$\rho(z)u_{tt} = \nabla(K(z)\nabla u) = K(z)u_{zz} + K'(z)u_z, \tag{2.56}$$

where $\nabla = (\partial_x, \partial_z)$ as $u(x, z, t)$. In the following sections, we study a velocity jump first and the general case of a piecewise velocity with a transition zone in two dimensions.

2.3.1 Two-dimensional velocity jump

First, we consider a 2D velocity jump

$$c(z) = \begin{cases} c_1 & z < z_0; \\ c_2 & z > z_0; \end{cases} \tag{2.57}$$

where c_1 and c_2 are constant. We discuss this velocity here in order to compare the reflection provided when a ramp is present versus when one is not present.

For this case, a plane wave moves downward across a velocity jump. We represent the wave in the top half by

$$u_{\text{top}}(x,z,t) = e^{i\omega(z/c_1 - t)} + Re^{i\omega(-z/c_1 - t)}, \tag{2.58}$$

and the wave which is transmitted across the velocity jump as

$$u_{\text{bottom}}(x,z,t) = Te^{i\omega(z/c_2 - t)}, \tag{2.59}$$

where R is the reflection coefficient and T is the transmission coefficient.

Again, we want to enforce displacement continuity across the jump as well as continuity of force. We write the continuity conditions as

$$u_{\text{top}} = u_{\text{bottom}} \quad \text{at } z = z_0; \tag{2.60}$$

$$K_{\text{top}} \nabla(u_{\text{top}}) = K_{\text{bottom}} \nabla(u_{\text{bottom}}) \quad \text{at } z = z_0; \tag{2.61}$$

this results in

$$e^{i\omega(z_0/c_1 - t)} + Re^{i\omega(-z_0/c_1 - t)} = Te^{i\omega(z_0/c_2 - t)}; \tag{2.62}$$

$$K_{\text{top}} \left[\frac{i\omega}{c_1} e^{i\omega(z_0/c_1 - t)} - \frac{i\omega}{c_1} Re^{i\omega(-z_0/c_1 - t)} \right] = K_{\text{bottom}} \left[T \frac{i\omega}{c_2} e^{i\omega(z_0/c_2 - t)} \right]; \tag{2.63}$$

which in turn implies

$$e^{i\omega z_0/c_1} + Re^{-i\omega z_0/c_1} = Te^{i\omega z_0/c_2}; \tag{2.64}$$

$$K_{\text{top}} \left[\frac{e^{i\omega z_0/c_1}}{c_1} - R \frac{e^{-i\omega z_0/c_1}}{c_1} \right] = K_{\text{bottom}} \left[\frac{e^{i\omega z_0/c_2}}{c_2} T \right]. \tag{2.65}$$

Recall that $K(z) = c^{\alpha}(z)$. Then, we find that

$$e^{i\omega z_0/c_1} + Re^{-i\omega z_0/c_1} = Te^{i\omega z_0/c_2} \tag{2.66}$$

$$c_1^{\alpha-1}\left[e^{i\omega z_0/c_1} - Re^{-i\omega z_0/c_1}\right] = c_2^{\alpha-1}\left[Te^{i\omega z_0/c_2}\right]. \tag{2.67}$$

This becomes the matrix problem

$$\begin{bmatrix} -e^{-i\omega z_0/c_1} & e^{i\omega z_0/c_2} \\ c_1^{\alpha-1}e^{-i\omega z_0/c_1} & c_2^{\alpha-1}e^{i\omega z_0/c_2} \end{bmatrix} \begin{bmatrix} R \\ T \end{bmatrix} = \begin{bmatrix} e^{i\omega z_0/c_1} \\ c_1^{\alpha-1}e^{i\omega z_0/c_1} \end{bmatrix} \tag{2.68}$$

which we solve to get the following solutions

$$R(\alpha,\omega) = -e^{i2z_0\omega/c_1}\frac{(c_1c_2^{\alpha} - c_1^{\alpha}c_2)}{(c_1c_2^{\alpha} + c_1^{\alpha}c_2)}; \tag{2.69}$$

$$T(\alpha,\omega) = e^{iz_0\omega/c_1}e^{-iz_0\omega/c_2}\frac{2c_1^{\alpha}c_2}{c_1c_2^{\alpha} + c_1^{\alpha}c_2}. \tag{2.70}$$

Given that we are considering the plane wave case in two dimensions where the wave is orthogonal to the velocity jump, it follows that the exact solutions for reflection and transmission coefficients would mimic the one-dimensional case. In fact, the two-dimensional reflection and transmission coefficient solutions given in Equations (2.69)) are identical to the one-dimensional reflection coefficient solution (cf. (2.31)) for this case.

2.3.2 Two-dimensional piecewise velocity ramp

We now examine what occurs for general two-dimensional piecewise velocity ramps

$$c(z) = \begin{cases} c_1 & z < z_0; \\ c_{\text{trans}}(z) & z_0 \le z \le z_0 + L; \\ c_2 & z_0 + L < z. \end{cases} \tag{2.71}$$

Unlike in the one-dimensional case, we now have to observe what occurs in two directions: the x-direction and the z-direction. The plane wave can hit the ramp at normal incidence or non-normal incidence.

In the normal incidence case, the plane wave is orthogonal to the transition zone of the velocity field. This means that the incident angle $\theta_1 = 0°$. For the two-dimensional problem of normal incidence, the plane wave is constant in the x-direction. The solution to Equation (2.56) is then, for $x \in \mathbb{R}$,

$$u(x,z,t) = \begin{cases} e^{i\omega(\pm z/c_1 - t)} & z < z_0; \\ Z(z)e^{-i\omega t} & z_0 \leq z \leq z_0 + L; \\ e^{i\omega(\pm z/c_2 - t)} & z_0 + L < z; \end{cases} \tag{2.72}$$

for $Z(z) = AZ_1(z) + BZ_1(z)$ for arbitrary $A,\ B \in \mathbb{C}$, where the real parameters k_x, k_z, ω satisfy the dispersion relation

$$k_x^2 + k_z^2 = \frac{\omega^2}{c_1^2}. \tag{2.73}$$

The regional solutions for this case are

$$u_{\text{top}}(x,z,t) = e^{i\omega(z/c_1 - t)} + Re^{i\omega(-z/c_1 - t)}; \tag{2.74}$$

$$u_{\text{trans}}(x,z,t) = AZ_1(z)e^{-i\omega t} + BZ_2(z)e^{-i\omega t}; \tag{2.75}$$

$$u_{\text{bottom}}(x,z,t) = Te^{i\omega(z/c_2 - t)}. \tag{2.76}$$

Applying the two-dimensional continuity conditions, we get the following system of equations

$$e^{i\omega z_0/c_1} + Re^{-i\omega z_0/c_1} = AZ_1(z_0) + BZ_2(z_0); \tag{2.77}$$

$$AZ_1(z_0+L) + BZ_2(z_0+L) = Te^{i\omega(z_0+L)/c_2}; \tag{2.78}$$

$$c_1^\alpha\left[\frac{i\omega}{c_1}e^{i\omega z_0/c_1} - \frac{i\omega}{c_1}Re^{-i\omega z_0/c_1}\right] = c_1^\alpha\left[AZ_1'(z_0) + BZ_2'(z_0)\right]; \tag{2.79}$$

$$c_2^\alpha\left[AZ_1'(z_0+L) + BZ_2'(z_0+L)\right] = c_2^\alpha T\frac{i\omega}{c_2}e^{i\omega(z_0+L)/c_2}; \tag{2.80}$$

which gives us the matrix problem

$$\begin{bmatrix} Z_1(z_0) & Z_2(z_0) & -e^{-i\omega z_0/c_1} & 0 \\ Z_1(z_0+L) & Z_2(z_0+L) & 0 & -e^{i\omega(z_0+L)/c_2} \\ Z_1'(z_0) & Z_2'(z_0) & \frac{i\omega}{c_1}e^{i\omega z_0/c_1} & 0 \\ Z_1'(z_0+L) & Z_2'(z_0+L) & 0 & -\frac{i\omega}{c_2}e^{i\omega(z_0+L)/c_2} \end{bmatrix} \begin{bmatrix} A \\ B \\ R \\ T \end{bmatrix} = \begin{bmatrix} e^{i\omega z_0/c_1} \\ 0 \\ \frac{i\omega}{c_1}e^{i\omega z_0/c_1} \\ 0 \end{bmatrix}. \tag{2.81}$$

Evaluating the matrix equation, we obtain

$$R(\alpha,\omega) = e^{i2\omega z_0/c_1}\frac{\omega^2 N_1 - c_1c_2N_2 + i\omega(c_1N_3 + c_2N_4)}{\omega^2 N_1 + c_1c_2N_2 + i\omega(c_2N_4 - c_1N_3)}; \tag{2.82}$$

$$T(\alpha,\omega) = -ie^{i2\omega(z_0+L)/c_2)}e^{i2\omega z_0/c_1}(1 + e^{i2\omega z_0/c_1})\frac{\omega c_2 N_5}{\omega^2 N_4 + c_1c_2N_2 + i\omega(c_2N_4 - c_1N_3)}. \tag{2.83}$$

These are the same N_i as defined in Equations 2.51.

For the non-normal incidence case, the plane wave hits the transition zone at an angle $\theta_1 \geq 0$. In normal incidence case, the wave was constant in the x-direction; however, that is not true for this case. The solution to Equation (2.56) is of the form

$$u(x,z,t) = \begin{cases} e^{i(k_x x \pm k_z z - \omega t)} & z < z_0; \\ Z(z)e^{i(k_x x - \omega t)} & z_0 \leq z \leq z_0 + L; \quad x \in \mathbb{R}; \\ e^{i(k_x x \pm k_z' z - \omega t)} & z_0 + L < z; \end{cases} \tag{2.84}$$

where $Z(z) = AZ_1(z) + BZ_2(z)$ with arbitrary $A,\ B \in \mathbb{C}$. The regional solutions are given by

$$u_{\text{top}}(x,z,t) = e^{i(k_x x + k_z z - \omega t)} + Re^{i(k_x x - k_z z - \omega t)}; \tag{2.85}$$

$$u_{\text{trans}}(x,z,t) = [AZ_1(z) + BZ_2(z)]e^{i(k_x x - \omega t)}; \tag{2.86}$$

$$u_{\text{bottom}}(x,z,t) = Te^{i(k_x x + k_z' z - \omega t)}. \tag{2.87}$$

Applying the two-dimensional continuity conditions and reducing, we obtain the system of equations

$$e^{ik_z z_0} + Re^{-ik_z z_0} = AZ_1(z_0) + BZ_2(z_0); \tag{2.88}$$

$$AZ_1(z_0 + L) + BZ_2(z_0 + L) = Te^{ik_z'(z_0+L)}; \tag{2.89}$$

$$ik_z e^{ik_z z_0} - iRk_z e^{-ik_z z_0} = AZ_1'(z_0) + BZ'(z_0); \tag{2.90}$$

$$AZ_1'(z_0 + L) + BZ_2'(z_0 + L) = ik_z' Te^{ik_z'(z_0+L)}; \tag{2.91}$$

which gives the matrix problem

$$\begin{bmatrix} Z_1(z_0) & Z_2(z_0) & -e^{-ik_z z_0} & 0 \\ Z_1(z_0+L) & Z_2(z_0+L) & 0 & -e^{ik_z'(z_0+L)} \\ Z_1'(z_0) & Z_2'(z_0) & ik_z e^{-ik_z z_0} & 0 \\ Z_1'(z_0+L) & Z_2'(z_0+L) & 0 & -ik_z' e^{ik_z'(z_0+L)} \end{bmatrix} \begin{bmatrix} A \\ B \\ R \\ T \end{bmatrix} = \begin{bmatrix} e^{ik_z z_0} \\ 0 \\ ik_z e^{ik_z z_0} \\ 0 \end{bmatrix}. \tag{2.92}$$

Solving this problem, we find the reflection and transmission coefficients have the form

$$R(\alpha,\omega) = e^{i2k_z z_0} \frac{N_2 + k_z k_z' N_1 + i(k_z' N_3 + k_z N_4)}{N_2 + k_z k_z' N_1 + i(k_z N_4 - k_z' N_3)}; \tag{2.93}$$

$$T(\alpha,\omega) = e^{-ik_z'(z_0+L)} e^{ik_z z_0} \frac{-i2k_z N_5}{N_2 + k_z k_z' N_1 + i(k_z N_4 - k_z' N_3)}. \tag{2.94}$$

These N_i are the same as the ones defined in Equations 2.51

2.4 Three-dimensional reflection coefficients

For the three-dimensional case, we denote the solutions to the three-dimensional acoustic wave equation

$$\rho(z)u_{tt} = \nabla(K(z)\nabla u), \tag{2.95}$$

where $\nabla = (\partial_x, \partial_y, \partial_z)$ as $u = u(x, y, z, t)$. We continue to use plane waves and consider a 1D Earth as with the previous cases. In the next few sections, we focus on finding reflection coefficients when given a three-dimensional velocity jump and a general three-dimensional velocity ramp.

2.4.1 Three-dimensional velocity jump

We begin with a three-dimensional velocity jump

$$c(z) = \begin{cases} c_1 & z < z_0; \\ c_2 & z > z_0; \end{cases} \tag{2.96}$$

where c_1 and c_2 are constant. For this case, a plane wave moves downward across a velocity jump. We represent the wave in the top half by

$$u_{\text{top}}(x, y, z, t) = e^{i\omega(z/c_1 - t)} + Re^{i\omega(-z/c_1 - t)}, \tag{2.97}$$

and the wave which is transmitted across the velocity jump as

$$u_{\text{bottom}}(x, y, z, t) = Te^{i\omega(z/c_2 - t)}, \tag{2.98}$$

where R is the reflection coefficient and T is the transmission coefficient.

These are the same regional solutions as the two-dimensional velocity jump; see Equations (2.58) and (2.59). The reflection and transmission coefficients will be the same for the three-dimensional as in two-dimensional case; see Equation (2.69). Let us compare the three-dimensional velocity ramp with the two-dimensional case as well.

2.4.2 Three-dimensional piecewise velocity ramp

In this section, we discuss the reflection coefficient given a general 3D piecewise velocity ramp

$$c(z) = \begin{cases} c_1 & z < z_0; \\ c_{\text{trans}}(z) & z_0 \le z \le z_0 + L; \\ c_2 & z_0 + L < z; \end{cases} \tag{2.99}$$

where $c_{\text{trans}}(z)$ is the velocity ramp and L is the length of the ramp.

In the normal incidence case, we get the solution, for $x, y \in \mathbb{R}$,

$$u(x, y, z, t) = \begin{cases} e^{i\omega(z/c_1 - t)} & z < z_0; \\ Z(z)e^{-i\omega t} & z_0 \le z \le z_0 + L; \\ e^{i\omega(z/c_2 - t)} & z_0 + L < z; \end{cases} \tag{2.100}$$

where $Z(z) = AZ_1(z) + BZ_2(z)$ for arbitrary A, $B \in \mathbb{C}$. The regional solutions for this case are

$$u_{\text{top}}(x, y, z, t) = e^{i\omega(z/c_1 - t)} + Re^{-i\omega(z/c_1 - t)}; \tag{2.101}$$

$$u_{\text{trans}}(x, y, z, t) = [AZ_1(z) + BZ_2(z)]e^{-i\omega t}; \tag{2.102}$$

$$u_{\text{bottom}}(x, y, z, t) = Te^{i\omega(z/c_2 - t)}; \tag{2.103}$$

and when we apply the three-dimensional continuity conditions, we get

$$e^{i\omega(z_0/c_1-t)} + Re^{-i\omega(z_0/c_1-t)} = [AZ_1(z_0) + BZ_2(z_0)]e^{-i\omega t}; \tag{2.104}$$

$$[AZ_1(z_0+L) + BZ_2(z_0+L)]e^{-i\omega t} = Te^{i\omega((z_0+L)/c_2-t)}; \tag{2.105}$$

$$\frac{i\omega z_0}{c_1}e^{i\omega(z_0/c_1-t)} - \frac{i\omega z_0}{c_1}Re^{-i\omega(z_0/c2-t)} = (AZ_1'(z_0) + BZ_2'(z_0))e^{-i\omega t}; \tag{2.106}$$

$$[AZ_1'(z_0+L) + BZ_2'(z_0+L)]e^{-i\omega t} = T\frac{i\omega(z_0+L)}{c_2}e^{i\omega((z_0+L)/c_2-t)}. \tag{2.107}$$

We reduce this to

$$e^{i\omega z_0/c_1} + Re^{-i\omega z_0/c_1} = AZ_1(z_0) + BZ_2(z_0); \tag{2.108}$$

$$AZ_1(z_0+L) + BZ_2(z_0+L) = Te^{i\omega(z_0+L)/c_2}; \tag{2.109}$$

$$\frac{i\omega z_0}{c_1}e^{i\omega z_0/c_1} - \frac{i\omega z_0}{c_1}Re^{-i\omega z_0/c_2} = AZ_1'(z_0) + BZ_2'(z_0); \tag{2.110}$$

$$AZ_1'(z_0+L) + BZ_2'(z_0+L) = T\frac{i\omega(z_0+L)}{c_2}e^{i\omega(z_0+L)/c_2}; \tag{2.111}$$

which becomes the matrix problem

$$\begin{bmatrix} Z_1(z_0) & Z_2(z_0) & e^{-i\omega z_0/c_1} & 0 \\ Z_1(z_0+L) & Z_2(z_0+L) & 0 & -e^{i\omega(z_0+L)/c_2} \\ Z_1'(z_0) & Z_2'(z_0) & \frac{i\omega}{c_1}e^{-i\omega z_0/c_2} & 0 \\ Z_1'(z_0+L) & Z_2'(z_0+L) & 0 & -\frac{i\omega}{c_2}e^{i\omega(z_0+L)/c_2} \end{bmatrix} \begin{bmatrix} A \\ B \\ R \\ T \end{bmatrix} = \begin{bmatrix} e^{i\omega z_0/c_1} \\ 0 \\ \frac{i\omega}{c_1}e^{i\omega z_0/c_1} \\ 0 \end{bmatrix}. \tag{2.112}$$

This is the same matrix problem as the one for the two-dimensional case (cf. Equation (2.81)). Therefore, the exact solution for the reflection coefficient will be the same.

In the non-normal incidence case, the three-dimensional solutions to the acoustic wave equation are of the form, for $x, y \in \mathbb{R}$,

$$u(x,y,z,t) = \begin{cases} e^{i(k_x x + k_y y \pm k_z z - \omega t)} & z < z_0; \\ [AZ_1(z) + BZ_2(z)]e^{i(k_x x + k_y y - \omega t)} & z_0 \le z \le z_0 + L; \\ e^{i(k_x x + k_y y \pm k'_z z - \omega t)} & z_0 + L < z. \end{cases} \tag{2.113}$$

The regional solutions for this case are

$$u_{\text{top}}(x,y,z,t) = e^{i(k_x x + k_y y + k_z z - \omega t)} + Re^{i(k_x x + k_y y - k_z z - \omega t)}; \tag{2.114}$$

$$u_{\text{trans}}(x,y,z,t) = [AZ_1(z) + BZ_2(z)]e^{i(k_x x + k_y y - \omega t)}; \tag{2.115}$$

$$u_{\text{bottom}}(x,y,z,t) = Te^{i(k_x x + k_y y + k'_z z - \omega t)}. \tag{2.116}$$

Applying the continuity conditions, we get the system

$$e^{i(k_x x + k_y y - \omega t)}[e^{ik_z z_0} + Re^{-ik_z z_0}] = [AZ_1(z_0) + BZ_2(z_0)]e^{i(k_x x + k_y y - \omega t)}; \tag{2.117}$$

$$[AZ_1(z_0 + L) + BZ_2(z_0 + L)]e^{i(k_x x + k_y y - \omega t)} = Te^{i(k_x x + k_y y + k'_z (z_0 + L) - \omega t)}; \tag{2.118}$$

$$c_1^{\alpha} e^{i(k_x x + k_y y - \omega t)}[ik_z e^{ik_z z_0} - ik_z Re^{-ik_z z_0}] = c_1^{\alpha}[AZ'_1(z_0) + BZ'_2(z_0)]e^{i(k_x x + k_y y - \omega t)}; \tag{2.119}$$

$$c_2^{\alpha}[AZ'_1(z_0 + L) + BZ'_2(z_0 + L)]e^{i(k_x x + k_y y - \omega t)} = c_2^{\alpha} ik'_z Te^{i(k_x x + k_y y + k'_z (z_0 + L) - \omega t)}; \tag{2.120}$$

which reduces to

$$e^{ik_z z_0} + Re^{-ik_z z_0} = AZ_1(z_0) + BZ_2(z_0); \tag{2.121}$$

$$AZ_1(z_0 + L) + BZ_2(z_0 + L) = Te^{ik'_z (z_0 + L)}; \tag{2.122}$$

$$ik_z e^{ik_z z_0} - ik_z Re^{-ik_z z_0} = AZ'_1(z_0) + BZ'_2(z_0); \tag{2.123}$$

$$AZ'_1(z_0 + L) + BZ'_2(z_0 + L) = ik'_z Te^{ik'_z (z_0 + L)}. \tag{2.124}$$

We transform these equations into the matrix problem

$$\begin{bmatrix} Z_1(z_0) & Z_2(z_0) & -e^{-ik_z z_0} & 0 \\ Z_1(z_0+L) & Z_2(z_0+L) & 0 & -e^{ik'_z(z_0+L)} \\ Z'_1(z_0) & Z'_2(z_0) & ik_z e^{-ik_z z_0} & 0 \\ Z'_1(z_0+L) & Z'_2(z_0+L) & 0 & -ik'_z e^{ik'_z(z_0+L)} \end{bmatrix} \begin{bmatrix} A \\ B \\ R \\ T \end{bmatrix} = \begin{bmatrix} e^{ik_z z_0} \\ 0 \\ ik_z e^{ik'_z z_0} \\ 0 \end{bmatrix}. \tag{2.125}$$

Recall this is the same matrix as for the two-dimensional case; see Equation (2.92). Therefore, the equation for the reflection and transmission coefficients will be the same.

2.5 Example of reflection coefficients for a two-dimensional velocity ramp

In the previous sections, we found the exact equation for reflection and transmission coefficients for multiple dimensions. We saw that the two-dimensional and three-dimensional exact solutions for the reflection coefficients mirrored their one-dimensional counterparts when considering the normal incidence case. In this section, we study the results of a non-normal incidence case for the velocity ramp

$$c(x,z) = \begin{cases} 1000 & z < 0; \\ 1000 + 100z & 0 \le z \le 10; \\ 2000 & 10 < z; \end{cases} \quad x \in \mathbb{R} \tag{2.126}$$

when $\alpha = 0$. We will find an equation for the reflection and transmission coefficients for arbitrary incidence angle θ_1 and then consider specific cases of θ_1.

We start with a simple wave equation in two dimensions where the velocity depends on depth z:

$$\frac{1}{c(z)^2}\frac{\partial^2 u}{\partial t^2} = \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial z^2}. \tag{2.127}$$

As we noted in the previous paragraph, we can set $\alpha = 0$ in Equation (2.56) to get Equation (2.127). Moreover, we employ the geophysical convention that z is the vertical dimension and increases as you go down.

Recall that the velocity field in this case is constant $c_1 = 1000\ m/s$ in the upper region $z < 0$, constant $c_2 = 2000\ m/s$ in the lower region $z > 10$ deeper than $L = 10$ m, and linear in the transition region $0 < z < 10$ with $c(z) = 1000 + mz$ where $m = (c_2 - c_1)/L = 100$ is the slope for the change in velocity with depth.

Using separation of variables, a basic solution in the upper region will have the form

$$u(x, z, t) = e^{i(k_x x + k_z z + \omega t)} \tag{2.128}$$

where the real parameters k_x, k_z, ω satisfy the dispersion relation; see Equation (2.73).

The solution in the lower region $z > 10$ will be similar except the dispersion relation will use velocity $2000 m/s$ instead.

The transition region is only slightly harder. Again, using separation of variables $u(x, z, t) = X(x)Z(z)T(t)$, we obtain the usual exponential solutions for X and T, $X(x) = \exp(ik_x x)$, $T(t) = \exp(i\omega t)$. Thus we are left with a single ODE for $Z(z)$ in the form

$$\frac{Z''}{Z} + \frac{\omega^2}{c(z)^2} = k_x^2. \tag{2.129}$$

To avoid any confusion, let us write out the velocity function explicitly, to see the ODE as

$$\frac{Z''}{Z} + \frac{\omega^2}{(1000 + 100z)^2} = k_x^2. \tag{2.130}$$

This equation can be solved explicitly using computational tools such as Wolfram Alpha, and the general solution is a linear combination of two Whittaker functions,

$$Z(z) = AM_{0,\beta}(2c(z)k_x/m) + BW_{0,\beta}(2c(z)k_x/m), \tag{2.131}$$

where M and W are the Whittaker functions and $\beta = \sqrt{100^2 - 4\omega^2}/200$. We see that β is a function of the ratio ω/s which is how the frequency dependence of our reflection coefficients will enter in the problem. This makes physical sense; the frequency ω, compared to the slope m of the transition zone, is what matters to the reflection.

The case $k_x = 0$ is special, corresponding to the normal incident case (plane waves that that have no x-dependence). In this case, the solutions are of the form

$$Z(z) = A(1000 + 100z)^{1/2+\beta} + B(1000 + 100z)^{1/2-\beta}, \tag{2.132}$$

using the same parameter β as above.

Now, we are ready to compute the plane wave reflection coefficients. The idea is similar to the previous case: Start with a plane wave in the upper region, add in its reflection in the upper region, then match coefficients at the line $z = 0$ to find equations to specify parameters A, B in the solution (cf. Equation (2.132)). You also need equations at the line $z = 10$ to connect the solution in the transition zone to the plane wave transmitted into the lower region $z > 10$.

We begin by fixing parameters k_x, k_z, ω, all positive and satisfying the relation given by Equation (2.73). We set the incident plane wave to be

$$u_{inc} = e^{i(k_x x + k_z z - \omega t)}. \tag{2.133}$$

Remembering our convention that z points down and that the parameters are positive, this is a wave traveling downwards in the direction of the vector $\mathbf{k} = (k_x, k_z)$.

From physical intuition, we expect the reflected wave to be going up, a mirror reflection of the plane wave:

$$u_{ref} = Re^{i(k_x x - k_z z - \omega t)}, \tag{2.134}$$

where we flipped the sign in front of the z term to get an up-going wave. We include the reflection coefficient R here, and note it may depend on the parameters k_x, k_z, ω.

First, in the upper region $z < 0$ we write the wave field as

$$u_{\text{top}} = u_{\text{inc}} + u_{\text{ref}} = e^{i(k_x x + k_z z - \omega t)} + R e^{i(k_x x - k_z z - \omega t)}. \tag{2.135}$$

Assume that the k_x and the ω parameters stay the same across all regions. Then, in the transition region we expect

$$u_{\text{trans}} = e^{i(k_x x - \omega t)} \left(A M_{0,\beta}(2c(z)k_x/s) + B W_{0,\beta}(2c(z)k_x/s) \right), \tag{2.136}$$

and in the lower region $z > 10$ we expect

$$u_{\text{bottom}} = T e^{i(k_x x + k_z' z - \omega t)}, \tag{2.137}$$

where notice we have a new parameter k_z' to find; this accounts for the fact that the direction of the plane wave will change when we get into the lower region.

Before we continue, we should discuss how to solve for k_z'. In fact, all that is required is to use Snell's Law. Define θ_1 as the angle of incidence and θ_2 as the angle of refraction. We can then write

$$k_x = \frac{\omega}{c_1} \sin\theta_1, \qquad k_z = \frac{\omega}{c_1} \cos\theta_1; \tag{2.138}$$

$$k_x' = \frac{\omega}{c_2} \sin\theta_2, \qquad k_z' = \frac{\omega}{c_2} \cos\theta_2. \tag{2.139}$$

Equating the two k_x's here gives us

$$\frac{\omega}{c_1} \sin\theta_1 = \frac{\omega}{c_2} \sin\theta_2, \tag{2.140}$$

or in other words

$$\frac{\sin\theta_1}{\sin\theta_2} = \frac{c_1}{c_2}, \tag{2.141}$$

which is Snell's Law.

Using Equation (2.93), we find the following reflection and transmission coefficients given the velocity ramp in Equation (2.126)

$$R(\alpha,\omega) = \frac{N_2 + k_z k_z' N_1 + i(k_z' N_3 + k_z N_4}{N_2 + k_z k_z' N_1 + i(k_z N_4 - k_z' N_3)}; \tag{2.142}$$

$$T(\alpha,\omega) = e^{-ik_z' 10} \frac{-i2k_z N_5}{N_2 + k_z k_z' N_1 + i(k_z N_4 - k_z' N_3)}; \tag{2.143}$$

where

$$N_1 = Z_1(10)Z_2(0) - Z_2(10)Z_1(0); \tag{2.144}$$

$$N_2 = Z_1'(10)Z_2'(0) - Z_2'(10)Z_1'(0); \tag{2.145}$$

$$N_3 = Z_1(10)Z_2'(0) - Z_2(10)Z_1'(0); \tag{2.146}$$

$$N_4 = Z_1'(10)Z_2(0) - Z_2'(10)Z_1(0); \tag{2.147}$$

$$N_5 = Z_1(10)Z_2'(10) - Z_2(10)Z_1'(10). \tag{2.148}$$

.

We can now apply Equation (2.142) to the solution of Equation (2.127) for the two-dimensional velocity ramp in Equation (2.126) and examine the reflection coefficient for different θ_1. Choosing a range of ω and specific θ_1, we find the remaining parameters k_x, k_z, and k_z' using Equation (2.138).

In Figure 2.1, we compare the results of six different values of the incident angle θ_1 for the reflection coefficient. We examine $\theta_1 = 5°, 10°, 15°, 20°, 25°$, and $30°$. Using Snell's Law (cf. Equation(2.141)), we find θ_2. In Figure (2.2), we compare the results of six different values of the incident angle θ for the transmission coefficient.

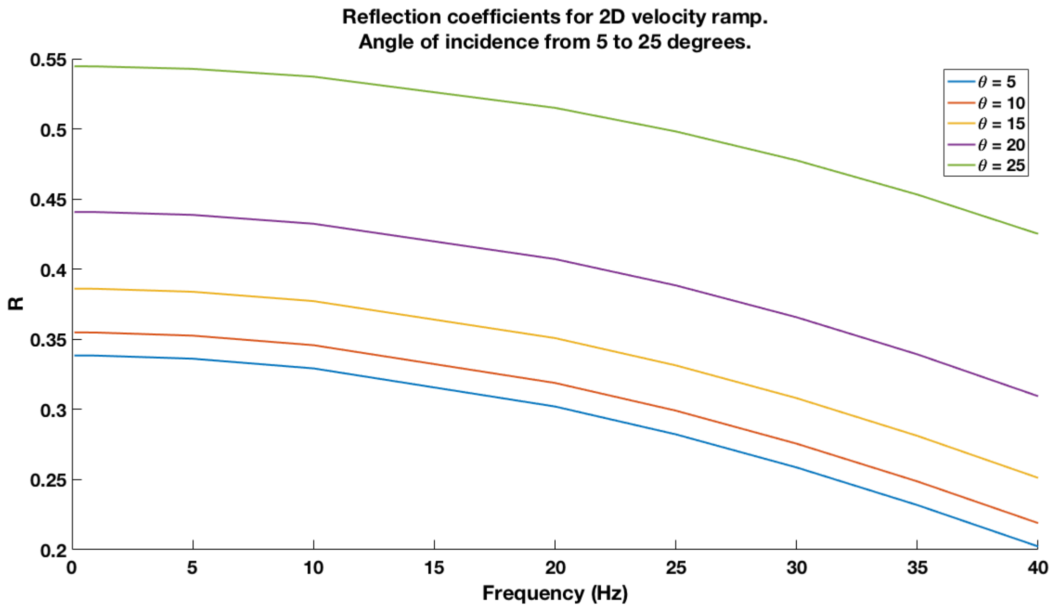


Figure 2.1: Reflection coefficients for non-normal incident case for the 2D velocity ramp.

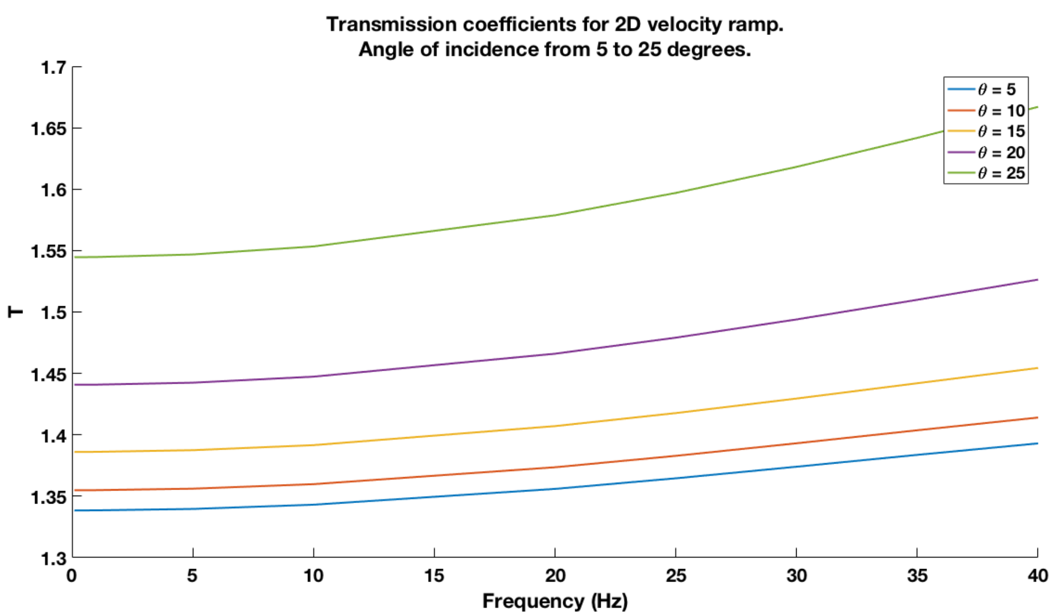


Figure 2.2: Transmission coefficients for non-normal incident case for the 2D velocity ramp.

2.6 Comparison of exact solutions to numerical results

In the geosciences, it is common to model solutions to the acoustic wave equation using numerical methods. In this section, we compare the exact solutions of the acoustic wave equation given a velocity ramp and some parameter α to numerical solutions modeled using finite difference methods. We employ the equations we found in the previous sections for the reflection and transmission coefficients in our finite difference scheme. We begin with the one-dimensional case. Then, we compare the exact and numerical solutions in the two-dimensional normal incidence case given a velocity ramp and some parameter α.

2.6.1 One-dimensional numerical results

Consider a velocity ramp

$$c(z) = \begin{cases} 1 & z < 1; \\ z & 1 \leq z \leq 2; \\ 2 & 2 < z. \end{cases} \tag{2.149}$$

For the one-dimensional model, we create a 1001×6001 grid for $z \in \{-1, ..., 5\}$ and $t \in \{0, ..., 4\}$. We use the explicit forward difference method. The time step is $dt = (4-0)/6000 = 1/1500$ and the spacial step is $dx = (5 - (-1))/1000 = 3/500$. We set the parameter $\alpha = 2$. For the initial and boundary conditions, we employed the boundary of the observed data generated by the reflection and transmission coefficients. Using the information from previous sections, we calculate the approximated solution given the reflection coefficient for the one-dimensional velocity ramp (cf. Equation (2.149)). Figure 2.3 shows the real part of the the exact and approximated solutions. Figure 2.4 depicts the imaginary part of the exact and approximated solutions. In both solutions, there is a distortion between $z = 1$ and $z = 2$. This distortion represents the velocity ramp described in Equation (2.149).

In Figure 2.5, we see the residuals for the real and imaginary parts. The difference between the exact and approximated solutions in both the real and imaginary parts is a

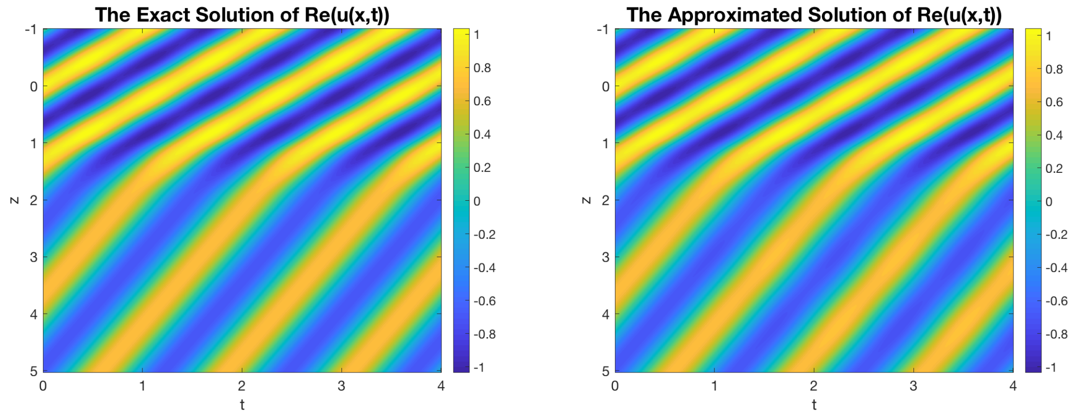


Figure 2.3: The time t is in units of seconds and the spacial term x is in units of meters.
(Left) The real part of the exact solution in one dimension.
(Right) The real part of the approximated solution in one dimension.

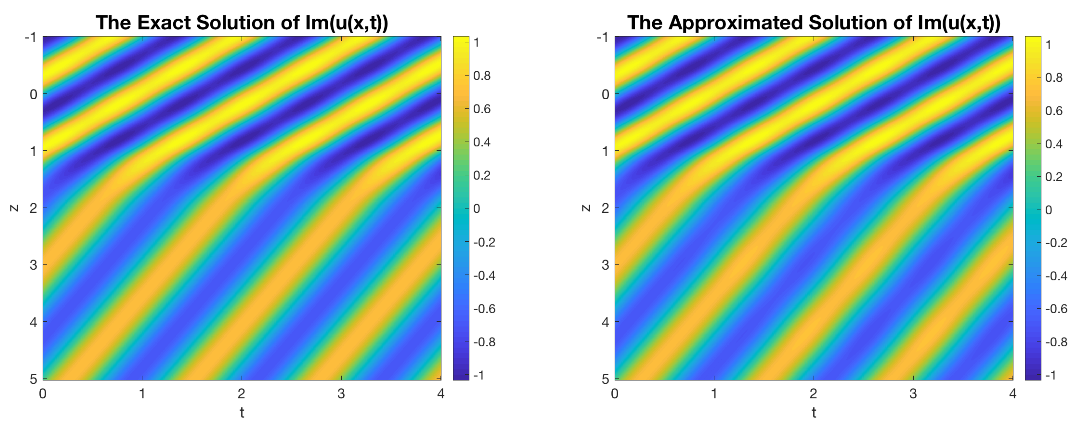


Figure 2.4: The time t is in units of seconds and the spacial term x is in units of meters.
(Left) The imaginary part of the exact solution in one dimension.
(Right) The imaginary part of the approximated solution in one dimension.

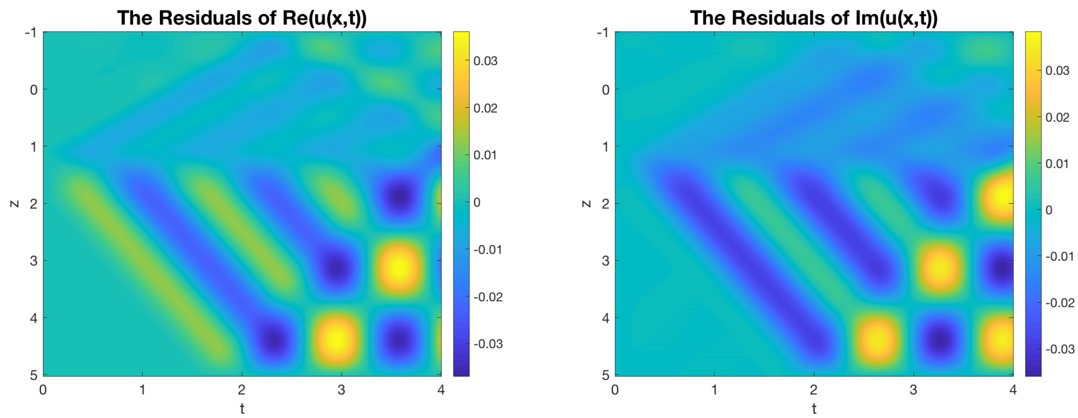


Figure 2.5: The time t is in units of seconds and the spacial term x is in units of meters. (Left) The residuals from the real part of the exact and approximated solutions. (Right) The residuals from the imaginary part of the exact and approximated solutions.

little greater than 0.03.

2.6.2 Two-dimensional numerical results

For the two-dimensional case, we extend the linear velocity ramp described in Equation (2.149) to

$$c(z) = \begin{cases} 1 & z < 1; \\ z & 1 \leq z \leq 2; \\ 2 & 2 < z, \end{cases} \quad x \in \mathbb{R}. \tag{2.150}$$

Given the two-dimensional velocity in Equation (2.150), we observe what occurs when the wave is normal to the velocity ramp described in Equation (2.150) in the two-dimensional case. We create a three-dimensional grid which is $101 \times 101 \times 601$ where $z \in \{-1, ..., 5\}$, $x \in \{-1, ..., 5\}$, and $t \in \{0, ..., 4\}$. The spacial steps are $dx = 3/50$ and $dz = 3/50$. The time step is $dt1/150$. Once again, we employed the explicite forward difference method. We set the parameter α equal to 2.

In Figures 2.6 and 2.7, we see the real part of the exact and approximated solutions to the two-dimensional acoustic wave equation at time $t = 0, 1, 2, 3,$ and 4 respectively. Figures

2.8 and 2.9 show the imaginary part of the exact and approximated solutions to the two-dimensional acoustic wave equation at $t = 0, 1, 2, 3$, and 4, respectively. In both the real and imaginary parts, we see the waves move further apart after passing through the transition zone $z = 1$ to $z = 2$ similar to the distortion that occurs in the transition zone in the one-dimensional case.

Figures 2.10 and 2.11 depict the difference between the real and imaginary parts of the exact and approximated solutions at $t = 0, 1, 2, 3$, and 4. In both cases, we see that the difference between the exact and approximated solutions is less than 0.2. As such, between one and two dimensions, the error increases by a factor of approximately 7.

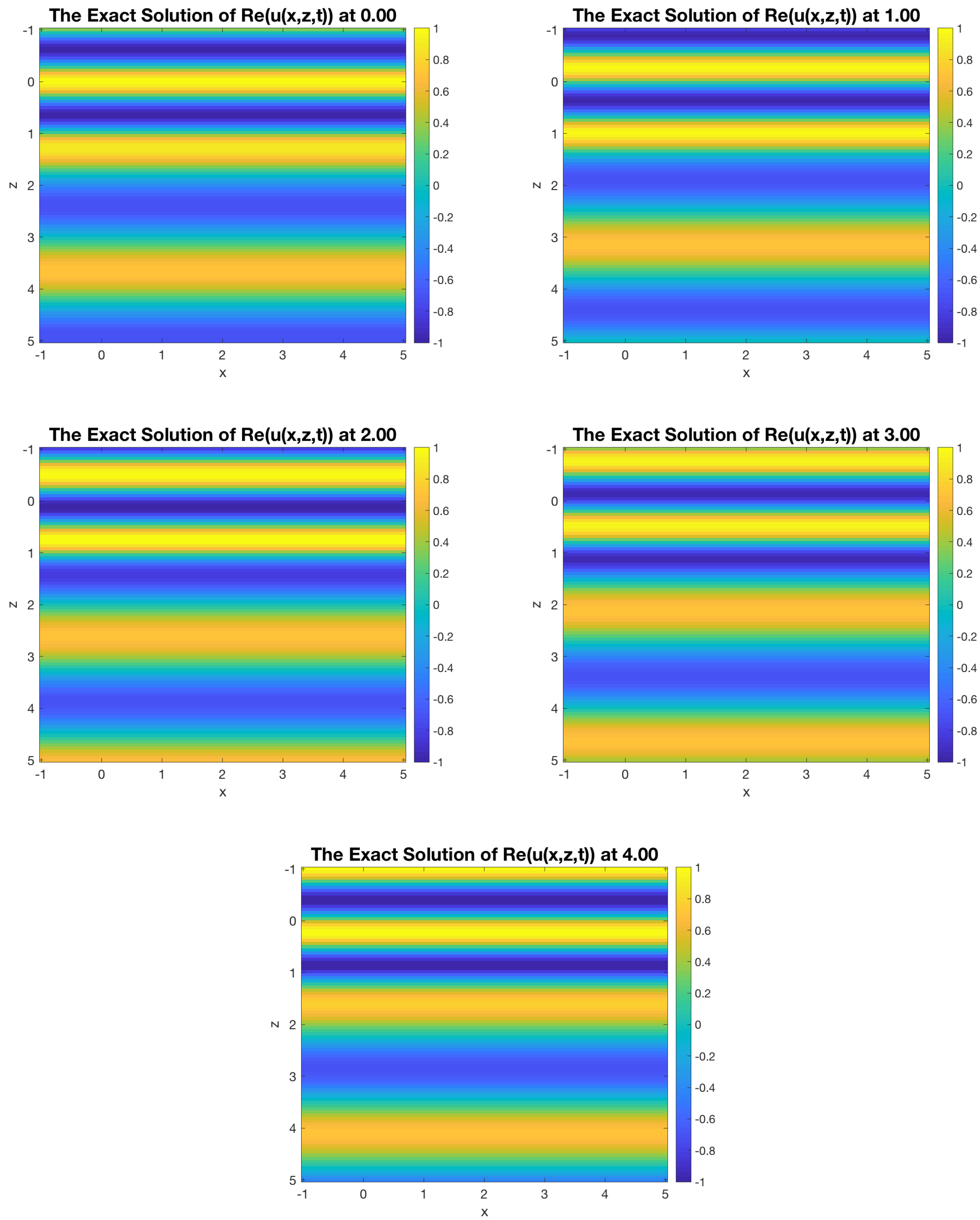


Figure 2.6: The real part of the exact solutions at (top left) $t = 0$, (top right) $t = 1$, (middle left) $t = 2$, (middle right) $t = 3$, and (bottom) $t = 4$.

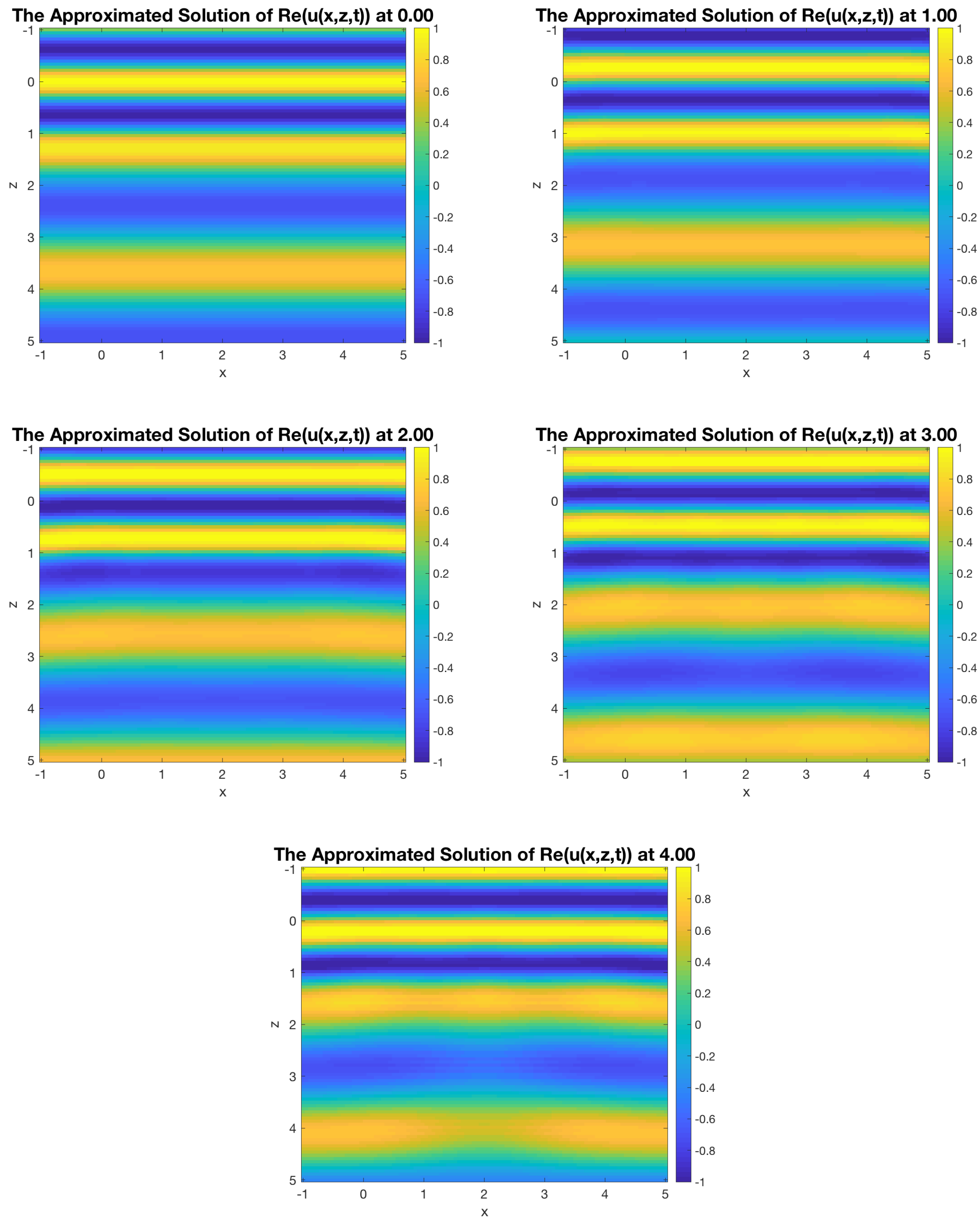


Figure 2.7: The real part of the approximated solutions at (top left) $t = 0$, (top right) $t = 1$, (middle left) $t = 2$, (middle right) $t = 3$, and (bottom) $t = 4$.

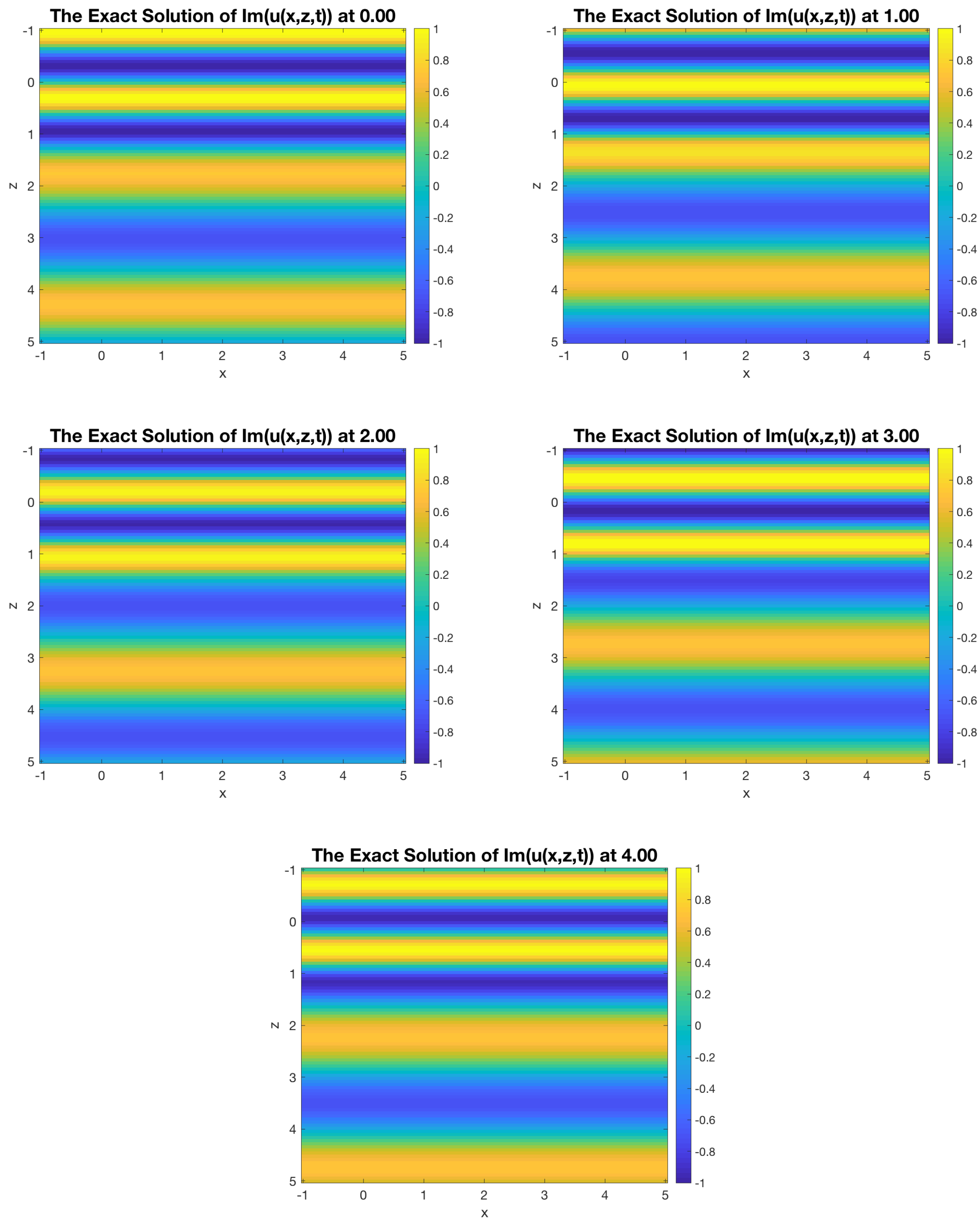


Figure 2.8: The time t is in units of seconds and the spacial terms x and z are in units of meters. The imaginary part of the exact solutions at (top left) $t = 0$, (top right) $t = 1$, (middle left) $t = 2$, (middle right) $t = 3$, and (bottom) $t = 4$.

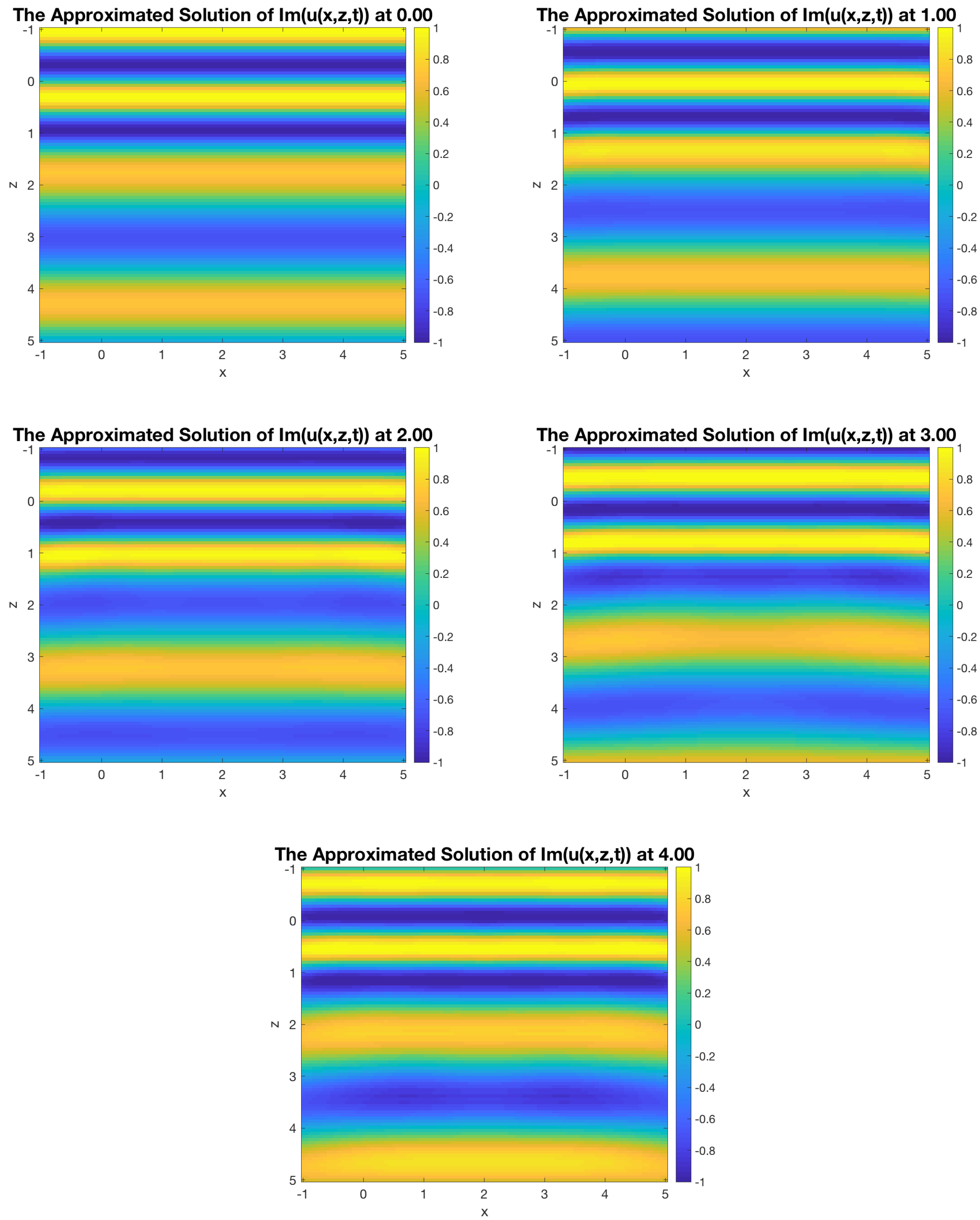


Figure 2.9: The time t is in units of seconds and the spacial terms x and z are in units of meters. The imaginary part of the exact and approximated solutions at (top left) $t = 0$, (top right) $t = 1$, (middle left) $t = 2$, (middle right) $t = 3$, and (bottom) $t = 4$.

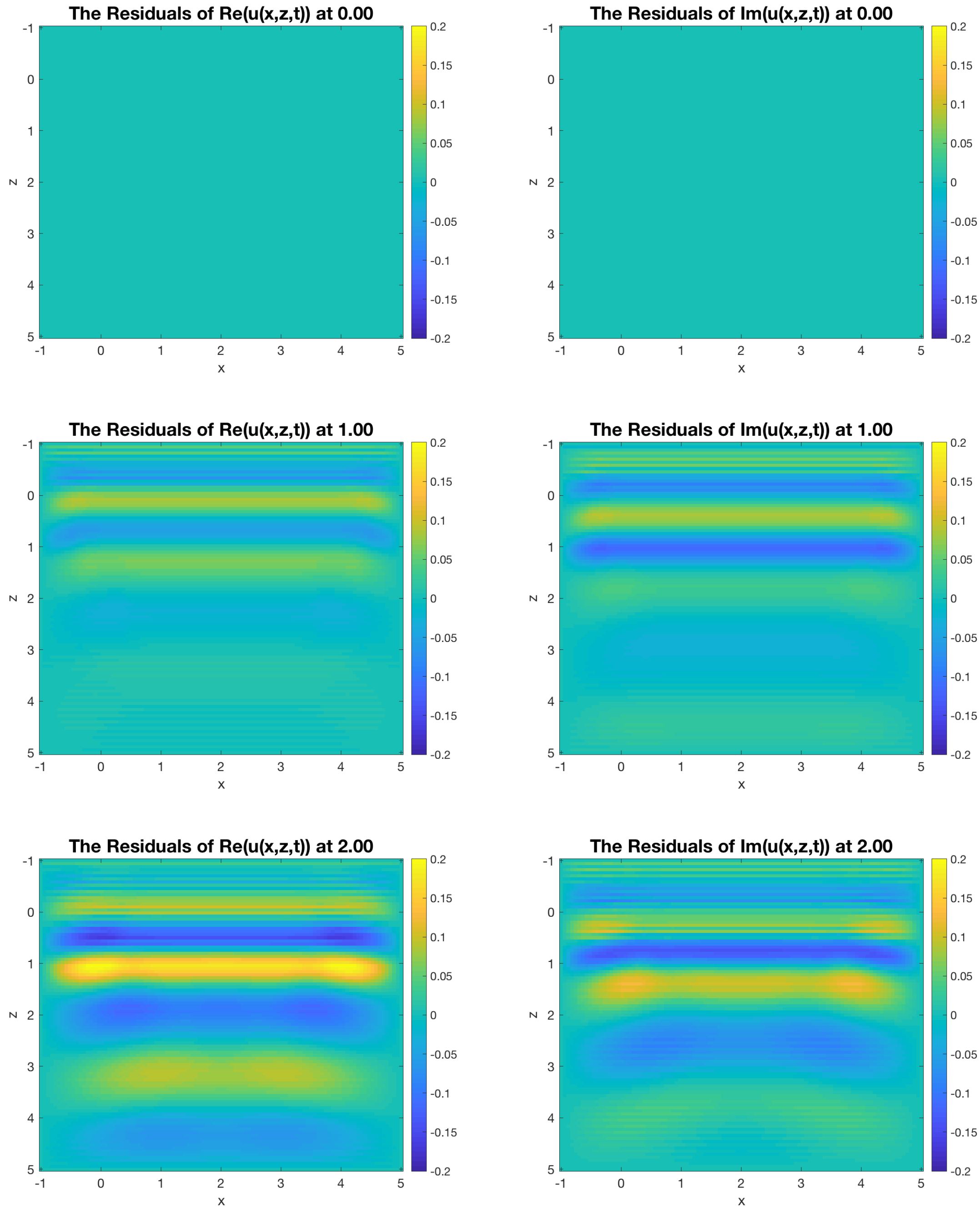


Figure 2.10: The time t is in units of seconds and the spacial terms x and z are in units of meters. The residuals from the real part of the exact and approximated solutions at (top left), (middle left), and (bottom left). The residuals from the imaginary part of the exact and approximated solutions (top right), (middle right), and (bottom right).

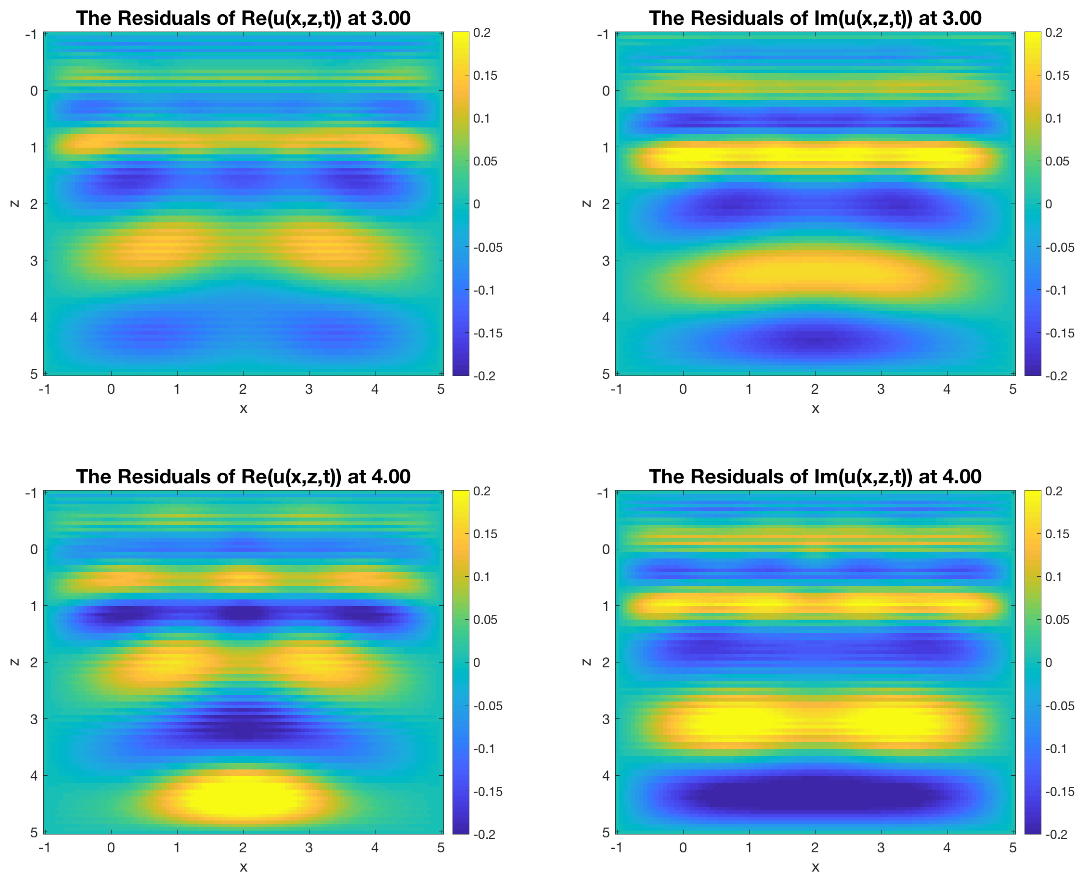


Figure 2.11: The time t is in units of seconds and the spacial terms x and z are in units of meters. The residuals from the real part of the exact and approximated solutions (top left), and (bottom left). The residuals from the imaginary part of the exact and approximated solutions (top right), and (bottom right).

Chapter 3

Introduction to wavelets

As we saw in the previous chapter, reflections are fundamental in seismic data. They provide information about what is under the earth's surface, how far down anomalies lie, and other related properties. In order to understand what reflections in seismic data mean, processing is required. One method of analyzing seismic data is through the use of wavelets; however, they are useful for more than simply processing seismic data (Fleet, 2008). We consider various applications of wavelets throughout the rest of this thesis; some applications involve seismic data sets, and others do not.

Our interest in wavelets arises from their ability to be windowed through large data and emphasize specific frequency bandwidths. Given the rise in data acquisition methods in the last decade, the amount of data in need of processing and interpretation has increased as well. Wavelets provide one answer to this problem; however, they also aid in other methods of data processing. We developed the inverted wavelet tree in order to apply machine learning techniques to large data. The inverted wavelet tree helps decrease the computation costs of detecting and identifying events in large data.

We begin the chapter by defining wavelets. We then examine the tree-structured wavelet transform and discuss how it motivates the inverted wavelet tree. Afterwards, we describe the inverted wavelet tree and how it will allow the application of neural networks for image recognition in large data with less computational costs.

3.1 Wavelets

In general terms, we can consider a wavelet as a small wave. In fact, the small nature of the wavelet is part of its appeal when it comes to analyzing data as the wavelet performs on a local level compared to Fourier transforms. For a more mathematical perspective, we define a wavelet as it is found in Mallat (2009).

Definition 3.1. A *wavelet* is a function $\psi \in \mathbf{L^2}(\mathbb{R})$ with a zero average:

$$\int_{-\infty}^{\infty} \psi(t)dt = 0. \tag{3.1}$$

Furthermore, wavelets are centered at $t = 0$ and satisfy the condition: $\|\psi(t)\| = 1$.

The condition $\|\psi(t)\| = 1$ means that wavelets are normal. A wavelet $\psi(t)$ satisfies the following properties (Fleet, 2008):

1. The function $\psi(t)$ is zero outside a finite interval;
2. The function $\psi(t)$ and its integer translates $\psi(t-k)$ for $k \in \mathbb{Z}$ form an orthonormal basis to $L^2(\mathbb{R})$.
3. The function $\psi(t)$ can be used to approximate smooth data, i.e. if $f \in L^2(\mathbb{R})$ is a smooth function, then there exists (ψ_n) of wavelets such that $\psi_n \longrightarrow f \in L^2(\mathbb{R})$.
4. For the function $\psi(t)$ and for any $k \in \mathbb{Z} \setminus \{0\}$,

$$\int_{\mathbb{R}} \psi(t)\psi(t-k)dt = 0. \tag{3.2}$$

5. The function $\psi(t)$ is a combination of dilations and translations of itself.

The first property of wavelets gives us the fact that a wavelet family contains a finite number of wavelets. This proves useful for maintaining a finite number of nodes containing data in the tree-structured wavelet transform. When considering discrete wavelet transforms, the integral in the fourth property establishes that the wavelet transform is orthogonal. We are most interested in the last property of wavelets. Given that a wavelet is a combination of dilations and translations of itself, then a wavelet transform can be built in such a way that it enables us to focus on specific segments of data. This beneficial aspect of wavelets is employed in the inverted wavelet tree.

With this in mind, the remainder of the chapter focuses on applications of wavelets to machine learning. It looks at using wavelet transforms to compress data to the size of a training image for neural networks. We will see further application of the following sections with regards to convolutional neural networks in Chapter 10.

3.2 Specific wavelets

In this thesis, we employ wavelets in spectral decomposition methods as well as machine learning. For the time-frequency analysis methods we will see in the next chapter, we utilize the Morlet wavelet (Ashmead, 2012).

Definition 3.2. The general form of the Morlet wavelet is given as:

$$\psi_{sl}(t) \equiv |s|^{-\frac{1}{2}} \left(e^{if\left(\frac{t-\ell}{s}\right)} - e^{-\frac{1}{2}f^2} \right) e^{-\frac{1}{2}\left(\frac{t-\ell}{s}\right)^2} \tag{3.3}$$

where the scale $s \in \mathbb{R}$ and the displacement $\ell \in \mathbb{R}$ are constants.

In Cohen (2019), the author explains that the Morlet wavelet preserves the temporal resolution of the original signal. In the frequency domain, it is shaped like a Gaussian curve. When convolved with the data, the smooth Gaussian minimizes ripple effects. This proves useful when we consider the new seismic attribute: the derivative of corrected phase as the diminished ripple affect means less misinterpretations as oscillations.

For wavelet analysis involving image recognition, we use the Haar wavelet as found in Mallat (2009).

Definition 3.3. The general form of the Haar wavelet is defined as:

$$\psi_{j,n} = \frac{1}{\sqrt{2^j}}\psi\Big(\frac{t-2^j n}{2^j}\Big) \tag{3.4}$$

for $(j,n) \in \mathbb{Z}^2$, where

$$\psi(t) = \begin{cases} 1 & \text{if } 0 \leq t \leq 1/2; \\ -1 & \text{if } 1/2 \leq t \leq 1; \\ 0 & \text{otherwise.} \end{cases} \tag{3.5}$$

The Haar wavelet family contains two wavelets: the approximation coefficients and the detailed coefficients. The approximation coefficients relate to the scaling function (Chan and Chen, 2017). It works similar to applying a low-pass filter to the data. The detail coefficients relate to the wavelet function which works similar to a high-pass filter which emphasizes the high-frequency information. In Chapter 10, we identify footsteps and digging which typically have low frequencies; therefore, we choose to use approximation coefficients of the Haar wavelet when considering the inverted wavelet tree in Section 3.4.

3.3 Tree-structured wavelet transforms

In neural networks used for image recognition, necessity requires training images to be small. If the images in a training set are too big, then the computational cost increases drastically. The small nature of the training images poses a potential issue, given that most data sets are much larger than a training image. Events in large data are often larger than the training images as well. It thus becomes necessary to window the data set in some way where the windows are the size of the training images, and also scales the events in the data to the size of the training image. One answer to this windowing issue is through tree-structured wavelet transforms.

Tree-structured wavelet transforms are introduced in Chang and Kuo (1993) for application in texture analysis. Tree-structured wavelet transforms are tree-structured algorithms which implement a wavelet transform. The top level of the tree is the data, and the tree then branches into several nodes depending on the number of wavelets in the wavelet family. Each subsequent level applies the wavelet transform to each node and then produces new nodes on the next level of the tree. Figure 3.1 is an example of a tree-structured wavelet transform, or a wavelet tree.

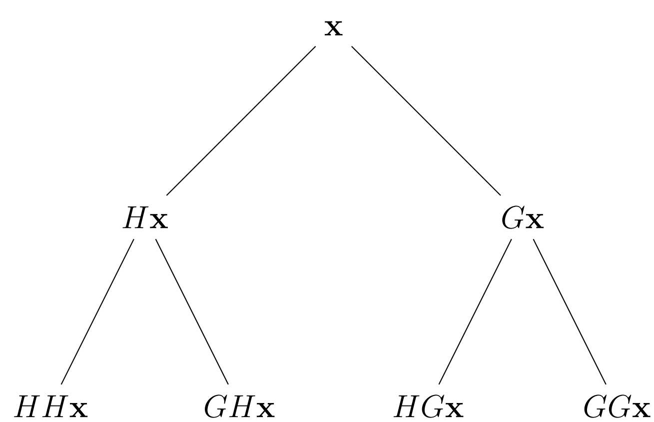


Figure 3.1: An example of a wavelet tree using the Haar transform. The data $\mathbf{x}$ resides at the top of the tree, the second row has a node $H\mathbf{x}$ for the approximation coefficients, and a node $G\mathbf{x}$ for the detailed coefficients. This process continues for subsequent levels.

The results at each node are the convolution of specific wavelets, chosen from within a family of wavelets, with the data. Each wavelet relates to a specific frequency band of the data (Bruns, 2004). In the case of the Haar transform, there is only the approximation coefficients and detail coefficients. Figure 3.2 presents an application of a tree-structured wavelet transform using the Haar transform to a chirp signal. We characterize a chirp signal by a sweeping frequency where the instantaneous frequency varies in time; see Mann and Haykin (1991) for more information. The data resides in the top row of the image, denoted as Level 0. Applying the first level of the tree-structured wavelet transform produces two

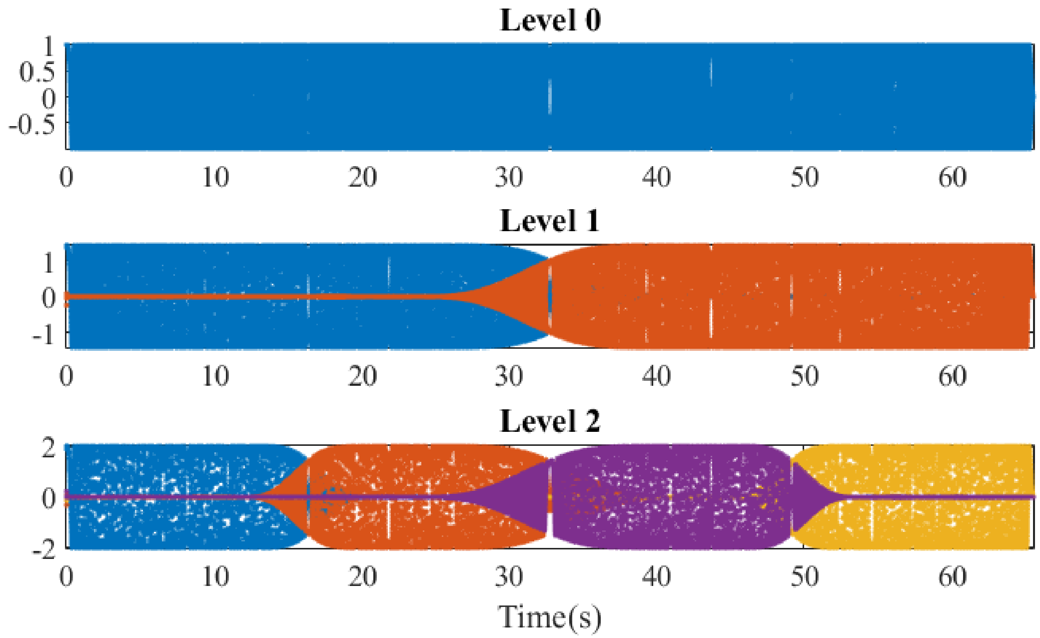


Figure 3.2: A tree-structured wavelet transform applied to a chirp signal using the Haar transform for three levels (McDonald, 2016a).

nodes as the Haar wavelet family contains two wavelets which we observe in Level 1, or the second row of Figure 3.2. The bottom row of Figure 3.2 depicts the last application of tree-structured wavelet transform for this example. Each of the nodes of Level 1 has both wavelets in the wavelet family convolved with it, thus producing four nodes in Level 2.

The size of the data is reduced by a power of two in the x-domain when the transform is applied. The results are two scaled processed versions of the original data set to which we assign to a node in the subsequent level of the wavelet tree. Scaling the data in this manner provides us with an answer to dealing with scale-invariance in image recognition techniques.

3.4 Inverted wavelet trees

Depending on the size of the data set, which generally can be quite large, the wavelet tree will need several levels in order to approach the size of a training image for the convolutional neural network. If a training image is a 128×128 matrix and the original data $\mathbf{X}$ has

$2^{14} = 16384$ bins in the x-domain, then the wavelet tree needs 7 levels in order to reduce the size of the data to 128 bins in the x-domain. The number of nodes N in a level is determined by the equation

$$N = Q^{m-1} \tag{3.6}$$

where Q is the number of wavelets in the wavelet family and m is the level of the tree the node resides. The seventh level of a wavelet tree using the Haar transform will have $2^6 = 64$ nodes. There are 6 other levels with 2^{m-1} nodes to take into account when considering the computational costs of this method. This approach can become computationally expensive very quickly.

An inverted tree-structured wavelet transform has the same basic principle as the tree-structured wavelet transform. The top level contains the original data and successive levels contain the results of applying a specific wavelet from the wavelet family to the data in the node. The main difference is that our top level has 2^{M-1} nodes, where M is the maximum number of levels in the tree. Each of the nodes in the top level holds a $N \times D$ strip of data, where N is some power of two that divides the x-domain of the data at least twice and D is the size of the data in the t-domain. The next level of the data is produced by pairing the first node with the second and then pairing each subsequent node to its successor if it is not already paired with the node that directly precedes it and applying the wavelet transform. The new node is the result of the wavelet transform applied to the concatenation of the data in the two nodes in the previous level. The next level of the tree contains 2^{M-m} nodes, where m is the level of the tree the node resides. If we have an inverted wavelet tree with a maximum of three levels, then the top level has $2^2 = 4$ nodes, the second level has $2^1 = 2$ nodes and the last level has $2^0 = 1$ nodes.

Mathematically, let $\mathbf{X}$ be a $P \times D$ matrix which represents the data set. The inverted tree-structured wavelet transform consists of creating a tree T with M levels and 2^{M-m} nodes in each level, where m is the current level of the tree T. For each node $a_{1,k}$ in level 1 with

$k \in \{1, ..., 2^{M-1}\}$, set

$$a_{1,k} := \{x_{i,j} : 1 + (k-1)L \leq i \leq kL, 1 \leq j \leq D\}, \tag{3.7}$$

where $L = 2^{\ell} < P/2$ for $\ell \in \mathbb{N}$. In levels $1 < m \leq M$, we define the nodes

$$a_{m,k} := W[a_{m-1,k}\ a_{m-1,k+1}], \tag{3.8}$$

where $a_{m-1,k}$ and $a_{m-1,k+1}$ are parent nodes from the previous level $m-1$. The term W indicates the wavelet used when applying the wavelet transform to the matrix $[a_{m-1,k}\ a_{m-1,k+1}]$. Figure 3.3 provides an example of a 3-level inverted wavelet tree.

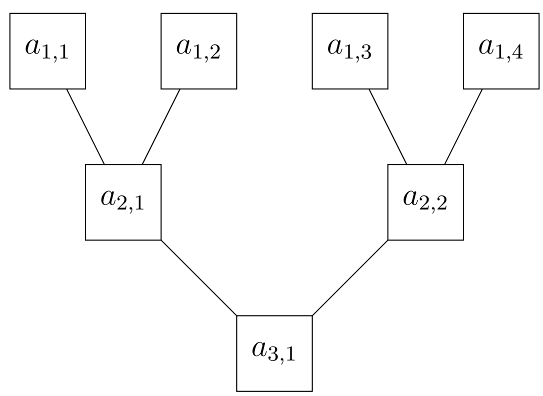


Figure 3.3: An example of a 3-level inverted wavelet tree.

Applying the inverted tree-structured wavelet transform to convolutional neural networks implies that L is equal to the dimension of the training images. Training images for convolutional neural networks are by necessity quite small in comparison to data sets. To limit the computational cost of applying inverted wavelet trees to the data in order to achieve scale invariance, we need to window the tree through the data. This idea entails redefining the nodes in the top level. The data in the first two nodes is thrown out and replaced with the data from their succeeding two nodes. After moving forward the final two nodes in the top level, the next portions of the original data set are then placed in the last two nodes.

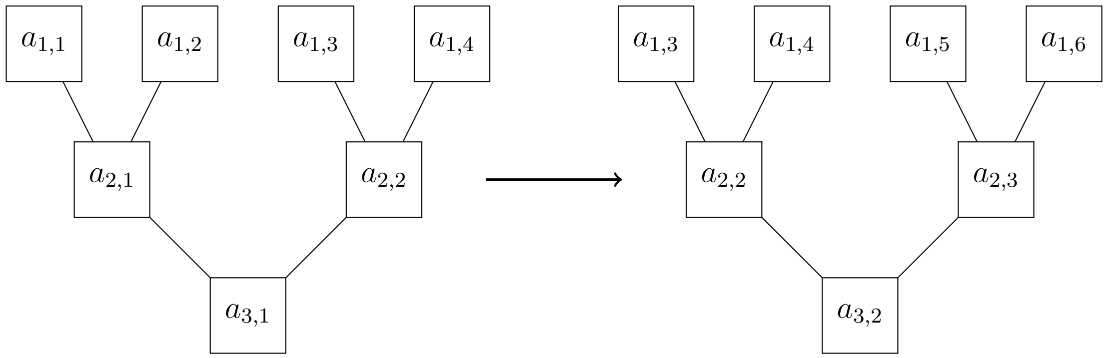


Figure 3.4: The left inverted wavelet tree shows the position of each node before windowing occurs. The right inverted wavelet tree shows the position of each node after windowing occurs. After windowing, the information in nodes $a_{1,1}$ and $a_{1,2}$ is thrown out along with their children nodes and the grandchild node. The information is then replaced with the data from nodes $a_{1,3}$ and $a_{1,4}$. The node $a_{2,2}$ follows its parents to replace node $a_{2,1}$. The nodes $a_{1,5}$ and $a_{1,6}$ represent the new data, which replaces the original information in these nodes. After applying a wavelet transform to the appropriate nodes, children nodes $a_{2,3}$ and $a_{3,2}$ are created.

The children nodes for each pair of nodes in the top level follow their parents over in their respective level. When we apply the wavelet transform, it need only be applied to the final two nodes of level 1 and the subsequent branch. Thus, a windowing of the data set is accomplished. Figure 3.4 depicts this windowing process for a 3-level inverted wavelet tree.

Mathematically speaking, suppose we have a tree T with n levels. Therefore, Level 1 of T has $N = 2^{n-1}$ nodes. The windowing for the tree T in the first level occurs by setting nodes $a_{1,k} = a_{1,k-2}$ and $a_{1,k+1} = a_{1,k-1}$, where n is the level and $k \in \mathbb{N}$. If $k - 2 \leq 0$, then the data from nodes $a_{1,k}$ and $a_{1,k+1}$ is thrown away. We define the final nodes in the level as

$$a_{1,N-1} = \{x_{i,j} : 1 + NL \leq i \leq (N+1)L, 1 \leq j \leq D\}, \tag{3.9}$$

and

$$a_{1,N} = \{x_{i,j} : 1 + (N+1)L \leq i \leq (N+2)L, 1 \leq j \leq D\}. \tag{3.10}$$

The children of nodes $a_{1,k}$ for $2 < k \leq N$ follow their parents in their respective level. The only computation necessary for updating the tree is the branch from the newly defined nodes $a_{1,N-1}$ and $a_{1,N}$. Applying the wavelet transform through this new branch produces the new inverted wavelet tree where the nodes take the form

$$a_{2,N^{1/2}} = W[a_{1,N-1}, a_{1,N}], \tag{3.11}$$

and

$$a_{3,N^{1/4}} = W[a_{2,N^{1/2}-1}, a_{2,N^{1/2}}]. \tag{3.12}$$

The windowing of the data set through the inverted tree-structured wavelet transform limits the size of the wavelet trees and thus reduces the computational cost it requires to process the data and also allows for translation-invariance. Nodes in subsequent levels are the results of applying any wavelet from a wavelet family to the node as long as the same wavelet is applied throughout the tree. We chose to use the approximation coefficients from the Haar transform in the experiments in Chapter 10.

3.5 Conclusions

In this chapter, we introduced wavelets. We began with a study of the tree-structured wavelet transform and saw how it motivates the inverted wavelet tree. Next, we developed the inverted wavelet tree and considered its application to convolutional neural networks for image recognition purposes. We investigate these ideas further in Chapter 10 where the inverted wavelet tree allows us to recognize events in large data sets, where these events are often larger than the training images we use to teach the convolution neural network.

Chapter 4

Using corrected phase to localize geological features in seismic data

Time-frequency analysis, or spectral decomposition, characterizes seismic signals with respect to frequency and time. From this information, we can derive seismic attributes for the purposes of localizing geological features in seismic data. In particular, we are interested in deducing rock properties about the seismic data using spectral decomposition methods. Popular time-frequency analysis methods include the Continuous Wavelet Transform (CWT) and the Short-Time Fourier Transform (STFT), while other spectral decomposition methods include the Synchrosqueezing Transform (SST) and Basis Pursuit (BP).

In Han et al. (2015), the authors apply all four of these methods to seismic data and compare the results experimentally. They found that basis pursuit provided the most promising results. Specifically, they discussed the amplitude and phase attributes. When examining these attributes, interpretation is necessary. We see later that sometimes an anomaly is difficult to decipher in data unless we know where it is ahead of time. To answer this, the authors of Han et al. (2015) proposed a derivative of corrected phase attribute which should outlines anomalies in the phase attribute of seismic data.

In this chapter, we apply the methods in Han et al. (2015) to new data and apply the derivative of corrected phase attribute to more spectral decomposition methods in order to extend the work of the authors. We focus on the time-frequency analysis methods, basis pursuit, continuous wavelet transform, and the synchrosqueezing transform. As the writers found in Han et al. (2015) when comparing the time-frequency methods on one data set, basis pursuit produced the best results for locating the geological structure in the new data set with regards to the phase attribute and the derivative of corrected phase attribute. The continuous wavelet transform performed adequately for highlighting the anomaly in both data sets at lower frequencies than those in basis pursuit. The synchrosqueezing transform proved impossible to locate the geological structure in both data sets when using the derivative of corrected phase attribute. In fact, the data appeared to become noisier after using the SST.

In the next section, we describe the time-frequency analysis methods: basis pursuit, continuous wavelet transform, and the synchrosqueezing transform. In the following section, we provide the data sets on which we test these time frequency analysis methods and the corrected phase method. Next, we examine the data sets and focus on the phase attribute and the corrected phase attribute. In the final section, we analyze the results of the derivative of the corrected phase method on other time-frequency analysis methods.

4.1 Spectral decomposition methods

We provide a summary of each method, based on the presentation and exposition in Han et al. (2015).

4.1.1 Basis Pursuit

The spectral decomposition method basis pursuit decomposes the signal from seismic data into individual atoms of a predefined dictionary. In time-frequency analysis, atoms are elementary waveforms which are discrete and populate a given dictionary (Tary et al., 2014). Specifically, the signal can be represented in series or matrix notation.

For a series representation, the signal is a convolution of the predefined wavelet family $\Psi(t,n)$ and the coefficient series $a(t,n)$ of these wavelets:

$$s(t) = \sum_{n=1}^{N} \Psi(t,n) * a(t,n) \tag{4.1}$$

where N represents the number of atoms, t is time, and n is an index to the dilation of the atom $\Psi(t,n)$ which determines its frequency. For the matrix representation of the signal, we have

$$\mathbf{s} = D\mathbf{a} + \eta = [\mathbf{\Psi_1}\ \mathbf{\Psi_2}\ \cdots\ \mathbf{\Psi_N}] \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_N \end{bmatrix} \tag{4.2}$$

where D is the wavelet dictionary, η is the noise, and $\mathbf{\Psi_n}$ is the convolution matrix of $\psi(t,n)$ with the dilation index n. Basis pursuit relates the time-frequency distribution to the set of weights $\mathbf{a}$ associated with the set of atoms $\psi(t,n)$ in the dictionary D.

Basis pursuit involves the following two steps:

1. A minimization term used to reduce the number of retrieved atoms as well as their magnitude, and

2. Simultaneously identifying all the atoms by applying a single inversion problem.

The method used in this chapter involves basis pursuit denoising. The object is to minimize the cost function:

$$J = \frac{1}{2}\|\mathbf{s} - D\mathbf{a}\|_2^2 + \lambda\|\mathbf{a}\|_1 \tag{4.3}$$

where the first term is the least-squares difference between the observed data $\mathbf{s}$ and the predicted data $D\mathbf{a}$, and the second term is the regularization term where λ controls the relative strength between the data misfit and the number of non-zero coefficients of $\mathbf{a}$.

Many optimization methods exists which can be used to solve this optimization problem.

Whatever method is employed, the optimization problem is guaranteed to converge eventually to a local minimum. The success of basis pursuit depends heavily on the predefined dictionary; however, while a larger wavelet dictionary provides better results, it also causes a longer computation time of the algorithm.

4.1.2 Continuous Wavelet Transform

Like basis pursuit, the continuous wavelet transform uses wavelets as the name suggests. While BP decomposes the signal from elements of a predefined wavelet dictionary, the CWT is the inner product of a wavelet family with the signal $f(t)$. We write this as

$$\begin{aligned} F_W(\sigma, \tau) &= \langle f(t), \psi_{\sigma,\tau}(t) \rangle && (4.4) \\ &= \int_{-\infty}^{\infty} f(t) \frac{1}{\sqrt{\sigma}} \bar{\psi} \left(\frac{t-\tau}{\sigma} \right) \mathrm{dt}, && (4.5) \end{aligned}$$

where $\psi_{\sigma,\tau}$ are the wavelets, $\bar{\psi}$ represents the complex conjugate of ψ, and $F_W(\sigma, \tau)$ is the time-scale map. The scale is a frequency band and extra work is necessary to build a time-frequency map. In order to achieve this mapping, several routes can be taken. In Sinha et al. (2005), the authors used the Morlet wavelet to compute a frequency spectrum of the signal $f(t)$. in particular, they described the time frequency map of the CWT as follows

$$\hat{f}(\omega, \tau) = \langle F_W(\sigma, \tau), \hat{\psi}_\omega(\sigma) \rangle, \tag{4.6}$$

where

$$\overline{\hat{\psi}_\omega}(\sigma) = \frac{\hat{\psi}(\sigma\omega) e^{-i\omega t}}{C_\psi \sigma^{3/2}} \tag{4.7}$$

is the scaled and modulated window and $\overline{\hat{\psi}_\omega}(\sigma)$ is the complex conjugate of $\hat{\psi}_\omega(\sigma)$.

The continuous wavelet transform differs from the short-time Fourier transform (STFT) in that its window has a variable length whereas the window length used in the STFT is fixed. The wavelets' ability to dilate in the CWT allows different frequencies to have different time support. In comparison to BP, the CWT does not depend on the convergence of a minimization problem. As a time-frequency mapping, the CWT improves the time-frequency resolution of a nonstationary signal which is advantageous when considering seismic data.

4.1.3 Synchrosqueezing Transform

Another wavelet-based spectral decomposition method of interest in this chapter is the synchrosqueezing transform (SST). Interestingly, it can be derived from the continuous wavelet transform (Herrera et al., 2014). Some issues arise when using the CWT with regards to smearing of the resolution. This occurs because the time resolution and the frequency resolution are inversely proportional. As the resolution for one increases, the other resolution decreases. The SST is meant to resolve this issue.

Recall that the CWT maps first to the time-scale domain and then the time-frequency domain. It is during this first mapping that the smearing often occurs. The synchrosqueezing transform attempts to improve the resolution of signal by mapping the time-frequency coefficients into instantaneous frequencies. The authors of Daubechies and Maes (1996) found that neglecting the smearing along the time axis enables the instantaneous frequency $\omega_s(\sigma, \tau)$ to be computed as the derivative of the CWT at any point (σ, τ) with respect to τ for all $F_W(\sigma, \tau) \neq 0$:

$$\omega_s(\sigma, \tau) = \frac{-j}{2\pi F_W(\sigma, \tau)} \left(\frac{\partial F_W(\sigma, \tau)}{\partial b} \right). \tag{4.8}$$

Mapping every point (b, a) to $(b, \omega_s(a, b))$ moves the points from the time-scale domain to the time-frequency domain. This conversion of points is called synchrosqueezing, and it is from this that the transform gets its name.

4.2 Data

We apply the methods in this chapter to the following two seismic data sets. Both data sets are post-stacked.

In Figure 4.1, the geological structure circled is a valley. This particular data set is from the CREWES Blackfoot data (CREWES, 2008); however, we refer it as the valley data set based on the geological feature we are interested in locating.

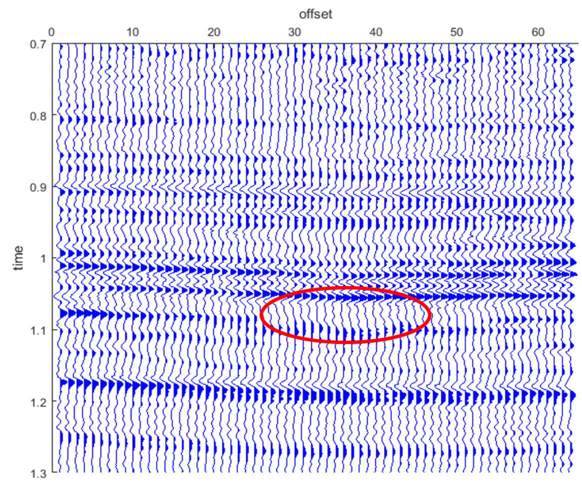


Figure 4.1: The valley lies between offsets $30 - -40$ and time $1.05s$. Time is measured in seconds.

In Figure 4.2, the goal is to identify the hydrocarbon reservoir circled in red. We call this data set the reservoir data set.

4.3 Phase attribute results from basis pursuit

We begin by applying basis pursuit to the geological data sets mentioned in the previous section and consulting the phase attribute. First, we consider the valley data set.

Figure 4.3 shows the phase attribute of the valley data set at specific frequency slices. Recall that the valley is located between offsets 30 and 40 and time $1.05s$. Referring to Figure 4.3 (upper left), it is evident that at lower frequencies basis pursuit does not provide a clear image of the valley. At about 26 Hz, distinguishing the valley becomes much easier as seen in Figure 4.3 (upper right). As the frequency increases to 32 Hz, the valley is still relatively clear to locate; however, as the frequency continues to increase, it becomes more

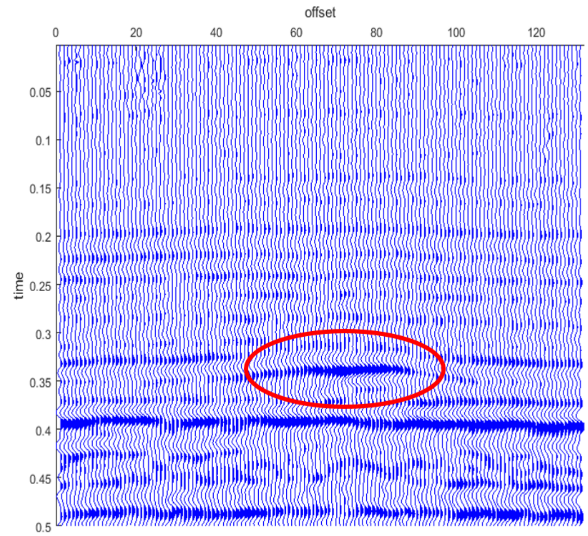


Figure 4.2: The hydrocarbon reservoir lies at around offsets $60 - -85$ and time $0.34s$. Time is measured in seconds.

difficult to localize the valley once again. Given the ability to locate the valley over a range of frequencies, we see that basis pursuit performs well on this data set.

Now, consider how well basis pursuit performs on the reservoir data set with regards to the phase attribute.

In contrast to the valley data set, basis pursuit provides better localization of the hydrocarbon reservoir at higher frequencies. These frequencies range from approximately 40 Hz to 60 Hz, as seen in the bottom row of Figure 4.4. At the lower frequencies, locating the reservoir proves more difficult as seen in the top row of Figure 4.4. Despite being able to find the reservoir at approximately 40 Hz and 60 Hz, it is not as visible as the BP results for the valley data set.

4.4 Derivative of the corrected phase method

Some of the difficulty in localizing the valley and hydrocarbon reservoir using the phase attribute from basis pursuit arises from the arbitrary coherent lines in the image which do not provide any information. These arbitrary lines result from the fact that the phase curve

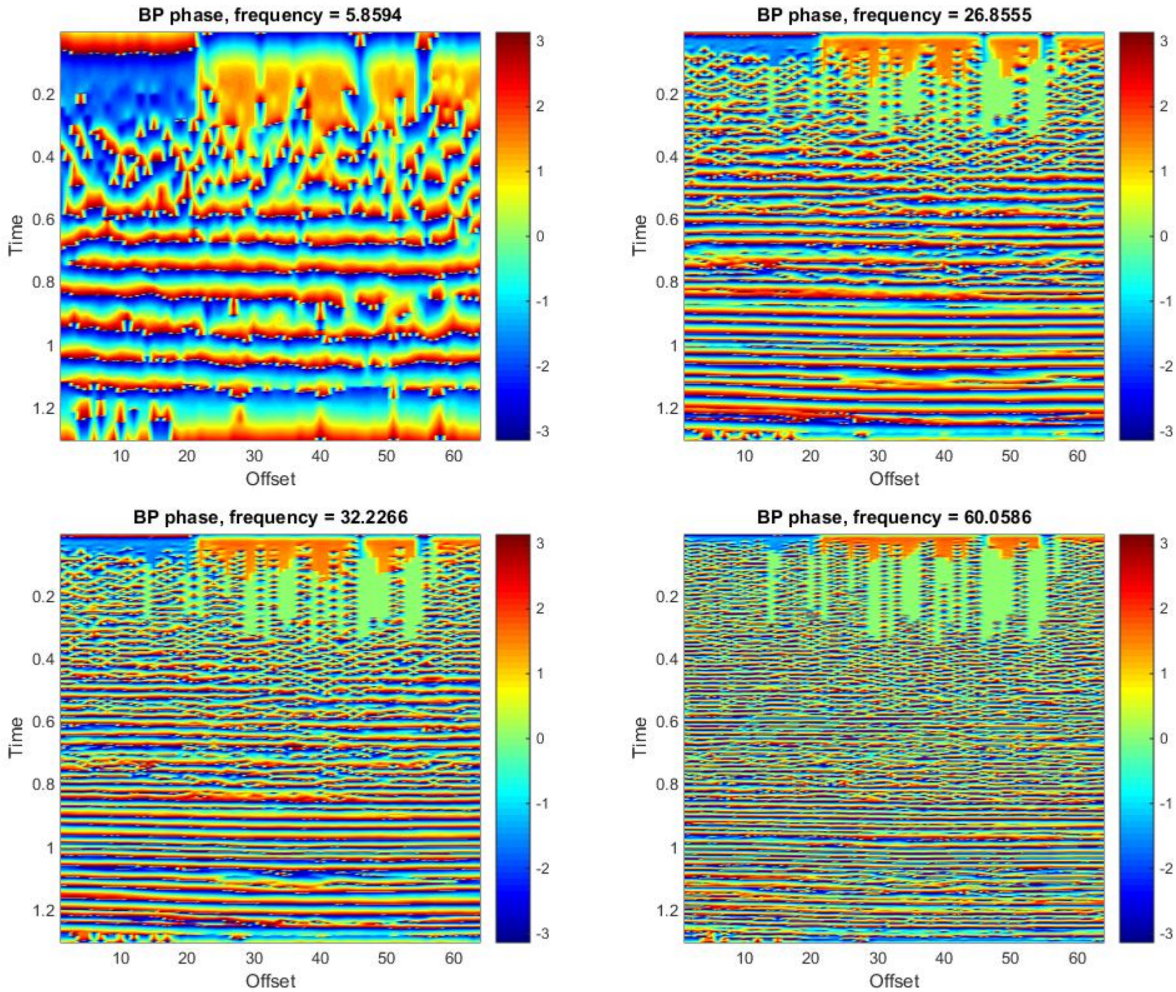


Figure 4.3: Phase attribute for valley post-stack data: Constant frequency slices obtained by applying BP to the valley data set at frequencies approximately 5 Hz, 26 Hz, 32 Hz, and 60 Hz, respectively. Time is measured in seconds.

of each offset is increasing. Figure 4.5 provides an example of a phase curve of an offset from the reservoir post-stack data set.

We can see that the curve increases quadratically and thus crosses a range of angles. Specifically, it goes through the cycle of 2π multiple times causing the ridges we see in the phase attribute since the colour bar is for angles between $-\pi$ and π. In Han et al. (2015), an attempt was made to remove these lines in a process called the corrected phase method.

This method involves unwrapping the time-dependent curve in the phase attribute and fitting it to a quadratic equation. Then, the quadratic is subtracted from the original curve.

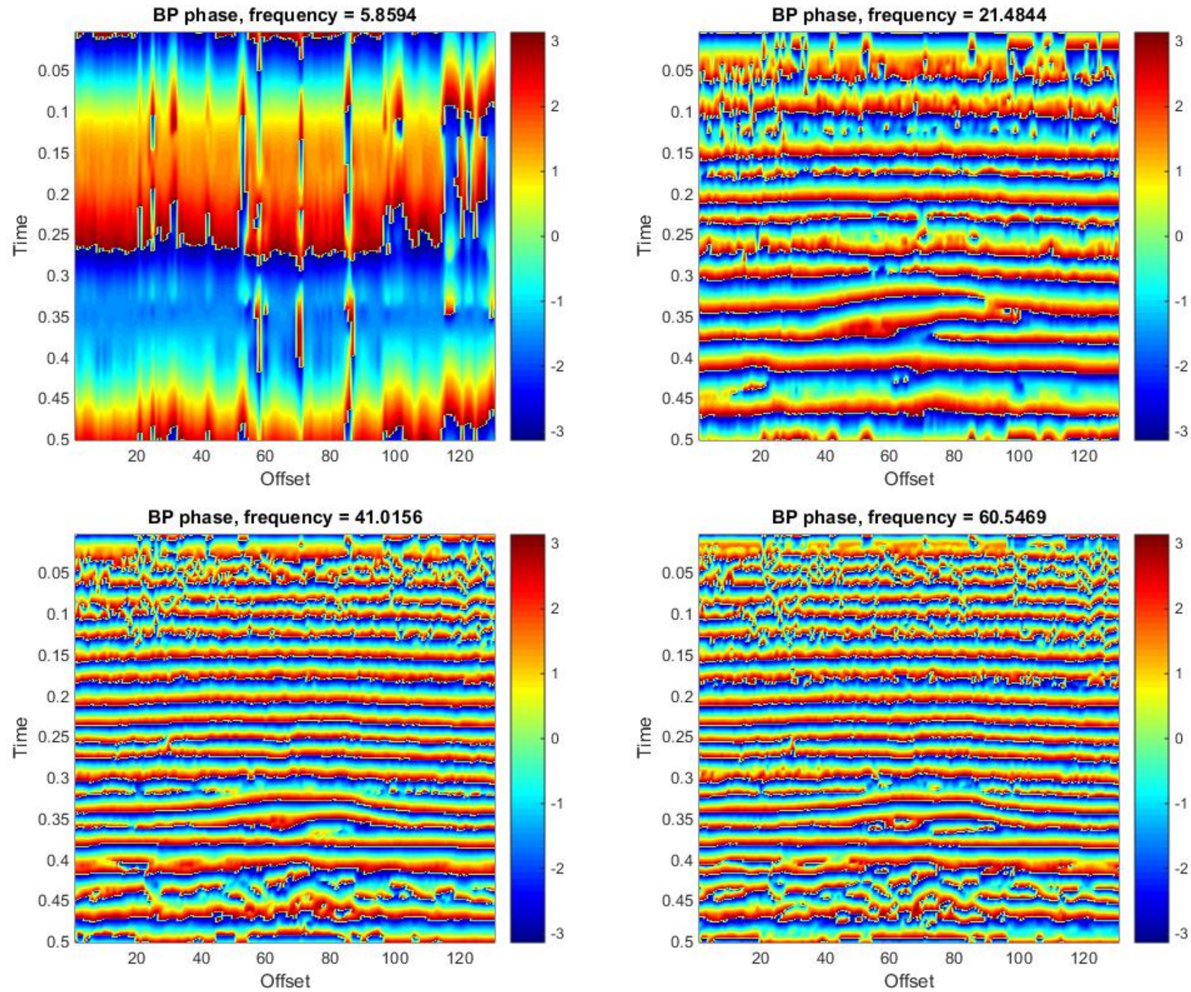


Figure 4.4: Phase attribute for reservoir post-stack data: Constant frequency slices obtained by applying BP to the reservoir data set at frequencies approximately 5 Hz, 20 Hz, 40 Hz, and 60 Hz, respectively. Time is measured in seconds.

The resulting curves are plotted in Figure 4.6.

In Figure 4.6, the top row shows the phase attribute on left and the corrected phase on the right for the valley data sets at approximately 26 Hz. The bottom row shows the phase attribute on the left and the corrected phase on the right for the reservoir data set at approximately 40 Hz. While this method removes the coherent lines from the phase attribute, it provides little extra information. In fact, this method provides less information than the original phase attribute. As such, another approach is necessary.

The authors in Han et al. (2015) take the process a step further by taking the derivative of

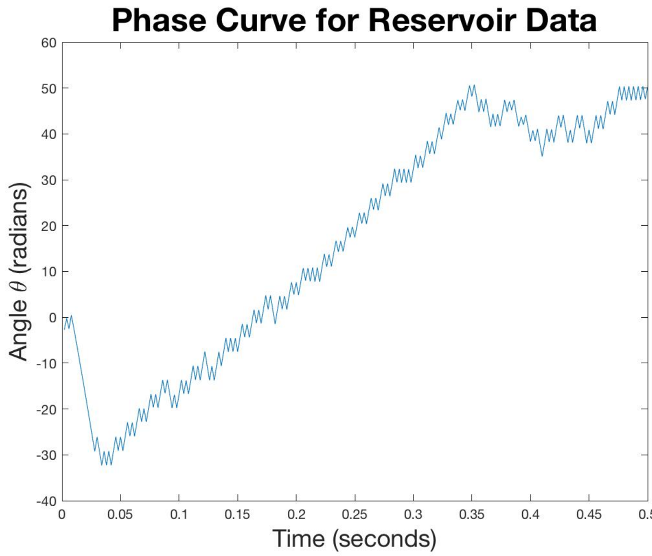


Figure 4.5: Phase curve of an offset from the reservoir data set.

the residuals and plotting these results. The derivative of the residual phase curves highlights the change in phase, which implies the boundary of the geological structure will be evident. These results are being produced from the information provided by running basis pursuit on the seismic data sets.

A study of Figure 4.7 shows the exact result achieved with regards to the valley data set. At the location of the valley, the boundary of the valley is evident at approximately 26 Hz and 32 Hz. This outline of the boundary enhances the ability to localize the valley.

In Figure 4.8, the results for the reservoir data set are promising; however, it is not as easy to locate the boundary of the reservoir as it was to localize the valley in Figure 4.7. We examine the derivative of the corrected phase attribute at approximately 40 Hz and 60 Hz since the reservoir was the most evident at the frequencies when considering the phase attribute. Notice in left column of Figure 4.8 that we can locate the reservoir; however, the shadowing at its location is not as dark as it was for the valley data set. This lack of shadowing potentially explains the absence of a clear boundary in the derivative of the corrected phase images in the right column of Figure 4.8 for the reservoir set. Despite this fact, the derivative of the corrected phase plots still produce a better localization of the

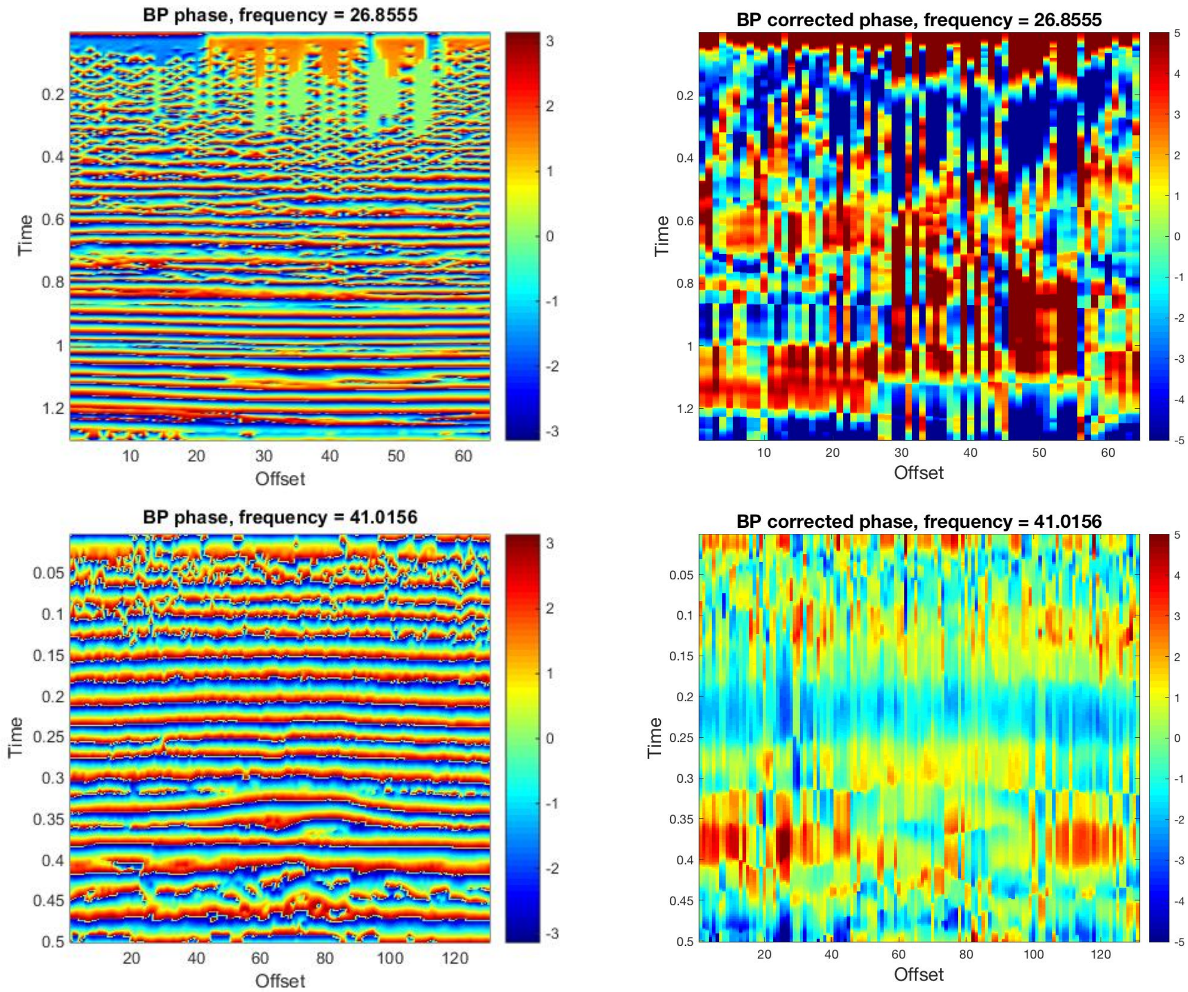


Figure 4.6: Comparison of (Top row) the phase attribute for the valley on the left to the corrected phase method on the right, and (Bottom row) the phase attribute for the reservoir on the left to the corrected phase method on the right. Time is measured in seconds.

hydrocarbon reservoir than the phase attribute. As seen in the right column of Figure 4.8, there is a distinct line located in the position of the reservoir outlining the bottom of it at both frequency slices: 40 Hz and 60 Hz.

The derivative of the corrected phase attribute holds promising results for localizing geological features in seismic data. In both seismic data sets, the geological structure was evident when this method was applied to it. While the derivative of the corrected phase method provided a better localization of the valley than the reservoir, it provided better results than the phase attribute in both cases.

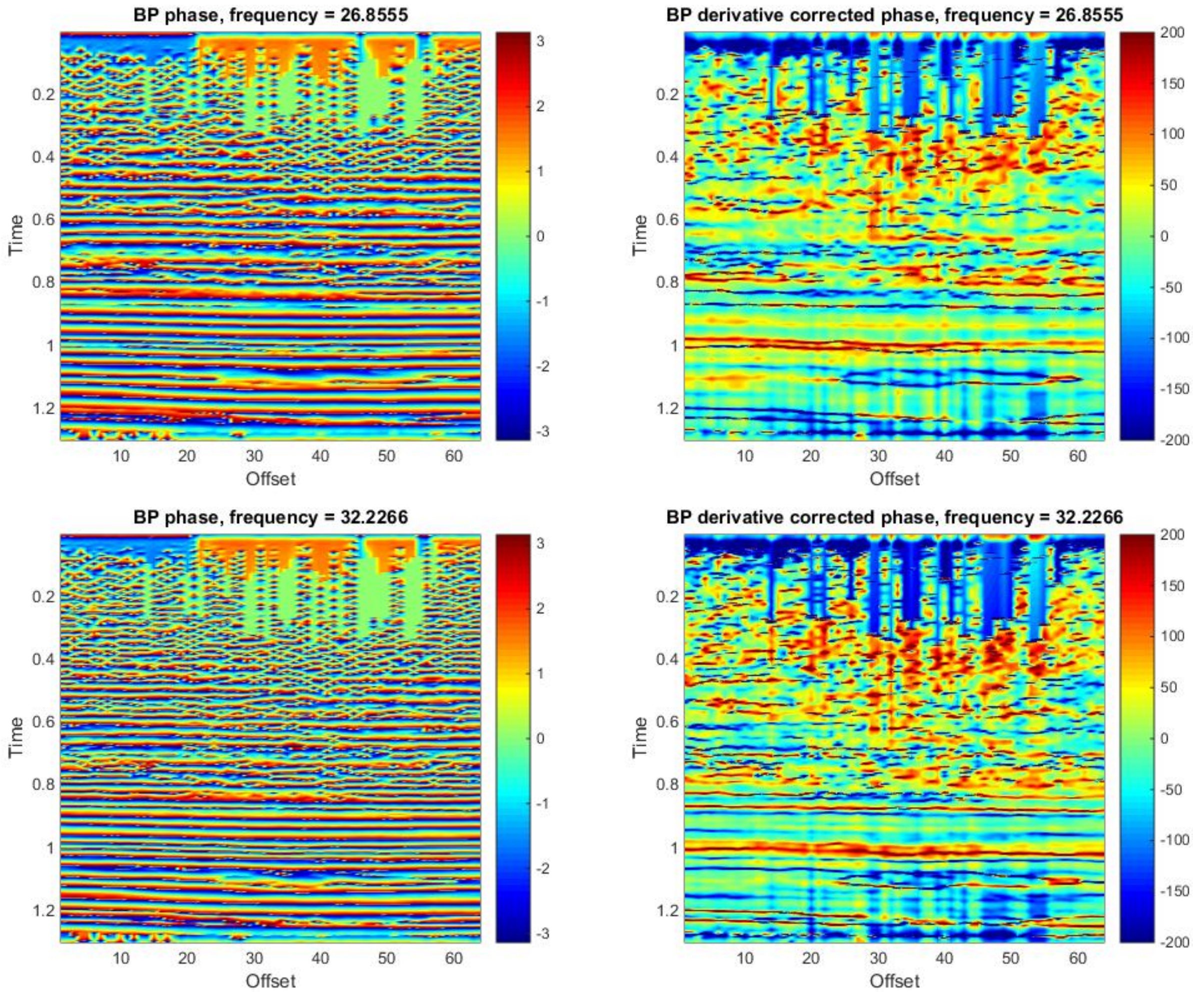


Figure 4.7: (Left column) Phase attribute of the valley data set at about 26 Hz and 32 Hz. (Right column): Derivative of the corrected phase method of the valley data set at about 26 Hz and 32 Hz. Time is measured in seconds.

4.5 Application of other time-frequency analysis methods

In this section, we consider the phase attribute produced by applying other time-frequency analysis methods to the valley and reservoir data sets. We also apply the derivative of the corrected phase method to these attributes and compare the results to those of basis pursuit.

We discuss the results of two spectral decomposition methods: Continuous Wavelet Transform (CWT) and Synchrosqueezing Transform (SST). For both methods, we choose

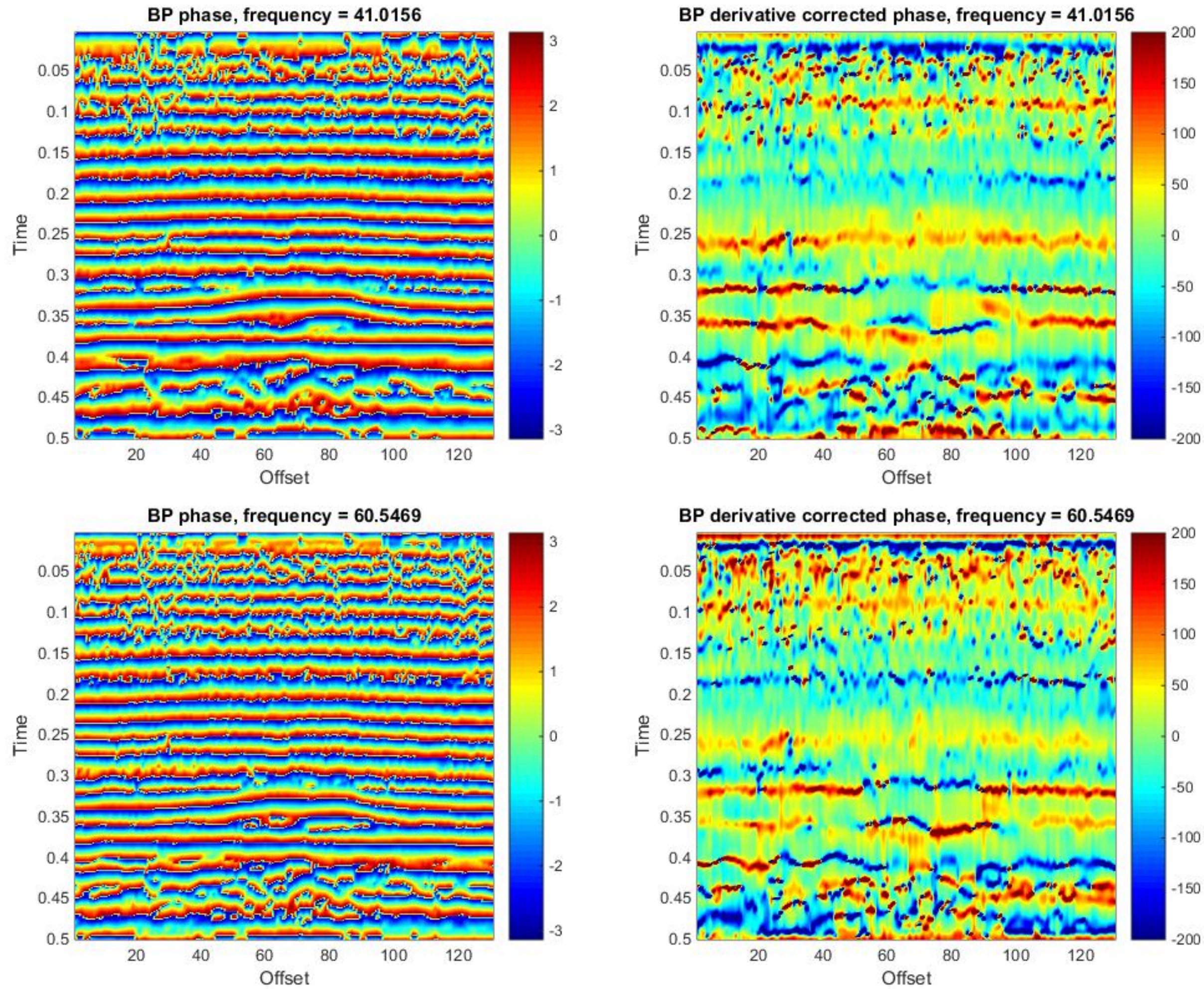


Figure 4.8: (Left column) Phase attribute of the reservoir data set at about 40 Hz and 60 Hz. (Right column) Derivative of the corrected phase method of the reservoir data set at about 40 Hz and 60 Hz. Tjme is measured in seconds.

the Morlet wavelet to be the mother wavelet. We start by examining the continuous wavelet transform.

In Han et al. (2015), the authors experimentally found that basis pursuit worked the best of the time-frequency methods they considered; however, CWT also showed promise. In Figure 4.9, the image on the left is the phase attribute produced by running CWT on the valley data set. The valley in this case can be localized with relative ease between 30 and 55 offset and at 1.01 secs. In fact, the image looks very similar to the results of BP; however, the frequency slice is much lower at about 6 Hz than that of basis pursuit. The image on the

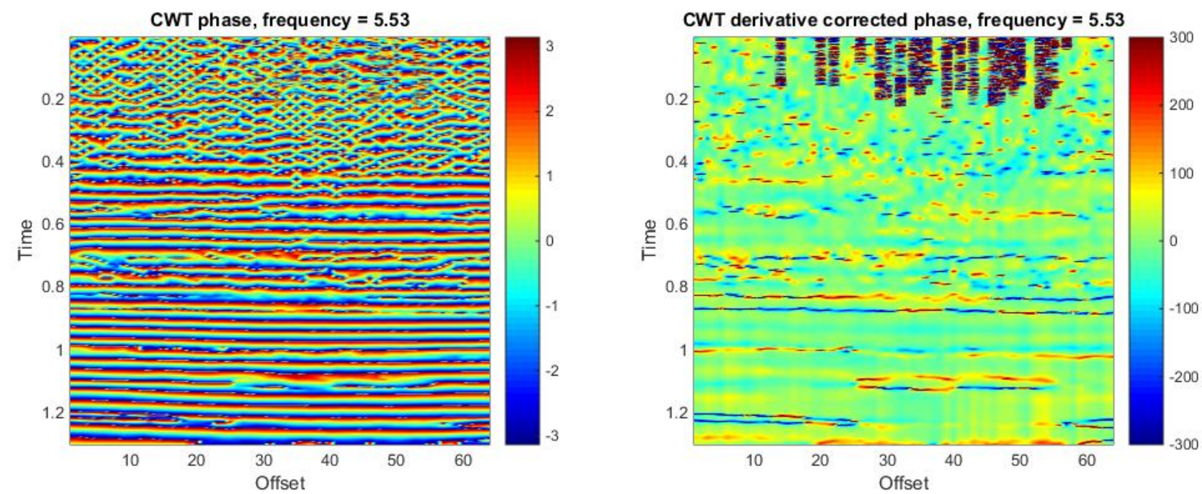


Figure 4.9: (Left) Phase attribute of the valley data set at about 6 Hz using CWT. (Right) Derivative of the corrected phase method of the valley data set at about 6 Hz using CWT. Time is measured in seconds.

right in Figure 4.9 are the results of the derivative of the corrected phase method applied to the phase attribute pictured on the left. As with BP, the boundary of the valley can be identified.

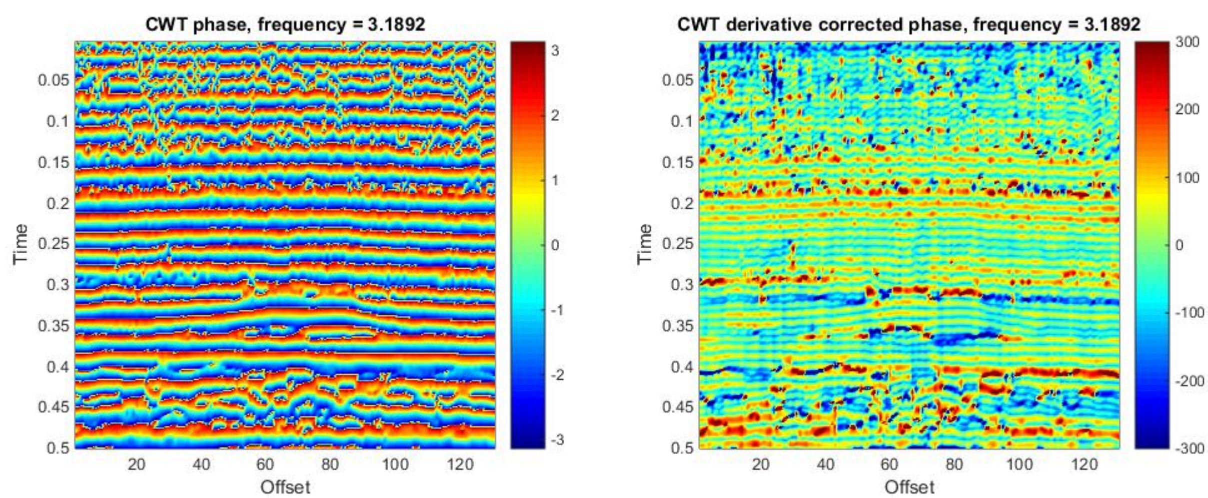


Figure 4.10: (Left) Phase attribute of the valley data set at about 3 Hz using CWT. (Right) Derivative of the corrected phase method of the valley data set at about 3 Hz using CWT. Time is measured in seconds.

Figure 4.10 shows similar results for the reservoir data set. While the reservoir can be localized in the phase attribute pictured on the left, it is not as easily located as the valley.

This is similar to the results of basis pursuit. Considering the derivative of the corrected phase method, there is an outline of where the reservoir, similar to the BP case; however, the outline is not as clear as it was when applying BP to the data set. While basis pursuit with the derivative of corrected phase method performed better on the reservoir data set, CWT is still a viable option as the derivative of the corrected phase method still provides better results than simply considering the phase attribute.

In Han et al. (2015), the authors found that the synchrosqueezing transform was the least helpful in located geological features in seismic data. As such, SST provides a study of how well the derivative of the corrected phase method works in identifying geological structures when produced from phase attributes where localizing the feature is difficult.

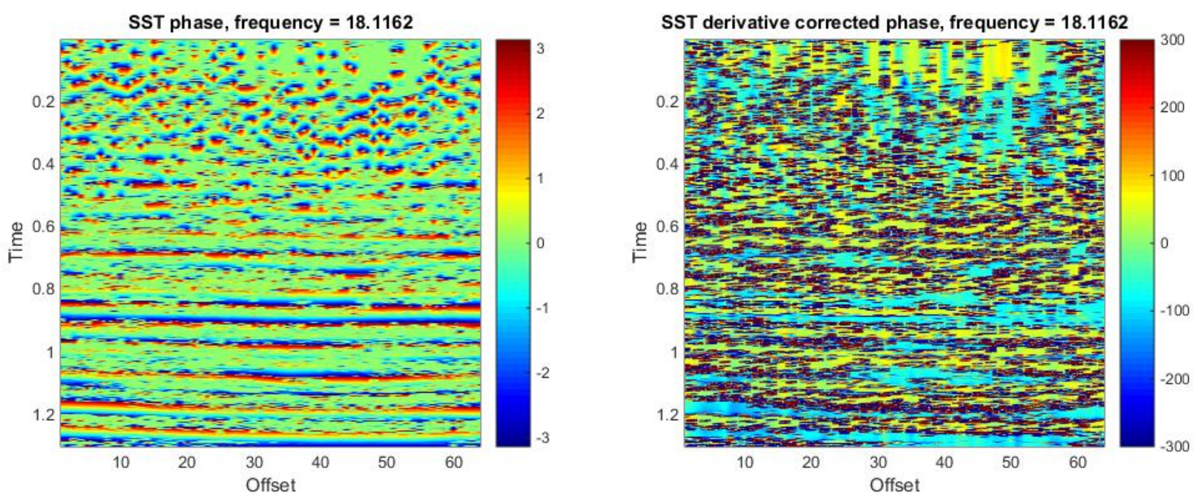


Figure 4.11: (Left) Phase attribute of the valley data set at about 18 Hz using SST. (Right) Derivative of the corrected phase method of the valley data set at about 18 Hz using SST. Time is measured in seconds.

Figure 4.11 shows the phase attribute and derivative of the corrected phase results for the valley data set. On the left, the valley can be localized to some extent; there is a shadow between 30 and 55 offset and at 1.02 seconds. The valley is not as clear as the results of BP and CWT. From the derivative of the corrected phase method pictured on the right in Fig. 4.11, the valley remains relatively difficult to locate. In fact, the corrected phase method does not remove the coherent lines as adequately for the phase attribute produced from SST as it does for BP and CWT. As a result, locating the reservoir in the derivative of the corrected phase proves difficult. This suggests that the success of the derivative of the corrected phase method depends to a degree on the success of localizing the geological feature in the phase attribute.

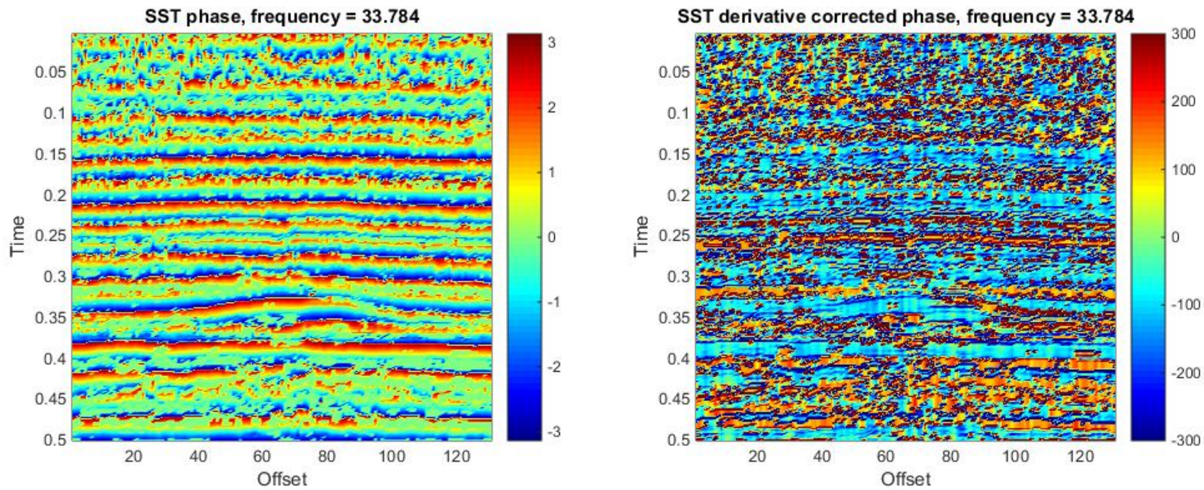


Figure 4.12: (Left) Phase attribute of the reservoir data set at about 33 Hz using SST. (Right) Derivative of the corrected phase method of the reservoir data set at about 33 Hz using SST. Time is measured in seconds.

As with BP and CWT, it proves more difficult for SST to localize the reservoir with regards to the phase attribute. Moreover, the corrected phase method does not remove the arbitrary lines from the phase attribute with respect to this data set as effectively as it did for BP and CWT. Consulting Figure 4.12, the image on the left shows the phase attribute produced from the results of SST at about 33 Hz. It is possible to locate the reservoir in this image; however, in the derivative of the corrected phase method pictured on the right, it is difficult to localize the reservoir. The method fails to highlight the change in boundary of the reservoir when using SST. As with the valley data set, there is a shadow present in the derivative of corrected phase image; however, identifying the hydrocarbon valley relies heavily on the reader's knowledge of its presence. The time-frequency analysis method SST with the derivative of the corrected phase method is not a promising combination for localizing geological features in seismic data.

4.6 Conclusions

We demonstrated the effectiveness of basis pursuit in localizing geological features when consulting the phase attribute. We also exhibited the extended capabilities of locating geological structures in seismic data once the phase is corrected and we considered the derivative of the corrected phase. We observed that to some degree the success of the derivative of the corrected phase method is dependent on how clearly the phase attribute localizes the geological feature. This result is evident with all time-frequency analysis methods we considered in the paper. As in Han et al. (2015), basis pursuit performed the best with respect to the phase attribute as well as the derivative of the corrected phase method. The continuous wavelet transform performed adequately for both data sets and both methods whereas the synchrosqueezing transform struggled to identify the valley and hydrocarbon reservoir.

Chapter 5

Distributed Acoustic Sensing

Distributed Acoustic Sensing (DAS) is a relatively new method for acquiring seismic data (Hartog, 2017); it uses fibre optic cables and laser interferometry for data acquisition. It has already shown promise for use in projects related to seismic acquisition, $C0_2$ monitoring, and smart cities (Cova et al., 2018). When comparing DAS to other seismic data acquisition methods such as geophones, fibre is inexpensive and provides a continuous receiver which can be oriented in many different configurations. Given its low cost, fibre optic cables can be installed permanently in seismic acquisition projects. Fibre-optic cables and DAS can be employed for more than seismic acquisition. It is often installed in construction projects in cities, and such installations provide the ability to monitor roads and city transportation using a DAS system.

For the remainder of this thesis, we focus on data acquired using distributed acoustic sensors. In this chapter, we explain how distributed acoustic sensing works and why it is a promising method for data acquisition and monitoring in seismic exploration. We also explore some open questions with regards to distributed acoustic sensing acquired data, some of which we examine in later chapters. For the explanation of how DAS works, we mainly follow Hardeman et al. (2017) and discussions with industry expert Matt McDonald from Fotech Solutions (McDonald, 2016b). More information about DAS and other similar acquisition methods can be found in Grattan and Meggitt (2000) and Hartog (2017).

5.1 Distributed acoustic sensing and fibre-optic cables

A distributed acoustic sensing system uses an interrogator and fibre-optic cable as a receiver to sense events. Inside the interrogator is a laser. When we connect the fibre-optic cable to the interrogator, a laser pulse is sent down the length of the cable. Strain on the fibre from nearby events, such as someone walking near the fibre-optic cable, induces reflections of the laser inside the cable, and these reflections cause the laser pulse to interfere with itself inside the fibre-optic cable. The interrogator retrieves these interferences to infer the location of the strain on the fibre. This process is called interferometry and it is the basis of how distributed acoustic sensing works. Interferometry uses the combination of waves to infer information about their state. We measure the intensity of this backscattered light as a function of time to determine the location of events at points along the fibre. Figure 5.1 gives a visual display of how interference occurs in the fibre-optic cable of a DAS sensor.

Let us consider a mathematical explanation of how interference occurs inside the fibre-cable. We assume that light inside the fibre-optic cable satisfies the wave equation

$$\frac{\partial^2 u}{\partial t^2} = c^2 \frac{\partial^2 u}{\partial x^2}, \tag{5.1}$$

where u is the amplitude of the light inside the fibre and c is its speed. For the mathematical analysis, it is convenient to express the general solution of the wave equation in complex form as

$$u(x,t) = e^{i(kx+\omega t)}, \tag{5.2}$$

for angular wavelength k and angular frequency ω. The parameter $k = \omega/c$ so we rewrite Equation 5.2 as

$$u(x,t) = e^{i\omega\left(\frac{x}{c}+t\right)}. \tag{5.3}$$

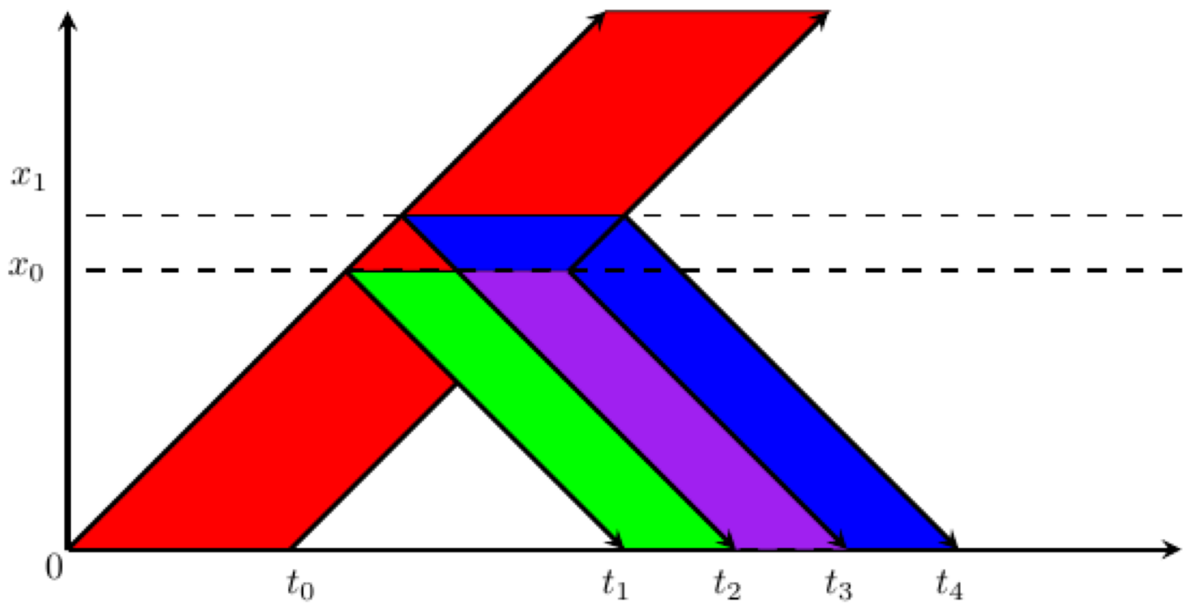


Figure 5.1: A visual representation of how distributed acoustic sensing works: A laser pulse is sent down the fibre and reflected back at points x_0 and x_1 along the length of the fibre. The red parallelogram represents the laser pulse. The green parallelogram represents the reflected portion of the laser at x_0 and the blue parallelogram represents the reflected portion of the laser pulse at x_1. The purple parallelogram represents the portion of the reflected laser pulse from x_1 which interferes with the portion of the reflected laser pulse from x_0. We use the purple parallelogram to extract information about events detected by the fibre (McDonald, 2016b).

Assume there are only two points x_1 and x_0 along the fibre which reflect the laser pulse back to the interrogator and let $\Delta x = x_1 - x_0$. To solve for Δx, we employ the intensity of the backscattered light I which is defined as

$$I = u_r(t)\overline{u_r}(t), \tag{5.4}$$

where u_r is the returning waveform from reflectors along the fibre and $\overline{u_r}$ is its complex conjugate. The phase delay at the first point x_0 is $2x_0/c$ seconds. Since we assumed $\Delta x = x_1 - x_0$, then we can write the phase delay for the second point x_1 in terms of x_0: $2x_1/c = 2(x_0 + \Delta x)/c$ seconds. Using Equation 5.3, we solve for the returning waveform u_r. The returning waveform of the first scatter point x_0, u_0, is

$$u_0(t) = r_0 e^{i\omega\left(2\frac{x_0}{c}+t\right)}, \tag{5.5}$$

where r_0 is the reflection coefficient at the point x_0 and the returning waveform of the second scatter point x_1, u_1, is

$$u_1(t) = r_1 e^{i\omega\left(2\frac{(x_0+\Delta x)}{c}+t\right)}, \tag{5.6}$$

where r_1 is the reflection coefficient at x_1. Summing the two returning waveforms from each scatter point gives us the entire returning waveform

$$\begin{aligned}
u_r(t) &= u_1(t) + u_2(t); && (5.7)\\
&= r_0 e^{i\omega\left(2\frac{x_0}{c}+t\right)} + r_1 e^{i\omega\left(2\frac{(x_0+\Delta x)}{c}+t\right)} && (5.8)\\
&= \left(r_0 + r_1 e^{i\omega 2\frac{\Delta x}{c}}\right) e^{i\omega\left(2\frac{x_0}{c}+t\right)} && (5.9)\\
&= \left(1 + \frac{r_1}{r_0} e^{i\omega 2\frac{\Delta x}{c}}\right) u_0(x,t). && (5.10)
\end{aligned}$$

The reflection coefficient $r_0 \neq 0$ since we assume a reflection occurs at x_0. The intensity of the returning waveform is then

$$\begin{aligned}
I &= u_r(t)\overline{u_r}(t) && (5.11)\\
&= \left(1 + \frac{r_1}{r_0} e^{i\omega 2\frac{\Delta x}{c}}\right) u_0(x,t) \left(1 + \frac{r_1}{r_0} e^{-i\omega 2\frac{\Delta x}{c}}\right) \overline{u_0}(x,t) && (5.12)\\
&= \left(r_0 + r_1 e^{i\omega 2\frac{\Delta x}{c}}\right)\left(r_0 + r_1 e^{-i\omega 2\frac{\Delta x}{c}}\right) && (5.13)\\
&= r_0^2 + r_1^2 + 2r_0 r_1 \left(e^{i\omega 2\frac{\Delta x}{c}} + e^{-i\omega 2\frac{\Delta x}{c}}\right) && (5.14)\\
&= r_0^2 + r_1^2 + 2r_0 r_1 \cos\left(2\omega\frac{\Delta x}{c}\right) && (5.15)\\
&= r_0^2 + r_1^2 + 2r_0 r_1 \cos\left(2k\Delta x\right), && (5.16)
\end{aligned}$$

where $k = \omega/c$. Equation 5.16 tells us that the measure of Δx depends on the stability of the wavelength k.

The method for which the parameter Δx is solved from here is proprietary to a given DAS company. Before we conclude this section, we discuss two aspects of distributed acoustic sensing that will appear in later chapters. One feature involves how long it takes for the laser pulse to reach the end of the fibre and return to the interrogator. The fastest rate at which this can be accomplished is known as the pulse repetition frequency, or the PRF. In Chapter 10, we consider data acquired with a DAS system at two different PRFs: 4kHz and 6kHz. Another aspect of a DAS system is the gauge length. The gauge length is a distance (length along the fiber in meters) specified by the manufacturer, which indicates the range over which point-wise strain measurements are averaged. It is limited by the physical design of the laser interferometer and DAS system. Typically, the gauge length is in the range of 1 to 25 meters. The gauge length is a property of the DAS system related to the pulse width of the laser interferometer. The gauge length is applied to the response of the fibre by integrating the gauge length along the length of the fibre. In other words, it acts as a moving integral filter in the spatial dimension of the fibre's response; cf. Hartog (2017). We study some visual effects of the gauge length on DAS data in Chapter 6. For more information on the effect of the gauge length, see Hartog (2017).

5.2 Applications

Distributed acoustic sensors began seeing commercial applications around 2005; cf. Hartog (2017). Such applications have since expanded to seismic acquisition, smart cities, security, and pipeline monitoring.

In the field of geophysics, distributed acoustic sensors have shown promise with regards to vertical seismic profiles; see Hardeman et al. (2017) and Hardeman et al. (2018). In vertical seismic profiling, sensors are placed along the inside of a well-borehole which can be several hundreds of meters in depth. These sensors detect reflections from anomalies near the well when a source is started within some proximity of the well-borehole. Geophones are discrete sensors which are slowly lowered into the well in a group. The length of this group is often much shorter than the well (Stewart, 2001). Acquisition of vertical sesimic profiles using geophones thus requires several sweeps with a source to acquire information about the entire well. When using a DAS system, fibre-optic cables can be installed into the well permanently, given their relative low cost, and fitted to the entire length of the well-borehole. The fibre provides a continuous receiver for the acquisition of vertical seismic profiles. The vertical seismic profiles are achieved quickly as the fibre does not need to be adjusted between sweeps of the source. While the amplitude response of a DAS system has not achieved the same strength as geophones, it is still possible to isolate events near the well; see Hardeman et al. (2017) and Hardeman et al. (2018).

In Chapter 8, we show the results of acquiring vertical seismic profiles at the Containment and Monitoring Institute's Field Research Station located in Newell County, AB. In both acquisitions discussed in the chapter, we locate the caprock which will store carbon-dioxide in future experiments at the site. The amount of time a DAS system saves in acquisition shows its promise, especially since work can be done to improve the response of DAS systems over time. For a comparison of a distributed acoustic sensor and geophones with regards to vertical seismic profiles, see Gordon and Lawton (2018).

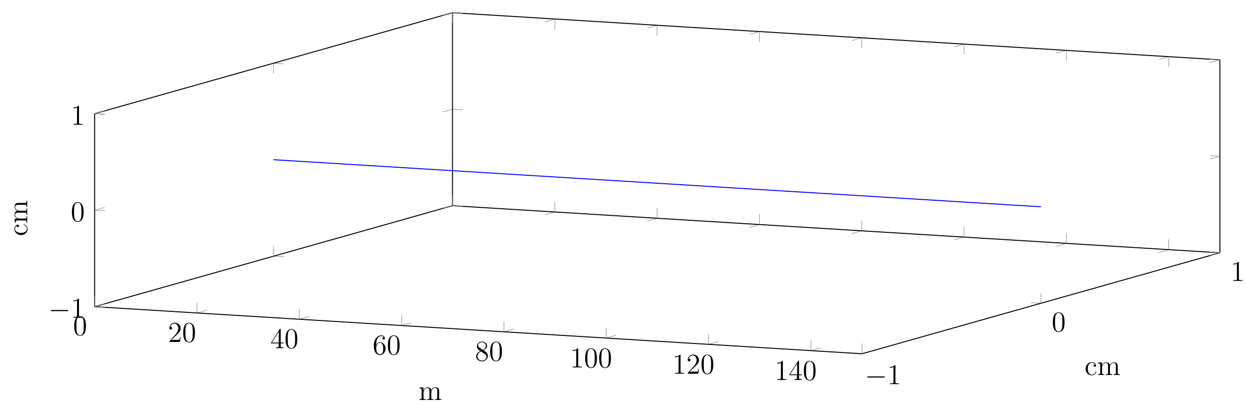
Further application of DAS systems finds itself in smart cities, and especially with regards to transportation monitoring. In Cova et al. (2018), the authors use data acquired using distributed acoustic sensing to determine the speed of vehicles detected by the sensor. In Chapter 9, we employ Gaussian mixture models to detect a vehicle driving along a fibre-optic cable in data acquired using DAS. Distributed acoustic sensors have also been utilized to monitor trains (Hartog, 2017).

A distributed acoustic sensor can also sense footsteps and digging. Chapter 10 presents an example of someone walking parallel to a fibre-optic cable. In the chapter, we train a convolutional neural network to detect these footsteps in the data. In another example, we display data acquired using a distributed acoustic sensor which shows someone walking and digging next to the fibre. Its effectiveness in detecting these different types of events proves especially useful in security and monitoring, and its ability to detect digging is particularly important when monitoring pipelines.

5.3 Open questions

One benefit of distributed acoustic sensors is the employment of fibre-optic cables. They are inexpensive, durable, and flexible. This flexibility allows them to be placed in different formations in order to better fit a monitoring project but also to increase the sensitivity of the sensor. When the fibre lies straight along a surface, it does not recover information about waves which hit the fibre perpendicularly; cf. Hartog (2017). The authors of Hornman (2016) and Hornman et al. (2013) solve this issue by winding the fibre helically; the direction of the helical wind also has an effect on what the sensor detects (Martin et al., 2014). Figure 5.2 shows an example of helically wound fibre and straight fibre. What events the fibre detects largely depends on the shape of the fibre. The helically wound fibre, by design, should be more sensitive than the straight fibre; however, in practice, we find that the response of the helically wound fibre is often not as strong as the straight fibre. In Chapter 7, we offer a means to compare straight and helically wound fibre using the model we create in Chapter 6 and using classical homotopy theory from algebraic topology.

While DAS systems detect events which occur near the fibre-optic cable, it is possible to observe where they happen along the fibre distance; however, there are not clear methods for discerning how far away these events lie from the fibre at a specific fibre distance. Knowing the exact location, or even an estimate of how far away an event is from the fibre, would be highly beneficial to monitoring projects. In Chapter 11, we produce a method for approximating the distance between the fibre and a detected event. The approach employs the hyperbolic shape of events in DAS-acquired data as well as the P-wave and S-wave responses



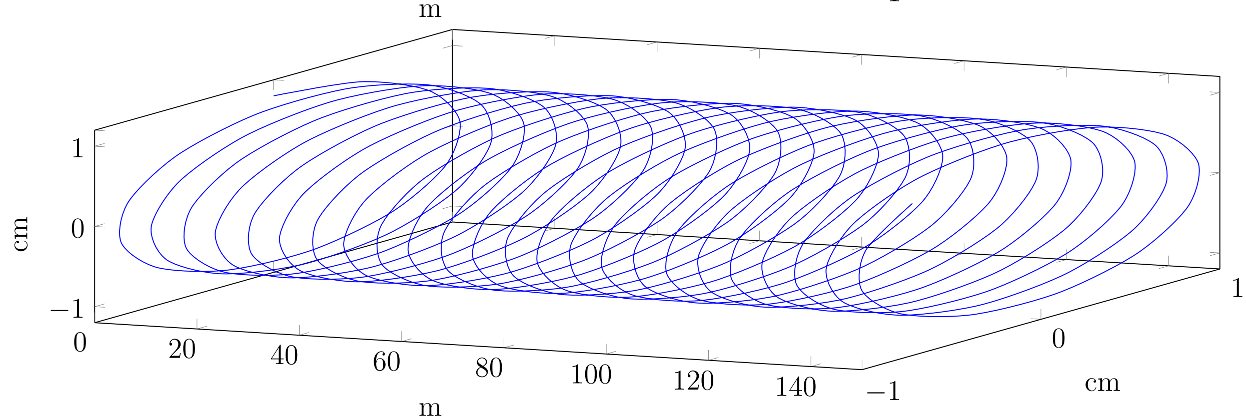


Figure 5.2: (Top) An example of a straight fibre optic cable in a DAS sensor. (Bottom) An example of a helically wound fibre-optic cable in a DAS sensor.

of the event to configure the distance between the source of the event and the fibre.

In the remaining chapters, we plan to supply insight some of these unanswered questions involving distributed acoustic sensors. We also provide results of various machine learning methods for detecting events in data acquired using DAS. In these investigations, we produce some conclusions for image registration of events in distributed acoustic sensing acquired data. We also consider means of detecting events over multiple windows of DAS data as well as emphasizing events in the data for better results. Before we address any of the issues, we begin in the next chapter by creating a means of modeling data from a distributed acoustic sensor.

Chapter 6

Modeling Distributed Acoustic Sensing Data

Given the myriad applications of distributed acoustic sensors we saw in Chapter 5, it is worthwhile to conduct a careful study of DAS acquired data. In many studies, the first step often deals with creating a model of the data. We model DAS data analytically, using the true seismic wavefield for constant velocity in three dimensions, with an arbitrary path for the placement of the DAS fibre in this chapter. The advantage of this approach is that we can rapidly simulate the true fibre response in three dimensions, in a variety of fibre trials, since we only need to compute data at points on the fibre. This is much faster than calculating a three dimensions finite difference simulation of a seismic wavefield or running a physical experiment with real fibre in the ground. It also provides a useful comparison to other simulation methods, such as the work of Eaid et al. (2017). The analytic method also allows an explicit separation of the P- and S-wave responses for the fibre.

In the following section, we describe the problem of acquiring seismic data using distributed acoustic sensing. We look at examples of the model used for the P-wave in two different media, and then move on to compare the results of the S-wave response of straight and helical fibre. Then, we show the full response of the straight and helically wound fibre given different gauge lengths.

For the P-wave example, we find that the slope of the hyperbolic response was much steeper for the saturated shale case than for the limestone case, since the P-wave moved slower through saturated shale versus limestone. In the S-wave comparison of helically wound fibre and straight fibre, we see that the straight fibre had a stronger response than the helical fibre for the S-wave; however, the helically wound fibre recovers more information for different source orientation vectors **A** unlike the straight fibre. When we consider how the gauge length affects the full waveform response of the fibre, we see that the hyperbolic response appears to flatten in both cases of the straight and helical fibre as the gauge length increased. Furthermore, some evidence of the helix is present in the helical fibre's full waveform response; however, it disappears as the gauge length is applied to the data.

6.1 Acquisition using distributed acoustic sensing and fibre-optic cables

In seismic acquisition, a fibre optic cable is installed in a specific configuration. In the examples in this chapter, we consider a fibre optic cable buried horizontally 10 meters beneath the Earth's surface over a distance of 200 m. Figure 6.1 is an example of full-waveform response of the fibre acquired from this formation. The figure shows the signal for the fibre sensor of a source detonated near the middle of the fibre.

To form the data set in Figure 6.1, a source is positioned in the centre of the fibre at the Earth's surface. Once the source detonates, it sends waves into the ground. When these waves hit the fibre, they stretch and strain the fibre optic cable. The strain on the fibre is recovered by an interrogator attached to the cable using a laser pulse that interacts with imperfections in the cable.

Figure 6.2 depicts the description of DAS acquistion. A source is represented by a red 'x.' The blue line represents the fibre 10m deep. The green half-sphere describes the wave

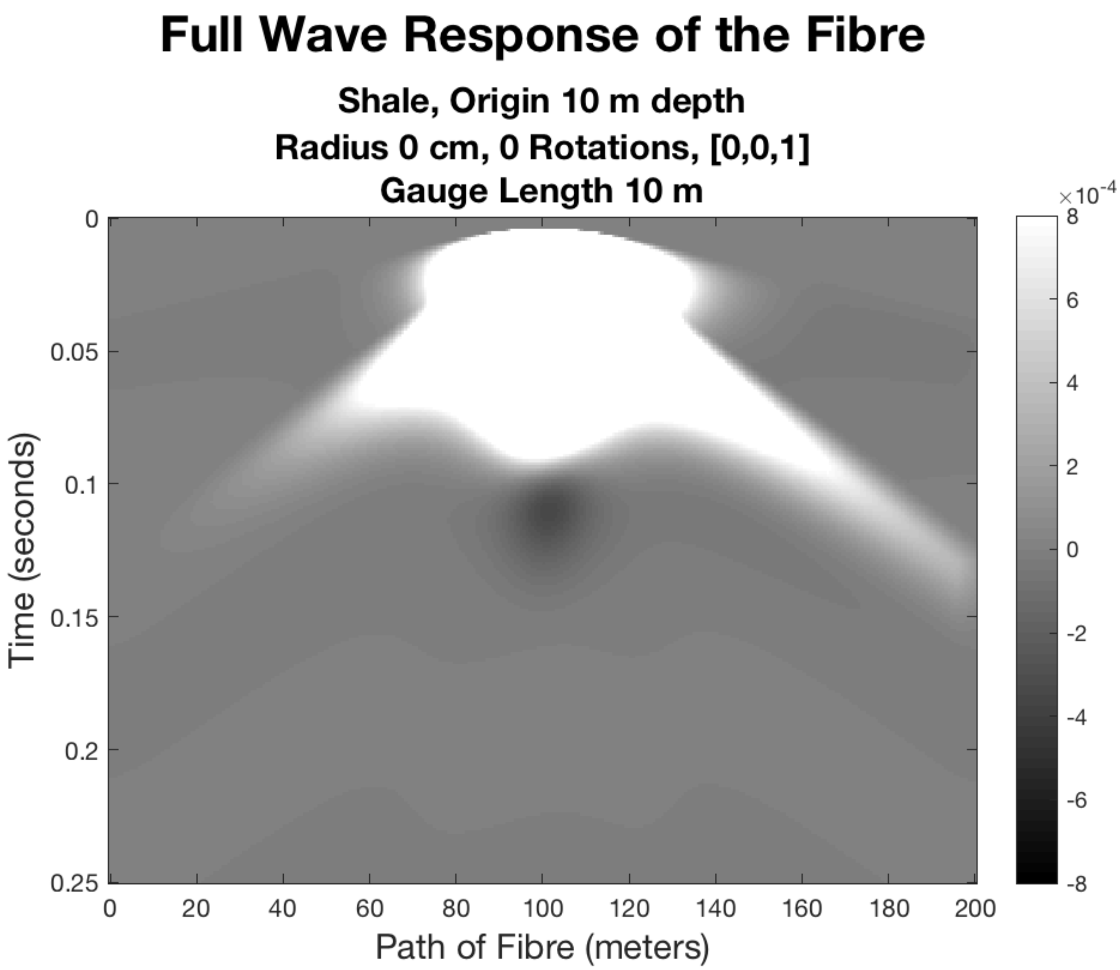


Figure 6.1: Full-waveform response of the straight fibre in saturated shale when the gauge length is set to 10m.

moving away from the source. The receiver is the point where the wave hits the fibre. We see that the wave hits two points on the fibre for this portion of the wave. As the wave made by the source moves further from the origin, the wave hits more of the fibre until either the fibre ends or the wave dissipates.

In the case of constant velocity, seismic waves move through the Earth's subsurface in a spherical shape. In fact, an exercise in calculus shows that the Laplace operator can be expressed in terms of radial derivatives:

$$\Delta\eta = \partial_r^2\eta + \frac{2}{r}\partial_r\eta \tag{6.1}$$

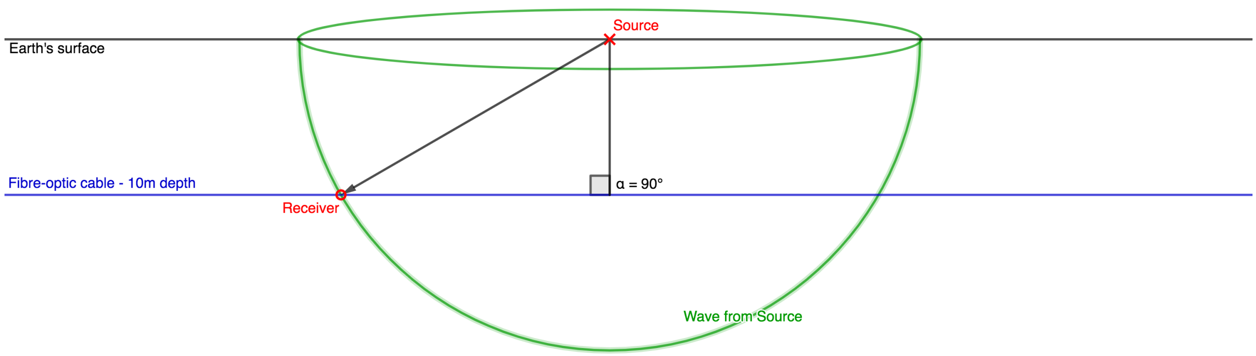


Figure 6.2: A physical model of the data acquisition using distributed acoustic sensing.

where $\eta = \eta(r)$ depends only on radial distance. We utilize the solution u of the elastic wave equation written in terms of the P-wave and S-wave components

$$\mathbf{u} = \nabla\Phi + \nabla \times \Psi \tag{6.2}$$

which we derive from the wave equation using the Helmholtz Decomposition Theorem; cf. Aki and Richards (2002) for details. Since derivatives are additive, the P-wave and S-wave components satisfy the wave equation and thus can be written using radial solutions. We will use this fact later when we begin creating our model.

An interrogator collects information from the fibre-optic cable when strain is placed on the fibre. This strain occurs when the seismic wave moves past the fibre and a portion of the laser pulse is returned to the interrogator. The strain of the fibre is the measure of how much the wave moves the fibre. The equation for strain is defined as follows:

$$\epsilon_{ij} = \frac{1}{2}(\partial_i \mathbf{u}_j + \partial_j \mathbf{u}_i), \tag{6.3}$$

where the strain ϵ is a symmetric 3×3 matrix with three eigenvectors and three eigenvalues which describe how much the material moves in each of three orthogonal directions.

In order to recover the response of the fibre, we consider the fact that the wave hits the fibre at any given point $\mathbf{p}(s_0)$ where $\mathbf{p} = (x, y, z)$ is a parameterization of the path of the fibre with real parameter s. At some point in space, the fibre-optic cable has a tangent vector $\mathbf{T_p}(s)$, normalized to unit length. The amount of stretching at this point is given by the product of the strain matrix with the tangent vector. Thus, we can determine the strain at this point on $\mathbf{p}$ by the following equation:

$$A(s,t) = \mathbf{T_p}^\top(s)\epsilon\Big(\mathbf{p}(s), t\Big)\mathbf{T_p}(s), \tag{6.4}$$

where the matrix $\epsilon(\mathbf{p}(s), t)$ is the strain at the point s on the path $\mathbf{p}$ of the fibre at time t.

To model the problem described above, we need to solve for the strain tensor. Using Equation 6.2, we solve for the strain produced by the full-waveform. Recall that

$$\epsilon_{ij} = \frac{1}{2}(\partial_i \mathbf{u}_j + \partial_j \mathbf{u}_i) \tag{6.5}$$

is the ij-th element of the strain tensor. Then, we have that

$$\begin{aligned}
\epsilon_{ij} &= \frac{1}{2}(\partial_i \mathbf{u}_j + \partial_j \mathbf{u}_i) && (6.6)\\
&= \frac{1}{2}(\partial_i((\nabla\Phi)_j + (\nabla \times \Psi)_j) + \partial_j((\nabla\Phi)_i + (\nabla \times \Psi)_i)) && (6.7)\\
&= \frac{1}{2}(\partial_i(\nabla\Phi)_j + \partial_i(\nabla \times \Psi)_j + \partial_j(\nabla\Phi)_i + \partial_j(\nabla \times \Psi)_i) && (6.8)\\
&= \frac{1}{2}(\partial_i(\nabla\Phi)_j + \partial_j(\nabla\Phi)_i) + \frac{1}{2}(\partial_i(\nabla \times \Psi)_j + \partial_j(\nabla \times \Psi)_i) && (6.9)\\
&= \epsilon^{\Phi}_{ij} + \epsilon^{\Psi}_{ij} && (6.10)
\end{aligned}$$

which implies that the full-waveform strain tensor is the sum of the P-wave strain tensor ϵ^{Φ} and the S-wave strain tensor ϵ^{Ψ}.

Solving for both tensors separately, we start with the P-wave strain tensor ϵ^{Φ}. For the purposes of this chapter, we work with a radial scalar potential

$$\Phi(x,y,z,t) = \frac{1}{r} F\Big(t - \frac{r}{\alpha}\Big) \tag{6.11}$$

where F is a minimum phase function, α is the velocity of the wave, and $r = \sqrt{x^2+y^2+z^2}$.

Taking the gradient of Φ, we find that

$$\nabla\Phi = \left(\frac{\partial}{\partial x}\frac{1}{r}F\Big(t-\frac{r}{\alpha}\Big), \frac{\partial}{\partial y}\frac{1}{r}F\Big(t-\frac{r}{\alpha}\Big), \frac{\partial}{\partial x}\frac{1}{r}F\Big(t-\frac{r}{\alpha}\Big)\right). \tag{6.12}$$

We solve for one component and the rest follow similarly. For the first component, we have

$$\frac{\partial}{\partial x}\frac{1}{r}F\Big(t-\frac{r}{\alpha}\Big) = \left(-\frac{1}{r^2}F + \Big(-\frac{1}{\alpha}\Big)\frac{1}{r}F'\right)\frac{\partial r}{\partial x} \tag{6.13}$$

$$= \left(-\frac{1}{r^2}F - \frac{1}{\alpha}\frac{1}{r}F'\right)\frac{x}{r} \tag{6.14}$$

$$= -\frac{x}{r^2}\left(\frac{1}{r}F + \frac{1}{\alpha}F'\right). \tag{6.15}$$

Hence, the gradient of Φ is

$$\nabla\Phi = (x,y,z)\left(-\frac{1}{r^2}\right)\left(\frac{1}{r}F + \frac{1}{\alpha}F'\right) \tag{6.16}$$

$$= (x,y,z)\left(-\frac{1}{r^2}\right)H_\alpha(r,t) \tag{6.17}$$

where

$$H_\alpha(r,t) = \left(\frac{1}{r}F\Big(t-\frac{r}{\alpha}\Big) + \frac{1}{\alpha}F'\Big(t-\frac{r}{\alpha}\Big)\right). \tag{6.18}$$

Then, we find the strain of $\nabla\Phi$ by for each component separately. Thus

$$\begin{aligned}
\epsilon_{11}^{\Phi} &= \frac{1}{2}(\partial_1(\nabla\Phi)_1 + \partial_1(\nabla\Phi)_1) &(6.19)\\
&= \frac{\partial(\nabla\Phi)_1}{\partial x} &(6.20)\\
&= \frac{\partial}{\partial x} x\left(-\frac{1}{r^2}\right) H_\alpha(r,t) &(6.21)\\
&= -\frac{1}{r^2}H_\alpha + x\frac{2}{r^3}\frac{\partial r}{\partial x}H_\alpha - x\frac{1}{r^2}\frac{\partial H_\alpha}{\partial x} &(6.22)\\
&= -\frac{1}{r^2}\left(H_\alpha + \left(\frac{1}{r}\frac{\partial H_\alpha}{\partial r} - \frac{2}{r^2}H_\alpha\right)x^2\right). &(6.23)
\end{aligned}$$

Similarly,

$$\begin{aligned}
\epsilon_{22}^{\Phi} &= -\frac{1}{r^2}\left(H_\alpha + \left(\frac{1}{r}\frac{\partial H_\alpha}{\partial r} - \frac{2}{r^2}H_\alpha\right)y^2\right); &(6.24)\\
\epsilon_{33}^{\Phi} &= -\frac{1}{r^2}\left(H_\alpha + \left(\frac{1}{r}\frac{\partial H_\alpha}{\partial r} - \frac{2}{r^2}H_\alpha\right)z^2\right). &(6.25)
\end{aligned}$$

For the off-diagonal terms, we compute

$$\begin{aligned}
\epsilon_{12}^{\Phi} &= \frac{1}{2}(\partial_1(\nabla\Phi)_2 + \partial_2(\nabla\Phi)_1) &(6.26)\\
&= \frac{1}{2}\left(\frac{\partial}{\partial x} y\left(-\frac{1}{r^2}\right) H_\alpha(r,t) + \frac{\partial}{\partial y} x\left(-\frac{1}{r^2}\right) H_\alpha(r,t)\right) &(6.27)\\
&= \frac{1}{2}\left(y\frac{2}{r^3}\frac{\partial r}{\partial x}H_\alpha - y\frac{1}{r^2}\frac{\partial H_\alpha}{\partial x} + x\frac{2}{r^3}\frac{\partial r}{\partial y}H_\alpha - x\frac{1}{r^2}\frac{\partial H_\alpha}{\partial y}\right) &(6.28)\\
&= \frac{1}{2}\left(xy\frac{2}{r^4}H_\alpha - xy\frac{1}{r^3}\frac{\partial H_\alpha}{\partial r} + xy\frac{2}{r^4}H_\alpha - xy\frac{1}{r^3}\frac{\partial H_\alpha}{\partial y}\right) &(6.29)\\
&= -\frac{xy}{r^2}\left(\frac{1}{r}\frac{\partial H_\alpha}{\partial r} - \frac{2}{r^2}H_\alpha\right). &(6.30)
\end{aligned}$$

By symmetry, we get

$$\epsilon_{21}^{\Phi} = -\frac{xy}{r^2}\left(\frac{1}{r}\frac{\partial H_\alpha}{\partial r} - \frac{2}{r^2}H_\alpha\right). \tag{6.31}$$

Similarly, we find

$$\epsilon_{13}^{\Phi} = \epsilon_{31}^{\Phi} = -\frac{xz}{r^2}\left(\frac{1}{r}\frac{\partial H_\alpha}{\partial r} - \frac{2}{r^2}H_\alpha\right); \tag{6.32}$$

$$\epsilon_{23}^{\Phi} = \epsilon_{32}^{\Phi} = -\frac{yz}{r^2}\left(\frac{1}{r}\frac{\partial H_\alpha}{\partial r} - \frac{2}{r^2}H_\alpha\right). \tag{6.33}$$

Therefore, the strain tensor ϵ^{Φ} is

$$\epsilon^{\Phi} = -\frac{1}{r^2}\Big(H_\alpha I + G_\alpha(r,t)\mathbf{R_\Phi}\Big), \tag{6.34}$$

where

$$G_\alpha(r,t) = \frac{1}{r}\frac{\partial H_\alpha}{\partial r} - \frac{2}{r^2}H_\alpha, \tag{6.35}$$

where I is the 3×3 identity matrix, and

$$\mathbf{R_\Phi} = \begin{bmatrix} x^2 & xy & xz \\ xy & y^2 & yz \\ xz & yz & z^2 \end{bmatrix} \tag{6.36}$$

$$= \begin{bmatrix} x \\ y \\ z \end{bmatrix} \begin{bmatrix} x & y & z \end{bmatrix} \tag{6.37}$$

$$= \mathbf{r}\mathbf{r}^{\top}, \tag{6.38}$$

where $\mathbf{r} = [x, y, z]^\top$. Using standard calculus and arithmetic techniques, we find that

$$\frac{\partial H_\alpha}{\partial r} = -\frac{1}{r^2}F\Big(t-\frac{r}{\alpha}\Big) - \frac{1}{r}\frac{1}{\alpha}F'\Big(t-\frac{r}{\alpha}\Big) + \frac{1}{\alpha^2}F''\Big(t-\frac{r}{\alpha}\Big) \tag{6.39}$$

$$= -\frac{1}{r}\Big(\frac{1}{r}F\Big(t-\frac{r}{\alpha}\Big) + \frac{1}{\alpha}F'\Big(t-\frac{r}{\alpha}\Big)\Big) - \frac{1}{\alpha^2}F''\Big(t-\frac{r}{\alpha}\Big) \tag{6.40}$$

$$= -\frac{1}{r}H_\alpha(r,t) - \frac{1}{\alpha^2}F''\Big(t-\frac{r}{\alpha}\Big) \tag{6.41}$$

$$\tag{6.42}$$

and hence

$$G_\alpha(r,t) = \frac{1}{r}\frac{\partial H_\alpha}{\partial r} - \frac{2}{r^2}H_\alpha(r,t) \tag{6.43}$$

$$= \frac{1}{r}\Big(-\frac{1}{r}H_\alpha(r,t) - \frac{1}{\alpha}F''\Big(t-\frac{r}{\alpha}\Big)\Big) - \frac{2}{r^2}H_\alpha(r,t) \tag{6.44}$$

$$= -\frac{1}{r^2}H_\alpha(r,t) - \frac{1}{r\alpha^2}F''\Big(t-\frac{r}{\alpha}\Big) - \frac{2}{r^2}H_\alpha(r,t) \tag{6.45}$$

$$= -\frac{3}{r^2}H_\alpha(r,t) - \frac{1}{r\alpha^2}F''\Big(t-\frac{r}{\alpha}\Big). \tag{6.46}$$

With this in mind, it is convenient to normalize $\mathbf{R}_{\boldsymbol{\Phi}}$ to obtain

$$\hat{\mathbf{R}}_\Phi = \frac{1}{r^2}\mathbf{R}_{\boldsymbol{\Phi}} = \frac{1}{r^2}\mathbf{r}\mathbf{r}^\top \tag{6.47}$$

as a rank one symmetric matrix with norm one. Then Equation 6.34 is written as

$$\epsilon^\Phi = -\frac{1}{r^2}\Big(H_\alpha(r,t)\mathbf{I} + \big(-3H_\alpha(r,t) - \frac{r}{\alpha^2}F''\Big(t-\frac{r}{\alpha}\Big)\big)\hat{\mathbf{R}}_\Phi\Big). \tag{6.48}$$

We take a moment to consider the physical significance of this expression. It indicates that the eigenvalue in the $\mathbf{r}$ direction is opposite in sign to the other two eigenvalues. This suggests that a reversal in polarity occurs in the direction of the fibre, relative to the source-receiver direction.

For the S-wave strain tensor, ϵ^{Ψ}, we define the vector potential as

$$\Psi(x, y, z, t) = \frac{1}{r} F \left(t - \frac{r}{\beta} \right) [A_x, A_y, A_z], \tag{6.49}$$

where β is the velocity of the S-wave, $r = \sqrt{x^2 + y^2 + z^2}$, and $\mathbf{A} = (A_x, A_y, A_z)$ is the source orientation vector for arbitrary A_x, A_y, $A_z \in \mathbb{R}$. The function F is the same as the function F for the P-wave; see Equation 6.11. We compute the curl

$$\begin{aligned} \nabla \times \Psi &= \begin{bmatrix} \widehat{i} & \widehat{j} & \widehat{k} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ \Psi_1 & \Psi_2 & \Psi_3 \end{bmatrix} & (6.50) \\ &= \left(\frac{\partial \Psi_3}{\partial y} - \frac{\partial \Psi_2}{\partial z}, \frac{\partial \Psi_1}{\partial z} - \frac{\partial \Psi_3}{\partial x}, \frac{\partial \Psi_2}{\partial x} - \frac{\partial \Psi_1}{\partial y} \right). & (6.51) \end{aligned}$$

Computing the partial derivatives of each $(\nabla \times \Psi)_n$ for $n = 1, 2, 3$, we have

$$\begin{aligned} (\nabla \times \Psi)_1 &= \frac{\partial \Psi_3}{\partial y} - \frac{\partial \Psi_2}{\partial z} & (6.52) \\ &= \frac{\partial}{\partial y} \left(\frac{1}{r} F \left(t - \frac{r}{\beta} \right) A_z \right) - \frac{\partial}{\partial z} \left(\frac{1}{r} F \left(t - \frac{r}{\beta} \right) A_y \right) & (6.53) \\ &= \left(-\frac{1}{r^2} \frac{y}{r} F - \frac{1}{r^2} \frac{y}{r} \frac{1}{\beta} F' \right) A_z - \left(-\frac{1}{r^2} \frac{z}{r} F - \frac{1}{r^2} \frac{z}{r} \frac{1}{\beta} F' \right) A_y & (6.54) \\ &= -\frac{1}{r^2} H_\beta(r, t)(A_z y - A_y z); & (6.55) \\ (\nabla \times \Psi)_2 &= \frac{\partial \Psi_1}{\partial z} - \frac{\partial \Psi_3}{\partial x} & (6.56) \\ &= -\frac{1}{r^2} H_\beta(r, t)(A_x z - A_z x); & (6.57) \\ (\nabla \times \Psi)_3 &= \frac{\partial \Psi_2}{\partial x} - \frac{\partial \Psi_1}{\partial y} & (6.58) \\ &= -\frac{1}{r^2} H_\beta(r, t)(A_y x - A_x y). & (6.59) \end{aligned}$$

Hence, the curl is

$$\nabla \times \Psi = -\frac{1}{r^2} H_\beta(r, t)[\mathbf{X}, \mathbf{Y}, \mathbf{Z}] \tag{6.60}$$

where

$$\mathbf{X} = A_z y - A_y z; \tag{6.61}$$

$$\mathbf{Y} = A_x z - A_z x; \tag{6.62}$$

$$\mathbf{Z} = A_y x - A_x y. \tag{6.63}$$

To solve for each component of the S-wave strain ϵ^{Ψ}, we have

$$\begin{aligned}
\epsilon_{11} &= \frac{\partial}{\partial x}(\nabla \times \Psi)_1 && (6.64)\\
&= \frac{\partial}{\partial x}\left(-\frac{1}{r^2}H_\beta(r,t)\right)\mathbf{X} && (6.65)\\
&= \frac{2}{r^3}\frac{\partial r}{\partial x}H_\beta \mathbf{X} + \left(-\frac{1}{r^2}\right)\frac{H_\beta}{\partial r}\frac{\partial r}{\partial x}\mathbf{X} + \left(-\frac{1}{r^2}\right)H_\beta \frac{\partial \mathbf{X}}{\partial x} && (6.66)\\
&= -\frac{1}{r^2}\left(\frac{1}{r}\frac{\partial H_\beta}{\partial r} - \frac{2}{r^2}H_\beta\right)x\mathbf{X} && (6.67)
\end{aligned}$$

since

$$\frac{\partial \mathbf{X}}{\partial x} = \frac{\partial}{\partial x}(A_z y - A_y z) = 0. \tag{6.68}$$

Let

$$G_\beta(r,t) = \frac{1}{r}\frac{\partial H_\beta}{\partial r} - \frac{2}{r^2}H_\beta. \tag{6.69}$$

By a similar argument,

$$\epsilon_{22} = -\frac{1}{r^2}G_\beta(r,t)y\mathbf{Y}; \tag{6.70}$$

$$\epsilon_{33} = -\frac{1}{r^2}G_\beta(r,t)z\mathbf{Z}. \tag{6.71}$$

For the off-diagonal cases, it suffices to find ϵ_{12}, ϵ_{13}, and ϵ_{23}. By symmetry, we then have

$$
\begin{aligned}
\epsilon_{12} &= \frac{1}{2}\left(\frac{\partial}{\partial x}(\nabla\times\Psi)_2 + \frac{\partial}{\partial y}(\nabla\times\Psi)_1\right) && (6.72)\\
&= \frac{1}{2}\left(\frac{\partial}{\partial x}\left(-\frac{1}{r^2}H_\beta(r,t)\mathbf{Y}\right) + \frac{\partial}{\partial y}\left(-\frac{1}{r^2}H_\beta(r,t)\mathbf{X}\right)\right) && (6.73)\\
&= \frac{1}{2}\left(-\frac{1}{r^2}\right)\left(\left(-\frac{2}{r}\frac{x}{r}H_\beta\mathbf{Y} + \frac{x}{r}\frac{\partial H_\beta}{\partial r}\mathbf{Y} + H_\beta\frac{\partial \mathbf{Y}}{\partial x}\right)\right. && (6.74)\\
&\quad \left. + \left(-\frac{2}{r}\frac{y}{r}H_\beta\mathbf{X} + \frac{y}{r}\frac{\partial H_\beta}{\partial r}\mathbf{X} + H_\beta\frac{\partial \mathbf{X}}{\partial y}\right)\right) && (6.75)\\
&= \frac{1}{2}\left(-\frac{1}{r^2}\right)\left(\left(-\frac{2}{r^2}H_\beta x\mathbf{Y} + \frac{1}{r}\frac{\partial H_\beta}{\partial r}x\mathbf{Y} + H_\beta(-A_z)\right)\right. && (6.76)\\
&\quad \left. + \left(-\frac{2}{r^2}H_\beta y\mathbf{X} + \frac{1}{r}\frac{\partial H_\beta}{\partial r}y\mathbf{X} + H_\beta A_z\right)\right) && (6.77)\\
&= \frac{1}{2}\left(-\frac{1}{r^2}\right)\left(\frac{1}{r}\frac{\partial H_\beta}{\partial r} - \frac{2}{r^2}H_\beta\right)(x\mathbf{Y} + y\mathbf{X}) && (6.78)\\
&= \frac{1}{2}\left(-\frac{1}{r^2}\right)G_\beta(r,t)\big(x\mathbf{Y} + y\mathbf{X}\big). && (6.79)
\end{aligned}
$$

Using a similar method, we get that

$$
\begin{aligned}
\epsilon_{13} &= \frac{1}{2}\left(-\frac{1}{r^2}\right)G_\beta(r,t)\big(x\mathbf{Z} + z\mathbf{X}\big); && (6.80)\\
\epsilon_{23} &= \frac{1}{2}\left(-\frac{1}{r^2}\right)G_\beta(r,t)\big(z\mathbf{Y} + y\mathbf{Z}\big). && (6.81)
\end{aligned}
$$

Symmetry gives ϵ_{21}, ϵ_{31}, and ϵ_{32} respectively. Therefore, the complete strain equation ϵ^Ψ as

$$
\epsilon^\Psi = \frac{1}{2}\left(-\frac{1}{r^2}\right)G_\beta(r,t)\mathbf{R}_{\boldsymbol{\Psi}}, \tag{6.82}
$$

where

$$\mathbf{R}_{\boldsymbol{\Psi}} = \begin{bmatrix} 2x\mathbf{X} & x\mathbf{Y}+y\mathbf{X} & x\mathbf{Z}+z\mathbf{X} \\ y\mathbf{X}+x\mathbf{Y} & 2y\mathbf{Y} & y\mathbf{Z}+z\mathbf{Y} \\ z\mathbf{X}+x\mathbf{Z} & z\mathbf{Y}+y\mathbf{Z} & 2x\mathbf{X} \end{bmatrix} \tag{6.83}$$

$$= \begin{bmatrix} x \\ y \\ z \end{bmatrix} \begin{bmatrix} \mathbf{X} & \mathbf{Y} & \mathbf{Z} \end{bmatrix} + \begin{bmatrix} \mathbf{X} \\ \mathbf{Y} \\ \mathbf{Z} \end{bmatrix} \begin{bmatrix} x & y & z \end{bmatrix} \tag{6.84}$$

$$= \mathbf{r}(\mathbf{r}\times\mathbf{A})^\top + (\mathbf{r}\times\mathbf{A})\mathbf{R}^\top, \tag{6.85}$$

where $\mathbf{r} = (x, y, z)$ is the position vector and $\mathbf{A} = (A_x, A_y, A_z)$ is the source orientation vector. We can also normalize the matrix $\mathbf{R}_{\boldsymbol{\Psi}}$ as

$$\hat{\mathbf{R}}_\Psi = \frac{1}{r^2}\mathbf{R}_\Psi \tag{6.86}$$

$$= \frac{1}{r^2}\big(\mathbf{r}(\mathbf{r}\times\mathbf{A})^\top + (\mathbf{r}\times\mathbf{A})\mathbf{r}^\top\big), \tag{6.87}$$

and so the strain matrix becomes

$$\epsilon^\Psi = -\frac{1}{2}G_\beta(r,t)\hat{\mathbf{R}}_\Psi \tag{6.88}$$

$$= -\frac{1}{2}\frac{1}{r^2}\Big(3H_\beta(r,t) + \frac{r}{\beta^2}F''\Big(t-\frac{r}{\beta}\Big)\Big)\hat{\mathbf{R}}_\Psi. \tag{6.89}$$

Before we continue, it is useful to consider the physical implications. The factor H_β is the dominant part. The symmetric matrix $\hat{\mathbf{R}_{\boldsymbol{\Psi}}}$ has two non-zero eigenvalues of opposite sign, so the direction of the fibre could interpolate between these two, depending on its direction. With helical fibres, we expect to see oscillations in the recorded signal.

We compute the strain tensor ϵ for the full-waveform to be

$$\begin{aligned}
\epsilon &= \epsilon^{\Phi} + \epsilon^{\Psi} && (6.90)\\
&= -\frac{1}{r^2}\Big(H_\alpha(r,t)\mathbf{I} + G_\alpha(r,t)\mathbf{R_\Phi}\Big) + \frac{1}{2}\left(-\frac{1}{r^2}\right) G_\beta(r,t)\mathbf{R_\Psi} && (6.91)\\
&= -\frac{1}{r^2}\Big(H_\alpha(r,t)\mathbf{I} + G_\alpha(r,t)\mathbf{R_\Phi} + \frac{1}{2}G_\beta(r,t)\mathbf{R_\Psi}\Big) && (6.92)\\
&= -\frac{1}{r^2}\Big(H_\alpha(r,t)\mathbf{I} + \big(-3H_\alpha(r,t) - \frac{r}{\alpha^2}F''\Big(t-\frac{r}{\alpha}\Big)\Big)\hat{\mathbf{R}}_\Phi && (6.93)\\
&\quad + \frac{1}{2}\Big(3H_\beta(r,t) + \frac{r}{\beta^2}F''\Big(t-\frac{r}{\beta}\Big)\Big)\hat{\mathbf{R}}_\Psi\Big) && (6.94)
\end{aligned}$$

These equations include both near field and far field terms.

Now that we have calculated the strain tensor results for the P-wave, S-wave, and full-waveform, we consider examples for P-wave and S-wave separately as well as the full-waveform. During these examples, we also compare responses given different minerals, observe differences between straight and helical fibre, and examine the effect of varying the gauge length.

6.2 Examples

In this section, we consider three examples. The first example looks at the P-wave response of a straight fibre for two different velocities. We will see the effect that different minerals have on the response of the fibre. The second example compares the S-wave response of the straight and helical fibre. We examine the differences between the two cases from the fibre in one mineral. In the final example, we analyze the results of the full-wave response of the fibre in one mineral given different gauge lengths.

For the models in this chapter, we used the source function

$$F(t) = \begin{cases} te^{-\sigma_1 t}\sin(\sigma_2 t) & t \geq 0 \\ 0 & t < 0, \end{cases} \tag{6.95}$$

where $\sigma_1 = 80$Hz and $\sigma_2 = 50$Hz. We chose this function in order to mimic a minimum-phase system for our models.

6.2.1 P-wave example

We begin with a quick study of the fibre's response given different media. Using the P-wave as an example, we examine two different media: saturated shale and limestone.

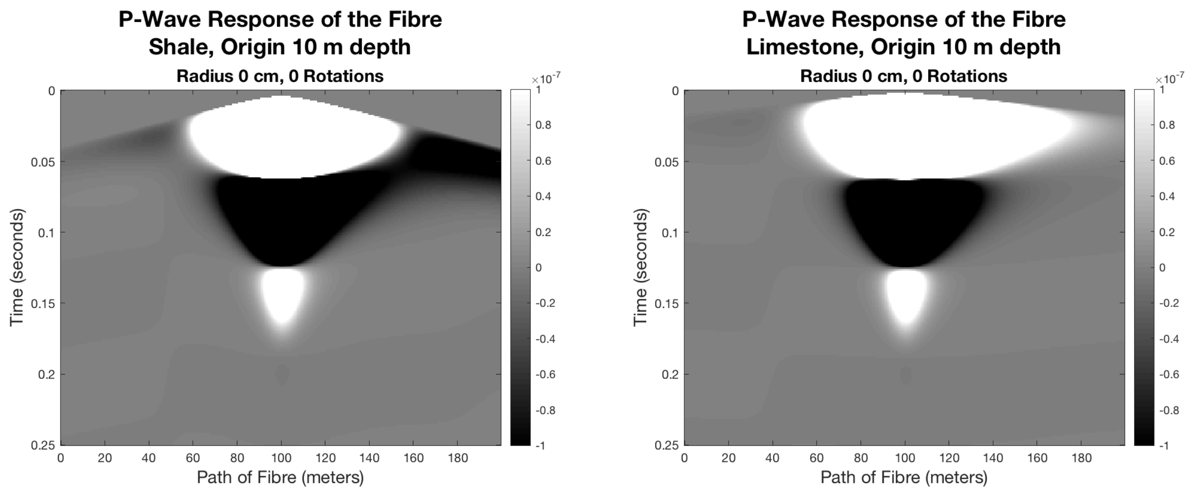


Figure 6.3: The P-wave response of the straight fibre in (left) saturated shale at 2500 m/s and (right) limestone at 6000 m/s

Figure 6.3 gives the results of the straight fibre's response for the P-wave in saturated shales (left) and limestone (right). Both responses are hyperbolic in shape; however, the hyperbola has a steeper slope in saturated shales as opposed to the more shallow slope of the limestone. This follows physically since the wave moves faster in limestone than it does in shale.

6.2.2 S-wave example

For the S-wave response, we examine the differences between straight fibre and helical fibre. For the helical fibre, the setup remains the same as the straight fibre, i.e., laid horizontally 10 meters deep; however, the fibre forms a helix of with radius 1cm and has 10 rotations per meter. Over the 200 meter distance we consider in the experiment, the fibre has 2000 rotations. We only look at the results in saturated shales. Given the results of the P-wave example, the S-wave response for limestone would be a shallow hyperbola in comparison to the steeper hyperbola we will see given by the saturate shales, as the velocity of the S-wave for limestone is 3300 m/s as opposed to the 800 m/s velocity of the S-wave for saturated shales. As with the P-wave example, we use the max velocity of the S-wave for the media as given in Bourbié et al. (1987).

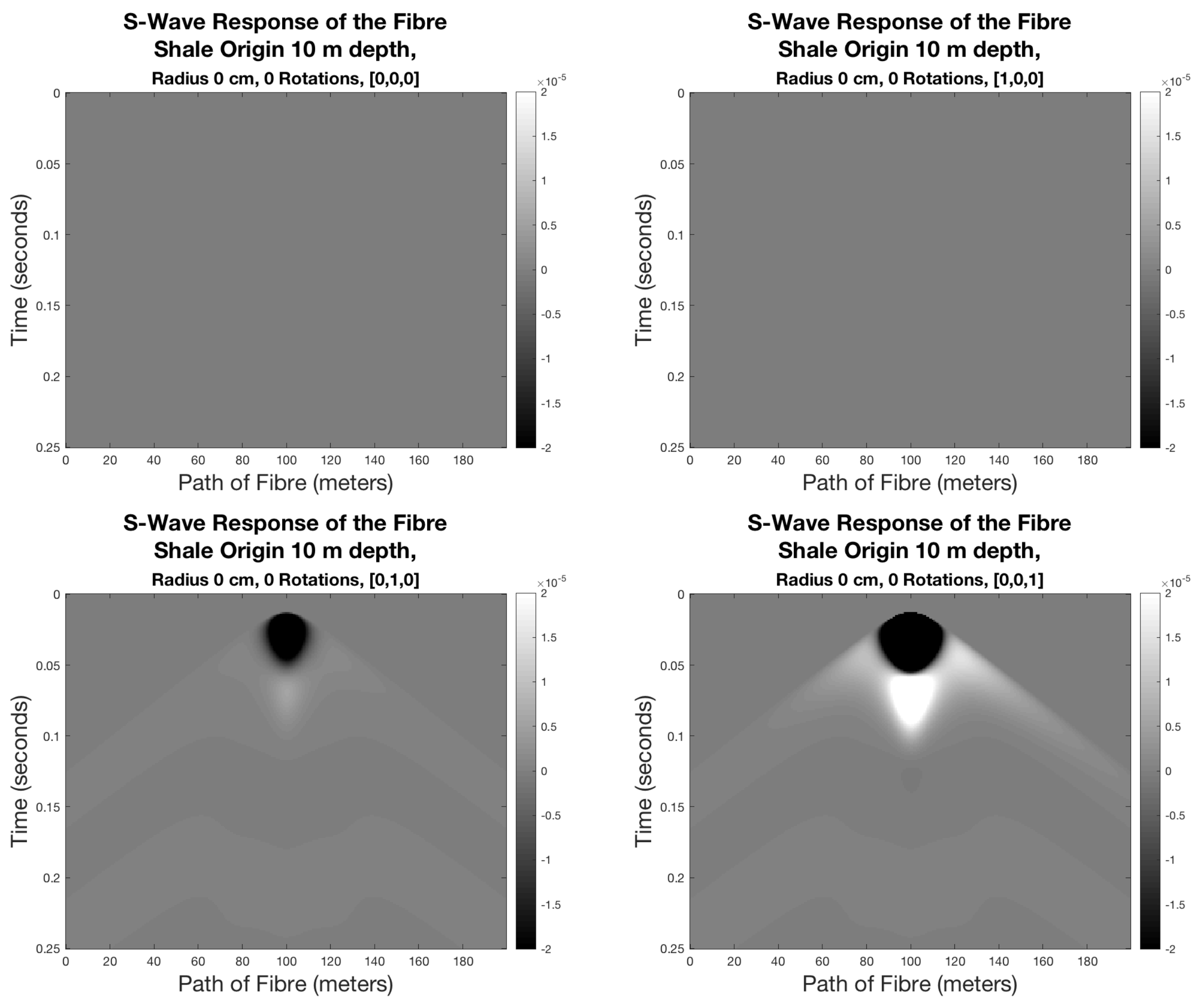


Figure 6.4: The S-wave response of the straight fibre in saturated shale when (top left) $\mathbf{A} = [0, 0, 0]$, (top right) $\mathbf{A} = [1, 0, 0]$, (bottom left) $\mathbf{A} = [0, 1, 0]$, and (bottom right) $\mathbf{A} = [0, 0, 1]$

Figures 6.4 and 6.5 show the S-wave response of the straight fibre in saturated shales

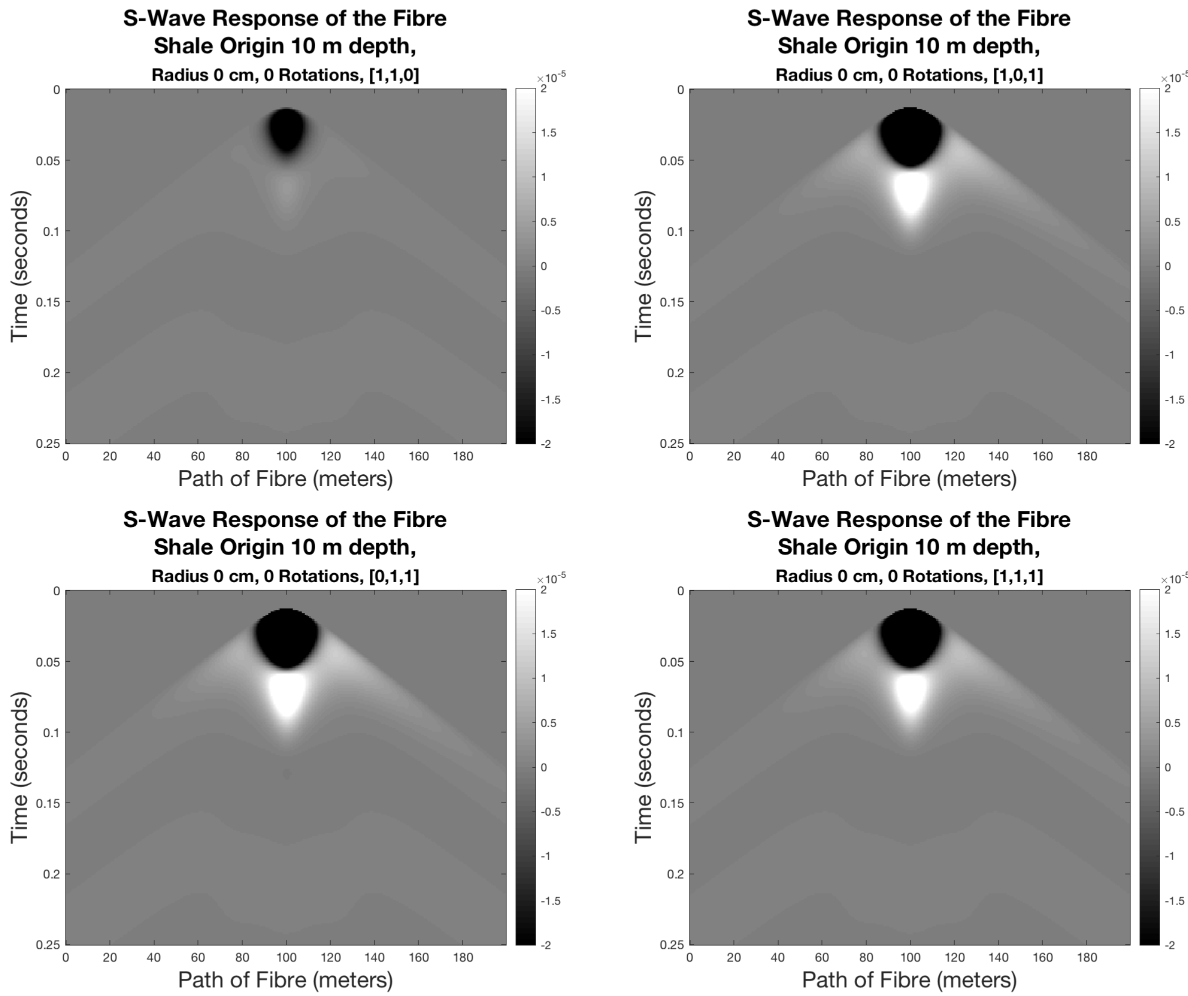


Figure 6.5: The S-wave response of the straight fibre in saturated shale when (top left) $\mathbf{A} = [1, 1, 0]$, (top right) $\mathbf{A} = [1, 0, 1]$, (bottom left) $\mathbf{A} = [0, 1, 1]$, and (bottom right) $\mathbf{A} = [1, 1, 1]$

for different vectors $\mathbf{A}$. When $\mathbf{A} = [0, 0, 0]$, the S-wave response is understandably 0. For $\mathbf{A} = [1, 0, 0]$, the fibre gives no response; yet, for all other vectors $\mathbf{A}$, the straight fibre does detect stress and strain. For vectors $\mathbf{A} = [0, 1, 0]$ and $A = [1, 1, 0]$, the response has a weaker amplitude than the other vectors $\mathbf{A}$ where the fibre picks up a signal. This result suggests that the straight fibre does not detect stress or strain in the x-direction.

Visually, the results for $\mathbf{A} = [0, 0, 1]$ look brighter than the results for the other vectors $\mathbf{A}$ with non-zero z components. A comparison shows a difference ranging between $(-2.01 \times 10^{-4}, 1.20 \times 10^{-4})$ for the response when $\mathbf{A} = [0, 1, 1]$ and $(-2.66 \times 10^{-4}, 1.69 \times 10^{-5})$ for the response when $\mathbf{A} = [1, 0, 1]$ whereas the difference ranges between -3.31×10^{-4} and 2.11×10^{-5} for the response when $\mathbf{A} = [1, 1, 1]$. These results imply that having non-zero $x-$ and $y-$ components affect the response of the fibre more than having only one of the $x-$ and $y-$ components be non-zero.

The addition of a non-zero $x-$ or $y-$ component to the vector A that has a non-zero z-component ($\mathbf{A} = [0, 1, 1]$, $\mathbf{A} = [1, 0, 1]$, $\mathbf{A} = [1, 1, 1]$) results in a different response whereas only one non-zero $x-$ or $y-$ component ($\mathbf{A} = [1, 0, 0]$,$\mathbf{A} = [0, 1, 0]$) does not guarantee a response from the straight fibre. This suggests that the formation of the fibre with respect to the orientation of the source affects the response of the fibre. In this case, the vector $\mathbf{A} = [1, 0, 0]$ causes the source for the S-wave to hit the straight fibre perpendicularly.

Figures 6.6 and 6.7 exhibit the S-wave response of the helical fibre in saturated shale for different vectors $\mathbf{A}$. As with the case for the straight fibre, the vector $\mathbf{A} = [0, 0, 0]$ provides a response of 0; the result for vector $\mathbf{A} = [1, 0, 0]$ shows that the fibre detects some strain from the wave. While it is not much, it is still more than the straight fibre provided. Table 6.1 shows the range of the amplitude for the straight and helical fibre response at different vectors $\mathbf{A}$.

Besides producing a response for the vector $\mathbf{A} = [1, 0, 0]$ unlike the straight fibre's response, the helical fibre produces different response ranges for each vector $\mathbf{A}$. The same cannot be said for the straight fibre as while providing different responses, the straight fi-

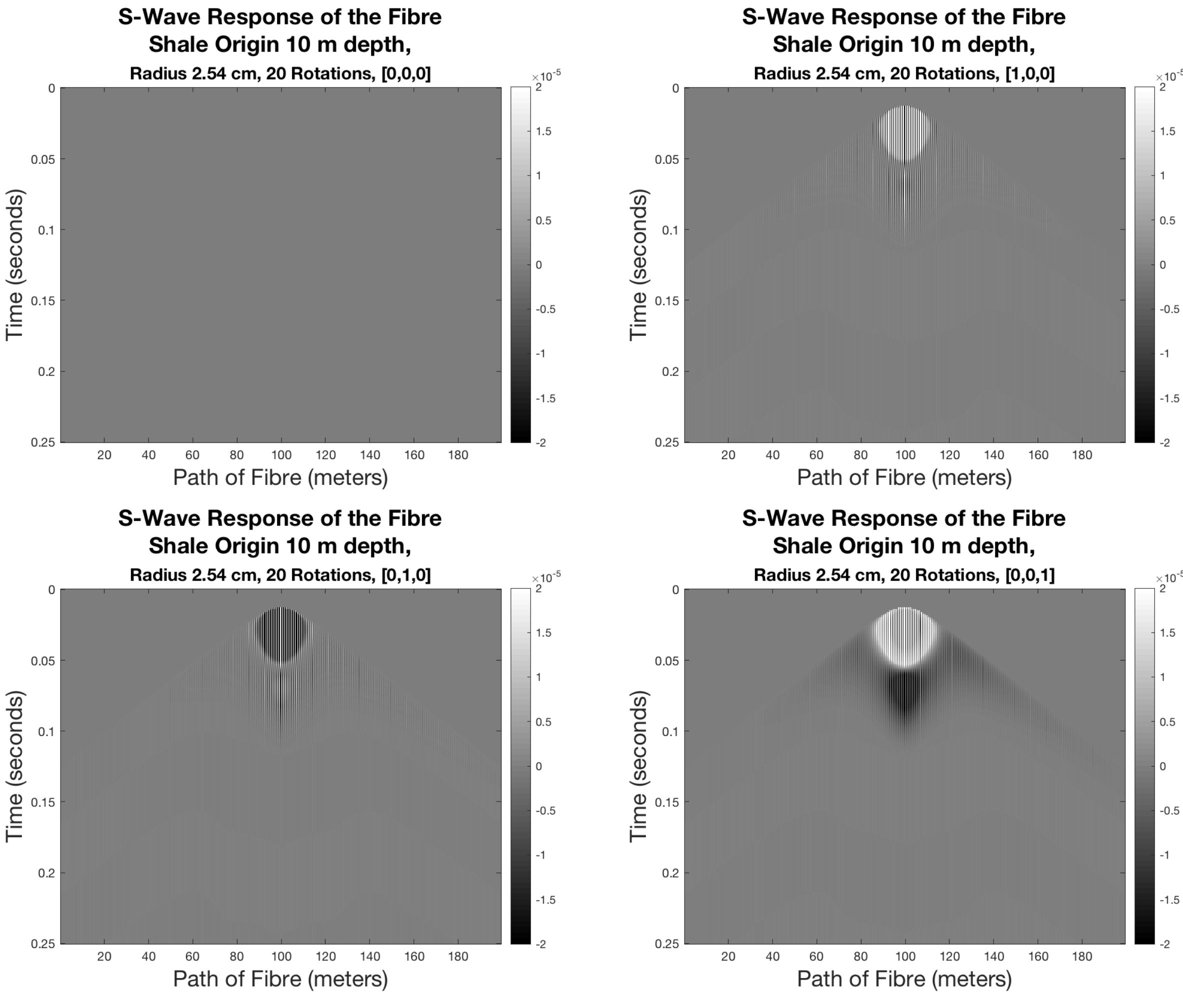


Figure 6.6: The S-wave response of the helical fibre in saturated shale when (top left) $\mathbf{A} = [0,0,0]$, (top right) $\mathbf{A} = [1,0,0]$, (bottom left) $\mathbf{A} = [0,1,0]$, and (bottom right) $\mathbf{A} = [0,0,1]$

Table 6.1: Amplitude range of the straight and helical fibre response for different vectors $\mathbf{A}$. The units are meters for each value in the table.

Vector $\mathbf{A}$	Straight Fibre Range	Helical Fibre Range
$[0,0,0]$	0	0
$[1,0,0]$	0	$(-4.14 \times 10^{-4}, 2.64 \times 10^{-5})$
$[0,1,0]$	$(-9.07 \times 10^{-5}, 5.77 \times 10^{-6})$	$(-3.01 \times 10^{-5}, 4.70 \times 10^{-4})$
$[0,0,1]$	$(-9.07 \times 10^{-5}, 5.77 \times 10^{-5})$	$(-3.55 \times 10^{-5}, 5.56 \times 10^{-4})$
$[1,1,0]$	$(-6.41 \times 10^{-5}, 4.08 \times 10^{-6})$	$(-2.77 \times 10^{-6}, 4.21 \times 10^{-5})$
$[0,1,1]$	$(-7.05 \times 10^{-4}, 4.49 \times 10^{-5})$	$(-4.64 \times 10^{-5}, 7.26 \times 10^{-4})$
$[1,0,1]$	$(-6.41 \times 10^{-4}, 4.08 \times 10^{-5})$	$(-6.66 \times 10^{-6}, 1.03 \times 10^{-4})$
$[1,1,1]$	$(-5.76 \times 10^{-4}, 3.66 \times 10^{-5})$	$(-2.27 \times 10^{-5}, 3.54 \times 10^{-4})$

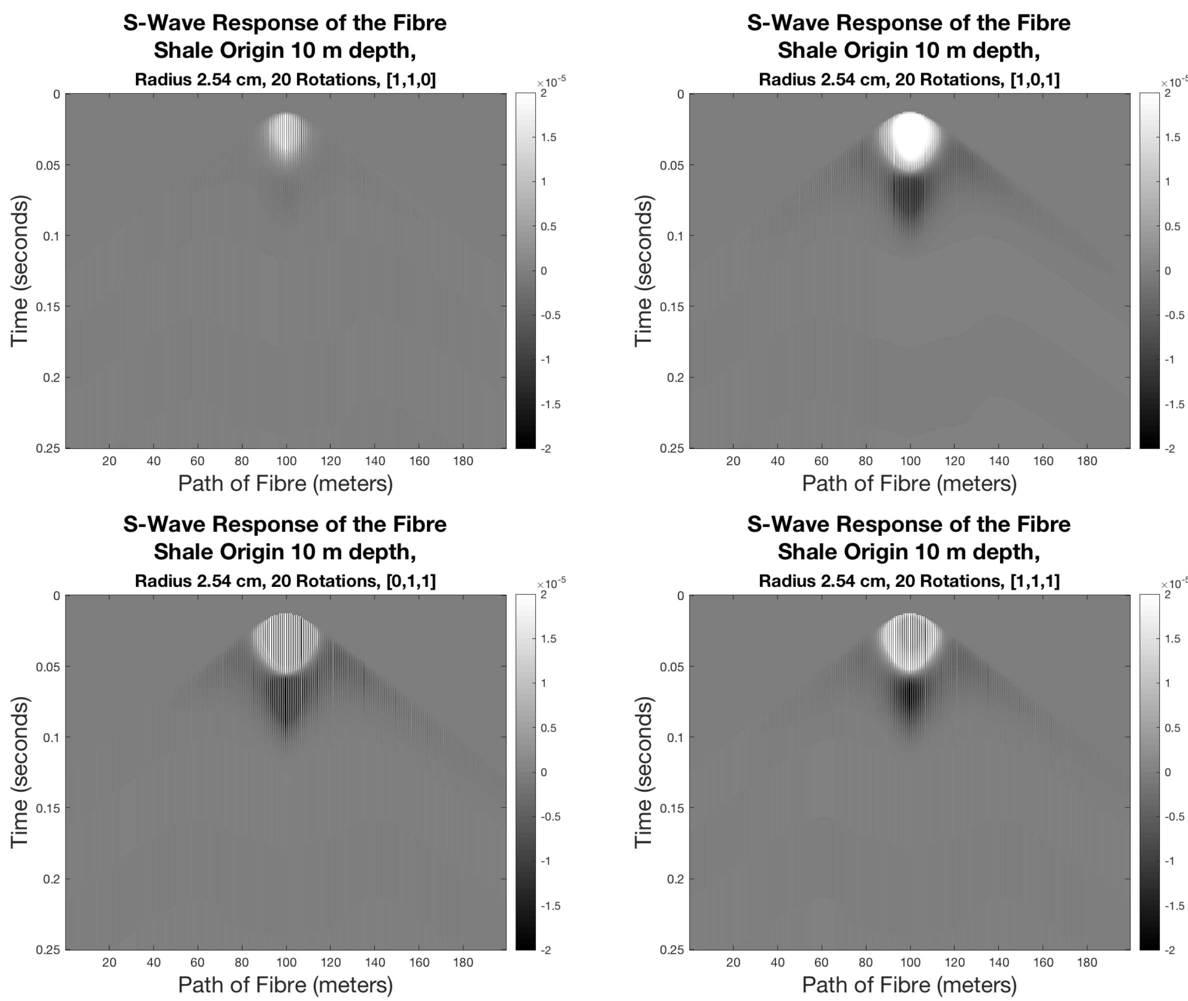


Figure 6.7: The S-wave response of the helical fibre in saturated shale when (top left) $\mathbf{A} = [1, 1, 0]$, (top right) $\mathbf{A} = [1, 0, 1]$, (bottom left) $\mathbf{A} = [0, 1, 1]$, and (bottom right) $\mathbf{A} = [1, 1, 1]$

bre produces equivalent range differences in the amplitude for the vectors $\mathbf{A} = [0, 1, 0]$ and $\mathbf{A} = [0, 0, 1]$, and $\mathbf{A} = [1, 1, 0]$ and $\mathbf{A} = [1, 0, 1]$ respectively. The maximum amplitude of the helical fibre is larger than the maximum amplitude of the straight fibre by a factor of 10^{-1} for most vectors $\mathbf{A}$.

More vectors $\mathbf{A}$ provide a response for the helical fibre, which supports the use of helical fibre in acquisition as it is more sensitive to strain in all three directions as opposed to only two directions we saw in the straight fibre. Not only is the helical fibre more sensitive, but the data set for the helical fibre is actually larger than the straight fibre. For all vectors $\mathbf{A}$, the data set of the straight fibre is 228×300 (Time t $\times$ Path of Fibre $\mathbf{p}(s)$) whereas the data set of the helical fibre is 228×1000. The difference results from the fact that the sample points are taken along the helical distance of the fibre. Since the sample points are 2/3m apart, the straight fibre only gives sample points equal to the length of the fibre divided by 2/3. When the fibre is helical, the length of the fibre over 200 meters increases. The sample point is taken at 2/3m distance along the arc length of the fibre. So, 2/3 m on the helical fibre is less linear ground distance than the 2/3m on the straight fibre, which means the helical fibre has more sample points than the straight fibre. The fact that helical fibre provides more sample points than straight fibre also supports the use of helically wound fibre in seismic acquisition.

6.2.3 Full-waveform example

Now we combine the results of the P-wave and S-wave examples to study the effect that different gauge lengths have on the full-waveform response. Recall that the gauge length is a property of the DAS system related to the pulse width of the laser interferometer. To replicate this in our model, we convolve the gauge length function

$$g(\tau) = \begin{cases} 1 & -\frac{\gamma}{2} \leq \tau \leq \frac{\gamma}{2} \\ 0 & \text{otherwise,} \end{cases} \tag{6.96}$$

where γ is the distance of the gauge length with the amplitude response of the fibre $A(s,t)$. This gives us the integral

$$\int_{s-\frac{\gamma}{2}}^{s+\frac{\gamma}{2}} A(\tau,t)d\tau \tag{6.97}$$

which we apply to the amplitude response of the fibre $A(s,t)$ for $0 < s \leq 200$ meters. We conduct the experiment for saturated shale. We also only consider the vector $\mathbf{A} = [0,0,1]$ as it produced the sharpest image for both straight and helical fibre. Over a distance of 200 meters, we compare the following gauge lengths: 5m, 10m, 20m and 25m.

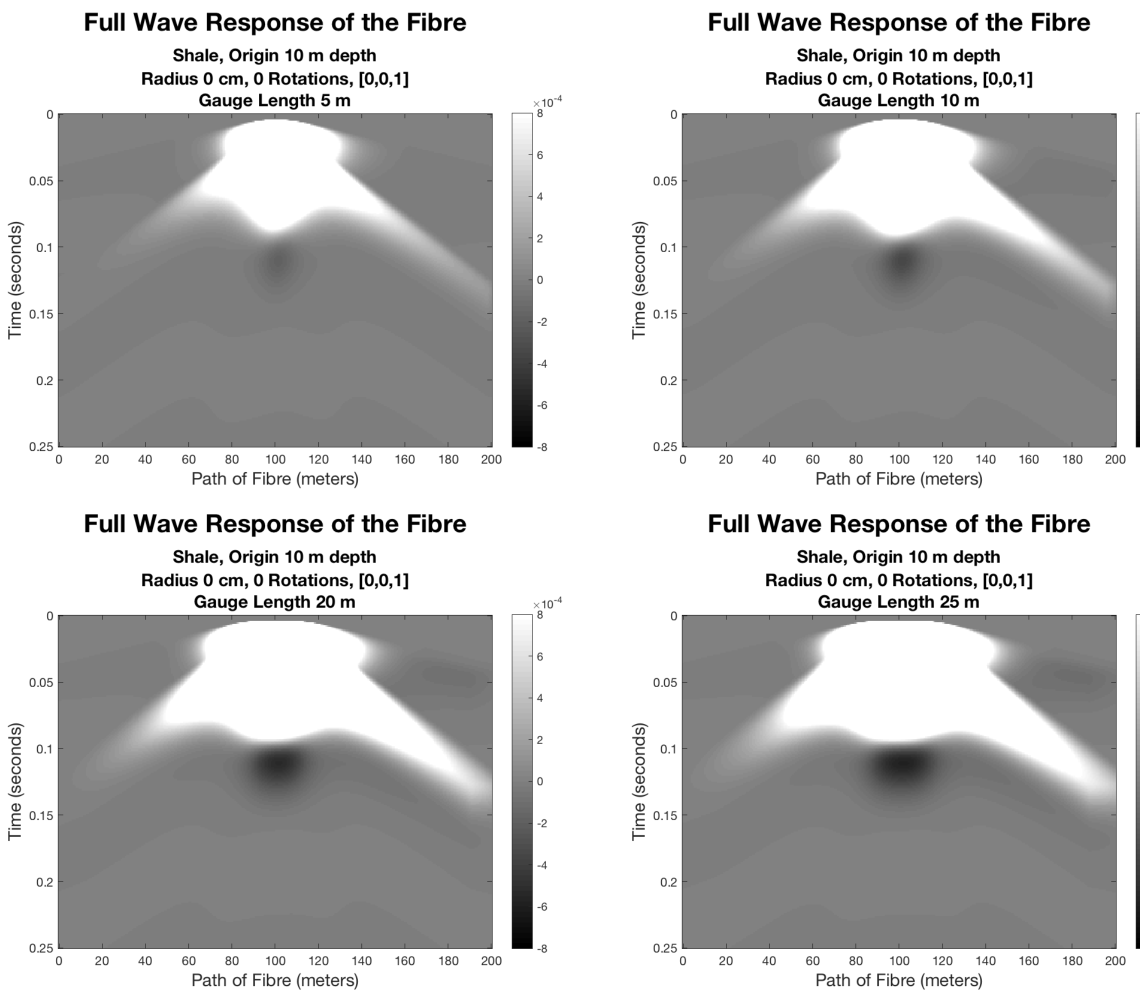


Figure 6.8: The full-wave response of the straight fibre in saturated shale when $\mathbf{A} = [0,0,1]$ for (top left) gauge length 5m, (top right) gauge length 10m, (bottom left) gauge length 20m, and (bottom right) gauge length 25m

Figure 6.8 shows the response of the straight fibre in shale for four different gauge lengths. The image is the sharpest for gauge length 5m with the result for gauge length 10m only slightly less sharp; however, the peak of the hyperbola starts to noticeably flatten for gauge length 20m and 25m.

Recall that the gauge length considers the results of a small portion of the signal at a time, i.e., for our model: 5m, 10m, 20m or 25m. During that portion, gauge length is contracted or stretched depending on the shape of the signal. Since the response of the fibre is hyperbolic, the larger gauge lengths contain a larger portion of the hyperbola. It holds more stretching and contracting information so that the two cancel each other out as seen in the flattening of the hyperbolae in the results for gauge length 20m and 25m.

Figure 6.9 (top) gives a physical representation of what we described in the previous paragraph. The gauge length 5m only contains a minimal amount of increasing and decreasing portion of the signal around the origin. The gauge length 10m contains more decreasing signal and increasing signal around the origin. Both gauge lengths 20m and 25m contain a lot of increasing and decreasing portions of the signal enough to cause a significant flattening of the hyperbola's peak as well as a widening of the signal. We see this occurring in Figure 6.9 (bottom right) for larger intervals around the origin as the gauge length increases, compared to the original signal at t_{10} seen in Figure 6.9 (bottom left). The gauge length also causes the amplitude to increase for all gauge lengths in comparison to the original signal; c.f. Figure 6.9 (bottom row).

Figure 6.9 only describes what occurs for the time step t_{10} of the full-waveform response of the straight fibre. It provides a good visual explanation for the spreading which occurs for the larger gauge lengths in the bottom two images of Figure 6.8.

Now consider what occurs to the helical fibre when we apply four different gauge lengths. Recall for the S-wave we saw some evidence of a helix; however, here in Figure 6.10, we see that helix is less evident as the gauge length is applied to the data.

Figure 6.11 (top) presents a comparison of the gauge lengths on the time step t_{10} of the

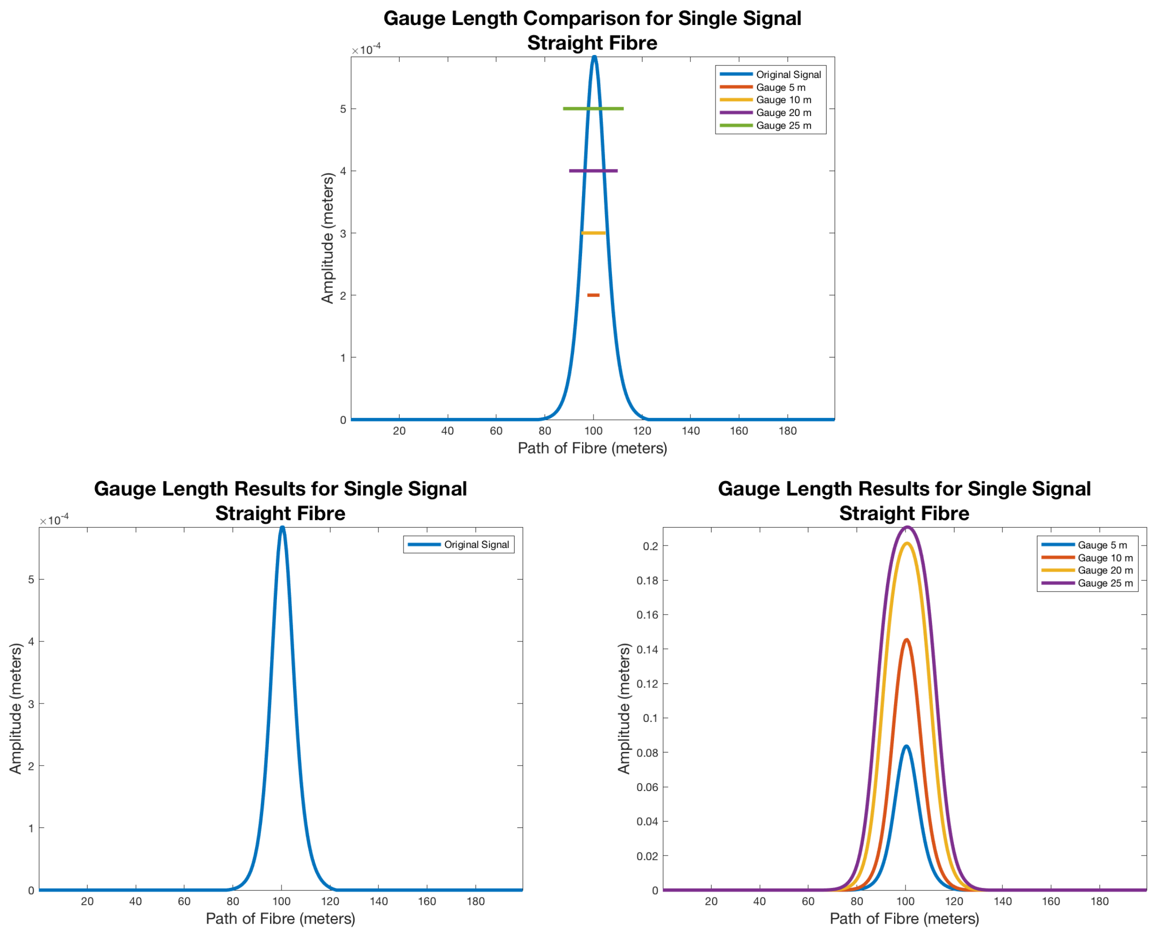


Figure 6.9: (Top) Comparison of the gauge lengths with respect to the full-waveform signal at t_{10} of the straight fibre: (Blue) Original signal at t_{10} (Red) gauge length 5m, (Yellow) gauge length 10m, (Purple) gauge length 20m, and (Green) gauge length 25m.
(Bottom Left) The original signal at t_{10}.
(Bottom Right) Comparison of the gauge lengths applied to the full-waveform signal at t_{10} of the straight fibre: (Blue) gauge length 5m, (Red) gauge length 10m, (Yellow) gauge length 20m, and (Purple) gauge length 25m.

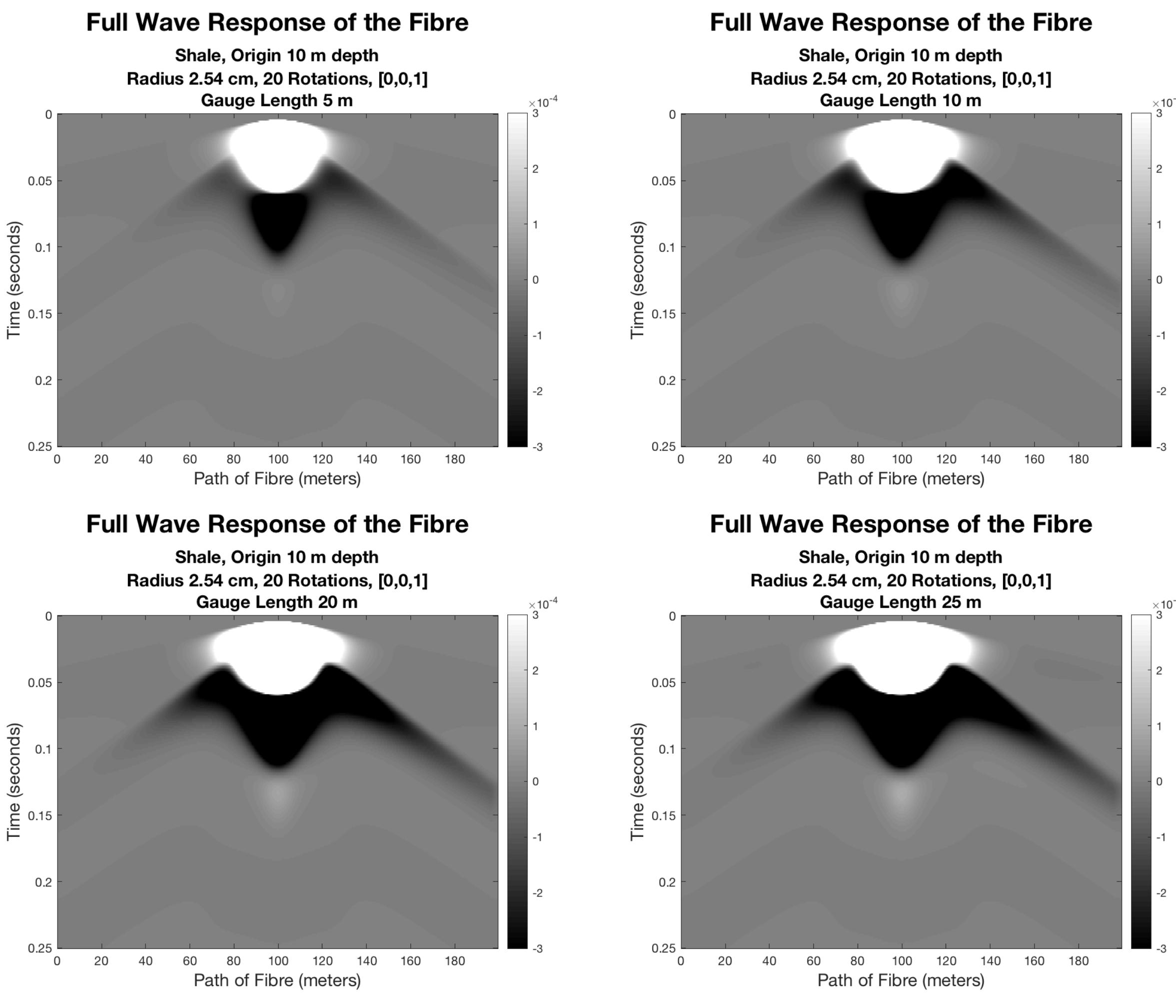


Figure 6.10: The full-wave response of the helical fibre in saturated shale when $\mathbf{A} = [0, 0, 1]$ for (top left) gauge length 5 m, (top right) gauge length 10 m, (bottom left) gauge length 20 m, and (bottom right) gauge length 25 m

full-waveform response of the helical fibre. In Figure 6.11 (bottom left), the original signal contains several oscillations between 80m and 120m. The 5m gauge length contains the fewest oscillations compared to the other gauge lengths. As the gauge length increases, the interval contains more oscillations which means that cancellations are more likely to occur given the presence of more increasing and decreasing data. Similar to the case when the fibre is straight, the larger gauge lengths increase the amplitude of the signal; however, the oscillations are lost.

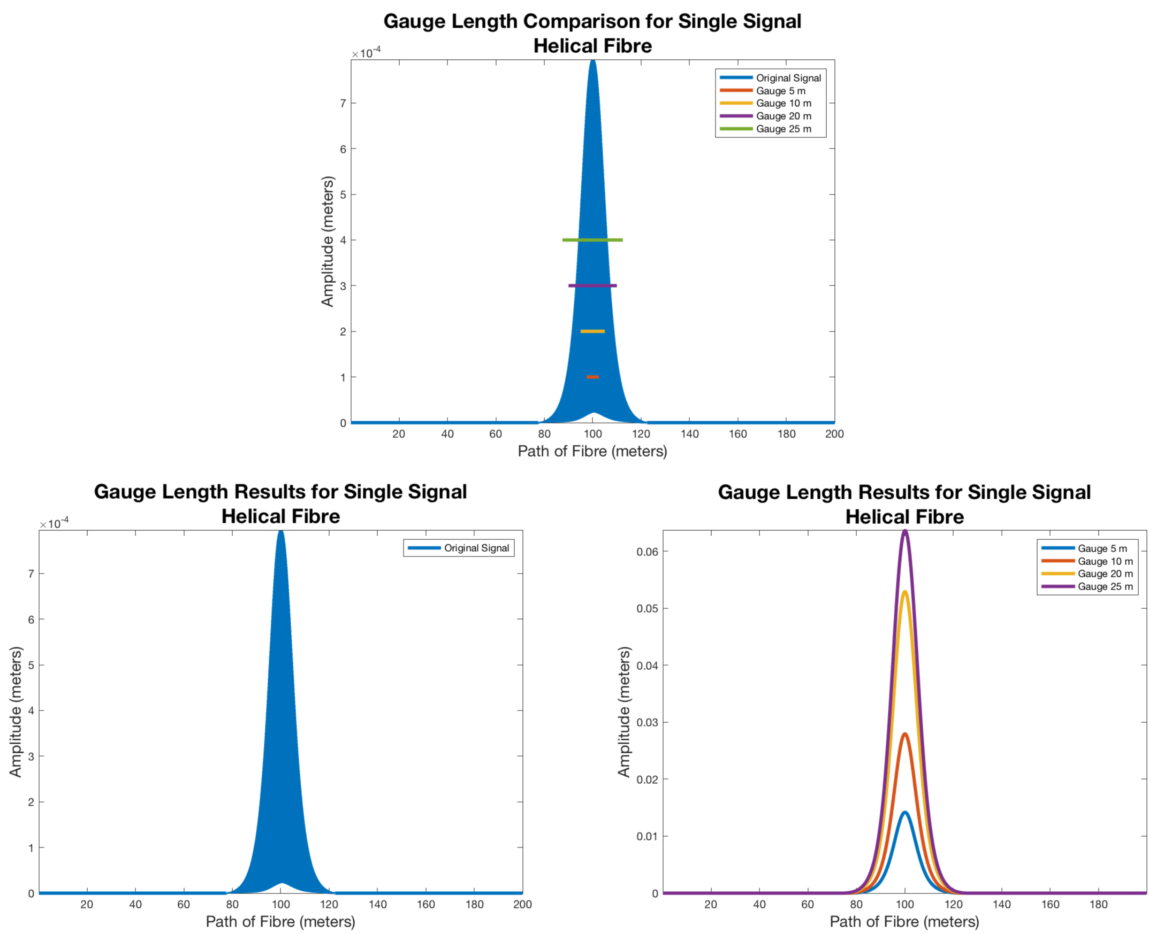


Figure 6.11: (Top) Comparison of the gauge lengths with respect to the full-waveform signal at t_{t0} of the helical fibre: (Blue) Original signal at t_{10} (Red) gauge length 5m, (Yellow) gauge length 10m, (Purple) gauge length 20m, and (Green) gauge length 25m.
(Bottom Left) The original signal at $t_1 0$.
(Bottom Right) Comparison of the gauge lengths applied to the full-waveform signal at t_1 of the helical fibre: (Blue) gauge length 5m, (Red) gauge length 10m, (Yellow) gauge length 20m, and (Purple) gauge length 25m.

6.3 Conclusions

We began with an explanation of the model and found the strain tensors for the P-wave, S-wave, and full-waveform responses. The results of the P-wave response in saturated shale and limestone were shown for straight fibre. Afterwards, the S-wave response in saturated shale was explored for straight and helical fibre. We saw that the helical fibre detects strain in more dimensions than the straight fibre; however, the helical fibre's response has a weaker amplitude in comparison to the straight fibre's response. Finally, we looked at the effects different gauge lengths had on the full-waveform response for straight and helical fibre. For both types of fibre, the larger gauge lengths produced a flattening in the response which was not found in the smaller gauge lengths. We noted that this is largely due to the amount of the signal contained in the larger gauge lengths which resulted in some cancellations. These cancellations also affected how evident the shape of the fibre was in the data as we saw in the helical full-wave response data.

Chapter 7

Investigation of fibre-optic cable formation in DAS acquistion

The flexiblity of fibre-optic cables provides more benefits than simply sensitivity to strain. In fact, this flexiblity allows it to be bent into many different formations in order to increase its sensitivity to strain. A fibre structure of particular interest is the helically wound fibre. Figure 7.1 provides an example of a helically wound fibre-optic cable. We saw in the S-wave example from Chapter 6, Section 6.2.2, that the helical fibre formed a more precise signal even with the gauge length applied than the straight fibre. One issue of the straight fibre is that it cannot detect the wave when it hits the straight fibre perpendicularly. Orienting

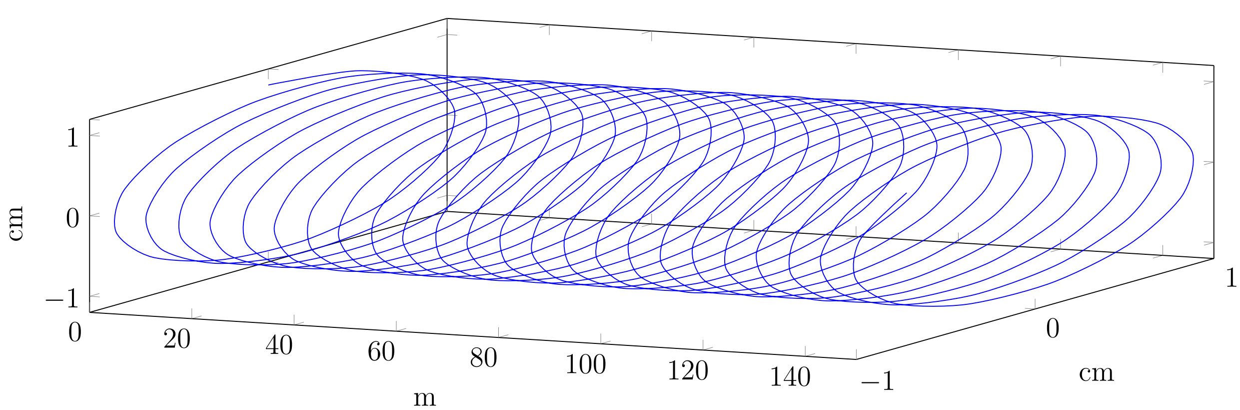


Figure 7.1: An example of a helically wound fibre-optic cable in a DAS sensor.

the fibre helically provides a way for the fibre to detect waves oriented perpendicular to the straight fibre.

In theory, the helically wound fibre should provide more information about the area around the fibre. However, in practice, it appears the helical fibre produces a much smaller signal than the straight fibre. As such, there are many questions as to why the helically wound fibre does not appear to perform as well as the straight fibre in real world applications. In this chapter, we attempt to offer insight into the differences between the signals in the straight and helically wound fibre and answer some of these questions.

In this chapter, we produce bounds for the amplitude response of the fibre for various configuration whether a gauge length is applied or not. Afterwards, we prove that the amplitude response of the helical fibre is connected via a homotopy to the amplitude response of the straight fibre. Employing this fact, we use the homotopy to model DAS data in order to compare the helically wound fibre and the straight fibre. The first model compares the results of a 2.54cm radius helical fibre laid horizontally along the earth's surface deforming into the straight fibre. We then use the homotopy to model helical fibre, laid vertically into the earth's surface 200 meters, deforming into straight fibre.

We found the bounds for the amplitude $A(s,t)$ of the data to be

$$\Lambda_{\min} \leq A(s,t) \leq \Lambda_{\max} \tag{7.1}$$

where $\Lambda_{\min}$ is the minimum eigenvalue for all eigenvalues $\lambda_{\min}$ for the strain matrix $\epsilon(\mathbf{p}(s),t)$ and $\Lambda_{\max}$ is the maximum eigenvalue for all eigenvalues $\lambda_{\max}$ for the strain matrix $\epsilon(\mathbf{p}(s),t)$. The bounds for the amplitude $(A*g)(s,t)$ of the data with the gauge length applied is

$$\gamma\Lambda_{\min} \leq (A*g)(s,t) \leq \gamma\Lambda_{\max} \tag{7.2}$$

where γ is the distance of the gauge length and g is the gauge length function

$$g(\tau) = \begin{cases} 1 & -\frac{\gamma}{2} \leq \tau \leq \frac{\gamma}{2}; \\ 0 & \text{otherwise.} \end{cases} \tag{7.3}$$

We also show that a homotopy exists between the amplitude response of the helically wound fibre and the amplitude response of the straight fibre. This means that the helically wound fibre can be deformed into the straight fibre. The homotopy allows us to compare the responses of the straight and helical fibre.

We used the homotopy to model helical fibre in two examples. The first example shows the results of a 2.54cm radius helix fibre, laid horizontally along the ground, deforming into the straight fibre. We consider the result at four locations and see that the L^2-norm of the data when the gauge length is applied increases across the homotopy from the helical fibre to the straight fibre; however, the norms of the data, without the gauge length applied, as well as the traces, determinants, and eigenvalues of the strain matrices decreases from the helical fibre to the straight fibre across the homotopy.

In the second example, we employ a homotopy to compare three formations of the fibre: a helix around a well-borehole with a 17.78cm radius; a helix beside a well-borehole with radius 2.54cm; and a straight fibre beside the well-borehole. In this case, we saw that all the norms decrease from the borehole helical fibre to the straight fibre. This information suggests that the results of the fibre formation are highly dependent on the location of the source with respect to the fibre.

7.1 Bounds on the amplitude of DAS data

In Chapter 6, we saw several examples where the helically wound fibre produced an amplitude response which was significantly smaller than the amplitude response of the straight fibre. The question arises as to whether this holds for all helical fibre or just the examples we saw earlier. Instead of using brute force to answer this question, we provide bounds on the amplitude, $A(s,t)$, and the amplitude of the data when a gauge length is applied, $(A*g)(s,t)$. These bounds provide us with the minimum and maximum amplitude for any choice of helical radius and number of rotations per meter.

In order to find these bounds, we need to use some results from linear algebra. We start by considering some definitions.

Definition 7.1. Let A be a $d \times d$ symmetric matrix. The matrix A is called *positive semi-definite* if all its eigenvalues are non-negative. This is denoted as $A \succeq 0$, where 0 denotes the $d \times d$ zero matrix. A positive definite matrix A satisfies the condition

$$A \succeq 0 \iff x^\top A x \geq 0 \quad \forall x \in \mathbb{R}^d. \tag{7.4}$$

We also need the following ordering for matrices.

Definition 7.2. For any two symmetric matrices A and B, the *Löwner ordering*, or *postive semi-definite ordering*, is defined as $A \succeq B$ if $A - B \succeq 0$.

Now we are prepared to prove the following lemmas in order to find bounds for the amplitude response of the fibre.

Lemma 7.3. *Let A be a symmetric matrix. Let $\lambda_{\min}$ and $\lambda_{\max}$ respectively denote the smallest and largest eigenvalues of A. Then*

$$\lambda_{\min} \cdot I \preceq A \preceq \lambda_{\max} I. \tag{7.5}$$

Proof. Let $\lambda_1, ..., \lambda_n$ be all the eigenvalues of A. The eigenvalues of $\lambda_{\min} \cdot I - A$ are $\{\lambda_{\min} - \lambda_1, ..., \lambda_{\min} - \lambda_n\}$.The maximum of these eigenvalues is

$$\max_j(\lambda_{\min} - \lambda_j) = \lambda_{\min} - \min_j \lambda_j = 0. \tag{7.6}$$

Thus, $\lambda_{\min} \cdot I - A \leq 0$ which implies the left half of Equation 7.5. Similarly, the eigenvalues of $\lambda_{\max} \cdot I - A$ are $\{\lambda_{\max} - \lambda_1, ..., \lambda_{\max} - \lambda_n\}$. Then, the minimum of these eigenvalues is

$$\min_j(\lambda_{\max}) = \lambda_{\max} - \max_j \lambda_j = 0. \tag{7.7}$$

Hence, $\lambda_{\max} \cdot I - A \geq 0$ which implies the right half of Equation 7.5. $\square$

The strain tensor ϵ is symmetric; and hence via application of Lemma 7.3, ϵ is bounded by,

$$\lambda_{\min} \cdot I \preceq \epsilon \preceq \lambda_{\max} \cdot I. \tag{7.8}$$

We also get the following lemma from linear algebra.

Lemma 7.4. *For symmetric A and B, $A \succeq B$ if and only if $v^\top Av \geq v^\top Bv$ for all vectors v.*

Proof. Let $A \succeq B$. Then, $A - B \succeq 0$. By definition of positive semi-definite, for all $v \in \mathbb{R}^d$

$$\begin{aligned} A - B \succeq 0 &\iff v^\top(A-B)v \geq 0 && (7.9)\\ &\iff v^\top Av - v^\top Bv \geq 0 && (7.10)\\ &\iff v^\top Av \geq v^\top Bv. && (7.11) \end{aligned}$$

Therefore, we have $A \succeq B$ if and only if $v^\top Av \geq v^\top Bv$. $\square$

Since we know that Equation 7.8 holds, then, by Lemma 7.4,

$$\begin{aligned} &\mathbf{T_p}^\top \lambda_{\min} \cdot I\mathbf{T_p} \leq \mathbf{T_p}^\top \epsilon \mathbf{T_p} \leq \mathbf{T_p}^\top \lambda_{\max} \cdot I\mathbf{T_p} && (7.12)\\ \implies &\mathbf{T_p}^\top \lambda_{\min} \cdot I\mathbf{T_p} \leq A(s,t) \leq \mathbf{T_p}^\top \lambda_{\max} \cdot I\mathbf{T_p} && (7.13)\\ \implies &\lambda_{\min} \left\|\mathbf{T_p}\right\|_2^2 \leq A(s,t) \leq \lambda_{\max} \left\|\mathbf{T_p}\right\|_2^2. && (7.14) \end{aligned}$$

Recall that since $\mathbf{T_p}$ is the unit tangent path of the fibre, $\|\mathbf{T_p}\|_2^2 = 1$. Hence, the bounds for the amplitude $A(s,t)$ will be

$$\lambda_{\min}(s,t) \leq A(s,t) \leq \lambda_{\max}(s,t). \tag{7.15}$$

The eigenvalues $\lambda_{\min}(s,t)$ and $\lambda_{\max}(s,t)$ are the minimum and maximum eigenvalues of the strain matrix $\epsilon(\mathbf{p}(s),t)$ at the point (s,t) on the fibre path $\mathbf{p}$, respectively. Let

$$\Lambda_{\min} = \min_{\substack{s\in\mathbb{R};\\ t\in[0,\infty)}} \lambda_{\min}(s,t) \tag{7.16}$$

and

$$\Lambda_{\max} = \max_{\substack{s\in\mathbb{R};\\ t\in[0,\infty)}} \lambda_{\max}(s,t). \tag{7.17}$$

We then have

$$\Lambda_{\min} \leq \lambda_{\min} \leq A(s,t) \leq \lambda_{\max} \leq \Lambda_{\max}, \tag{7.18}$$

from whence we derive

$$\Lambda_{\min} \leq A(s,t) \leq \Lambda_{\max}. \tag{7.19}$$

When applying the gauge length to the waveform response, we convolve amplitude of the waveform $A(s,t)$ with the gauge length function

$$g(\tau) = \begin{cases} 1 & -\frac{\gamma}{2} \leq \tau \leq \frac{\gamma}{2}; \\ 0 & \text{otherwise}, \end{cases} \tag{7.20}$$

where γ is the distance of gauge length. We calculate the data with the gauge length applied via the convolution

$$(A * g)(s,t) = \int_{-\infty}^{\infty} A(\tau,t)g(s-\tau)d\tau \tag{7.21}$$

$$= \int_{s-\frac{\gamma}{2}}^{s+\frac{\gamma}{2}} A(\tau,t)d\tau. \tag{7.22}$$

Recall from calculus that the following inequality holds for integrals:

Lemma 7.5. *If $m \leq f(x) \leq M$ for $a \leq x \leq b$ then*

$$m(b-a) \leq \int_a^b f(x)dx \leq M(b-a). \tag{7.23}$$

Using Lemma 7.5, we then find that the data with the gauge length applied $A_G(s,t)$ is bounded between

$$\int_{s-\frac{\gamma}{2}}^{s+\frac{\gamma}{2}} \Lambda_{\text{min}} d\tau \leq \int_{s-\frac{\gamma}{2}}^{s+\frac{\gamma}{2}} A(\tau,t)d\tau \leq \int_{s-\frac{\gamma}{2}}^{s+\frac{\gamma}{2}} \Lambda_{\text{max}} d\tau \tag{7.24}$$

$$\implies \Lambda_{\text{min}} \int_{s-\frac{\gamma}{2}}^{s+\frac{\gamma}{2}} d\tau \leq (A * g)(s,t) \leq \Lambda_{\text{max}} \int_{s-\frac{\gamma}{2}}^{s+\frac{\gamma}{2}} d\tau \tag{7.25}$$

$$\implies \Lambda_{\text{min}} \left(s + \frac{\gamma}{2} - \left(s - \frac{\gamma}{2} \right) \right) \leq (A * g)(s,t) \leq \Lambda_{\text{max}} \left(s + \frac{\gamma}{2} - \left(s - \frac{\gamma}{2} \right) \right) \tag{7.26}$$

$$\implies \gamma \Lambda_{\text{min}} \leq (A * g)(s,t) \leq \gamma \Lambda_{\text{max}}. \tag{7.27}$$

With these arguments, we have proved the following results with regards to the bounds on the amplitude response of the data both when a gauge length is applied to the data, and when it is not applied to the data.

Theorem 7.6. *The amplitude response $A(s,t)$ of the data without a gauge length applied satisfies the bound*

$$\Lambda_{\text{min}} \leq A(s,t) \leq \Lambda_{\text{max}}, \tag{7.28}$$

for

$$\Lambda_{\text{min}} = \min_{\substack{s \in \mathbb{R}; \\ t \in [0,\infty)}} \lambda_{\text{min}}(s,t) \tag{7.29}$$

and

$$\Lambda_{\text{max}} = \max_{\substack{s \in \mathbb{R}; \\ t \in [0,\infty)}} \lambda_{\text{max}}(s,t), \tag{7.30}$$

where λ_{min} is the minimum eigenvalue of the strain matrix ϵ at (s,t) and λ_{max} is the maximum *eigenvalue of the strain matrix* ϵ *at* (s, t)*.*

Theorem 7.7. *Define the gauge length function* g *as*

$$g(\tau) = \begin{cases} 1 & -\frac{\gamma}{2} \leq \tau \leq \frac{\gamma}{2}; \\ 0 & \textit{otherwise}, \end{cases} \tag{7.31}$$

and γ *is the distance of the gauge length. The amplitude response of the data with a gauge length applied,* $(A * g)(s, t)$*, satisfies the bound*

$$\gamma \Lambda_{\min} \leq (A * g)(s, t) \leq \gamma \Lambda_{\max}, \tag{7.32}$$

where

$$\Lambda_{\min} = \min_{\substack{s \in \mathbb{R}; \\ t \in [0, \infty)}} \lambda_{\min}(s, t) \tag{7.33}$$

and

$$\Lambda_{\max} = \max_{\substack{s \in \mathbb{R}; \\ t \in [0, \infty)}} \lambda_{\max}(s, t), \tag{7.34}$$

where $\lambda_{\min}$ *is the minimum eigenvalue of the strain matrix* ϵ *at* (s, t) *and* $\lambda_{\max}$ *is the maximum eigenvalue of the strain matrix* ϵ *at* (s, t)*.*

7.2 Comparison of straight and helically wound fibre

One way to compare the straight and helically wound fibre directly is by using a homotopy.

Definition 7.8. Let X, Y be topological spaces, and $f, g : X \longrightarrow Y$ continuous maps. A *homotopy* from f to g is a continuous function $F : X \times [0, 1] \longrightarrow Y$ satisfying

$$F(x, 0) = f(x) \quad \text{and} \quad F(x, 1) = g(x), \tag{7.35}$$

for all $x \in X$. If a homotopy exists, we say that f is homotopic to g and write $f \simeq g$.

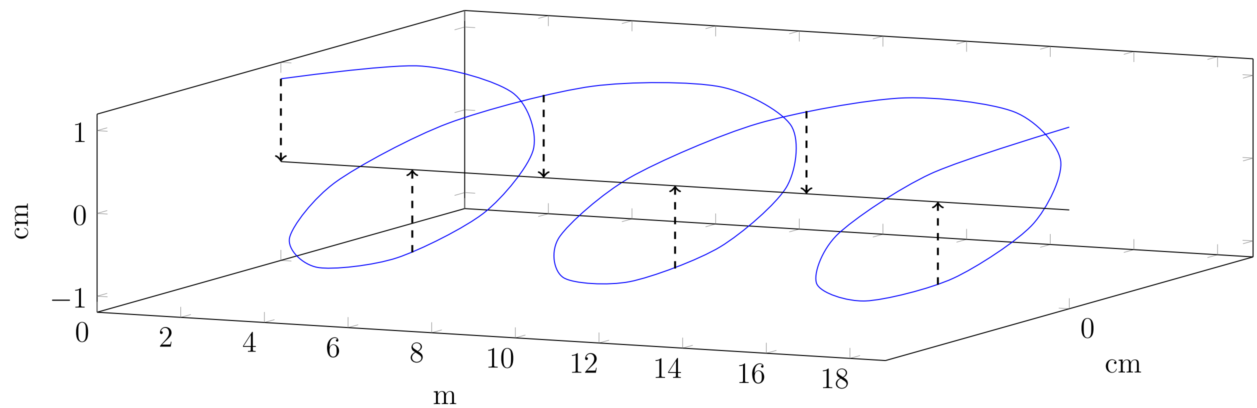


Figure 7.2: A helix deforming to a straight line.

Figure 7.2 gives a visual idea of what the homotopy does with regards to the fibre-optic cable. As we move over the homotopy, the shape of the helix deforms along the black arrows in the figure to the straight fibre. In Figure 7.3, a 2π cross-section of a helical fibre with radius h_r is considered. For certain values of $w \in [0, 1]$, the radius of the helix shrinks. Figure 7.3 shows what occurs given four values of w: 0, w_1, w_2, and 1.

We show that the helically wound fibre deforms into the straight fibre using the following homotopy. We assume the path of the fibre is always smooth, since we cannot use an interrogator on a broken fibre-optic cable.

Lemma 7.9. *Let $p_{\text{str}} : \mathbb{R} \longrightarrow \mathbb{R}^3$ and $p_{\text{hel}} : \mathbb{R} \longrightarrow \mathbb{R}^3$ represent the smooth functions for the straight and helical fibre respectively. The linear function $F : \mathbb{R} \times [0, 1] \longrightarrow \mathbb{R}^3$*

$$F(s, w) = (1 - w)p_{\text{hel}}(s) + wp_{\text{str}}(s) \tag{7.36}$$

is a homotopy which deforms the helical fibre into the straight fibre.

Proof. Let p_{str} and p_{hel} represent the functions for the straight and helical fibre respectively. Both functions map from $\mathbb{R} \longrightarrow \mathbb{R}^3$. Consider the w-linear function $F : \mathbb{R} \times [0, 1] \longrightarrow \mathbb{R}^3$

$$F(s, w) = (1 - w)p_{\text{hel}}(s) + wp_{\text{str}}(s). \tag{7.37}$$

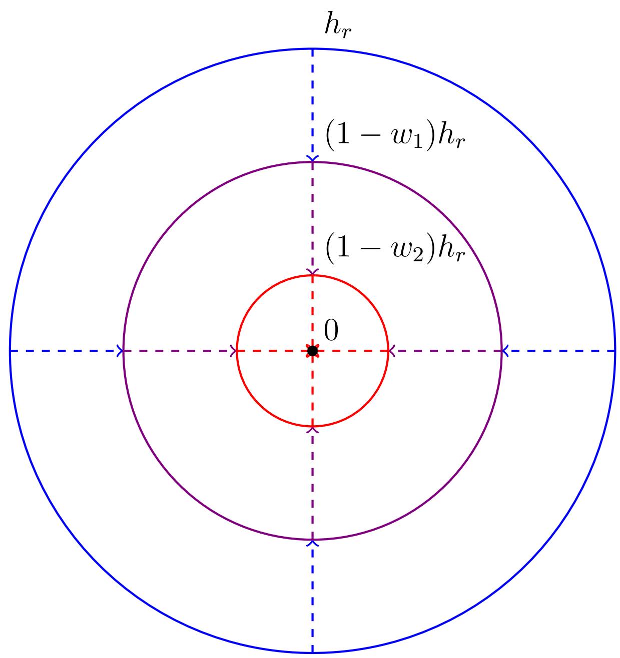


Figure 7.3: A 2π cross-section of a fibre-optic cable with helical radius h_r deforming into a straight fibre over four choices of w in the homotopy: 0 (blue), w_1 (purple), w_2 (red), and 1 (black).

Since F is linear w, then F is continuous in w. The functions p_{str} and p_{hel} are smooth; therefore, p_{str} and p_{hel} are continuous in s and it follows that $F : \mathbb{R} \times [0, 1] \longrightarrow \mathbb{R}^3$ is continuous in both s and w. At $w = 0$, we have

$$F(s, 0) = 1 \cdot p_{\text{hel}}(s) + 0 \cdot p_{\text{str}}(s) = p_{\text{hel}}(s) \tag{7.38}$$

and at $w = 1$, we find

$$F(s, 1) = 0 \cdot p_{\text{hel}}(s) + 1 \cdot p_{\text{str}}(s) = p_{\text{str}}(s). \tag{7.39}$$

Hence, $p_{\text{str}} \simeq p_{\text{hel}}$. $\square$

Now that we have a homotopy between the paths of straight and helical fibre, we can use it to find a homotopy between the strain tensor of the straight fibre and the strain tensor of the helical fibre. Before we continue however, we should discuss some issues. The function

F works as a homotopy for our model because the model does not assume there is anything supporting the helical structure of the fibre. Therefore, we do not need to worry about any objects impeding the deformation of the helical fibre into the straight fibre. In reality, some material (foam, aluminum, etc.) is needed to preserve the fibre's helical shape. Whatever the material, it would cause an obstruction in our ability to deform the helical fibre into straight fibre. In terms of homotopy theory, we could consider the material employed to keep the helical shape of the fibre to be a hole, then let the hole shrink to zero.

For now, producing a homotopy between the straight strain component and the helical strain component the way our model does provides stepping stones for extending to harder cases, i.e. when materials preserve the helix. Recall the following theorem from algebraic topology (Rotman, 1988).

Theorem 7.10. *If $f_1, g_1 : X \longrightarrow Y$ are homotopic,and $f_2, g_2 : Y \longrightarrow Z$ are homotopic, then their compositions $f_2 \circ f_1$ and $g_2 \circ g_1 : X \longrightarrow Z$ are also homotopic.*

Using a composition of the fibre path $\mathbf{p}$ with the strain function ϵ, we can prove that a homotopy exists between the strain of the helically wound fibre and the strain of the straight fibre. Let $U \subseteq \mathbb{R}^3$ and $I \subseteq [0, \infty)$. Recall that the strain function $\epsilon : U \times I \to \mathbb{R}^3$ is used to define the strain tensor of the straight fibre and the helical fibre. In particular, the function ϵ is the same for both straight and helical fibre; it is the input p_{str} and p_{hel} that give the structure of the fibre for which ϵ gives the strain.

Lemma 7.11. *Let $U \subseteq \mathbb{R}^3$ and $I \subseteq [0, \infty)$. Let $\epsilon : U \times I \to \mathbb{R}^3$ be the strain function which produces the strain matrix for the fibre at the point $(\mathbf{p}(s), t)$ for $s \in \mathbb{R}$ on the fibre path $\mathbf{p}$ and for $t \in I$. The strain of the helically wound fibre is homotopic to the strain of the straight fibre.*

Proof. Let $\mathbf{p}_{\text{str}}$ be the path of the straight fibre and $\mathbf{p}_{\text{hel}}$ be the path of the helical fibre. In terms of Theorem 7.10, we have $f_1 = \mathbf{p}_{\text{str}}$ and $g_1 = \mathbf{p}_{\text{hel}}$ and $f_2 = \epsilon$ and $g_2 = \epsilon$. By Lemma 7.9,

$\mathbf{p}_{str} \simeq \mathbf{p}_{hel}$; given that $f_2 = g_2$, then they are homotopic by the constant function $F(x, w) = \epsilon$ for all $w \in [0, 1]$. Hence, by Theorem 7.10, we have that $\epsilon(\mathbf{p}_{\mathrm{str}}, t) \simeq \epsilon(\mathbf{p}_{\mathrm{hel}}, t)$. □

Recall from Chapter 6 that $A(s,t) = \mathbf{T_p}^{\top}(s)\epsilon(\mathbf{p}(s), t)\,\mathbf{T_p}(s)$ is the amplitude response of the fibre where $\mathbf{T_p}$ is the tangent path of the fibre. Given that the fibre path $\mathbf{p}$ is smooth, $\mathbf{T_p}$ exists for all (s, t). Now, we use a similar argument to Lemma 7.9 to prove the tangent path of the helical fibre is homotopic to the tangent path of the straight fibre.

Lemma 7.12. *Let $\mathbf{p}_{\mathrm{str}} : \mathbb{R} \to \mathbb{R}^3$ and $\mathbf{p}_{\mathrm{hel}} : \mathbb{R} \to \mathbb{R}^3$ represent the smooth functions for the straight and helically wound fibre, respectively. Let $\mathbf{T_p} : \mathbb{R} \to \mathbb{R}^3$ represent the unit tangent path of the fibre path $\mathbf{p}$. The w-linear function $G : \mathbb{R} \times [0, 1] \to \mathbb{R}^3$*

$$G(s, w) = (1 - w)\,\mathbf{T}_{\mathbf{p}_{\mathrm{hel}}}(s) + w\,\mathbf{T}_{\mathbf{p}_{\mathrm{str}}}(s) \tag{7.40}$$

is a homotopy which deforms the unit tangent path of the helical fibre into the unit tangent path of the straight fibre, i.e. $\mathbf{T}_{\mathbf{p}_{\mathrm{hel}}} \simeq \mathbf{T}_{\mathbf{p}_{\mathrm{str}}}$.

Proof. Let $\mathbf{p}_{\mathrm{str}} : \mathbb{R} \to \mathbb{R}^3$ and $\mathbf{p}_{\mathrm{hel}} : \mathbb{R} \to \mathbb{R}^3$ represent the smooth functions for the straight and helically wound fibre, respectively. Consider the w-linear function $G : \mathbb{R} \times [0, 1] \to \mathbb{R}^3$

$$G(s, w) = (1 - w)\,\mathbf{T}_{\mathbf{p}_{\mathrm{hel}}}(s) + w\,\mathbf{T}_{\mathbf{p}_{\mathrm{str}}}(s). \tag{7.41}$$

Since G is linear in w, then G is continuous in w. The functions $\mathbf{T}_{\mathbf{p}_{\mathrm{str}}}$ and $\mathbf{T}_{\mathbf{p}_{\mathrm{hel}}}$ are continuous since $\mathbf{p}_{\mathrm{str}}$ and p_{hel} are both smooth and $\mathbf{p}'(s) \neq [0, 0, 0]$ for the path of both the straight and helical fibre by the orientation of the fibre-optic cable. Thus, G is continuous in both s and w. At $w = 0$, we have

$$G(s, 0) = 1 \cdot \mathbf{T}_{\mathbf{p}_{\mathrm{hel}}}(s) + 0 \cdot \mathbf{T}_{\mathbf{p}_{\mathrm{str}}}(s) = \mathbf{T}_{\mathbf{p}_{\mathrm{hel}}}(s), \tag{7.42}$$

and at w = 1,

$$G(s, 1) = 0 \cdot \mathbf{T}_{\mathbf{p}_{\mathrm{hel}}}(s) + 1 \cdot \mathbf{T}_{\mathbf{p}_{\mathrm{str}}}(s) = \mathbf{T}_{\mathbf{p}_{\mathrm{str}}}(s). \tag{7.43}$$

Hence, $\mathbf{T}_{\mathbf{p}_{\mathrm{hel}}} \simeq \mathbf{T}_{\mathbf{p}_{\mathrm{str}}}$. $\square$

By a similar argument, we get that $\mathbf{T}_{\mathbf{p}_{\mathrm{hel}}}^{\top} \simeq \mathbf{T}_{\mathbf{p}_{\mathrm{str}}}^{\top}$. Now, we have the building blocks to show that the amplitude response of the helical fibre is homotopic to the amplitude response of the straight fibre.

Theorem 7.13. *Define the amplitude response of the fibre* $A : \mathbb{R} \times [0, \infty) \to \mathbb{R}$ *as*

$$A(s,t) = \mathbf{T}_{\mathbf{p}}^{\top}(s)\epsilon(\mathbf{p}(s),t)\,\mathbf{T}_{\mathbf{p}}(s) \tag{7.44}$$

where $\mathbf{T}_{\mathbf{p}}$ *is the unit tangent path of the fibre path* $\mathbf{p}$ *and* ϵ *is the strain function of the fibre. The amplitude response of the helical fibre,* A_{hel}*, is homotopic to the amplitude response of the straight fibre,* A_{str}*, i.e.* $A_{\mathrm{hel}} \simeq A_{\mathrm{str}}$.

Proof. The inner product is a continuous function which means that

$$A =< \mathbf{T}_{\mathbf{p}}(s), \epsilon(\mathbf{p}(s),t), \mathbf{T}_{\mathbf{p}}(s) > \tag{7.45}$$

is continuous. Since homotopies are preserved under composition by Theorem 7.10 and the product of continuous functions is continuous, then $A_{\mathrm{hel}} \simeq A_{\mathrm{str}}$. $\square$

Typically an interrogator applies a gauge length before it outputs the data. As such, we show that the amplitude response of the helical fibre when the gauge length is applied is homotopic to the amplitude response of the straight fibre when the gauge length is applied.

Theorem 7.14. *Define the amplitude response of the fibre* $A : \mathbb{R} \times [0, \infty) \to \mathbb{R}$ *as*

$$A(s,t) = \mathbf{T}_{\mathbf{p}}^{\top}(s)\epsilon(\mathbf{p}(s),t)\,\mathbf{T}_{\mathbf{p}}(s) \tag{7.46}$$

where $\mathbf{T}_{\mathbf{p}}$ *is the unit tangent path of the fibre path* $\mathbf{p}$ *and* ϵ *is the strain function of the fibre.*

Let $(A * g)(s,t)$ be the amplitude response of the fibre when the gauge length is applied where

$$g(\tau) = \begin{cases} 1 & -\frac{\gamma}{2} \leq \tau \leq \frac{\gamma}{2}; \\ 0 & \textit{otherwise}, \end{cases} \tag{7.47}$$

is the gauge length function and γ is the distance of the gauge length. Then the amplitude response of the helical fibre when the gauge length is applied is homotopic to the amplitude response of the straight fibre when the gauge length is applied, i.e. $(A_{\text{hel}} * g)(s,t) \simeq (A_{\text{str}} * g)(s,t)$.

Proof. The w-terms from the homotopies are constant in τ. Therefore, they act as constant coefficients of the terms in the homotopy that are dependent on τ. When these τ dependent terms are integrated, the w-terms remain. As such, at $w = 0$, we have the integrated term for the helical fibre's response and at $w = 1$, we have the integrated term of the straight fibre's response. By Theorem 7.13, we know that $A_{\text{hel}}(s,t)$ and $A_{\text{str}}(s,t)$ are continuous. Moreover, by the Fundamental Theorem of Calculus, $(A_{\text{hel}} * g)(s,t)$ and $(A_{\text{str}} * g)(s,t)$ are continuous. Thus, $(A_{\text{hel}} * g)(s,t) \simeq (A_{\text{str}} * g)(s,t)$, as was to be shown. □

Now that we have shown that a homotopy exists between the amplitude responses of the helical and straight fibres, we can now use the homotopy to compare the two formations of fibre. We will specifically compare the norms of the data over the homotopy in the following section of two examples.

7.2.1 Homotopy model

With a homotopy between the amplitude response of the straight fibre and the amplitude response of the helical fibre, we create a model to study what occurs as the fibre deforms from helically wound to straight. Consider the following fibre configuration:

$$\mathbf{p}(s) = [s; 100 + h_r \cos(2\pi h_n s); -0.5 + h_r \sin(2\pi h_n s)] \tag{7.48}$$

where h_r is the helical radius and h_n is the number of rotations per meter. We find the equation for the straight fibre $\mathbf{p}_{\text{str}}(s)$ by setting the helical radius $h_r = 0$ and the number of rotations $h_n = 0$, which gives

$$\mathbf{p}_{\text{str}}(s) = [s; 100; -0.5] \tag{7.49}$$

and

$$\mathbf{p}_{\text{hel}}(s) = [s; 100 + h_r \cos(2\pi h_n s); -0.5 + h_r \sin(2\pi h_n s)] \tag{7.50}$$

where we choose $h_r = 2.54$cm and $h_n = 10$ rotations. Therefore, the homotopy F becomes

$$F(s, w) = (1 - w)\mathbf{p}_{\text{hel}}(s) + w\mathbf{p}_{\text{str}}(s) \tag{7.51}$$

$$= (1 - w)[s; 100 + h_r \cos(2\pi h_n s); -0.5 + h_r \sin(2\pi h_n s)] + w[s; 100; -0.5] \tag{7.52}$$

$$= [(1 - w)s + ws; (1 - w)(100 + h_r \cos(2\pi h_n s)) + w100; \tag{7.53}$$

$$(1 - w)(-0.5 + h_r \sin(2\pi h_n s)) + w(-0.5)] \tag{7.54}$$

$$= [s; 100 + (1 - w)h_r \cos(2\pi h_n s); -0.5 + (1 - w)h_r \sin(2\pi h_n s)]. \tag{7.55}$$

The w from the homotopy only affects the sine and cosine from the helical fibre. This follows given that the homotopy deforms helix into a straight line.

We use $F(s, w)$ to represent the fibre path in our model. The values of w vary in order to show how $A(s, t)$ deforms over w from the helical $A_{\text{hel}}(s, t)$ to the straight $A_{\text{str}}(s, t)$. Specifically, we choose

$$w \in \left\{0, \frac{1}{7}, \frac{1}{4}, \frac{1}{2}, \frac{3}{5}, \frac{3}{4}, 1\right\}. \tag{7.56}$$

To give examples outside multiples of 1/4, we pick values of $w = 1/7$ and $w = 3/5$. We position the source at four different locations on the fibre to study how the position of the source affects the response of the fibre. We choose the following four locations with respect to the fibre in our model:

1. Location 1 - [101,105,0];

2. Location 2 - [106,105,0];

3. Location 3 - [111,105,0];

4. Location 4 - [100,111,0].

Let us begin by considering the results for all w when the source is at Location 1. Figures 7.4 – 7.7 show the fibre deforming from helical fibre of radius 2.54cm for 10 rotations per meter of fibre to straight fibre. We plug the different values of w into the homotopy F from Equation 7.55 and use it for the path of the fibre. The left column of each set of figures shows the data without the gauge length applied and the right column depicts the data with the 10 meter gauge length applied.

Figure 7.4 shows the 2.54cm helical fibre at the top and the 2.54cm radius diminished by 1/7th of the radius on the bottom row. The data without the gauge length applied as seen on the left shows some evidence of the helical fibre given the presence of oscillations in the signal, whereas the version of the data with the gauge length applied on the right does not have these oscillations. At least for this size of helix and gauge length, the gauge length causes the parts of the data which highlight the helix to disappear. Moreover, the signal of the data with the gauge length applied looks brighter and larger because the gauged data applies a sampling rate of $\Delta x = 2/3$m which is typical for real data. The data without a gauge length applied has a sampling rate of $\Delta x = 1/3$m because we need at least three points per rotation in order to realize the helix.

In Figure 7.5, the helical shape of the fibre becomes much more prominent in the data without a gauge length applied on the left as the radius decreases; however, applying the gauge length causes the helix to disappear from the data. The prominence of the helix may be due to the number of sample points employed to create the helical data set. The smaller helix benefits from the fewer sample points, while the larger helix did not. It may also be due to how close the source is to the fibre.

Focusing now on the data with the 10 meter gauge length applied, notice that the helical

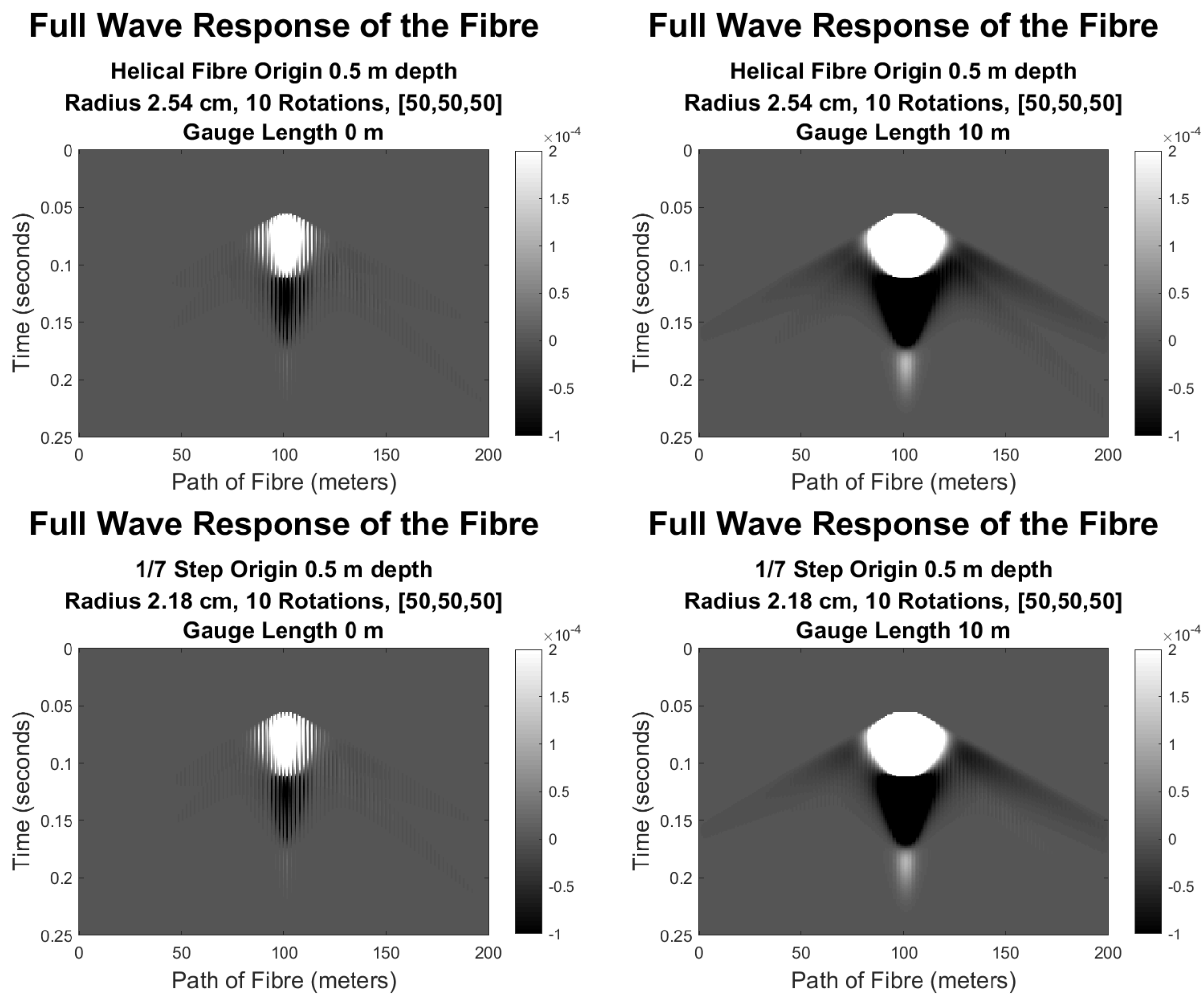


Figure 7.4: (Left column) The data without the gauge length applied (top) helical fibre and (bottom) helical fibre diminished by 1/7-th of the radius when the source is at Location 1. (Right column) The 10 m gauge length applied to data seen in the left column.

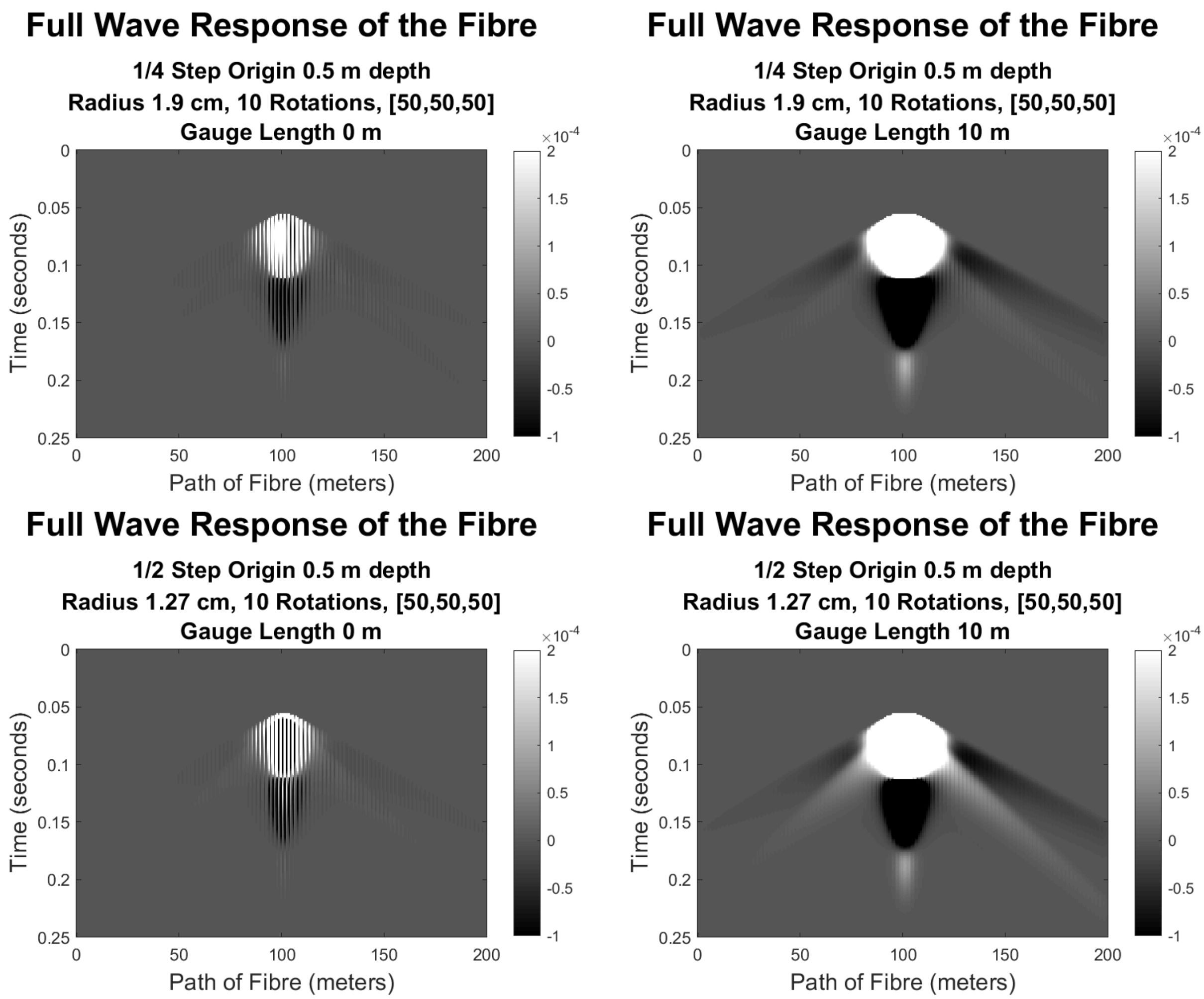


Figure 7.5: (Left column) The data without the gauge length applied (top) helical fibre diminished by 1/4-th of the radius and (bottom) helical fibre diminished by 1/2-th of the radius when the source is at Location 1. (Right column) The 10m gauge length applied to data seen in the left column.

fibre of radius 2.54cm decreased by 1/4-th of the radius (top right) does not have distinguishable S-wave response, but decreasing the radius by 1/2-th of the radius (bottom right) produces a visible S-wave.

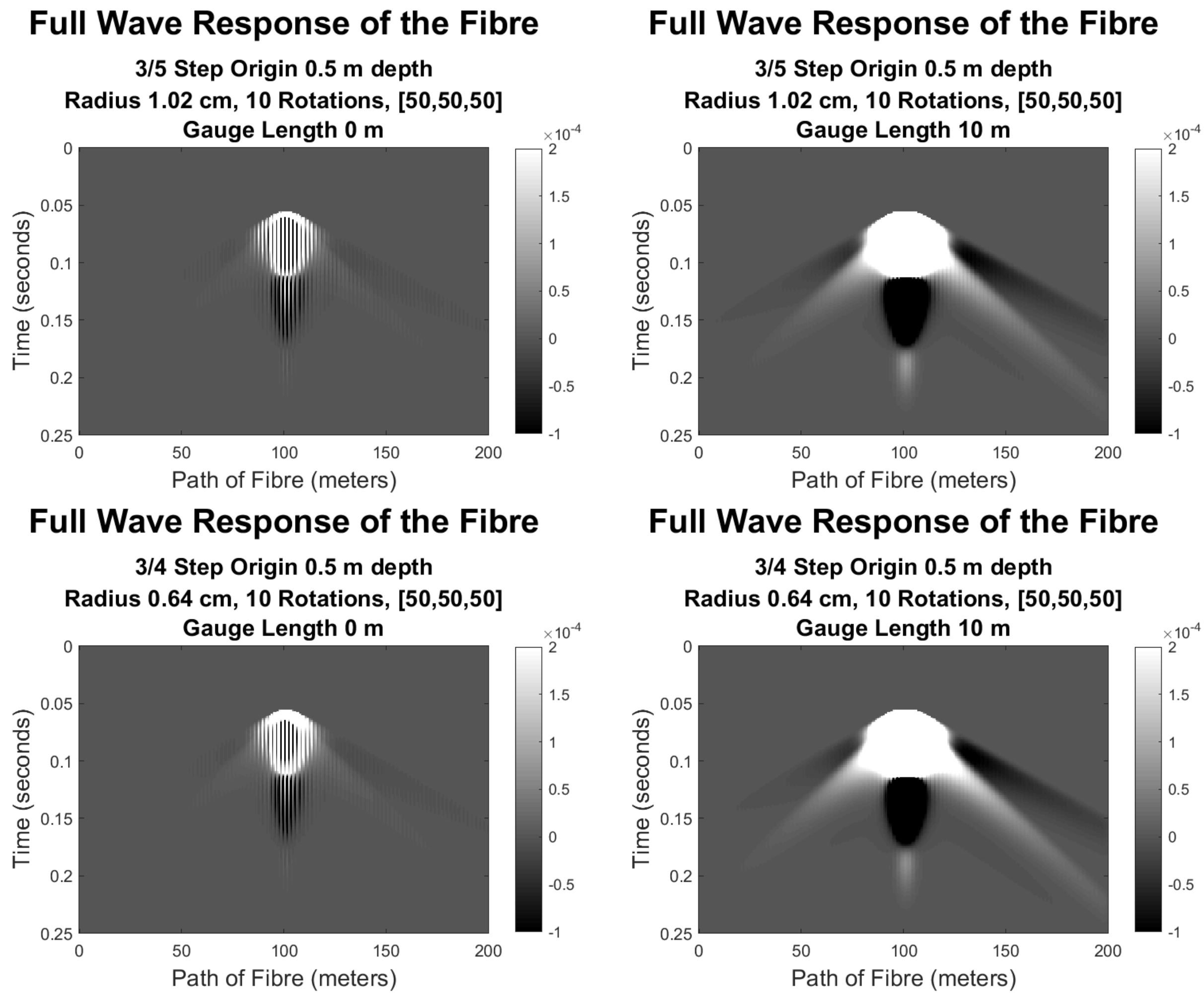


Figure 7.6: (Left column) The data without the gauge length applied (top) helical fibre diminished by 3/5-th of the radius and (bottom) helical fibre diminished by 3/4-th of the radius when the source is at Location 1. (Right column) The 10m gauge length applied to data seen in the left column.

Interestingly, Figure 7.6 provides even more evidence of the S-wave response as the radius of the helix decreases. The helix can still be distinguished in the data without the gauge length applied on the left; however, it does not appear in the data with a gauge length applied on the right, which is to be expected given the results found in Figure 7.4 and 7.5.

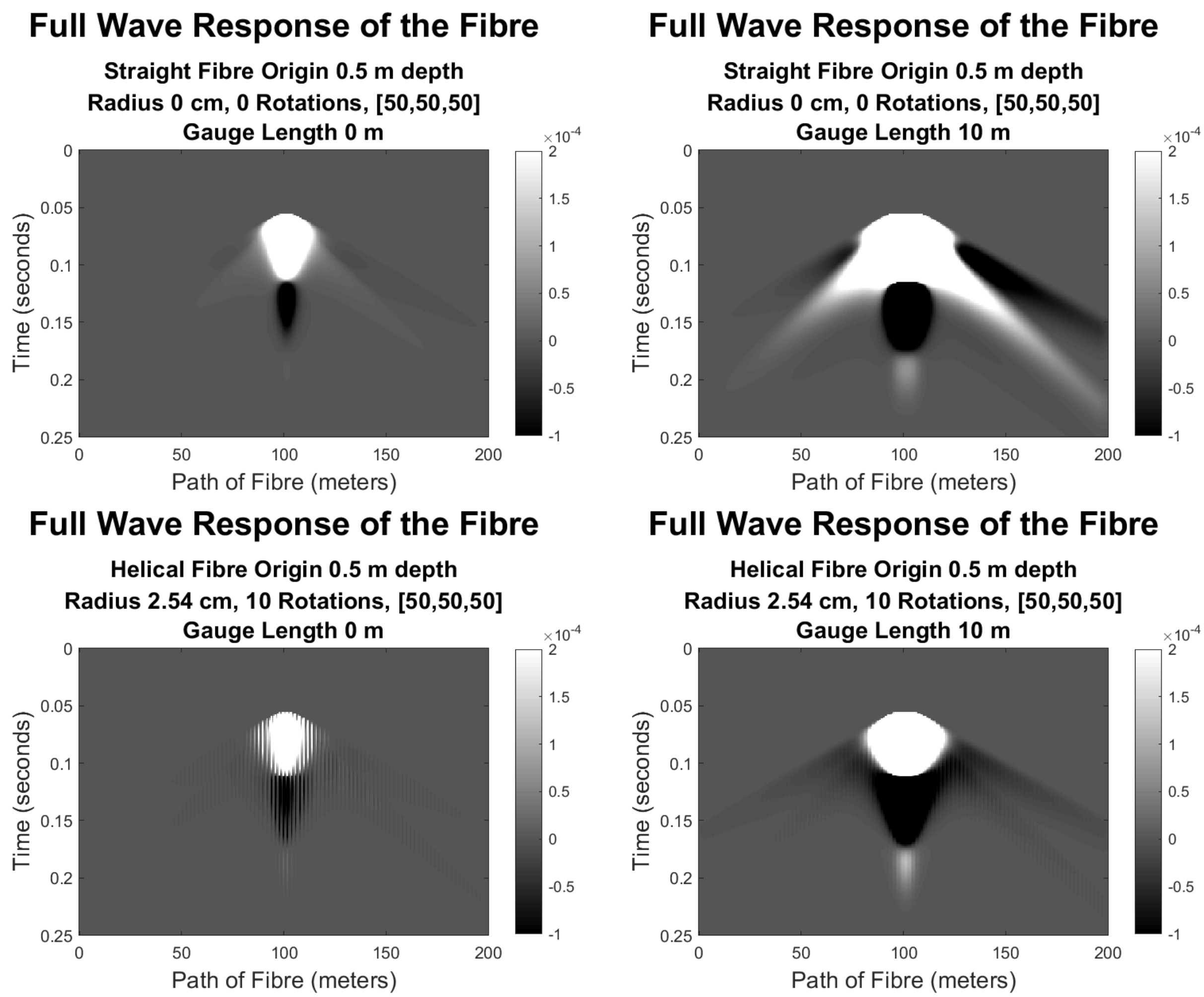


Figure 7.7: (Left column) The data without the gauge length applied (top) straight fibre and (bottom) helical fibre at Location 1. (Right column) The 10m gauge length applied to data seen in the left column at Location 1.

Figure 7.7 shows this pattern continuing. We compare the final results of the homotopy by including the results from the helical fibre in the bottom row of Figure 7.7. The S-wave is present for the straight fibre whereas it is not evident for the helical fibre of radius 2.54cm. the signal from the straight fibre possesses a flatter peak than the helical fibre. The tail of the signal for the straight fibre is wider. While the signal of the helical fibre is smaller than the signal of the straight fibre, it appears to be more precise with regards to the shape of the hyperbola; however, it is much more difficult to distinguish the S-wave response in the helical fibre response than it is in the straight fibre response. Recall from Chapter 6 that the helical fibre's S-wave response had a smaller amplitude than the straight fibre's S-wave response. the amplitude of the helical fibre's S-wave response is present; however, it may be small enough that it is not visible using the current color axis.

Figure 7.8 depicts the results of the straight fibre (top) and the helical fibre (bottom) when the source is at Location 2. The hyperbola is to right of where it was in the data for Location 1. We see that applying the gauge length to the data (right) has similar results to the data at Location 1. The peak of the straight fibre's response is flattened. The helical fibre's response no longer has evidence of a helix once the gauge length is applied.

In Figure 7.9, the results for the data at Location 3 produce a similar outcome to the previous two locations. The hyperbola response has once again moved toward the right. This is because Location 3 occurs at 111 meters down the fibre.

Figure 7.10 presents the results for Location 4. Recall that Location 4 is located at the center of the fibre but 11 meters away from the fibre formation, unlike the previous three locations which were 5 meters away from the fibre. The amplitude response of both the helical and straight fibre is much smaller for the new location, both when the gauge length is applied to the data and when it is not, than for the previous four locations. Once again when we apply the gauge length to the straight and helical fibre, it is difficult to distinguish the S-wave response in the helical fibre, while we can locate it in the straight fibre's response.

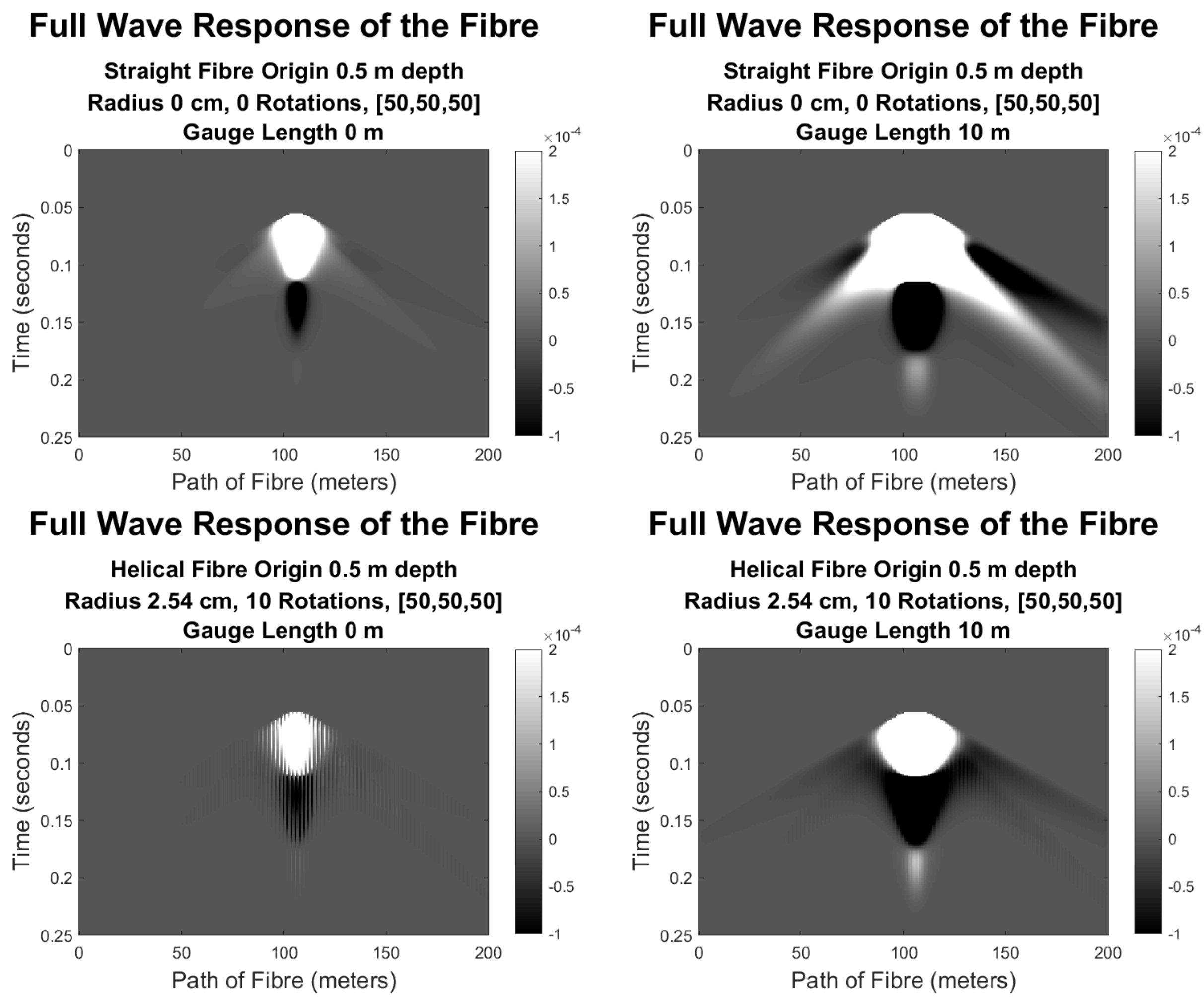


Figure 7.8: (Left column) The data without the gauge length applied (top) straight fibre and (bottom) helical fibre at Location 2. (Right column) The 10m gauge length applied to data seen in the left column at Location 2.

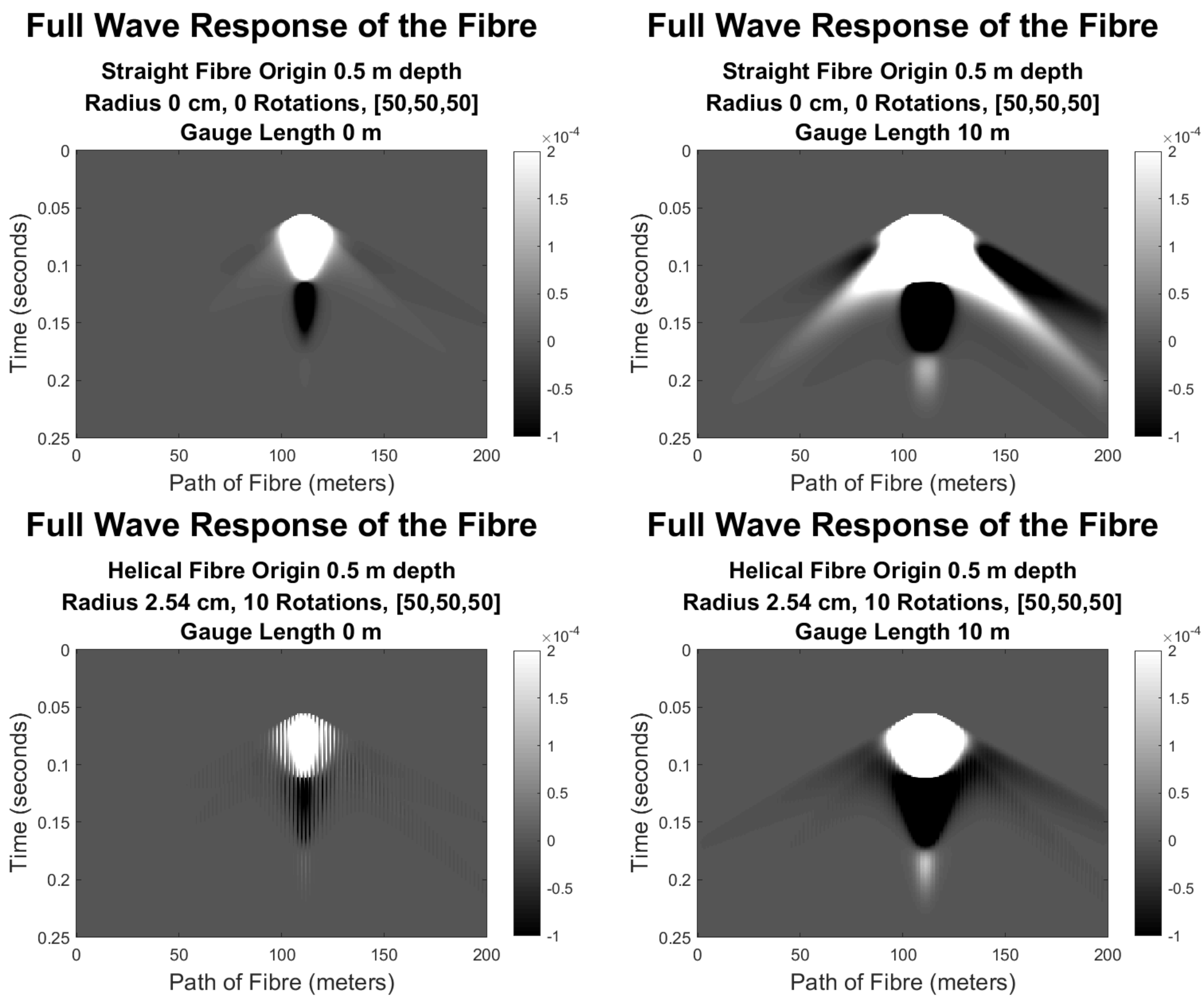


Figure 7.9: (Left column) The data without the gauge length applied (top) straight fibre and (bottom) helical fibre at Location 3. (Right column) The 10m gauge length applied to data seen in the left column at Location 3.

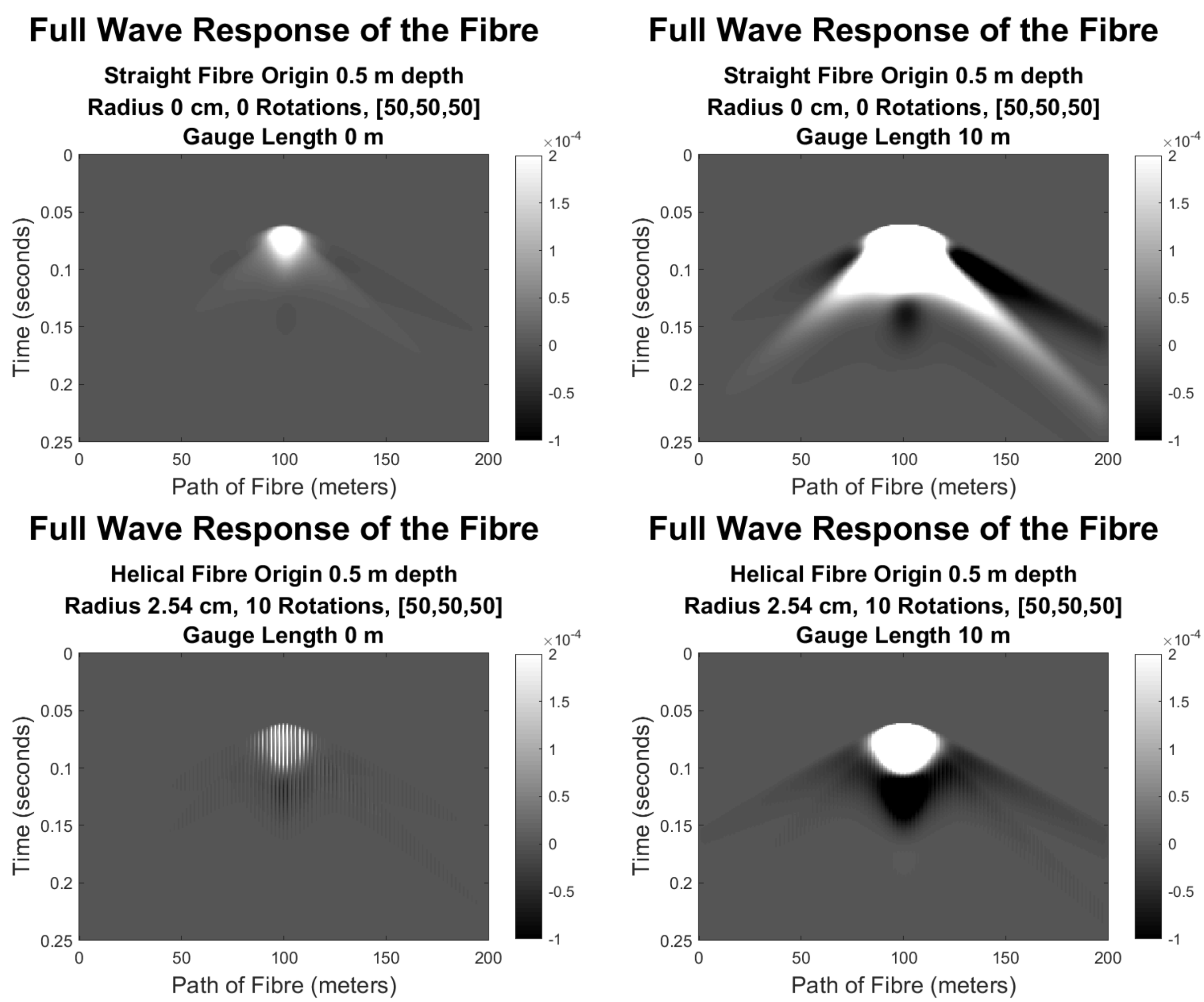


Figure 7.10: (Left column) The data without the gauge length applied (top) straight fibre and (bottom) helical fibre at Location 4. (Right column) The 10m gauge length applied to data seen in the left column at Location 4.

Analysis of the data along the homotopy

Investigating whether the results increase or decrease over w would provide insight into how the helical fibre compares to the straight fibre. Given that the output of the homotopy model is a matrix for each value of w, we use the L^2-norm to offer a comparison of each w along the homotopy. We first study the L^2-norm of the data without a gauge length applied and the L^2-norm of the data with a gauge length of 10m applied. Afterwards, we examine the effects on the trace, determinant, and eigenvalues of the strain matrix at each point $(\mathbf{p}(s), t)$ on the fibre path $\mathbf{p}$ at time t. We also include eight different S-wave orientation vectors to study the effect that the orientation of the S-wave has on the amplitude response of the fibre over the homotopy.

The results above consider a uniform response for the S-wave orientation vector $\mathbf{A} = [50, 50, 50]$. While not included here, we also conducted an experiment where we looked at less uniform values of $\mathbf{A}$ which we group into two sets. The first set of S-wave orientation vectors is $\mathbf{A} = [50, 20, 20]$, $\mathbf{A} = [20, 50, 20]$ and $\mathbf{A} = [20, 20, 50]$. We also studied vectors A where the values were between 0 and 1. The second set of S-wave orientation vectors $\mathbf{A}$ we examine are the vectors $\mathbf{A} = [1, 1, 1]$, $\mathbf{A} = [0.60, 1, 0.60]$, $\mathbf{A} = [0.60, 0.60, 1]$, and $\mathbf{A} = [1, 0.60, 0.60]$. While the data sets are not presented here, we do include information about the data sets in the our following comparisons.

Our first comparison looks at the norm of the data without the gauge length applied. We also include the results of applying the norm to the data when the gauge length is applied. In both cases, we applied the norm at each point w. In Figure 7.11, we see the results for the vectors $\mathbf{A} = [50, 50, 50]$, $\mathbf{A} = [50, 20, 20]$, $\mathbf{A} = [20, 50, 20]$, and $\mathbf{A} = [20, 20, 50]$ for each of the four locations. Figure 7.12 shows the norms when the gauge length is applied for these vectors A. Figure 7.13 and Figure 7.14 provides a similar setup for the vectors $\mathbf{A} = [1, 1, 1]$, $\mathbf{A} = [0.60, 1, 0.60]$, $\mathbf{A} = [0.60, 0.60, 1]$, and $\mathbf{A} = [1, 0.60, 0.60]$.

In Figures 7.11 to 7.14, for all S-wave orientation vectors $\mathbf{A}$ when taking the L^2-norm of the data without the gauge length applied, the response of the helically wound fibre is larger

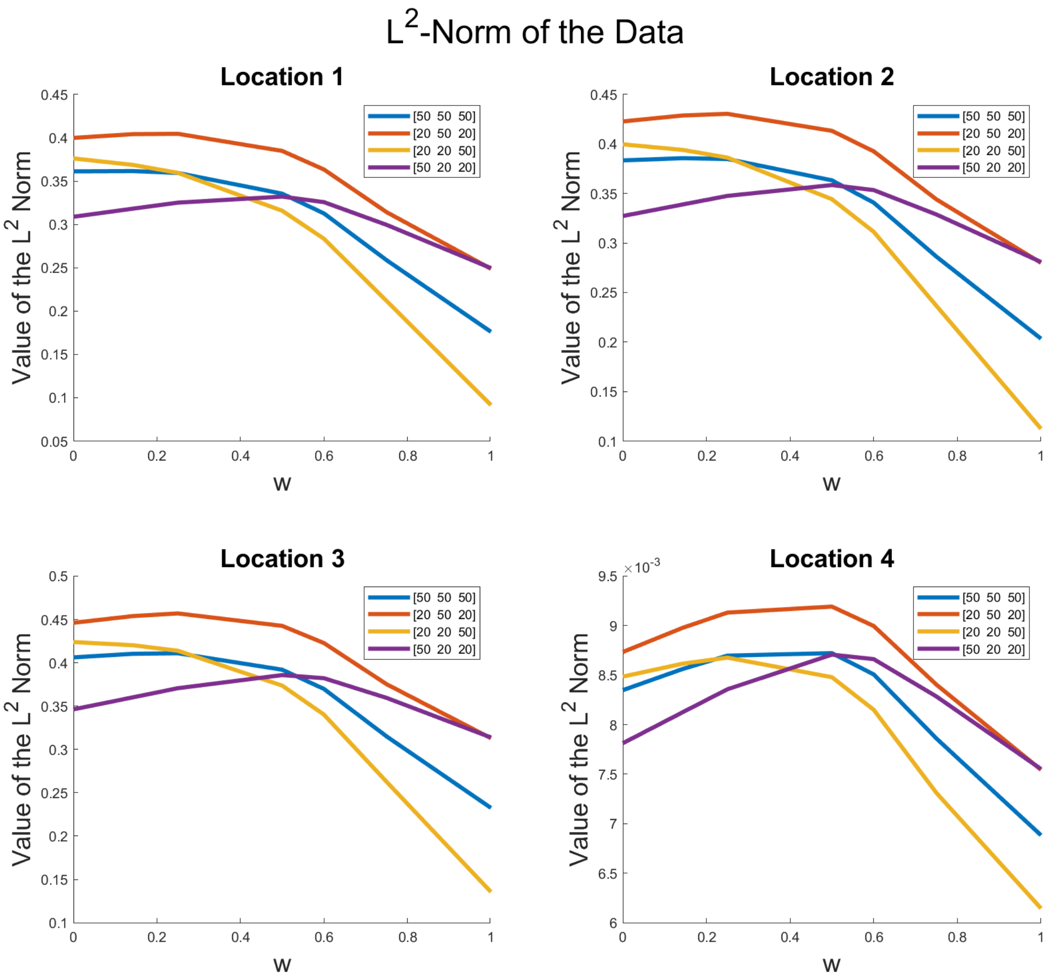


Figure 7.11: The L^2-norm of the data without the gauge length applied at (top left) Location 1, (top right) Location 2, (bottom left) Location 3, and (bottom right) Location 4 for S-wave orientation vectors $\mathbf{A} = [50, 50, 50]$ (blue), $\mathbf{A} = [20, 50, 20]$ (red), $\mathbf{A} = [20, 20, 50]$ (yellow), and $\mathbf{A} = [50, 20, 20]$ (purple).

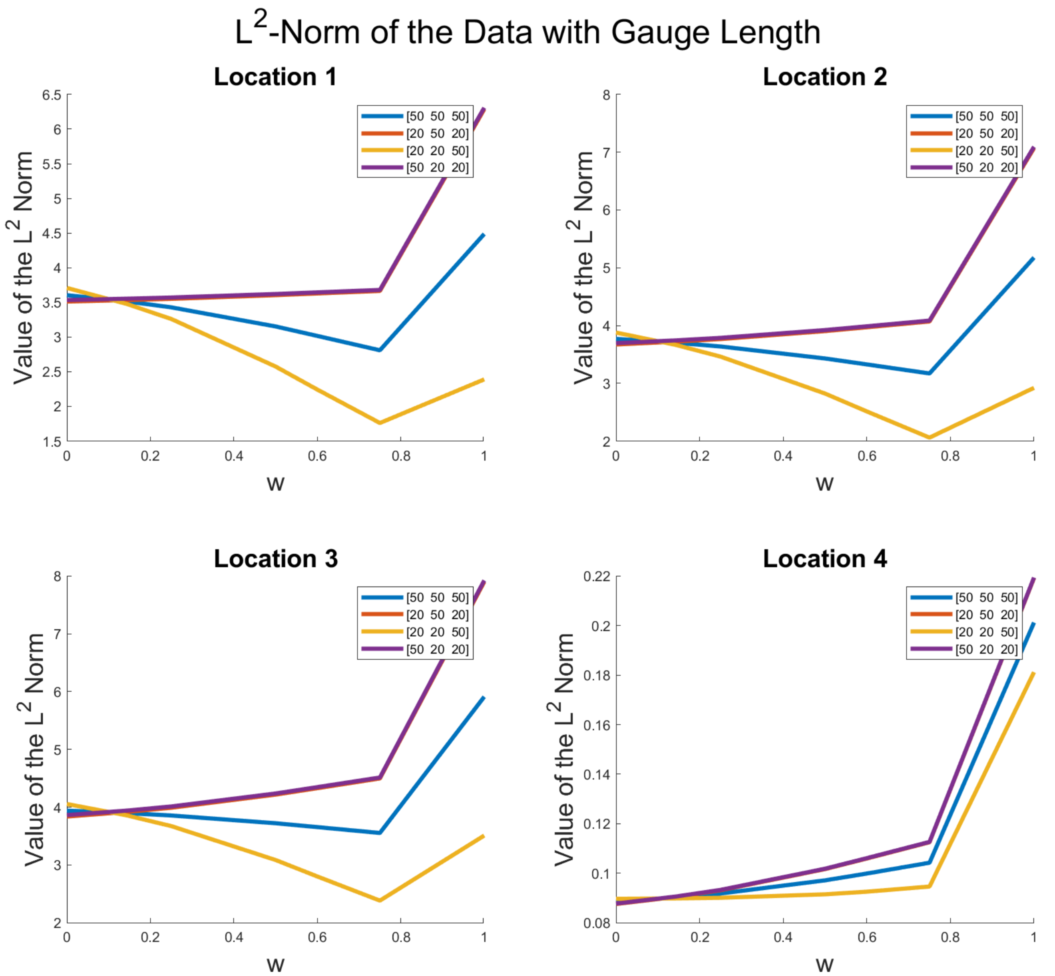


Figure 7.12: The L^2-norm of the data with the gauge length applied at (top left) Location 1, (top right) Location 2, (bottom left) Location 3, and (bottom right) Location 4 for S-wave orientation vectors $\mathbf{A} = [50, 50, 50]$ (blue), $\mathbf{A} = [20, 50, 20]$ (red), $\mathbf{A} = [20, 20, 50]$ (yellow), and $\mathbf{A} = [50, 20, 20]$ (purple).

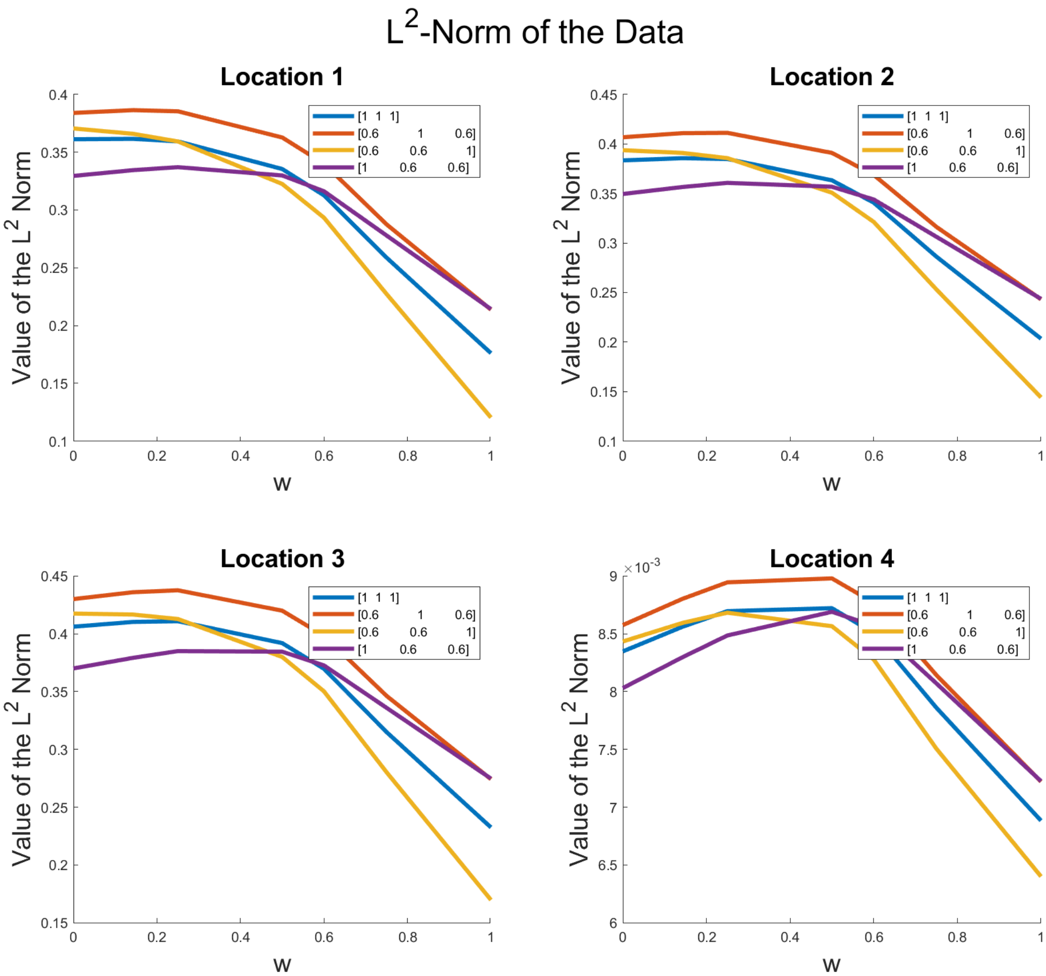


Figure 7.13: The L^2-norm of the data without the gauge length applied at (top left) Location 1, (top right) Location 2, (bottom left) Location 3, and (bottom right) Location 4 for S-wave orientation vectors $\mathbf{A} = [1, 1, 1]$ (blue), $\mathbf{A} = [0.6, 1, 0.6]$ (red), $\mathbf{A} = [0.6, 0.6, 1]$ (yellow), and $\mathbf{A} = [1, 0.6, 0.6]$ (purple).

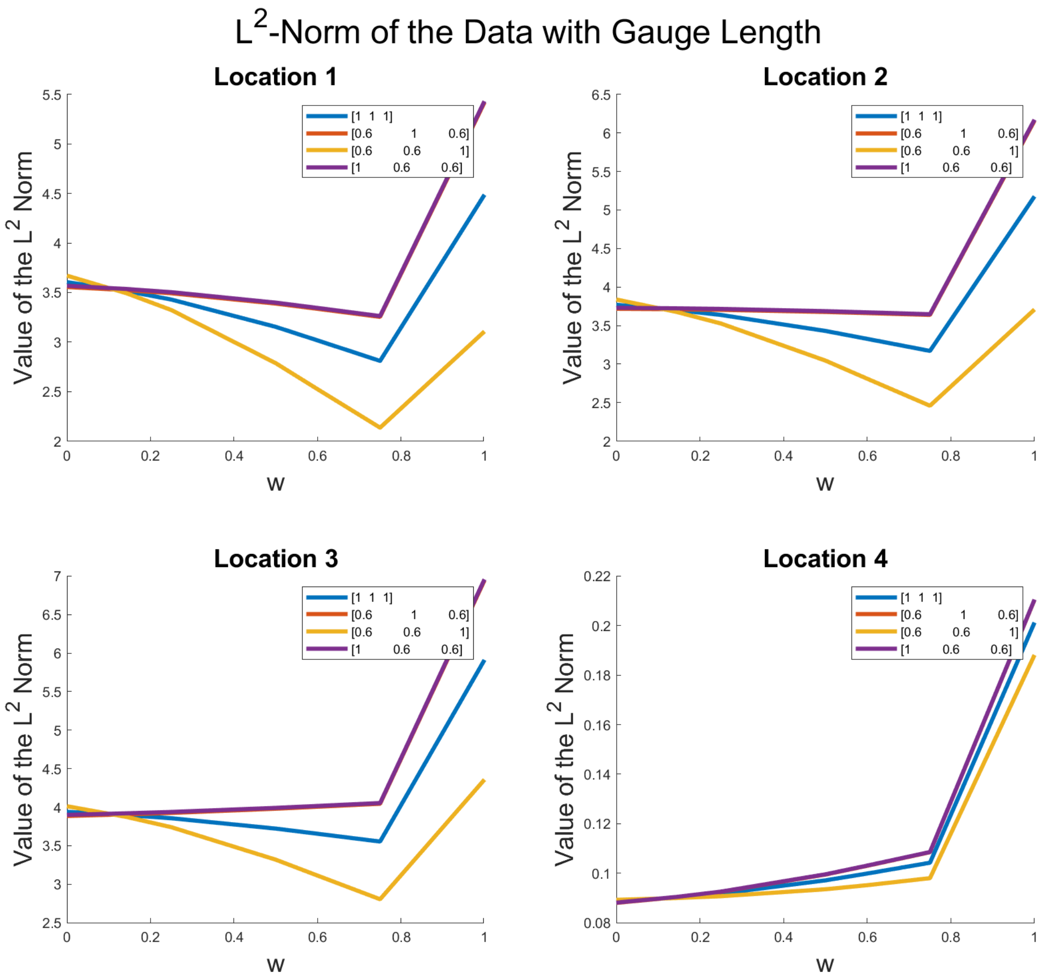


Figure 7.14: The L^2-norm of the data with the gauge length applied at (top left) Location 1, (top right) Location 2, (bottom left) Location 3, and (bottom right) Location 4 for S-wave orientation vectors $\mathbf{A} = [1, 1, 1]$ (blue), $\mathbf{A} = [0.6, 1, 0.6]$ (red), $\mathbf{A} = [0.6, 0.6, 1]$ (yellow), and $\mathbf{A} = [1, 0.6, 0.6]$ (purple).

than the response of the straight fibre. Once the gauge length is applied to the data, the norm of the response for the helical fibre is smaller than the norm of the response for the straight fibre. Given that Location 4 is 11 meters away from the fibre compared to Locations 1 to 3, the norm of the helical fibre's response appears to decrease more in comparison to the norm of the straight fibre's response for all S-wave orientation vectors $\mathbf{A}$. Only two cases of S-wave orientation vectors $\mathbf{A}$ appear to have a helical response which is larger than the straight fibre's response when the gauge length is applied: the cases when $\mathbf{A} = [20, 20, 50]$ and $\mathbf{A} = [0.6, 0.6, 1]$. This only holds true for the first three locations, and not for Location 4. It suggests there are some formations of fibre where the helical fibre provides a better response than the straight fibre once a gauge length is applied to the data; however, it depends on the orientation of the S-wave response and the location of the source.

With regards to the choice of orientation, the S-wave orientation vectors $\mathbf{A} = [20, 50, 20]$ and $\mathbf{A} = [50, 20, 20]$ vary when no gauge length is applied to the data; however, they appear to give very similar results once the gauge length is applied. The same can be said for the S-wave orientation vectors $\mathbf{A} = [0.6, 1, 0.6]$ and $\mathbf{A} = [1, 0.6, 0.6]$. This may be attributed to the formation of the fibre with regards to the S-wave orientation vector. The uniform S-wave orientation vectors $\mathbf{A} = [50, 50, 50]$ and $\mathbf{A} = [1, 1, 1]$ give responses which lie between the results of the other S-wave orientation vectors $\mathbf{A}$ when the gauge length is applied to the data and when it is not. Given the different results for each vector $\mathbf{A}$, it suggests that the success of the fibre depends on the orientation of the source which follows as the amplitude response is linear with respect to the vector $\mathbf{A}$.

We also observe that for some S-wave orientation vectors $\mathbf{A}$, the 2.54cm helical radius response ($w = 0$) does not give the largest norm of the data without the gauge length applied across the homotopy. It occurs for the S-wave orientation vectors $\mathbf{A} = [50, 20, 20]$ and $\mathbf{A} = [1, 0.6, 0.6]$ in Figures 7.11 and 7.13, respectively. When the gauge length is applied to the data, the norm of the data at $w = 0$ is not the minimum over the homotopy for S-wave orientation vectors $\mathbf{A} = [50, 50, 50]$, $\mathbf{A} = [20, 20, 50]$, $A = [1, 1, 1]$, and $\mathbf{A} = [0.6, 0.6, 1]$

in Locations 1, 2, and 3, whereas for S-wave orientation vectors $\mathbf{A} = [0.6, 1, 0.6]$ and $\mathbf{A} = [1, 0.6, 0.6]$, it only occurs at Location 1. Given these varying results, it supports the fact that the amplitude of the source affects the succes of the fibre formation.

Analysis of the strain matrices: Trace

Now, let us explore the effects on the strain matrix ϵ along the homotopy. We begin by calculating the trace of the strain matrix at each point (s, t) and assign it to the corresponding entry of a matrix M_{tr}. Then we take the L^2-norm of the matrix M_{tr} for each w. From linear algebra, the trace simply sums the diagonal of the strain matrix. The general form of the trace is

$$\operatorname{tr}(\epsilon) = -\Big(G_\alpha + \frac{1}{r^2} H_\alpha\Big), \tag{7.57}$$

where α is the velocity of the P-wave. Interestingly, the trace is only dependent on the P-wave which means that the trace does not depend on the orientation of the S-wave response.

Figures 7.15 and 7.16 show the results of computing the norm on the traces. The trace varies over each location. It does decrease from the helically wound fibre's response to the straight fibre's response; however, it appears to stay unchanged between the different S-wave orientation vectors $\mathbf{A}$. Given that the trace only depends on the P-wave, it follows that the S-wave orientation vector $\mathbf{A}$ would not affect the trace, since the vector $\mathbf{A}$ shows up only in the S-wave response. The norm of the traces remains unaffected by the amplitude of the source, but it does depend on the location of the source since the value of the norm of the traces changes between each location.

Analysis of the strain matrices: Determinants

Another method of comparison using the homotopy involves looking at the determinant of the strain tensor. We calculate the general determinant of the strain tensor to be the

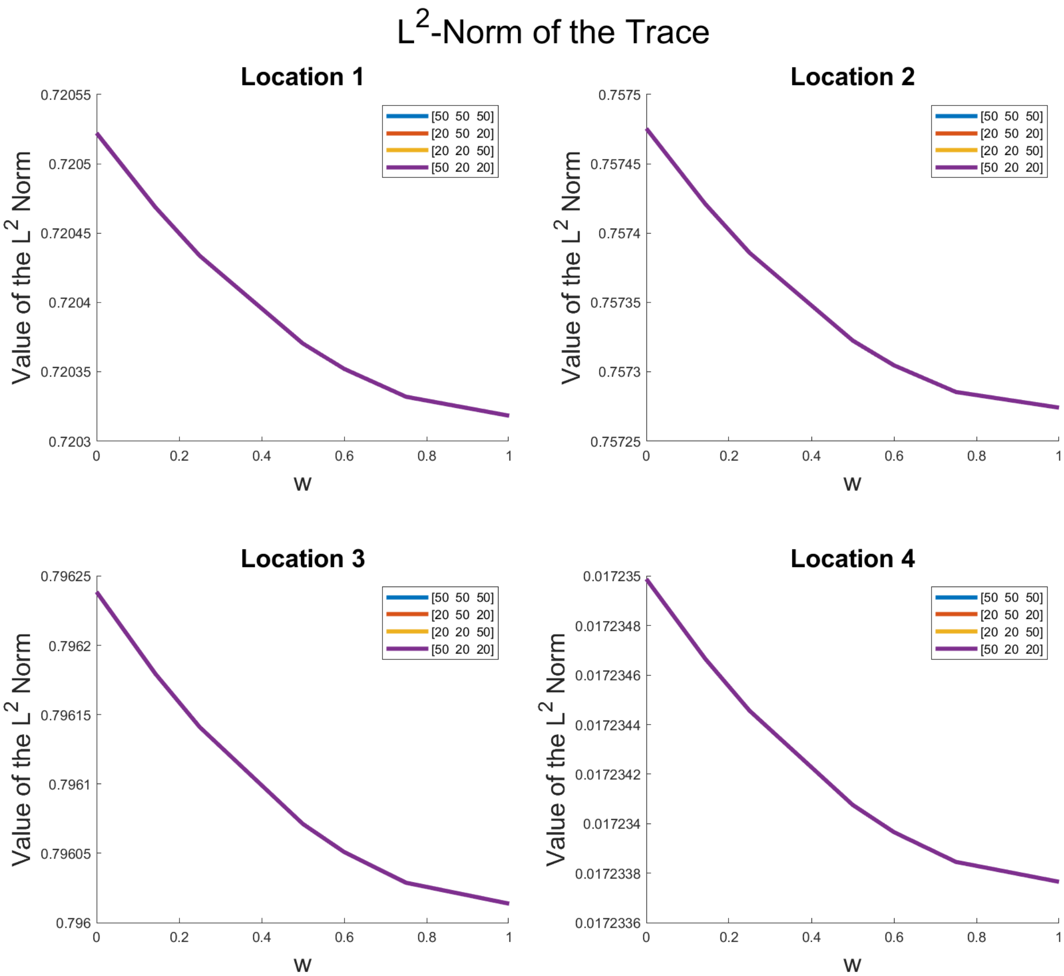


Figure 7.15: The L^2-norm of the traces of the strain matrices of the data at (top left) Location 1, (top right) Location 2, (bottom left) Location 3, and (bottom right) Location 4 for S-wave orientation vectors $\mathbf{A} = [50, 50, 50]$ (blue), $\mathbf{A} = [20, 50, 20]$ (red), $\mathbf{A} = [20, 20, 50]$ (yellow), and $\mathbf{A} = [50, 20, 20]$ (purple).

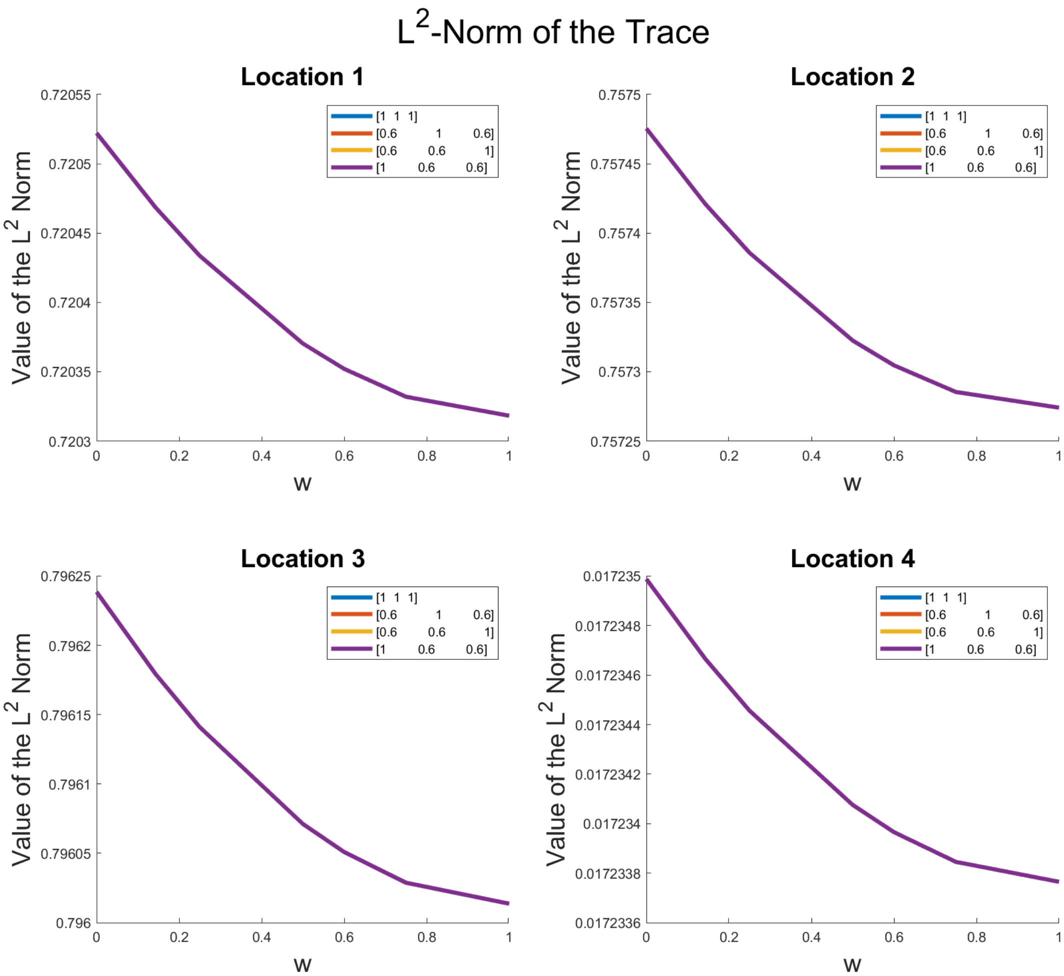


Figure 7.16: The L^2-norm of the traces of the strain matrices of the data at (top left) Location 1, (top right) Location 2, (bottom left) Location 3, and (bottom right) Location 4 for S-wave orientation vectors $\mathbf{A} = [1, 1, 1]$ (blue), $\mathbf{A} = [0.6, 1, 0.6]$ (red), $\mathbf{A} = [0.6, 0.6, 1]$ (yellow), and $\mathbf{A} = [1, 0.6, 0.6]$ (purple).

following equation:

$$\det(\epsilon) = \frac{1}{8r^6}\Big(8G_\alpha^3 A + 4G_\alpha^2 G_\beta B + G_\alpha G_\beta^2 C + 8G_\alpha H_\alpha r^2 + G_\beta^3 D + G_\beta^2 H_\alpha E + 8H_\alpha\Big) \tag{7.58}$$

where

$$A = x^2y^2z(y-z); \tag{7.59}$$

$$B = -xy\Big(2\mathbf{X}yz(z-y) + \mathbf{Y}xz(2z-3y) + \mathbf{Z}xy(2z-y)\Big); \tag{7.60}$$

$$C = 2\Big(\mathbf{X}^2y^3z + \mathbf{Y}^2x^2z(3y-z) - \mathbf{Z}^2x^2y^2\mathbf{XY}z(-\mathbf{XY}z + 4xy^2 - 3xyz) \tag{7.61}$$

$$+ \mathbf{XZ}xy^2(2y-3z) + \mathbf{YZ}x^2y(2y-3z)\Big); \tag{7.62}$$

$$D = \mathbf{XY}\big(\mathbf{Y}y(\mathbf{X}yz - 2xy) - z^2(\mathbf{X}y + \mathbf{Y}x)\big) + \mathbf{XZ}y^2\big(\mathbf{X}y - (\mathbf{X}z + \mathbf{Z}x)\big) \tag{7.63}$$

$$+ \mathbf{YZ}x^2\big(\mathbf{Y}y - (\mathbf{Y}z + \mathbf{Z}y)\big) + \mathbf{XYZ}y(y - xz) + \mathbf{Y}^3x^2z; \tag{7.64}$$

$$E = -2\Big((\mathbf{X}y - \mathbf{Y}x)^2 + (\mathbf{X}z - \mathbf{Z}x)^2 + (\mathbf{Y}z - \mathbf{Z}y)^2\Big). \tag{7.65}$$

The functions G and H depend on the minimum-phase source function used. The subscripts on the functions G and H denote whether the P-wave (α) or S-wave (β) velocity is used in the function. The terms $\mathbf{X}$, $\mathbf{Y}$, and $\mathbf{Z}$ are the same as the equations found in Chapter 6; c.f. Equation 6.61.

Given the complicated nature of the equation for the determinant, a numerical comparison of the strain determinants at each point along the fibre would provide more insight at this time. Numerical comparison seems much more feasible since it would be unrealistic to simplify the equation by setting some values of $\mathbf{A}_i = 0$ as the vector $\mathbf{A}$ represents the orientation of the S-wave response. Let us consider the determinant of the strain matrix at each the point (s,t) and take the norm over all (s,t). Again, we look at the results for all S-wave orientation vectors $\mathbf{A}$ that we studied for the previous comparisons.

As with the trace example, we find the determinant of the strain matrix for each point (s,t) and then assign it to a matrix M_{det}. There will be a matrix M_{det} for each w. We take

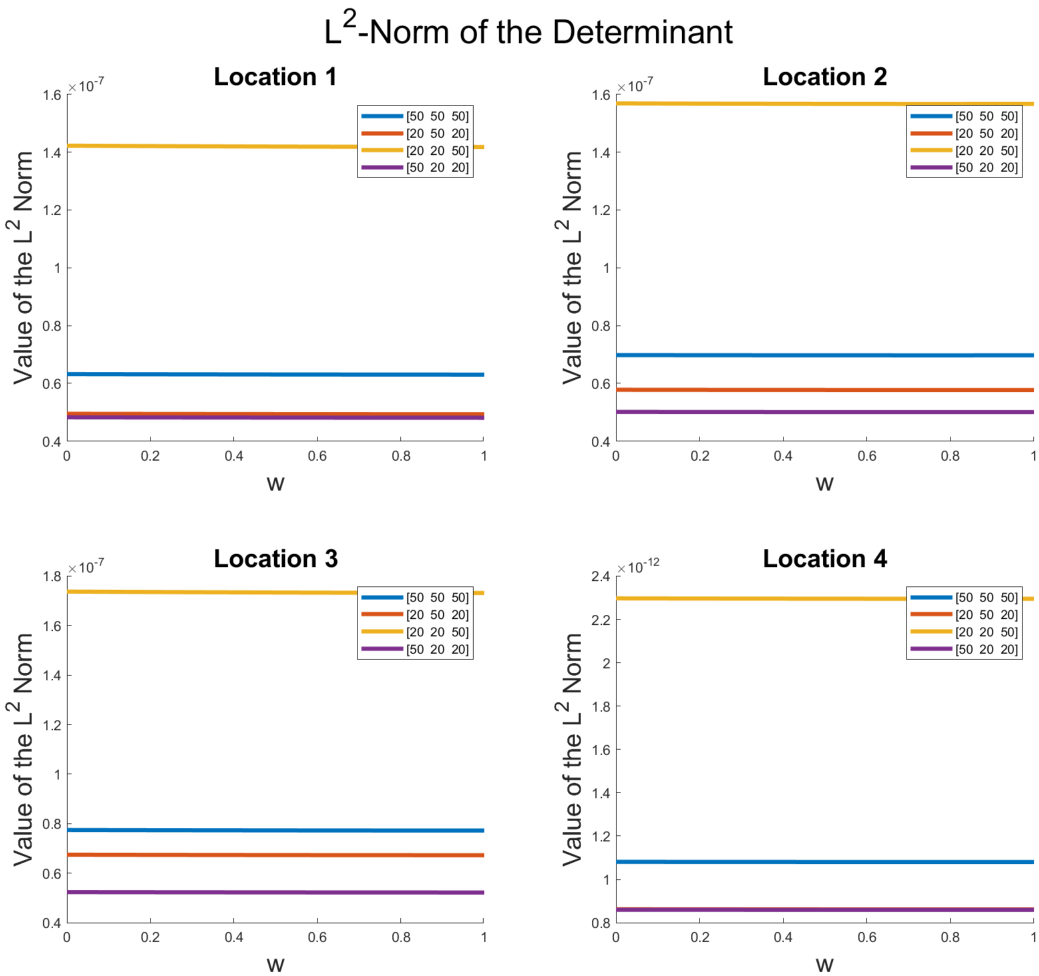


Figure 7.17: The L^2-norm of the determinants of the strain matrices of the data at (top left) Location 1, (top right) Location 2, (bottom left) Location 3, and (bottom right) Location 4 for S-wave orientation vectors $\mathbf{A} = [50, 50, 50]$ (blue), $\mathbf{A} = [20, 50, 20]$ (red), $\mathbf{A} = [20, 20, 50]$ (yellow), and $\mathbf{A} = [50, 20, 20]$ (purple).

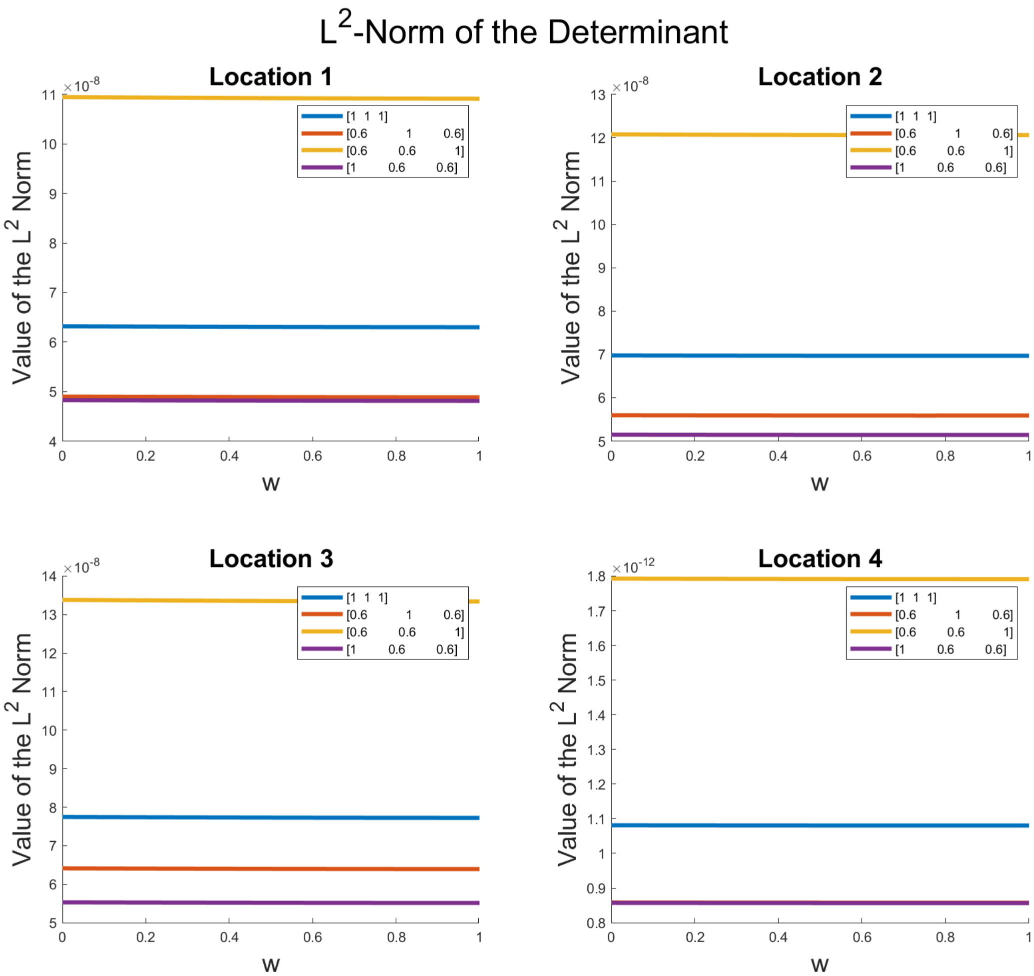


Figure 7.18: The L^2-norm of the determinants of the strain matrices of the data at (top left) Location 1, (top right) Location 2, (bottom left) Location 3, and (bottom right) Location 4 for S-wave orientation vectors $\mathbf{A} = [1, 1, 1]$ (blue), $\mathbf{A} = [0.6, 1, 0.6]$ (red), $\mathbf{A} = [0.6, 0.6, 1]$ (yellow), and $\mathbf{A} = [1, 0.6, 0.6]$ (purple).

the L^2-norm of each M_{det} to generate the Figures 7.17 and 7.18. Figure 7.17 shows the L^2-norm of the determinants for all (s, t) for the first set of S-wave orientation vectors $\mathbf{A}$ at each location. Figure 7.18 depicts the norm of the determinant for all (s, t) for the second set of S-wave orientation vectors $\mathbf{A}$. Interestingly, the value of the norm of the determinants does not appear to change over the homotopy for any S-wave orientation vector $\mathbf{A}$ or location; however, they each decrease by a small increment from the helically wound fibre's response to the straight fibre's response. Consistently across all four locations, the S-wave orientation vectors $\mathbf{A} = [20, 20, 50]$ and $\mathbf{A} = [0.6, 0.6, 1]$ provide the highest norms of the determinants out of the four vectors in their group over each location. Both sets of S-wave orientation vectors $\mathbf{A}$ keep the same order from greatest norm to least norm across all four locations.

Analysis of the strain matrices: Eigenvalues

Finally, we compare the eigenvalues of the strain matrices at points (s, t). Given that ϵ is a 3×3 symmetric matrix, there will be three eigenvalues. In Chapter 6, we discussed the eigenvalues for the P-wave and S-wave strain matrices separately, and employed the eigenvectors to deduce information about the fibre sensor. From Theorem 7.6 and 7.7, we know that the eigenvalues play an important role with respect to the amplitude of the response. In this section, we calculate the eigenvalues for each strain matrix, group the eigenvalues into three groups, and then compute the L^2-norm over all (s, t) for three different sets of eigenvalues.

In Figure 7.19, we see the results of the L^2-norm applied to the first group of eigenvalues for the S-wave orientation vectors $\mathbf{A} = [50, 50, 50]$, $\mathbf{A} = [20, 50, 20]$, $\mathbf{A} = [20, 20, 50]$, and $\mathbf{A} = [50, 20, 20]$ at each location, and in Figure 7.20, we find the results of the norm applied to the first group of eigenvalues for the S-wave orientation vectors $\mathbf{A} = [1, 1, 1]$, $\mathbf{A} = [0.6, 1, 0.6]$, $\mathbf{A} = [0.6, 0.6,]$, and $\mathbf{A} = [1, 0.6, 0.6]$ at each location. As with the norm of the determinants, the first eigenvalues do not appear to change over the homotopy. There is a slight change in some of the norms; however, this change is smaller than a factor of 10^{-4} in some cases. As

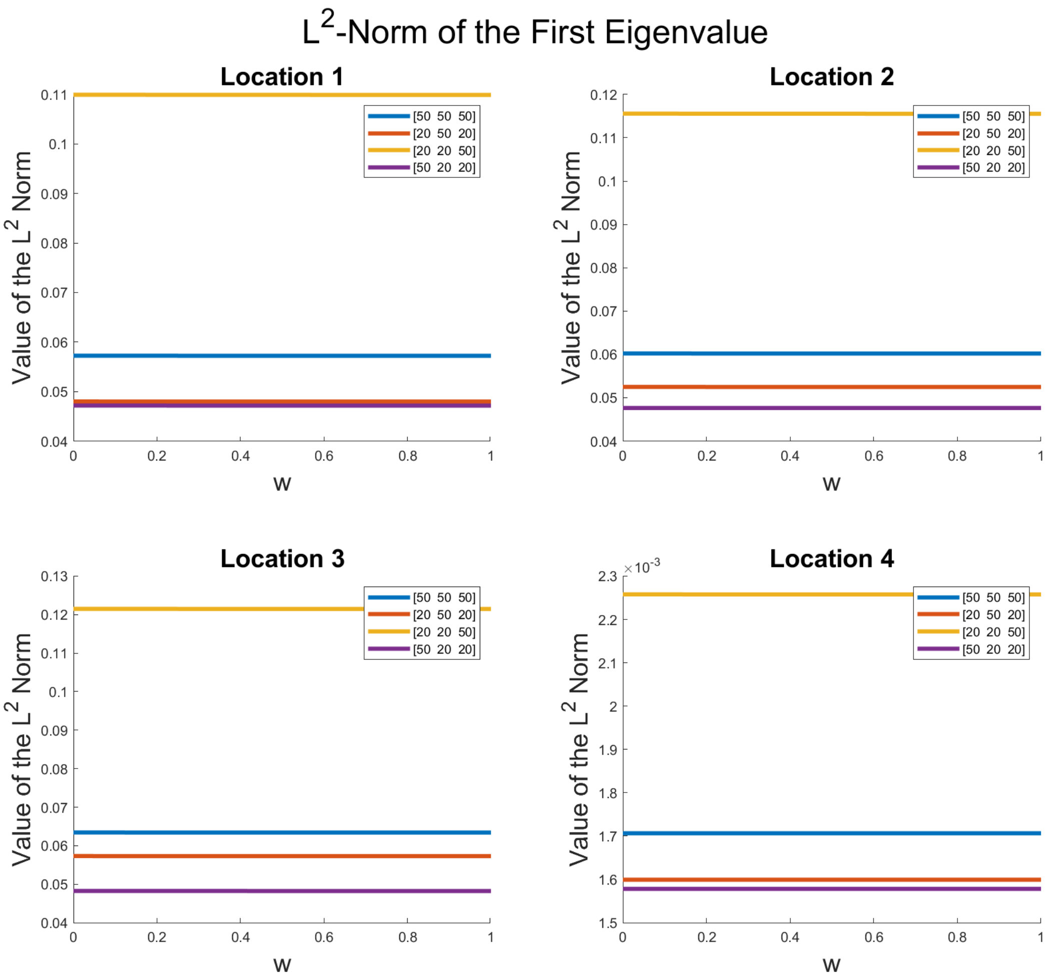


Figure 7.19: The L^2-norm of the first eigenvalues of the strain matrices of the data at (top left) Location 1, (top right) Location 2, (bottom left) Location 3, and (bottom right) Location 4 for S-wave orientation vectors $\mathbf{A} = [50, 50, 50]$ (blue), $\mathbf{A} = [20, 50, 20]$ (red), $\mathbf{A} = [20, 20, 50]$ (yellow), and $\mathbf{A} = [50, 20, 20]$ (purple).

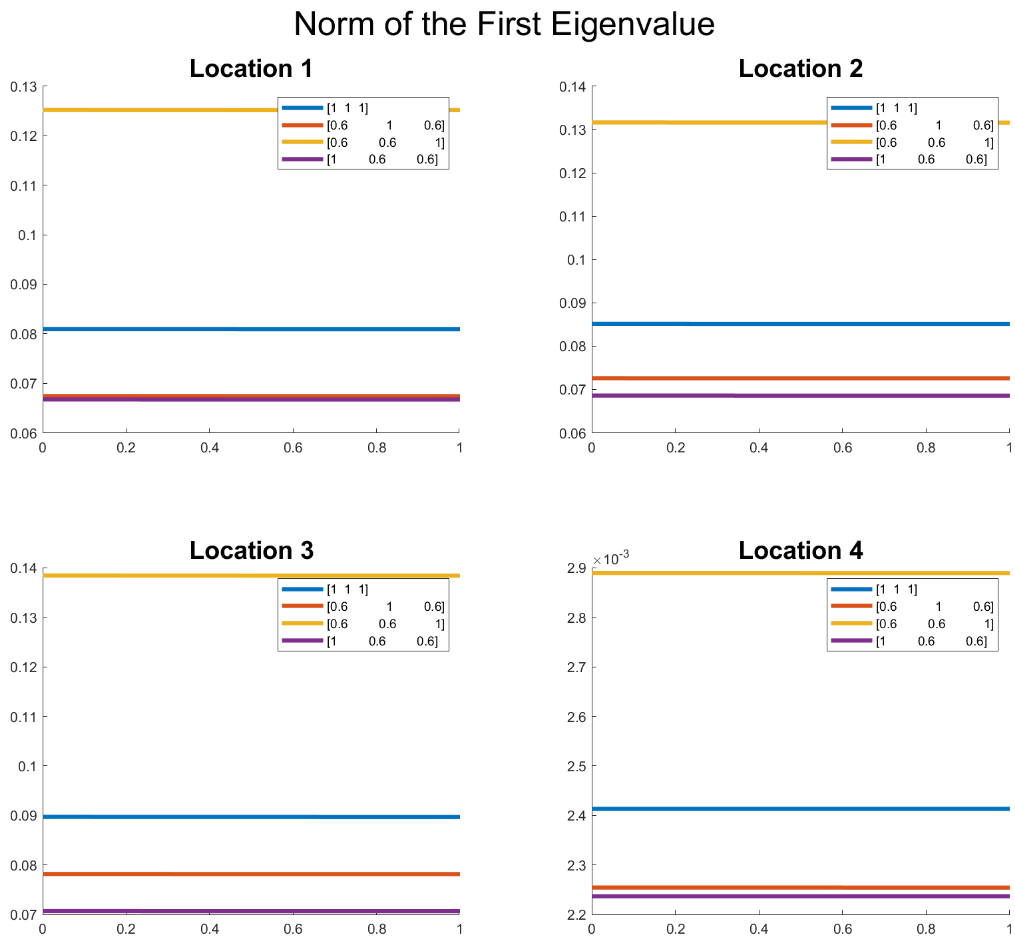


Figure 7.20: The L^2-norm of the first eigenvalues of the strain matrices of the data at (top left) Location 1, (top right) Location 2, (bottom left) Location 3, and (bottom right) Location 4 for S-wave orientation vectors $\mathbf{A} = [1, 1, 1]$ (blue), $\mathbf{A} = [0.6, 1, 0.6]$ (red), $\mathbf{A} = [0.6, 0.6, 1]$ (yellow), and $\mathbf{A} = [1, 0.6, 0.6]$ (purple).

with the case for the norm of the determinants, we see that both sets of S-wave orientation vectors have the same ordering as the norms of the first eigenvalues from greatest to least. In fact, the ordering is the same as the case for the determinants.

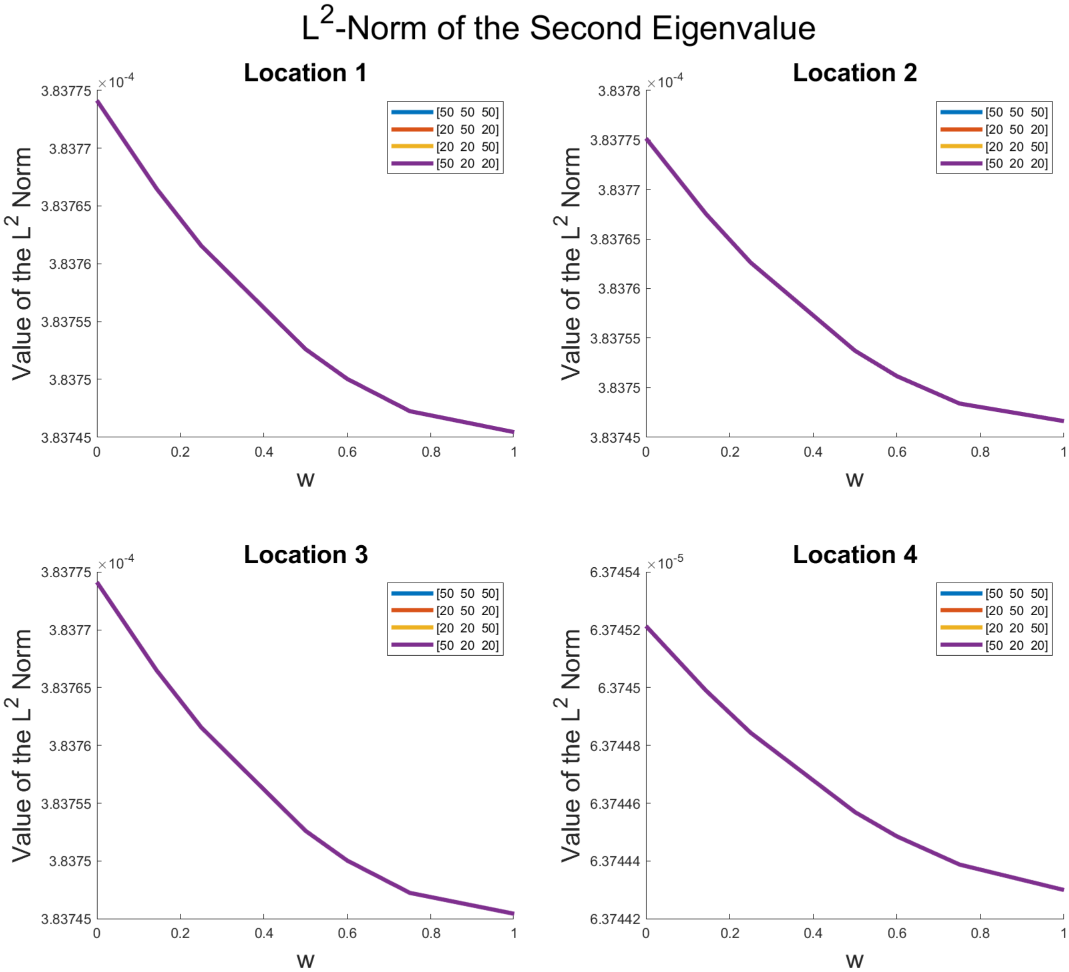


Figure 7.21: The L^2-norm of the second eigenvalues of the strain matrices of the data at (top left) Location 1, (top right) Location 2, (bottom left) Location 3, and (bottom right) Location 4 for S-wave orientation vectors $\mathbf{A} = [50, 50, 50]$ (blue), $\mathbf{A} = [20, 50, 20]$ (red), $\mathbf{A} = [20, 20, 50]$ (yellow), and $\mathbf{A} = [50, 20, 20]$ (purple).

While the L^2-norms of the second eigenvalues change over each location, we see in Figures 7.21 and 7.22 that the norm stays the same for each S-wave orientation vector $\mathbf{A}$ in the two sets of S-wave orientation vectors. Any difference between the norms appears to occur at a

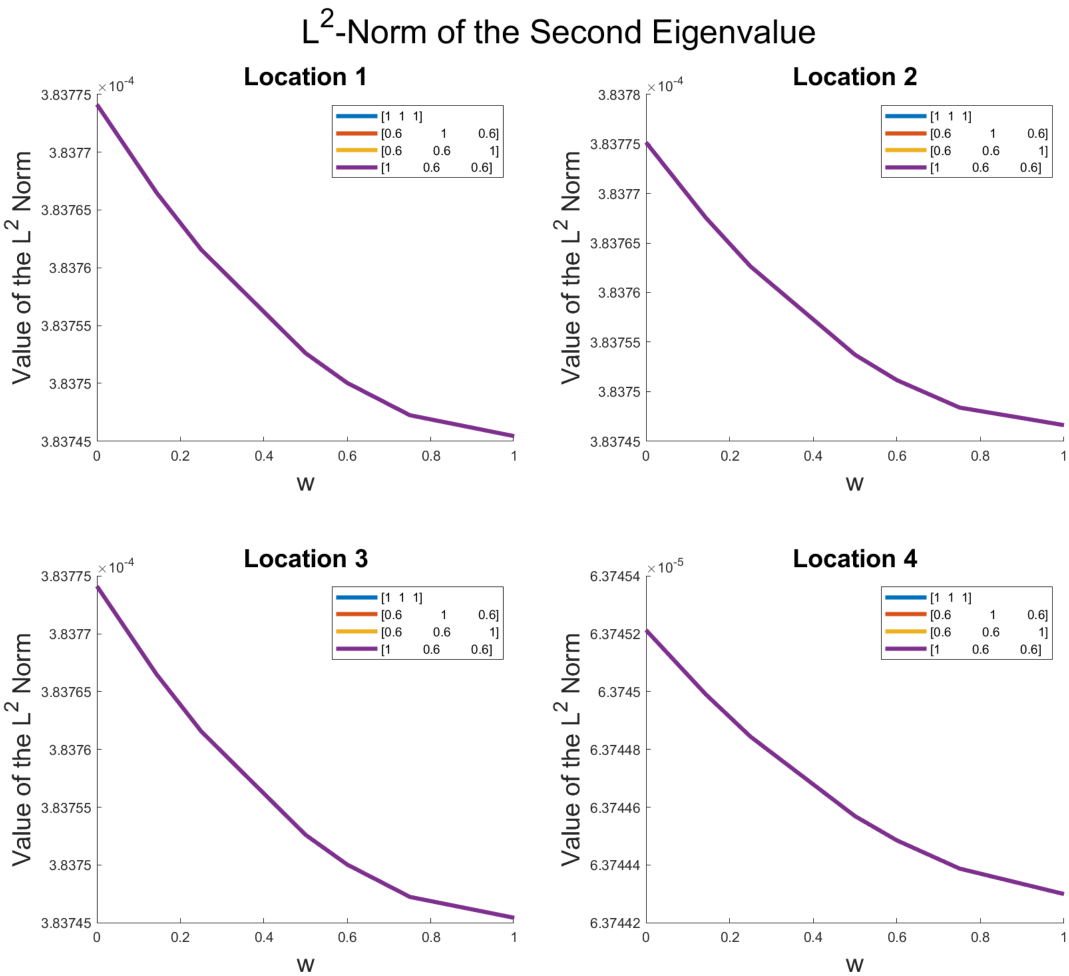


Figure 7.22: The L^2-norm of the second eigenvalues of the strain matrices of the data at (top left) Location 1, (top right) Location 2, (bottom left) Location 3, and (bottom right) Location 4 for S-wave orientation vectors $\mathbf{A} = [1, 1, 1]$ (blue), $\mathbf{A} = [0.6, 1, 0.6]$ (red), $\mathbf{A} = [0.6, 0.6, 1]$ (yellow), and $\mathbf{A} = [1, 0.6, 0.6]$ (purple).

factor of 10^{-17} or less which is negligible.

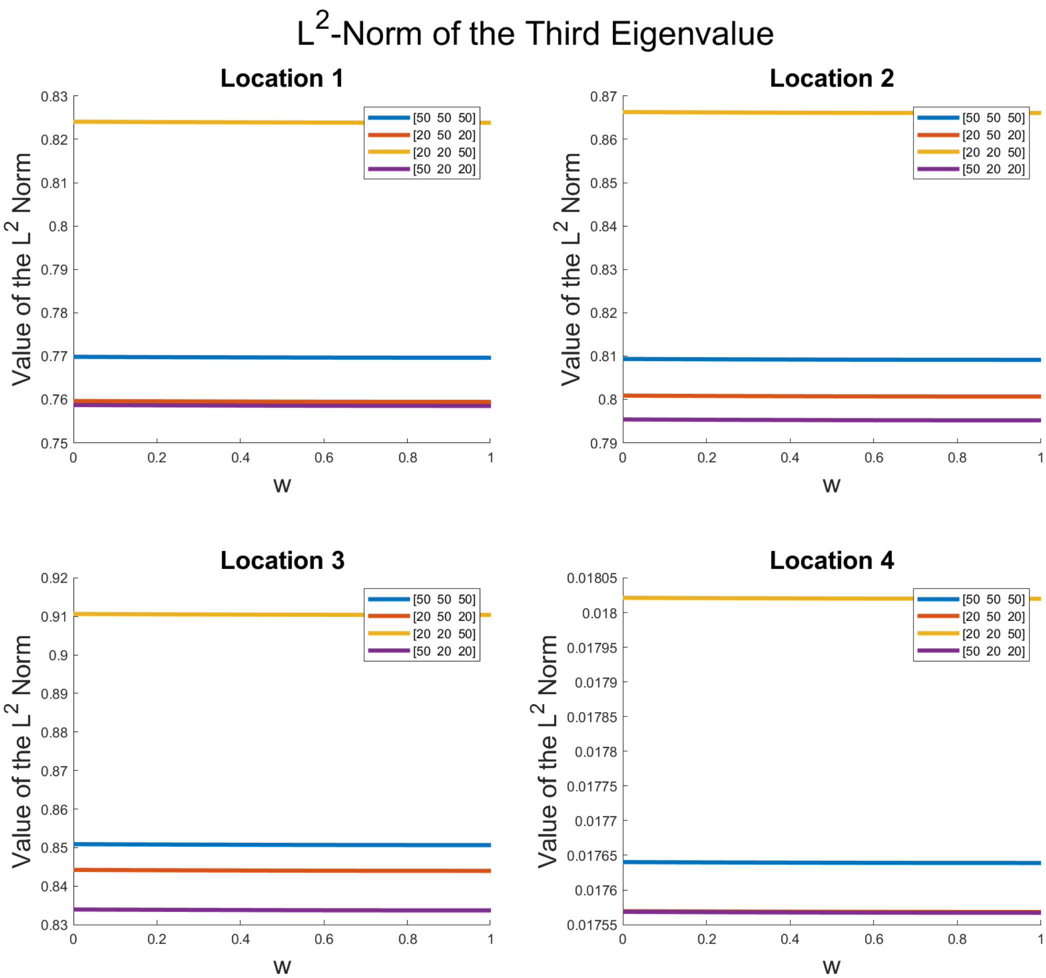


Figure 7.23: The L^2-norm of the third eigenvalues of the strain matrices of the data at (top left) Location 1, (top right) Location 2, (bottom left) Location 3, and (bottom right) Location 4 for S-wave orientation vectors $\mathbf{A} = [50, 50, 50]$ (blue), $\mathbf{A} = [20, 50, 20]$ (red), $\mathbf{A} = [20, 20, 50]$ (yellow), and $\mathbf{A} = [50, 20, 20]$ (purple).

In Figures 7.23 and 7.24, the L^2-norm of the third eigenvalues follows a similar pattern to norm of the first eigenvalues; however, the norms of the third eigenvalues are higher at all four locations than the norm of the first eigenvalues. While difficult to detect in the figures, the norm of the third eigenvalues decrease over the homotopy from the helically wound fibre to the straight fibre.

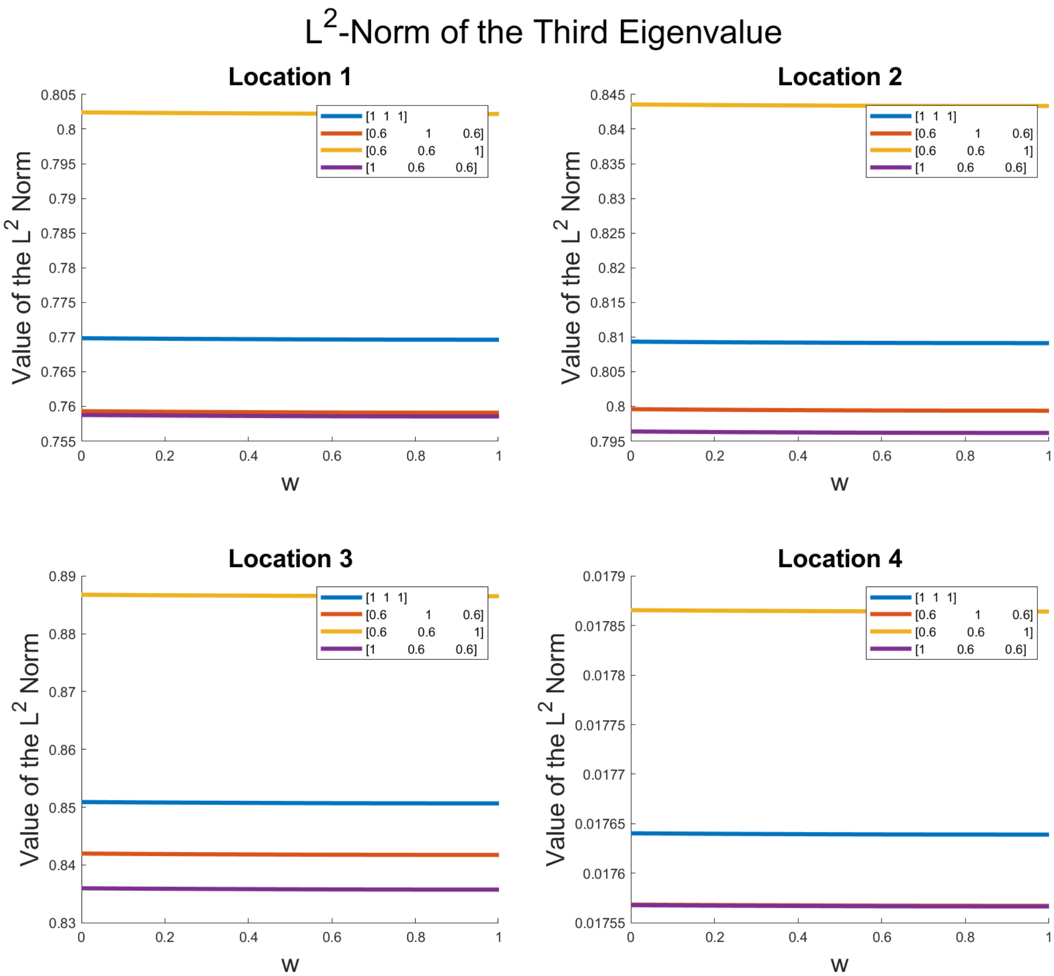


Figure 7.24: The L^2-norm of the third eigenvalues of the strain matrices of the data at (top left) Location 1, (top right) Location 2, (bottom left) Location 3, and (bottom right) Location 4 for S-wave orientation vectors $\mathbf{A} = [1, 1, 1]$ (blue), $\mathbf{A} = [0.6, 1, 0.6]$ (red), $\mathbf{A} = [0.6, 0.6, 1]$ (yellow), and $\mathbf{A} = [1, 0.6, 0.6]$ (purple).

While the norm of the data with the gauge length applied showed an increase from the helical fibre's response to the straight fibre's response across the homotopy, in every other comparison, we saw the helical fibre achieved a higher norm than the straight fibre. This suggests that the gauge length has a significant effect on the performance of the fibre. In fact, the gauge length appears to destroy a lot of information that the helically wound fibre holds. For this study, we used a gauge length of 10 meters. For future work, an investigation of how smaller and larger gauge lengths affect the helical fibre should be conducted. We should also examine how the gauge length affects the trace, determinants, and eigenvalues of the strain matrices at each point (s, t).

7.2.2 Homotopy comparison of three types of fibre formation

In Chapter 8, we will see how fibre-optic cables and distributed acoustic sensing are used to acquire vertical seismic profiles from well-boreholes. We motivate the model for this example based on this type of seismic acquisition. Specifically, we model our data based on fibre in a well borehole. A helix of 17.78cm radius depicts a fibre wrap around the outside of the borehole. We refer to it as the borehole helical fibre. The helix of radius 2.54cm represents a well borehole with a helical fibre cable laid along the well-borehole which we refer to as the cable helical fibre. The straight fibre models a straight fibre laid along the outside of the well-borehole. We simply call this case the straight fibre.

To obtain these values for these radii, we chose the values of w to be 0, 6/7, and 1, where $w = 0$ gives the borehole helical fibre, $w = 6/7$ gives the cable helical fibre, and $w = 1$ gives the straight fibre. The source is located 10 meters from the top of the fibre.

Figure 7.25 displays the results of the data for each fibre formation without the gauge length applied on the left and the results of the data for each fibre formation with the gauge length applied on the right. The first example had the fibre laid horizontally along the earth's surface. In this case, the fibre is laid vertically into the ground 200m deep. We only see half a hyperbola for the source instead of the whole hyperbolic signal from the last example. In

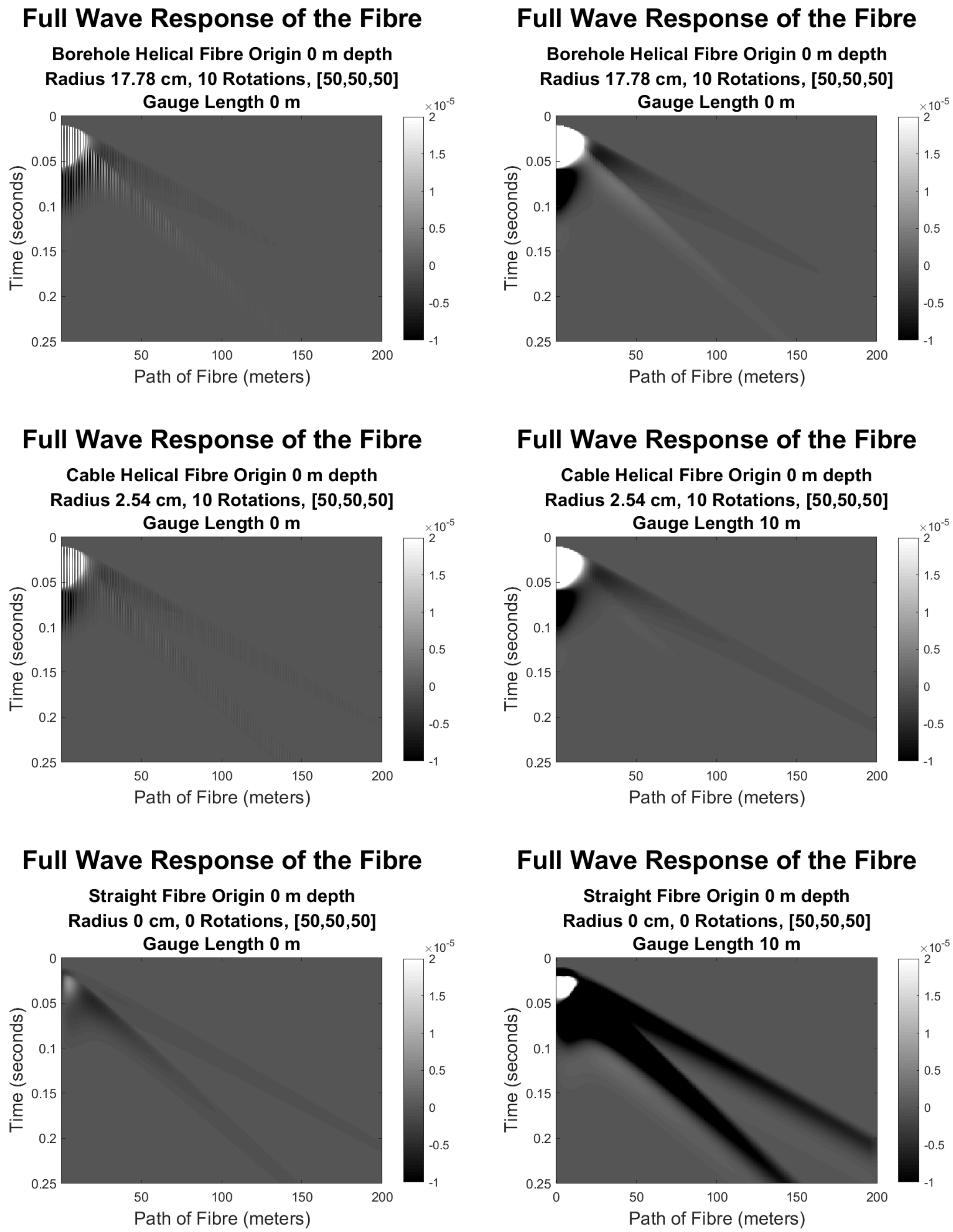


Figure 7.25: (Left column) The data without the gauge length applied (top) borehole helical fibre, (middle) cable helical fibre, and (bottom) helical fibre. (Right column) The 10m gauge length applied to data seen in the left column.

the top row of Figure 7.25, we find the results for the borehole helical fibre. We see some oscillations in the data on the left; however, once the gauge length is applied, the oscillation disappears. We also see oscillations in the cable helical fibre data before the gauge length is applied as seen in the middle row of Figure 7.25 on the right. Comparing the top and middle rows of Figure 7.25, the S-wave response is more evident in the data produced by borehole helical fibre, as opposed to the cable helical fibre.

As with the previous example, the signal detected by the straight fibre is more visible when the gauge length is applied than the signal detected by the helical fibres in the right column of Figure 7.25. Let us examine the trace, determinant, and eigenvalues for each S-wave orientation vector $\mathbf{A}$ and fibre formation to see how the results compare to the previous example.

Figure 7.26 presents the results of applying the L^2-norm over the homotopy of the data without the gauge length applied (left) and with the gauge length applied (right). For all eight S-wave orientation vectors $\mathbf{A}$, the norms of data with and without the gauge length applied decreases from the borehole helical fibre to the straight fibre. Unlike the horizontal fibre example, the helical fibre remains larger than the straight fibre for both the data with the gauge length applied and without the gauge length applied. This result supports the idea that the success of the fibre formation is heavily dependent on the location of the source with respect to the fibre.

Recall that the trace of the strain matrix depends only on the P-wave; see Equation 7.57. In Figure 7.27, the norm of the traces for all (s, t) does not change with the amplitude, which is what we expect. As with the previous method, we see that the trace decreases from the borehole helical fibre to the straight fibre.

Now consider the L^2-norm of the determinants for all S-wave orientation vectors $\mathbf{A}$. In Figure 7.28, we observe that the norm of the determinants decreases from the borehole helical fibre to the straight fibre. In this case, the norms of the determinants for the S-wave orientation vectors $\mathbf{A} = [20, 20, 50]$ and $\mathbf{A} = [0.6, 0.6, 1]$ have the largest value. Unlike the previous

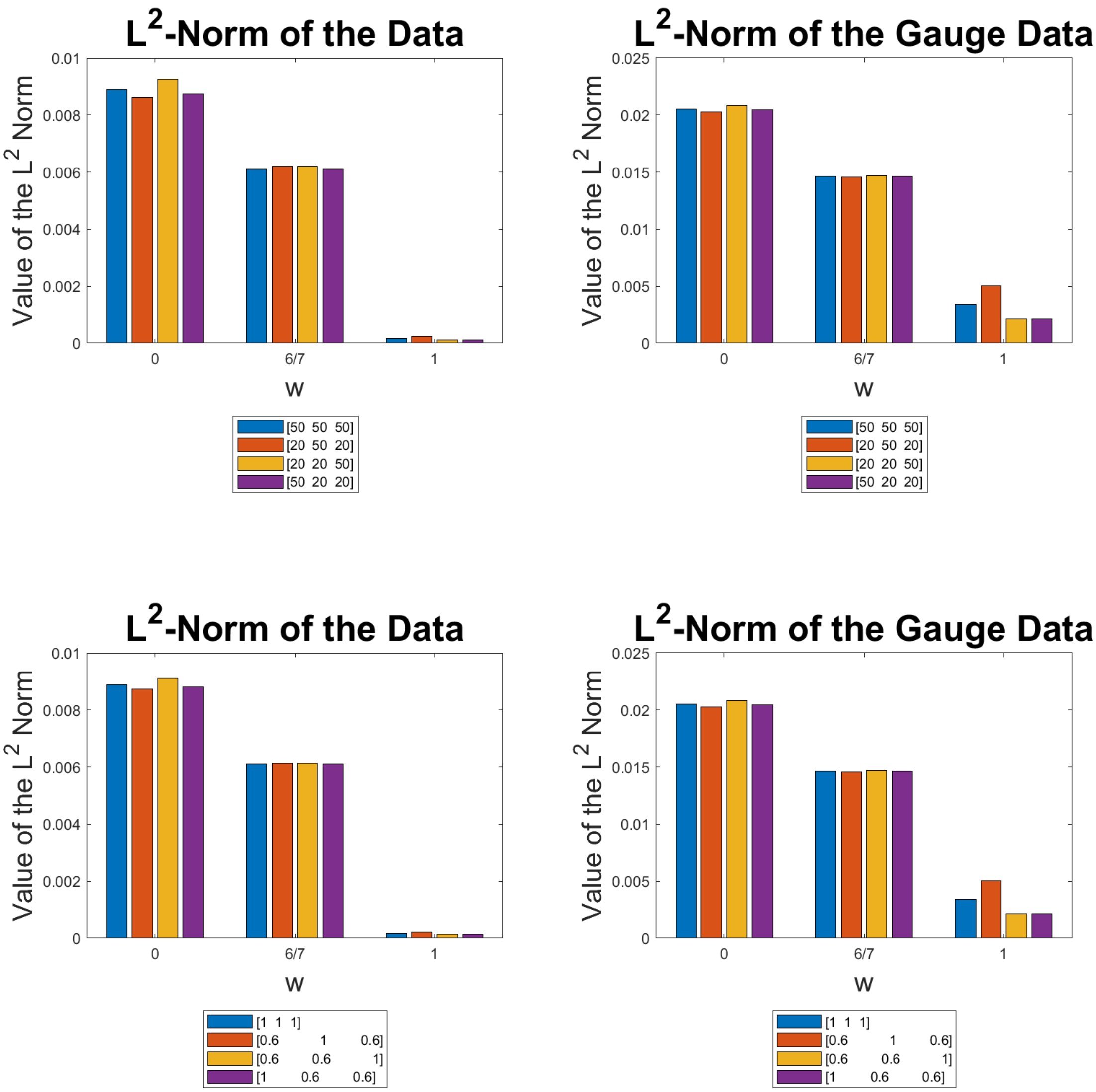


Figure 7.26: (Left column) The data without the gauge length applied (top) straight fibre and (bottom) helical fibre at Location 4. (Right column) The 10m gauge length applied to data seen in the left column.

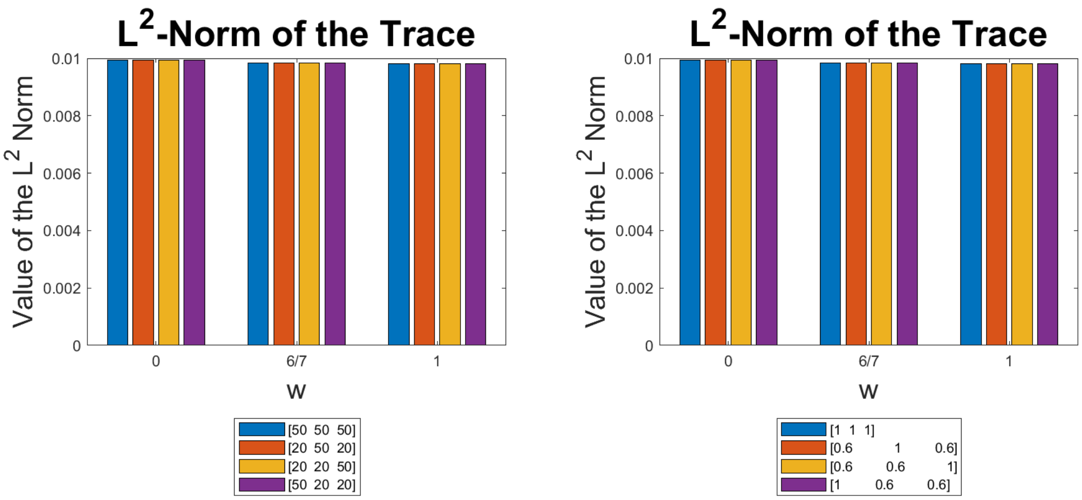


Figure 7.27: The L^2-norm of the traces of the strain matrices of the data for S-wave orientation vectors (left) $\mathbf{A} = [50, 50, 50]$ (blue), $\mathbf{A} = [20, 50, 20]$ (red), $\mathbf{A} = [20, 20, 50]$ (yellow), and $\mathbf{A} = [50, 20, 20]$ (purple) as ell as (right) $\mathbf{A} = [1, 1, 1]$ (blue), $\mathbf{A} = [0.6, 1, 0.6]$ (red), $\mathbf{A} = [0.6, 0.6, 1]$ (yellow), and $\mathbf{A} = [1, 0.6, 0.6]$ (purple).

example, the norm of the determinants for the S-wave orientation vectors $\mathbf{A} = [20, 50, 20]$ and $\mathbf{A} = [0.6, 1, 0.6]$ follows the norm of the determinants for the S-wave orientation vectors $\mathbf{A} = [20, 20, 50]$ and $\mathbf{A} = [0.6, 0.6, 1]$, respectively, suggesting that the source's orientation affects the vertical fibre similarly for these specific S-wave orientation vectors $\mathbf{A}$.

For the norm of the eigenvalues, the L^2-norm decreases from the borehole helical fibre to the straight fibre for all three eigenvalues. The norms for the S-wave orientation vector $\mathbf{A} = [20, 50, 20]$ and $\mathbf{A} = [20, 20, 50]$ follow a similar path versus the previous example when the S-wave orientation vectors $\mathbf{A} = [20, 50, 20]$ and $\mathbf{A} = [50, 20, 20]$ followed a path similar to the norms of the determinants; see Figures 7.29 – 7.31. For both examples, the norm of the second eigenvalues appears to be the same for all the S-wave orientation vectors; see Figure 7.30.

It is interesting to see similarities to the previous example as well as where they differ. In particular, it shows how much of an effect the S-wave orientation vector $\mathbf{A}$ in combination with the location of the source, and the location and formation of the fibre has on the data.

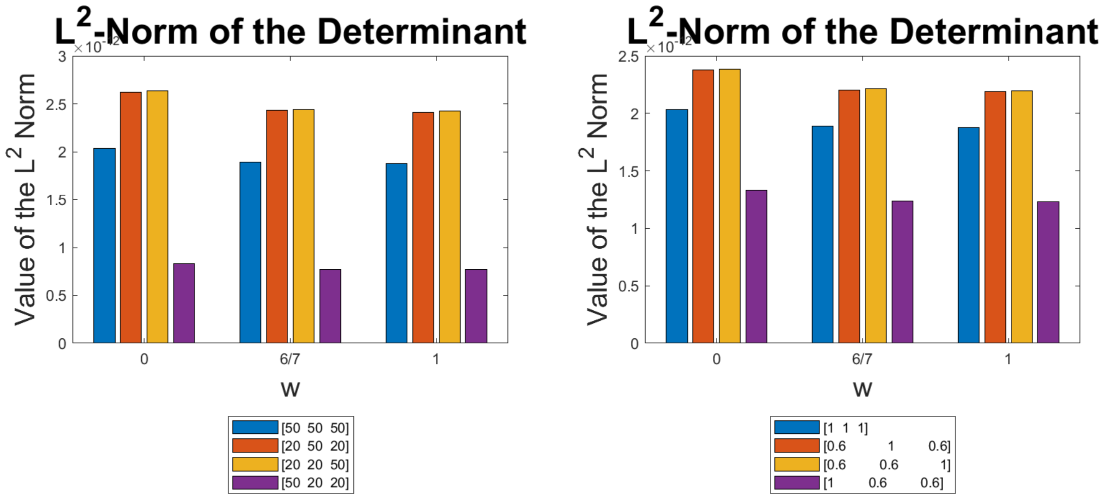


Figure 7.28: The L^2-norm of the determinants of the strain matrices of the data for S-wave orientation vectors (left) $\mathbf{A} = [50, 50, 50]$ (blue), $\mathbf{A} = [20, 50, 20]$ (red), $\mathbf{A} = [20, 20, 50]$ (yellow), and $\mathbf{A} = [50, 20, 20]$ (purple) as ell as (right) $\mathbf{A} = [1, 1, 1]$ (blue), $\mathbf{A} = [0.6, 1, 0.6]$ (red), $\mathbf{A} = [0.6, 0.6, 1]$ (yellow), and $\mathbf{A} = [1, 0.6, 0.6]$ (purple).

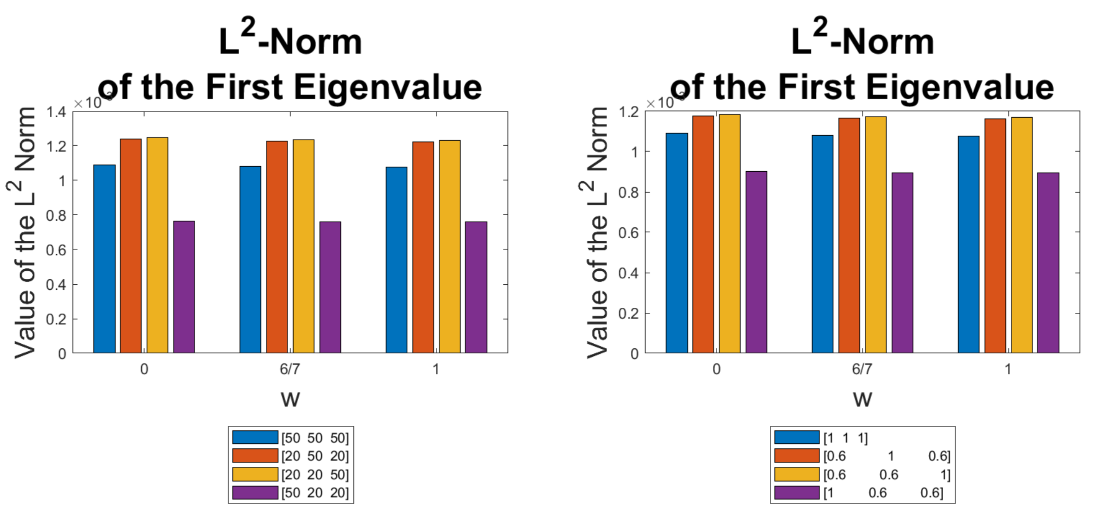


Figure 7.29: The L^2-norm of the first eigenvalues of the strain matrices of the data for S-wave orientation vectors (left) $A = [50, 50, 50]$ (blue), $A = [20, 50, 20]$ (red), $A = [20, 20, 50]$ (yellow), and $A = [50, 20, 20]$ (purple) as ell as (right) $A = [1, 1, 1]$ (blue), $A = [0.6, 1, 0.6]$ (red), $A = [0.6, 0.6, 1]$ (yellow), and $A = [1, 0.6, 0.6]$ (purple).

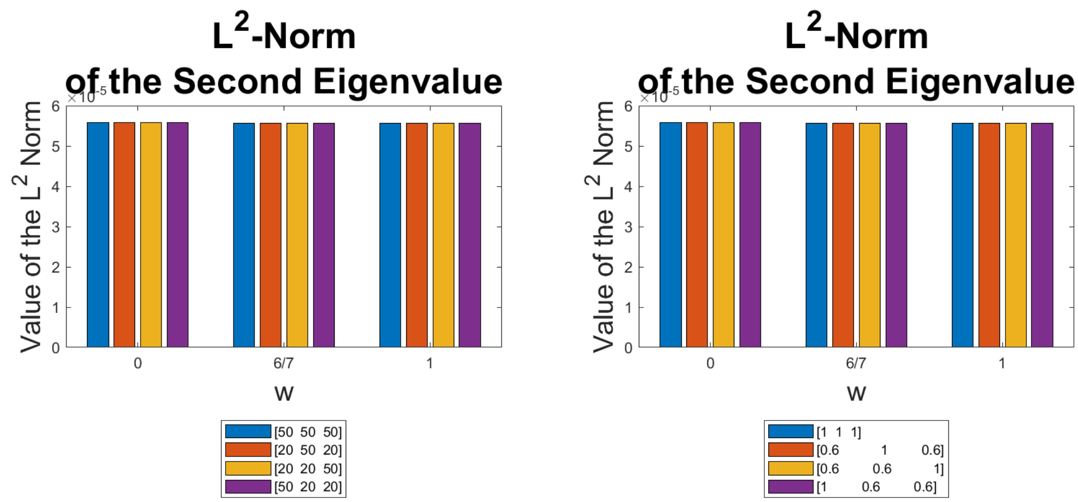


Figure 7.30: The L^2-norm of the second eigenvalues of the strain matrices of the data for S-wave orientation vectors (left) $A = [50, 50, 50]$ (blue), $A = [20, 50, 20]$ (red), $A = [20, 20, 50]$ (yellow), and $A = [50, 20, 20]$ (purple) as ell as (right) $A = [1, 1, 1]$ (blue), $A = [0.6, 1, 0.6]$ (red), $A = [0.6, 0.6, 1]$ (yellow), and $A = [1, 0.6, 0.6]$ (purple).

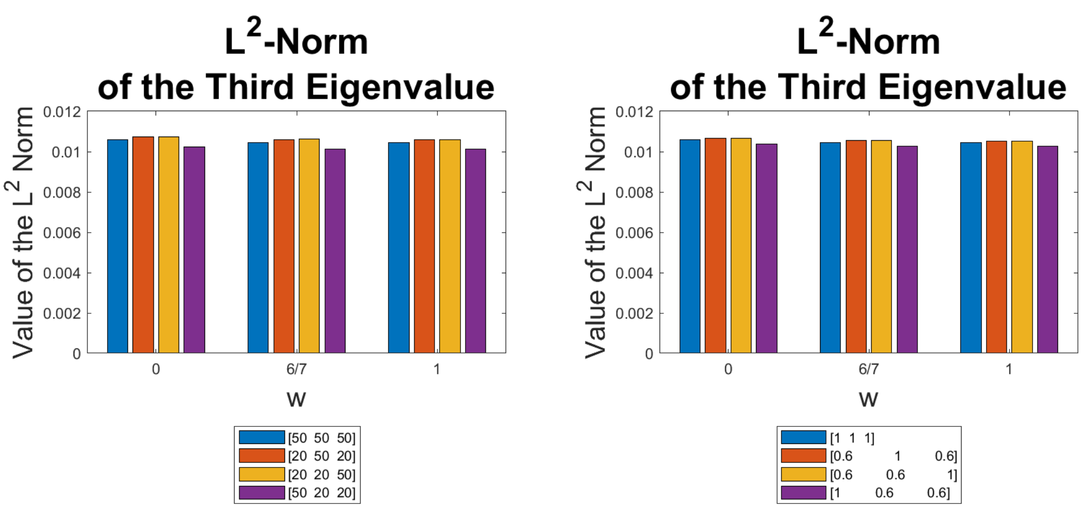


Figure 7.31: The L^2-norm of the third eigenvalues of the strain matrices of the data for S-wave orientation vectors (left) $A = [50, 50, 50]$ (blue), $A = [20, 50, 20]$ (red), $A = [20, 20, 50]$ (yellow), and $A = [50, 20, 20]$ (purple) as ell as (right) $A = [1, 1, 1]$ (blue), $A = [0.6, 1, 0.6]$ (red), $A = [0.6, 0.6, 1]$ (yellow), and $A = [1, 0.6, 0.6]$ (purple).

Here we saw that the borehole helical fibre consistently provided a stronger signal than the other fibre formations. Comparing the vertical cable helical fibre to the horizontal cable helical fibre, we found that the gauge length amplifies the vertical helical more than the vertical straight whereas the opposite occurs for the horizontal case. For the vertical case, the norm consistently decreases as the borehole fibre deforms to the straight fibre. It would be interesting to see how different gauge lengths and source locations affect these results. Specifically, we would want to check if the helical fibre would still perform better than the straight fibre with these changes in the model.

7.3 Conclusions

We found bounds for the amplitude response of the fibre, both when the gauge length is applied and when it is not. We also showed that the amplitude response of the helical fibre is homotopic to the straight fibre by building homotopies between the path of the helical and straight fibres, the strain tensor, and the tangent path of the helical and straight fibres. The existences of this homotopy between the two fibre formations provides mathematical support for comparing the response of each formation in a distributed acoustic sensor. As such, we used the homotopy for two examples to compare the straight and helical fibre. In the first example, we looked at the results of a helical fibre of radius 2.54cm as it deforms into straight fibre. The norms of the data when the gauge length was applied showed the norm increasing from the helical fibre to the straight fibre along the homotopy; however, for the L^2-norm of the data when the gauge length was not applied as well as the L^2-norms of the traces, determinants, and eigenvalues, the L^2-norm decreased over the homotopy from the helical fibre to the straight fibre. We then employed a homotopy to compare the results of a borehole helical fibre, cable helical fibre, and straight fibre laid vertically in the earth. In this example, the L^2-norm always decreased from the borehole helical fibre to the straight fibre for all comparisons.

For both examples of homotopy comparisons of the helical and straight fibre, it is interesting to see that the L^2-norm of the traces, determinants, and the eigenvalues showed the norms decreasing over the homotopy from the borehole helical fibre to the straight fibre. It was only when we applied the gauge length to the data that we saw a difference in the results for the two cases. For the horizontal fibre, the norm of the data with the gauge length applied increased from the helical fibre of radius 2.54cm to the straight fibre; however, for the vertical fibre, the norm of the data with the gauge length applied decreased from the borehole helical fibre to the straight fibre. This result suggests that the success of the fibre formations is highly dependent on where the source is located with respect to the fibre. Further investigation of the vertical fibre with different locations for the source is merited. Finding other means besides the norm to compare the helical and straight fibre over the homotopy is also worth investigating.

Chapter 8

Vertical Seismic Profiles using Distributed Acoustic Sensing

Interest in acquiring geophysical data using fibre-optic cables has increased over the last several years. While several techniques are available for data acquisition using fiber-optic cables, we focus on the use of distributed acoustic sensing (DAS) in this chapter. We apply this process to acquiring vertical sesimic profiles in boreholes located at the Field Research Station of the Containment and Monitoring Institute (CaMI) site in Newell County, AB.

The Containment and Monitoring Institute is a research station which focuses on subsurface monitoring (CMC Research Institutes, 2015). Specifically, at CaMI, they study methods for the containment of carbon and other subsurface fluids. Such studies require monitoring which can be done by considering the veritcal seismic profiles of the two wells at the site. When using geophones to acquire vertical seismic profiles in wells, several sweeps must be done in order to cover the entire length of the well (Stewart, 2001). Typically a tool with 5 to 80 geophones is moved from the bottom of the well to the surface. A source is denoted on the surface whenever it reaches a new location in the well borehole. The extent to which the well is covered depends on funding and time-constraints of such a project.

Distributed acoustic sensing is often used for monitoring purposes. Given that fibre optic cables are cheap and durable, they can be installed permanently in well-boreholes to monitor CO_2 storage. The fibre-optic cable can cover the entire length of the well. Using a fibre-optic cable as the receiver provides more receiver points along the depth of the well, unlike the discrete number of geophones that may be used in an experiment. The two observation wells at the Containment and Monitoring Institute's Field Research Station also have some geophones permenantly installed in the wells (Gordon and Lawton, 2018). The authors of Gordon and Lawton (2018) compare the results of a walk-away VSP using the fibre-optic cables set up in the wells to the results of the geophones.

In this chapter, we study how successful using distributed acoustic sensing acquired data is for monitoring the location of the future CO_2 plume at the CaMI Field Research Station in Newell County, AB when applying standard vertical seismic profile processing techniques. We discuss the schematic of the fibre at the CaMI Field Research Station. We then consider two experiments: one conducted in 2017 and another done in 2018. We find that the results in the 2017 experiment to be less promising when standard processing techniques are applied; however, it is still possible to identify the caprock used for storing the CO_2. Some extra processing is required. The 2017 case study employed fewer source locations than the 2018 study; the 2018 experiment provides better results with regards to the caprock being much more identifiable after standard processing.

8.1 Containment and Monitoring Institute Field Research Station in Newell County, AB

Figure 8.1 shows a full-sweep of the fibre at the site from work done during the second experiment we will discuss in this chapter. The source is located at the center of the trench, which allows us to see the full trench and both wells. In comparison to the data in Hardeman et al. (2017), the data was collected along the opposite direction of the fibre loop. From left to right in Figure 8.1, we see the results from the trench to Well 1, to the straight fibre in

Well 2, to the helically wound fibre in Well 2.

We provide a schematic of the Field Research Station's fibre loop in Figure 8.2. In the schematic, the straight fibre is designated by the green lines and the helical fibre is represented by the blue lines. The orange lines depict the connecting fibre. The full schematic begins with the fibre leaving the shed and connecting to the straight fibre for half the length of the trench. The fibre then goes along the trench helically for 1.1 kilometers and then returns halfway along the trench straight. Afterwards, the fibre connects to Well 1 where it descends into the earth straight for approximately 300 meters before returning the surface. It proceeds to Well 2 where it goes down into the Well and returns to the surface as straight fibre. Then it descends again into Well 2, helically wound, before returning to the surface, at which point it returns to the shed.

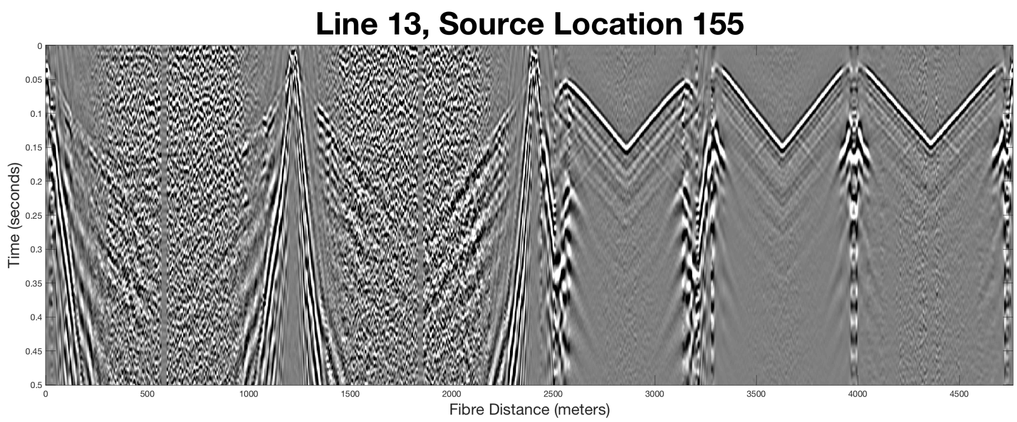


Figure 8.1: Full fibre data for Line 13 Flag 155

8.2 Examples from 2017 acquisition

In 2017, Fotech Solutions utilized their interrogator to collect DAS data during an experiment conducted by CaMI in Newell County, AB. In the next subsection, we discuss specifics of the experiment. We show a full sweep of the fibre at a source location near the center of the trench. Afterwards, we consider the results from the straight fibre in Wells 1 and 2. Finally,

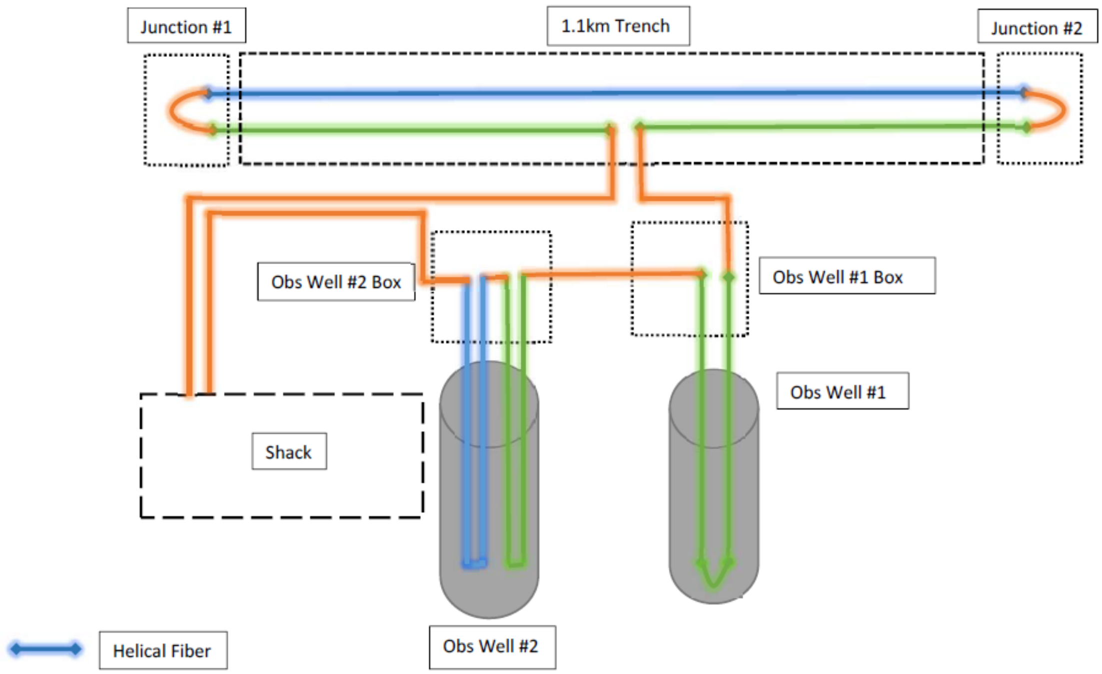


Figure 8.2: A schematic of the fibre at the site in Newell County, AB.

we employ Source Location (SL) 103 to process the vertical seismic profiles for both wells.

8.2.1 The 2017 experiment

The experiment consisted of 270 shots over 49 locations along two full lines. There were five source locations, numbered 101 to 105, along the length of the trench on Line 35, and 44 locations along a a line intersecting Well 1 on Line 23. Only Line 35 is considered in this chapter. The Source Location 103 resided between Wells 1 and 2 and was approximately 500 meters from source locations 101 and 105. The wells reach a depth of approximately 300 meters. Processing is applied to the raw backscatter data to obtain the optical phase, and then each shot is cross-correlated with the pilot sweep and stacked. There are 10 shots per stack.

Figure 8.3 shows the resulting data when the vibroseis truck was at flag 103 for the full fibre. From approximately 0 to 750 meters, the helical fibre in Well 2 senses much more

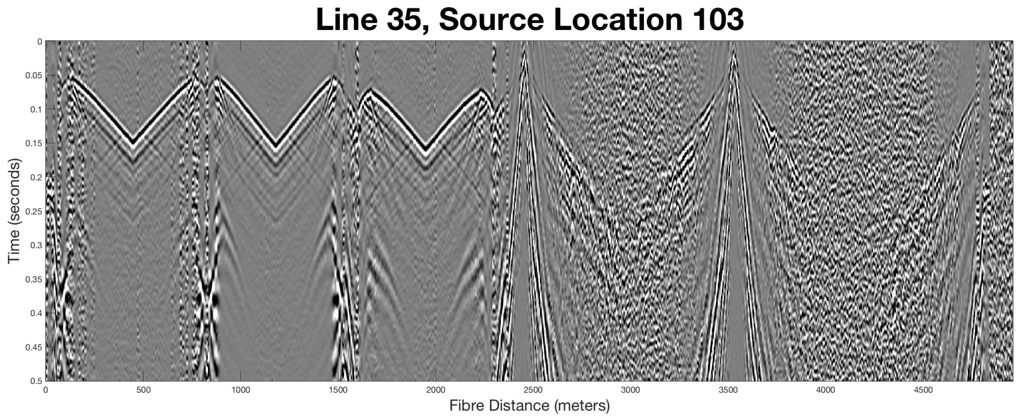


Figure 8.3: Full fibre data for Line 35 Flag 103

of the surface noise than the straight fibres in Wells 1 and 2, which go from around 800 to 2400 meters. On the day of acquisition, there was an extreme wind warning in the area, which likely accounts for the source of the noise. The reason why the helical fibre appears more sensitive to the wind noise than the straight fibre is unknown, but it could be due to differences in the coupling between the two fibres and the well. Counter-intuitively perhaps, it seems as though the helical fibre is less sensitive to the vibe signal both in the wells and along the trench.

Figures 8.4–8.6 show the results at each source location of Well 1 and Well 2. In this case, it focuses on the fibre which goes straight down Well 2 before returning to the surface, moving to Well 1, and extending down the straight fibre in Well 1 before returning. From these images, we can clearly see the P and S wave response of the fibre as well as AVO effects and the position of several prominent reflectors. The reflector located at approximately 0.25 seconds is of particular interest since it is the caprock where the CO_2 will be stored.

8.2.2 Vertical seismic profiles for Well 1

We start with the vertical seismic profiles for Well 1 using Source Location 103. Figure 8.7 (left) shows Well 1 when the vibroseis truck was at source location 103. Since the fibre goes down and back up the Well, we fold the well data in half, which can be seen on the right in

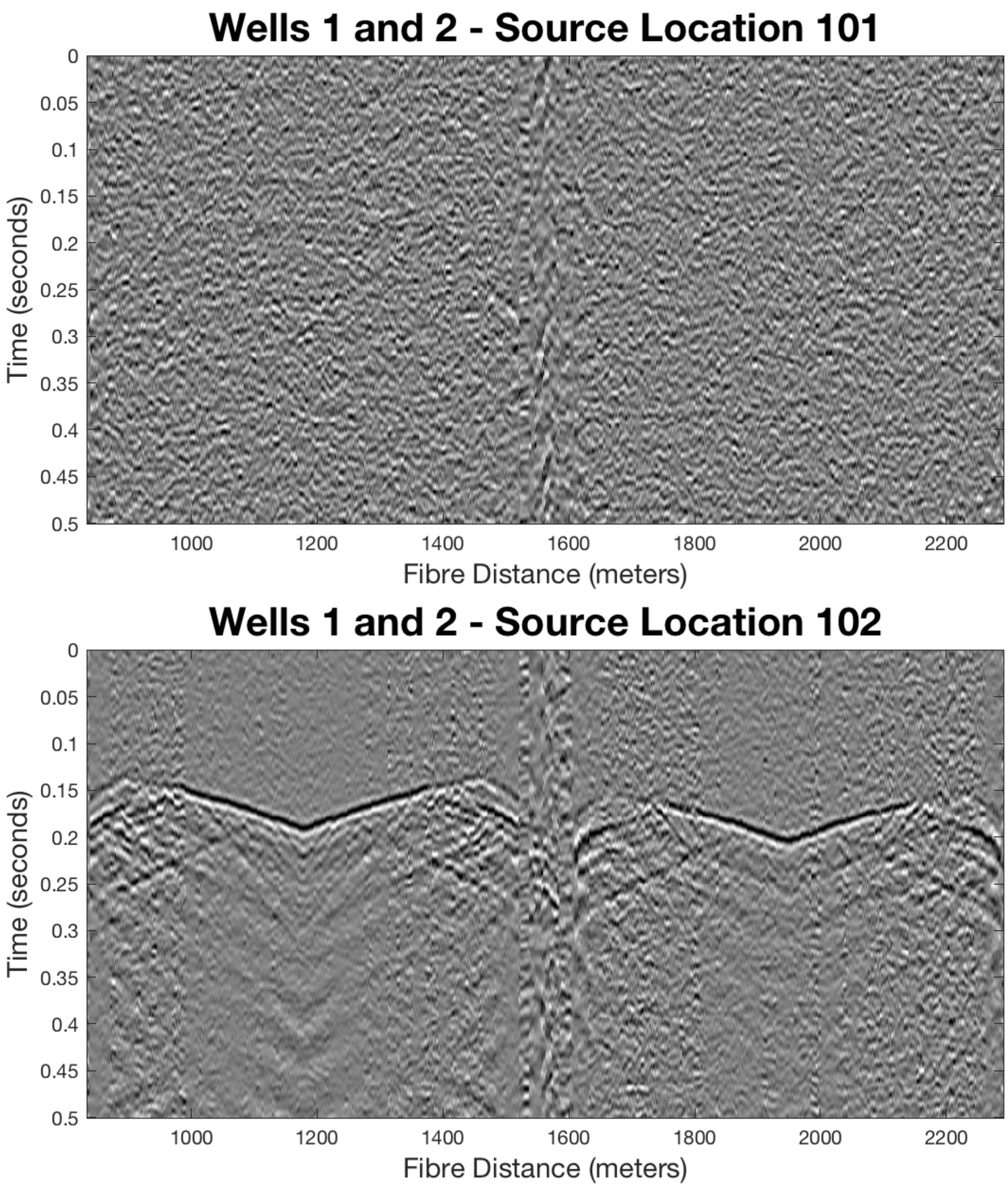


Figure 8.4: The straight-fibre from Well 2 to the straight fibre in Well 1 acquired when the vibroseis truck was at Source Locations 101 (top) and 102 (bottom).

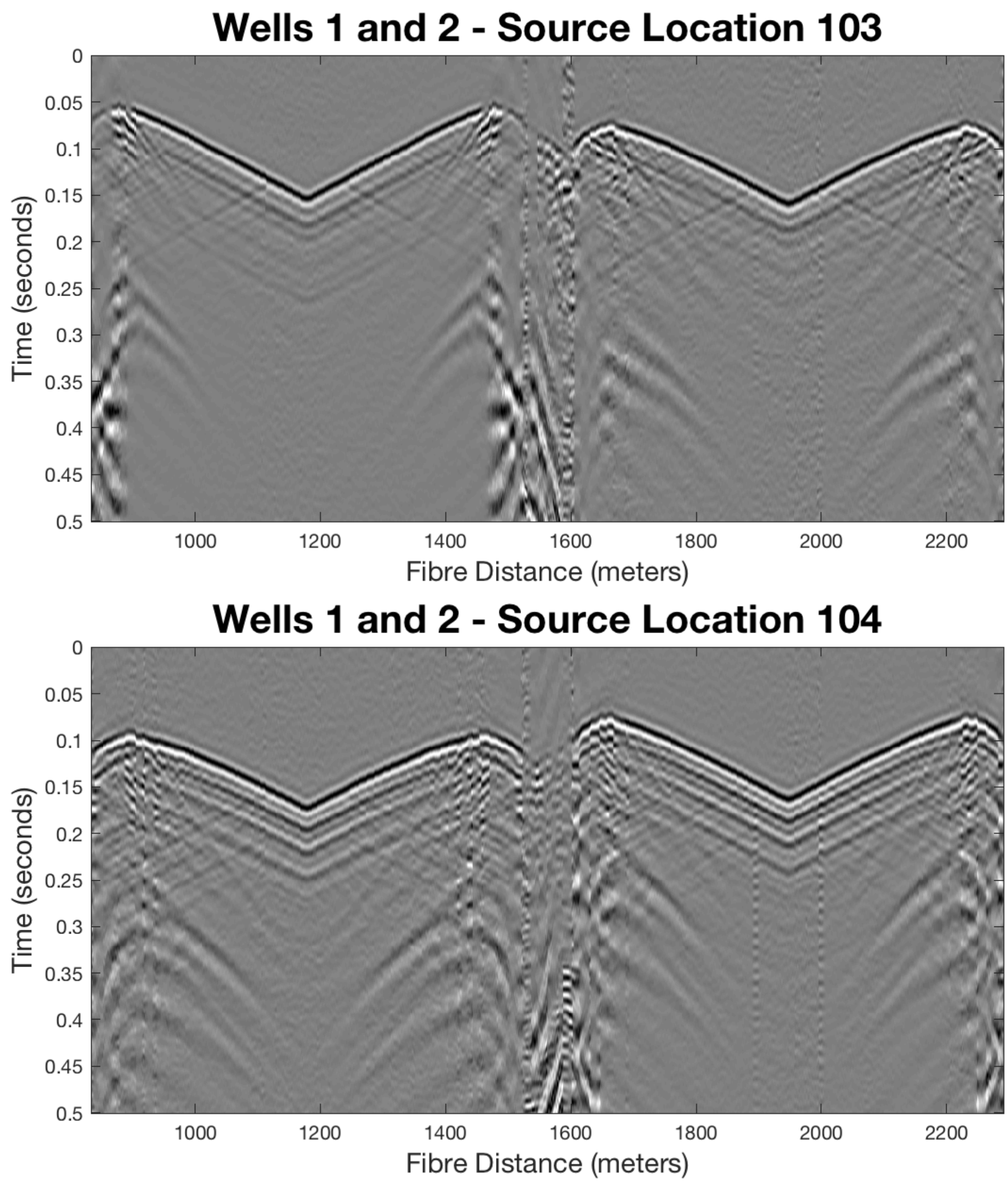


Figure 8.5: The straight-fibre from Well 2 to the straight fibre in Well 1 acquired when the vibroseis truck was at Source Locations 103(top) and 104 (bottom).

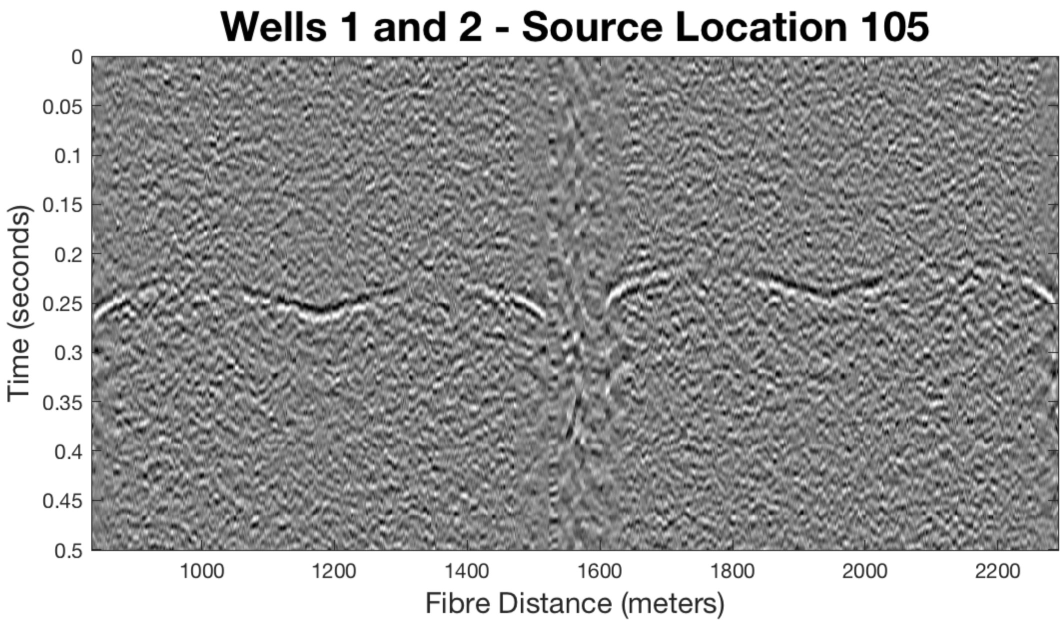


Figure 8.6: The straight-fibre from Well 2 to the straight fibre in Well 1 acquired when the vibroseis truck was at Source Location 105.

Figure 8.7, and then apply standard VSP processing techniques to the folded data.

We separate the upgoing and downgoing wavefields using an $f - k$ filter. The upgoing wavefield contains reflections from anomalies near the well including the caprock. Figure 8.8 shows the upgoing and downgoing wavefields.

We employ the CREWES public toolbox to find the first break pick and use it to flatten the upgoing wavefield for Well 1 at Source Location 103 (Margrave, 2018). Flattening the upgoing wavefield shows where the reflections occur in time. Figure 8.9 depicts the first break pick on the right and the flattened upgoing wavefield on the left. We observe a prominent reflector at 0.25 seconds. This is the caprock above where the CO_2 is stored. Its prominence is promising for monitoring purposes.

Finally, we deconvolve the upgoing wavefield using the first break pick. Applying the deconvolution to the data is meant to improve the image's resolution. Figure 8.10 provides the results of the upgoing wavefield (left) next to the results of the deconvolved upgoing wavefield (right). The caprock reflection near 0.25 seconds is more distinguishable in the

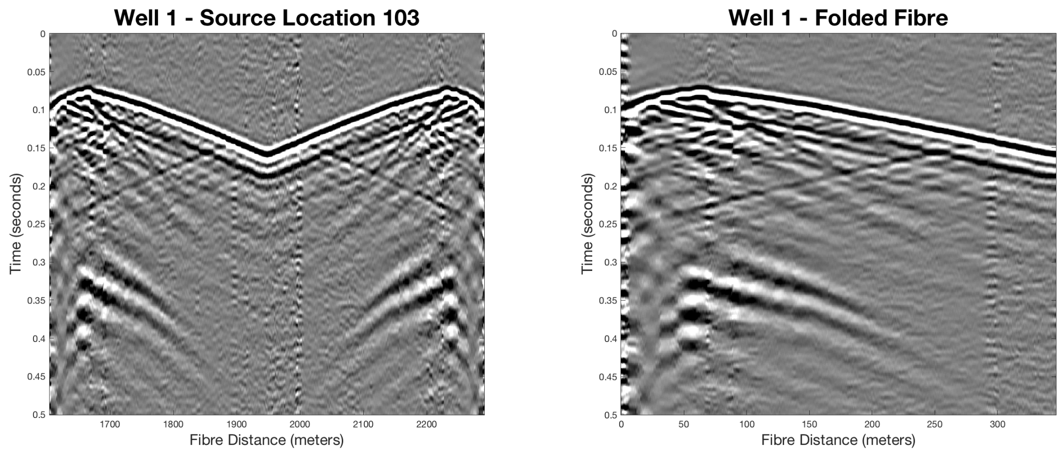


Figure 8.7: (Left) Well 1 when the vibroseis truck is at Source Location 103. (Right) The straight fibre in Well 1 folded.

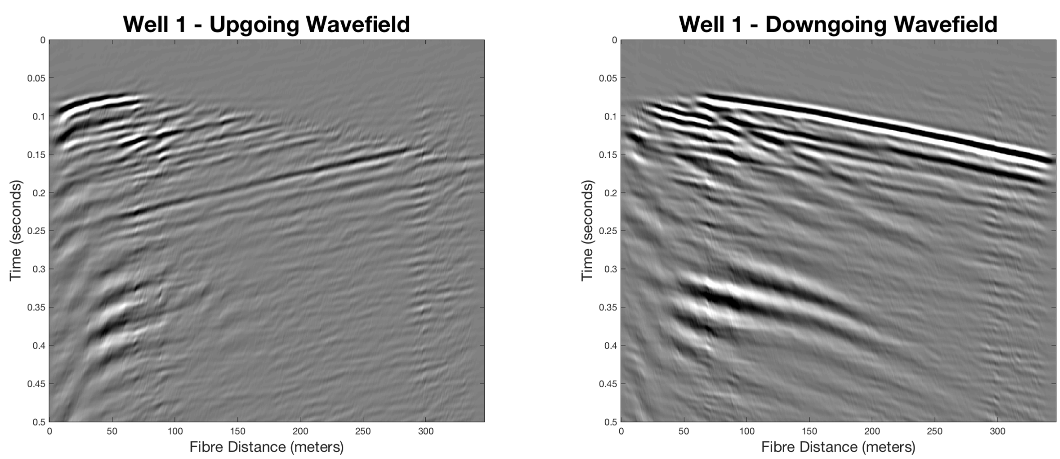


Figure 8.8: (Left) The upgoing wavefield for Well 1. (Right) The downgoing wavefield for Well 1.

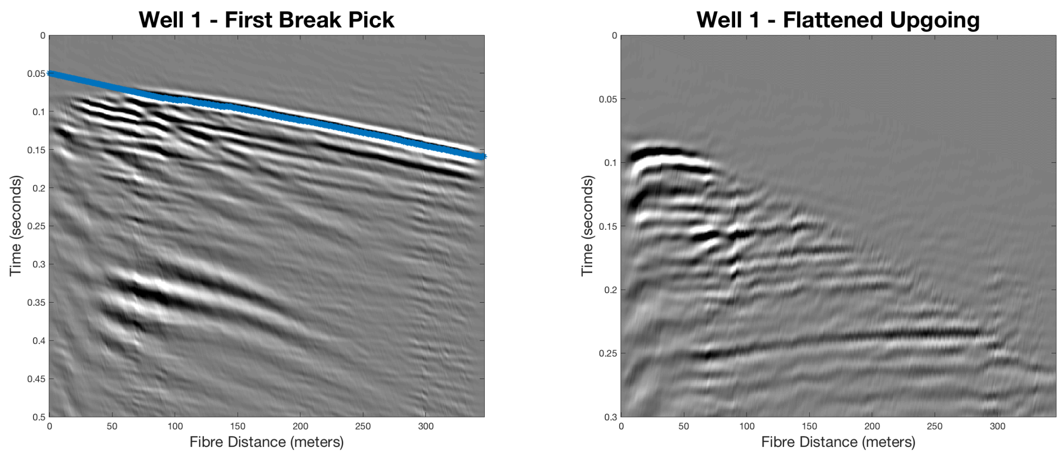


Figure 8.9: (Left) The first-break picks for Well 1 at Source Location 103. (Right) The flattened upgoing wavefield for Well 1.

deconvolved upgoing wavefield than in the unprocessed upgoing wavefield.

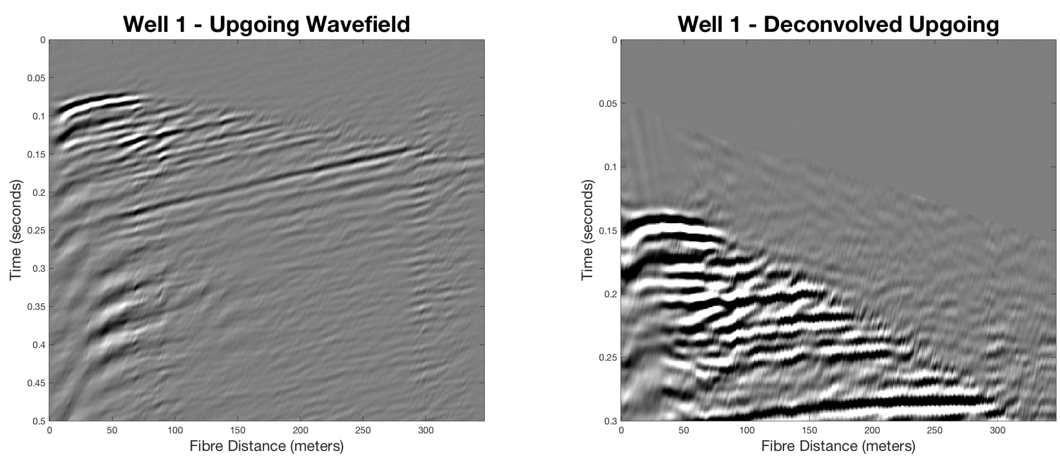


Figure 8.10: (Left) The upgoing wavefield for Well 1. (Right) The deconvolved upgoing wavefield for Well 1.

8.2.3 Vertical seismic profiles for Well 2

We also utilize Source Location 103 to process the vertical seismic profiles for Well 2. Figure 8.11 shows the results of the Well 2 at Source Location 103 on the left. We repeat the process of folding the data in half, as seen on the right in Figure 8.11, which is what we did for the data from Well 1. We then use the folded fibre to compute the vertical seismic profile for the well.

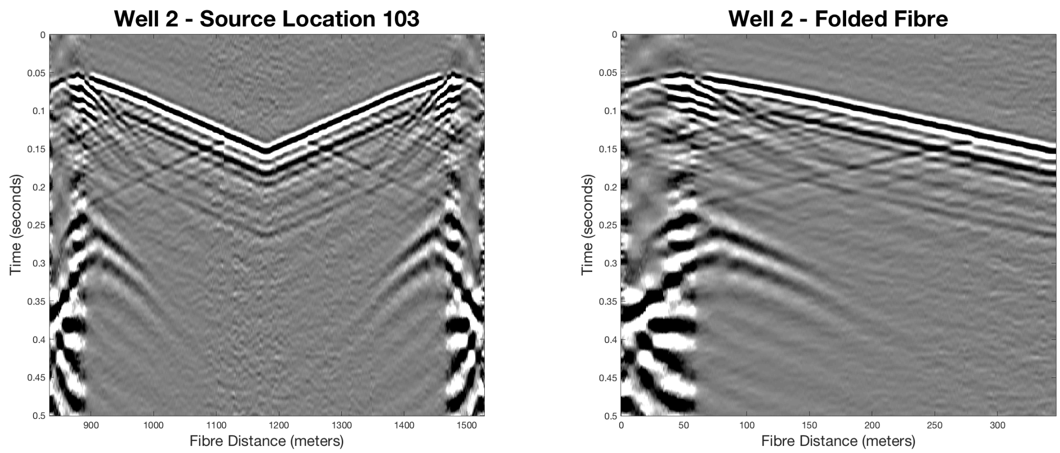


Figure 8.11: (Left) Well 2 when the Vibroseis truck is at Source Location 103. (Right) The straight fibre in Well 2 folded.

As with Well 1, we begin by computing the upgoing and downgoing wavefields for Well 2. We provide the results of this computation in Figure 8.12.

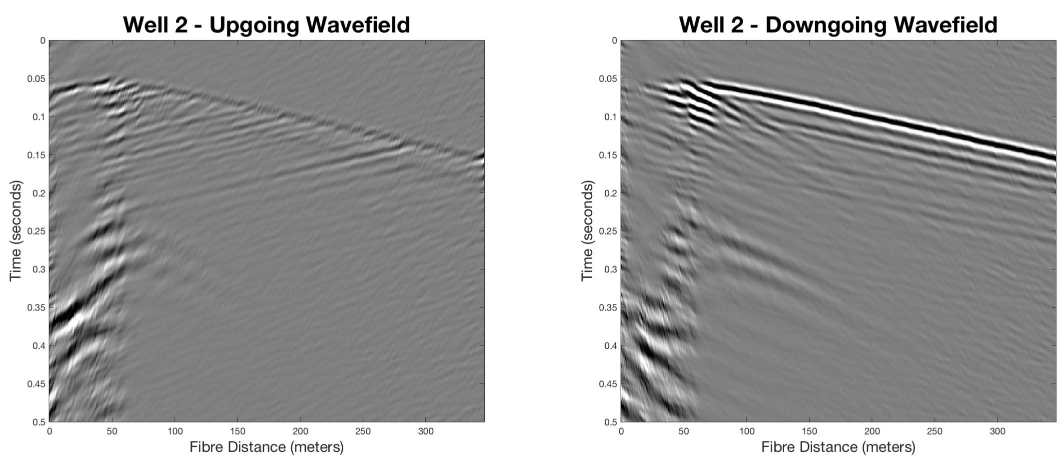


Figure 8.12: (Left) The upgoing wavefield for Well 2. (Right) The downgoing wavefield for Well 2.

After completing these computations, we chose the first break which can be seen on the left in Figure 8.13. Using the first break pick, we flatten the upgoing wavefield using the

CREWES public toolbox; see Figure 8.13 (right). While less prominent in Well 2, we can still distinguish the caprock at 0.25 seconds.

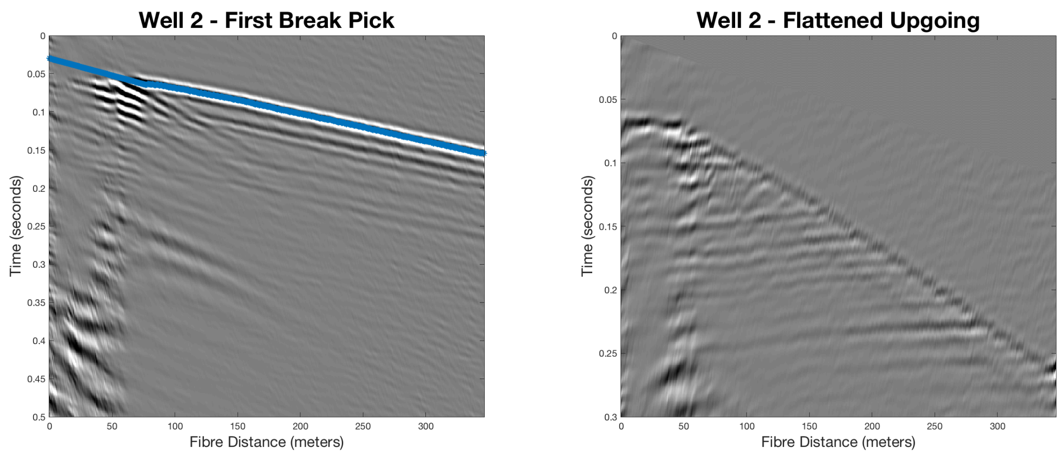


Figure 8.13: (Left) The first-break picks for Well 2 at source location 103. (Right) The flattened upgoing wavefield for Well 2.

We also utilize the CREWES public toolbox to deconvolve the upgoing wavefield (Margrave, 2018). In Figure 8.14, the upgoing wavefield is on the left and the deconvolved upgoing wavefield is on the right.

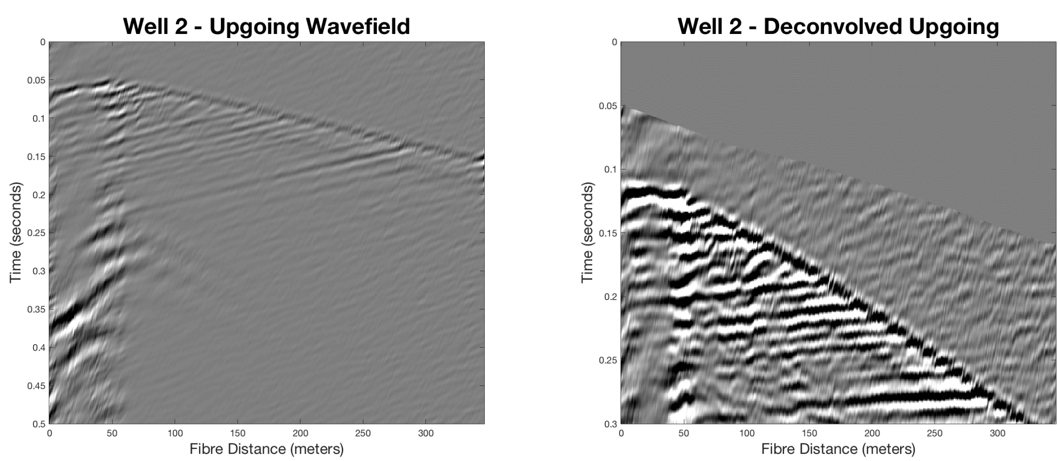


Figure 8.14: (Left) The upgoing wavefield for Well 2. (Right) The deconvolved upgoing wavefield for Well 2.

Before continuing to the next experiment, notice that deconvolving the upgoing wavefield for both wells saw the reflectors shift from their locations in the flattend upgoing wavefield. In particular, the caprock reflector at 0.25 seconds shifted to 0.3 seconds in these two cases. This shows that deconvolution is not always helpful when processing vertical seismic profiles. Further processing is required in this case for interpretation purposes. A far-offset deconvolution might provide better results. Due to time constraints, we have not done further processing on this data set at this time. Note that we may still compare the results from the flattened upgoing wavefields in this experiment and also consider how the deconvolution works with a source location that is closer to the well in the next experiment.

8.3 Examples from the 2018 acquisition

In August 2018, another experiment was conducted at the CaMI Field Research Station in Newell County, AB. Fotech Solutions acquired seismic data during the experiment using their interrogator and the installed fibre-optic cable system at the site. In this section, we provide the results of the straight fibre in the two wells. We first explain the experiment conducted in August 2018, and we also show the results of the full fibre at one source location. Afterwards, we look at the results of the straight fibre in Wells 1 and 2 for eleven source locations. Finally, we process the zero-offset VSP for Well 1 and Well 2.

8.3.1 The 2018 experiment

For the experiment, a Vibroseis truck generated 110 source locations, numbered 101 to 210, along the 1.1 kilometers of the trench. The source locations were approximately 10 meters apart. A total of 448 shots were acquired on one full line, Line 13. The raw backscattered data was processed to obtain the optical phase. Each shot was cross-correlated with the pilot sweep and then stacked. Each stack consisted of 4 shots.

Figure 8.15 shows an aerial view of the site with each of the source locations along the trench marked as well as the two wells. The red dots signify the source locations we will consider in the following section to show what occurs in the wells as the source location is moved along the trench. The blue dots represent Well 1 and Well 2, and the blue balloons show the source locations used to process the vertical seismic profiles for each well.

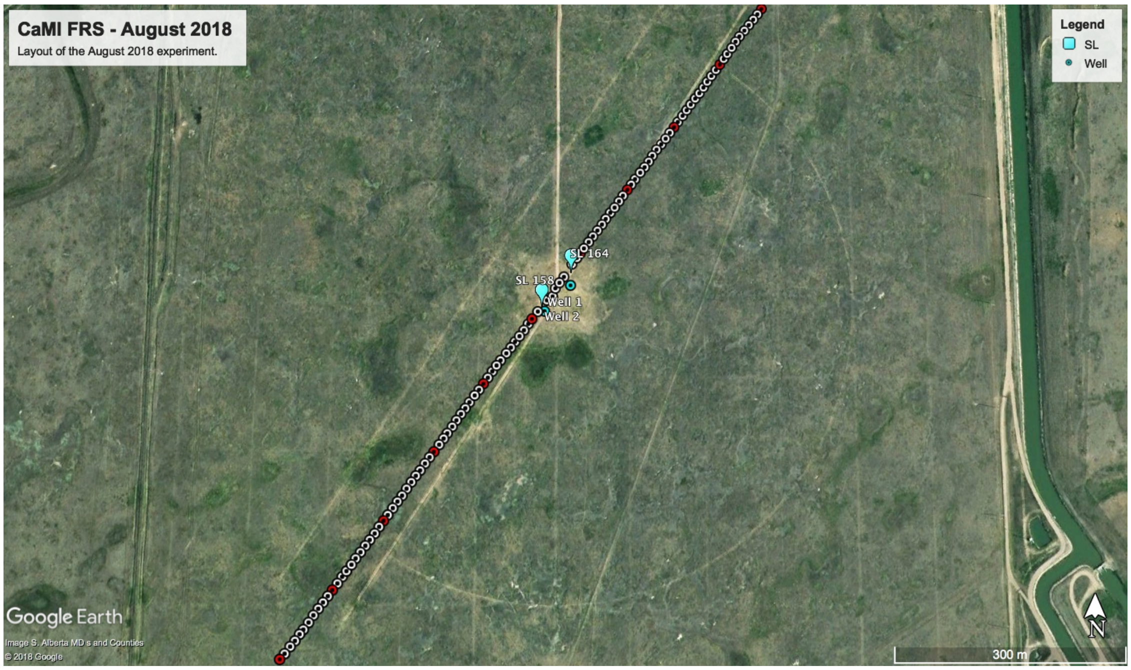


Figure 8.15: A aerial view of the experiment at the site in Newell County, AB.

In this experiment, the interrogator was connected to the opposite side of the fibre. From Figure 8.1, we observe that the fibre leaves the shack and goes to the half-trench before moving to the helically wound fibre along the full trench. It then returns up the straight fibre in the half trench before connecting to the straight fibre in Well 1. The fibre goes down the well before returning to the surface and going to Well 2. In the second well, the fibre descends twice into the well, first as straight fibre, and then as a helically wound fibre, before it returns to the shack.

Figures 8.16 to 8.21 show the straight fibre in Wells 1 and 2 at eleven different source locations along the trench. In particular, we consider the results for source locations 101, 112, 123, 134, 145, 156, 167, 178, 189, 200, and 210. The eleven source locations were chosen so that consecutive flags were approximately 110 meters apart.

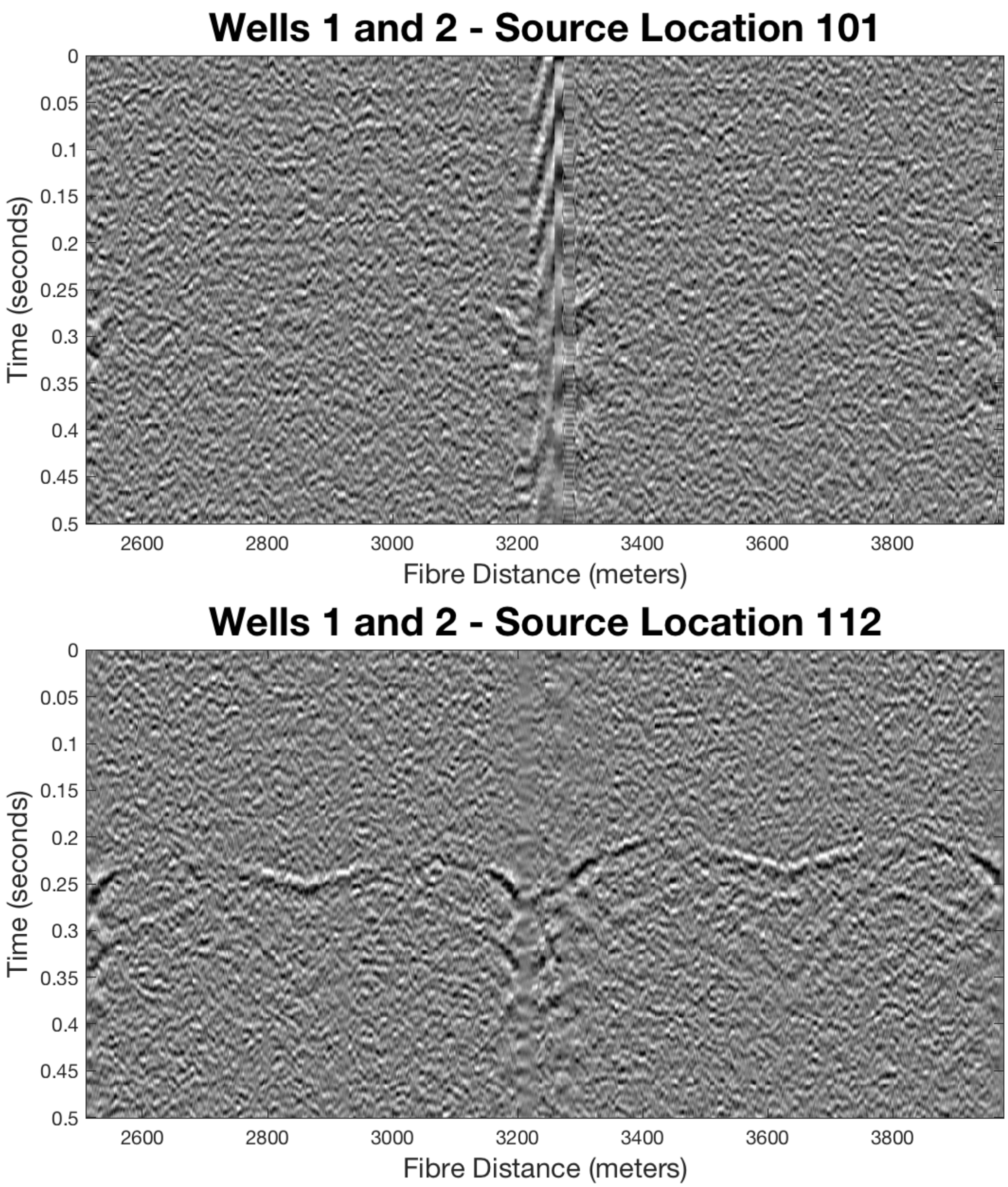


Figure 8.16: The straight-fibre from Well 1 to the straight fibre in Well 2 acquired when the vibroseis truck was at Source Locations 101 (top) and 112 (bottom)

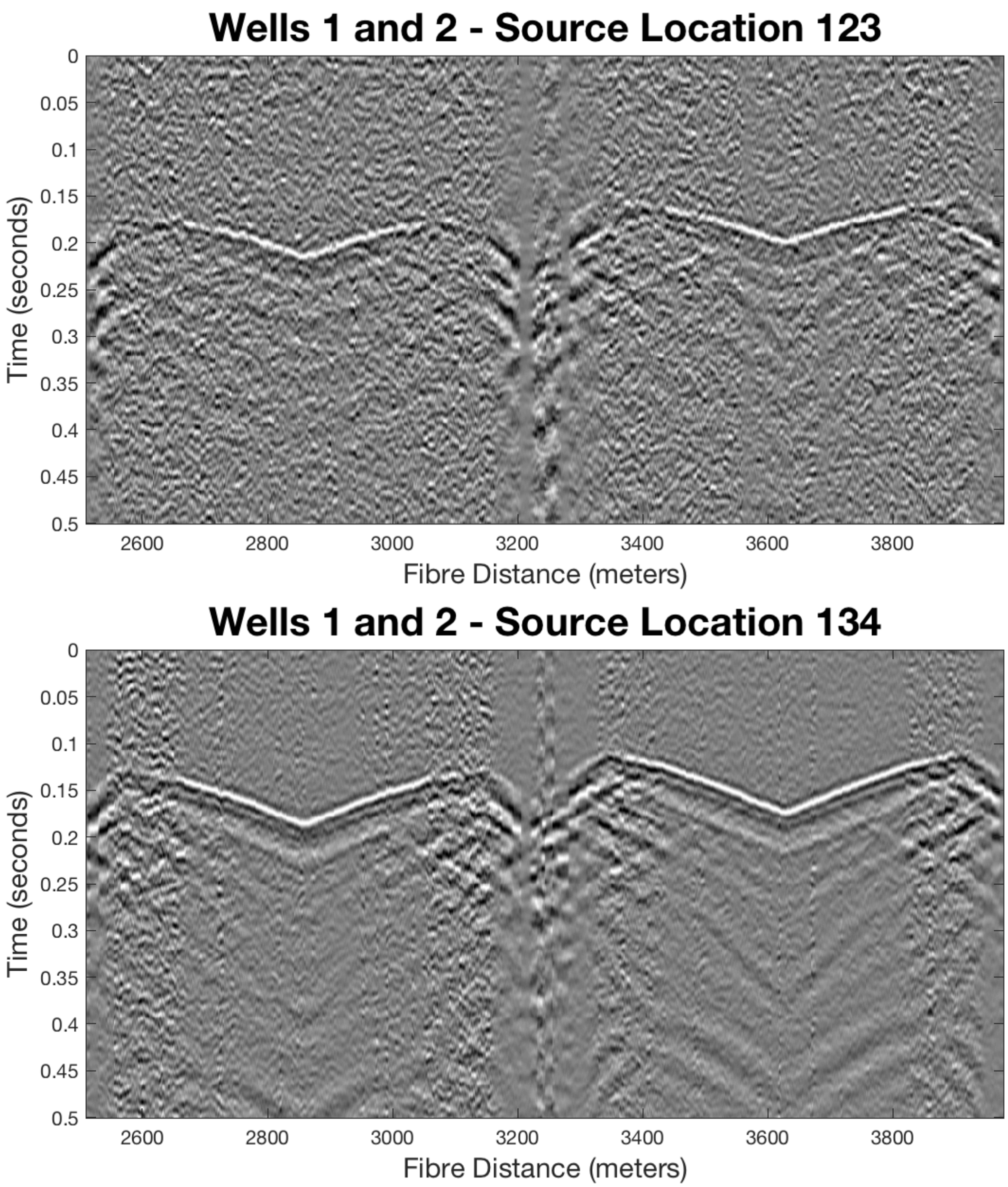


Figure 8.17: The straight-fibre from Well 1 to the straight fibre in Well 2 acquired when the vibroseis truck was at Source Locations 123 (top) and 134 (bottom)

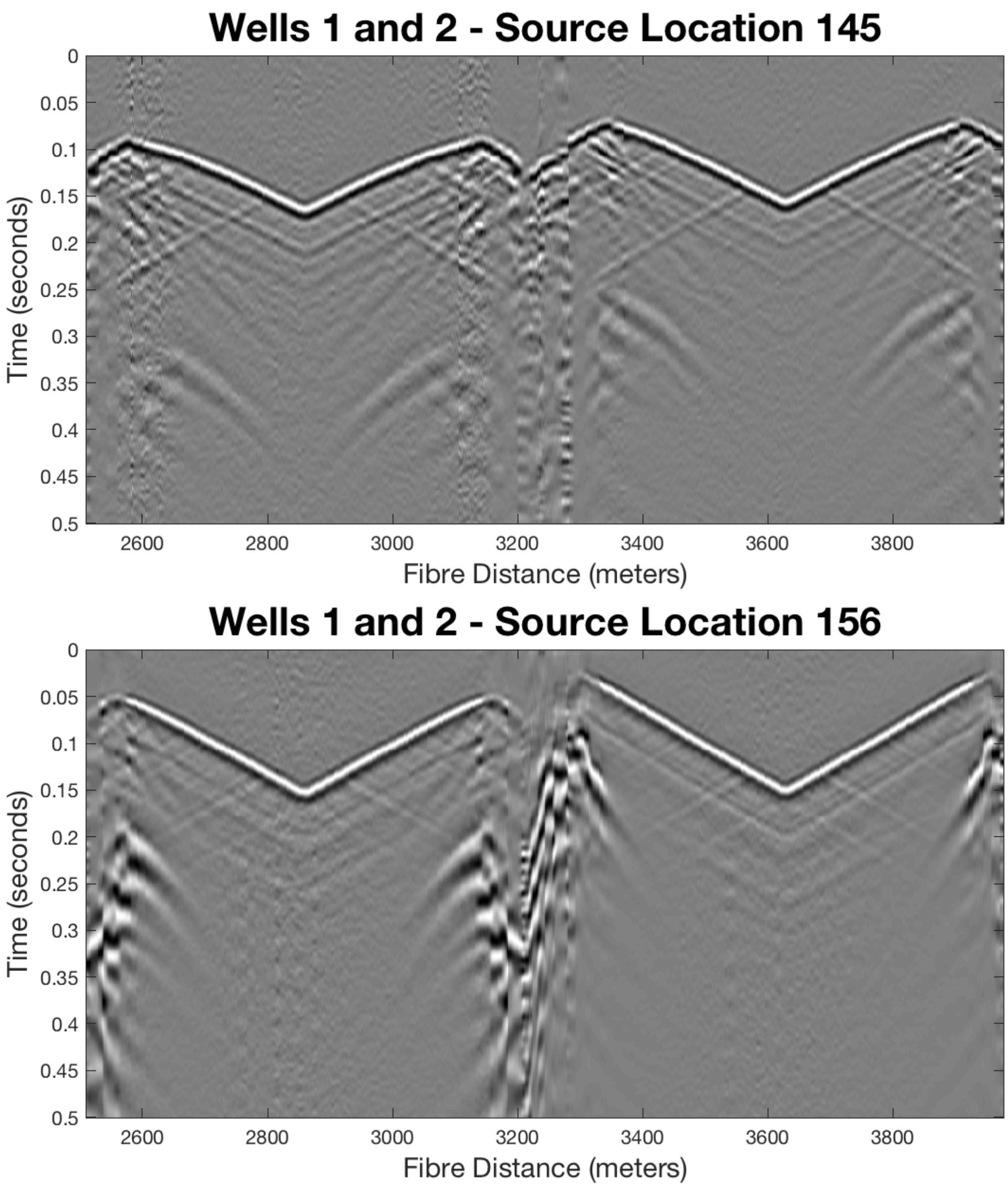


Figure 8.18: The straight-fibre from Well 1 to the straight fibre in Well 2 acquired when the vibroseis truck was at Source Locations 145 (top) and 156 (bottom)

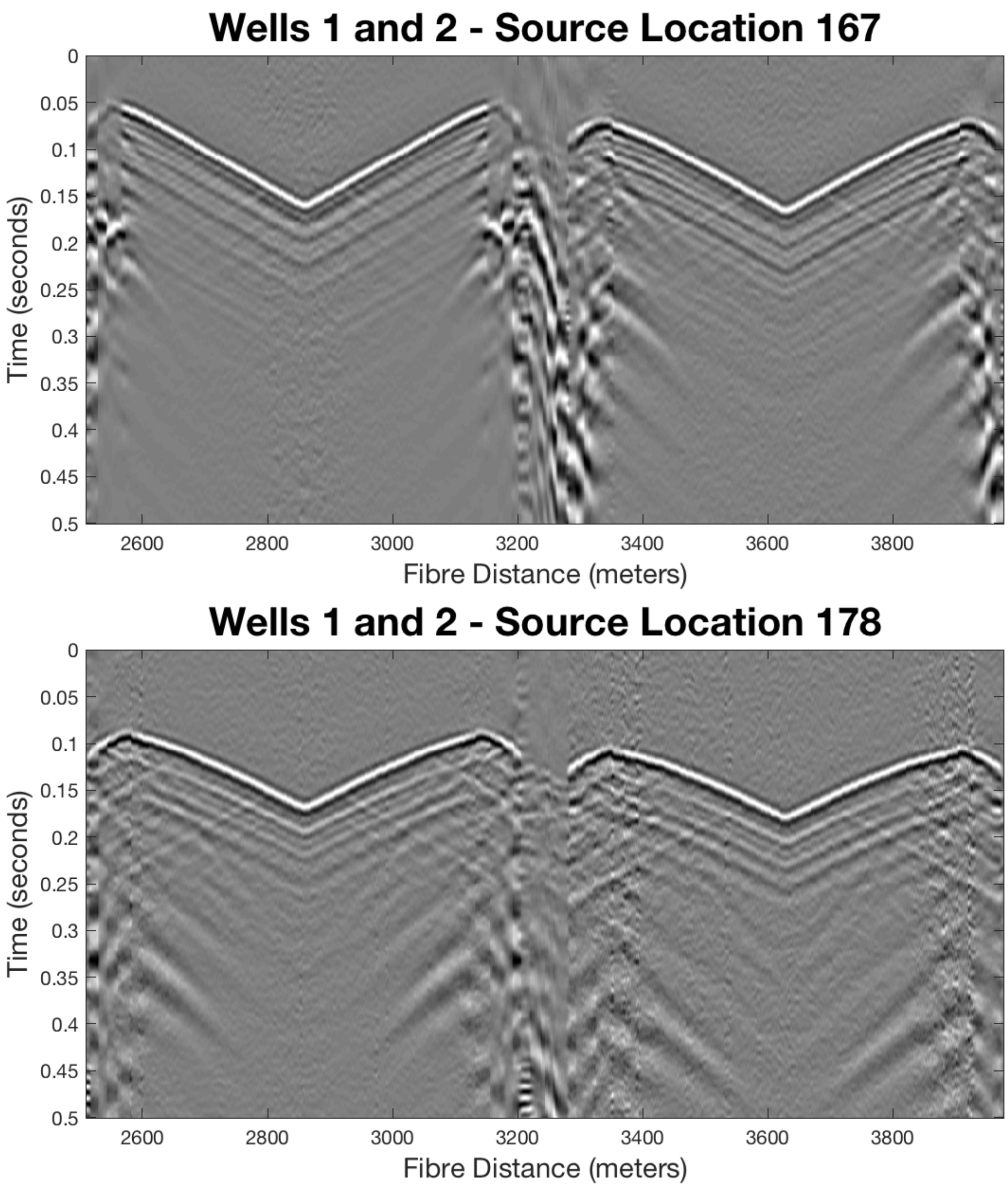


Figure 8.19: The straight-fibre from Well 1 to the straight fibre in Well 2 acquired when the vibroseis truck was at Source Locations 167 (top) and 178 (bottom)

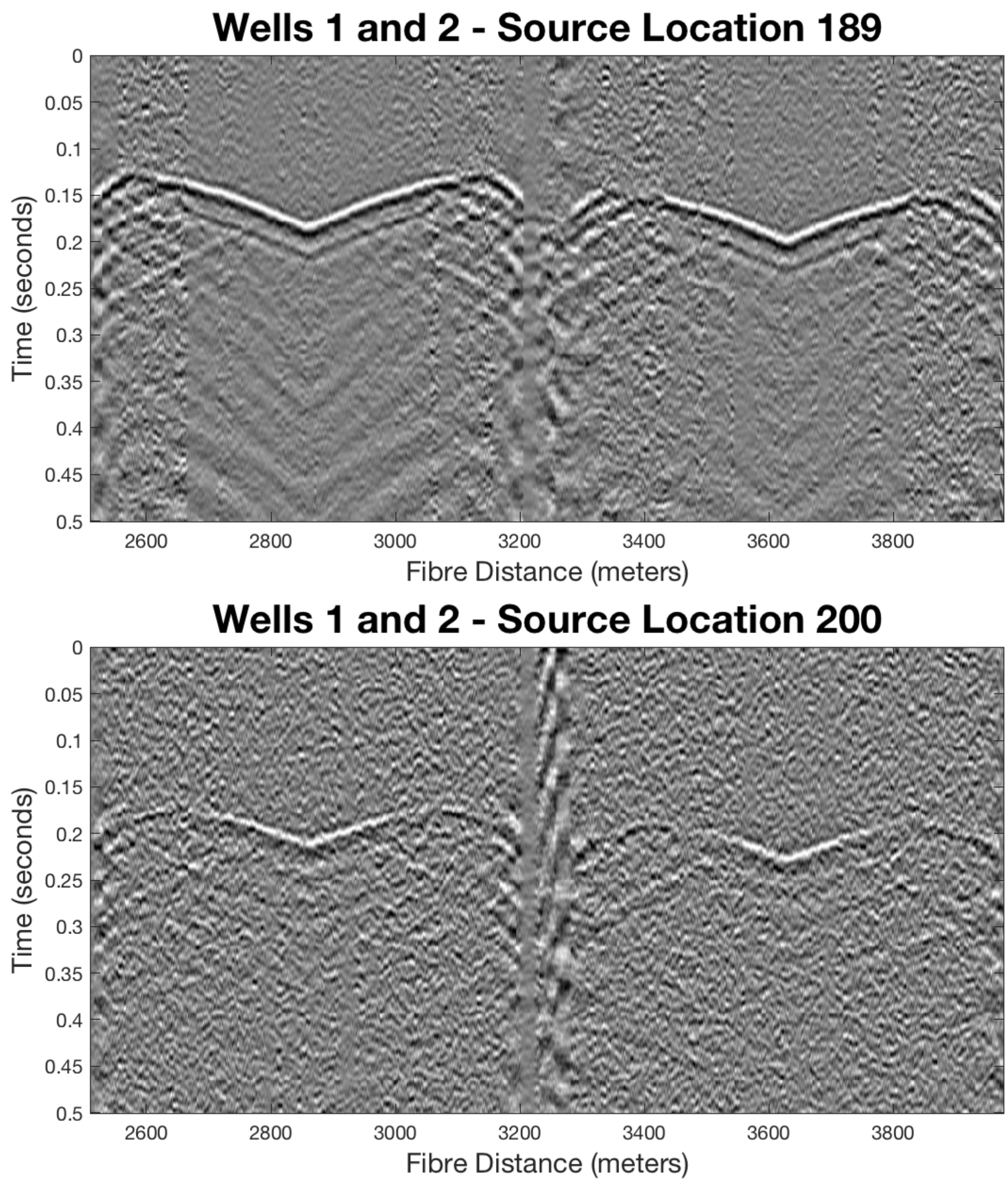


Figure 8.20: The straight-fibre from Well 1 to the straight fibre in Well 2 acquired when the vibroseis truck was at Source Locations 189 (top) and 200 (bottom)

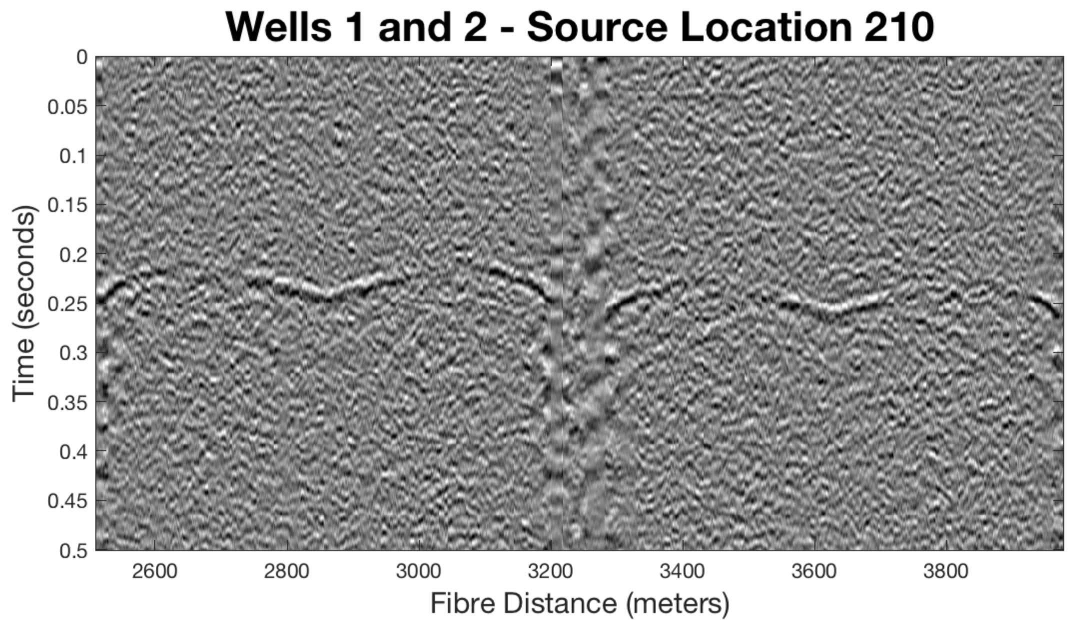


Figure 8.21: The straight-fibre from Well 1 to the straight fibre in Well 2 acquired when the vibroseis truck was at Source Location 210.

The top image of Figure 8.16 depicts the results from the wells when the source location is the farthest southwest on the trench. No conclusive information is acquired by the straight fibre in Wells 1 and 2 at Source Location 101. When the vibroseis truck is moved closer to the wells at Source Location 112, the fibre picks up details about the wells. This pattern continues as the source locations move closer to the center of the trench. Figures 8.18 and 8.19 provide the best resolution for the wells. Some P-wave and S-wave responses are evident in all four source locations. The clarity of the data results from these source locations being the closest to the center of the trench which means they are closest source locations to the wells out of the eleven source locations shown. After Source Location 178, the fibre picks up less detail about the wells from the straight fibre as the vibrosesis moves to source locations further from the center of the trench. Note that the straight fibre acquires more information when the truck is at Source Location 210 (the farthest northeast on the trench) than it does at Source Location 101 (the farthest southwest).

8.3.2 Vertical seismic profiles for Well 1

We now consider the vertical seismic profiles for Well 1; note that we used source location 163 as it was closest to Well 1. Figure 8.22 (left) shows Well 1 when the vibroseis truck was at source location 163. Since the fibre goes down and back up the well, we fold the well data in half which can be seen on the right in Figure 8.22.

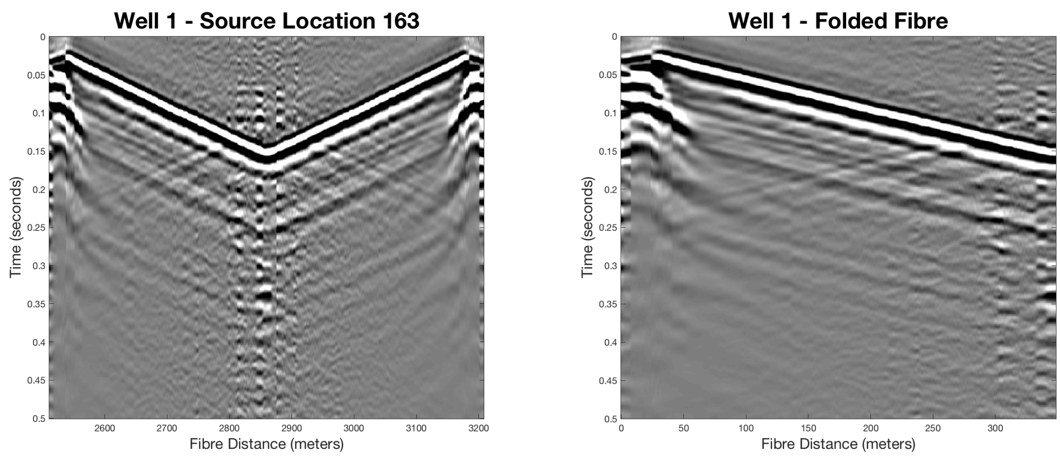


Figure 8.22: (Left) Well 1 when the vibroseis truck is at Source Location 163. (Right) The straight fibre in Well 1 folded.

Figure 8.23 depicts the upgoing and downgoing wavefields.

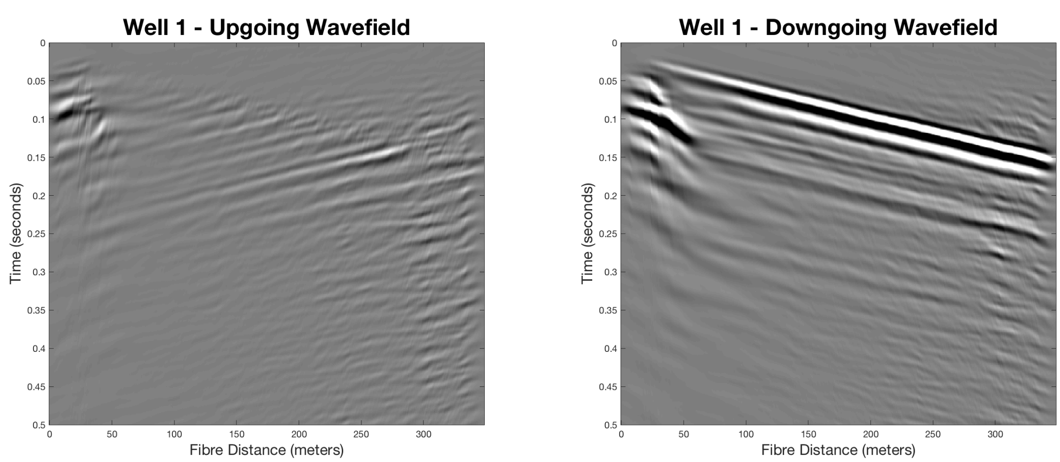


Figure 8.23: (Left) The upgoing wavefield for Well 1. (Right) The downgoing wavefield for Well 1.

We employ the CREWES public toolbox to find the first break pick and flatten the upgoing wavefield for Well 1 at source location 163 in order to see where the reflectors occur in time (Margrave, 2018). Figure 8.24 shows the first break pick on the right and the flattened upgoing wavefield on the left.

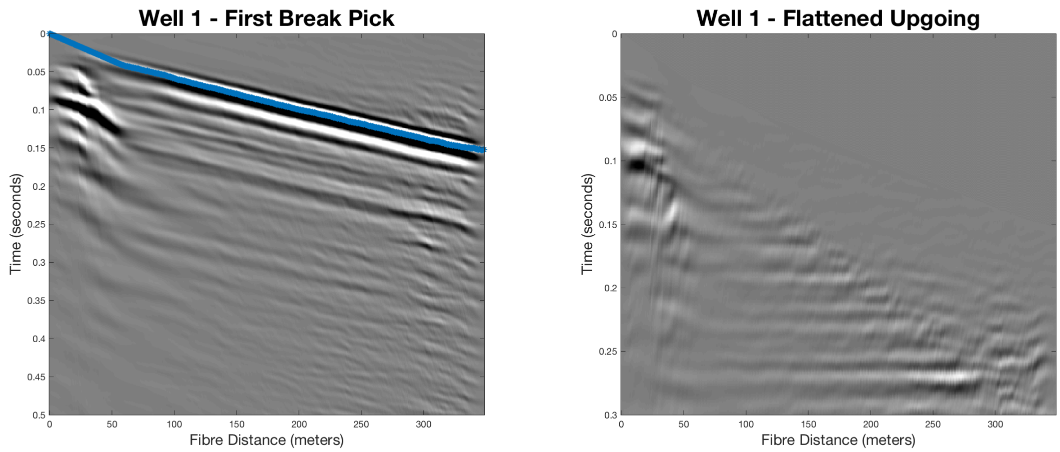


Figure 8.24: (Left) The first-break picks for Well 1 at Source Location 163. (Right) The flattened upgoing wavefield for Well 1.

Finally, we deconvolved the upgoing wavefield to improve the resolution of the data (Margrave, 2006). Figure 8.25 provides the results of the upgoing wavefield (left) next to the results of the deconvolved upgoing wavefield (right). We see the caprock reflector at approximately 0.25 seconds in Figure 8.24 (right) and Figure 8.25 (right), similar to the 2017 experiment's results.

8.3.3 Vertical seismic profiles for Well 2

For Well 2, we find Source Location 158 to be the closest source to the well and use it to process the vertical seismic profiles. Figure 8.26 shows the results of the Well 2 at Source Location 158 on the left. Given that the fibre goes straight down Well 2 and returns to the surface, we fold the data in half as seen on the right in Figure 8.26. We then use the folded fibre to compute the vertical seismic profile for the well.

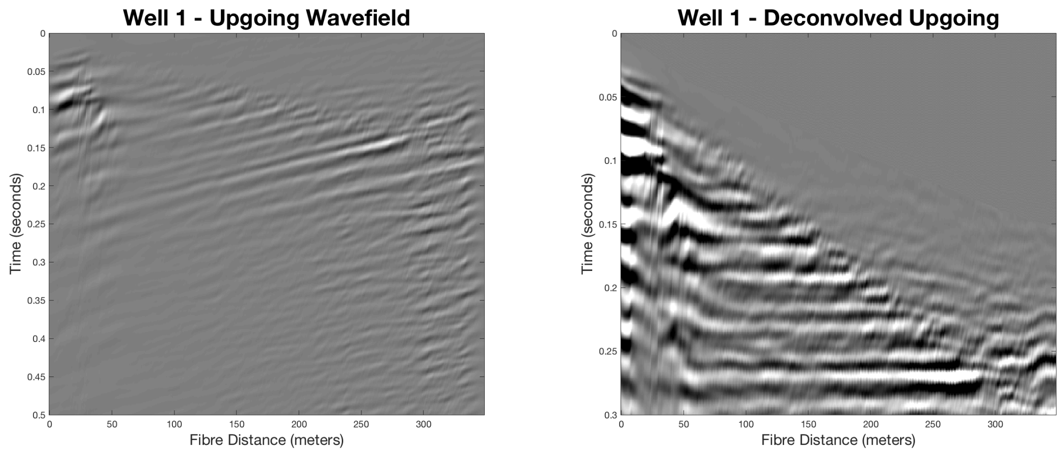


Figure 8.25: (Left) The upgoing wavefield for Well 1. (Right) The deconvolved upgoing wavefield for Well 1.

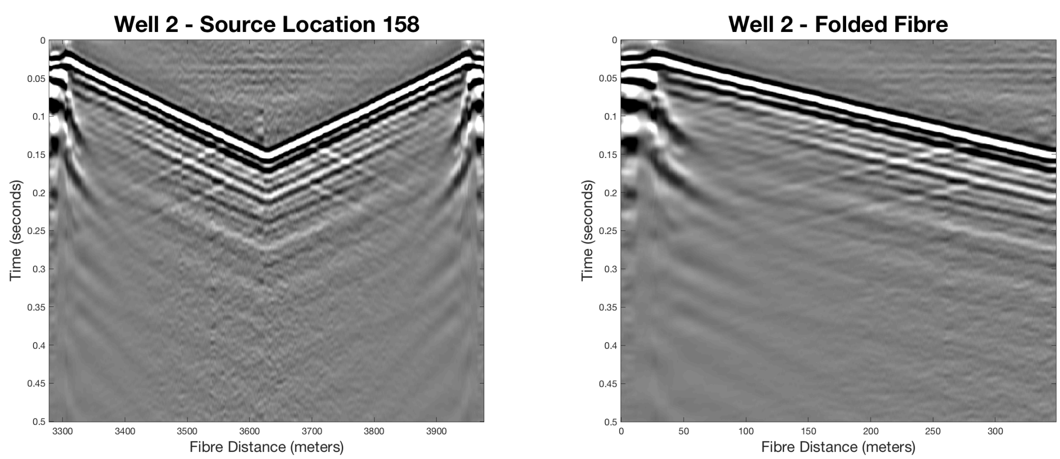


Figure 8.26: (Left) Well 2 when the Vibroseis truck is at Source Location 158. (Right) The straight fibre in Well 2 folded.

Following the same process as with Well 1, we start by computing the upgoing and downgoing wavefields for Well 2. We show the results of this computation in Figure 8.27.

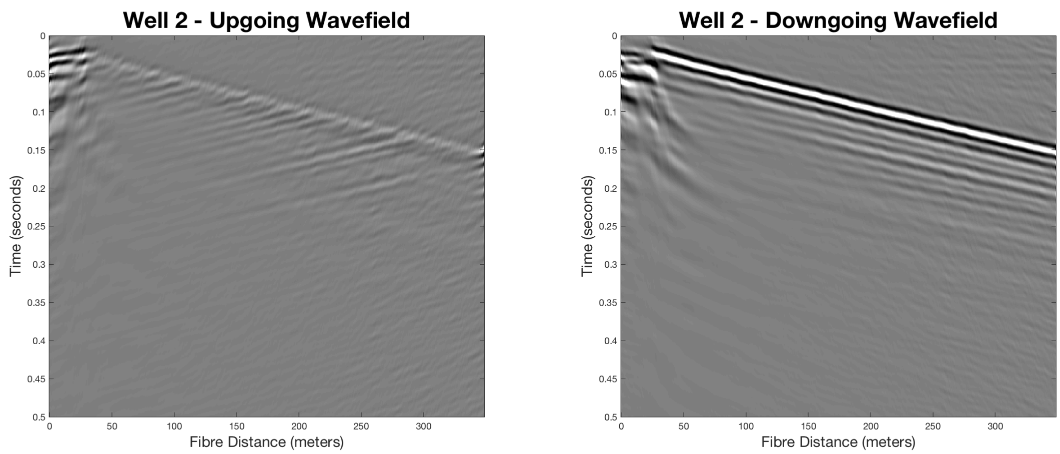


Figure 8.27: (Left) The upgoing wavefield for Well 2. (Right) The downgoing wavefield for Well 2.

We then pick the first break which can be seen on the left in Figure 8.28. We employ the first break pick to flatten the upgoing wavefield; see Figure 8.28 (right). The CREWES public toolbox was utilized to find both the first break pick and the flattened upgoing wavefield.

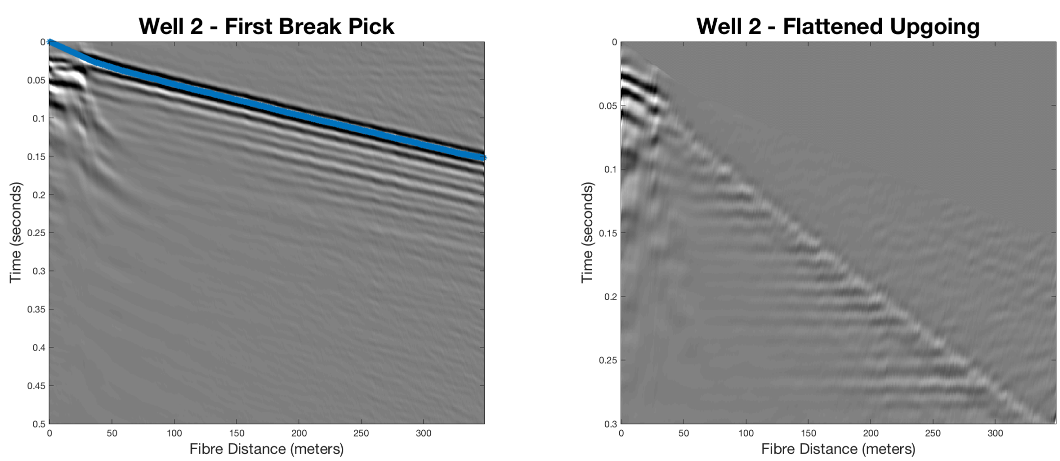


Figure 8.28: (Left) The first-break picks for Well 2 at Source Location 158. (Right) The flattened upgoing wavefield for Well 2.

We also use the CREWES public toolbox to deconvolve the upgoing wavefield. Figure

8.29 displays the upgoing wavefield is on the left and the deconvolved upgoing wavefield is on the right. The reflector for the cap rock is less distinguishable in the flattened upgoing wavefield for Well 2; however, we can identify it in the deconvolved upgoing wavefield at approximately 0.25 seconds.

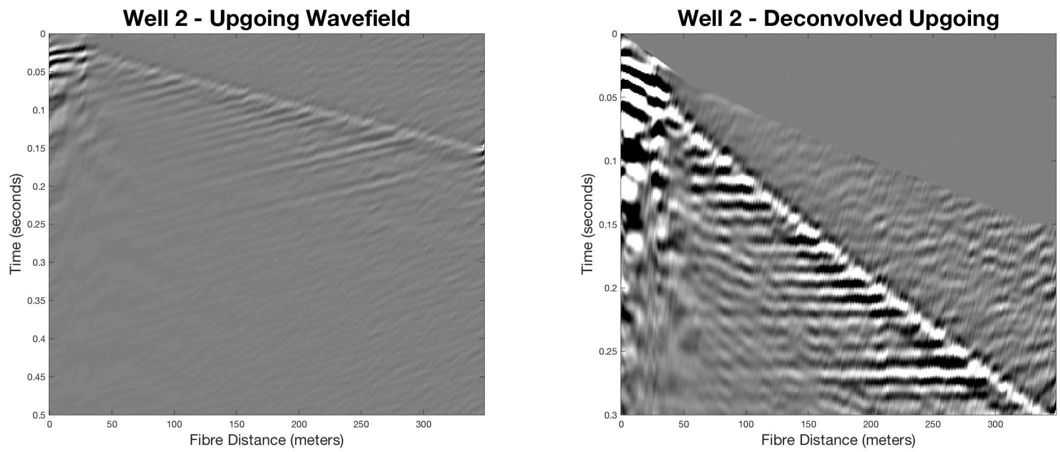


Figure 8.29: (Left) The upgoing wavefield for Well 2. (Right) The deconvolved upgoing wavefield for Well 2.

8.4 Conclusions

We discussed the Containment and Monitoring Institute's Field Research Station and focused on the schematic of the fibre loop at the site. We then considered the data acquired from the CaMI site in Newell County, AB in 2017 and looked at the full fibre data for Line 35 at Source Location 103 as well as the data from the two wells. With regard to the Wells 1 and 2, we compared the results when the fibre goes straight down Well 2 and returns to the surface before going straight down Well 1 and returning to the surface at each source location. The Source Location 103 provided the best results of the subsurface in all three types of data. This fact is probably because it resides in the middle of both wells and beside the trench. With regards to the vertical seismic profile results, we were able to identify the caprock at approximately 0.25 seconds in the flattened upgoing wavefield but struggled to do so in the deconvolved upgoing wavefield for both wells. In fact, the reflector appeared to occur near 0.3 seconds in the deconvolved upgoing wavefield for the 2017 experiment in both wells.

We then explained the experiment conducted in August 2018. We looked at the straight fibre in Wells 1 and 2 at 11 different source locations along the trench. The source locations between 145 and 178 provided the best results for the wells. We also noted that Source Location 210 provided more information about both the wells than Source Location 101. Then, we used standard processing techniques to find the vertical seismic profiles of Well 1 at Source Location 163 and Well 2 at Source Location 158. The caprock was less prominent in the flattened upgoing wavefield for the 2018 experiment; however, it could be identified at approximately 0.25 seconds in the deconvolved upgoing wavefield for both wells.

Recall that the caprock reflector for the 2017 experiment appeared at 0.3 seconds in the deconvolved upgoing wavefield whereas, in the 2018 experiment, it occurred at 0.25 seconds in the deconvolved upgonig wavefield. The difference is likely due to the fact that the source locations used in the 2018 experiment were closer to the wells than the source locations used in the 2017 experiment. It also shows that the deconvolution does not always improve the quality of the resolution for the upgoing wavefield. The prominence of the reflector in the results of both experiments is still very promising for using distributed acoustic sensing when monitoring CO_2 storage.

Chapter 9

Locating events using independent component analysis and Gaussian mixture models

Distributed acoustic sensors have practical use in security and monitoring (Hartog, 2017). Detecting events in the DAS-acquired data becomes a major issue, especially since such data can cover time-intervals spanning hours. It thus becomes essential to be able to detect events within terabytes of data. In this chapter, we offer an analysis of one technique for locating these events inspired by the authors of Shamsa and Paydayesh (2019): Gaussian mixture models.

In Shamsa and Paydayesh (2019), the authors employ independent component analysis (ICA) with a Gaussian mixture model (GMM) to detect events in microseismic data. They utilize ICA to enhance the feature they wish to detect in the microseismic signal with the GMM. We apply a similar idea to a data set acquired using a DAS system installed next to a road, and used to detect a vehicle driving along a road. The GMM extracts the foreground of the data which contains the vehicle's signal. As such, we will refer to it as the vehicle detector. We compare the use of ICA beforehand in a variety of different methods.

Independent component analysis was created as an answer to the *cocktail party problem*, where the objective is to differentiate between two speakers based on several microphones collecting information. Given that a DAS system only possesses one receiver, we use several data sets of the same region of the fibre to act as the additional microphones in this case. After a brief introduction to the methodology of the chapter, Gaussian mixture models, and independent component analysis, we begin by applying an ICA to segments of multiple data sets to enhance the vehicle signal over the noise. In the first half of the chapter, we chose the multiple data sets to contain a majority of data with the vehicle's signal and a minority which contained only noise. Independent component analysis gives two independent components: one for the vehicle signal and one for the noise. Afterwards, we then train a GMM on each set of independent components separately before applying the two GMMs to a test data set. We also provide the results of training the GMM on a training data set. We compare the effectiveness of each method by calculating the percentage of overlap the detector produces for locating the vehicle's signal in the data. For the first case, the GMM without an ICA applied beforehand performed better, achieving between approximately 91% and 93% overlap of the signal as opposed to the GMM with an ICA, which achieved between approximately 20% and 43% overlap of the vehicle signal in the test data. This is likely due the fact that ICA cannot order the independent components, i.e. the vehicle signal may not be assigned to the same component for each segment (Hyvärinen and Oja, 2000).

In the second example, we apply an ICA to a larger segment of the multiple data sets and then train the GMM on a sliding window (the same size as the segments from the first example) across the larger independent component results. In this case, the vehicle signal and noise were delegated to two separate independent components unlike the overlap we saw in the previous case. We window a GMM across both independent components, and compare the results of the GMM without an ICA applied. The GMM taught using the noisy independent component achieved a overlap between approximately 59% and 80% whereas the GMM trained on the vehicle independent component earned a overlap between approximately 90% and 93%. The GMM taught on the training set acquired a overlap between approximately 97% and 98%. The increase in overlap between the two cases likely stems from the fact that more frames were employed to train the GMMs. This improved the results of the GMM taught on the training data. The consistent signal for the independent components also factor in the significant increase in overlap.

We also applied the GMM to multiple data sets where the majority contained noise and only one data set contained the vehicle's signal. We repeated the experiment for large 70 second portions of data and multiple 5 second segments of data to estimate the independent components. The vehicle detectors trained on independent components only achieved around 81% which was in the cases when we considered the 70s segments of the data sets. The results in these cases only achieved a higher percentage than the results from the first example where we trained on 5s of multiple data sets where a majority of the segments contained noise. Despite this result, we concluded that using multiple data sets where a majority contain the vehicle's signal for estimating the independent components produced better results than any other case involving the use of independent component analysis.

While the GMMs trained on the independent components did not outperform the GMM trained on a training set in any case, the GMMs trained on the independent components often provided a more precise detection of the event unlike the GMM taught using the training data. The success of the Gaussian mixture model in detecting the vehicle's signal across both methods still shows that this path of event detection holds promise. In the future, using more specific data sets to enhance the vehicle signal with the independent component analysis is worth investigation. Employing fewer data sets by choosing ones which contain no vehicle signal and a single data set which contains the vehicle signal may improve the results of the ICA to extract the vehicle signal from the noise. This works on the idea that the noise should be similar across the DAS system's monitoring of the road. Investigation of other methods for feature extraction besides ICA should also be considered.

9.1 Methodology

In Shamsa and Paydayesh (2019), the authors outline a method which uses a Gaussian mixture model (GMM) and independent component analysis (ICA) to separate the signal from the noise in microseismic data. They found that applying an independent component analysis from Guoshen (2012) to the data prior helped with detection and used this method on data being acquired in real time. The authors saw improvement when more data was fed through the ICA and GMM as it allowed the ICA to better distinguish between noise and the signal; however, it was possible to provide too much data and cause the error from the ICA to grow too large.

Following a similar vein, we utilize the fastICA described in Gävert et al. (2005) and employ it with regards to the foreground detector from MATLAB following the example from MATLAB (2019). In the MATLAB example, the Gaussian mixture model differentiates the foreground from the background to identify cars in video data. The first few frames of the video contain no images of cars in order to define the background for the Gaussian mixture model. As the cars arrive in the frames, the GMM distinguishes them from the background. We apply this methodology to data of a vehicle driving along a road, acquired using a distributed acoustic sensor. While the example's application is to images which contain cars, we use the foreground detector on DAS-acquired data to indicate the signal of a vehicle. We refer to it as the vehicle detector in the applications later in this chapter. We also administer additional morphological cleaning to the data than what is found in the MATLAB example.

In the following sections, we provide a brief study of Gaussian Mixture Models and their use in foreground detection. We then examine the algorithm behind independent component analysis and its ability to distinguish between a signal and noise in data. Afterwards, we apply the methodology described in Shamsa and Paydayesh (2019) to DAS-acquired data of a vehicle driving along a road.

9.2 Gaussian mixture models

Traffic monitoring often employs Gaussian mixture models for foreground detection where the cars lie in the foreground. We follow an example from MATLAB that applies GMMs to data from a video data of cars driving down a road (MATLAB, 2019). It bases the foreground detector functions on the Gaussian mixture models described in Stauffer and Grimson (1999) and KaewTraKulPong and Bowden (2001). The authors of Stauffer and Grimson (1999) sought to create a foreground detector which could deal with arbitrary variations in the background; in particular, those found in video data such as the shadows from the sun's movement, moving tree branches, etc. Their answer to address these changes in the background of video data involves describing each pixel of an image via some number of Gaussians distributions, often ranging between three and five.

Distributed acoustic sensors do not produce video data; however, they are used in monitoring projects over large periods of time. Segmenting the data in the form of a video makes the data more manageable. It also lowers the computational costs of applying processing techniques. Employing a Gaussian mixture model on DAS-acquired data enables the distinction between the background noise and the foreground where events such as walking or digging may occur. Despite the lack of visual concerns such as shadows due to lighting or moving tree branches, data often contains noise or other information which can overpower events in the data that we wish to detect. The GMM treats the noise as the background and produces a foreground where the events reside. In Section 9.4, a DAS system is employed to detect a vehicle driving by the road. The goal of the GMM is to distinguish between the vehicle's signal and the background noise.

The authors of Stauffer and Grimson (1999) follow two basic steps in their proposed GMM:

1. Model the values of a specific pixel as a mixture of Gaussian distributions.

2. Assign the pixel to the foreground or background based on persistence and variance of

each Gaussian distribution in the mixture.

Each pixel in the image undergoes a "pixel process." At any point in time, the pixel history is known, i.e.

$$\{X_1, ..., X_t\} = \{\mathbf{I}(x_0, y_0, i) : 1 \leq i \leq t\} \tag{9.1}$$

where X_i is the pixel value and $\mathbf{I}$ is the image sequence. The recent history of each pixel is modeled by a mixture of K Gaussian distributions. Let $w_{i,t}$ be the weight estimate of the ith Gaussian in the mixture at time t, $\mu_{i,t}$ be the mean value of the ith Gaussian in the mixture at time t, and $\Sigma_{i,t} = \sigma_i^2 I$ is the covariance matrix of the ith Gaussian in the mixture at time t. We calculate the probability of the current pixel value occurring via the equation:

$$P(X_t) = \sum_{i=1}^{K} w_{i,t} * \eta(X_t, \mu_{i,t}, \Sigma_{i,t}) \tag{9.2}$$

where the Gaussian probability density function

$$\eta(X_t, \mu_{i,t}, \Sigma_{i,t}) = \frac{1}{(2\pi)^{\frac{n}{2}} |\Sigma_{i,t}|^{\frac{1}{2}}} \exp\Big(-\frac{1}{2}(X_t - \mu_{i,t})^T \Sigma_{i,t}^{-1} (X_t - \mu_{i,t})\Big). \tag{9.3}$$

To maximize the likelihood of the observed data, the authors execute an online K-means approximation since an exact expectation-maximization (EM) algorithm would be computationally expensive. Each new pixel is compared to the existing K Gaussian distributions until we find one within 2.5 standard deviations of a distribution. When this occurs, we have a match. If the pixel is not within 2.5 standard deviations of any of the K Gaussian distributions, then we replace the least probable distribution with the current pixel value as well as its mean value, an initially high variance, and the low prior weight. We update the weight using the following update step:

$$w_{i,t} = (1 - \alpha) w_{i,t-1} + \alpha M_{i,t}, \tag{9.4}$$

where α is the learning parameter and

$$M_{i,t} = \begin{cases} 1 & \text{for a model which matched the current pixel value;} \\ 0 & \text{for remaining models.} \end{cases} \tag{9.5}$$

The parameters μ and σ are updated as follows:

i. For all unmatched Gaussian distributions, μ and σ remain the same;

ii. For the matched distribution,

$$\mu_{i,t} = (1-\rho)\mu_{i,t-1} + \rho X_t; \tag{9.6}$$

$$\sigma^2_{i,t} = (1-\rho)\sigma^2_{i,t-1} + \rho(X_t - \mu_{i,t})^T(X_t - \mu_{i,t}), \tag{9.7}$$

where

$$\rho = \alpha\eta(X_t|\mu_{i,t}, \sigma_{i,t}). \tag{9.8}$$

Once each pixel is assigned to a Gaussian distribution in the mixture, we order the Gaussian distributions using the parameter w/σ which puts the most probable background distributions at the top of the list and the least probable background distributions at the bottom to eventually be replaced by new distributions. Let

$$B = \operatorname{argmin}_b \Big(\sum_{k=1}^{b} w_k > T \Big), \tag{9.9}$$

where T is the measure of the minimum portion of the data for which the background should be accounting. We choose the first B Gaussian distributions in our list to model the background of the data.

We use a two-pass, connected component algorithm to segment labeled foreground pixels into regions. In Stauffer and Grimson (1999), the authors employ a linearly predictive multiple hypothesis tracking algorithm to correlate connected component between frames of data. The algorithm utilizes the position and size from the previous frame to make its judgment about the next. Matching the models to connected components typically requires comparing each model against the current pool of connected components.

In KaewTraKulPong and Bowden (2001), the authors improved the work in Stauffer and Grimson (1999) by adjusting the update step for the weights w, and the parameters μ and σ, as well as including a threshold for shadow detection. Instead of using a K-means approximation, the authors of KaewTraKulPong and Bowden (2001) implemented an EM-algorithm for the first L windows using the following update equations:

$$w_{i,t+1} = w_{i,t} + \frac{1}{t+1}\Big(M_{i,t+1} - w_{i,t}\Big); \tag{9.10}$$

$$\tag{9.11}$$

$$\mu_{i,t+1} = \mu_{i,t} + \frac{M_{i,t+1}}{\sum_{k=1}^{t+1} M_{i,k}}\Big(X_{t+1} - \mu_{i,t}\Big); \tag{9.12}$$

$$\tag{9.13}$$

$$\Sigma_{i,t+1} = \Sigma_{i,t} + \frac{M_{i,t+1}}{\sum_{k=1}^{t+1} M_{i,k}}\Big((X_{t+1} - \mu_{i,t})(X_{t+1} - \mu_{i,t})^T - \Sigma_{i,t}\Big), \tag{9.14}$$

where w_i is the k-th Gaussian component and the parameter $1/\alpha$ defines the time constant. Then, they apply a L-recent windows update using the following equations:

$$w_{i,t+1} = w_{i,t} + \frac{1}{L}\Big(M_{i,t+1} - \omega_{i,t}\Big); \tag{9.15}$$

$$\tag{9.16}$$

$$\mu_{i,t+1} = \mu_{i,t} + \frac{1}{L}\Big(\frac{M_{i,t+1}X_{t+1}}{\omega_{i,t+1}} - \mu_{i,t}\Big); \tag{9.17}$$

$$\tag{9.18}$$

$$\Sigma_{i,t+1} = \Sigma_{i,t} + \frac{1}{L}\Big(\frac{M_{i,t+1}}{\omega_{i,t+1}}(X_{t+1} - \mu_{i,t})(X_{t+1} - \mu_{i,t})^T - \Sigma_{i,t}\Big). \tag{9.19}$$

They achieved the shadow detection in KaewTraKulPong and Bowden (2001) using a position vector at the RGB mean of the pixel background, E, an expected chromaticity distortion, d, and a brightness threshold, τ. We calculate the brightness distortion a and the colour distortion c:

$$a = \text{argmin}_z (X_t - zE)^2; \tag{9.20}$$

$$c = \left\| X_t - aE \right\|_2, \tag{9.21}$$

for an observed pixel value, X_t. The authors define a moving shadow if the brightness distortion a is within 2.5 standard deviations and $\tau < c < 1$.

While this Gaussian mixture model is geared toward video data, we adapt it work on DAS-acquired data in Section 9.4 to detect a vehicle driving along a fibre-optic cable receiver. In the next section, we discuss indepedent component analysis.

9.3 Independent component analysis

To help the GMM differentiate between the vehicle's signal and the background noise of the data, we utilize independent component analysis for the purposes of feature extraction. Independent component analysis offers a solution to the cocktail party problem which asks the solver to distinguish between two individuals speaking simultaneous given two recordings from two microphones at different locations in the room.

Mathematically, we can write this problem as

$$v_1(t) = a_{11}s_1(t) + a_{12}s_2(t); \tag{9.22}$$

$$v_2(t) = a_{21}s_1(t) + a_{22}s_2(t), \tag{9.23}$$

where v_1 and v_2 are the recordings from the microphones, a_{ij} for $1 \leq i, j \leq 2$ are the parameters describing the distance between each speaker and the microphone, and s_1 and

s_2 are the speech signals from each speaker respectively. This problem can be written as

$$\mathbf{v} = A\mathbf{s}. \tag{9.24}$$

Independent component analysis provides an estimate of the independent components $\mathbf{s}$ based solely on information from $\mathbf{v}$. It does this by solving for an approximation of A which it then computes the inverse of A to solve

$$\mathbf{s} = A^{-1}\mathbf{v}. \tag{9.25}$$

Before the ICA algorithm is applied, the data $\mathbf{v}$ must be whitened (Hyvärinen and Oja, 2000). This process involves transforming $\mathbf{v}$ into a new vector whose components are uncorrelated and the variance of these components equals 1, i.e., $\mathbf{x} = M\mathbf{v}$ where $\mathbf{x}$ is the correlation matrix satisfying $E\{\mathbf{x}\mathbf{x}^T\} = 1$. The method we employ later for the ICA algorithm uses principal component analysis to whiten $\mathbf{x}$.

We implement a fixed-point ICA algorithm for estimating multiple independent components based on the work in Hyvärinen (1999). The authors of Hyvärinen (1999) define the algorithm as follows: Assume we have a sample of the prewhitened random vector $\mathbf{x}$:

1. Take a random initial vector $\mathbf{w}(0)$ where $\|\mathbf{w}\| = 1$. Let $k = 1$.

2. Let

$$\mathbf{w}(k) = E\{\mathbf{x}(\mathbf{w}(k-1)^T\mathbf{w}(k-1))^3\} - 3\mathbf{w}(k-1). \tag{9.26}$$

   Estimate the expectation E using a large sample of vectors $\mathbf{x}$.

3. (a) Update $\mathbf{w}(k)$ to the quantity $\mathbf{w}(k) - \overline{\mathbf{B}}\overline{\mathbf{B}}^T\mathbf{w}(k)$ and replace $\mathbf{w}(k)$ with the update. Symbolically,

$$[\text{Updated}]\ \mathbf{w}(k) := \mathbf{w}(k) - \overline{\mathbf{B}}\overline{\mathbf{B}}^T\mathbf{w}(k). \tag{9.27}$$

(b) Compute

$$\frac{\mathbf{w}(k)}{\|\mathbf{w}(k)\|}. \tag{9.28}$$

4. For $|\mathbf{w}(k)^T\mathbf{w}(k-1)|$ not close to 1, let $k = k+1$ and return to 2. Otherwise, output the vector $\mathbf{w}(k)$ into a column of the orthogonal mixing matrix B.

The parameter $\overline{\mathbf{B}}$ is the matrix whose columns are taken from the previously found columns of the orthogonal mixing matrix $\mathbf{B}$, where $\mathbf{s} = \mathbf{B}^T\mathbf{x}$. The columns of $\mathbf{B}$ are the independent components computed by the algorithm. Once the solution $\mathbf{w}(k)$ reaches the basins of attractions of one of the fixed points, we drop step 3(a) in order to prevent an estimation error from building in $\overline{\mathbf{B}}$.

We employ MATLAB code written by Hugo Gävert, Jarmo Hurri, Jaakko Särelä, and Aapo Hyvärinen to apply the FastICA algorithm to the examples in Section 9.4 (Gävert et al., 2005). The authors based their `fastICA` function on the work in Hyvärinen (1999) and Hyvärinen and Oja (2000).

A significant factor to address with regards to independent component analysis is that the independent components must be non-Gaussian (Hyvärinen and Oja, 2000). If both components are Gaussian, then the matrix A cannot be estimated. While the combination of independent component analysis with a mixture model which estimates pixel values based on Gaussians seems problematic, the foreground detector in a GMM employs a mixture of Gaussians to estimate each pixel which does not mean each pixel itself is Gaussian. We later apply the GMM to the independent components generated by the ICA so the foreground detector using Gaussian distributions to estimate the points in the independent components. Because we apply the ICA to the data first and then the GMM to the independent components, this will not affect how the ICA works with the data.

9.4 Examples

For the following sections, we examine different methods for applying the GMM and ICA to a data set of a vehicle driving along a road next to a fibre-optic sensor. Our attention resides on a portion of the road that is approximately 65m in length, and focus on it for about 70 seconds. We start with the application of a Gaussian mixture model to a single data set. We consider the results of administering an ICA to segments of multiple data sets of the same portion of road. Finally, we apply an ICA to multiple large data sets and then apply the GMM to the resulting independent components. Recall that we refer to the foreground detector from the literature as a vehicle detector in this chapter given that we use the GMM to detect the vehicle's signal.

Figure 9.1 depicts the data we use to train the GMM in the following examples when we do not apply an ICA first. A Fotech DAS system collected the data of a vehicle driving on a road with a fibre-optic cable laid beside it. The data is approximately 135 seconds in length and considers a 65m stretch of fibre. It shows a vehicle driving along the fibre-optic cable between 90 and 105 seconds.

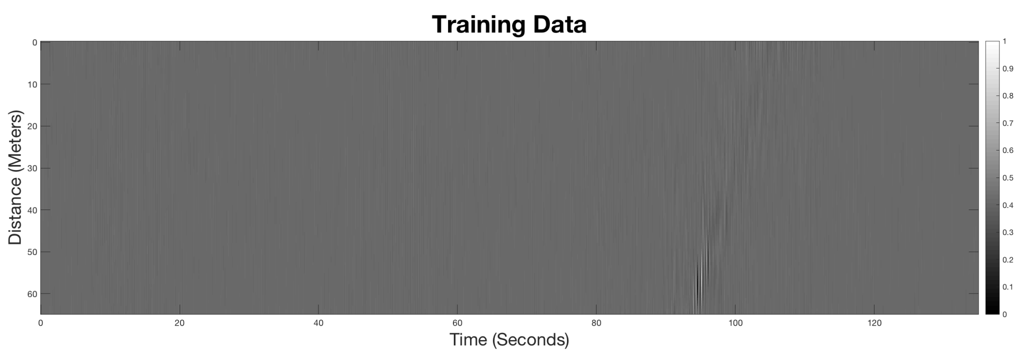


Figure 9.1: The training data set: A 65 meter section of fibre with a vehicle driving along a road next to a DAS system. The vehicle starts at 90 seconds and goes to 105 seconds.

In order to offer a comparison of the effectiveness of each method's vehicle detector, we apply the resulting vehicle detectors in each section to a test data set; c.f. Figure 9.2. The figure shows the data set of a vehicle driving along a road beside a fibre-optic cable. It starts with no vehicle present. After 10 seconds, the vehicle appears and drives along the 65m of fibre for 60 seconds before the data ends. In each of the following sections, we observe how well the vehicle detectors detect the vehicle in the test data set from Figure 9.2.

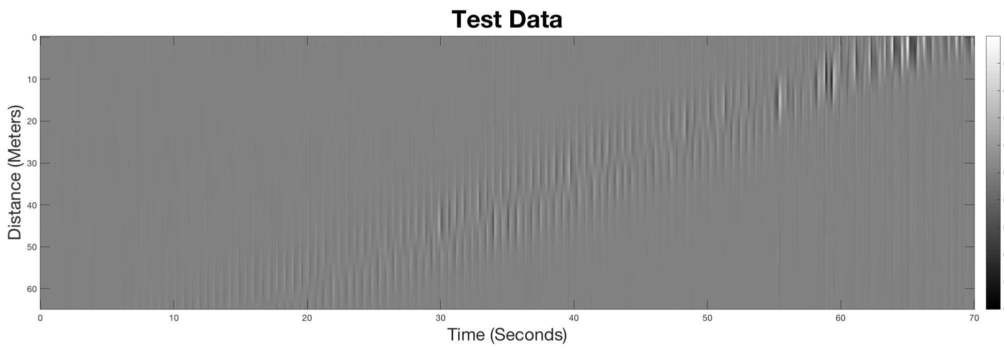


Figure 9.2: The test data set: A 65 meter section of fibre with a vehicle driving along a road next to a DAS system. The vehicle starts at 10 seconds and goes to 70 seconds.

There is only one DAS system in this setup, which means there is only one microphone for the independent component analysis in each example. Given that the background noise should be similar over the long period of time, we generate more "microphones" by considering data sets acquired at different times for the same section of fibre. While this may provide us with a good estimation of the noise, it may cause issues with how well the ICA depicts the signal. A problem may arise if we use data with a variety of vehicle signals, such as the vehicle driving from 0m to 65m and the vehicle driving from 65m to 0m. We keep this in mind as we consider the following examples.

9.4.1 Case 1: GMM applied to independent components generated from segments of multiple data sets

The authors of Shamsa and Paydayesh (2019) found that segmenting the data sets before performing an independent component analysis to distinguish the microseismic event signal from the noise improved their results as it allowed less room for accumulating error. We begin in a similar fashion by applying an independent component analysis on segments of

20 data sets of the vehicle driving along the same 65m length of road next to a fibre-optic cable. Each segment contains 5 seconds and the entire 65 meters of fibre. We remove every 10th shot from the data to form a new data set. Then, we segment the new data. For the independent component analysis, we only consider two independent components: the signal of the vehicle and noise. In this case, we use segments from 20 data sets which show the data from the same segment of fibre in time to find the two independent components. We then use the GMM on the independent components containing the signal of the vehicle for each segment of the data.

Let us consider the values we chose for three parameters of the vehicle detector: the number of Gaussians modes, the number of training frames, and the learning rate. For this case, we decided to use three Gaussian modes. Given that our data is finite in time, we only get 27 segments from the each data set. Each independent component has 27 frames. We set the number of training frames for the vehicle detector to be 10. We use approximately 1/3 of the frames to train the vehicle detector. Finally, we chose the learning rate to be 0.0001. The learning rate affects how fast the model adapts to change in the parameters.

Before we examine the results from the vehicle detectors taught on the independent components, we train on the training set to get some idea of how the GMM should perform visually. Given that the training set is the same size as the data sets we used to determine the independent components, we set the number of training frames to be 10 in this case as well. This is approximately 1/3 of the frames from the training set, keeping in line with the percentage of frames used for the vehicle detector of the independent components.

After the detector is trained, the MATLAB example applies a morphological cleaning to the proposed masks of the vehicle for each frame. For our experiment, we apply an additional cleaning to the vehicle mask which bridges any 0-valued pixels in the mask which are surrounded by two or more 1-valued pixels as well as determining the value of a pixel by taking into account the majority of values around said pixel and forcing it to match. We wish to detect the vehicle the entire time so bridging and filling holes in the mask is not going to be an issue. Figure 9.3 presents the results of the vehicle detector trained on the training data set. The top figure shows the results after the initial cleaning and the bottom figure displaces the results after the second cleaning.

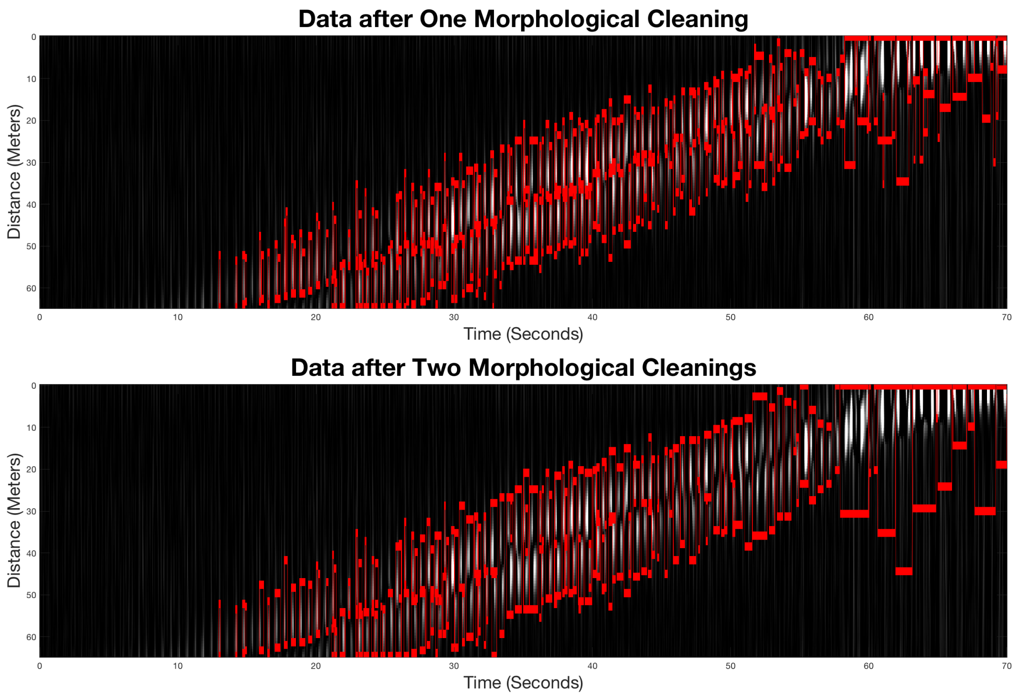


Figure 9.3: The vehicle detector trained on the training data applied to the test data after (top) one morphological cleaning and (bottom) two morphological cleanings.

The vehicle detector taught on frames from the training data set does not start detecting events until about 5 seconds after the vehicle begins to show up in the data. The first cleaning shows smaller boxes for events along the vehicle signal toward the beginning and larger bounding boxes at the end. Also, the size of the bounding boxes around events changes sporadically. The second cleaning shows that some of the smaller boxes around events in the signal have disappeared. The boxes still jump significantly in size toward the end of the data. While the signal still appears within the bounding boxes like we want, the large bounding boxes suggests that the vehicle is anywhere within an almost 30m distance. More precision in detection is desirable.

We consider the set of first independent components. Figure 9.4 shows the resulting first independent components from the segmented data. The yellow vertical lines distinguishes the different frames. With so many segments, it is difficult to determine anything specific about each segment. On close observation, some frames show signs of diagonal movement which would likely be the signal of the vehicle. We will consider the first four segments later in Figure 9.8.

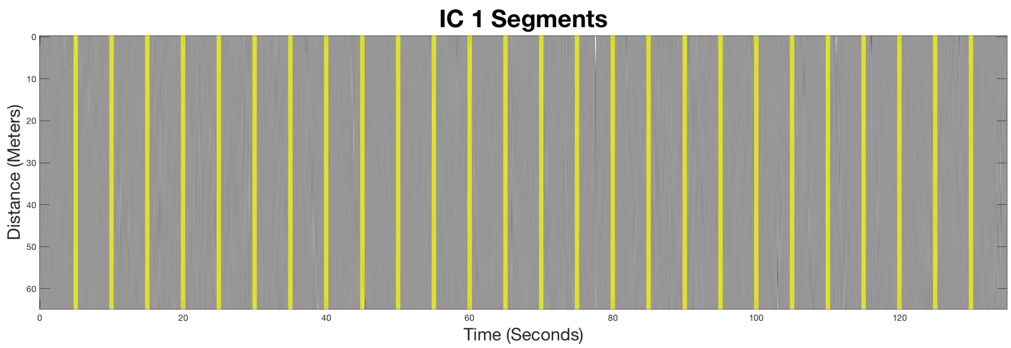


Figure 9.4: The 27 segments of the first independent component generated by performing independent component analysis on segments from 20 data sets containing the same 65m segment of fibre over 135 seconds.

Now, we study the results of the applying the vehicle detector trained on the first independent component to the test data. Figure 9.5 displays the results after the first morphological cleaning on the top and the second morphological cleaning on the bottom. Neither case detects the vehicle's signal as successfully as the test case (seen in Figure 9.3). The vehicle detector produced a tighter fit to the vehicle's signal. Most of the bounding boxes around the signal are approximately the same size, which implies that this result gives a better idea of where the vehicle is located along the 65m stretch. The second morphological cleaning shows that the vehicle detector located more of the vehicle's signal after a second cleaning.

Given that we performed the ICA on segments of data, we also consider how the GMM acts on the second independent components from the segmented data as we cannot guarantee the vehicle signal was only delegated to the first component in each segment. Figure 9.6 presents the second independent components of each segment where the yellow vertical lines

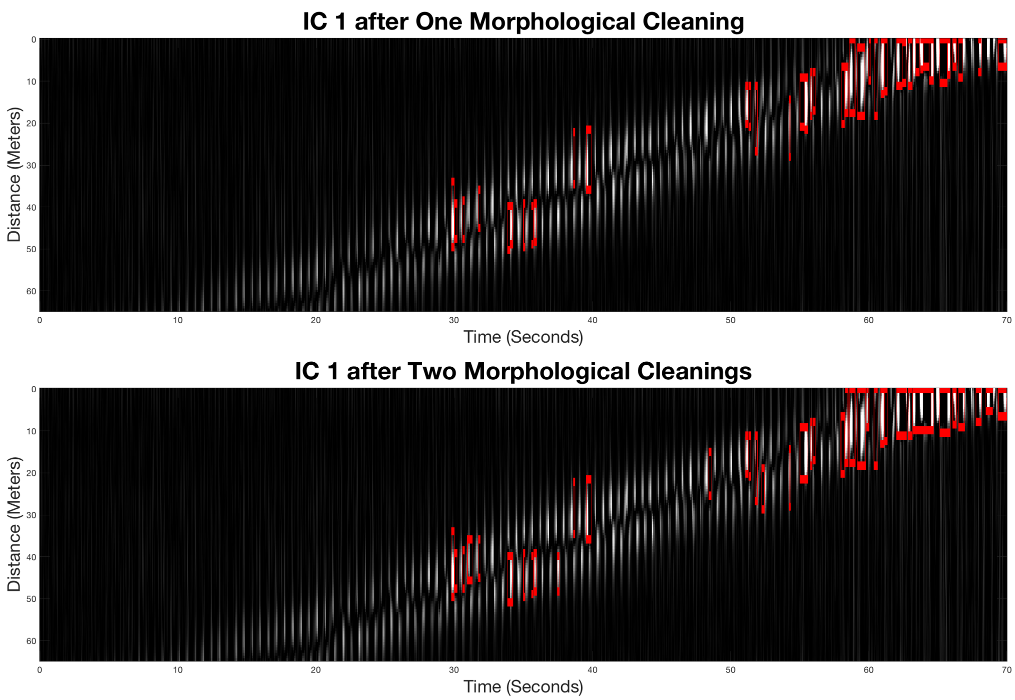


Figure 9.5: The vehicle detector trained on the first independent component for each segment of data applied to the test data after (top) one morphological cleaning and (bottom) two morphological cleanings.

distinguish between the second independent component frames. While difficult to see once again, we can observe some diagonal movement in some of the segments which potentially represents the vehicle moving along the fibre. As with the first independent component, we will consider a few frames of the second component later in the section.

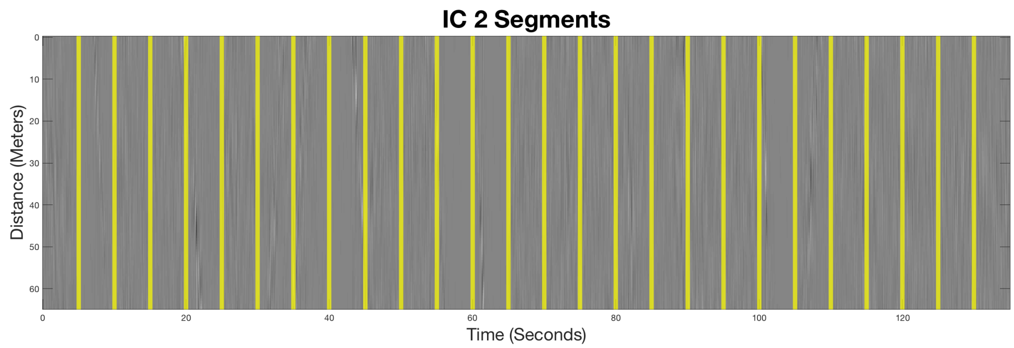


Figure 9.6: The 27 segments of the second independent component generated by performing independent component analysis on segments from 20 data sets containing the same 65m segment of fibre over 135 seconds.

Figure 9.7 displays the results of the vehicle detector trained on the second independent component segments applied to the test data set. The top image shows the results after the first morphological cleaning and the bottom row presents the results after the second cleaning.

The vehicle detector taught using the second independent component segments appears to perform similarly to the vehicle detector trained on the first independent component segments on the test data set. Both sets begin to detect the vehicle's signal at around 30 seconds. The bounding boxes are similarly compact around the signal of the vehicle for this case as well, giving a better estimate of where the vehicle is along the 65m segment of fibre. For the second independent component, it is more evident that the second morphological cleaning improved the results of the vehicle detector as more of the signal is detected between 30 and 45 seconds in that case. The relative success of the vehicle detector for the second independent component on the test data supports our hypothesis that some of the vehicle's

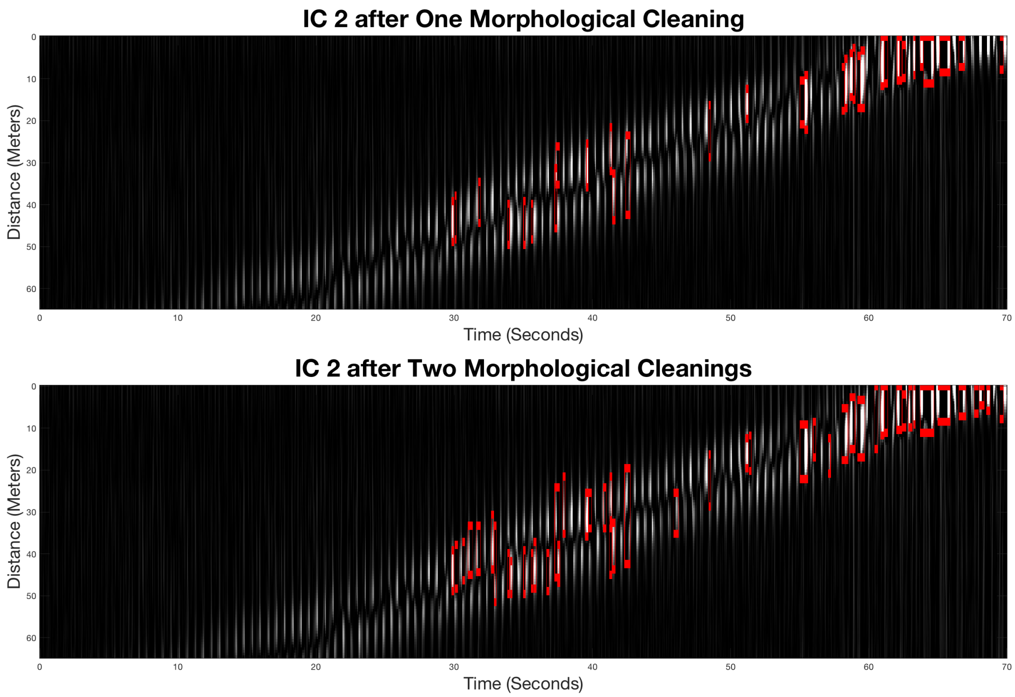


Figure 9.7: The vehicle detector trained on the second independent component for each segment of data applied to the test data after (top) one morphological cleaning and (bottom) tow morphological cleanings.

signal was delegated to the second component during the independent component analysis of segments from the twenty data sets.

Figure 9.8 presents the two independent components for the first five segments of data. The first independent components are in the left column and the second independent components are on the right column. We chose two independent components because we wanted to distinguish between the noise and the vehicle signal. The first three frames of the first independent component clearly shows the signal of the vehicle. While the first and third frames of the second independent component are noise, the second frame of the second independent component depicts the signal of the vehicle moving at a higher velocity along the 65m segment of fibre than the vehicle signal in the first component.

The vehicle signal carries through consecutive frames in the same independent component with relative continuity; however, the vehicle signal is not necessarily delegated to the same

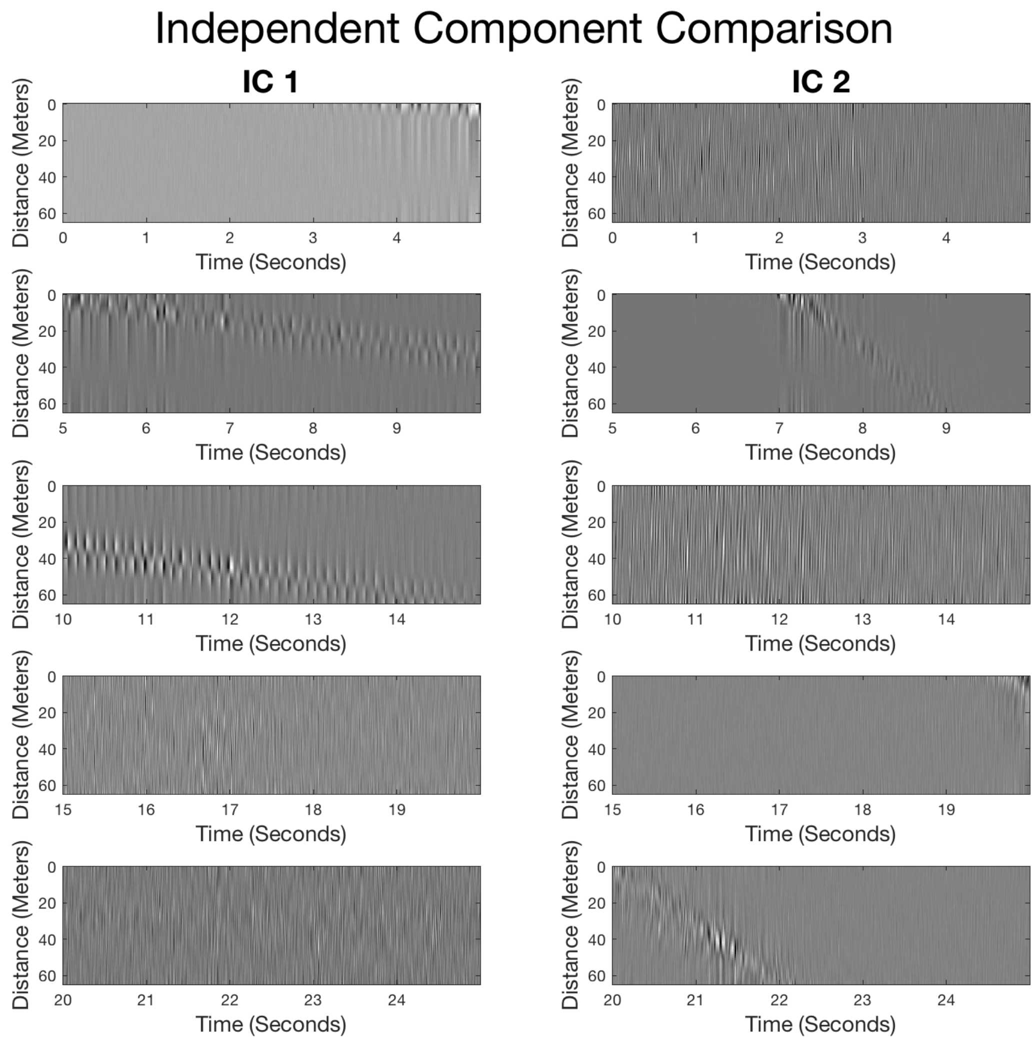


Figure 9.8: A comparison of the first five frames of the two independent components. (Left column) The first five frames of the first independent component, covering the time between 0 and 20 seconds. (Right column) The first five frames of the second independent component, covering the time between 0 and 20 seconds.

independent component once that journey along the fibre is complete as can be seen from the fourth and fifth frame of the second independent component. This supports our claim from earlier that training on the second independent component was appropriate. It should also be noted that while these conditions hold true for the first five frames of the independent components, it is not necessary that these conditions hold true for the remaining twenty-two frames.

Another clear issue with this example is the amount of frames on which the vehicle detector can train the background model. As we discussed in Section 9.1, the vehicle detector needs a number of frames containing only the background before introducing the vehicle signal. Given the performance of all three vehicle detectors in this example, it is essential that we include more training frames, but we also need to ensure that the vehicle detector trains only on background data for the first few frames. While we could employ data sets which gathered data over a longer time period, an easier method finds itself in using the large portion of the data sets to find the independent components and overlapping frames of the data instead of using 5 second disjoint segments. In the next section, we consider the results of independent components from a large segment of multiple data sets instead of multiple segments.

9.4.2 Case 2: GMM applied to independent components from multiple data sets

We now calculate the first and second independent components by applying an ICA to a portion of all twenty data sets which is the same size as the training data and the test data before applying a GMM to distinguish the vehicle from the background. Instead of the entire 135 seconds of data, we only consider 70 seconds of data. Figure 9.9 shows the two independent components determined by the ICA.

The first independent component displays the signal of the vehicle driving along the fibre whereas the second independent component contains the noise. We follow a similar

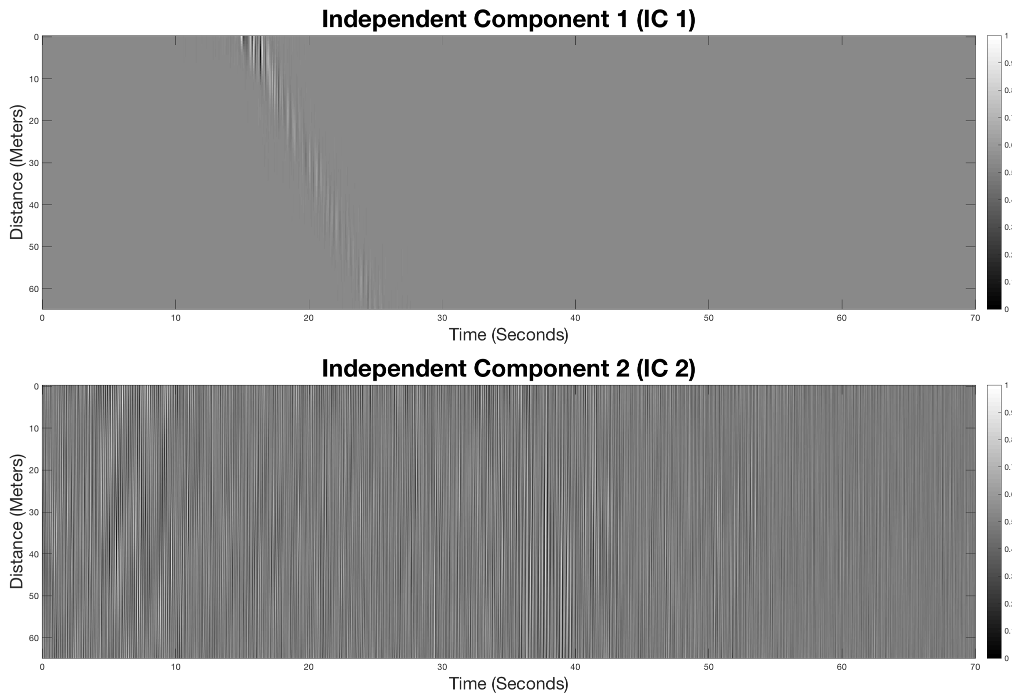


Figure 9.9: (Top) The first independent component produced by conducting an independent component analysis on a large segment of twenty data sets generated by the same section of fibre. (Bottom) The second independent component produced by conducting an independent component analysis on a large segment of twenty data sets generated by the same section of fibre.

pattern to the previous example. We start by considering the results of the vehicle detector trained on the training data set, then the results of the vehicle detector trained on the first independent component, and finally the results of the vehicle detector trained on the second independent component. Afterwards, we set the parameters for each vehicle detector the same as the previous case; however, each detector is taught using 1857 frames, which was approximately 1/3 of the total frames available for each case.

Figure 9.10 presents the results of the vehicle detector trained on the training data. In this case, we train the vehicle detector on overlapping frames of the training data and then consider how well it performs on the fourteen segments of the test data as we did in the previous case. Each of the overlapping frames is the same size as the segments from the previous example: 5 seconds by 65 meters.

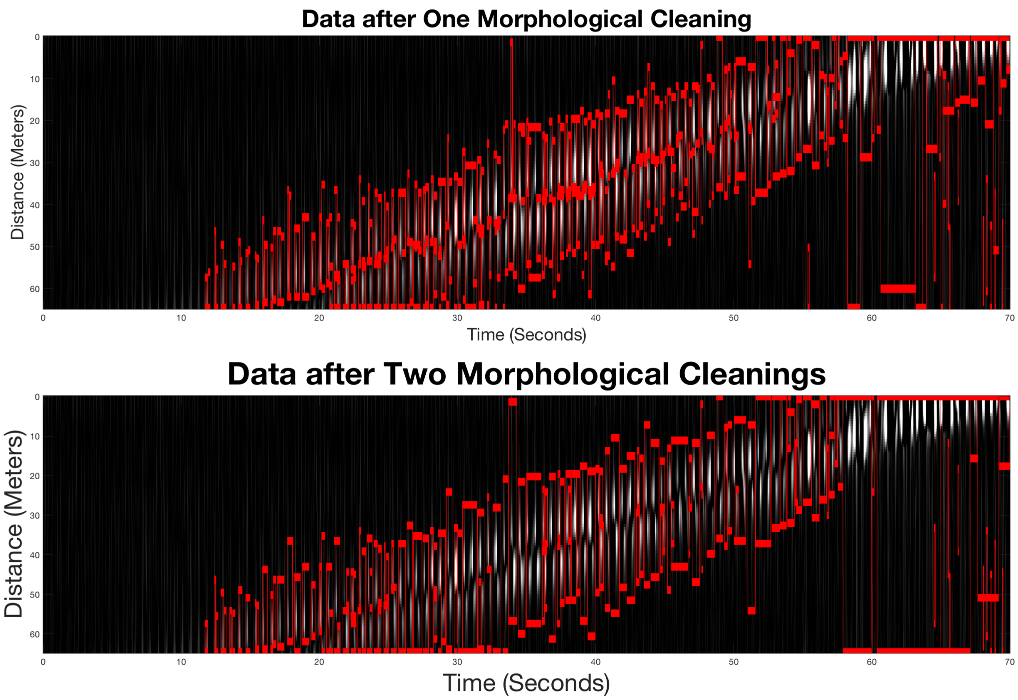


Figure 9.10: The vehicle detector trained on the training data applied to the test data after (top) one morphological cleaning and (bottom) two morphological cleanings.

The top row of Figure 9.10 shows the results of the vehicle detector after one morphological cleaning, and the bottom depicts the results after two morphological cleanings. While the vehicle detector does a decent job of detecting the signal from 15 seconds onward for both cleanings, the bounding boxes are quite large especially towards the end. More precision around the signal would be convenient for a better judgment of exactly where the vehicle is along the fibre. The second morphological cleaning reduced the number of boxes by combining them together; however, that is about all that can be determined from observation.

Now, let us consider how the vehicle detector taught on overlapping segments of the first independent component fares for finding the signal of the vehicle; c.f. Figure 9.11.

This detector performs relatively similarly to the last detector which was trained on the training data. One difference between the two cases is that the bounding boxes are smaller for this vehicle detector, which is especially visible in the final few boxes. Given that the bounding boxes are more concentrated around the vehicle signal suggests that the ICA

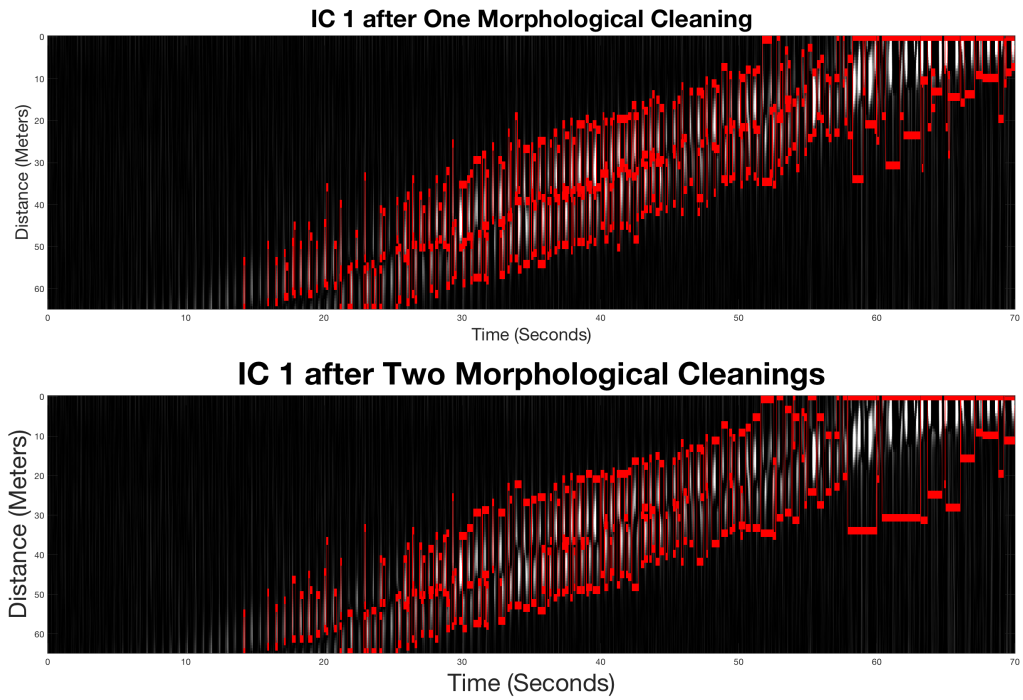


Figure 9.11: The vehicle detector trained on a window moving over the first independent component applied to the test data after (top) one morphological cleaning and (bottom) two morphological cleanings.

performed well with regards to separating the signal from the noise. The main difference between the top row and bottom row of Figure 9.11 is that the bottom row has fewer boxes. In particular, the second morphological cleaning forces smaller boxes to combine.

For the final comparison, let us consider how the vehicle detector which trained on the second independent component performs on the test data; see Figure 9.12. In the first morphological cleaning, the detector appears to observe something other than the signal of the vehicle between 5 and 10 seconds. After the second morphological cleaning, this bounding box disappears. Perhaps there is more noise than signal in this segment of data, which may be explained by the fact that this detector was trained on the second independent component which contained more noise than the first independent component. This detector did not perform as well with regards to detecting the vehicle's signal, unlike the vehicle detector trained on the first independent component. The detector for the first independent component consistently determines the vehicle signal at around 15 seconds whereas the vehicle detector for the second independent component does not consistently detect the vehicle signal until after about 25 seconds.

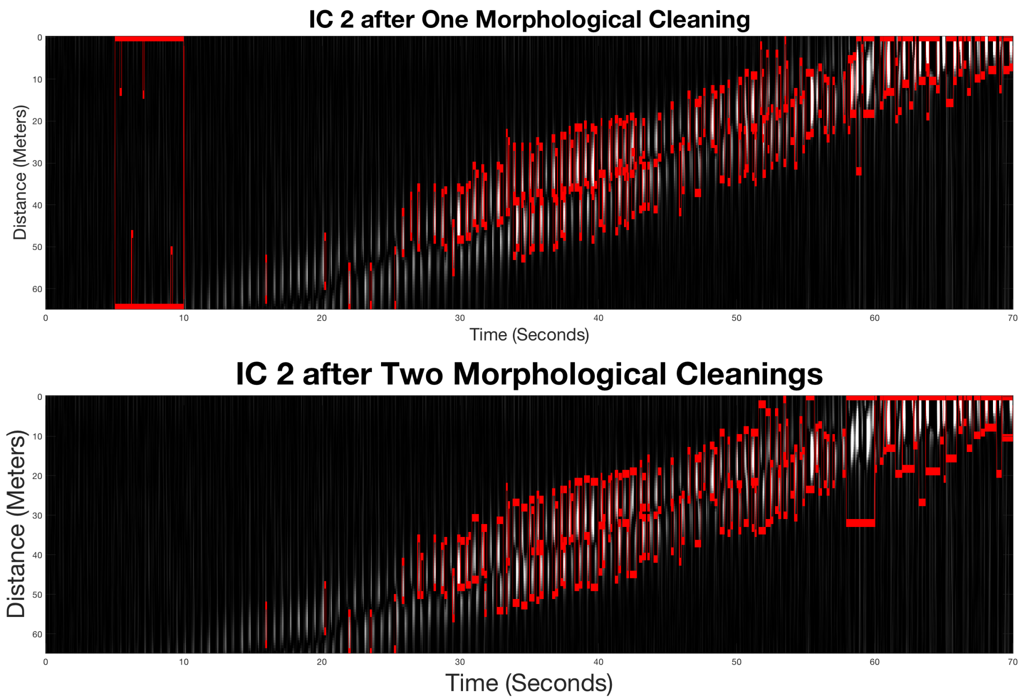


Figure 9.12: The vehicle detector trained on a window moving over the second independent component applied to the test data after (top) one morphological cleaning and (bottom) two morphological cleanings.

9.4.3 Case 3a: GMM applied to independent components from segments of multiple specific data sets

For Cases 1 and 2, we employed twenty data sets where only five contain noise and the remaining fifteen contains the vehicle's signal. Using so many data sets with the vehicle's signal may have given the independent component analysis some trouble, despite it being the same vehicle generating the signal at different time periods. In the following sections, we eliminate this potential issue for the ICA by determining the independent components based on multiple data sets where only one data set contains the vehicle's signal but the remaining data sets contain only noise. We repeat the same approach as we did in the previous two sections: we start by repeating Case 1 from Section 9.4.1. In that case, we applied the ICA to several segments of multiple data sets. We segment five data sets and apply the ICA to the segments: one data set contains the vehicle's signal and the remaining four data sets contain only noise.

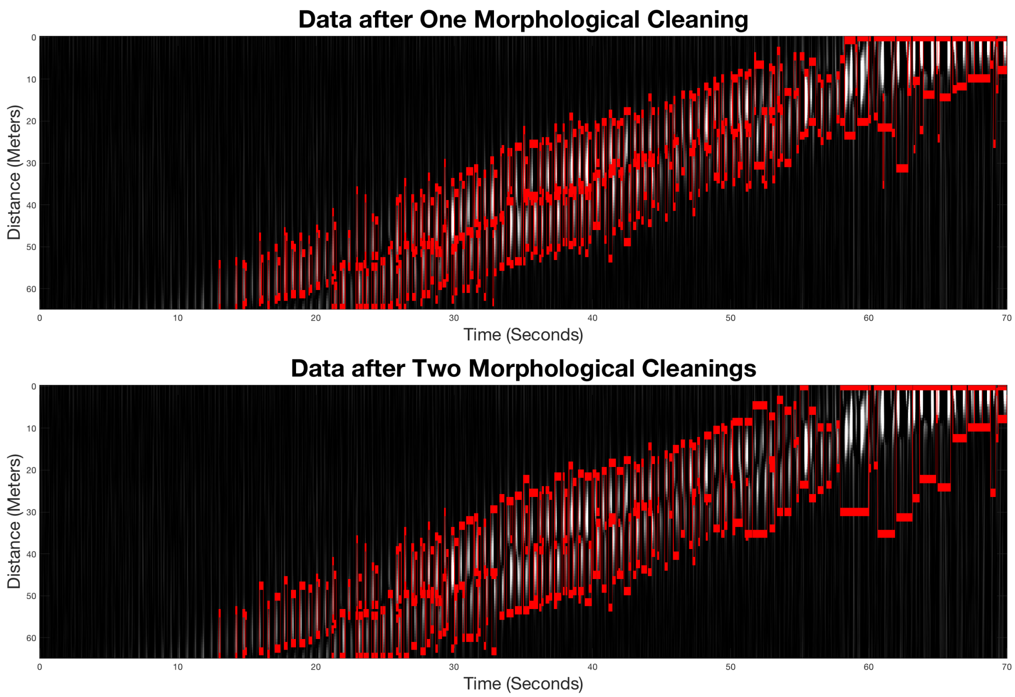


Figure 9.13: The vehicle detector trained on a window moving over the second independent component applied to the test data after (top) one morphological cleaning and (bottom) two morphological cleanings.

Figure 9.13 shows the results of applying the GMM to the training data. As in Case 1, the training data is cut into 27 segments and the GMM is trained on 10 of these segments. Visually, the results look very similar to the results from Case 1.

The ICA generated two independent components for each of the twenty-seven segments. In Figure 9.14, we find the results for the first independent component. We distinguish some diagonal movement, which would represent the vehicle in some segments between 20s and 40s.

Figure 9.15 shows the results of the GMM trained on the first independent components seen in Figure 9.14 applied to the test data. In this case, we employed five Gaussian modes

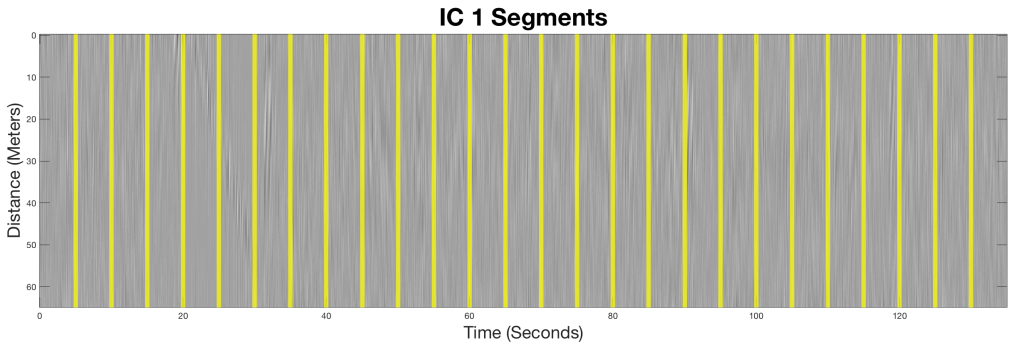


Figure 9.14: The 27 segments of the first independent component generated by performing independent component analysis on segments from 5 data sets containing the same 65m segment of fibre over 135 seconds.

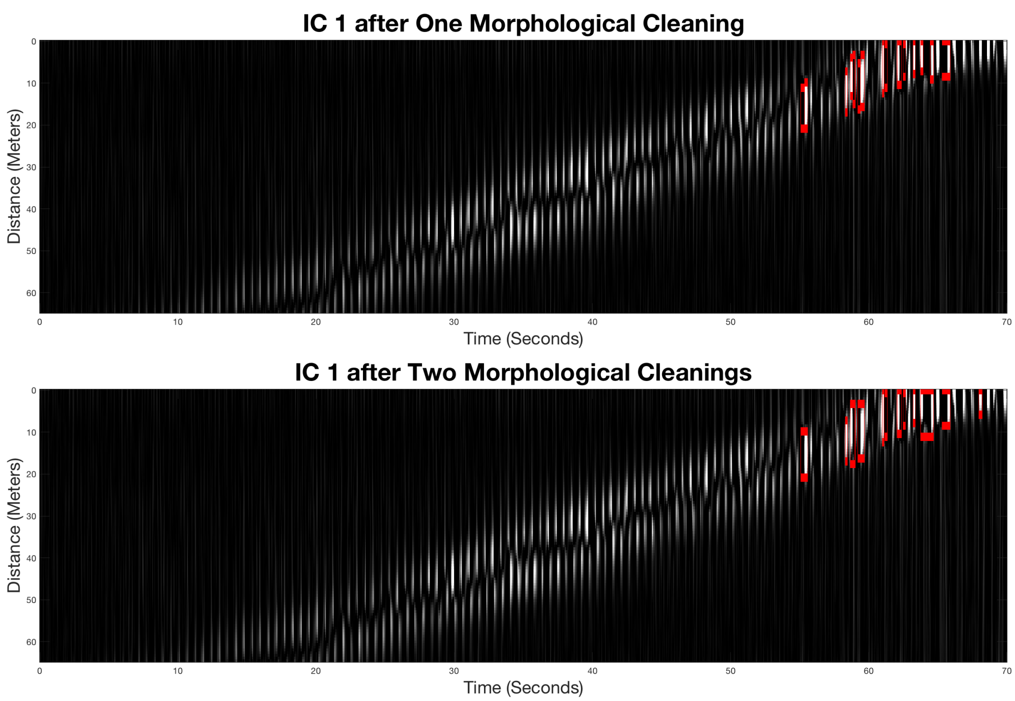


Figure 9.15: The vehicle detector trained on a window moving over the first independent component applied to the test data after (top) one morphological cleaning and (bottom) two morphological cleanings.

instead of the three we used in the previous cases. We also applied a stricter morphological cleaning this time because when we utilized the same cleaning as in the previous cases, we found that the GMM was prone to detecting the noise over the signal. This suggests that generating independent components from data sets where a majority train only on noise and only one contains the vehicle's signal may bias the GMM to detect noise over the signal.

Consider the second independent component for each segment in Figure 9.16. It is difficult to identify any segments which contain diagonal movement that would represent the vehicle. This suggests that the ICA mainly allocated the noise to the second independent component.

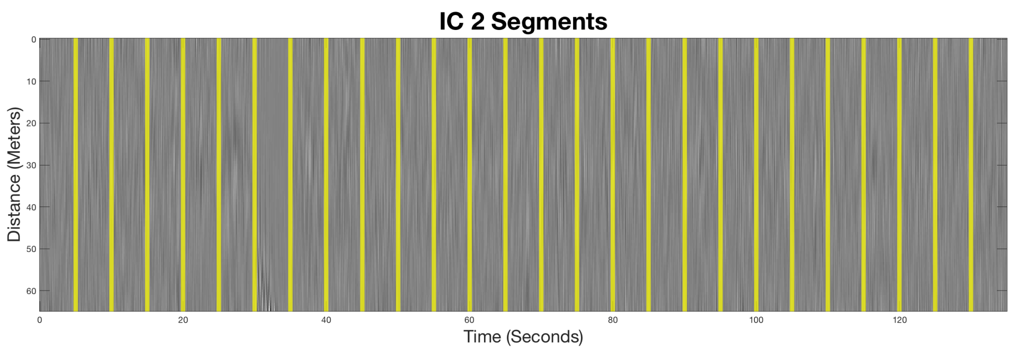


Figure 9.16: The 27 segments of the second independent component generated by performing independent component analysis on segments from 5 data sets containing the same 65m segment of fibre over 135 seconds.

Figure 9.17 provides the results of the GMM trained on the second independent component of each segment. The GMM provides some detection of the vehicle signal; however, it does not detect the entire signal. Given its ability to detect any of the vehicle's signal, some of the signal must have been delegated by the ICA to the second independent component.

The use of more noisy data sets than data sets containing the vehicle's signal has evidently caused an issue with how well the GMM performs when detecting the vehicle signal in the test data. Figure 9.18 presents the first five segments of both independent components. The first independent component is in the left column and the second independent component is in the right column.

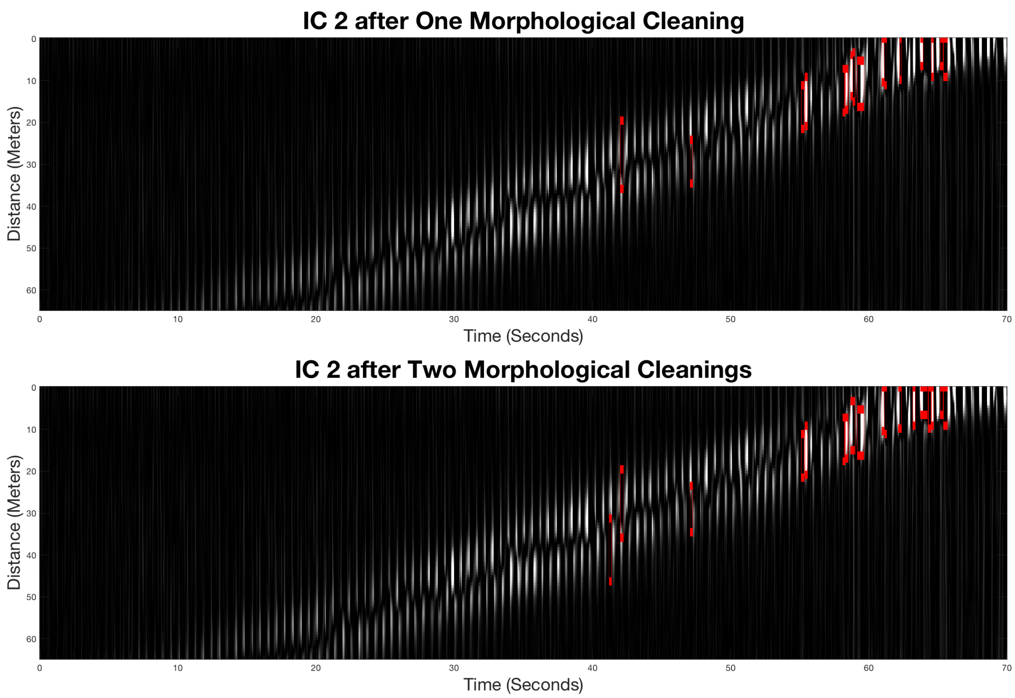


Figure 9.17: The vehicle detector trained on a window moving over the second independent component applied to the test data after (top) one morphological cleaning and (bottom) two morphological cleanings.

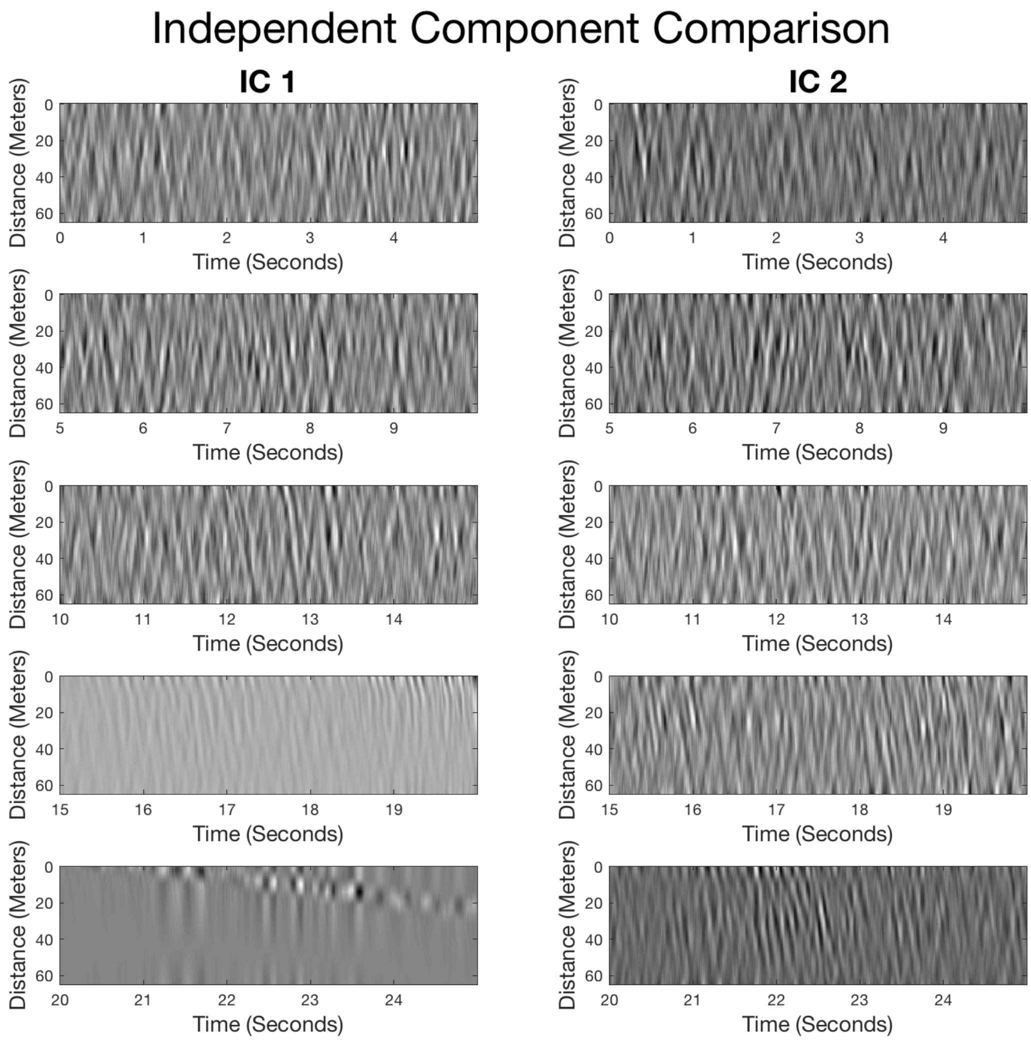


Figure 9.18: A comparison of the first five frames of the two independent components. (Left column) The first five frames of the first independent component, covering the time between 0 and 20 seconds. (Right column) The first five frames of the second independent component, covering the time between 0 and 20 seconds.

The vehicle signal appears in the last two frames for the first independent component, so the ICA is able to estimate the vehicle's signal and the noise segment. As with the previous cases, the vehicle signal is assigned to the same independent component for consecutive frames, but this is not a guarantee given that the ICA does not order the independent components (Hyvärinen and Oja, 2000). Using fewer data sets which contain the vehicle's signal means that more frames of both independent components contain only noise.

In the next case, we estimate the vehicle and noise signal based on only two data sets: one containing the vehicle's signal and one containing only noise. One issue that may continue to arise is that noise is assigned to both components when no vehicle signal is present in the noise; however, having an equal number of data sets for both cases may allow the GMM have less of a bias towards the noise.

Case 3b: Example using only two data sets

We consider the same segmentation of the data and apply the ICA to these segments as we did in the last case to estimate the vehicle's signal and the noise; however, this time we only include one data set with noise and one data set with the vehicle's signal. Figure 9.19 shows the results of applying the vehicle detector trained on the training data. The vehicle's signal is clearly detected after about 15 seconds in both morphological cleanings. As we saw previously, the second cleaning combines many of the bounding boxes from the previous cleaning. Like in Case 3a, we employ five Gaussian modes and apply a stricter morphological cleaning for all the vehicle detectors in this section. It was necessary to use such strict parameters in order to remove the noise, as otherwise, the vehicle detector was biased towards detecting the noise when trained on the independent components. We will see in a later section how it affects the GMM trained on the training data in each case.

We now observe how well training the GMM on the first independent component works for providing a more precise detection of the vehicle's signal. Figure 9.20 shows the first independent component that the ICA estimated from the 5s segments taken from the two

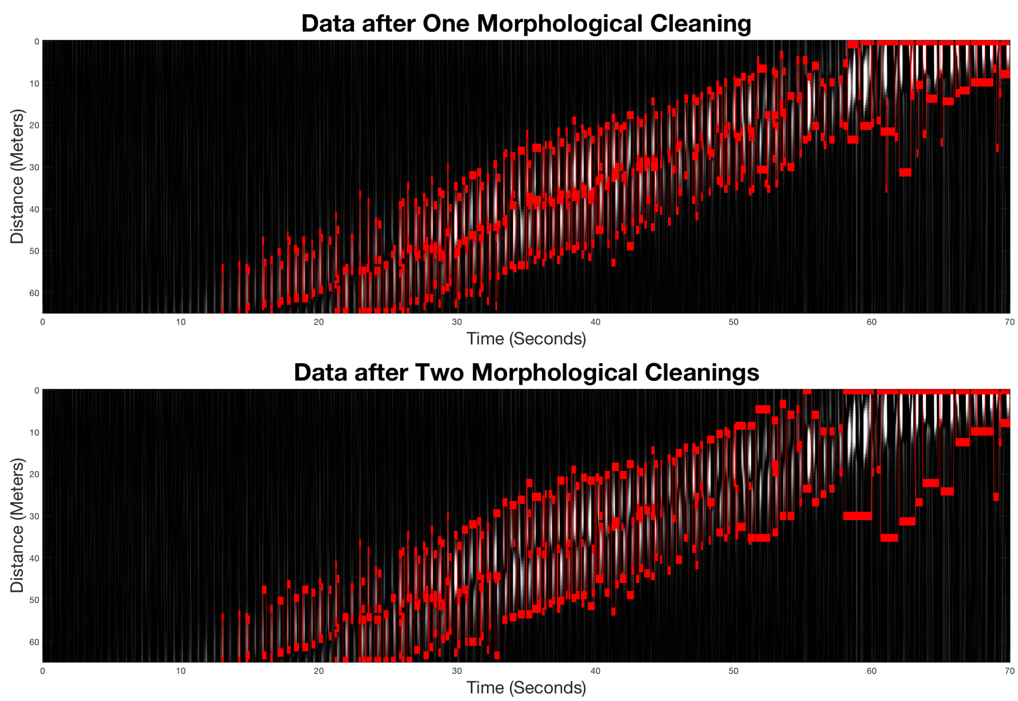


Figure 9.19: The vehicle detector trained on a window moving over the second independent component applied to the test data after (top) one morphological cleaning and (bottom) two morphological cleanings.

data sets. As with Case 3a, the vehicle's signal is present in frames residing between 20s and 40s.

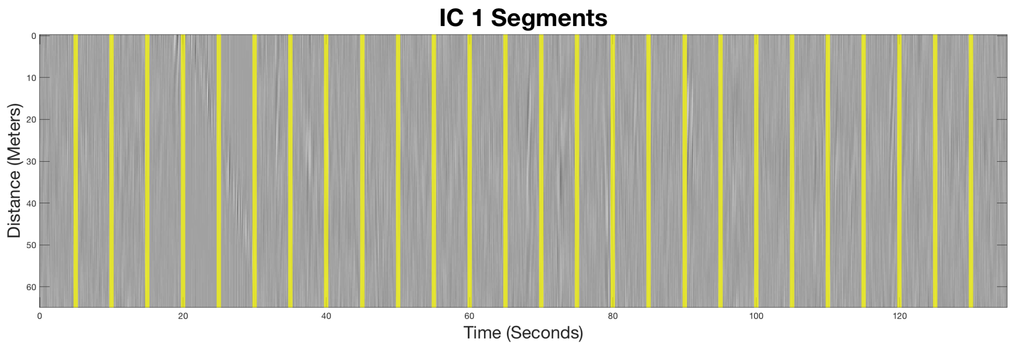


Figure 9.20: The 27 segments of the first independent component generated by performing independent component analysis on segments from two data sets containing the same 65m segment of fibre over 135 seconds.

Figure 9.21 presents the results of the GMM trained on the first independent component from each of the segments. We use 10 of the 27 frames to teach the GMM and then apply the vehicle detector to the test data. While the portions of the vehicle's signal that the GMM detects gives a better idea of the vehicle's location along the fibre, it does not detect nearly as much of the signal as the previous cases. In this case the results of the GMM trained on the training data is more beneficial despite its imprecision as it detects the vehicle for most of its journey across the 65m of fibre.

We examine the results of the vehicle detector taught on the second independent component. In Figure 9.22, we find the second independent component estimated by the ICA using the two data sets in this case. Unlike the first independent component, it is difficult to differentiate any segements which contain diagonal movement to signify the vehicle's signal, suggesting the noise was often allocated to this signal. We expect the GMM trained on the second independent component to provide poorer results than the one trained on the first independent component.

In Figure 9.23, this vehicle detector provides a comparable result to the vehicle detector

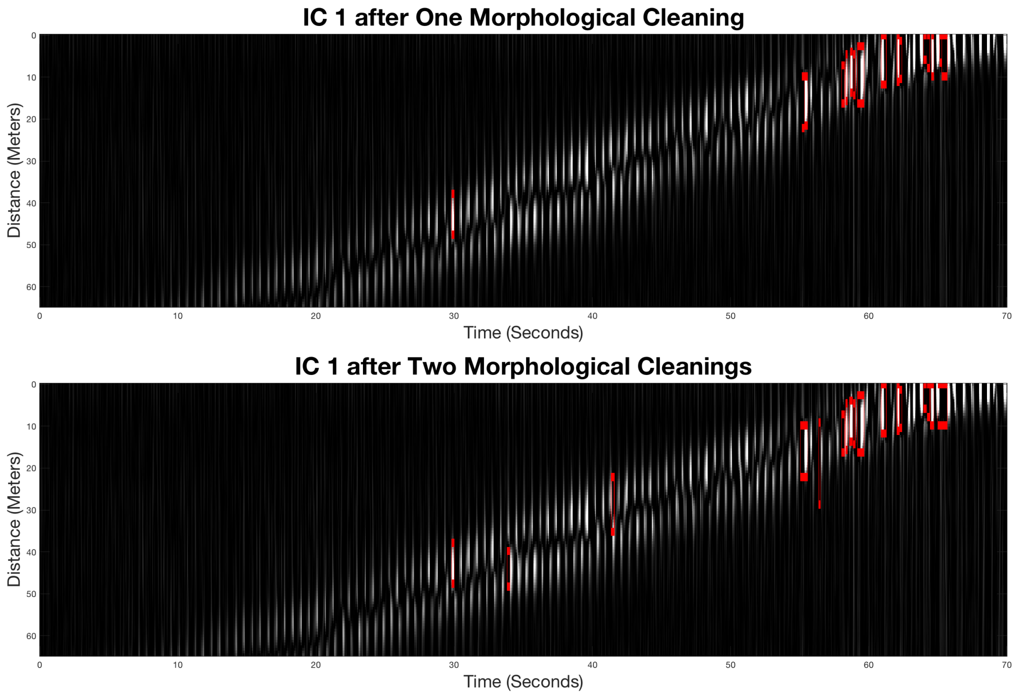


Figure 9.21: The vehicle detector trained on a window moving over the first independent component applied to the test data after (top) one morphological cleaning and (bottom) two morphological cleanings.

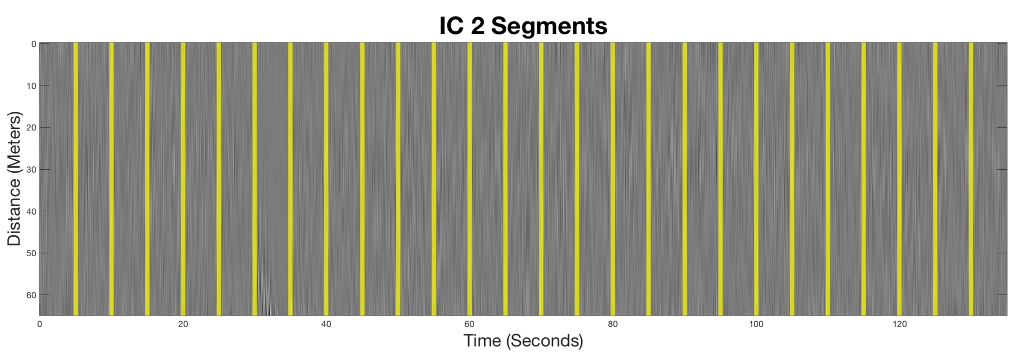


Figure 9.22: The twenty-seven segments of the first independent component generated by performing independent component analysis on segments from two data sets containing the same 65m segment of fibre over 135 seconds.

trained on the first independent component. The second morphological cleaning does make more noise evident given the two bounding boxes which do not appear near the vehicle's signal. This result offers some support for our hypothesis that the noise was delegated to the second independent component.

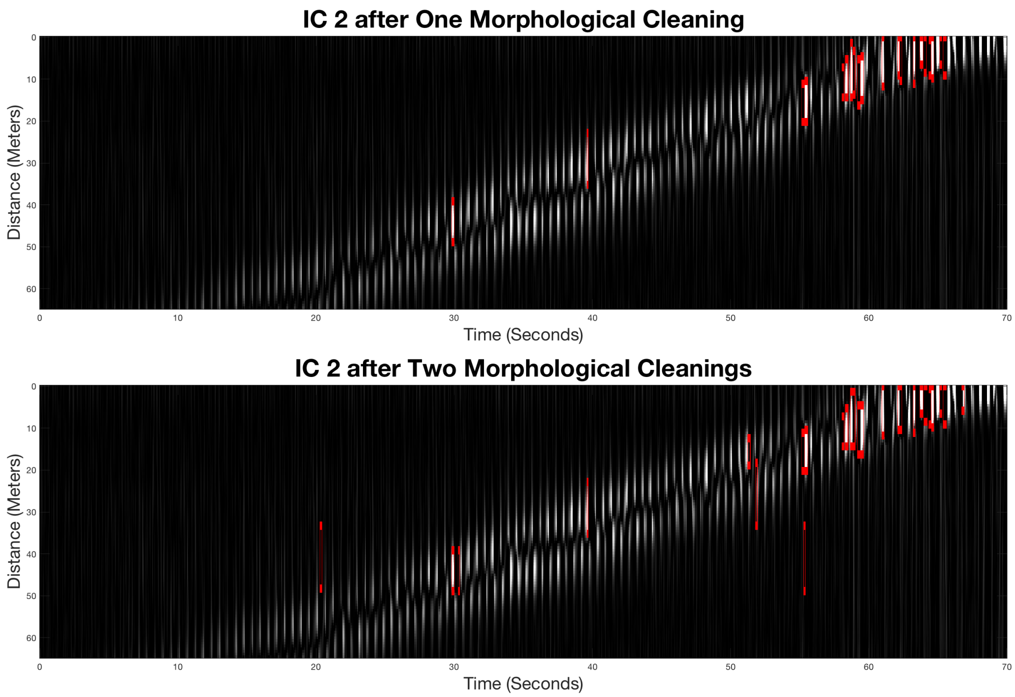


Figure 9.23: The vehicle detector trained on a window moving over the second independent component applied to the test data after (top) one morphological cleaning and (bottom) two morphological cleanings.

Figure 9.24 displays the first five frames of the independent components estimated by the ICA using the two data sets. Once again, the vehicle's signal is allocated the first independent component in the last two frames.

When the vehicle's signal is not present in a segment of the data, the ICA assigns noise to both independent components. Given how poorly the vehicle detectors trained on each independent components perform on the test data, it suggests that there are too few frames to train the GMMs adequately. This follows from the results in Case 1. Also, the fact that we had to make the parameters for the vehicle detectors in both Case 3a and 3b stricter suggests that the GMMs obtained a bias for detecting noise during their training. In the

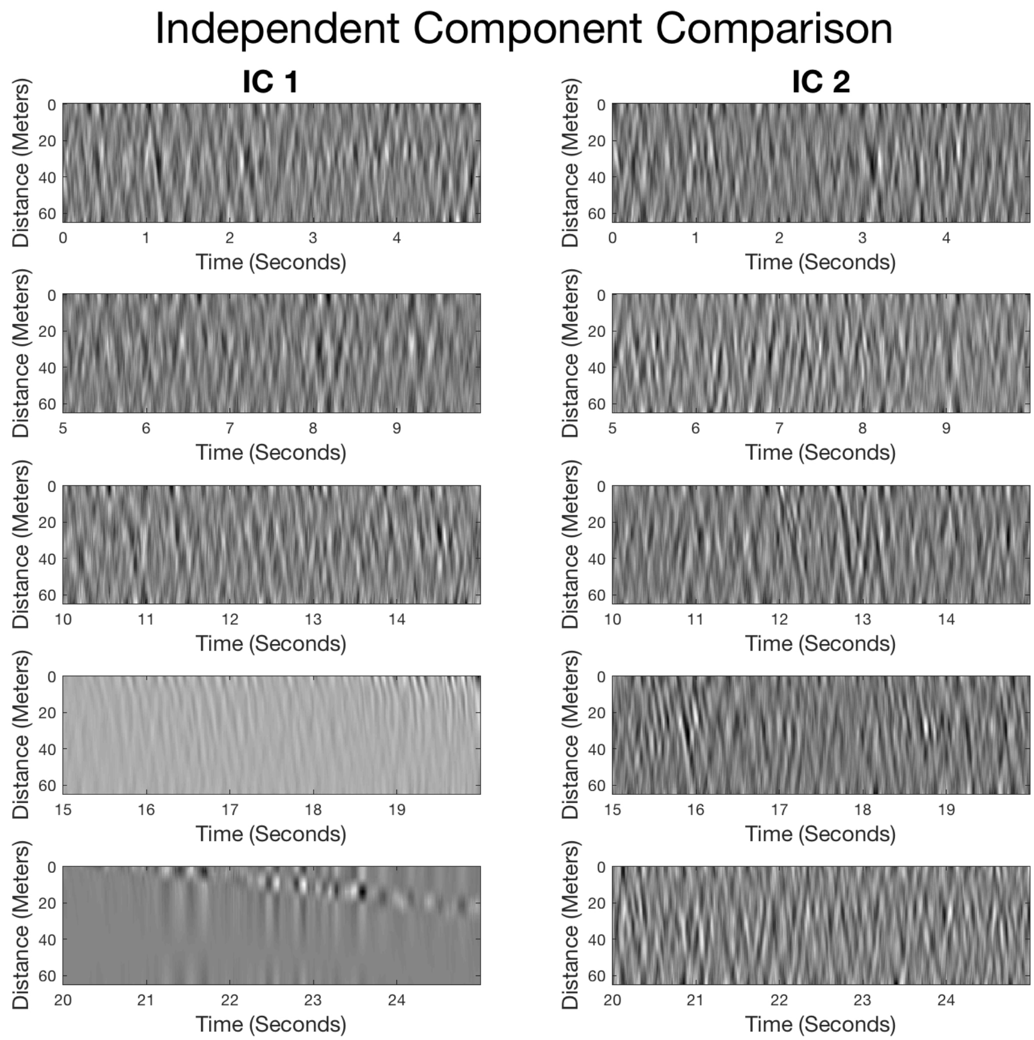


Figure 9.24: A comparison of the first five frames of the two independent components. (Left column) The first five frames of the first independent component, covering the time between 0 and 20 seconds. (Right column) The first five frames of the second independent component, covering the time between 0 and 20 seconds.

next section, we discuss two cases where we apply the ICA to a larger portion of multiple data sets.

9.4.4 Case 4a: GMM applied to independent components from multiple specific data sets

Given how poorly segmenting the data works for detecting the vehicle's signal, we now consider applying the ICA to a portion of the multiple data sets. We repeat Case 3a by using five data sets where one contains the vehicle's signal and the rest only contain noise. As with Case 2, we examine the results of a 70s portion of each data set. We use 1417 frames of data to train the GMM which correlates to approximately 10/27th of the available frames. Each frame is 5s by 65m. We return to our original parameters for the GMM: three Gaussian modes and the original morphological cleaning parameters.

Figure 9.25 provides the two independent components estimated from the five data sets by the independent component analysis. The ICA assigns the vehicle signal to the first independent component and the noise to the second independent component. We begin by considering how well a GMM trained on the training data performs on the test data, and we use this to see how well the GMMs trained on the first and second independent component, respectively, perform in comparison.

As with the previous cases, we see that the vehicle detector trained on the training data detects the vehicle signal relatively well. It begins detecting the signal at around 12s in both morphological cleanings. Towards the end of the test data, the bounding boxes are much larger. We study how the GMMs taught using the independent components performs with respect to locating the vehicle signal. We start with the first independent component which we saw in the top row of Figure 9.25. This component contains the vehicle's signal so it should produce a good detection of the signal.

Figure 9.27 shows the results of the vehicle detector taught using the first independent component. As with Case 2, employing a larger segment of the data to estimate the first

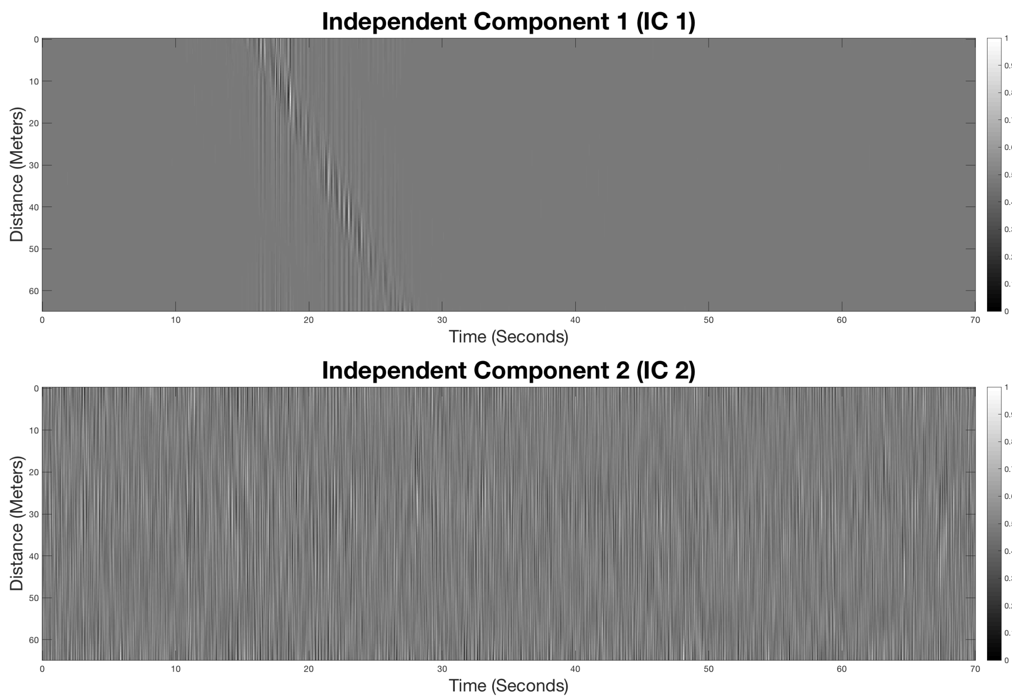


Figure 9.25: (Top) The first independent component produced by conducting an independent component analysis on a large segment of five data sets generated by the same section of fibre. Four of the data sets contained only noise and the remaining data set contained the vehicle's signal. (Bottom) The second independent component produced by conducting an independent component analysis on a large segment of five data sets generated by the same section of fibre. One data set contained only noise and the remaining data set contained the vehicle's signal.

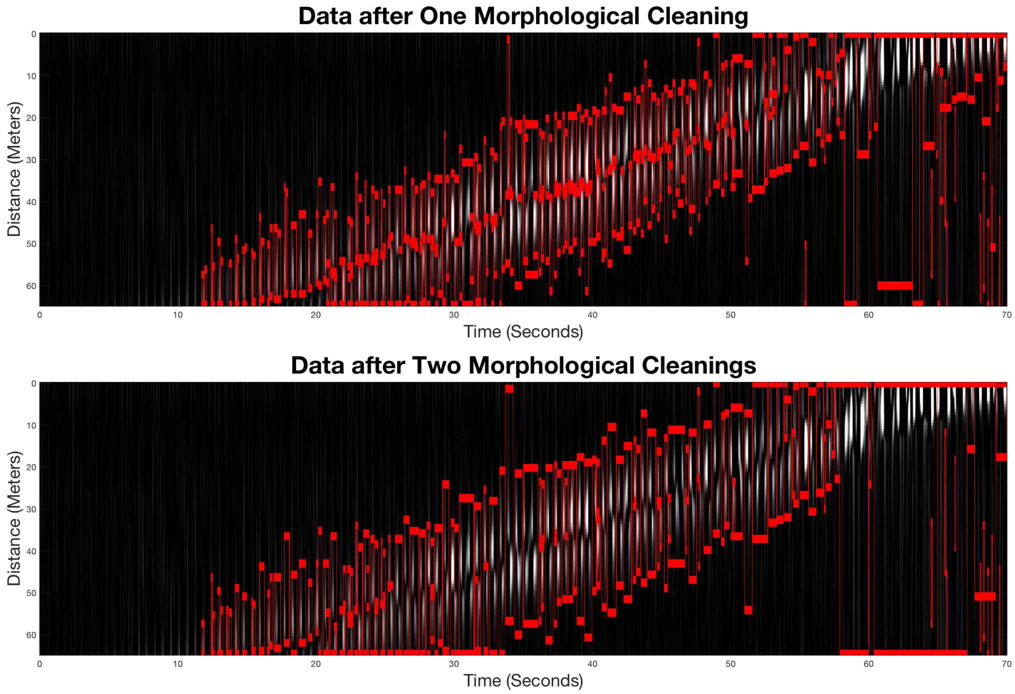


Figure 9.26: The vehicle detector trained on the training data applied to the test data after (top) one morphological cleaning and (bottom) two morphological cleanings.

and second independent component provided better results. The detector starts consistently detecting the signal after about 25s. The second morphological cleaning makes the bounding boxes larger as well as detects more of the vehicle's signal between 15s and 20s than the first morphological cleaning. While less of the signal is detected than when the GMM is taught using the training data, it still provides a better estimate of where the vehicle is than the segmented examples we saw in Cases 1, 3a, and 3b.

In Figure 9.28, we provide the results of the vehicle detector trained on the second independent component for completeness. The ICA assigned the noise to the second independent component; we see how that affects this vehicle detector's ability to detect the vehicle's signal. After the first morphological cleaning, a large bounding box is present at the beginning of the test data where there is no vehicle signal. This likely results from the vehicle detector being trained on the noise independent component. These bounding boxes disappear after the second morphological cleaning. Observe that the detector is still able to detect the ve-

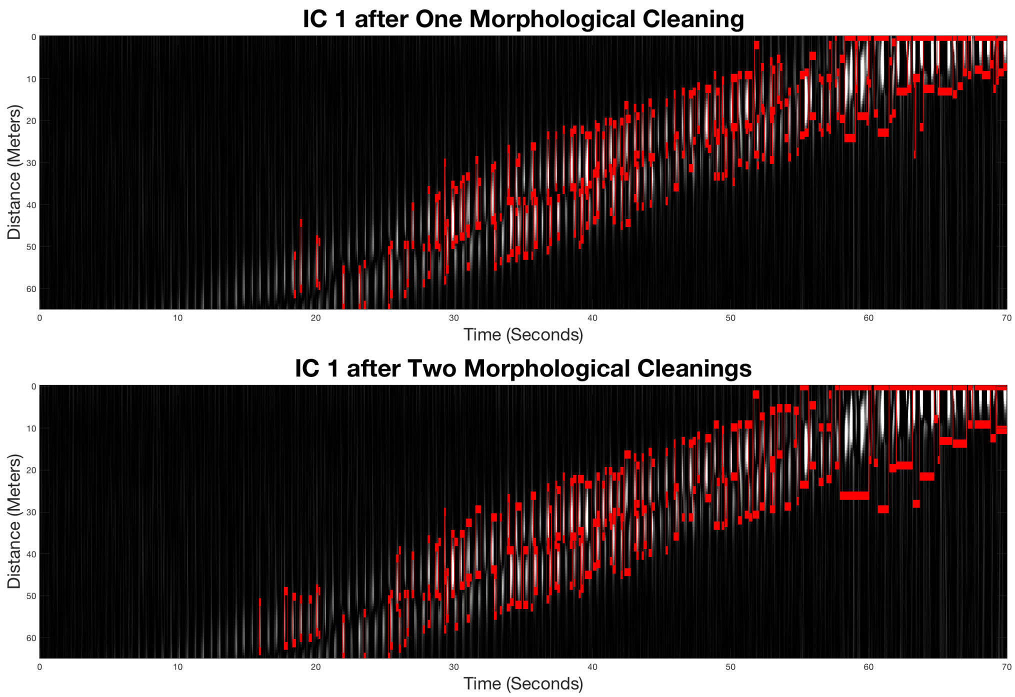


Figure 9.27: The vehicle detector trained on a window moving over the first independent component applied to the test data after (top) one morphological cleaning and (bottom) two morphological cleanings.

hicle's signal at around 25s and onward for both cleanings; however, it does not appear to be as consistent as the vehicle detector trained on the first independent component.

Case 4b: Example using only two data sets

In Case 4a, the ICA estimates the two independent components from five data sets where four data sets only contain noise and one data set contains the vehicle's signal. In this section, we study how well the vehicle detectors locate the vehicle's signal in the test data when the only two data sets are used to estimate the independent components. As we attempted in Case 3b, potentially only employing a balance of vehicle signal to noise will allow the vehicle detectors to remain unbiased towards detecting noise. We utilize the same parameters to train the GMMs as we did in Case 4a.

Figure 9.29 presents the ICA's estimates for the first independent component on the top row and the second independent component on the bottom row. Once again, the ICA

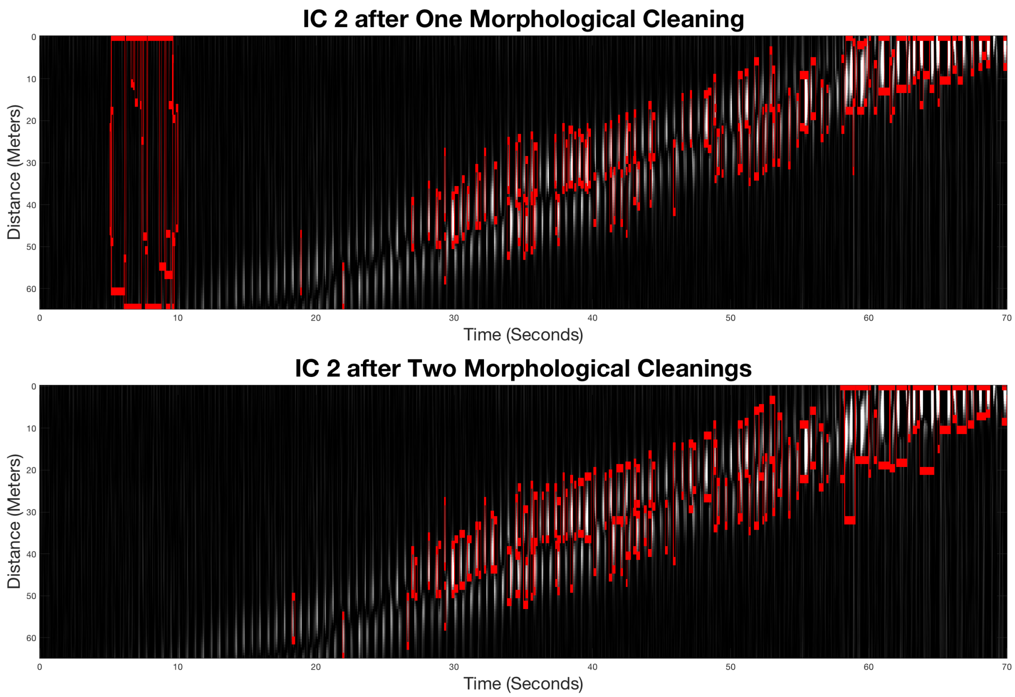


Figure 9.28: The vehicle detector trained on a window moving over the second independent component applied to the test data after (top) one morphological cleaning and (bottom) two morphological cleanings.

allocates the vehicle's signal to the first independent component. The vehicle's signal is evident between approximately 18s and 27s. Before we train the vehicle detectors on the independent components, we train a GMM on the training data for comparison. We train it on the final 70s of the training data for consistency with the length of the independent components in time. We chose the final 70s as it is the portion which contains the vehicle's signal.

In Figure 9.30, the results of the vehicle detector trained on the final 70s of the training data appears to have similar accuracy for detecting the vehicle's signal as the Case 2 and 4a. In Section 9.5, we offer a more definitive comparison of their results. For now, we note that the GMM still provides large bounding boxes at the end of the test data and a second morphological cleaning reduces the number of bounding boxes.

We now consider the results of the vehicle detector taught using the first independent component in Figure 9.31. This is the component that the ICA allocated the vehicle's signal

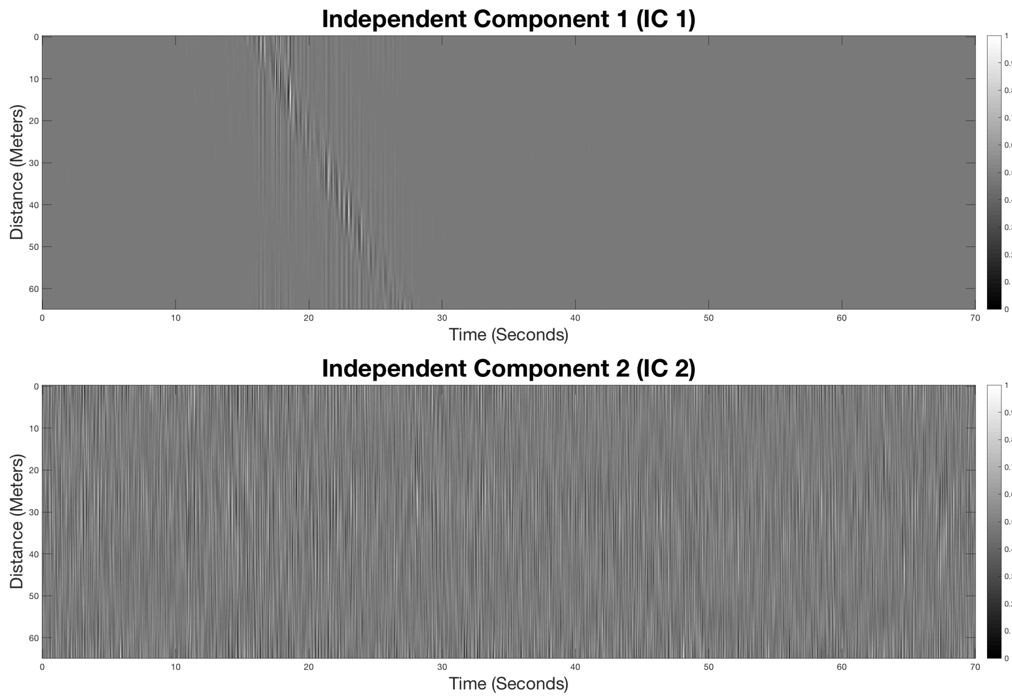


Figure 9.29: (Top) The first independent component produced by conducting an independent component analysis on a large segment of two data sets generated by the same section of fibre. One data set contained only noise and the remaining data set contained the vehicle's signal.
(Bottom) The second independent component produced by conducting an independent component analysis on a large segment of two data sets generated by the same section of fibre. One data set contained only noise and the remaining data set contained the vehicle's signal.

in Figure 9.29. The vehicle detector begins to notice the vehicle's signal consistently at around 25s. As with Case 4a, more bounding boxes appear before 20s once the second morphological cleaning is applied. In general, these results appear very similar to those from Case 4a for the first independent component.

Figure 9.32 shows that the results for the vehicle detector trained on the second independent component are also similar to those from Case 4b. Despite offering some balance in the number of data containing the vehicle's signal versus data which does not contain any vehicle signal, it did not appear to improve the results in this case. We do notice a drastic difference in the results between applying the ICA to large segments of the data versus smaller segments when we choose specific data sets to estimate the independent components.

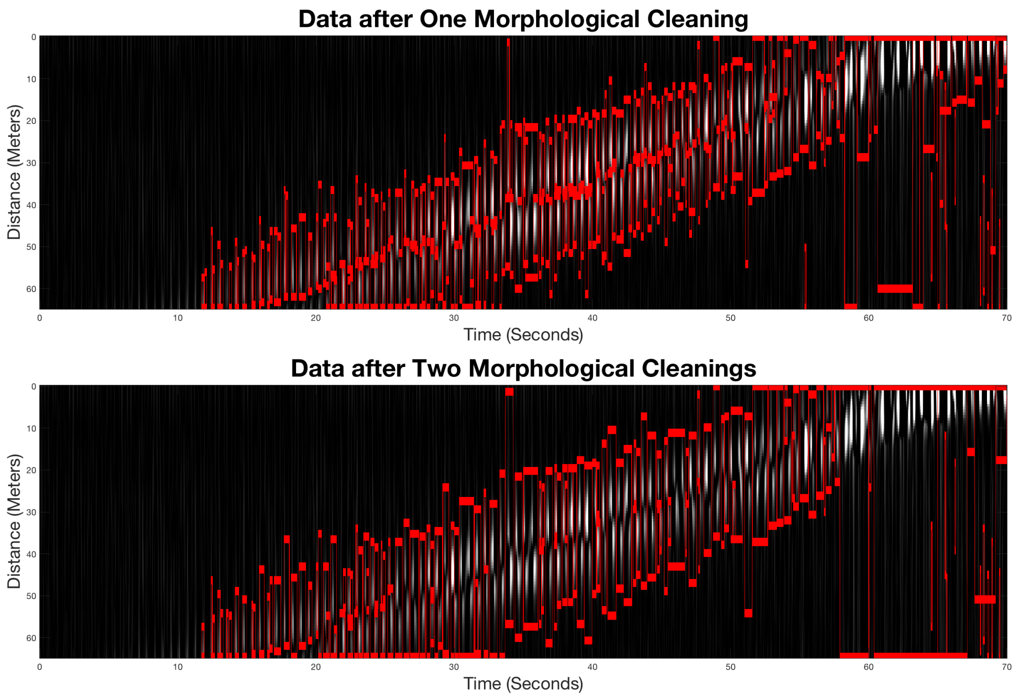


Figure 9.30: The vehicle detector trained on the training data applied to the test data after (top) one morphological cleaning and (bottom) two morphological cleanings.

In the next section, we offer a comparison of all the method discussed in this chapter and provide a more definitive idea of the most effective approach for detecting the vehicle.

9.5 Comparison of methods

Before we conclude, we offer a quantitative comparison of the six cases. In order to provide this comparison, we calculate how many points from the foreground mask are contained within the bounding boxes determined by each vehicle detector and divide the results by the total number of points in the foreground mask for each detector, which we call the overlap percentage. Figure 9.33 depicts six bar graphs comparing the percentage of points overlapped by each vehicle detector in each case. The case described in Section 9.4.1 is on the top left and the case from Section 9.4.2 is on the top right. Cases 3a and 4a are on the second row, left and right respectively. The bottom row contains the results from Cases 3b and 4b. We

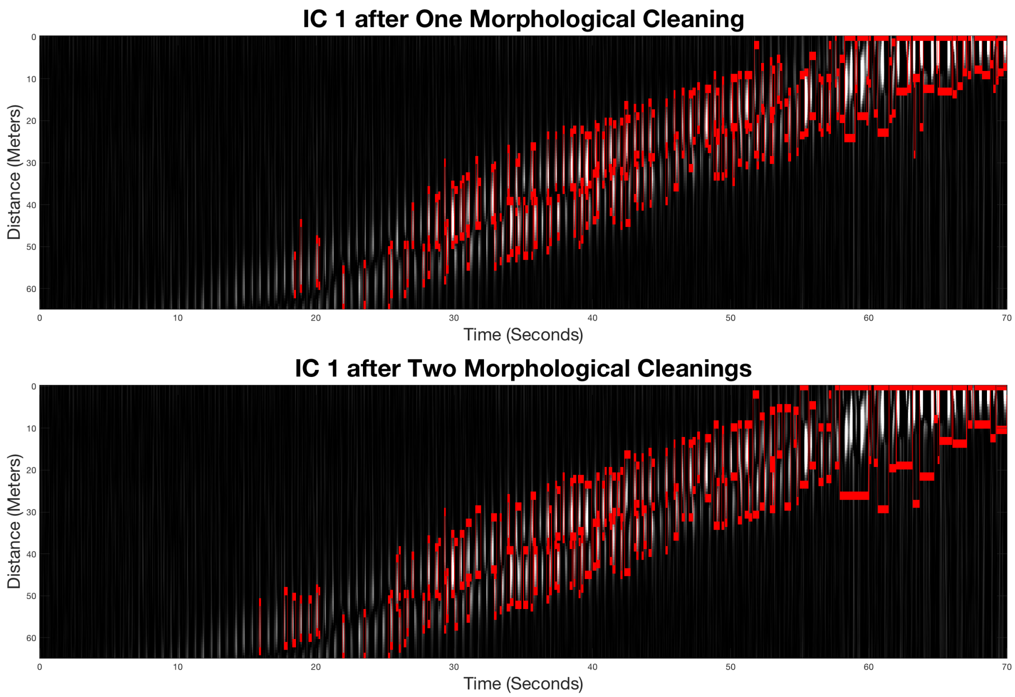


Figure 9.31: The vehicle detector trained on a window moving over the first independent component applied to the test data after (top) one morphological cleaning and (bottom) two morphological cleanings.

order the final two rows this way to provide a comparison of the small segmented data to the large segment of data given different number of data sets employed for the ICA estimates. The height of the bar in the graph provides an idea of how well the method performed on the data, i.e. the taller the bar in the graph, the more points from the vehicle's signal the method detected. The six bars in each graph depict the results of each application of the vehicle detectors to the data:

- (IC1-C1) the first independent component vehicle detector after one morphological cleaning,
- (IC2-C1) the second independent component vehicle detector after one morphological cleaning,
- (D-C1) the training data vehicle detector after one morphological cleaning,

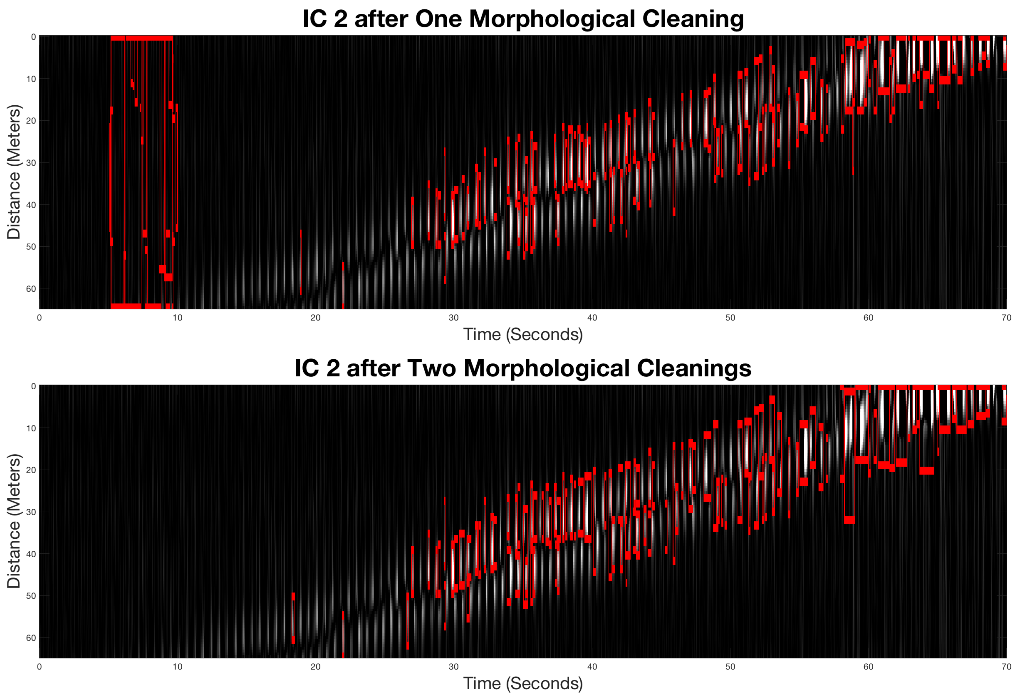


Figure 9.32: The vehicle detector trained on a window moving over the second independent component applied to the test data after (top) one morphological cleaning and (bottom) two morphological cleanings.

- (IC1-C2) the first independent component vehicle detector after two morphological cleanings,
- (IC2-C2) the second independent component vehicle detector after two morphological cleanings, and
- (D-C2) the training data vehicle detector after two morphological cleanings.

The training data vehicle detectors for all six cases performed better than any other method. The vehicle detector trained on the segmented training data did not detect the vehicle's signal as well as the vehicle detector trained on overlapping frames of the training data for either morphological cleaning. This follows given that the training set for the first data vehicle detector contained only 10 images compared to the 1857 training frames for the second data detector. This suggests that the number of training frames is important. The vehicle detector for the first independent component performed better than the vehicle

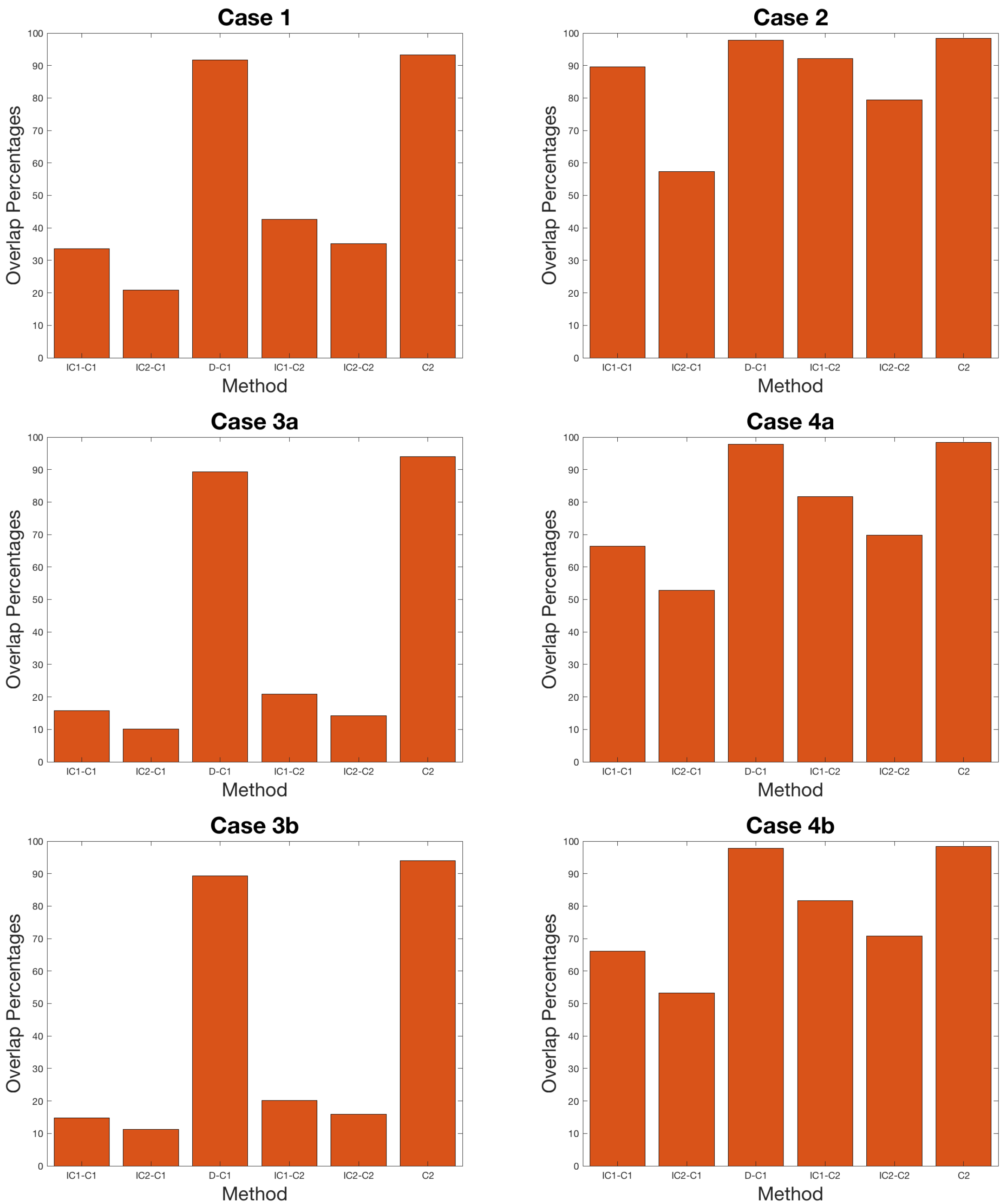


Figure 9.33: (Left) The percentage of points in the foreground which are contained by a bounding box for each of the six methods when we train on independent components generated by applying the ICA (top left) Case 1, (top right) Case 2, (middle left) Case 3a (middle right) Case 4a, (bottom left) Case 3b, and (bottom right) Case 4b.

detector for the second independent component in all six cases as well. These results support our visual observation that the vehicle's signal was often delegated to the first independent component in many of the cases.

9.6 Conclusions

In this chapter, we employed the ideas from Shamsa and Paydayesh (2019), which used Gaussian mixture models and independent component analysis on microseismic data to highlight events, in applications to data acquired a distributed acoustic sensor. The data contained the signal of a vehicle driving along a road. In each section, we applied independent component analysis to a different ratio of data sets containing the vehicle's signal and noise respectively to compute two independent components. The independent component analysis separated the vehicle's signal from the noise. In some cases, we computed the two independent components from non-overlapping segments, and in other cases, we performed the ICA on a 70s segment of data. Then, we trained a GMM on a training set and the two independent components.

We found that the vehicle detector trained on the training data statistically performed better than the detectors taught using the independent components. The vehicle detector trained on the independent component containing the vehicle's signal performed the second best when the independent components were generated from 70s segments of data. Performing ICA on 5s segments of data provided the worst results in every case.

For future work, we plan to explore how well this method works when multiple signals are present in the data; such as, data containing multiple vehicles or a vehicle and pedestrians. We expect that the independent component analysis will be beneficial for separating the different types of signal and will allow the GMM to isolate the specific signal in the data. Without separating the data with multiple signals, the GMM would likely detect each different signal at the same time without distinguishing between them. As such, the independent component analysis would provide the ability to train the GMM on a specific type of signal, which would be beneficial in monitoring projects among other applications.

Chapter 10

Image-recognition using convolutional neural network on DAS data

Using distributed acoustic sensing and fibre-optic cables to acquire seismic data enables the generation of terabytes worth of data in a short amount of time. While a plethora of data provides an abundance of information, it is necessary to interpret that data. In particular, seismic imaging and interpretation depends on finding events and anomalies in seismic data. With terabytes of data to consider, the creation of a new method beyond human capabilities becomes essential for analyzing data efficiently. One answer for this interpretation problem is through machine learning and image-recognition.

Over the last century, the advancement of computers led to the question of how to teach computers to perform tasks. Recently, image-recognition joined the ranks of such tasks for computers to learn. One aspect of machine learning that is especially useful in image-recognition are neural networks. Given the large nature of seismic data sets, convolutional neural networks are among the most effective.

In this chapter, we combine a convolutional neural network with an inverted wavelet tree in order to classify events in DAS events. While convolutional neural networks are very efficient, events in DAS data can be larger than a typical training image. They require some help with regards to reducing the scale of events. One answer to this problem is the tree-structured wavelet transforms we saw in Chapter 3. Recall that these transformationss apply a wavelet transform across a tree where each node of the tree represents the data convolved with a wavelet in the wavelet family. The wavelets themselves correspond to a specific frequency band, and consequently scale the original data to smaller sizes. Therefore, pairing convolutional neural networks with tree-structured wavelet transforms provides a scale-invariant image-recognition method. In particular, the inverted wavelet tree proves to be very useful for addressing scale-invariance.

In the following section, we introduce neural networks with a focus on convolutional neural networks (CNN). In the next section, we discuss two applications of convolutional neural networks to two similar image-recognition problems. In both examples, we discuss the architecture of the CNN created to detect events. From there, we consider the training set and test it on a portion of known events. For the first example, we introduce the microseismic data we use as the outside-samples for testing the first convolutional neural network (CNN) we built with the help of code from the UFLDL Deep Learning Tutorial at Stanford University; cf. Ng et al. (2013) and Yang (2014). We then apply both the inverted tree-structured wavelet transform and CNN to both sets of microseismic data and compare the results on each level of the tree. For the second data set, we then introduce the CNN we trained to identifying walking, digging, and noise events in DAS data. Next, we consider application to two data sets acquired at different pulse-repetition-frequencies using a Fotech Solutions interrogator which contain walking and digging events. Finally, we conclude.

The first convolutional neural network obtains an 85% accuracy when identifying single hyperbolic events. We find that using hyperbolic events such as walking steps to identify microseismic events is effective; however, decimation of the data affects the success of the convolutional neural network in distinguishing events. For multiple hyperbolic events such as walking and digging, the convolutional neural network drops to an 80% accuracy potentially due to the similarities between walking and digging data sets. We saw that the pulse-repetition-frequency (PRF) at which the data was collected also affects the success of the convolutional neural network. With a variety of possible obstacles affecting the outcome of the convolutional neural network, procuring a method which normalizes the distributed acoustic sensing acquired data is required in order to improve the success of a convolutional neural network for image-recognition purposes on DAS acquired data if we want it to work for data acquired using different PRF or other methods.

10.1 Convolutional neural networks

The natural world provides inspiration for mathematical and computational methods, and neural networks hold a strong resemblance to the human body's nervous system (Wu, 1992). Just as the body's nervous system interprets outside stimuli to determine how the body should react or decipher information, an artificial neural network utilizes 'neurons' to make decisions, or output answers, about information fed into the network.

An activation function $f : \mathbb{R} \to \mathbb{R}$ is utilized by the neural network to make decisions (Adcock and Dexter, 2019). There are standard activation functions used in the literature such as the sigmoid function, the identity, the arctangent function, and the hyperbolic tangent function to name a few. We use the sigmoid function as the activation function for our neural networks in this chapter as they do in Ng et al. (2013). The activation function is scaled using a weight W and translated using a bias b. The parameters W and b in each node of the neural network are optimized in a hypothesis function $h_{W,b}$ using a cost function J as part of the neural network's training. The hypothesis function $h_{W,b}$ is defined as

$$h_{W,b}(x) = f\left(\sum_{i=1}^{m} W_i x_i + b_i\right) \tag{10.1}$$

for m examples (x, y). For this chapter, the cost function J is the average of the least squares difference between the hypotheses $h_{W,b}$ of each sample x and the label y for that sample with a regularization term (Ng et al., 2013). Mathematically, we write the cost function J as follows:

$$J(W,b;x,y) = \frac{1}{2}\left\| h_{W,b}(x) - y \right\|^2 \tag{10.2}$$

where given m examples of samples paired with labels (x, y), we expand the cost function J to be:

$$J(W,b) = \left[\frac{1}{m}\sum_{i=1}^{m} J(W,b;x^{(i)},y^{(i)})\right] + \frac{\lambda}{2}\sum_{l=1}^{n_l-1}\sum_{i=1}^{s_l}\sum_{j=1}^{s_l+1}\left(W_{ji}^{(l)}\right)^2 \tag{10.3}$$

$$= \left[\frac{1}{m}\sum_{i=1}^{m} \frac{1}{2}\left\| h_{W,b}(x^{(i)}) - y^{(i)} \right\|^2\right)\right] + \frac{\lambda}{2}\sum_{l=1}^{n_l-1}\sum_{i=1}^{s_l}\sum_{j=1}^{s_l+1}\left(W_{ji}^{(l)}\right)^2. \tag{10.4}$$

The term λ is the weight decay parameter which helps prevent overfitting by decreasing the magnitude of the weights.

For the most part, the architecture of a neural network can be generalized to a connected graph with layers and units in each layer. Several different types of layers exist for neural networks, each of which performs different tasks. The type of neural network is dependent on how the units in these layers are connected to units of other layers. For larger data sets, it is important to limit these connections between units, as this decreases computational cost. One way to address the need for limiting connections is a convolutional neural network.

In a convolutional neural network, there is at least one convolutional layer. A convolutional layer has a number of units equal to the number of filters defined when training the network. The choice of filters determines the number of weights in the layer; it is also necessary to fix a dimension for the filters. For improved accuracy, the dimension of the filters should be the size of the event being trained to detect. In one of the examples in Section 10.2, we wish to detect hyperbolas extracted from a data set of someone walking parallel to a fibre-optic cable. The hyperbolas are approximately 91×91 matrix entries in size. As such, we set the dimension of the filters to be 91×91 when training the convolutional neural network. These filters are convolved through the data and produce a smaller set of convolved features. These convolved features save the computational cost of processing the entire data set in a fully-connected network.

After the convolutional layer, the results from each unit are directed to the next layer. This layer can be another convolutional layer or a pooling layer. Depending on the size of the data, it may be prudent to have another convolutional layer. A pooling layer takes the convolved features and pools them together using different methods to produce what is called a pooled feature. This layer condenses the data further, cutting down on computational costs and allowing for much faster classification. Some methods employed for pooling include max-pooling or mean-pooling. The max-pooling involves taking the maximum of each convolved feature and the mean-pooling takes the average of each convolved feature. Not only does pooling reduce the dimension of the data, but it also allows for less overfitting, which in turn improves results.

Depending on the complexity of the convolutional neural network, the pooling layer can feed into another convolutional layer or into another pooling layer. Our convolutional neural network will feed into a densely-connected output layer. The output layer gives the results or predictions of the neural network. For example, in hand-written digit recognition, there would be 10 output units: one for each integer between 0 and 9. In this chapter, one of our CNNs has two output units in the output layer: one for if an image has a hyperbola and another if it does not have a hyperbola. The other CNN has three outputs for walking, digging, and noise events.

Networks are trained by inputting learned parameters. The weights involved in the convolutional layer are part of these learned parameters. In order to learn the parameters for a neural network, it is necessary to have a training set. It follows that if a network must learn to predict events it needs training objects containing samples of the events which should be predicted. In image recognition, the training set contains images of events the user wants to detect.

Parameters can be taught using gradient methods. These gradient methods must converge to a local optima which is why batch methods are often used for convolutional neural networks. We used the stochastic gradient descent (SGD) method to learn our parameters. The update step of the SGD method for a pair $(x^{(i)}, y^{(i)})$ is

$$\theta = \theta - \alpha \nabla_\theta J(\theta; x^{(i)}, y^{(i)}) \tag{10.5}$$

where α is the learning rate, x is the sample, and y is the label assigned to x. The learning rate α is chosen to be small prior to computation, and is updated after each iteration. The stochastic gradient descent method overcomes the slow calculation common to batch methods (Ng et al., 2013). It learns parameters by computing the gradient of the cost function $J(\theta)$ over a few batches of training images as opposed to the entire training set to update the parameters after each iteration.

For the SGD method, it is also necessary to pick a momentum. The momentum helps push the SGD objective down any shallow ravines present in the calculation in order to improve performance of the method. As with the learning rate, the momentum is updated after each iteration. The initial choice of the momentum is around 0.5 before being chosen between 0.9 and 1. The SGD outputs learned parameters after a user-determined number of loops through the training set. The user then employs the learned parameters in the neural network to classify events.

For a detailed tutorial on neural networks and other deep learning methods, see Ng et al. (2013).

10.2 Applications

We now apply the methods discussed earlier in the chapter to build two different convolutional neural networks. We develop our CNNs using supplementary code from Ng et al. (2013) and Schmidt (2005). We start by training a CNN to distinguish between when an image is a hyperbola and when it is not, given that seismic events are often hyperbolic in shape (as we will see with the microseismic data later) (Gulati et al., 1997). We then train a CNN to identify specific seismic events.

10.2.1 Convolutional neural networks with two classes

Our first convolutional neural network learned to identify whether an image was a hyperbola or not a hyperbola in order to identify events in microseismic data. In the following sections, we analyze the structure of the CNN we built. We consider the training set we created to train the convolutional neural network to detect hyperbolas and test the trained CNN on a test set of images. Finally, we apply the trained CNN to detect hyperbolas in microseismic data.

Convolution neural network architecture

After some experimentation, we chose the following architecture for this example. The convolutional neural network has three layers, not including the input layer. The first layer is a convolutional layer which utilizes 20 filters of dimension 91×91 matrix entries. The second layer is a pooling layer that applied a mean-pooling. The outputs propagate into a softmax regression with cross entropy objective. The dimension of the pooling features was 2×2. The final layer is the output layer which contains two units determining one of two cases: whether there was a hyperbola present or whether there was not a hyperbola present in the image.

We employ a stochastic gradient descent method using momentum in order to optimize the parameters of a given objective. This function involves using the minFunc function created by Mark Schmidt (Schmidt, 2005). The SGD method calculates the learned parameters to train the convolutional neural network to detect hyperbolas in data sets based on the cost function $J(\theta)$ and allows for a faster convergence. We set the momentum at 0.99. To train the neural network, we use the learning rate $\alpha = 0.01$. The SGD subsamples over batches containing 32 training samples. The training ran through the training set three times in order to educate the convolutional neural network.

Training set

Training sets are an important component of neural networks. In image recognition, these sets comprise of various images which are used to train the neural network to identify specific objects. We taught the convolutional neural network to identify hyperbolas given that geophysical anomalies in microseismic often take the shape of a hyperbola.

We created a training set by extracting images from real data. The real data training set was segemented from the data in Figure 10.1. This image is of someone walking parallel to a length of fibre-optic cable and was acquired using distributed acoustic sensing. Upon close examination of each step individually, one finds a hyperbola. We remove each of these hyperbolas to create a training set. This training set contains 528 images of size 128×128. The set contains 371 hyperbolas and 157 non-hyperbolas. We define an image as a non-hyperbola if it did not contain a complete hyperbola in the window. The top row of Figure 10.2 shows an example of an image in the real data training set which is considered a hyperbola while the bottom row depicts examples of images which are considered non-hyperbolae.

Test set

To help determine the accuracy of our trained convolutional neural network, we train the CNN on 400 images in the training set and test it on the remaining 128 images afterwards. Figure 10.3 displays the probabilities that an image is a hyperbola or not a hyperbola for each image in the test set. As the reader can see, the CNN remains confident for determining hyperbolae giving a probability of approximately 1 for many of the images. For about 13 images, the CNN decides they were not hyperbolas. Comparing the predictions of the CNN to the manually applied labels, we get an accuracy of approximately 85%.

Figure 10.4 (left) gives an example of a image the CNN determined was a hyperbola whereas Figure 10.4 (right) gives an example of an image the CNN predicted was not a hyperbola. In both cases, the CNN predicted the image accurately.

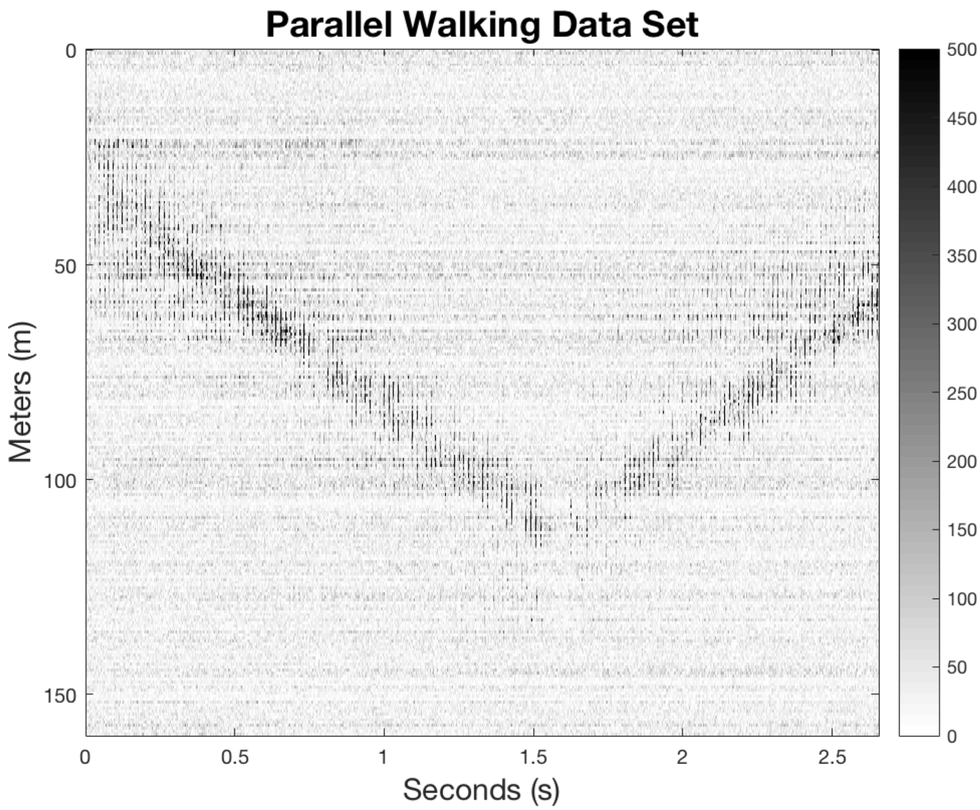


Figure 10.1: An image of someone walking parallel to a fibre-optic cable acquired using distributed acoustic sensing. The parallel walking data set is employed to create the training set for the convolutional neural network.

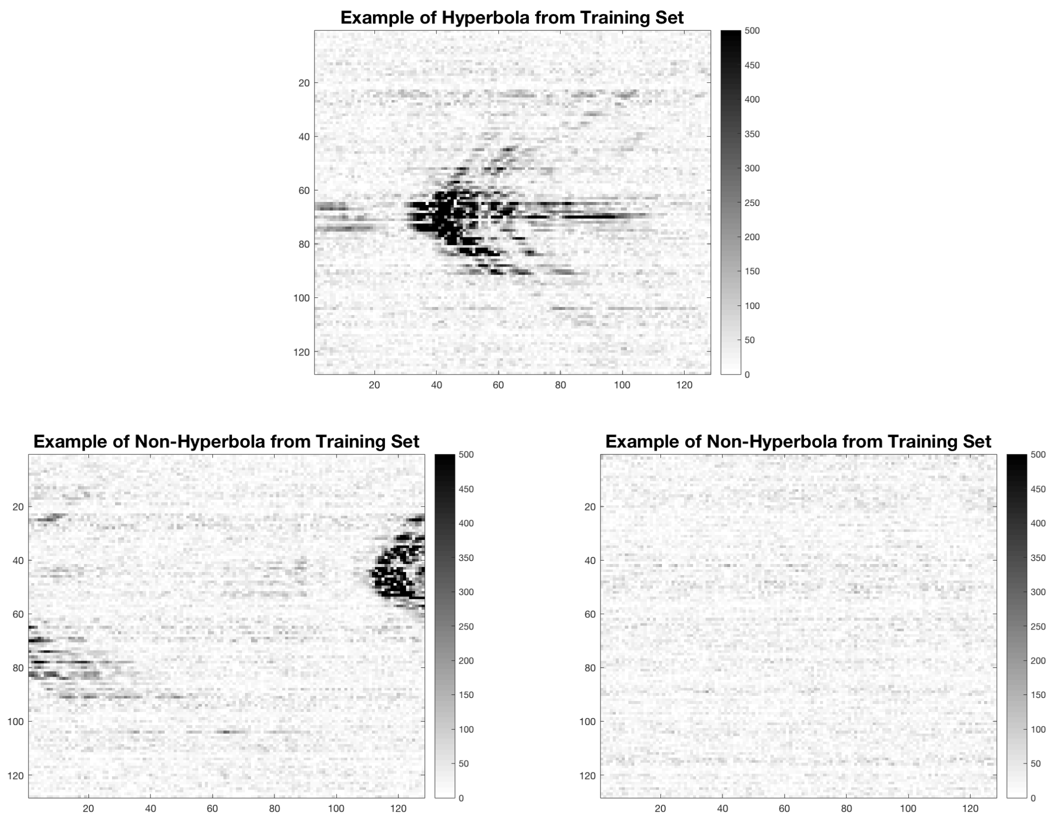


Figure 10.2: (Top) An image of a hyperbola from the training set. (Bottom) Examples of non-hyperbolae from the training set.

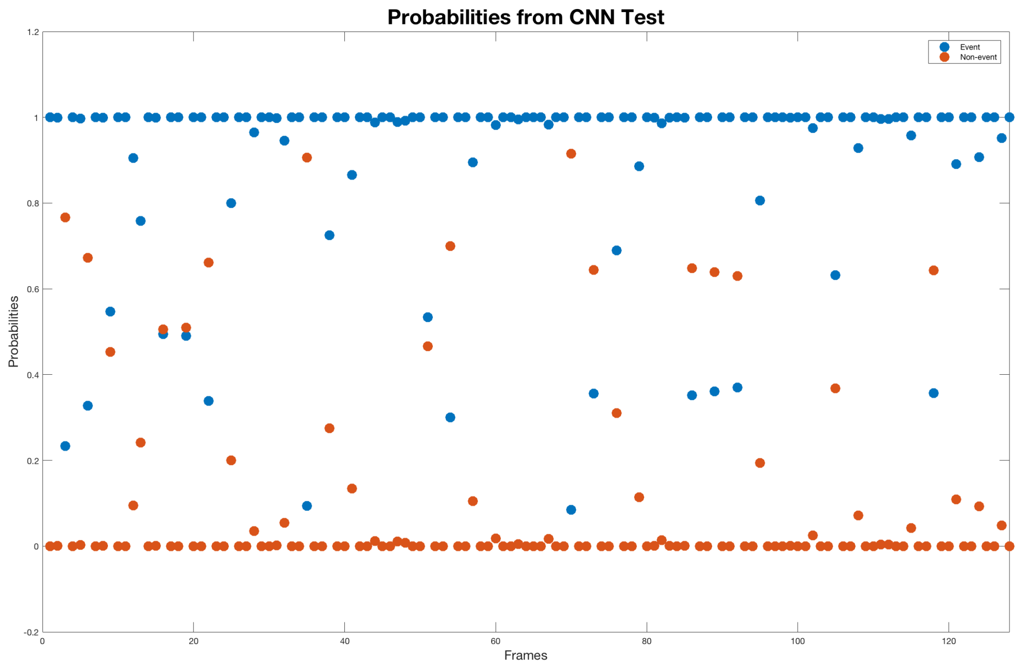


Figure 10.3: The probability that each image in the remaining 128 images of the training set is a hyperbola (described by a blue dot) or a non-hyperbola (described by an red dot).

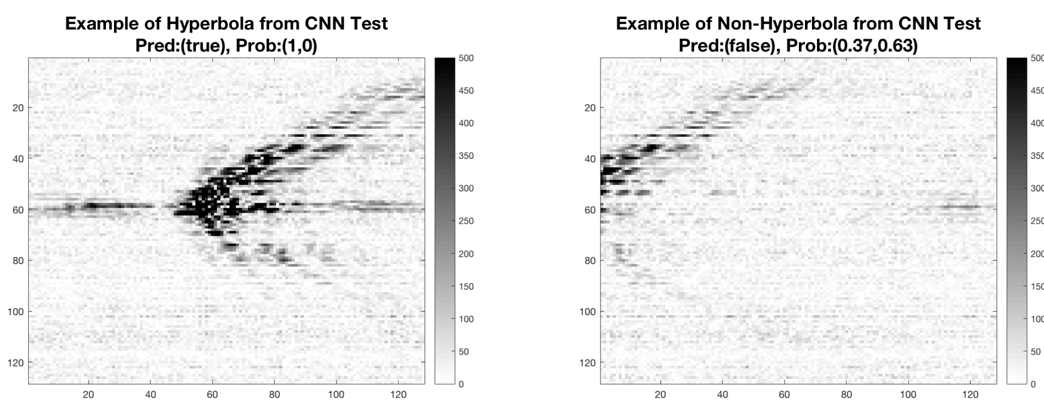


Figure 10.4: (Left) An example of a hyperbola the CNN predicted from the test set images. (Right) An example of a non-hyperbola the CNN predicted from the test set images.

Testing on outside sample

Now, we apply the convolutional neural network that we have trained and tested previously to two microseismic data sets acquired by Fotech Solutions using distributed acoustic sensing and fibre optics. Figures 10.5 and 10.6 show the microseismic data sets to which we apply the CNN from this section.

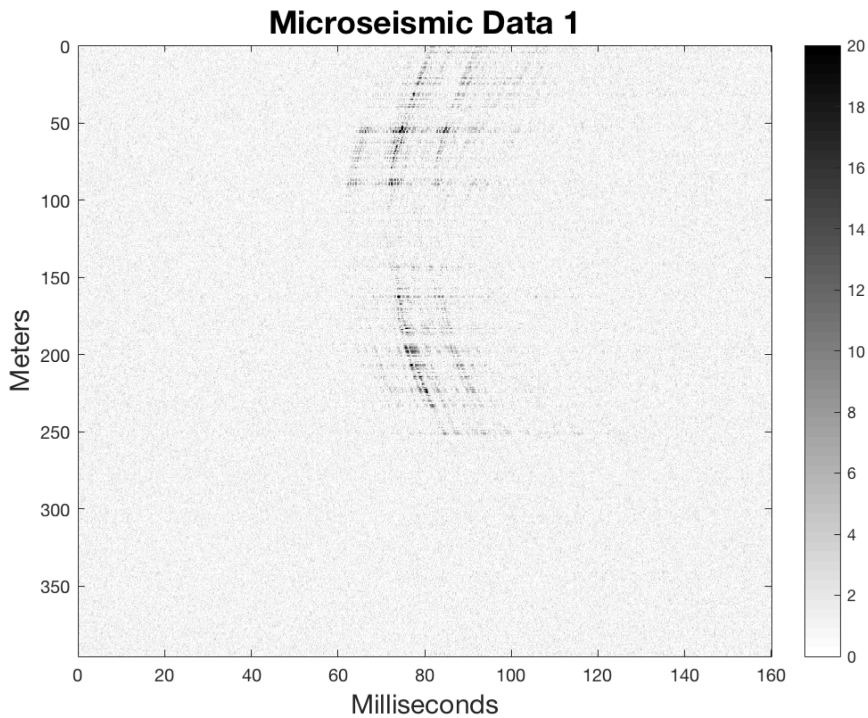


Figure 10.5: The first microseismic data set to which we apply the CNN.

We begin by considering the microseismic data set from Figure 10.5. We now need to implement the inverted wavelet tree in order to apply the CNN. The microseismic data set is a whole set as opposed to the test case in the previous section where we had a number of segemented images the size of the training images. We need to window the CNN through the microseismic data. Furthermore, given the size of the data set, it is necessary to downsample the data set in order to analyze it using the CNN.

Figure 10.7 provides an image of the downsampled version of the data pictured in Figure 10.5. In this case, we downsampled by 75%.

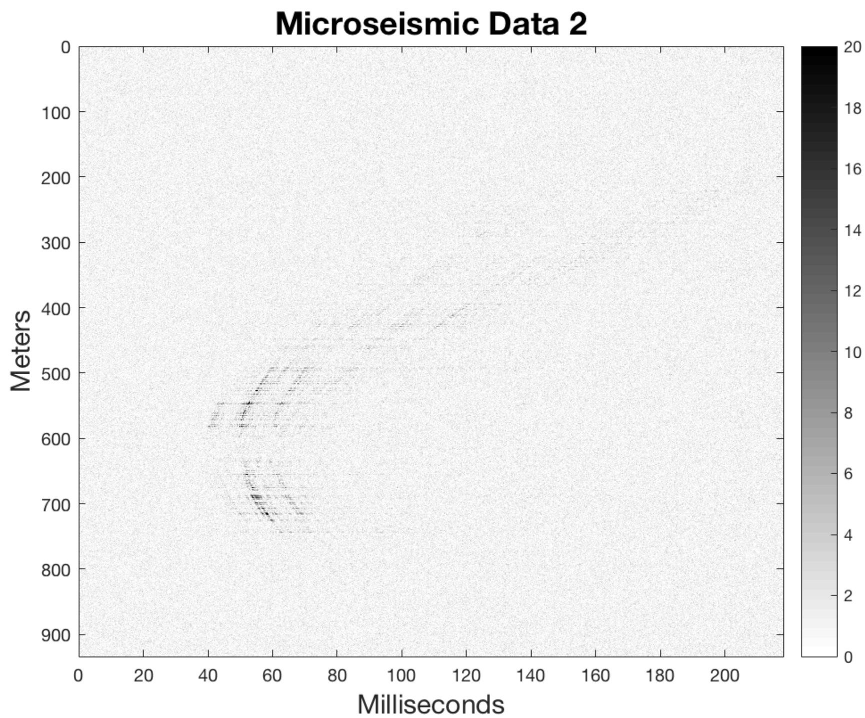


Figure 10.6: The second microseismic data set to which we apply the CNN.

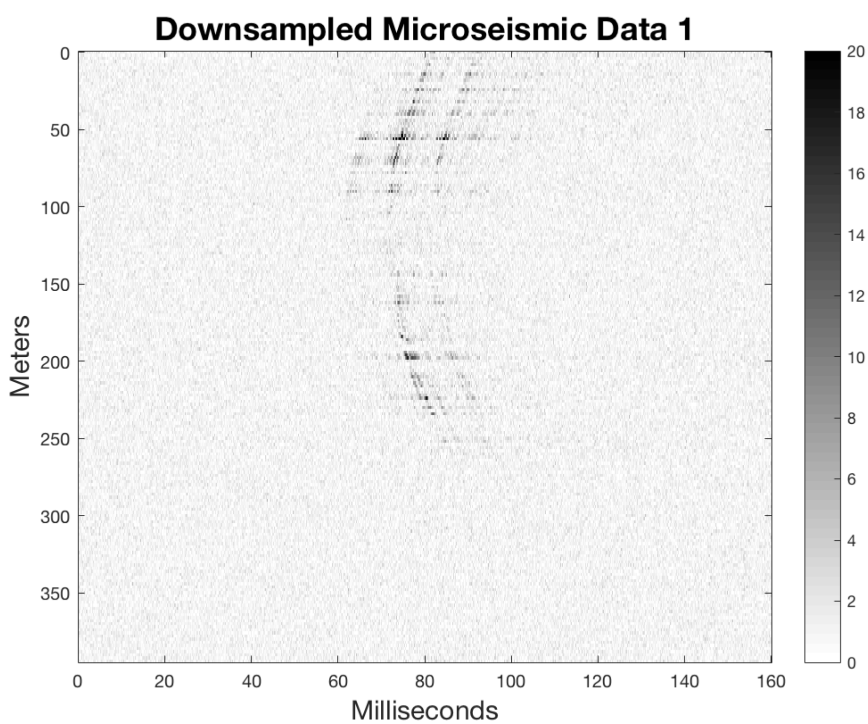


Figure 10.7: The downsampled version of the microseismic data from Figure 10.5.

Given the size of the downsampled data, we built a three-level inverted wavelet tree which applies the Haar transform producing the approximation coefficients at the nodes in the inverted wavelet tree. Figure 10.8 gives the probabilities of whether or not a hyperbola is present in the frame at each level. The red dots represent the probability that a hyperbola is not present and the blue dots represent the probability that a hyperbola is present. There are only predictions for hyperbolae in Level 1 and Level 2, which gives some idea as to the size of the events in the data. The location of these predictions suggest that there are hyperbolae in the data correlating to Frames $40 - 55$ in Level 1 and Frames $20 - 27$ in Level 2. Moreover, the scale of the events is somewhere between the scale in Level 1 and the scale in Level 2. While Level 3 has no predictions for hyperbolae, some variation in prediction resides between Frames 9 and 16 which may suggest an event; however, the event does not fit the scale provided in Level 3.

Figure 10.9 gives examples of what our trained CNN determined was a hyperbola versus a non-hyperbola. As the reader can see in Figure 10.9, we have programmed the code to say 'true' if a hyperbola is present and 'false' if a hyperbola is not present.

We also applied the CNN to the second microseismic data set (Figure 10.6). As with the first data set, we need to utilize an inverted wavelet tree as well as downsample the data by 75% in order to apply the CNN; see Figure 10.10.

As with the previous data set, we built a wavelet tree with three levels and applied the Haar transform, taking the approximation coefficients for the nodes of the tree. Figure 10.11 shows the probabilities for identifying hyperbolas in the second downsampled microseismic data set at each level. None of the levels of the wavelet tree predict a hyperbola in the data set. There is significant activity in Level 1 between Frames 20 and 60. The main activity in Level 2 is between Frames 10 and 15; and in Level 3, the activity is isolated around Frame 6. This could suggest that there is some event in the corresponding locations of the data set; however, it is not the definitive answer that we need. Potentially, the lack of hyperbolae could mean that the events did not fit in any level for the inverted wavelet tree's scaling.

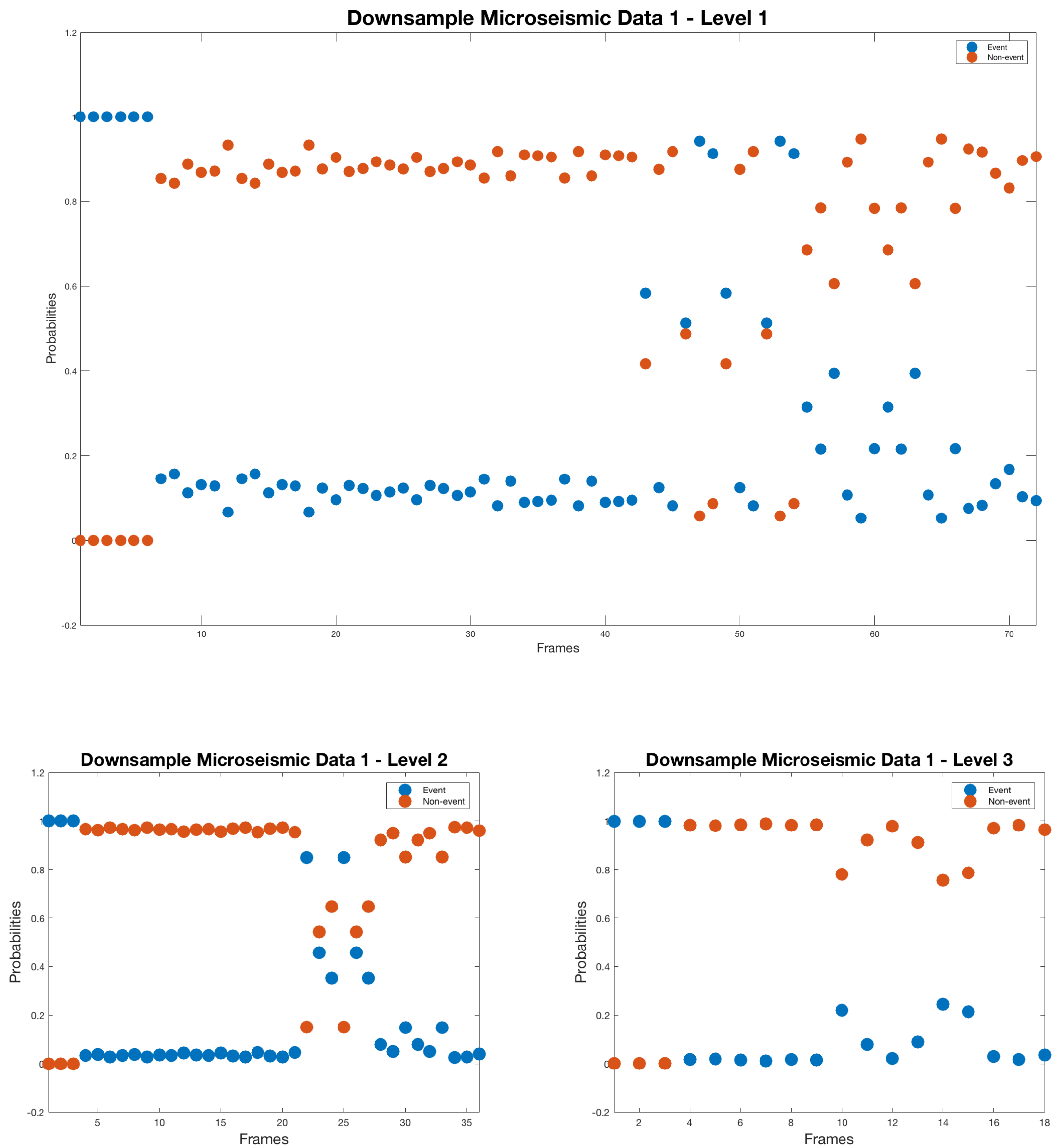


Figure 10.8: The probabilities in each frame of a hyperbola (blue) being present and a non-hyperbola (red) being present on (top) Level 1, (bottom left) Level 2, and (bottom right) Level 3.

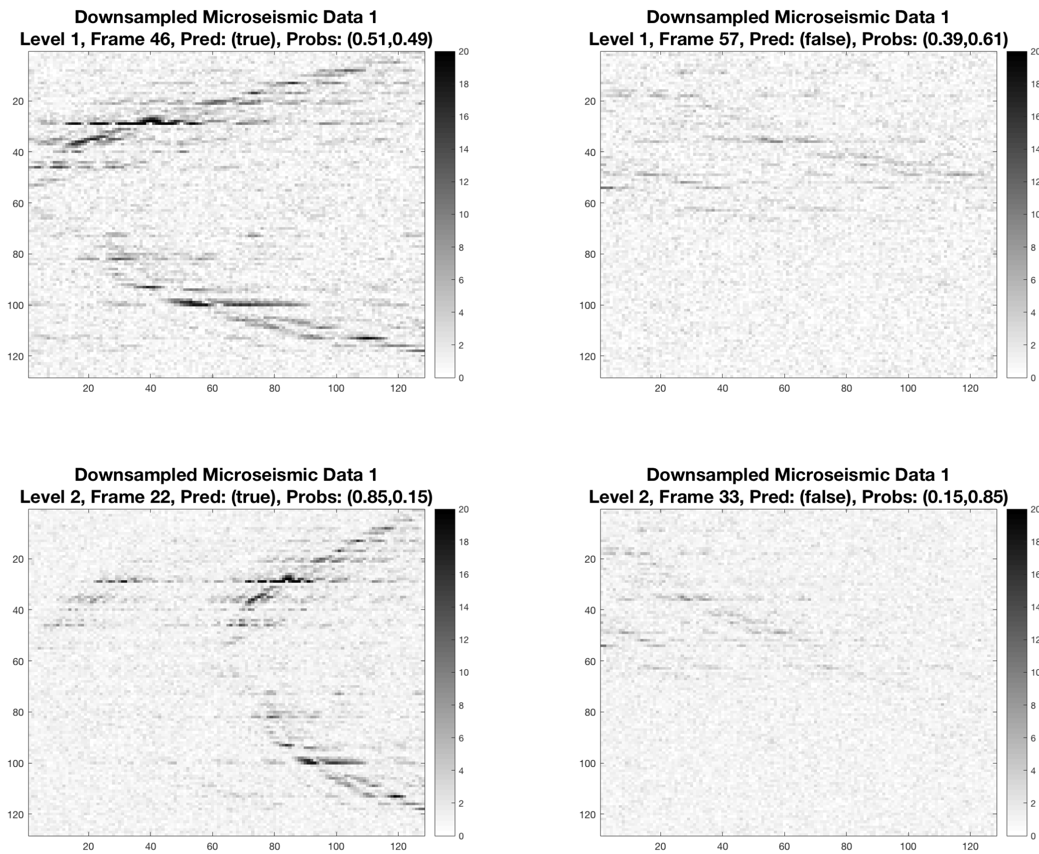


Figure 10.9: Images of hyperbolas and non-hyperbolas from the first downsampled microseismic data set predicted by the trained CNN.
(Top left) An example of a hyperbola on Level 1 of the inverted wavelet tree.
(Top right) An example of a non-hyperbola on Level 1 of the inverted wavelet tree.
(Bottom left) An example of a hyperbola on Level 2 of the inverted wavelet tree.
(Bottom right) An example of a non-hyperbola on Level 2 of the inverted wavelet tree.

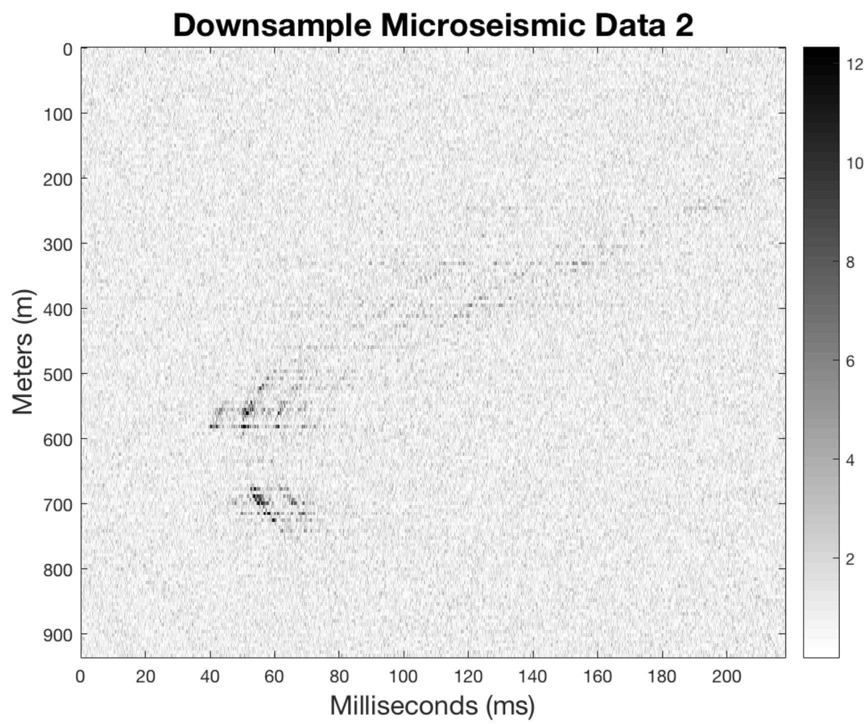


Figure 10.10: The downsampled version of the microseismic data from Figure 10.6.

It may be necessary to downsample the microseismic data by more than 75% in this case, unlike in the previous case. Upon visual examination of both data sets, it is possible to observe some difference in the degree of curvature of the hyperbolic events in both data sets. Perhaps less downsampling is necessary for the second data set given the degree of curvature for the hyperbolic events in this case is larger than the ones in the first data set. It merits further investigation.

10.2.2 Convolutional neural network with three classes

We now train a convolutional neural network to distinguish between three different events: someone walking next to the fibre, someone digging next to the fibre, and noise. Distributed acoustic sensing describes both digging and walking as hyperbolic events. To differentiate between these events, we train on images which contain multiple events instead of just one event (as was the case in the previous example). We hypothesize that the distance between

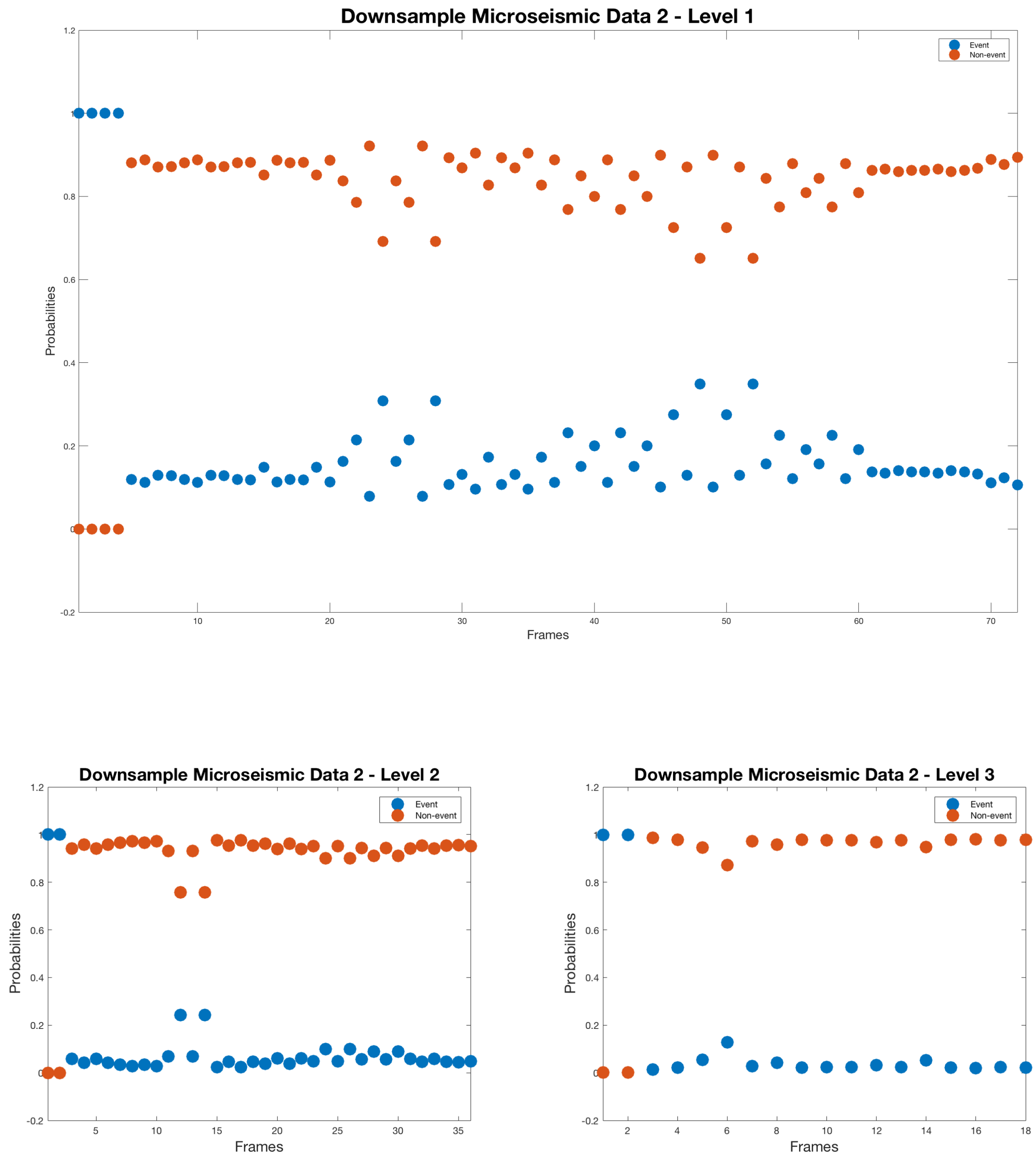


Figure 10.11: Probabilities of hyperbolae and non-hyperbolae from applying the trained CNN to the second downsampled microseismic data set. The probabilities in each frame of a hyperbola (blue) being present and a non-hyperbola (red) being present on (top) Level 1, (bottom left) Level 2, and (bottom right) Level 3.

the steps and the digging will help differentiate between the two cases.

We take this experiment a step further than the last example in that we will apply the neural network to varying downsampled versions of the data set and compare the results. Specifically, we run the neural network on the data first without any downsampling, then when downsampled by 10 shots, and finally when downsampled by 50 shots.

Architecture of the convolutional neural network

The key feature we changed from the previous neural network's architecture is the number of classes it needs to identify. Previously, we had two classes: a hyperbola and a non-hyperbola. For this experiment, we have three classes: walking, digging, and noise. Given that we include multiple hyperbolic events to help distinguish between walking and digging, it is necessary to increase the filter dimension to 101×101. Otherwise, we kept the same parameters. Leaving as much similarity as possible allows us to judge how well convolutional neural networks perform when answering the equivalent of a yes-no question about an event versus providing a specific name for an event.

Training set

As with the previous CNN, we extract images from real data to create a training set. Instead of working with only one data set, we employ 207 data sets: 96 contain walking events and 111 contain digging events. For noise events, we remove images from the data sets which did not overlap with walking or digging events in the data. Given that noise is more prevalent in data, we limit the number of noise events to a similar amount of digging and walking events in order to avoid giving the neural network a bias toward identifying noise.

Figures 10.12 and 10.13 provides an example of data sets containing walking and digging respectively. Within both of these data sets, we see examples of noise. The walking is located between 50 and 175 meters and 2000 to 11000 milliseconds in Figure 10.12. The digging events are located between 50 and 100 meters and 3000 and 10000 milliseconds in

Figure 10.13.

The training set contains 3232 images of size 128×128. The set has 1002 walking events, 1142 digging events, and 1086 noise events. The top row of Figure 10.14 exhibits an example of a walking event on the left and a digging event on the right. The bottom row of Figure 10.14 provides two examples of a noise event. Observe from the bottom row of Figure 10.14 that noise events vary drastically. This may prove to cause the CNN some difficulty in distinguishing between the three different classes.

Figure 10.12: An image of someone walking parallel to a fibre-optic cable acquired using distributed acoustic sensing. The footsteps are located between 50 and 175 meters and 3000 and 11000 milliseconds. The walking data set is employed to create the walking events in the training set for the convolutional neural network.

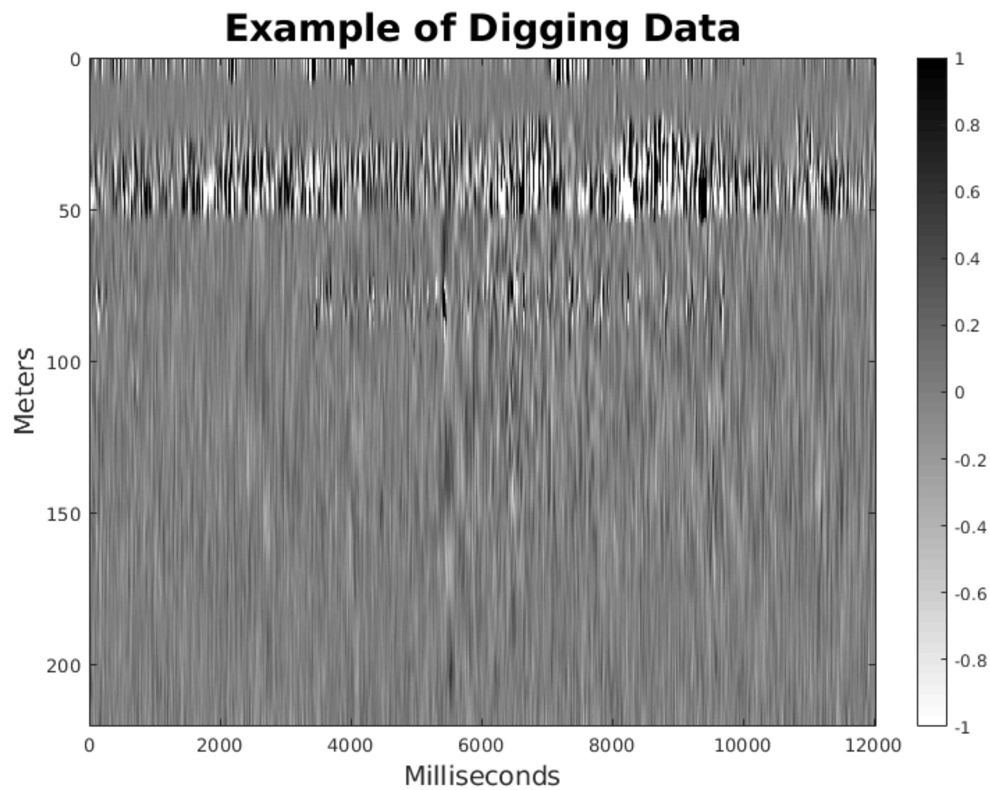


Figure 10.13: An image of someone digging next to a fibre-optic cable acquired using distributed acoustic sensing. The digging events are located between 50 and 100 meters and 3000 and 10000 milliseconds. The digging data set is one of several data sets employed to create the digging events in the training set for the convolutional neural network.

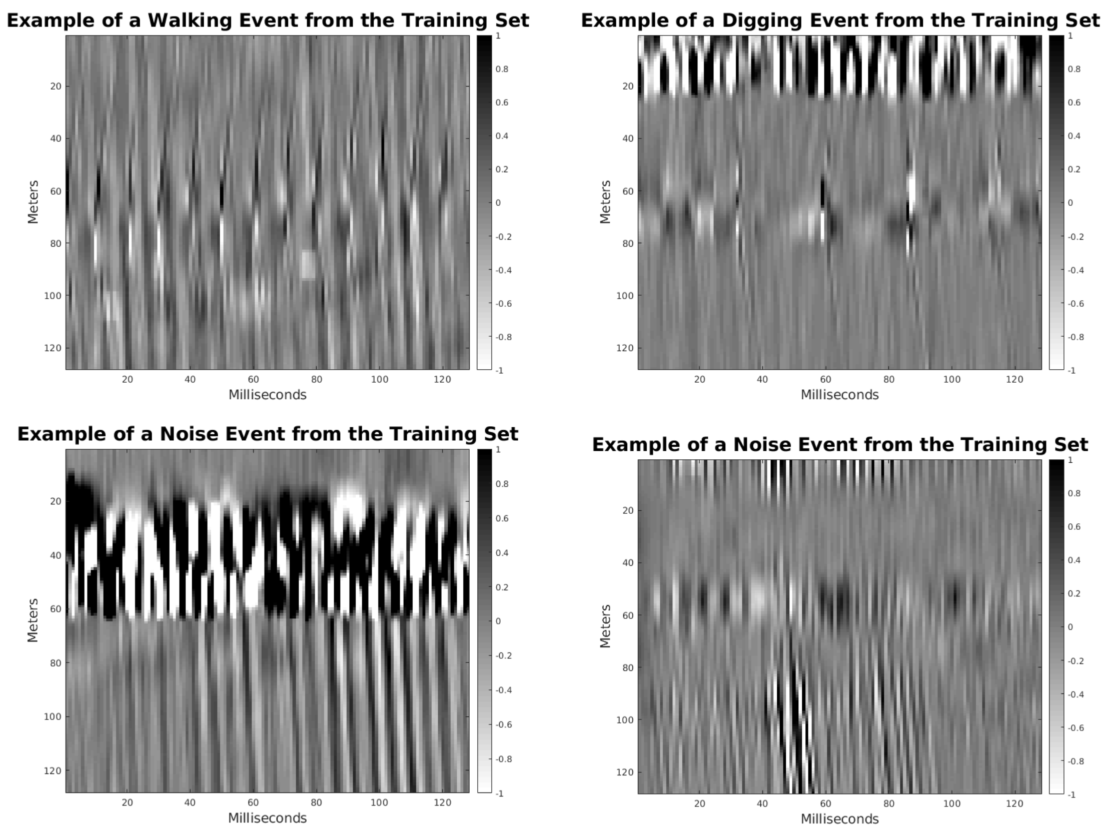


Figure 10.14: (Top Left) An image of a walking event from the training set.
(Top Right) An image of a digging event from the training set.
(Bottom) Examples of noise events from the training set.

Test set

We train the CNN on 3232 images of the training set and test the results on a testing set of 2155 images. In Figure 10.15, we show the probabilities that an image is someone walking, someone digging, or noise in the data. Figure 10.16 provides some examples of what the CNN decided represented a walking event, a digging event, and a noise event.

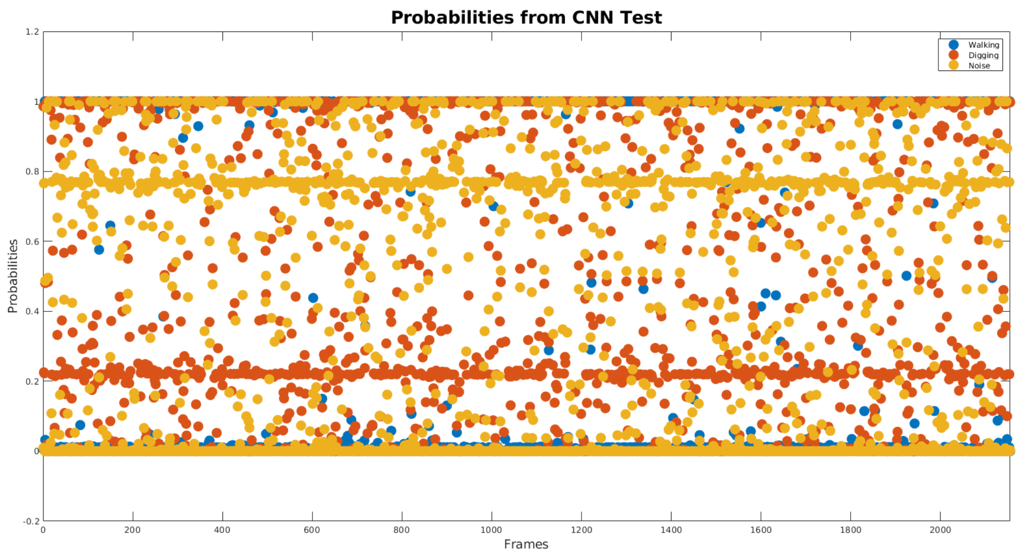


Figure 10.15: The probability that each image in the 2155 images of the test set is a walking event (blue), digging event (red), or a noise event (yellow).

In Figure 10.15, the prevalence of red and yellow dots strictly greater than 0 and strictly less than 1 suggests that the neural network struggled to distinguish between digging and noise. The CNN did not have confidence in what was noise versus what is digging. Fewer blue dots representing walking events strictly greater than 0 and strictly less than 1 suggests that the CNN possessed more confidence in what was a walking event.

The CNN has an accuracy of approximately 80% which is a little more than 5% less than the neural network from the previous experiment. The decrease in accuracy may be attributed to the fact that noise is the third class given that noise often varies in appearance. The similarities between walking and digging may also contributed to the accuracy's decline.

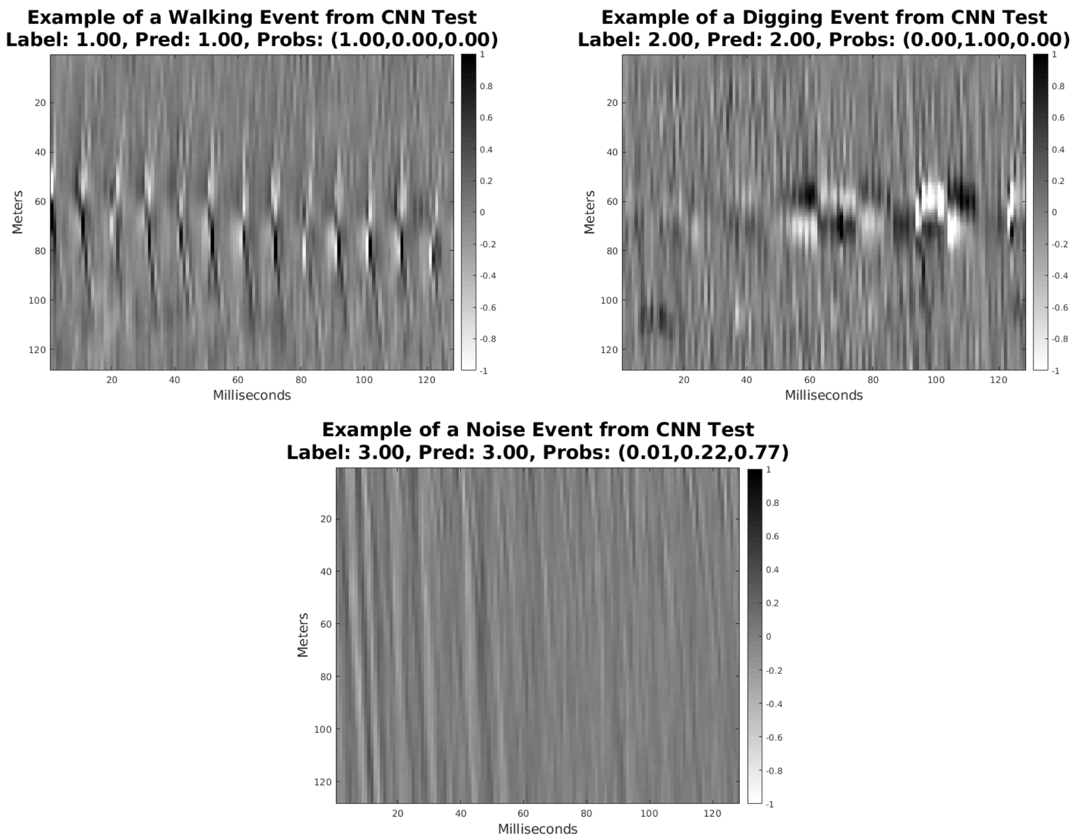


Figure 10.16: (Top left) An example of a walking event the CNN predicted from the test set images.
(Top right) An example of a digging event the CNN predicted from the test set images.
(Bottom) An example of a noise event the CNN predicted from the test set images.

Notice that the CNN possessed more than 5 times the images with which to train than the CNN from the previous example. Intuitively, it seems as if more training images would be helpful for teaching a CNN; however, we see in this experiment that that more training images does not always help increase the accuracy of a training set.

While a significant decline occurred in accuracy between the CNNs in the two experiments, 80% accuracy is still promising for applications to other data. We explore the success of applying the CNN to two data sets which contain walking and digging events in the following section.

Test on outside samples

To create outside examples, we glued a data set containing walking and a data set containing digging together. We did this for two data sets acquired at different pulse repetition frequencies - 4kHz and 6kHz. Recall from Chapter 5 that the pulse repetition frequency is the rate at which a laser pulse can be sent down the fibre and all the backscattered light exits before we launch the next pulse. Given how large DAS data can be, we also apply the CNN to three different decimations of the two data sets — no decimation, decimation by 10 shots, and decimation by 50 shots. We created the training set from data decimated by 50 shots in order to keep the image dimension of the training images 128×128. This fact suggests that the CNN will do the best on the data decimated by 50 shots for both data sets.

Figure 10.17 displays the walking and digging data set with no decimation acquired at a pulse repetition frequency of 4kHz. The walking events lie between 50 and 175 meters and 0.25×10^4 and 1.2×10^4 milliseconds. The digging events are located between 50 and 100 meters and 1.5×10^4 and 2×10^4 milliseconds.

Given how the inverted wavelet tree creates levels after the data in previous levels, it is difficult to calculate the accuracy of the CNN on real data at these levels; however, we can judge the accuracy of the first level. As such, we only consider the probabilities of

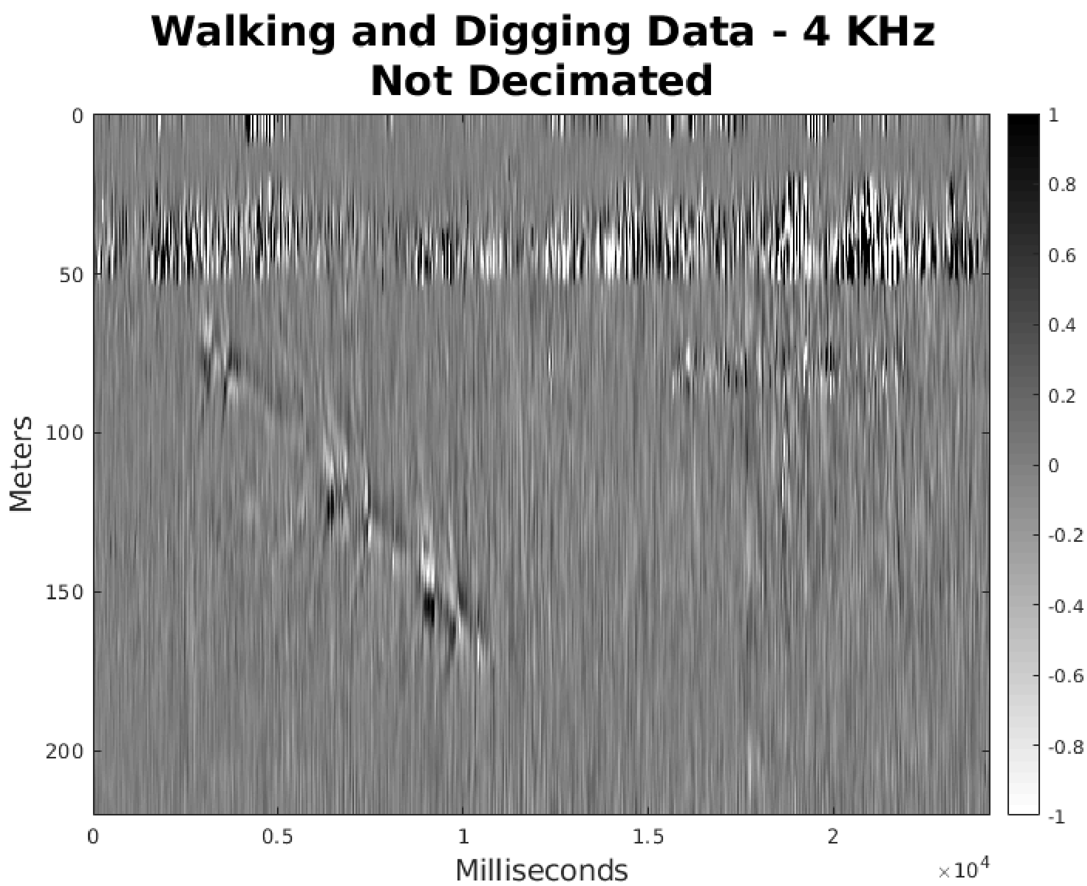


Figure 10.17: An image of someone walking parallel to the fibre-optic cable and someone digging next to the fibre. The data was acquired at a pulse repetition rate of 4kHz. No decimation has been applied to the data.

walking, digging, and noise events in the first level of the wavelet tree. Figure 10.18 shows the probabilities a frame of the data is walking, digging, or noise. Because the data is not decimated, it has more frames. Furthermore, the walking and digging events were localized within the data so many frames will contain noise events which explains why the orange dots are more prevalent in Figure 10.18.

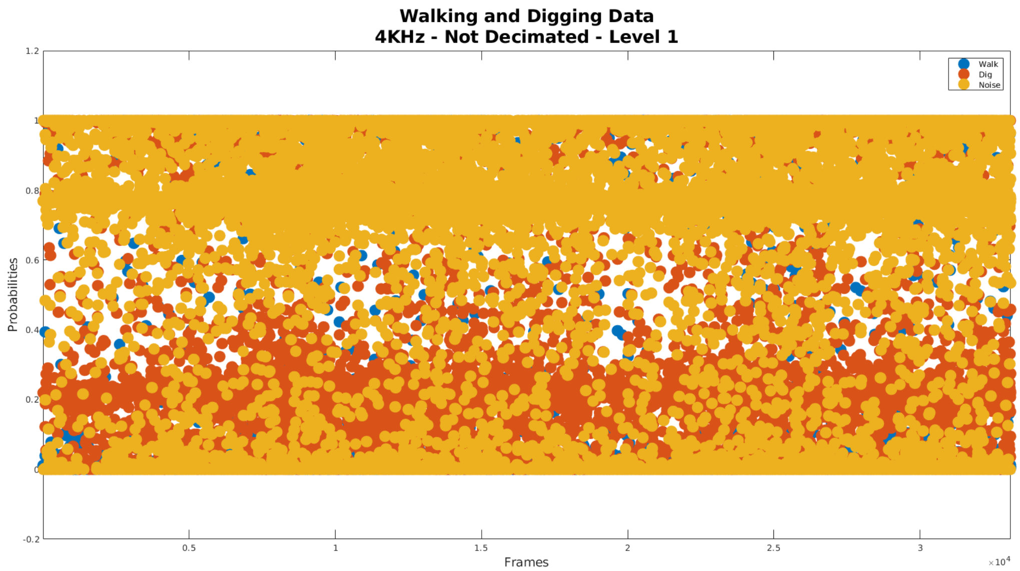


Figure 10.18: Probabilities of walking, digging and noise events from applying the trained CNN to the 4kHz PRF walking and digging data set which is not decimated. The probabilities in each frame of a walking event (blue), digging event (red), and noise event (yellow) being present on Level 1.

Let us consider decimating the data by 10 shots; see Figure 10.19. Visually, it is difficult to distinguish any differences in the two data sets from Figures 10.17 and 10.19. Given that we removed every 10th shot from the data set to create the image seen in Figure 10.19, there will be fewer frames for the CNN to identify as walking, digging, or noise. We see this reflected in the significant decrease in the density of the probabilities in Figure 10.20.

While the number of frames drops between the zero decimation and the decimated by 10 shots cases, the probabilities for digging and noise show prominence in the area strictly greater than 0 and strictly less than 1; however, we witness less confidence with regards

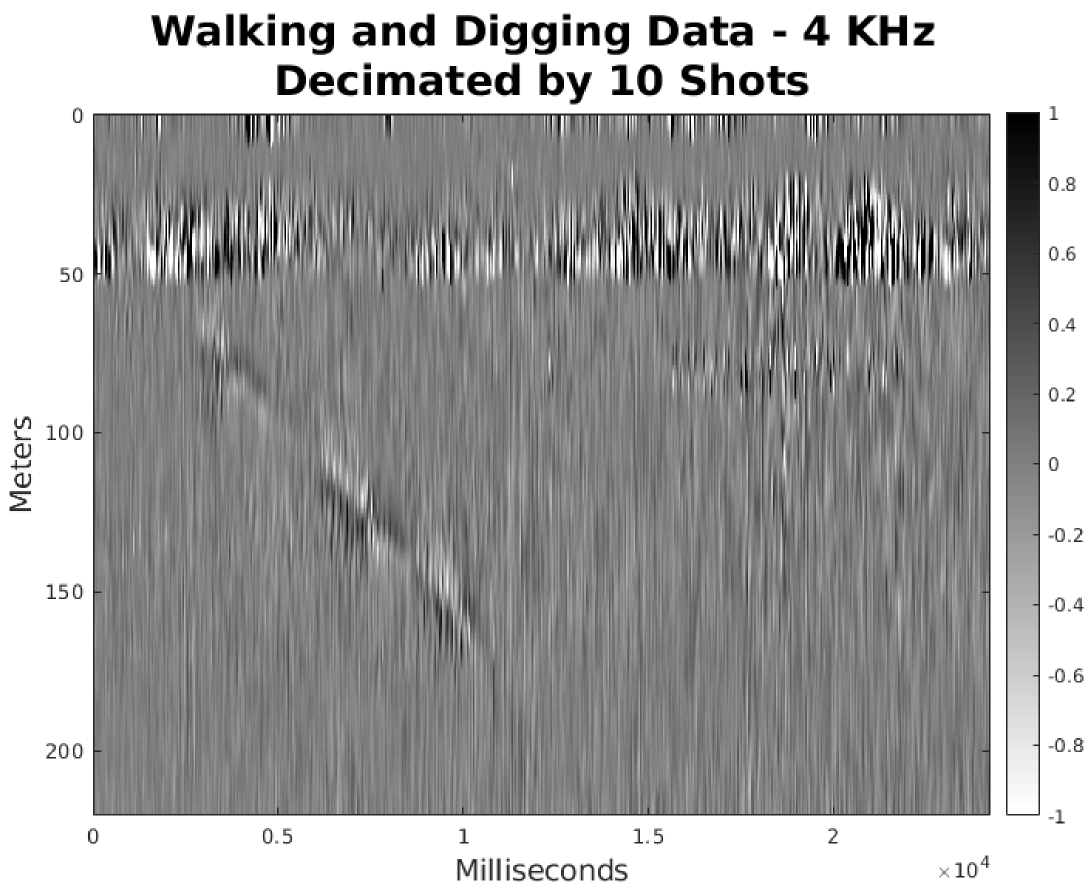


Figure 10.19: The 4kHz PRF walking and digging data set decimated by 10 shots. The walking events are located between 50 and 175 meters and 0.25×10^4 and 1.1×10^4 milliseconds. The digging events reside between 50 and 100 meters and 1.5×10^4 and 2.25×10^4 milliseconds.

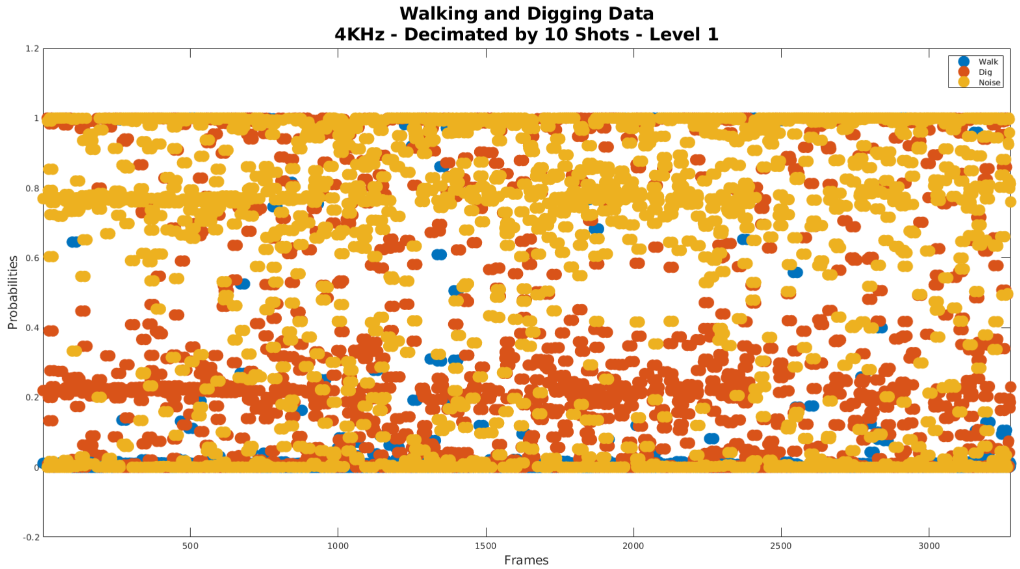


Figure 10.20: Probabilities of walking, digging, and noise events from applying the trained CNN to the 4kHz PRF walking and digging data set, decimated by 10 shots. The probabilities in each frame of a walking event (blue), digging event (red), and noise event (yellow) being present on Level 1.

to some walking events (blue dots). We should note that in both cases, a large cluster of noise probabilities landed around 0.8 and a large cluster of digging probabilities resided at approximately 0.2, which suggests that the CNN had some difficulty distinguishing between noise and digging.

Now, we examine the case when we decimate the 4kHz PRF walking and digging data by 50 shots; see Figure 10.21. Here the walking events are located between 50 and 175 meters and 0.25×10^4 and 1.1×10^4 milliseconds. The digging events reside between 50 and 100 meters and 1.5×10^4 and 2.25×10^4 milliseconds.

As with the 10 shot case, there are fewer frames. Figure 10.22 depicts the probabilities for the walking, digging, and noise. The fewer frames allows us to decipher more details from the probabilities. Recall from the image of the data that the walking occurs at the beginning of the data set and the digging occurs towards the end. Notice in Figure 10.22 that more frames identify as walking in the first 300 frames vs the last 300 frames. Similarly,

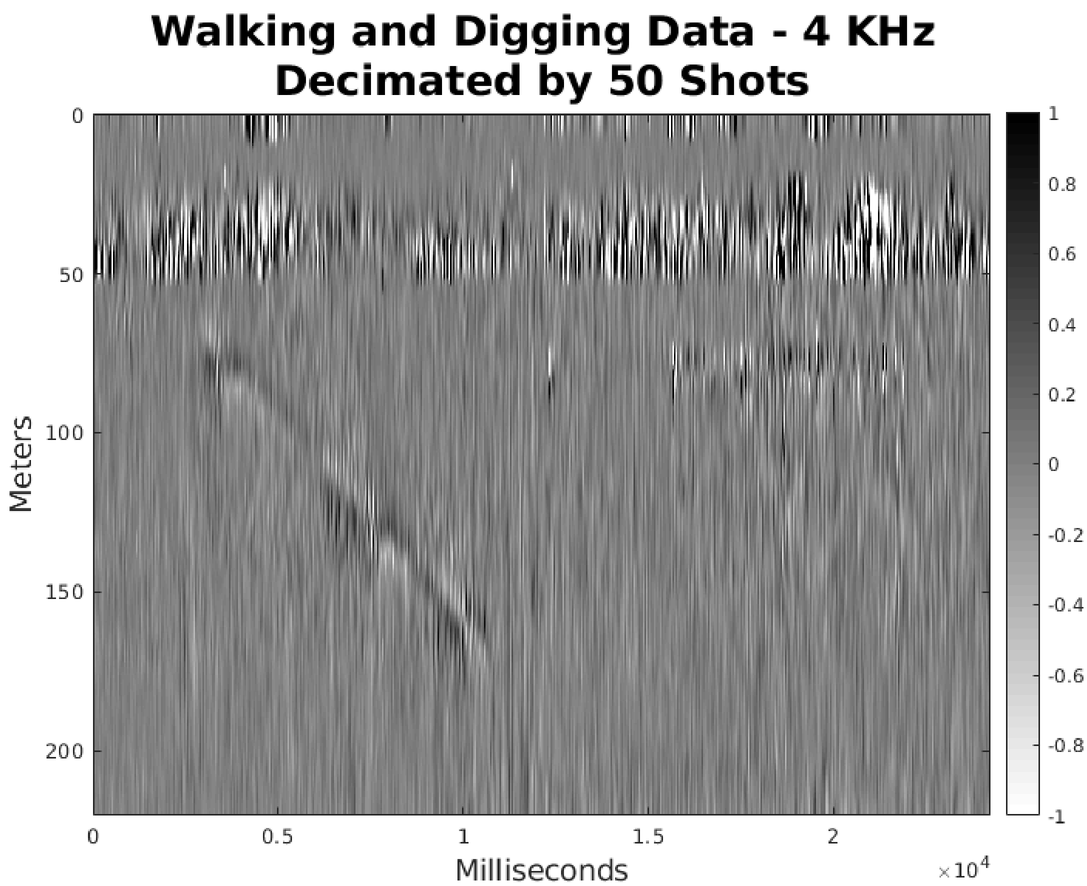


Figure 10.21: The 4kHz PRF walking and digging data set decimated by 50 shots. The walking events are located between 50 and 175 meters and 0.25×10^4 and 1.1×10^4 milliseconds. The digging events reside between 50 and 100 meters and 1.5×10^4 and 2.25×10^4 milliseconds.

more frames have a higher probability of being a digging event in the last 300 frames than in the first 300.

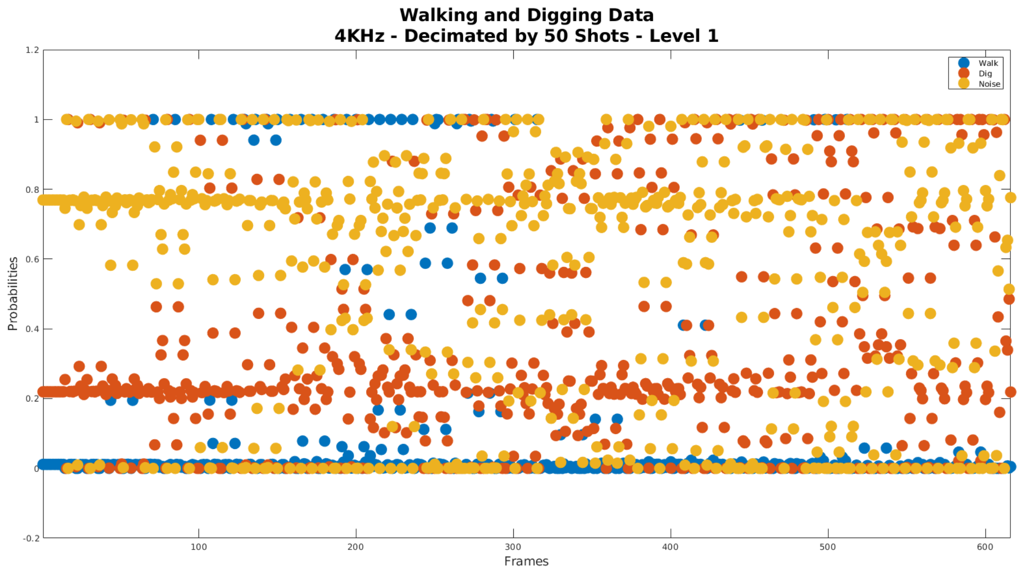


Figure 10.22: Probabilities of walking, digging, and noise events from applying the trained CNN to the 4kHz PRF walking and digging data set, decimated by 50 shots. The probabilities in each frame of a walking event (blue), digging event (red), and noise event (yellow) being present on Level 1.

Let us consider the data acquired using a pulse repetition frequency of 6kHz. Figure 10.23 presents the image of someone walking next to the fibre and then someone digging next to the fibre acquired at a pulse repetition frequency of 6kHz with zero decimation of the data. As with the 4kHz case, we consider the CNN results from the data which was not decimated, the data decimated by 10 shots, and then the data decimated by 50 shots. For all cases, we find the walking events between 25 and 150 meters and 0.25×10^4 and 1×10^4 milliseconds. The digging events are located between 25 and 75 meters and 1.5×10^4 and 2×10^4 milliseconds.

In Figure 10.24, we see the probabilities that an event is walking, digging, or noise for the data which was not decimated. The results are similar to the 4kHz PRF case; however, the density of probabilities in the last few frames is less spread out than in the 4kHz case.

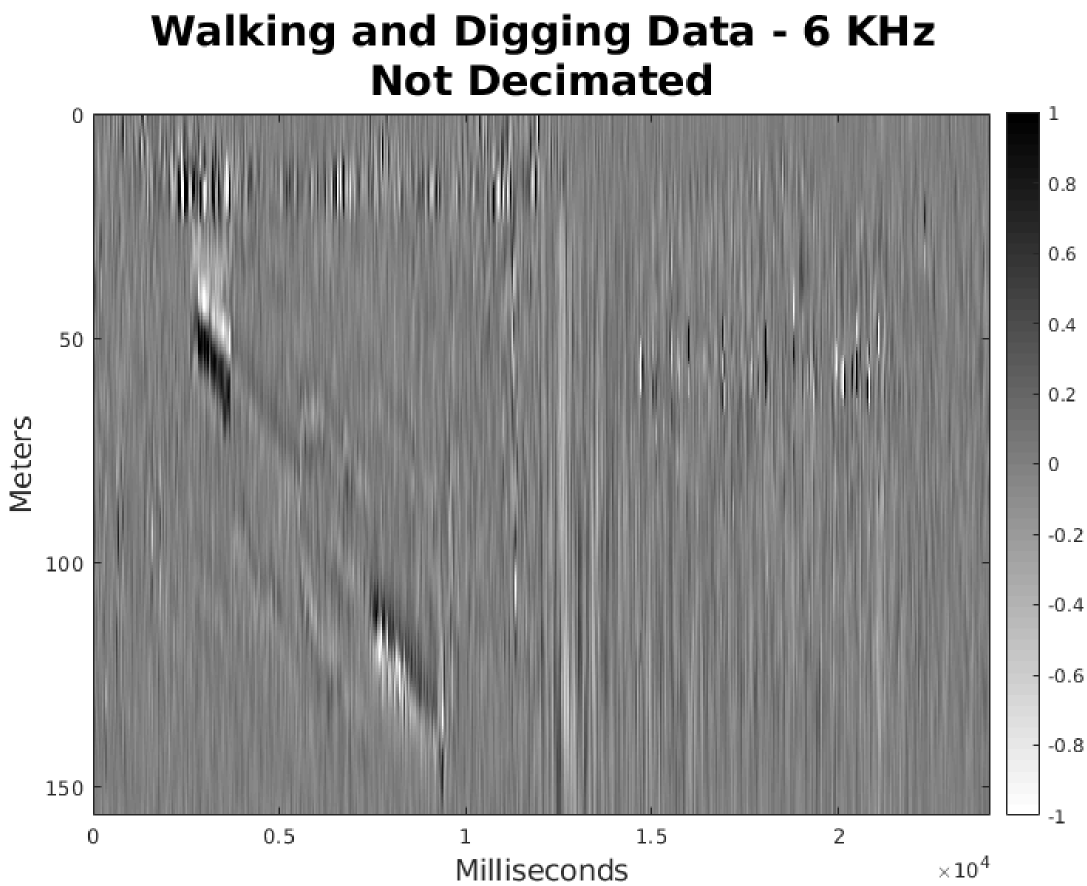


Figure 10.23: The 6kHz PRF data set which has not been decimated of someone walking next to the fibre (between 25 and 150 meters and 0.25×10^4 and 1×10^4 milliseconds) and someone digging next to the fibre (between 25 and 75 meters and 1.5×10^4 and 2×10^4 milliseconds).

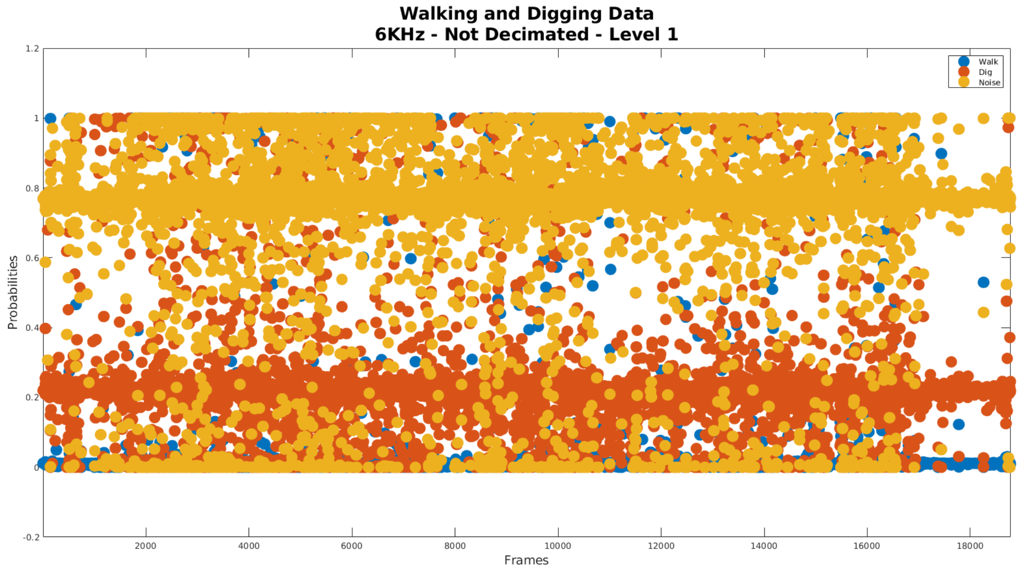


Figure 10.24: Probabilities of walking, digging, and noise events from applying the trained CNN to the 6kHz PRF walking and digging data set with no decimation. The probabilities in each frame of a walking event (blue), digging event (red), and noise event (yellow) being present on Level 1.

Figure 10.25 provides an image of the walking and digging data acquired at a PRF of 6kHz decimated by 10 shots.

As with the 4kHz, we see fewer frames as witnessed in the decrease in the dot density seen in Figure 10.26. Interestingly, some frames near the beginning of the data appear to be identified as digging events given the cluster of red dots near probability 1 in this region. Since digging events occur at the end of the data set, this suggests that the CNN struggled with identifying frames for this data set.

Finally, we consider the 6kHz PRF walking and digging data decimated by 50 shots as seen in Figure 10.27.

This case has significantly fewer frames. In Figure 10.28, we see a cluster of walking events in the first 300 frames which correlates to where the walking is in the data set. The CNN identifies a few frames near the beginning as digging events, signified by the red dots. Recall that the digging occurs near the end of the data set so these frames were mis-identified;

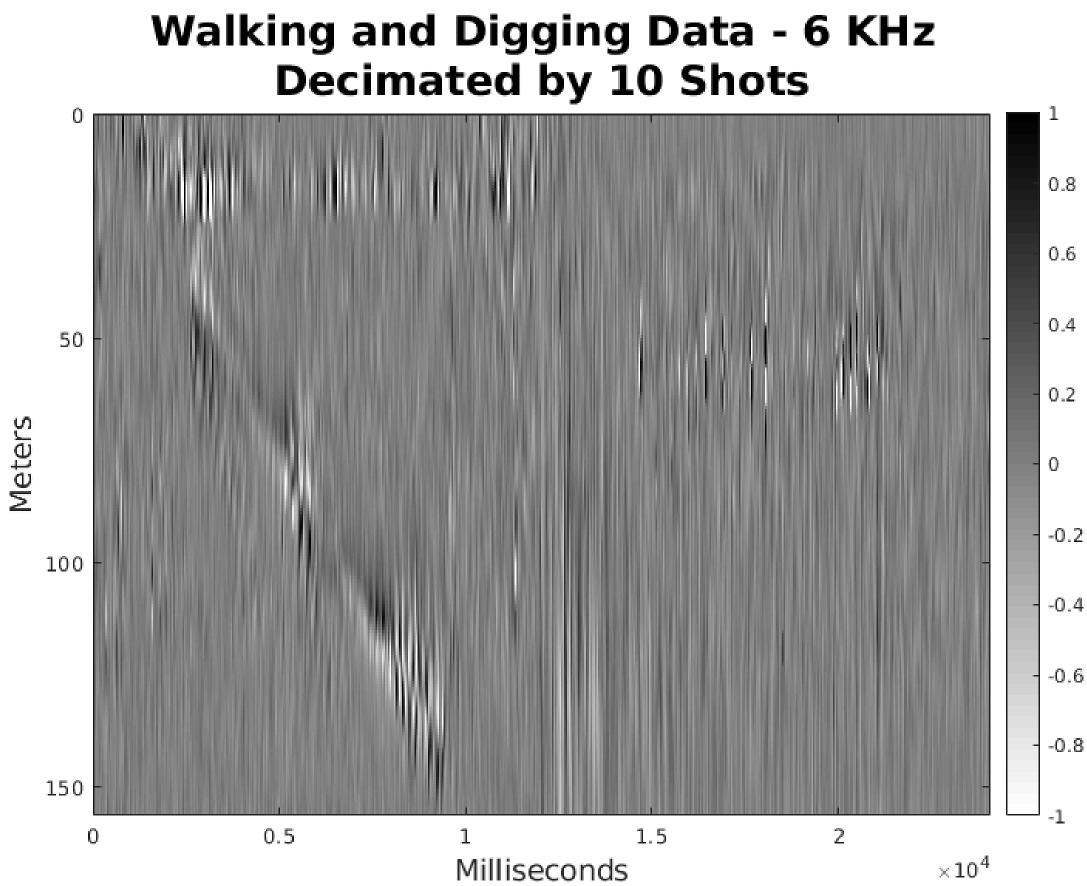


Figure 10.25: The 6kHz PRF data set, decimated by 10 shots, of someone walking next to the fibre and someone digging next to the fibre. The walking events reside between 25 and 150 meters and 0.25×10^4 and 1×10^4 milliseconds. The digging events are located between 25 and 75 meters and 1.5×10^4 and 2×10^4 milliseconds.

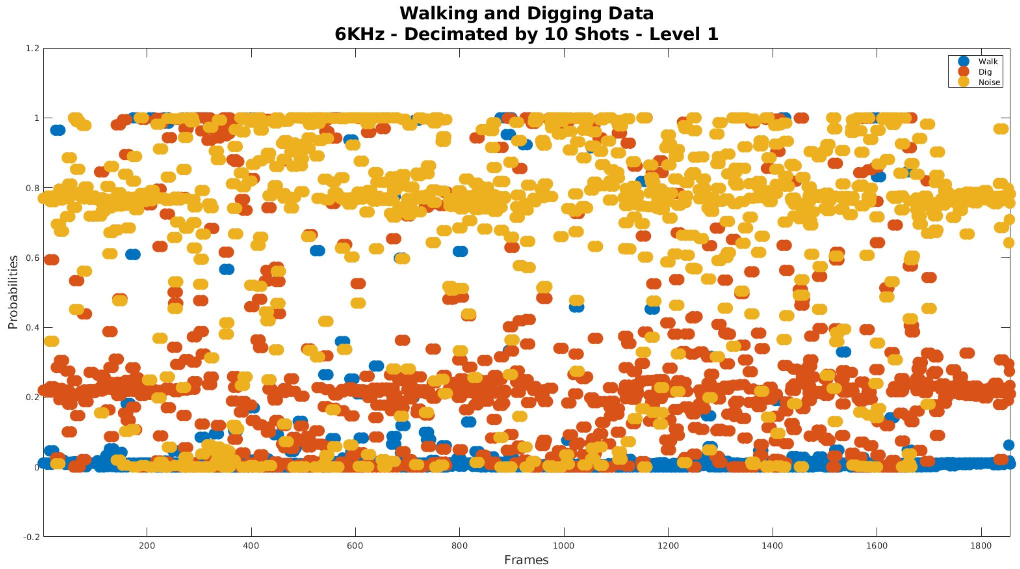


Figure 10.26: Probabilities of walking, digging and noise events from applying the trained CNN to the 6kHz PRF data set, decimated by 10 shots. The probabilities in each frame of a walking event (blue), digging event (red), and noise event (yellow) being present on Level 1.

however, a small cluster of digging events can be seen in the last few frames. Given the prominent cluster of walking events around probability 1 during the first 300 frames, the CNN likely performed adequately with regards to identifying walking; however, it does not appear to have done as well with regards to digging events.

Table 10.1 shows the accuracy of the CNN at identifying walking and digging events in each of the data sets. Specifically, we only looked at the results when a frame of data overlapped with known locations of walking or digging events.

For both PRFs, the CNN had the highest accuracy when we decimated the data by 50 shots. This result follows given that we built the training set using 50 shot decimated data. Interestingly, the CNN performed significantly better on the 4kHz data than the 6kHz data. Potentially, we employed more 4kHz data sets to build the training set than 6kHz data; given that the CNN's accuracy never achieves higher than 40% for the 6kHz case, this is likely the case. The accuracy for the 4kHz case dips drastically when the data is decimated by 10

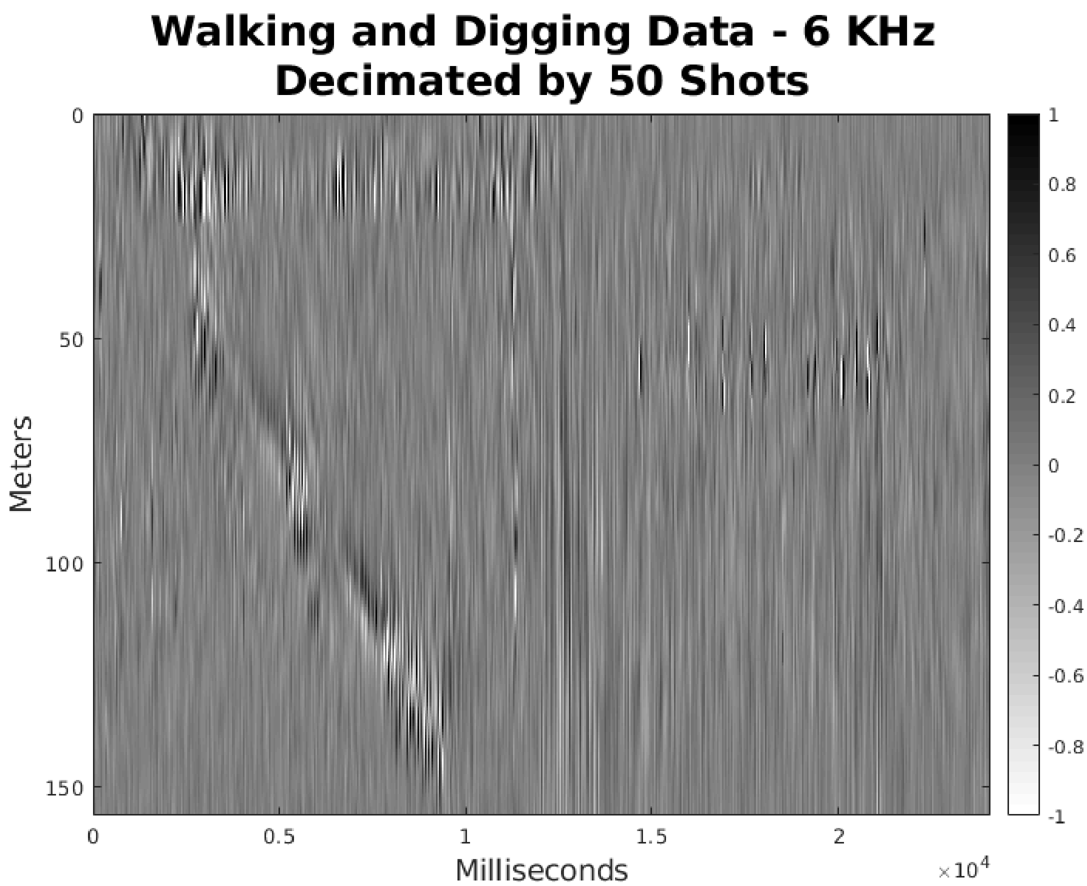


Figure 10.27: The 6kHz PRF data set, decimated by 50 shots, of someone walking next to the fibre and someone digging next to the fibre. The walking events are between 25 and 150 meters and 0.25×10^4 and 1×10^4 milliseconds. The digging events are located between 25 and 75 meters and 1.5×10^4 and 2×10^4 milliseconds.

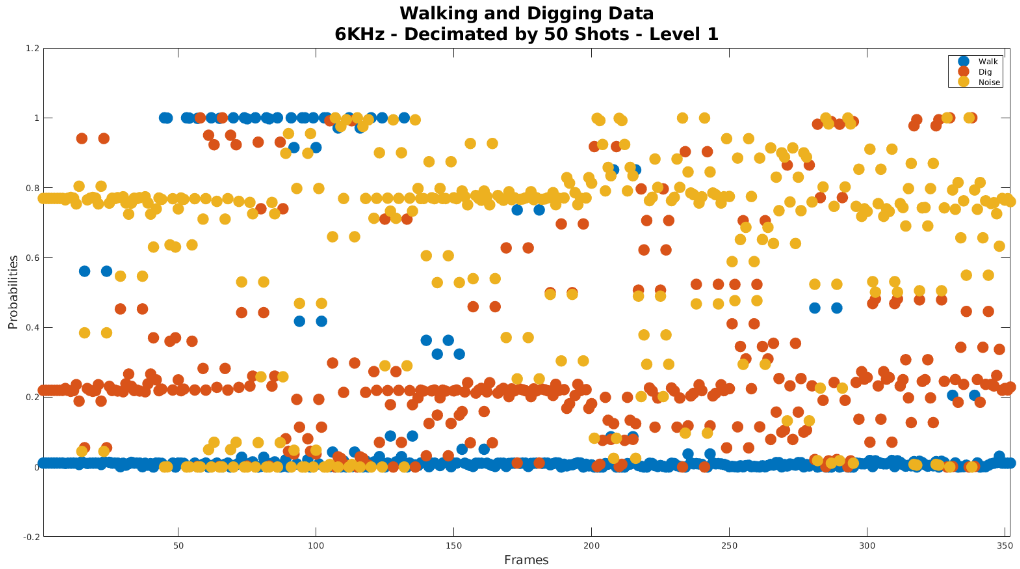


Figure 10.28: Probabilities of walking, digging, and noise events from applying the trained CNN to the 6kHz PRF walking and digging data set, decimated by 50 shots. The probabilities in each frame of a walking event (blue), digging event (red), and noise event (yellow) being present on Level 1.

shots or less, which supports our theory that the CNN would work most successfully on the data decimated by 50 shots. Furthermore, the 4kHz case, when decimated by 50 shots, had approximately 77% accuracy which is around 3% worse than how well the CNN performed on the test set.

Table 10.1: The accuracy of the convolutional neural network for each type of 4kHz PRF and 6kHz PRF acquired walking and digging data.

	No Decimation	10 Shots Decimation	50 Shots Decimation
4kHz	11%	34%	77%
6kHz	9%	0%	39%

Figure 10.29 provides examples of frames the CNN identified correctly (left column) and incorrectly (right column) of walking, digging, and noise events in the 4kHz PRF acquired data. Interestingly, the CNN identified some walking as digging as seen in the middle row of the figure. Moreover, in the bottom row, we see that the CNN distinguished a noise event which clearly has walking in it.

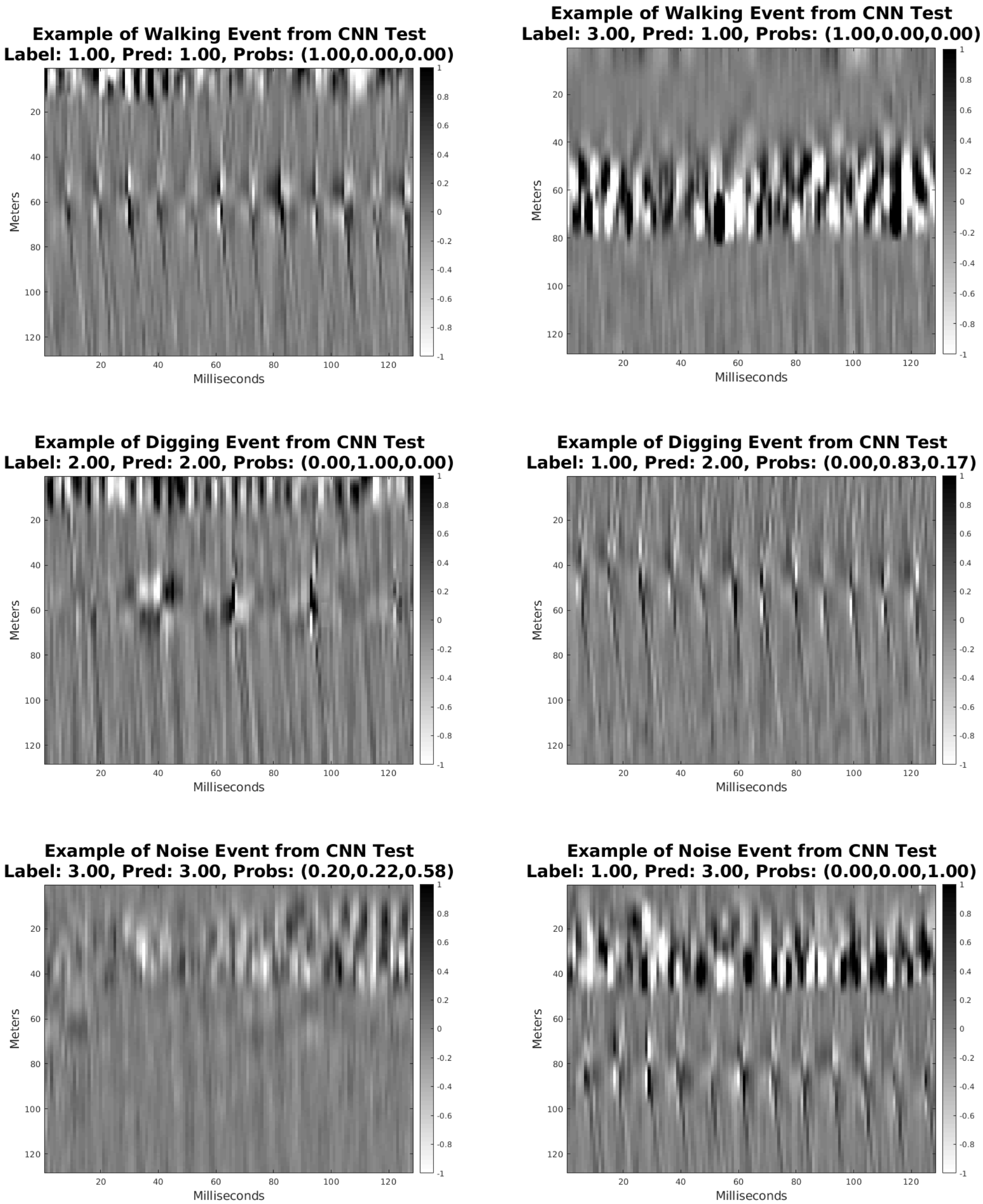


Figure 10.29: Images of walking (top row), digging (middle row), and noise (bottom row) events from the 4kHz PRF acquired data when decimated by 50 shots. The left column shows accurately defined events and the right column displays inaccurately defined events.

Before we conclude, we observe that the drastic difference in accuracy between the 4kHz case and the 6kHz case suggests that convolutional neural networks would struggle in application to identifying seismic events which contain multiple hyperbolic events in data acquired using distributed acoustic sensing and fibre-optics. Even within the more successful 4kHz case, we saw a dip in success as the number of shots we decimated by decreased. The success of the CNN depends heavily on how similar the data is to the training images from which the CNN was taught.

The experiments in this chapter applied convolutional neural networks to DAS data shown in the space-time domain. We have shown throughout this section that DAS acquired data can differ enough to cause a CNN to struggle with image-recognition. Changing the domain of the DAS data to depend on frequency may allow the convolutional neural network to be more successful. Given the opportunity, this is definitely a route to explore with regards to using convolutional neural networks for image recognition in DAS data.

The results of this experiment raise the question of feature-based image registration for data acquired using a distributed acoustic senors. Image registration is an important part of image processing or image recognition. It entails aligning two or more images into one coordinate system when the images are taken at different times, from different sensors, or from different viewpoints (Islam and Kabir, 2013). For example, features for a photograph of a person would include scaling and skewing the image as well as shifting, flipping, and rotating the person inside the image. For DAS-acquired data, the general shape of an event is hyperbolic. The features for image registration of such data would involve different methods for changing the shape or location of the hyperbolic response. Some clear attributes include scaling and shifting the response in the data. From Chapter 6, we know that the velocity of the source changes the shape of the hyperbolic response as well as the size of the gauge length applied to the data. The experiment in this section suggests that the pulse-repetition-frequency also affects the shape of the response. Not only does information about the signal determine features for image registration, but also how the distributed acoustic sensor collects the data, such as the gauge length and PRF employed. As such, all these aspects must be considered when developing image recognition methods for data collected using a DAS system.

10.3 Conclusions

We began with a short study of neural networks with an emphasis on convolutional neural networks for the purposes of image-recognition. We then focused on two different experiments. In the first experiment, we considered the architecture of the convolutional neural network we used for application to microseismic data. We introduced the training set for the neural network and the parameters we chose to train the CNN. We saw that the CNN had approximately 85% accuracy for detecting hyperbolas in the testing set. Afterwards, we applied the CNN and the inverted wavelet to two sets of microseismic data. In one case, we were able to detect events in the data set; in the other data set, we were unable to detect events. We also considered an example where we trained a CNN with similar architecture to distinguish between walking, digging, and noise. Once trained, we saw that the CNN had approximately 80% accuracy when identifying events. We applied the CNN to two data sets containing walking and digging events: one collected using a pulse repetition frequency of 4kHz and one using 6kHz. Given the relative accuracy between the two data sets, we noted that some equalizer is necessary for DAS data before image-recognition using convolutional neural networks can be generally successfully, instead of successful on isolated cases.

Chapter 11

Hyperbolic methods for determining the distance of events from the fibre

Beyond seismic applications, a fascinating development with distributed acoustic sensing is that vehicles and trains, as well as pedestrians, can be detected using DAS on fibre-optic cables installed underground in cities. Given that telecommunications widely utilize fibre-optics, these cables often cover entire cities. Since attaching an interrogator to the cable leaves the fibre's original purpose unaffected, it allows for the easy implementation of DAS to collect data in these cities. These data sets can be used to monitor vehicular traffic as well as city transit; cf. Cova et al. (2018). Unfortunately, this benefit to DAS comes at a cost: While fibre-optic cables cover most cities for different purposes, these installations often neglect to record the cable's exact location, and knowledge of its location would significantly improve information provided by DAS acquired data. Even when the fibre location is known, it would be beneficial to know the distance between the fibre sensor and an event, especially when using a DAS system for monitoring purposes.

In this chapter, we develop a method to calculate the distance between an event and the sensor in a DAS system. The method involves distinguishing between the P-wave and S-wave responses in the seismic event in order to calculate the velocity of each response and its first response in time. Seismic events are hyperbolic; see Gulati et al. (1997). Using that fact, we outline the P-wave and S-wave response in events and use the points to describe a hyperbola. This information allows us to determine how far the event is from the fibre sensor. Specifically, we use the equation:

$$D_{\mathbf{fib}} = \frac{t_{0_p} - t_{0_s}}{m_p - m_s}, \tag{11.1}$$

where t_{0_p} and t_{0_s} are the first time responses of the P-wave and S-wave, respectively, m_p is the slowness of the P-wave, and m_s is the slowness of the S-wave. Using the hyperbolic shape of the events gives the method its name: the Hyperbola Method.

We begin producing a hyperbolic event using the modeling techniques described in Chapter 6. Then, we employ the change in the slope of the event in time to determine the P-wave and S-wave response. No further processing is applied to the data in this case. We apply three different optimization methods for fitting the points to a hyperbola: a simplex search method, the BFGS method, and an 'interior-point' algorithm; cf. Lagarias et al. (1998), Kroon (2010), Byrd et al. (1999), Byrd et al. (2000), and Waltz et al. (2006).

We examine two examples: one of synthetic data without a gauge length and one of synthetic data with a 10m gauge length applied when the event starts at $t = 0$ seconds. Each optimization method gives between 11% and 33% error for locating the event with respect to fibre. When we added a time offset of $t = 0.1$ seconds to the event, the BFGS method saw a substantial increase in error compared to the case when the time offset was $t = 0$ seconds, showing that the initial estimate for the P-wave and S-wave coefficients matter significantly.

In the following examples, we attempt to lower the error by deconvolving the data and applying our methods to the deconvolved data set. The deconvolution emphasizes the P-wave and S-wave response and allows for easier separation. When we deconvolved the data before applying the Hyperbola Method, the error for our initial estimate of the coefficients drops to approximately 9% error for one example and stays below a 20% error for every other deconvolution example. The results of each optimization method showed an error bounded above 35%; however, for most deconvolved examples, the error for each method was between 7% and 16%. While the BFGS method improved when a deconvolution was applied to the data, it struggled to estimate the hyperbola coefficients for the P-wave response when the signal was offset by $t = 0.1$s as it did for the case when the data with a signal offset by $t = 0.1$ was not deconvolved. Evenso, many of these optimization techniques produced a better fit for the hyperbolae of the P-wave and S-wave response when we employed a deconvolution. We hypothesize that when using the Hyperbola Method on real data, deconvolving the data will produce better results than not implementing the deconvolution.

Let us observe some caveats of the distance between an event and the fibre this method calculates. Either the location of the event or of the fibre must be known in order to find the other. This method gives a radial distance around a specific point along the fibre. The event or fibre lays somewhere in a radial disc of length $D_{\mathbf{fib}}$. Ideally, the exact location of the event or fibre would be deduced; however, this method gives us a significant step in that direction.

11.1 Model of events

For this experiment, we use synthetic models, where a 500m straight fibre-optic cable is laid horizontally, buried 0.5m below the Earth's surface. We created a model of a source located approximately 26m from the fibre, midway down the fibre distance. Figure 11.1 (left column) shows when the source starts at $t = 0$ seconds whereas Figure 11.1 (right column) presents the data when the source begins at $t = 0.1$ seconds. The top row is the data without a gauge length applied and the bottom row is the data with a 10m gauge length applied.

We used Equation 6.95 as the source function for the models in this chapter, which we

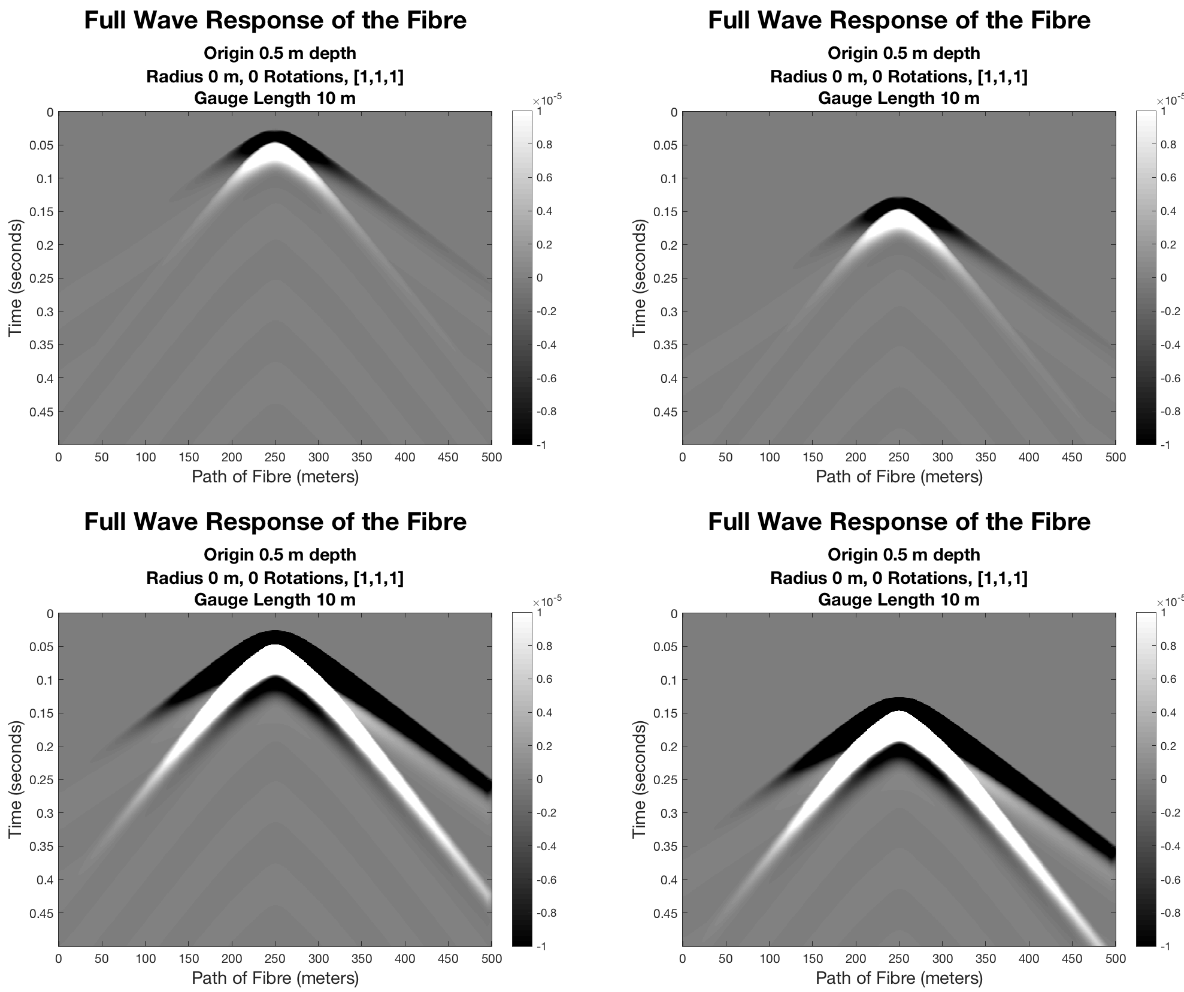


Figure 11.1: (Top row) The data which starts at $t = 0$ seconds (left) and at $t = 0.1$ seconds (right). (Bottom row) The data with a 10m gauge length applied which starts at $t = 0$ seconds (left) and at $t = 0.1$ seconds (right).

include again below as a reminder:

$$F(t) = \begin{cases} te^{-\sigma_1 t}\sin(\sigma_2 t) & t \geq 0; \\ 0 & t < 0; \end{cases} \tag{11.2}$$

where $\sigma_1 = 80$Hz and $\sigma_2 = 50$Hz. We choose the P-wave velocity of the event to be 1000m/s and the S-wave velocity to be 600m/s.

11.2 Hyperbola Method

In this section, we describe the method we implemented for finding the distance between the fibre and the event. The hyperbolic structure of the event proves fundamental to this method. Recall that the equation for a vertical hyperbola is written as

$$h_s = t_0 - m\sqrt{(s - s_0)^2 + a^2}, \tag{11.3}$$

where t_0 is the vertex of the hyperbola in time, s_0 is the position along the fibre, m is the slowness of the wave, and a is the distance between the peak and the asymptote of the hyperbola.

First, we need to find the points which describe the hyperbolic wave-front of the P-wave and S-wave. In an offset, the first waveform is the P-wave as it is faster than the S-wave. The secondary waveform is the S-wave. Since our model is based on a minimum phase system, the front of the P-wave waveform has a higher amplitude spike in the data followed by a similar spike in amplitude for the S-wave. We calculate the change in the slope of each offset in time. The higher slopes represent the location of the P-wave and S-wave respectively as the slope of a spike is sharp in comparison to the rest of the signal. With this information, we create a mask of the time-derivative data set by choosing a threshold for the slope. The mask will emphasize all slopes above the threshold. If the threshold is chosen well, it produces a clear visual of the P-wave and S-wave.

The mask is a binary matrix where every value greater than the threshold equals 1 and every value below the threshold equals zero. We apply a morphological cleaning to the mask, which is an image processing technique that removes isolated 1s and 0s. It can also be employed to ensure the P-wave and S-wave hyperbolae are connected, respectively. The `bwboundaries` function provides the boundary of binary objects in the mask. Employing the `bwboundaries` function from MATLAB, we obtain the boundary of objects in the mask to find the hyperbola of the P-wave and the hyperbola of the S-wave. We will see in the next section that this method varies depending on whether or not a deconvolution is applied to the data in time before we create the mask.

We estimate the slowness m by using the hyperbola information from the boundaries to find the slope of the hyperbola. Choosing points near the end instead the middle of the hyperbolas provides a better estimate of the wave's velocity as this is where the most change occurs. The point s_0 is chosen based on where the peak of the hyperbola occurs along the distance of the fibre. The point t_0 can be approximated based on the peak of the hyperbola's sharpness and where it occurs in time. Using the above estimations, we choose the term a by finding when the peak of the hyperbola in time t_{0_h} is equal to the asymptote of the hyperbola generated by m, s_0, and t_0 in space, call the point s_{0_a}, and then solving for the difference between s_{0_a} and s_0.

After estimating the vector $\mathbf{V_h} = [s_0, t_0, m, a]$, we optimize the cost function

$$C(\mathbf{V_h}) = \|t_h - h_{s_h}(\mathbf{V_h})\|_2 \tag{11.4}$$

for each vector $\mathbf{V_h}$, where h_{s_h} is the hyperbola described in Equation 11.3 for points s_h of the hyperbola and the variable t_h is the points of the hyperbola in time.

In this chapter, we compare the results of three optimization methods: a simplex search method, the BFGS method, and an 'interior-point' algorithm. For the simplex search method, we utilize the MATLAB function `fminsearch`, which implements the simplex search method from Lagarias et al. (1998). For the BFGS method, we employ the `fminlbfgs` function, written for MATLAB by Dirk-Jan Kroon (Kroon, 2010). Finally, the MATLAB function `fmincon` solves an unconstrained optimization problem using an 'interior-point' algorithm based on the results of Byrd et al. (1999), Byrd et al. (2000), and Waltz et al. (2006). We choose to use the `fmincon` function as it allows us to set upper and lower bounds for the vector $\mathbf{V_h}$ being optimized. We optimize the same objective function for each search method in the following examples.

Optimizing $C(\mathbf{V_h})$ provides the coefficients to the hyperbola equation describing the wave response. Using the optimized slowness m and the peak of the hyperbola in time $t_{0_h} = \min(t_h)$ of the P-wave and S-wave responses, we solve for the distance from the fibre $D_{\mathbf{fib}}$ using Equation 11.1, which is defined as follows:

$$D_{\mathbf{fib}} = \frac{t_{0_p} - t_{0_s}}{m_p - m_s}, \tag{11.5}$$

where t_{0_p} is the peak of the P-wave hyperbola in time, t_{0_s} is the peak of the S-wave in time, m_p is the slowness of the P-wave, and m_s is the slowness of the S-wave.

In the following applications of the Hyperbola Method to synthetic data, we setup a basic premise for determining the initial values of $\mathbf{V_h}$ which we do not change between each example. Beyond setting specific thresholds for each case, the work below is relatively automated; however, in some cases, we need to improve how the points of the P-wave and S-wave hyperbolae are defined. In each case, we provide a description of why we need to redefine the points of the hyperbolae and how we achieve it. Future work would entail automating this portion of determing the P-wave and S-wave hyperbolae, especially since many of the issues which arise will likely be found in real data as well.

11.3 Application of the methodology to synthetic data

We now apply the method described in the last section to the models seen in Section 11.1. We start by considering the results from the synthetic data without a gauge length applied for two different time offsets, and then from data with a 10m gauge length applied for two different time offsets. Recall that we choose the gauge length of 10m as that is a typical value used in DAS systems. Finally, we deconvolve the data in the models from Section 11.1 and apply the methods from the previous section to the deconvolved data sets.

11.3.1 Data without a gauge length applied

We start with the model seen in top left of Figure 11.1, which is the data without a gauge length applied whose source starts at $t = 0$ seconds. Taking the derivative of the offsets in time, we make a mask of the change in slope of the data. Figure 11.2 shows the derivative of the data in time on the left and the mask created using the threshold 1.00×10^{-11} on the right.

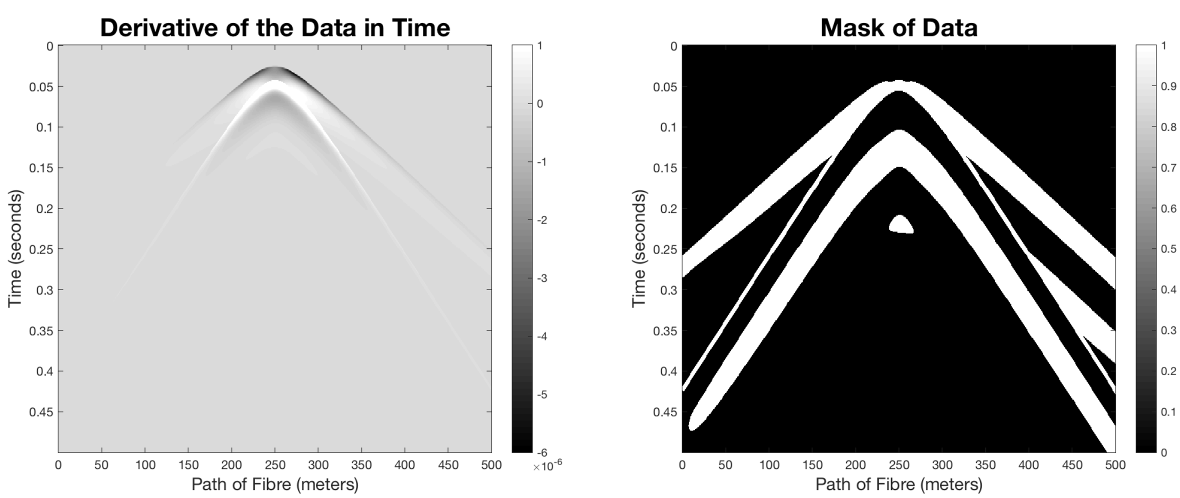


Figure 11.2: (Left) The derivative of the data in time.
(Right) The mask of the slope data using a threshold of 1.00×10^{-11}.

In the mask of the data seen in Figure 11.2, we see three separate white hyperbolic shapes. In the top hyperbolic shape, the top boundary portrays the P-wave and the bottom boundary gives an outline of the S-wave. Removing this top shape by isolating the boundaries using the `bwboundaries` function from MATLAB, we solve for the points of the hyperbolae. Figure 11.3 depicts the boundary of the top hyperbolic white shape removed from the mask of the data, and the right presents the P-wave and S-wave hyperbola determined by the boundary seen on the left.

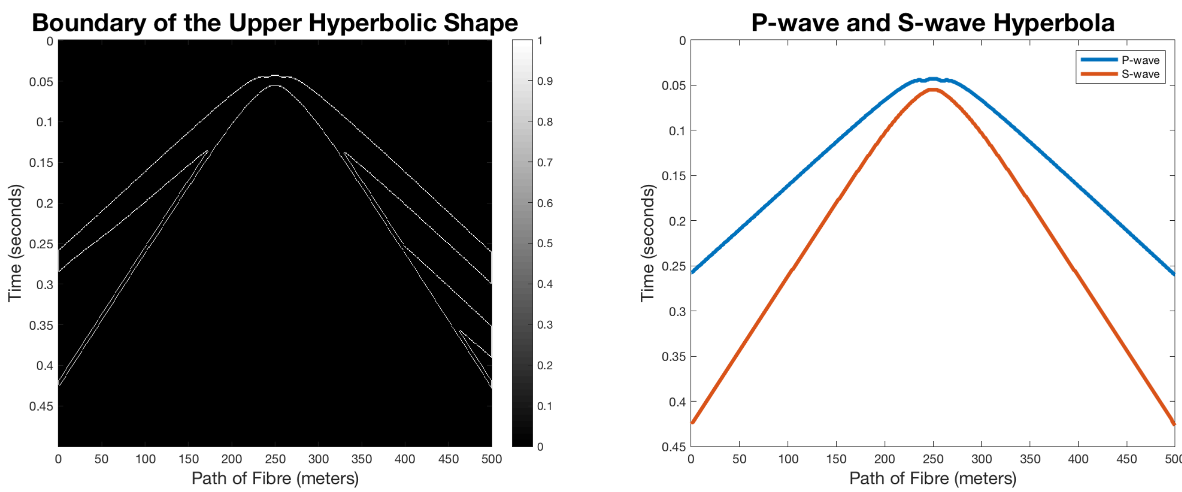


Figure 11.3: (Left) The boundary of the upper hyperbolic shape from the mask of the data. (Right) The P-wave hyperbola (blue) and the S-wave hyperbola (red) generated from the boundary of the upper hyperbolic shape from the mask of the data.

Using the points from the boundary seen in Figure 11.3, we now solve for the estimates of the vector $\mathbf{V_h} = [s_0, t_0, m, a]$ for the coefficients of the hyperbola equation described by these points. We assume that $t_0 = 0$ for both the P-wave and S-wave coefficients, given that the hyperbolae lie between $t = 0$ and $t = 0.5$. The P-wave and S-wave are generated from the source at the same point in time. While we do know from creating the model that this choice of t_0 is correct, we will see in the examples where time is offset by $t = 0.1$ that choosing an initial value of $t_0 = 0$ proves adequate for optimizing the coefficients.

We find s_0 by considering the minimum value of the points in time, t_h, for both the P-wave and the S-wave hyperbolae and relating them to their respective point s_h along the fibre. Given that the same source generates both the P-wave and S-wave, s_0 and t_0 should be the same for both the P-wave and the S-wave. This calculation gives us $s_0 = 249.7$m for this case.

To solve for slowness m, we calculate the slope of each hyperbola by determining the change in time over the change in the path of the fibre. In this case, we examine the change between $x_1 = 71$m and $x_2 = 213.7$m along the fibre path. Finding the relevant t_1 and t_2 gives an estimate of $m_p = 9.00 \times 10^{-4}$s/m for the P-wave and $m_s = 1.60 \times 10^{-3}$s/m for the S-wave. The estimated P-wave velocity is $1/m_p = 1111$m/s and the estimated S-wave velocity is $1/m_s = 625$m/s, which are both relatively close the actual values.

Finally, we define the asymptotes of the hyperbolae using s_0, t_0 and m. For the P-wave, the asymptote of the hyperbola is

$$y_P = m_p(s_p - s_0) + t_0 \tag{11.6}$$

and

$$y_S = m_s(s_s - s_0) + t_0 \tag{11.7}$$

where s_p is the spacial points along the fibre path of the P-wave hyperbola and s_s is the spacial points along the fibre path of the S-wave hyperbola. Now, we solve $y_P = t_{0_p}$ and $y_S = t_{0_s}$ for their respective s_p and s_s values where $t_{0_p} = 0.0429$ seconds is the time at the peak of the P-wave hyperbola and $t_{0_s} = 0.055$ seconds is the time at the peak of the S-wave hyperbola. We denote the peak of the P-wave wave and S-wave in space as s'_p and s'_s, respectively. It then follows that $a_p = |s'_p - s_0| = 45.3$m and $a_s = |s'_s - s_0| = 34.7$m; we take the absolute value of the difference because a_p and a_s describe distances.

Now, we have our initial estimates for the P-wave and S-wave coefficients. Specifically, the P-wave estimate is

$$\mathbf{V_p} = [s_0, t_0, m_p, a_p] \tag{11.8}$$

$$= [249.7, 0, -9.00 \times 10^{-4}, 45.3] \tag{11.9}$$

and the S-wave estimate is

$$\mathbf{V_s} = [s_0, t_0, m_s, a_s] \tag{11.10}$$

$$= [249.7, 0, -1.60 \times 10^{-3}, 34.7]. \tag{11.11}$$

We choose the following upper and lower bounds, u_b and l_b respectively, for the 'interior-point' algorithm with respect to the P-wave vector $\mathbf{V_p}$:

$$0 \leq s_0 \leq \infty; \tag{11.12}$$

$$0 \leq t_0 \leq t_{0_p}; \tag{11.13}$$

$$-\infty \leq m_p \leq \infty; \tag{11.14}$$

$$0 \leq a_p \leq \infty. \tag{11.15}$$

We pick these bounds based on the following reasoning: s_0 and a_p describe distances so they are both be greater than 0, the vertex t_0 of the hyperbola must be greater than 0 since it describes a time when the event started but also less than the peak of the P-wave hyperbola in time t_{0_p}, and m_p is the slowness of the P-wave. We use some of the optimized values from the P-wave coefficients to produce bounds for the S-wave initial estimate $\mathbf{V_s}$ which we set as follows:

$$0 \leq s_0 \leq \infty; \tag{11.16}$$

$$0 \leq t_0 \leq t_{0_p}; \tag{11.17}$$

$$-\infty \leq m_s \leq \infty; \tag{11.18}$$

$$0 \leq a_p \leq \infty; \tag{11.19}$$

where the parameter t_{0_p} is the peak of the P-wave hyperbola in time. We use the bounds for every remaining example in this chapter.

Tables 11.1 and 11.2 presents the optimized coefficients from the respective optimization methods with the initial estimate for comparison. For each optimization method, we force the P-wave and S-wave s_0s to be equivalent by taking the average of the two values if they are not already equal. We also force t_0 to be greater than 0 as we assume that the source is not detonated until after the distributed acoustic sensor is turned on. We apply these conditions to the remaining examples in this chapter as well.

	P-wave Coefficients
$\mathbf{V_p}$	$[s_0, t_0, m_p, a_p]$
Input	$[249.7, 0, -9.00 \times 10^{-4}, 45.3]$
Simplex Search	$[249.8, 5.70 \times 10^{-3}, -1.00 \times 10^{-3}, 35.3]$
BFGS	$[249.8, 0, -1.00 \times 10^{-3}, 45.3]$
'Interior-point'	$[249.8, 5.80 \times 10^{-3}, -1.00 \times 10^{-3}, 35.3]$

Table 11.1: The optimized P-wave coefficients of the hyperbolae functions describing the P-wave respectively for each optimization method.

	S-wave Coefficients
$\mathbf{V_s}$	$[s_0, t_0, m_s, a_s]$
Input	$[249.7, 0, -1.60 \times 10^{-3}, 34.7]$
Simplex Search	$[249.8, 6.20 \times 10^{-3}, -1.70 \times 10^{-3}, 29.7]$
BFGS	$[249.8, 0, -1.70 \times 10^{-3}, 34.6]$
'Interior-point'	$[249.8, 9.90 \times 10^{-3}, -1.7 \times 10^{-3}, 27.0]$

Table 11.2: The optimized S-wave coefficients of the hyperbolae functions describing the S-wave respectively for each optimization method.

Figure 11.4 shows the data with the P-wave (blue) and S-wave (red) hyperbola function generated by each optimization method. For the most part, the estimates fit the hyperbola event for each optimization method; however, each method struggled to fit the peak of the P-wave hyperbola to the data. The optimized coefficients for the S-wave appear to be more successful for all three optimization methods than the optimized P-wave coefficients. Given how difficult fitting the peak of the hyperbolic event, this suggests that the slope near the peak of the P-wave was not sharp enough to be detected by our threshold. This may be due in some part to interference from the S-wave response as there exists an overlap in the responses. The overlap of the two responses may have caused the P-wave to lose some of its energy.

Using the optimized P-wave and S-wave coefficients, we calculate the fitted peaks of the P-wave and S-wave hyperbolae and employ Equation 11.1 to find the distance between the event and the fibre. Table 11.3 presents the results for $D_{\mathbf{fib}}$ and the relative error in comparison to the actual closest location of the fibre with respect to the source. We calculate the error using the following equation:

$$\text{Error} = \frac{|L - D_{\mathbf{fib}}|}{|L|} \times 100, \tag{11.20}$$

where L is the actual distance between the fibre and the source. For these models, $L = 26$m. All the optimization techniques except the BFGS method have an error less than 30%. The simplex search method performed the best with an error less than 17%.

	Distance along fibre s_0	Offset	Distance from Fibre $D_{\mathbf{fib}}$	Error
Event Location	250m		26m	
Input	249.7m	0.3m	18.80m	28%
Simple Search	249.8m	0.2m	21.83m	16%
BFGS	249.8m	0.2m	17.40m	33%
'Interior-Point'	249.8m	0.2m	20.66m	21%

Table 11.3: The estimated distance and error of three optimization methods in estimating the distance between the fibre and the event when the event starts at $t = 0$ seconds.

Events in DAS acquired data do not always occur at $t = 0$ seconds. As such, we include an example where we apply the methods describe for the previous example to a case when the source occurs at $t = 0.1$ seconds. We follow the same method outlined for the case when the source occurs at $t = 0$ seconds.

We use the same threshold (1.00×10^{-11}) from the case when the source started at $t = 0$ seconds. Figure 11.5 displays the time-derivative of the data when the source starts at $t = 0.1$ seconds on the left and the mask of the data on the right. The clearest distinction between the two data sets is that the peak of the hyperbolic response occurs at a later time than the

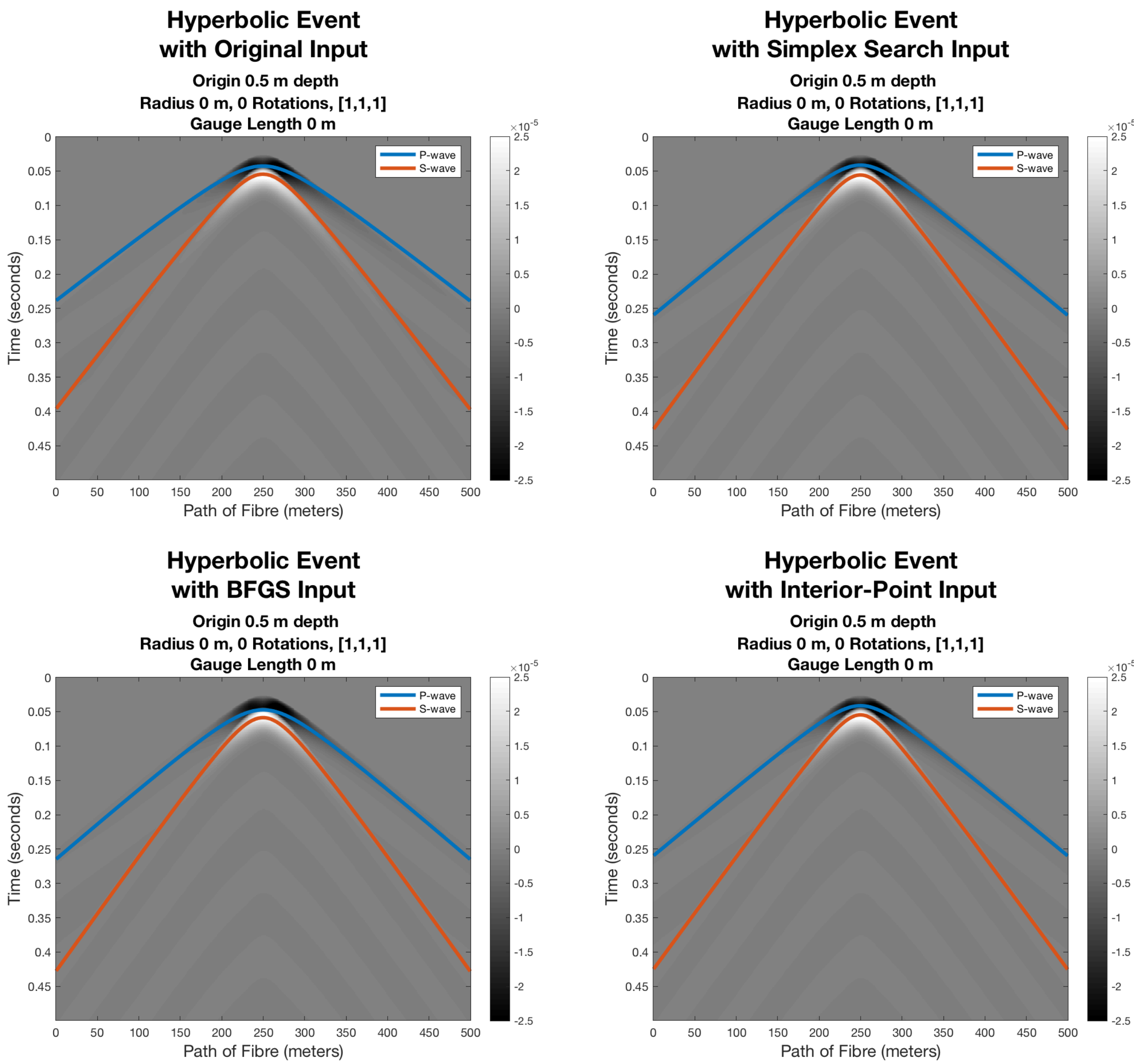


Figure 11.4: A hyperbola fitted to the P-wave (blue) and S-wave (red) response of the event when it starts at $t = 0$ seconds. The hyperbola fitting for (top left) the original input vector, (top right) the optimized vector from the simplex search method, (bottom left) the optimized vector from the BFGS method, and (bottom right) the optimized vector from the ‘interior-point’ algorithm.

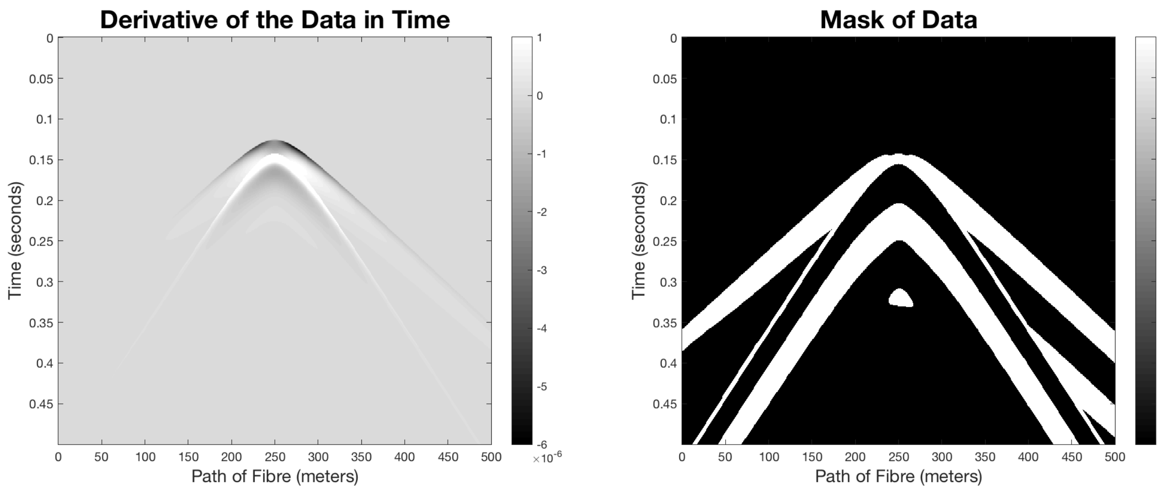


Figure 11.5: (Left) The time-derivative of the data when the source starts at $t = 0.1$ seconds. (Right) The mask of derivative data using a threshold of 1.00×10^{-11}.

case pictured in Figure 11.2.

Once again, we find the boundary of the upper white hyperbolic shape in the mask, where the top boundary provides points of the P-wave hyperbola and the bottom boundary provides the points of the S-wave hyperbola. Figure 11.6 shows the boundary of the upper hyperbolic shape in the mask in the top row, and the P-wave and S-wave hyperbolae determined by the boundary on the bottom row.

The S-wave does not fit the full path of the fibre. Given that we choose points along the bottom boundary at each offset for the S-wave hyperbola points, a sharp decrease is seen at around offset 20 for the hyperbola points of the bottom boundary and a sharp increase at around offset 480; c.f. Figure 11.6 (bottom left). Taking the derivative of the S-wave points in time, which we denote as t_s, and finding the maximum and minimum slopes gives the offsets where the S-wave is no longer visible; see Figure 11.6 (bottom right).

We employ the same strategy to estimate the initial P-wave and S-wave coefficient vectors $\mathbf{V_p}$ and $\mathbf{V_s}$ that we used for the data when the source begins at $t = 0$ seconds. For the upper and lower bounds in the 'interior-point' algorithm, see Equations 11.12 to 11.19. Tables 11.4 and 11.5 presents the initial estimate for $\mathbf{V_p}$ and $\mathbf{V_s}$ as well as the results of each optimization

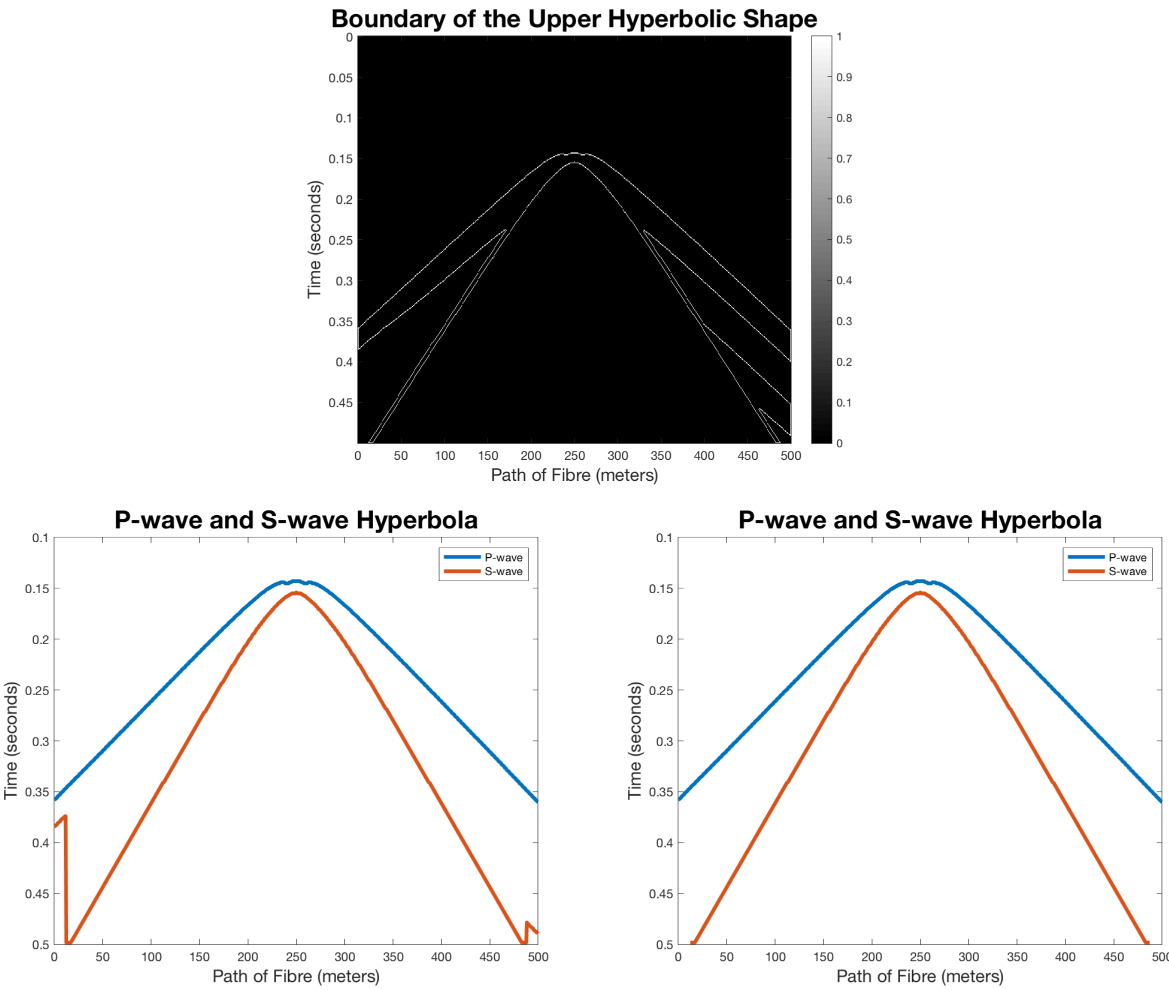


Figure 11.6: (Top) The boundary of the upper hyperbolic shape from the mask of the data when the source starts at $t = 0.1$ seconds.
(Bottom left) The P-wave hyperbola(blue) and the S-wave hyperbola (red) generated by the boundary of the upper hyperbolic shape from the mask of the data.
(Bottom right) The P-wave hyperbola (blue) and the S-wave hyperbola (red) generated from the boundary of the upper hyperbolic shape from the mask of the data after removing points which were not part of the S-wave response.

method.

	P-wave Coefficients
$\mathbf{V_p}$	$[s_0, t_0, m_p, a_p]$
Input	$[249.7, 0, -9.00 \times 10^{-4}, 152]$
Simplex Search	$[249.8, 0.1057, -1.00 \times 10^{-3}, 35.4]$
BFGS	$[249.8, 0, -1.50 \times 10^{-3}, 152]$
'Interior-point'	$[249.8, 0.1042, -1.00 \times 10^{-3}, 37.1]$

Table 11.4: The optimized P-wave coefficients of the hyperbolae functions describing the P-wave respectively for each optimization method. These hyperbola coefficients were generated for the data when the source starts at $t = 0.1$ seconds.

	S-wave Coefficients
$\mathbf{V_s}$	$[s_0, t_0, m_s, a_s]$
Input	$[249.7, 0, -1.60 \times 10^{-3}, 98.7]$
Simplex Search	$[249.8, 0.1067, -1.70 \times 10^{-3}, 29.4]$
BFGS	$[249.8, 0.1064, -1.70 \times 10^{-3}, 29.7]$
'Interior-point'	$[249.8, 0.0977, -1.7 \times 10^{-3}, 35.8]$

Table 11.5: The optimized S-wave coefficients of the hyperbolae functions describing the S-wave respectively for each optimization method. These hyperbola coefficients were generated for the data when the source starts at $t = 0.1$ seconds.

The BFGS method suffers the most from the choice of $t_0 = 0$ for both the initial estimates for the P-wave coefficients; however, it manages to perform well for the S-wave coefficients. In comparison, the simplex search method and 'interior-point' algorithm perform well despite the poor initial choice of t_0 for both the P-wave and S-wave.

Figure 11.7 depicts the data when the source starts at $t = 0.1$ seconds using the coefficients for the P-wave and S-wave from Tables 11.4 and 11.5. The input estimates for $\mathbf{V_p}$ and $\mathbf{V_s}$ clearly show a poor fit to the data in the top left of Figure 11.7 whereas the BFGS method provides a decent fit for the S-wave response, but does not fare as well for the P-wave response in the signal. The methods in simplex search method and 'interior-point' algorithm struggle with the peak of the P-wave response, as in the previous case, but fit the S-wave response relatively well.

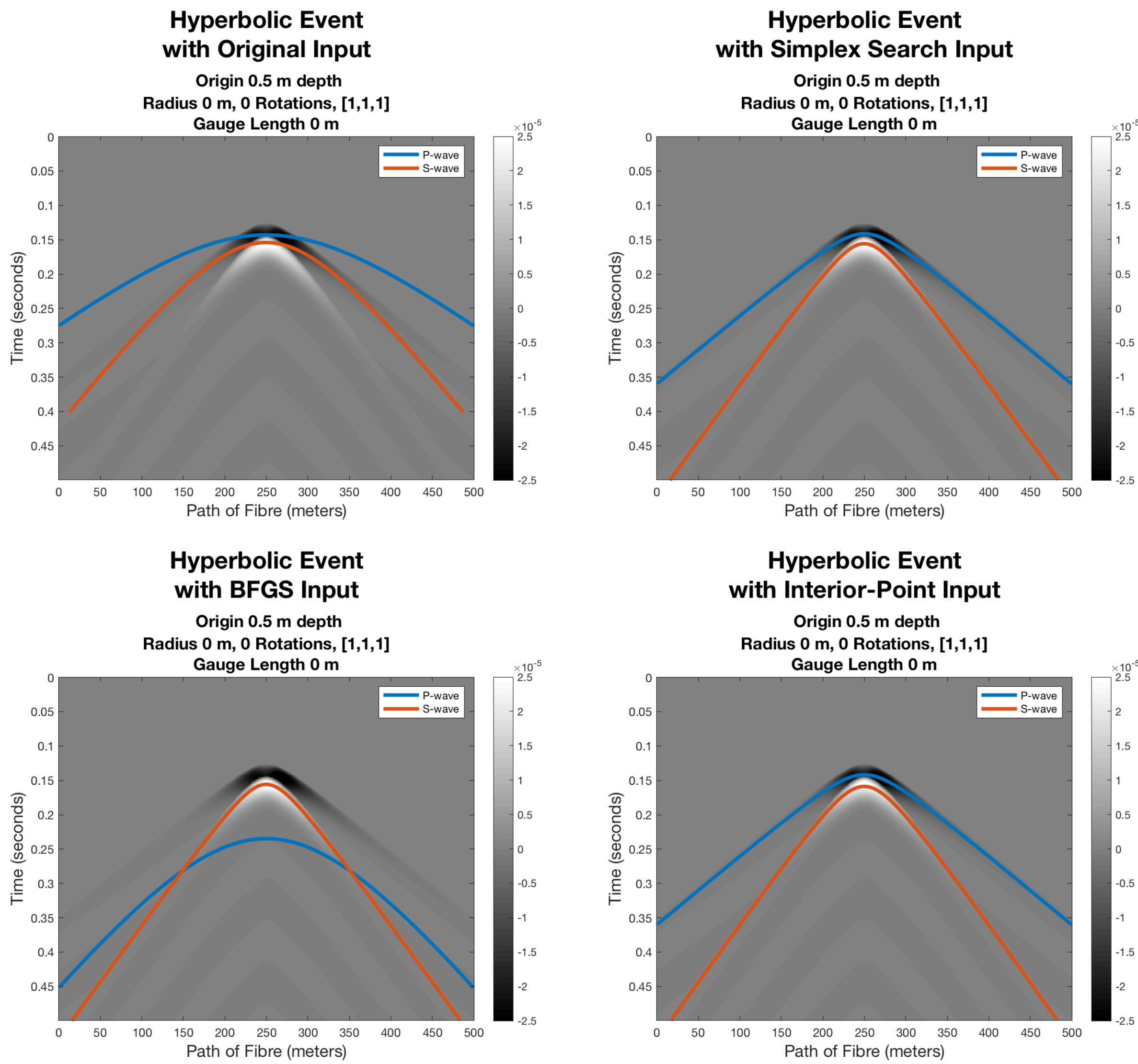


Figure 11.7: A hyperbola fitted to the P-wave (blue) and S-wave (red) response of the event when it starts at $t = 0.1$ seconds. The hyperbola fitting for (top left) the original input vector, (top right) the optimized vector from the simplex search method, (bottom left) the optimized vector from the BFGS method, and (bottom right) the optimized vector from the 'interior-point' algorithm.

Table 11.6 provides the resulting fibre distance using the P-wave and S-wave coefficients estimated in Tables 11.4 and 11.5. Recall that the BFGS method produced a poor fit of the hyperbolae to the hyperbolic response of the event especially for the P-wave coefficients. In fact, Tables 11.4 and 11.5 shows that the BFGS method optimized the P-wave velocity to be $1/m_p = 666$m/s, which is very close to the estimated S-wave velocity. As such the 2385% error follows from the poor fit to the data especially since the P-wave hyperbola maps to points below the S-wave hyperbola in Figure 11.7 (bottom left). Interestingly, the 'interior-point' algorithm saw a decrease in error from the previous case, whereas the simplex search method saw a 0.4% increase in error.

	Distance along fibre s_0	Error	Distance from Fibre $D_{\mathbf{fib}}$	Error
Event Location	250m		26m	
Input	249.7m	0.3m	17.80m	32%
Simplex Search	249.8m	0.2m	21.73m	16%
BFGS	249.8m	0.2m	646.4m	2385%
'Interior-Point'	249.8m	0.2m	24.56m	6%

Table 11.6: The estimated distance and accuracy of three optimization methods in estimating the distance between the fibre and the event when the event starts at $t = 0.1$ seconds.

11.3.2 Data with a 10m gauge length applied

Typically, a DAS system applies a gauge length to the data, which we discussed in Chapter 6. As such, we now consider what occurs when the data has a 10 meter gauge length applied to it. We set the gauge length equal to 10m as that is a standard choice in DAS systems. Also, we saw in Section 6.2.3 that the 10m gauge length provides some flattening of the hyperbolic response, but less than that of a 20m or 25m gauge length. This choice allows us to see the effect the gauge length has on estimating the fibre distance, given that it does affect the appearance of the data. We follow the exact same strategy we say in the previous section. Figure 11.8 shows the time-derivative of the data with a 10m gauge length applied on the left, and the mask created using the threshold 1.00×10^{-9} on the right.

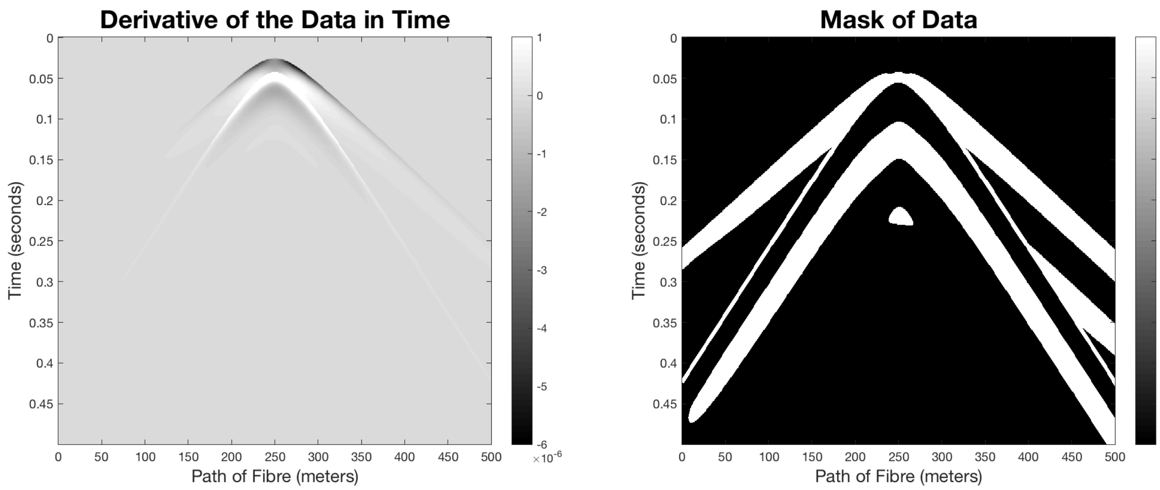


Figure 11.8: (Left) The time-derivative of the data with a 10m gauge length applied. (Right) The mask of the derivative of the data with a 10m gauge length applied using a threshold of 1.00×10^{-9}.

After a morphological cleaning, we take the upper white hyperbolic shape and find the boundaries as it provides a decent outline of the P-wave and S-wave hyperbolae. Figure 11.9 shows the boundary on the left and the hyperbolae for the P-wave (blue) and S-wave (red) deduced from the boundary on the right.

Using these points found in Figure 11.9, we estimate the coefficients for the hyperbola function describing the P-wave and S-wave respectively. We use the same choice of upper and lower bounds for the 'interior-point' algorithm from Equations 11.12 to 11.19. Tables 11.7 and 11.8 presents the initial coefficients chosen for the P-wave and S-wave in the top row and the optimized results in the next three rows.

In Figure 11.10, we include the hyperbolae generated by the coefficients in Tables 11.7 and 11.8 with the data that has a 10m gauge length applied to it. The P-wave hyperbola is blue and the S-wave hyperbola is red. As with the data without a gauge length applied, fitting the peak of the hyperbolic event for the P-wave response proves to be difficult. This is likely due to the slope changing by a degree smaller than the threshold we chose for this case. The fit to the S-wave response is less precise than for the case where the data did not have a gauge length applied.

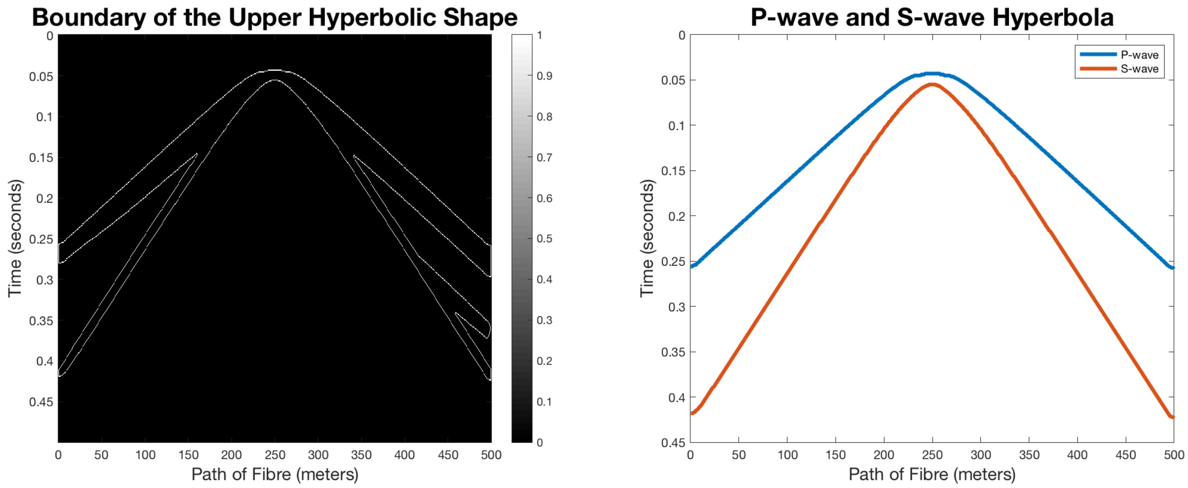


Figure 11.9: (Left) The boundary of the upper hyperbolic shape from the mask of the data when a 10m gauge length is applied.
(Right) The P-wave hyperbola (blue) and the S-wave hyperbola (red) generated from the boundary of the upper hyperbolic shape from the mask of the data with a 10m gauge length applied.

	P-wave Coefficients
$\mathbf{V_p}$	$[s_0, t_0, m_p, a_p]$
Input	$[249.7, 0, -9.00 \times 10^{-4}, 45.3]$
Simplex Search	$[249.8, 6.80 \times 10^{-3}, -1.00 \times 10^{-3}, 34.1]$
BFGS	$[249.8, 0, -1.00 \times 10^{-3}, 45.3]$
'Interior-point'	$[249.8, 7.20 \times 10^{-3}, -1.00 \times 10^{-3}, 33.7]$

Table 11.7: The optimized P-wave coefficients of the hyperbolae functions describing the P-wave respectively for each optimization method. These hyperbola coefficients were generated for the data with a 10m gauge length applied.

	S-wave Coefficients
$\mathbf{V_s}$	$[s_0, t_0, m_s, a_s]$
Input	$[249.7, 0, -1.6 \times 10^{-3}, 34.7]$
Simplex Search	$[249.8, 9.30 \times 10^{-3}, -1.70 \times 10^{-3}, 27.9]$
BFGS	$[249.8, 7.00 \times 10^{-4}, -1.70 \times 10^{-3}, 34.5]$
'Interior-point'	$[249.8, 9.40 \times 10^{-3}, -1.70 \times 10^{-3}, 27.9]$

Table 11.8: The optimized S-wave coefficients of the hyperbolae functions describing the S-wave respectively for each optimization method. These hyperbola coefficients were generated for the data with a 10m gauge length applied.

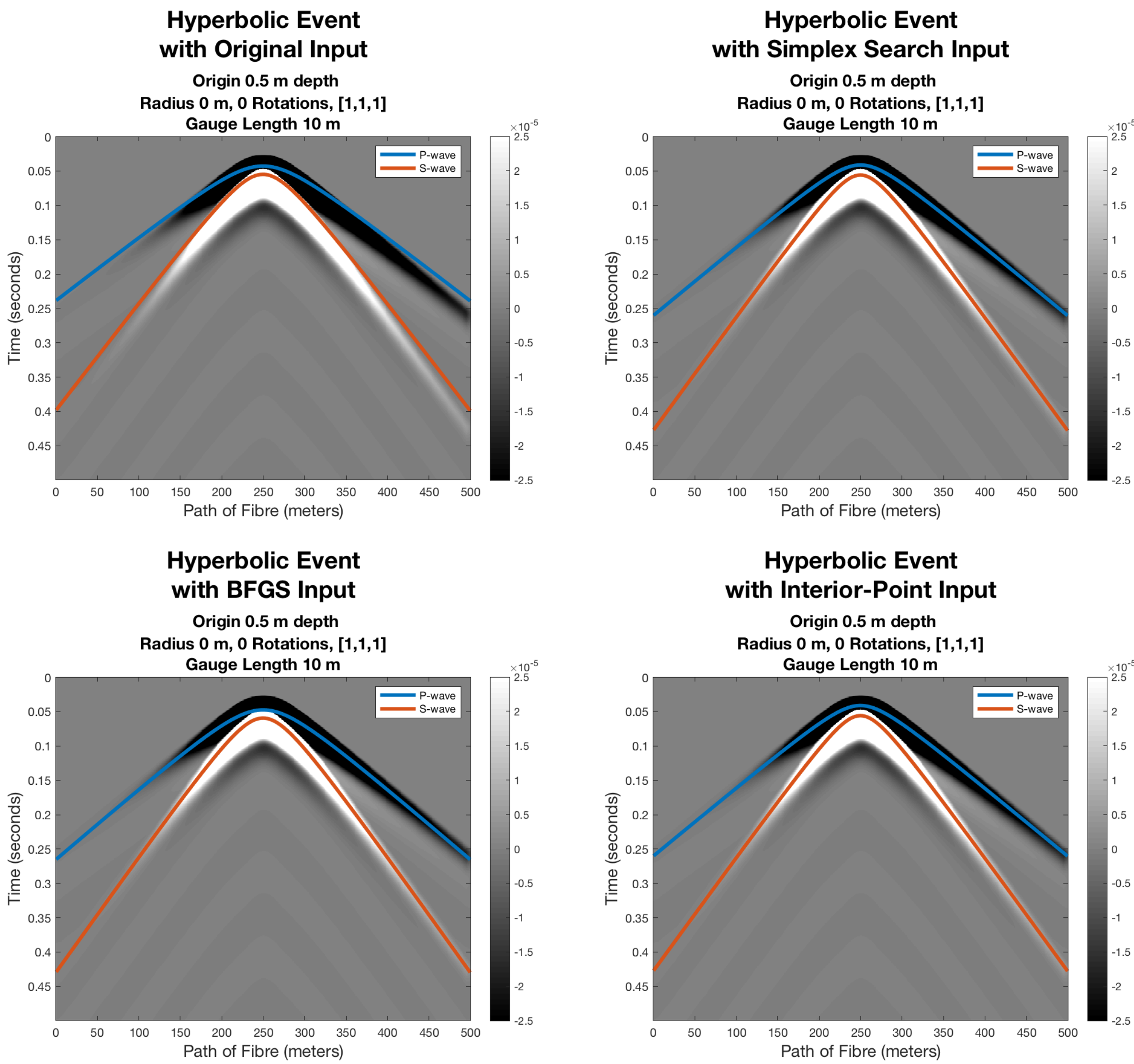


Figure 11.10: A hyperbola fitted to the P-wave (blue) and S-wave (red) response of the event when a 10m gauge length is applied to the data. The event starts at $t = 0$ seconds. The hyperbola fitting for (top left) the original input vector, (top right) the optimized vector from the simplex search method, (bottom left) the optimized vector from the BFGS method, and (bottom right) the optimized vector from the 'interior-point' algorithm.

Table 11.9 presents the distance between the fibre and the event as well as the error with respect to the actual distance between the source and the fibre in the model. Interestingly, applying a gauge length to the data causes the error to decrease for the initial input and the results of all three optimization methods. In fact, they each gave an error of less than 30%. The simplex search method and the 'interior-point' algorithm reach errors below 15%.

	Distance along fibre s_0	Error	Distance from Fibre $D_{\mathbf{fib}}$	Error
Event Location	250m		26m	
Input	249.7m	0.3m	18.99m	27%
Simplex Search	249.8m	0.2m	22.25m	14%
BFGS	249.8m	0.2m	18.26m	30%
'Interior-Point'	249.8m	0.2m	22.31m	14%

Table 11.9: The estimated distance and error of three optimization methods in estimating the distance between the fibre and the event when the event starts at $t = 0$ seconds in the data with a 10m gauge length applied.

Let us now investigate how well our method works for estimating the distance between the fibre and the source when the data has a gauge length of 10m applied to it and the source starts at $t = 0.1$ seconds. This allows us to see which method performs the best given poor initial conditions. We employ the same techniques that we used for the previous cases to estimate the initial values of the $\mathbf{V_p}$ and $\mathbf{V_s}$ and only provide the initial estimates and the results of the optimization methods later in the section.

To create the mask of the time-derivative data, we set the threshold equal to 1.00×10^{-9}. Figure 11.11 (right) shows the mask of the time-derivative data.

We take the boundary of the upper white hyperbolic shape in the mask in order to find the points of the P-wave and S-wave hyperbolae. For the P-wave, we use points from the top boundary of the upper hyerpbolic shape in the mask and for the S-wave, we take points from the bottom boudnary. The S-wave response does not fit the full length of the fibre path. As such, we remove the excess points like we did for the case when the data did not have a gauge length applied to it and the source started at $t = 0$ seconds. Figure 11.12 (left) provides a visual of the boundary of the upper hyperbolic shape in the mask, and Figure 11.12 (right)

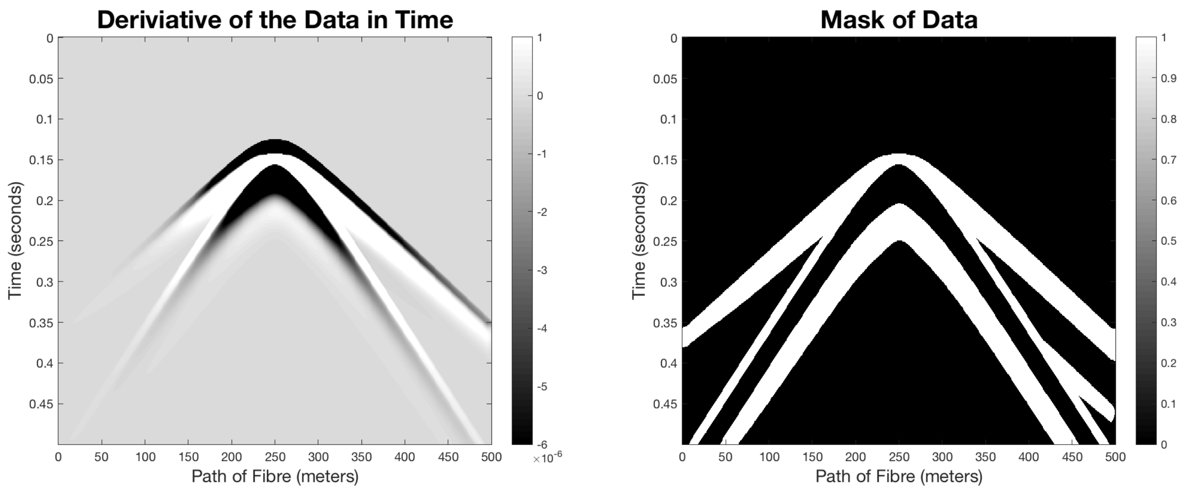


Figure 11.11: (Left) The time-derivative of the data with a 10m gauge length applied when the source starts at $t = 0.1$ seconds.
(Right) The mask of the time-derivative of the data with a 10m gauge length applied using a threshold of 1.00×10^{-9}.

shows the resulting P-wave hyperbola (blue) and S-wave hyperbola (red).

With the points from Figure 11.12, we estimate the coefficients of the P-wave and S-wave hyperbolae using the same strategy as before. Tables 11.10 and 11.11 yields the initial value of the P-wave and S-wave coefficients for this case as well as the results of the three different optimization methods.

	P-wave Coefficients
$\mathbf{V_p}$	$[s_0, t_0, m_p, a_p]$
Input	$[249.7, 0, -9.00 \times 10^{-4}, 152]$
Simplex Search	$[249.8, 0.1068, -1.00 \times 10^{-3}, 34.1]$
BFGS	$[249.7, 0, -1.50 \times 10^{-3}, 152]$
‘Interior-point’	$[249.8, 0.1054, -1.00 \times 10^{-3}, 35.7]$

Table 11.10: The optimized P-wave coefficients of the hyperbolae functions describing the P-wave respectively for each optimization method. These hyperbola coefficients were generated for the data with a 10m gauge length applied when the source starts at $t = 0.1$ seconds.

For the two experiments where the source starts at $t = 0.1$ seconds instead of $t = 0$ seconds, we chose $t_0 = 0$ for both the P-wave and S-wave coefficients. This affects the initial choices of the P-wave and S-wave coefficient in the fourth term of the vector; see a_p and

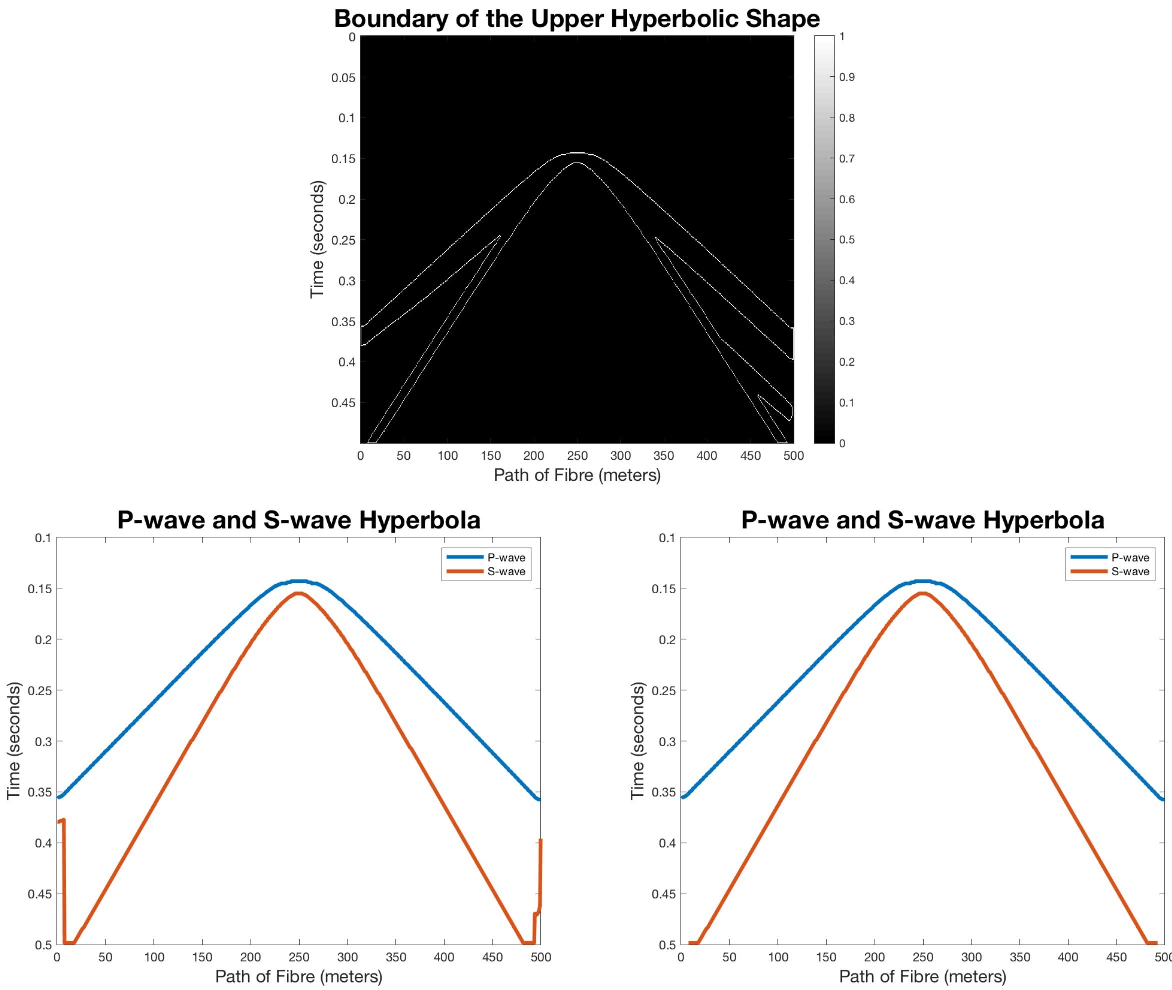


Figure 11.12: (Top) The boundary of the upper hyperbolic shape from the mask of the data with a 10m gauge length applied where the source starts at $t = 0.1$ seconds.
(Bottom left) The P-wave hyperbola (blue) and the S-wave hyperbola (red) generated from the boundary of the upper hyperbolic shape from the mask of the data with a 10m gauge length applied.
(Bottom right) The P-wave hyperbola (blue) and the S-wave hyperbola (red) generated from the boundary of the upper hyperbolic shape from the mask of the data with a 10m gauge length applied after removing points which were not part of the S-wave response.

	S-wave Coefficients
$\mathbf{V_s}$	$[s_0, t_0, m_s, a_s]$
Input	$[249.7, 0, -1.6 \times 10^{-3}, 98.0]$
Simplex Search	$[249.8, 0.1115, -1.70 \times 10^{-3}, 26.6]$
BFGS	$[249.7, 0, -2.10 \times 10^{-3}, 98.0]$
'Interior-point'	$[249.8, 0.1088, -1.70 \times 10^{-3}, 28.7]$

Table 11.11: The optimized S-wave coefficients of the hyperbolae functions describing the S-wave respectively for each optimization method. These hyperbola coefficients were generated for the data with a 10m gauge length applied when the source starts at $t = 0.1$ seconds.

a_s. For the case when the time was offset by $t = 0$ seconds, the initial estimates for the a_p and a_s terms were much smaller than in this case. We solved for these by finding when the peak of the hyperbola in time intersected the asymptote of the hyperbola. With our initial guess of $t_0 = 0$ for both the P-wave and S-wave coefficients in the case when the source starts at $t = 0.1$ seconds, the a_p and a_s terms are much larger. The simplex search method and 'interior-point' algorithm overcame the poor choice of t_0, a_p, and a_s whereas the BFGS method did not.

Using the coefficients from Tables 11.10 and 11.11, we see how well the P-wave and S-wave hyperbolae for each set of coefficients fit the data in Figure 11.13.

As to be expected from the results in Tables 11.10 and 11.11, the P-wave coefficients estimated by the BFGS method provide a poor fit for the hyperbolic response of the event; however, the S-wave coefficients appear to adequately fit the data; c.f. Figure 11.13 (bottom left). For the initial guess of the P-wave and S-wave coefficients, neither set provides a good fit for the event; however, the fit for the results of the simplex search method and 'interior-point' algorithm produce an efficient fit of the data despite the recurring issue of a poor fit to the peak of the P-wave response.

Both the simplex search method and the 'interior-point' algorithm have the lowest error with regards to determining the distance between the event and the fibre as seen Table 11.12. This follows as they provided the best fit for the data; c.f. Figure 11.13. The BFGS method has the largest error at over 60%, which is a significant improvement to the case when

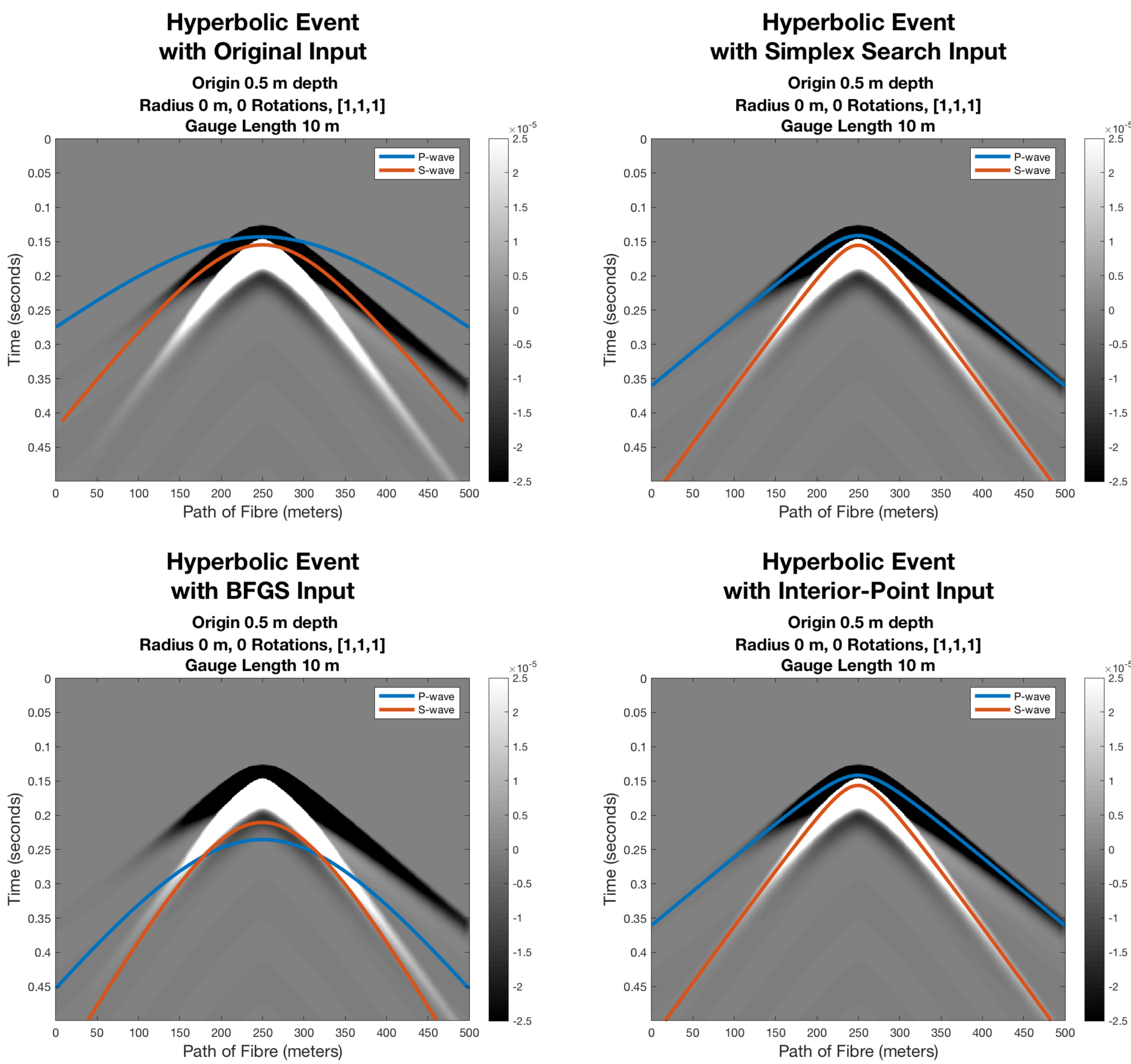


Figure 11.13: A hyperbola fitted to the P-wave (blue) and S-wave (red) response of the event when a 10m gauge length is applied to the data. The event starts at $t = 0.1$ seconds. The hyperbola fitting for (top left) the original input vector, (top right) the optimized vector from the simplex search method, (bottom left) the optimized vector from the BFGS method, and (bottom right) the optimized vector from the 'interior-point' algorithm.

	Distance along fibre s_0	Error	Distance from Fibre $D_{\mathbf{fib}}$	Error
Event Location	250m		26m	
Input	249.7m	0.3m	18.38m	29%
Simplex Search	249.8m	0.2m	22.10m	15%
BFGS	249.7m	0.3m	41.70m	60%
'Interior-Point'	249.8m	0.2m	22.94m	12%

Table 11.12: The estimated distance and error of three optimization methods in estimating the distance between the fibre and the event when the source starts at $t = 0.1$ seconds. The data has a 10m gauge length applied to it.

the data did not have a gauge length applied. Once again, the error of the 'interior-point' algorithm decrease from the error in the case when the events start at $t = 0$ seconds for the data with a 10m gauge length applied. The BFGS's consistently high error suggests that it is not the best method fitting the data in order to use the hyperbolic function to judge the difference between the fibre sensor and the event.

One approach which may help with fitting the data is making the distinction between the P-wave and S-wave more evident before we create the mask to find the points of the P-wave and S-wave hyperbolae. In the next two examples, we show how applying a deconvolution to the data helps achieve this outcome.

11.3.3 Deconvolved data without a gauge length applied

A distinct separation of the P-wave and S-wave in the hyperbolic response of the event would be ideal for determining the points of the P-wave and the S-wave hyperbolae. As we saw for both cases when the S-wave response did not fit the full path of the fibre, our current method for distinguishing the P-wave and S-wave hyperbolas allows for some errors to occur when finding the S-wave hyperbola. While employing a larger threshold when creating the mask may work, it would also give us fewer points of the hyperbola and likely cause the peak of the hyperbolic response to become less visible. Given that our current threshold has difficulty determining the peak of the hyperbola response, choosing a larger threshold may increase the error of the Hyperbola Method for determining the distance between the event and the fibre-optic cable. To address this issue, we apply a deconvolution to each offset of the data to highlight the peaks of the P-wave and S-wave responses in each offset.

Figure 11.14 depicts the event when the source starts at $t = 0$ seconds next to the deconvolved data of the event. We employ the `deconf` function from the CREWES Toolbox to deconvolve each offset of the dataset in time (Margrave, 2018). In Figure 11.14 (right), there appears to be less overlap of the P-wave and S-wave as opposed to the data which is not deconvolved (left).

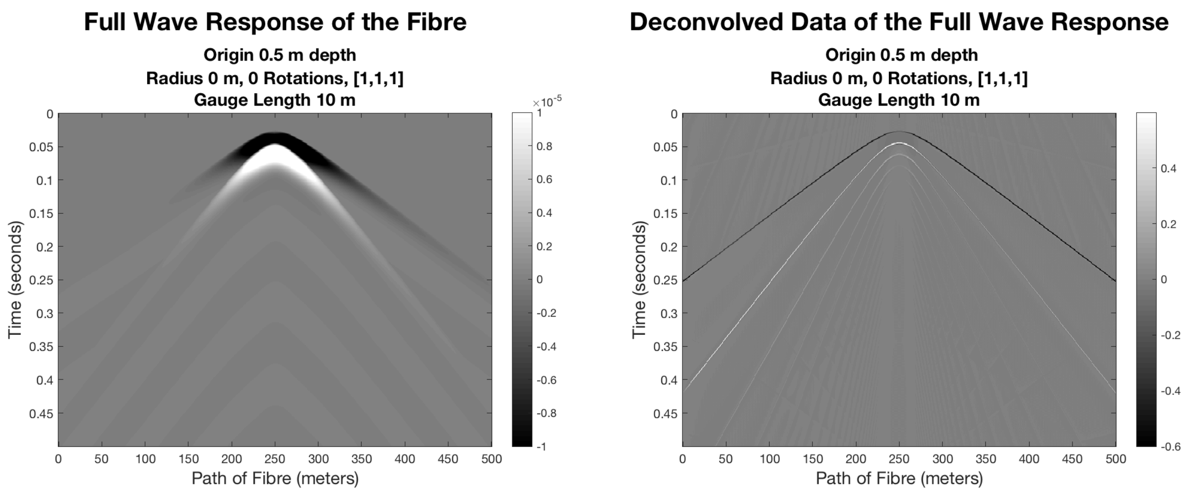


Figure 11.14: (Left) The data of the event when the source starts at $t = 0$ seconds. (Right) The deconvolved data of the event when the source starts at $t = 0$ seconds.

Taking the derivative of the deconvolved data in time, we set a threshold of 0.1 to create a mask. We set the threshold this high because the deconvolution increases the P-wave and S-wave amplitudes at the first time arrivals while decreasing their response everywhere else. Figure 11.15 presents the deconvolved data on the left and the mask of the deconvolved data on the right. In Figure 11.15, the P-wave and S-wave responses are separable unlike in the previous examples. We highlight the boundaries of the P-wave and S-wave response in the mask using the `bwboundaries` function from MATLAB as we did with the previous cases. Given that we are choosing the initial points of our P-wave and S-wave hyperbolae based on the boundary of the mask, we must chose the threshold carefully in order to guarantee that the top of the boundary in each case is connected. We apply a bridge morphological cleaning to the mask before finding the boundary to ensure that the P-wave and S-wave responses are connected in the mask.

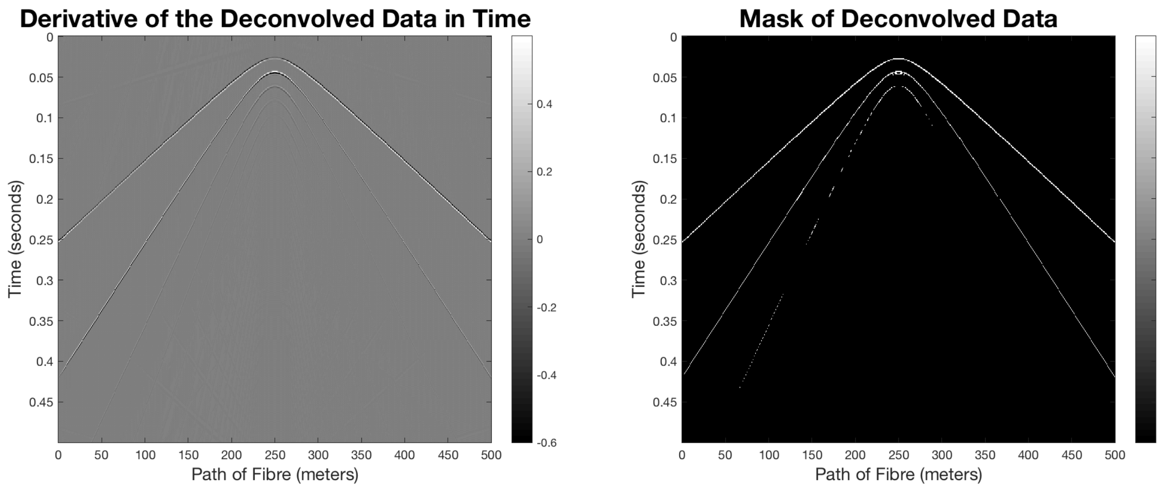


Figure 11.15: (Left) The time-derivative of the deconvolved data when the event begins at $t = 0$ seconds.
(Right) The mask of the time-derivative of the deconvolved data using a threshold of 0.1.

The top hyperbolic shape in the mask is the P-wave as it is the faster wave in the response; dually, the bottom hyperbolic shape is the S-wave as it moves slower than the P-wave. Using the boundaries of the two hyperbolic shapes in the mask, we separate the P-wave and S-wave responses and calculate the points of the P-wave and S-wave hyperbolae based off the points at the top of each boundary respectively. Figure 11.16 shows the results of separating the P-wave and S-wave responses and the resulting points of the hyperbolae.

The hyperbola points for the P-wave and S-wave both have smoother peaks than the previous cases when we did not deconvolve the data. Let us see how well the different optimization methods perform with the points from this new technique. We employ the same methods for estimating the initial input for the P-wave and S-wave coefficients as we did in Section 11.3.1. Moreover, we use the upper and lower bounds from Equations 11.12 to 11.19 for the 'interior-point' algorithm in this example as well.

In Figure 11.17, we observe how well the hyperbola functions defined by the coefficients

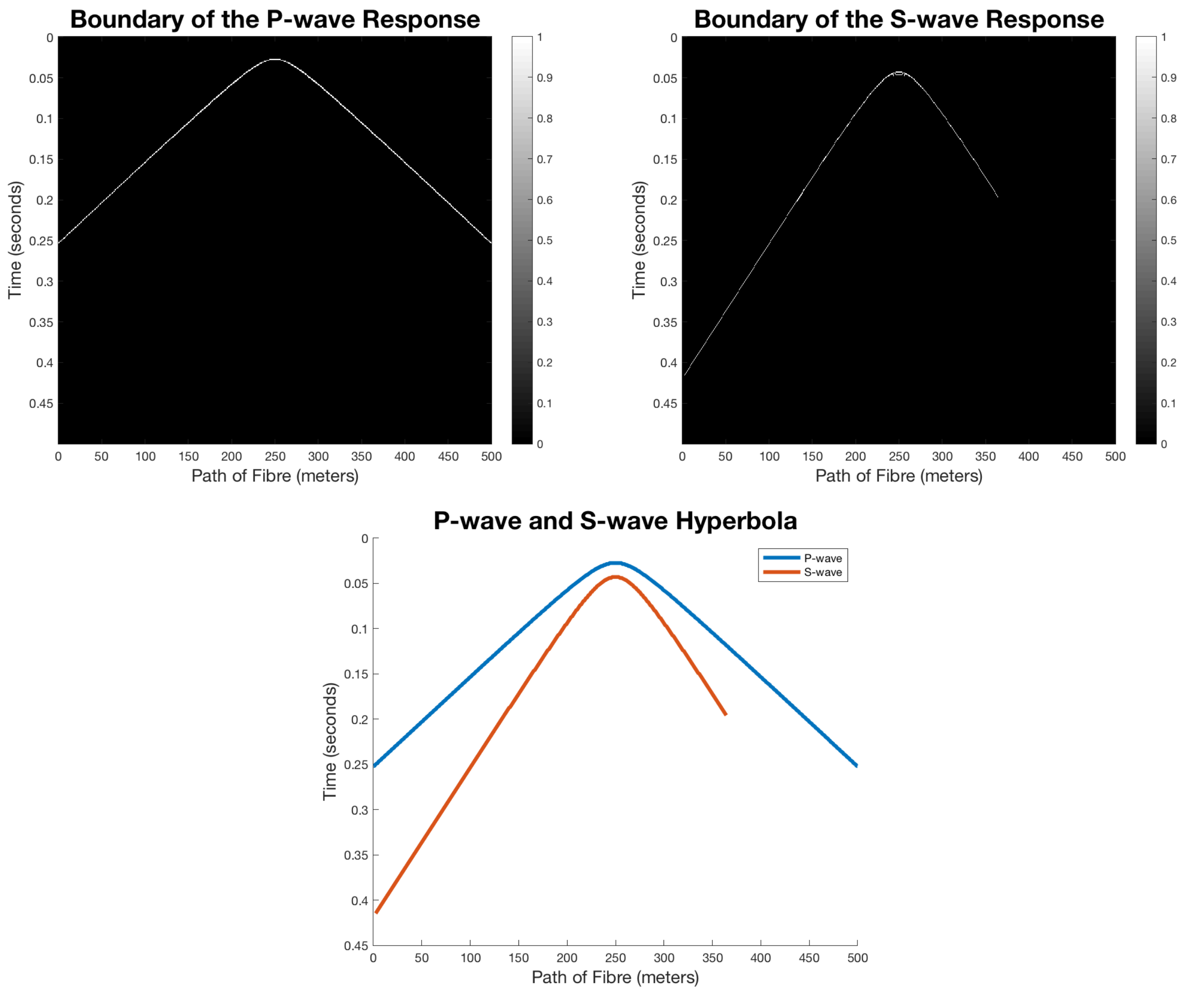


Figure 11.16: (Top left) The boundary of the P-wave from the mask.
(Top right) The boundary of the S-wave from the mask.
(Bottom) The P-wave hyperbola (blue) and the S-wave hyperbola (red) generated from the boundary of the P-wave and S-wave from the mask.

	P-wave Coefficients
$\mathbf{V_p}$	$[s_0, t_0, m_p, a_p]$
Input	$[249.7, 0, -1.00 \times 10^{-3}, 28.7]$
Simplex Search	$[250.0, 1.20 \times 10^{-3}, -1.00 \times 10^{-3}, 26.1]$
BFGS	$[249.8, 0, -1.00 \times 10^{-3}, 28.7]$
‘Interior-point’	$[250.0, 1.30 \times 10^{-3}, -1.00 \times 10^{-3}, 26.0]$

Table 11.13: The optimized P-wave coefficients of the hyperbolae functions describing the P-wave respectively for each optimization method. These hyperbola coefficients were generated from the deconvolved data when the source starts at $t = 0$ seconds.

	S-wave Coefficients
$\mathbf{V_s}$	$[s_0, t_0, m_s, a_s]$
Input	$[249.7, 0, -1.60 \times 10^{-3}, 26.0]$
Simplex Search	$[250.0, 0, -1.70 \times 10^{-3}, 25.8]$
BFGS	$[249.8, 0, -1.70 \times 10^{-3}, 26.0]$
'Interior-point'	$[250.0, 2.00 \times 10^{-4}, -1.70 \times 10^{-3}, 25.6]$

Table 11.14: The optimized S-wave coefficients of the hyperbolae functions describing the S-wave respectively for each optimization method. These hyperbola coefficients were generated from the deconvolved data when the source starts at $t = 0$ seconds.

in Tables 11.13 and 11.14 fit to the hyperbolic response in the data. Visually, the P-wave response has a better fit for all three optimization methods as well as a better initial estimate (especially with regards to the peak of the response). There also appears to be less of a delay in the fit of the S-wave response from all three methods and the initial input, unlike with the results generated from the data without a deconvolution applied to it.

Table 11.15 uses the hyperbolic functions created by the coefficients in Tables 11.13 and 11.14 and Equation 11.1 to find the distance between the event and the fibre. Every method performed better with regards to judging the distance between the fibre sensor and the event location. Including the initial coefficient estimate, every method was within approximately 4m of the actual distance. Applying the deconvolution to the data before solving for the points caused a significant drop in error for each method. While all the methods seem to provide a good fit to the data in Figure 11.17, the simplex search method provides the best fit with an error of approximately 9%; however every other error remains at less than 16%.

	Distance along fibre s_0	Offset	Distance from Fibre $D_{\mathbf{fib}}$	Error
Event Location	250m		26m	
Input	249.7m	0.3m	22.27m	14%
Simplex Search	250.0m	0.0m	23.65m	9%
BFGS	249.8m	0.2m	22.05m	15%
'Interior-Point'	250.0m	0.0m	23.47m	10%

Table 11.15: The estimated distance and error of three optimization methods in estimating the distance between the fibre and the event when the data is deconvolved in time and the event starts at $t = 0$ seconds.

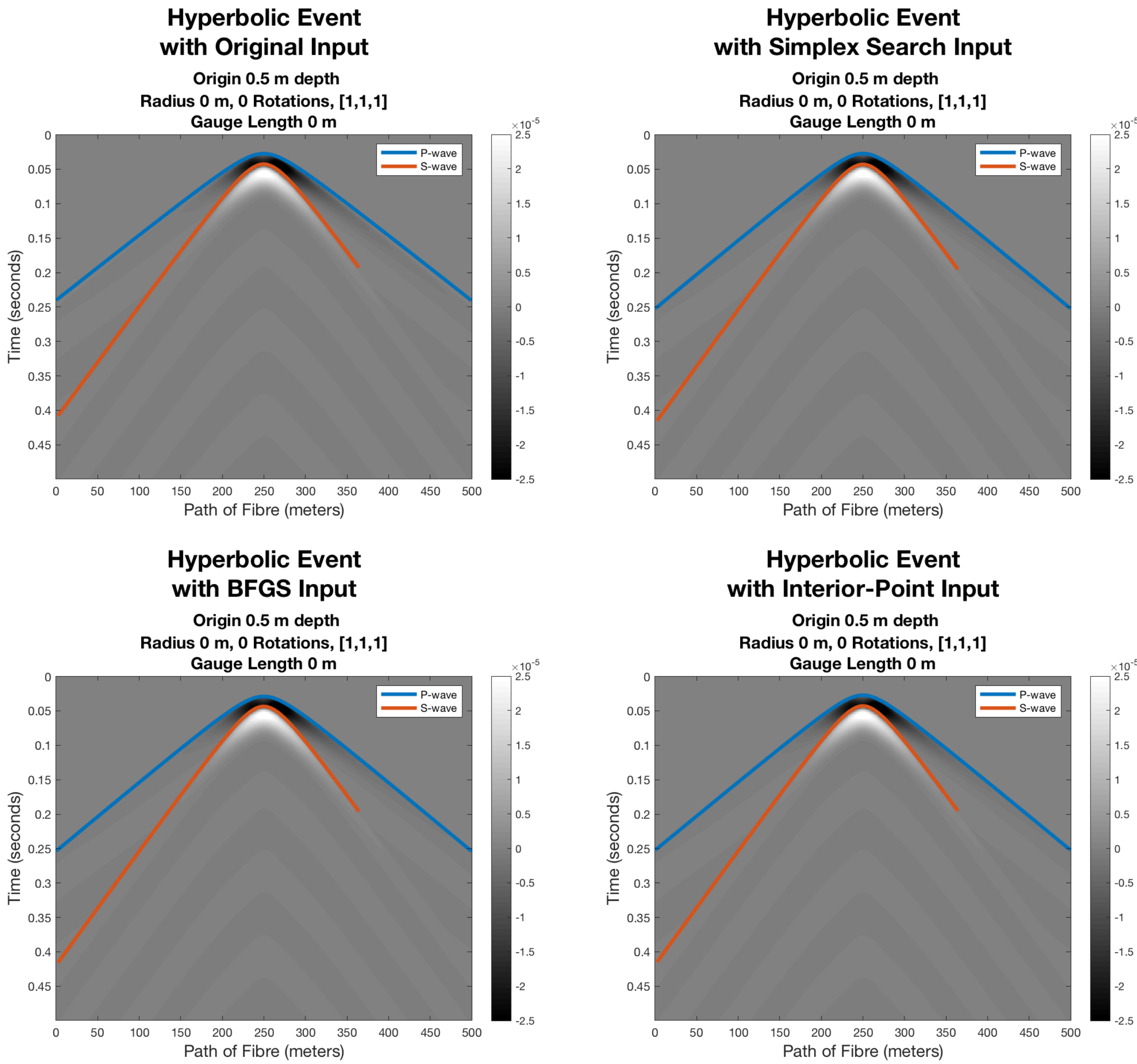


Figure 11.17: A hyperbola fitted to the P-wave (blue) and S-wave (red) response of the event which starts at $t = 0$ seconds and the points of the hyperbolae estimated from the deconvolved data. The hyperbola fitting for (top left) the original input vector, (top right) the optimized vector from the simplex search method, (bottom left) the optimized vector from the BFGS method, and (bottom right) the optimized vector from the 'interior-point' algorithm.

While our initial estimate for the P-wave and S-wave coefficients had an error of less than 15% for calculating the distance between the fibre and the event in this case, we should consider what occurs when the initial input for the coefficients is not well chosen. As such, we deconvolve the data without a gauge length applied when the event starts at $t = 0.1$ seconds. As with the previous time offsets cases, we choose $t_0 = 0$ for our initial estimate of the P-wave and S-wave coefficients. Figure 11.18 shows the results of the data when the source starts at $t = 0.1$ seconds next to the deconvolved data of the event.

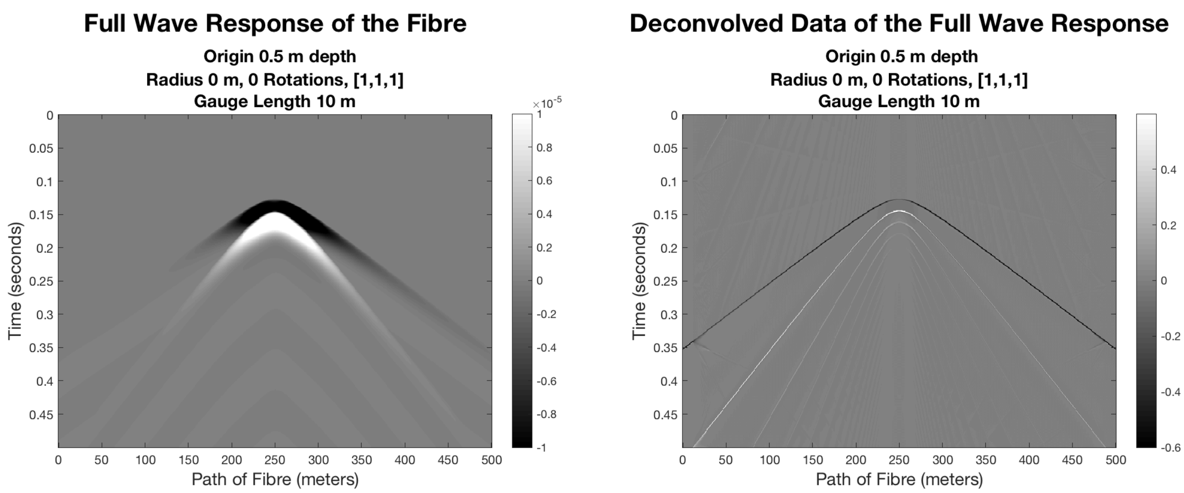


Figure 11.18: (Left) The event starts at $t = 0$ seconds.
(Right) The deconvolved data of the event which starts at $t = 0.1$ seconds.

Again, the deconvolution produces a clear separation of the P-wave and S-wave responses in the event. Taking the time-derivative of the deconvolved data, we set the threshold equal to 0.05 for this example. We ensure each component is connected by applying a bridge morphological cleaning to the mask. Figure 11.19 provides an image of the resulting mask.

We now separate the P-wave and the S-wave responses from the mask of the deconvolved data using the `bwboundaries` function as we have in previous cases. Figure 11.20 displays these results and the P-wave and S-wave hyperbola points determined by the top of the P-wave and S-wave boundaries, respectively.

With these points, we make an initial estimate of the P-wave hyperbola and S-wave

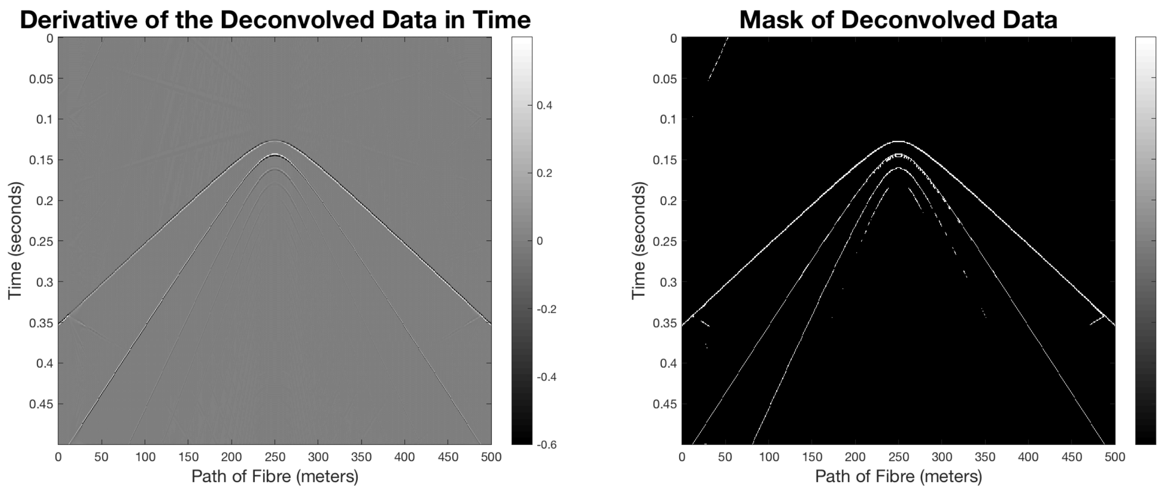


Figure 11.19: (Left) The time-derivative of the deconvolved data when the event begins at $t = 0.1$ seconds.
(Right) The mask of the time-derivative of the deconvolved data using a threshold of 0.05.

hyperbola coefficients using the same method as the previous cases. Again, we set $t_0 = 0$ for both the P-wave and S-wave input as the window around an event in real-data may not allow for a good estimate of the vertex t_0 of each hyperbola. Tables 11.16 and 11.17 provides the coefficient results of each optimization method.

	P-wave Coefficients
$\mathbf{V_p}$	$[s_0, t_0, m_p, a_p]$
Input	$[249.7, 0, -9.00 \times 10^{-4}, 134.7]$
Simplex Search	$[250.0, 0.1011, -1.00 \times 10^{-3}, 26.2]$
BFGS	$[249.7, 0, -1.50 \times 10^{-3}, 134.7]$
‘Interior-point’	$[250.0, 0.1005, -1.00 \times 10^{-3}, 27.0]$

Table 11.16: The optimized P-wave coefficients of the hyperbolae functions describing the P-wave respectively for each optimization method. These hyperbola coefficients were generated from the deconvolved data when the source starts at $t = 0.1$ seconds.

As with the previous cases where the time was offset by 0.1 seconds, the BFGS method had difficulty defining a_p for the P-wave hyperbola coefficient. It also struggled with the slowness of the P-wave m_p. We see the poor results of the BFGS method reflected in its fitting to the hyperbolic response of the event in Figure 11.21.

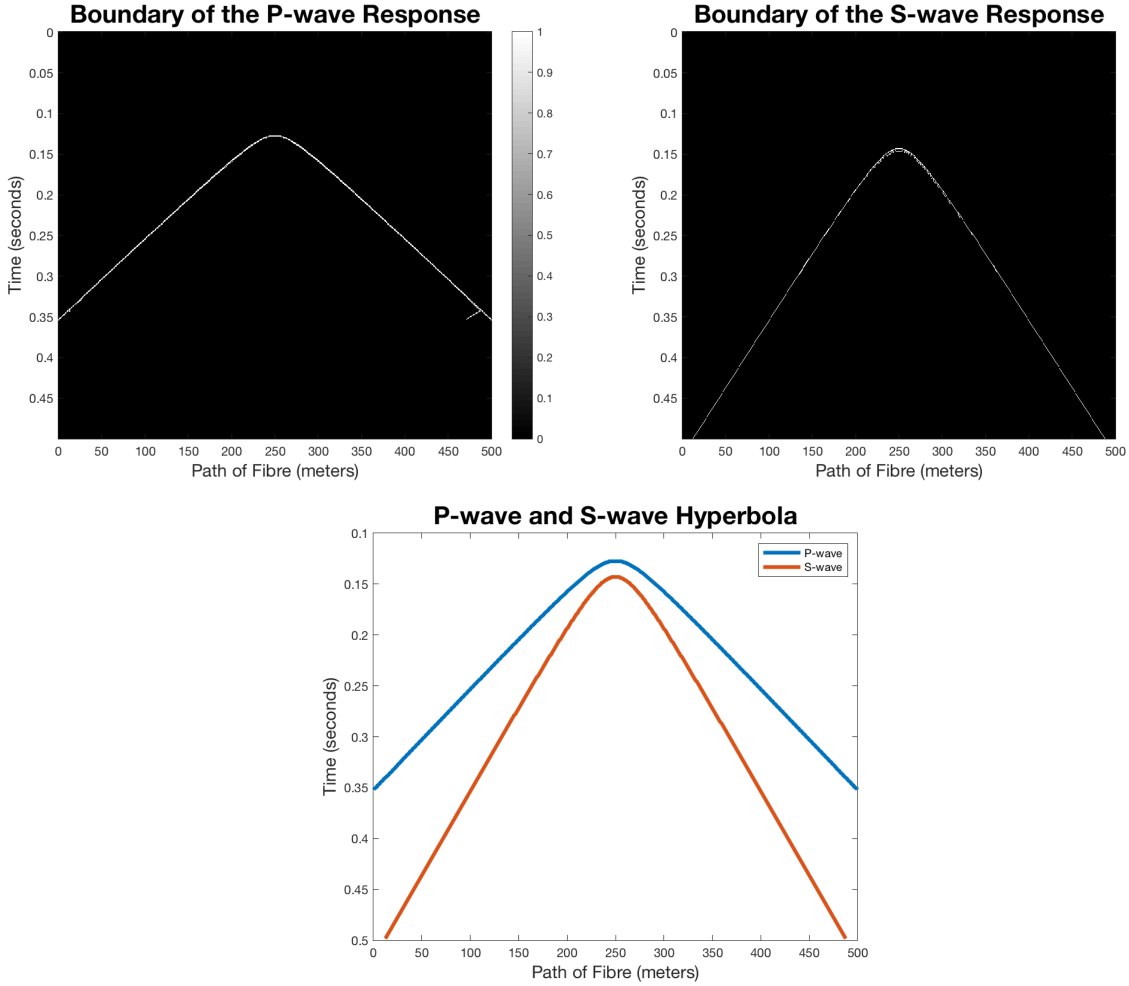


Figure 11.20: (Top left) The boundary of the P-wave from the mask.
(Top right) The boundary of the S-wave from the mask.
(Bottom) The P-wave hyperbola (blue) and the S-wave hyperbola (red) generated from the boundary of the P-wave and S-wave from the mask.

	S-wave Coefficients
$\mathbf{V_s}$	$[s_0, t_0, m_s, a_s]$
Input	$[249.7, 0, -1.60 \times 10^{-3}, 90.0]$
Simplex Search	$[250.0, 0.0996, -1.70 \times 10^{-3}, 26.0]$
BFGS	$[249.7, 0, -2.10 \times 10^{-3}, 90.0]$
'Interior-point'	$[250.0.0992, -1.70 \times 10^{-3}, 26.3]$

Table 11.17: The optimized S-wave coefficients of the hyperbolae functions describing the S-wave respectively for each optimization method. These hyperbola coefficients were generated from the deconvolved data when the source starts at $t = 0.1$ seconds.

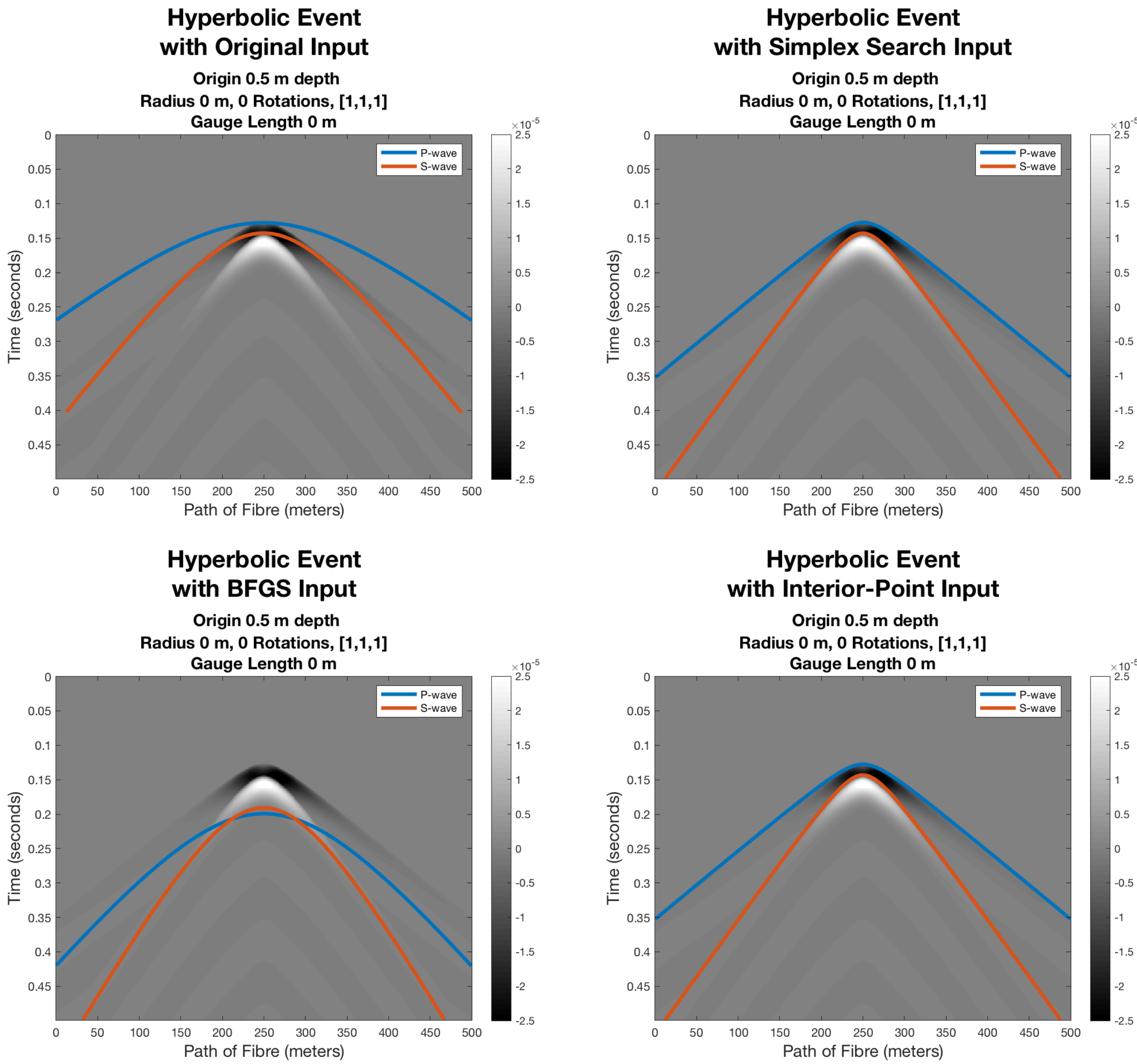


Figure 11.21: A hyperbola fitted to the P-wave (blue) and S-wave (red) response of the event which starts at $t = 0.1$ seconds and the points of the hyperbolas were estimated from the deconvolved data. The hyperbola fitting for (top left) the original input vector, (top right) the optimized vector from the simplex search method, (bottom left) the optimized vector from the BFGS method, and (bottom right) the optimized vector from the 'interior-point' algorithm.

From Figure 11.21, the simplex search method and the 'interior-point' algorithm prove the best for optimizing the coefficients of the P-wave and S-wave hyperbola to best fit the hyperbolic response. In Table 11.18, we observe that the simplex search method has the lowest error for determining the distance between the fibre and the event; however, the error for the initial estimate and the 'interior-point' algorithm are within 1% of the error for the simplex search method.

	Distance along fibre s_0	Offset	Distance from Fibre $D_{\mathbf{fib}}$	Error
Event Location	250m		26m	
Input	249.7m	0.3m	23.49m	10%
Simplex Search	250.0m	0.0m	23.56m	9%
BFGS	249.7m	0.3m	13.42m	48%
'Interior-Point'	250.0m	0.0m	23.40m	10%

Table 11.18: The estimated distance and error of three optimization methods in estimating the distance between the fibre and the event when the data is deconvolved in time and the event starts at $t = 0.1$ seconds.

While the input estimate does not provide a good choice of t_0 for both the P-wave and the S-wave as well as a_p and a_s, it does preserve the distance between the P-wave and the S-wave as well as estimate the slowness of the P-wave and S-wave correctly. Despite the poor fit to the data, it was thus able to provide a good estimate of the distance between the fibre sensor and the event. The simplex search method provides the best optimized coefficients for the P-wave and S-wave hyperbola function which follows from how well the respective hyperbolae fit the data. As such, it produces the best estimate of the distance between the fibre-optic cable and the event out of all three optimization methods at less than 3m from the actual distance. Despite fitting the S-wave response well, the BFGS method was less successful with the P-wave response. As such, its high error with respect to the other methods is likely due to the fact that its hyperbolic functions do not maintain the distance between the P-wave and S-wave responses as well as the fact that its estimate for the P-wave velocity is more than 300m/s less than the actual velocity of the P-wave in the model. We do see a decrease in error for the BFGS method in this case compared to the previous two examples where the source was offset by 0.1s.

11.3.4 Deconvolved data with a 10m gauge length applied

Now let us consider how well the Hyperbola Method performs when we deconvolve the data with a gauge length applied. We use the 10m gauge length for comparison to the case described in Section 11.3.2. The first case we study is when the source starts at $t = 0$ seconds and examine the case when the source starts at $t = 0.1$ seconds afterwards.

We begin by deconvolving each offset of the data with the 10m gauge length applied. Figure 11.22 shows the data with the 10m gauge length applied next to the deconvolved data. As with the previous example, the deconvolution makes it easier to distinguish the P-wave response from the S-wave response.

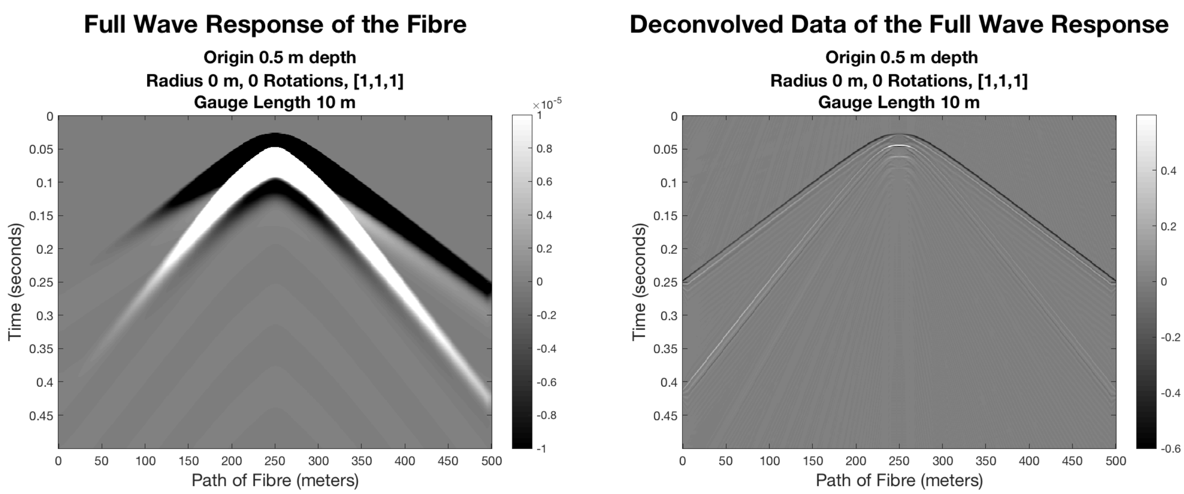


Figure 11.22: (Left) The event with a 10m gauge length applied to the data starts at $t = 0$ seconds.
(Right) The deconvolved event of the data with a 10m gauge length applied starts at $t = 0$ seconds.

Now, we create a mask of the deconvolved data using the threshold 0.08 on the time-derivative of the deconvolved data; see Figure 11.23. The image of the mask is found on the right in Figure 11.23.

The effect of the gauge length on the data is evident in the the mask of the time-derivative

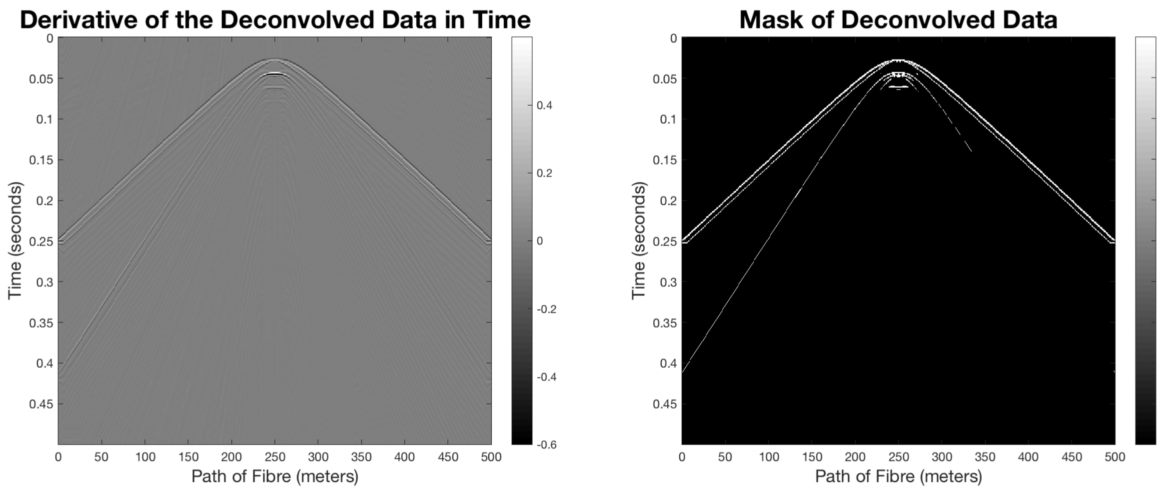


Figure 11.23: (Left) The time-derivative of the deconvolved data when the event begins at $t = 0$ seconds.
(Right) The mask of the time-derivative of the deconvolved data using a threshold of 0.05.

of the deconvolved data. Instead of one clear outline of the P-wave and S-wave response, we see a second hyperbolic outline underneath each response. This highlights the flattening of the hyperbolic response that we saw in Chapter 6.

Despite the echo that the gauge length causes to appear in the mask, we perform a morphological cleaning on the data to bridge any gaps in order to avoid using the wrong hyperbolic set of points. Once again, we employ `bwboundaries` to separate the P-wave and S-wave response, taking the top hyperbolic shape for the P-wave and the bottom hyperbolic shape for the S-wave. We present the P-wave and S-wave response from the mask in the top row of Figure 11.24.

Figure 11.24 (bottom) displays the result of defining the points of the P-wave and S-wave hyperbola by the top points of the boundary for the P-wave and S-wave masks respectively. As with the previous deconvolution example, the peak of both hyperbolae is smooth in comparison to the examples where we did not deconvolve the data. It appears that the gauge length has an effect on the S-wave response as the S-wave is missing the portion from about 300m to 500m. This does not affect our method as the S-wave peak is still evident and we estimate the slope of the hyperbola using the left side.

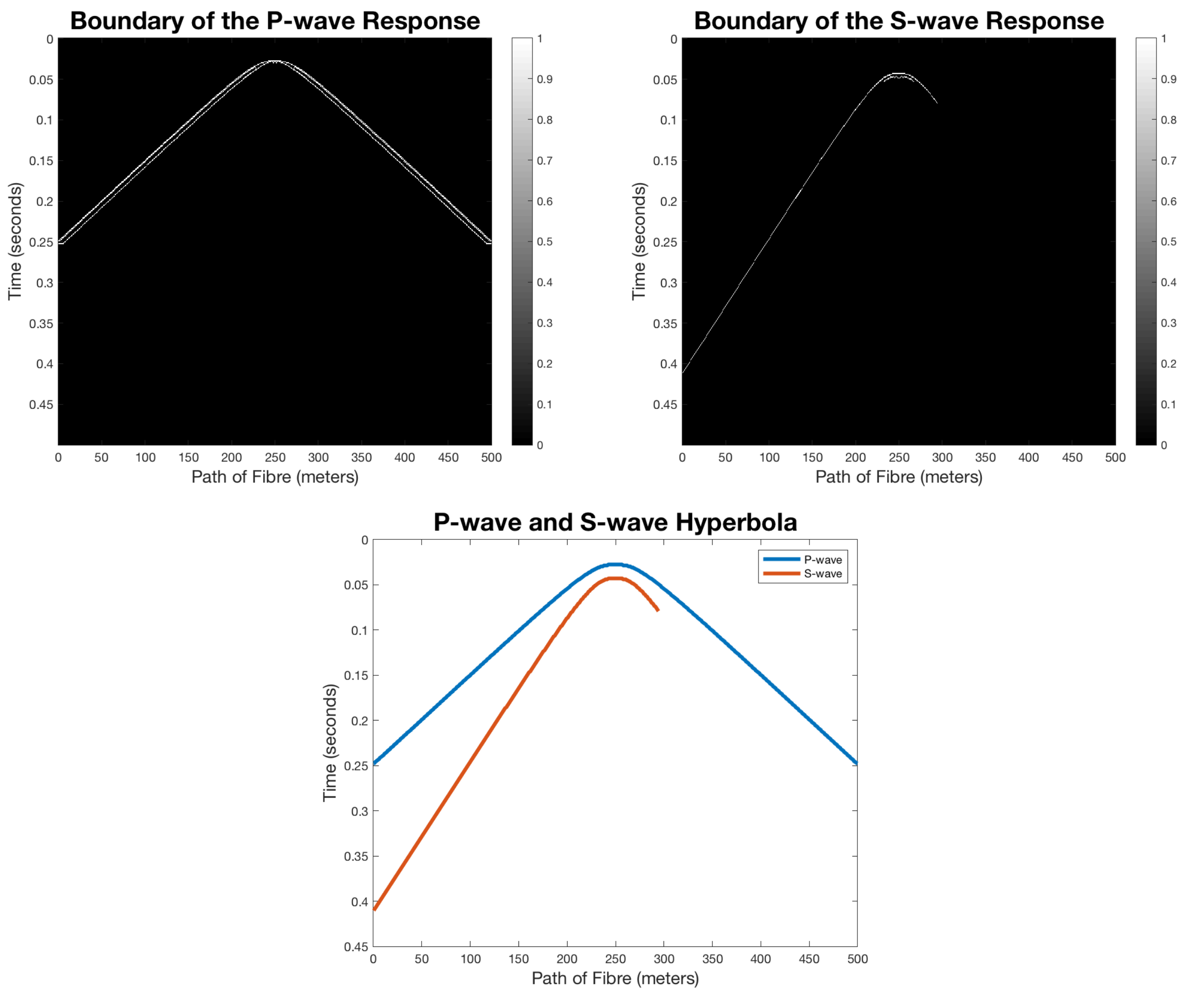


Figure 11.24: (Top left) The boundary of the P-wave from the mask.
(Top right) The boundary of the S-wave from the mask.
(Bottom) The P-wave hyperbola (blue) and the S-wave hyperbola (red) generated from the boundary of the P-wave and S-wave from the mask.

We choose the initial value of the P-wave and S-wave hyperbola coefficients the same way we chose them for the last few examples. Tables 11.19 and 11.20 shows the optimized output coefficients for each method next to the initial estimates.

	P-wave Coefficients
$\mathbf{V_p}$	$[s_0, t_0, m_p, a_p]$
Input	$[249.7, 0, -9.00 \times 10^{-4}, 28.7]$
Simplex Search	$[250.1, 0, -1.00 \times 10^{-3}, 30.5]$
BFGS	$[249.5, 0, -1.00 \times 10^{-3}, 28.8]$
'Interior-point'	$[249.3, 6.40 \times 10^{-3}, -1.00 \times 10^{-3}, 21.2]$

Table 11.19: The optimized P-wave coefficients of the hyperbolae functions describing the P-wave respectively for each optimization method. These hyperbola coefficients were generated from the deconvolved data with a 10m gauge length applied when the source starts at $t = 0$ seconds.

	S-wave Coefficients
$\mathbf{V_s}$	$[s_0, t_0, m_s, a_s]$
Input	$[249.7, 0, -1.60 \times 10^{-3}, 26.0]$
Simplex Search	$[250.1, 0, -1.70 \times 10^{-3}, 32.2]$
BFGS	$[249.5, 0, -1.70 \times 10^{-3}, 26.9]$
'Interior-point'	$[249.3, 1.10 \times 10^{-3}, -1.60 \times 10^{-3}, 22.6]$

Table 11.20: The optimized S-wave coefficients of the hyperbolae functions describing the S-wave respectively for each optimization method. These hyperbola coefficients were generated from the deconvolved data with a 10m gauge length applied when the source starts at $t = 0$ seconds.

Figure 11.25 provides a glimpse of how well the hyperbolic functions defined by the coefficients in Tables 11.19 and 11.20 fit the hyperbolic response of the event.

Visually, each method, including the initial estimate of the coefficients, produced coefficients which fit the data well for both the P-wave and S-wave response. Table 11.21 gives a numerical idea of how well each set of coefficients worked for fitting the hyperbolic shape of the data. Our initial estimate for the P-wave and S-wave coefficients has an error below 15%. It also shows the distance between the source and the fibre to be less than 4 meterS. As with the other deconvolution case when the source starts at $t = 0$ seconds, the

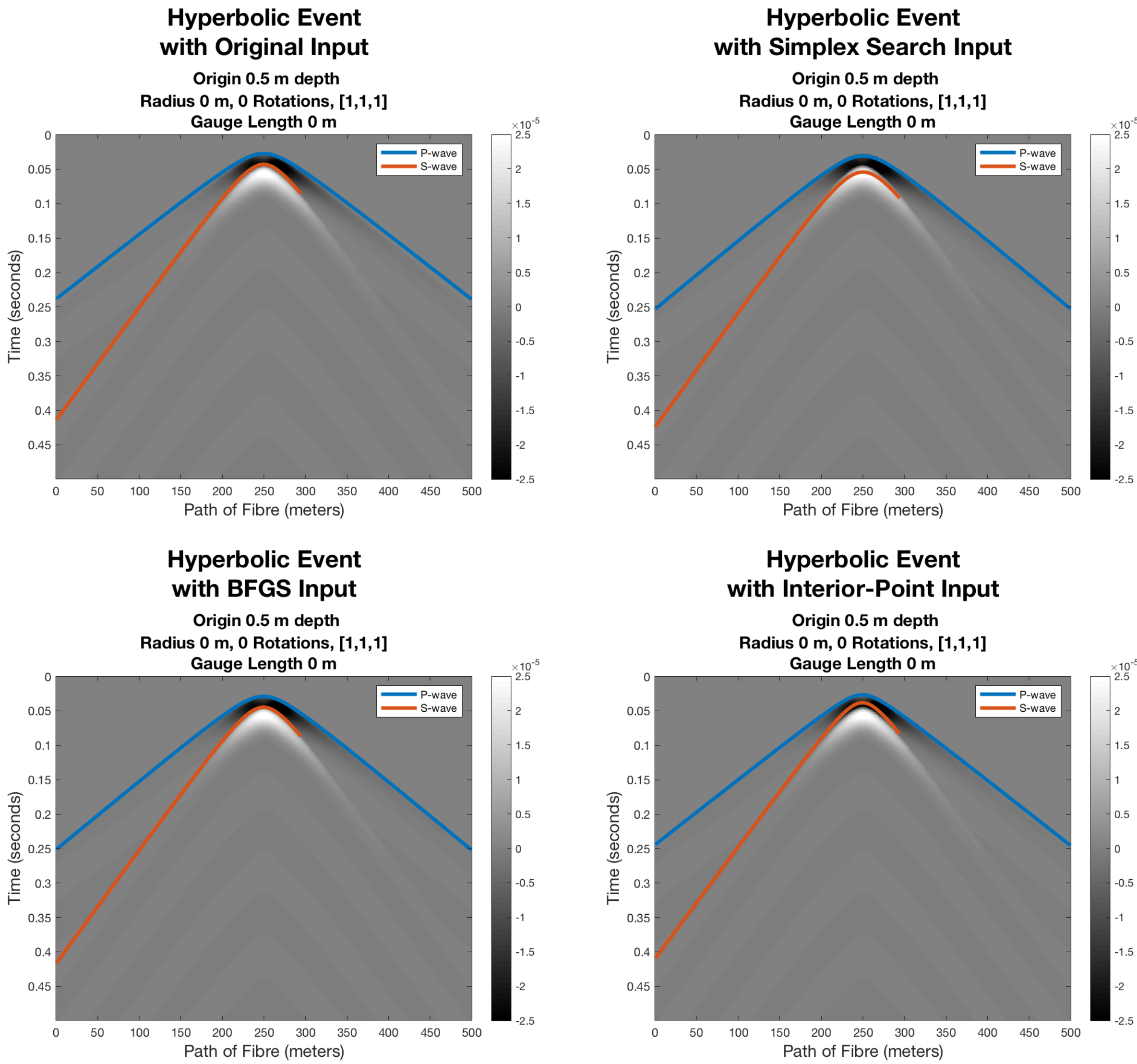


Figure 11.25: A hyperbola fitted to the P-wave (blue) and S-wave (red) response of the event, with a 10m gauge length applied to the data, which starts at $t = 0$ seconds. The points of the hyperbolae were estimated from the deconvolved data. The hyperbola fitting for (top left) the original input vector, (top right) the optimized vector from the simplex search method, (bottom left) the optimized vector from the BFGS method, and (bottom right) the optimized vector from the 'interior-point' algorithm.

BFGS method also performed well. Interestingly, in comparison to the example where we did notdeconvolve the data with a 10m gauge length applied, the only errors that decreased are the input estimate and the coefficients optimized by the BFGS method. The other two methods saw an increase in the error by more than double for both cases. This suggest that the echo of the hyperbolic response that we saw in the mask of the time-derivative of the deconvolved data has a greater effect on the shape of the signal than we originally thought. The increase in error for the simplex search method and the 'interior-point' algorithm may also be attributed to the fewer points generated from the S-wave response.

	Distance along fibre s_0	Offset	Distance from Fibre $D_{\mathbf{fib}}$	Error
Event Location	250m		26m	
Input	249.7m	0.3m	22.40m	14%
Simplex Search	250.0m	0.0m	34.55m	33%
BFGS	249.5m	0.5m	24.03m	8%
'Interior-Point'	249.3m	0.7m	16.87m	35%

Table 11.21: The estimated distance and error of three optimization methods in estimating the distance between the fibre and the event when the data is deconvolved in time and the event starts at $t = 0$ seconds.

For the final example in this chapter, we want to see how well the optimization methods perform when the initial value is not chosen well for the deconvolved data with a 10m gauge length applied. We now deconvolve the data when a gauge length of 10m is applied to the data and the source starts at $t = 0.1$ seconds.

Figure 11.26 displays the results of applying the deconvolution to the offsets of the data. We then take the derivative in time of the data and create a mask using the threshold 0.08; c.f. Figure 11.27. Even though we apply a morphological cleaning to bridge the mask, there are some irregularities in the P-wave response between 10m to 30m and 470m to 490m along the path of the fibre in the mask. While the left side is completely disconnected, the right side is still connected; so, the points of the P-wave will be less hyperbolic for this example.

We separate the P-wave and S-wave responses as before using the `bwboundaries` function in Matlab. Figure 11.28 (bottom left) shows how the irregularities affect the hyperbolic

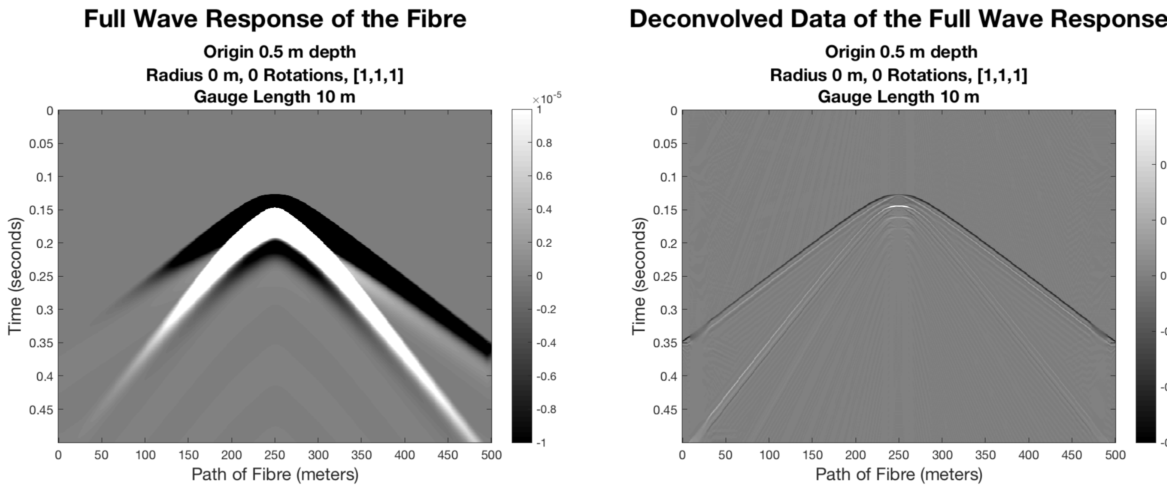


Figure 11.26: (Left) The event with a 10m gauge length applied to the data which starts at $t = 0.1$ seconds.
(Right) The deconvolved event of the data with a 10m gauge length applied which starts at $t = 0.1$ seconds.

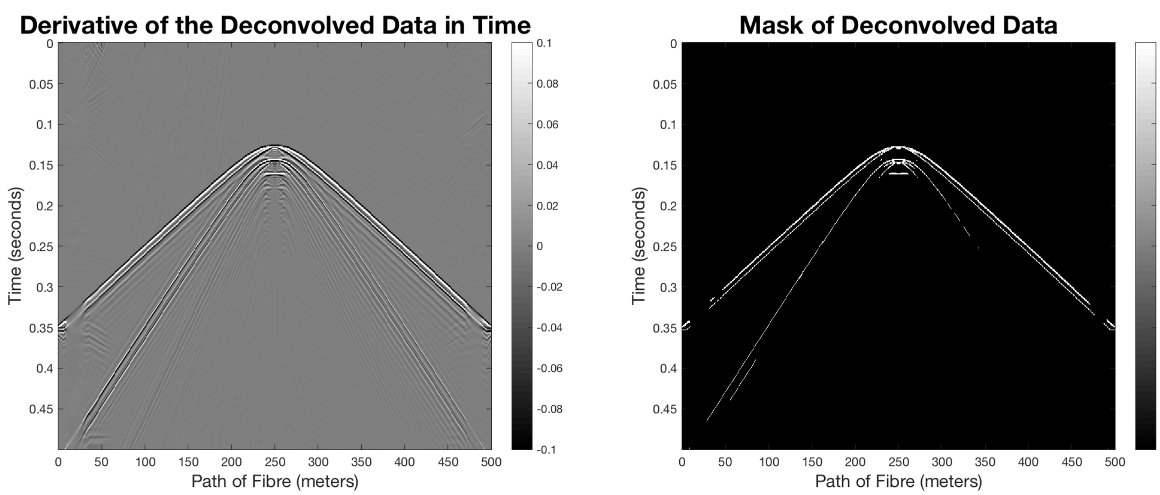


Figure 11.27: (Left) The time-derivative of the deconvolved data when the event begins at $t = 0.1$ seconds and a 10m gauge length has been applied to the data.
(Right) The mask of the time-derivative of the deconvolved data using a threshold of 0.08.

points that outline the P-wave and S-wave response in the event. This example requires a little more cleaning, which is provided on the bottom right of the figure. We remove these points by taking the derivative of the P-wave hyperbola in time, t_p, and only considering the portion of t_p between the minimum and maximum peaks of the derivative of t_p. The minimum peak highlights the sharp decrease in time from 0.34s to 0.33s at around 25m and the maximum peak highlights the sharp increase in time from 0.33s to 0.34 at around 475m. The improved P-wave hyperbola can be seen in Figure 11.28 (bottom right).

We now estimate the initial coefficients for the P-wave and S-wave hyperbolae based on the method we used in Section 11.3.1; also choose $t_0 = 0$ for both the initial P-wave and S-wave estimates. We define the upper and lower bounds for the 'interior-point' algorithm based on Equations 11.12 to 11.19. Tables 11.22 and 11.23 presents the results of the optimization methods applied to the initial estimate of the coefficients.

	P-wave Coefficients
$\mathbf{V_p}$	$[s_0, t_0, m_p, a_p]$
Input	$[249.7, 0, -9.00 \times 10^{-4}, 140.7]$
Simplex Search	$[250.1, 0.0949, -1.00 \times 10^{-3}, 31.4]$
BFGS	$[253.8, 0, -1.60 \times 10^{-3}, 140.7]$
'Interior-point'	$[250.6, 0.0939, -1.00 \times 10^{-3}, 32.5]$

Table 11.22: The optimized P-wave coefficients of the hyperbolae functions describing the P-wave respectively for each optimization method. These hyperbola coefficients were generated from the deconvolved data with a 10m gauge length applied when the source starts at $t = 0.1$ seconds.

From a glance, the initial estimates and the results of the BFGS method does the worst job of fitting the points to the hyperbolic response of the data; however, a consideration of Table 11.24 shows that the initial estimate appears to give a distance which is within 5m of the actual distance. The success of the initial estimate despite its poor fit to the data likely stems from the preservation of the distance between the P-wave and the S-wave, which the BFGS method loses. Once again, the BFGS method produces an increadibly high error. Its consistent high error for the case when the initial estimate is poorly chosen suggests

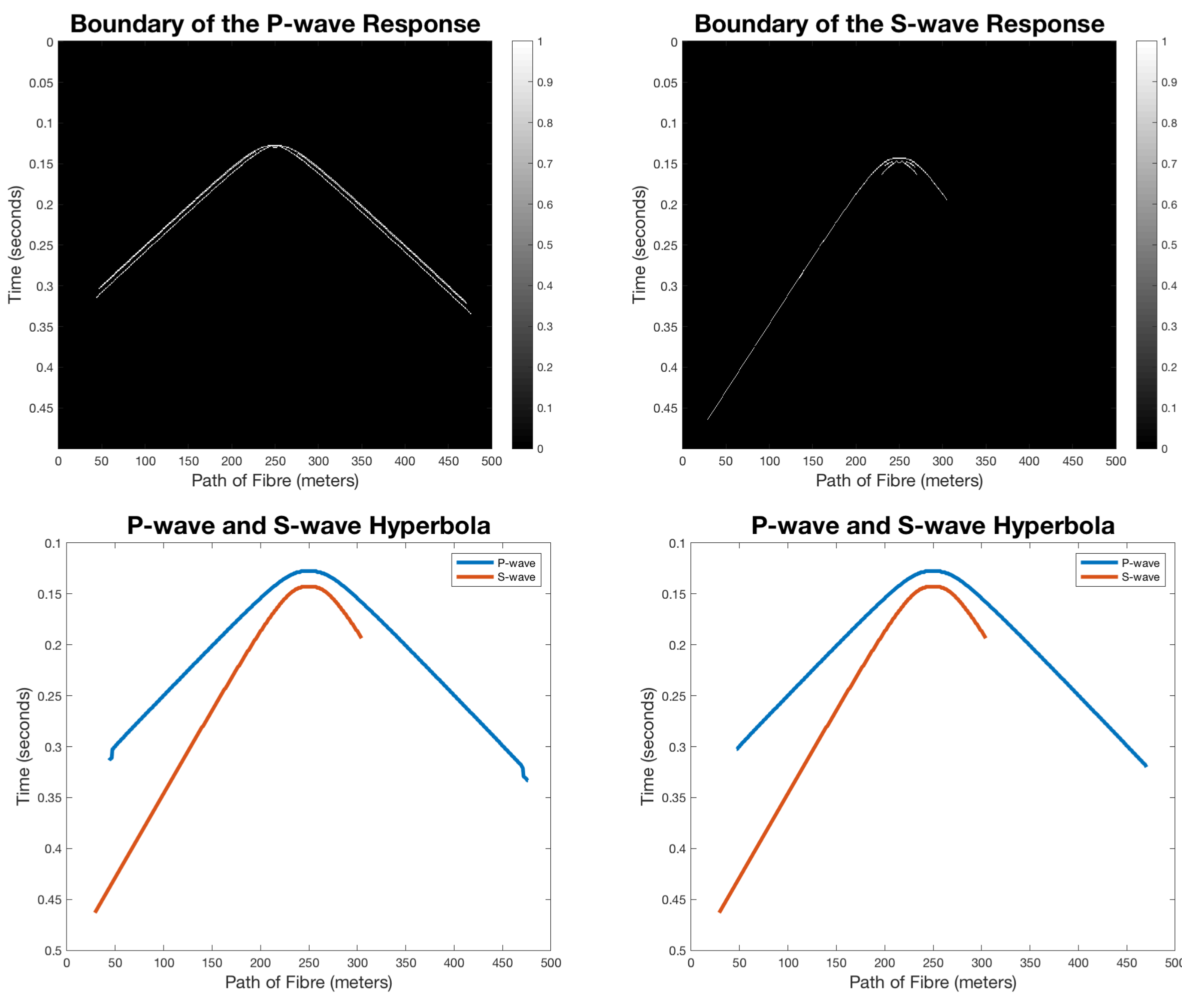


Figure 11.28: (Top left) The boundary of the P-wave from the mask.
(Top right) The boundary of the S-wave from the mask.
(Bottom left) The P-wave hyperbola (blue) and the S-wave hyperbola (red) generated from the boundary of the P-wave and S-wave in the mask.
(Bottom right) The P-wave hyperbola (blue) and the S-wave hyperbola (red) generated from the boundary of the P-wave and S-wave in the mask after removing the irregularities from the P-wave hyperbola.

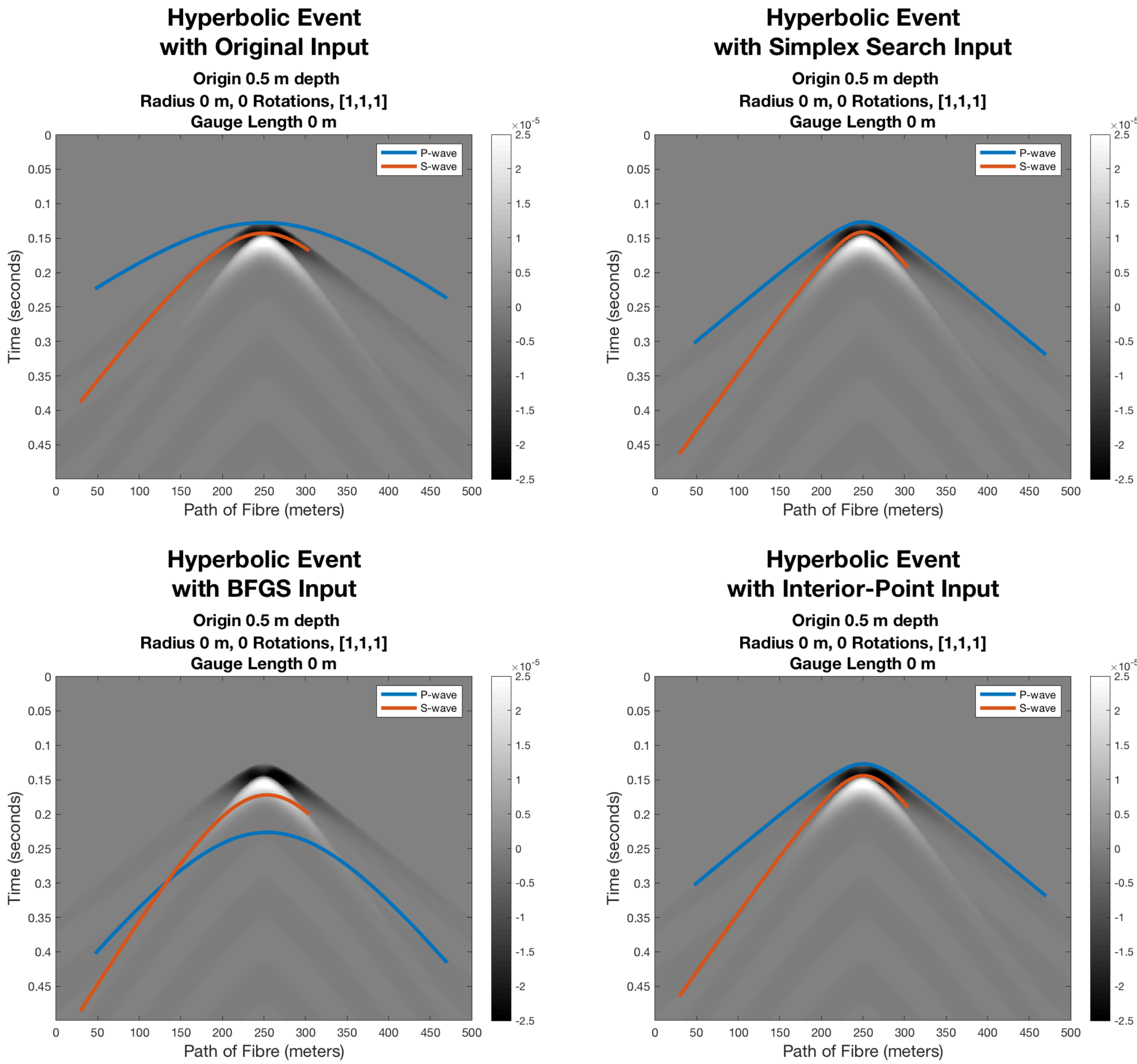


Figure 11.29: A hyperbola fitted to the P-wave (blue) and S-wave (red) response of the event, with a 10m gauge length applied to the data, which starts at $t = 0.1$ seconds. The points of the hyperbolae were estimated from the deconvolved data. The hyperbola fitting for (top left) the original input vector, (top right) the optimized vector from the simplex search method, (bottom left) the optimized vector from the BFGS method, and (bottom right) the optimized vector from the 'interior-point' algorithm.

	S-wave Coefficients
$\mathbf{V_s}$	$[s_0, t_0, m_s, a_s]$
Input	$[249.7, 0, -1.60 \times 10^{-3}, 87.3]$
Simplex Search	$[250.1, 0.0861, -1.70 \times 10^{-3}, 32.7]$
BFGS	$[253.8, 0, -2.00 \times 10^{-3}, 85.1]$
'Interior-point'	$[250.6, 0.0726, -1.70 \times 10^{-3}, 41.1]$

Table 11.23: The optimized S-wave coefficients of the hyperbolae functions describing the S-wave respectively for each optimization method. These hyperbola coefficients were generated from the deconvolved data with a 10m gauge length applied when the source starts at $t = 0.1$ seconds.

that is would not be a beneficial method to employ in future experiments. The simplex search method still provides a better fit to the data as well as a lower error than the initial estimate and the BFGS method. Only two methods show a decrease in error from having the data deconvolved before finding the points of the hyperbola: the initial estimate and the 'interior-point' algorithm. While the deconvolution only removes a small percent of the error for each method, we hypothesize that applying a deconvolution to real data before we use the Hyperbola Method will provide better results than if we do not apply a deconvolution before using the Hyperbola Method on real data.

	Distance along fibre s_0	Offset	Distance from Fibre $D_{\mathbf{fib}}$	Error
Event Location	250m		26m	
Input	249.7m	0.3m	21.05m	19%
Simplex Search	250.1m	0.1m	21.7m	17%
BFGS	253.8m	3.8m	131.4m	405%
'Interior-Point'	250.6m	0.6m	23.7m	9%

Table 11.24: The estimated distance and error of three optimization methods in estimating the distance between the fibre and the event when the data is deconvolved in time and when the event starts at $t = 0.1$ seconds.

11.4 Conclusions

In this chapter, we proposed a method for distinguishing the distance between events in DAS-acquired data and the fibre-optic sensor from the DAS sensor that acquired the data. We began the chapter by defining the Hyperbola Method, and we then described the models we created in order to investigate how well the Hyperbola Method worked on synthetic data. We created two sets of models where we applied a gauge length of 10m to one set and we refrained from applying a gauge length to the other set. In each set, one model had an event starting at $t = 0$ seconds and another model where the event started at $t = 0.1$ seconds. We also compared the results of three different optimization methods.

In the first four examples, we applied the Hyperbola Method to each model without any further processing. The methods performed adequately for isolating the distance between the fibre and the event with an error less than 30% in most cases. The models where the event was offset by $t = 0.1$ seconds generally saw an increase in error as we purposefully chose poor initial estimates for the P-wave and S-wave estimates (given that we will not necessarily be able to make a good estimate when we apply the Hyperbola Method to real data). The BFGS method consistently provided a high error for the data offset by $t = 0.1$s.

In order to decrease the error, we deconvolved each model in time and then applied the Hyperbola Method. The deconvolution improved the ability to separate the P-wave and S-wave. For most cases, the deconvolution reduced the error for determining the distance between the fibre-optic cable and the event in the DAS-acquired data in most cases. We especially saw improvement for the models where the time of the event began at $t = 0.1$ seconds for most every method except the BFGS method. A substantial decrease in the error for the initial estimate occurred after deconvolution for both sets of models and both time offsets for the source. This is promising with regards to application to real data as it shows that a poor choice of initial estimate will not cripple the results. Table 11.25 provides the errors for each model when the event started at $t = 0$ seconds. There was a significant difference between the error of the deconvolved cases and the cases which we refrained from deconvolving the data; however, the error did not decrease for the simplex search method and the 'interior-point' algorithm for the case when we deconvolved the data with a 10m gauge length applied. We expect that applying a deconvolution when considering real data would provide better results than if we did not. Similarly, Table 11.26 shows the errors for each model when the event started at $t = 0.1$ seconds. In this case, the deconvolution provided a significant improvement for most every method but the BFGS method did not perform as well.

	Original Input	Simplex Search	BFGS	'Interior-Point'
Model 1	28%	16%	33%	21%
Model 2	27%	14%	30%	14%
Deconvolved Model 1	14%	9%	15%	10%
Deconvolved Model 2	14%	33%	8%	35%

Table 11.25: Comparison of each optimization method's error in estimating the distance between the fibre and the event when the event starts at $t = 0$ seconds. Model 1 is the data without a gauge length applied. Model 2 is the data with a 10m gauge length applied.

Method	Original Input	Simplex Search	BFGS	'Interior-Point'
Model 1	32%	16%	2385%	6%
Model 2	29%	15%	60%	12%
Deconvolved Model 1	10%	9%	48%	10%
Deconvolved Model 2	19%	17%	405%	9%

Table 11.26: Comparison of each optimization method's error in estimating the distance between the fibre and the event when the event starts at $t = 0.1$ seconds. Model 1 is the data without a gauge length applied. Model 2 is the data with a 10m gauge length applied.

In each example, we also fit the resulting hyperbola functions from the different optimization methods to the data. Interestingly, the error seen in Tables 11.25 and 11.26 did not always reflect how well the P-wave and S-wave hyperbolae fit to the data. Given that the 'interior-point' algorithm achieved an error of less than 15% for all but two of the eight examples, the Hyperbola Method is a feasible approach for determining the distance between the event and the fibre.

For future work, we plan to apply the Hyperbola Method to real DAS-acquired data instead of synthetic data. In each example in this chapter, at least one method found a distance between the fibre and the even to be less than 5 meters. We are interested in seeing how well the Hyperbola Method fares in comparison to the synthetic data. The inclusion of noise and anomalies would potentially show that deconvolving the data would be essential for real data. Given that most of the optimization methods were often within 5m of the actual distance between the event and the fibre senor, it shows that the Hyperbola Method has substantial promise for application to real data.

We also plan to automate this method more than is currently included in this chapter. For the most part, the only variables which saw change between each example were the thresholds for the masks; however, in some cases, it was necessary to investigate methods for improving how the points of the P-wave and S-wave hyperbolae were chosen from the mask. Future work for the Hyperbola Method would include providing a more automated approach for determining if the points of the P-wave and S-wave hyperbolae were adequate for estimating the distance between the fibre sensor and an event.

Chapter 12

Conclusion

Throughout this thesis, we studied and developed methods for detecting and classifying events in data. Data acquired using distributed acoustic sensors became a major focus of our work given that DAS systems are used in various applications and produce large amounts of data. Detecting and classifying events in such data is essential to its applications. With this in mind, recall that we divided the thesis into three parts: seismic processing techniques, distributed acoustic sensing and how it works, and applications to DAS-acquired data.

In the first part, we considered several seismic processing techniques, such as wavelets and time-frequency analysis. We provided insight into a new wavelet transform: the inverted wavelet tree, which proved important for applications to neural networks in Chapter 10. We also focused on a vital component of events: reflections. In Chapter 2, we found exact solutions for reflection and transmission coefficients. These solutions were employed to create seismic models and compared to numerical results.

We saw in Chapter 5 the important role that reflections play in how a distributed acoustic sensor works. Reflections of the laser in the fibre-optic cable of a DAS system signify where an event occurs along the fibre. We produced an analytical model of DAS data which allows the user to move from straight to helically wound fibre easily. This ease enabled us to develop a homotopy to compare the straight and helically wound fibre; c.f. Theorems 7.13 and

7.14. Specifically, we used classical homotopy theory to deform a helically wound fibre into straight fibre and investigated the results of the data along that deformation. In Chapter 7, we employed the homotopy for two models: one comparing a horizontal straight and helically wound fibre and another comparing three types of vertical fibre which may be found in a well-borehole. Our comparison gave us significant insight into the relationship between the source and the fibre as well as the formation of the fibre. During our investigation in this chapter, we saw that the model of DAS data from Chapter 6 also allowed us to find bounds for the amplitude results of a DAS sensor, providing further insight into DAS systems; see Theorems 7.6 and 7.7.

With an understanding of how DAS systems work and motivation for its applications, we studied different methods for analyzing DAS data. In Chapter 8, the fibre-optic cable proved fruitful for detecting a future CO_2 storage location. Such detection supported the use of DAS in monitoring the site. In Chapters 9 and 10, we focused on several machine learning techniques as distributed acoustic sensors are capable of capturing large quantities of data. Gaussian mixture models proved beneficial for locating events in DAS-acquired data. Employing independent component analysis to separate the event signal from the noise provided increased precision with regards where the event occurs along the fibre sensor, but detects less of the signal. Utilizing convolutional neural networks allowed us to locate events in microseismic data, but also distinguish between walking events and digging events in monitoring data. It also raised the question of what features should be accounted for in image recognition. Our work suggests that we need to consider not only information about the source of the event such as its location or its velocity in the Earth's surface, but also information about how the distributed acoustic sensor collects the data such as the gauge length or pulse-repetition-frequency used when employing feature-based image recognition on DAS data.

Our final experiment of this thesis consisted of providing the distance between an event and the DAS sensor in synthetic data. We used the hyperbolic shape of events in DAS-acquired data, giving the method its name: the Hyperbola Method. In the method, the P-wave and S-wave responses of an event are determined. We used the points along these responses to optimize the hyperbola function describing the P-wave and S-wave responses. These functions provided information about the P-wave and S-wave velocities, which we employed with their respective first time-arrivals, to determine the distance between the event and the DAS sensor. In every model we studied, at least one optimization method estimated the distance to be within 5m of the actual distance.

Overall, we investigated and produced new methods for locating and identifying events in data with a focus on applying these techniques to distributed acoustic sensing acquired data. We accomplished these goals and more throughout this work. Not only did we apply detection and recognition methods applied to DAS-acquired data, but we also provided further insight into the expected amplitude responses of DAS systems as well as producing a means of comparing different formations of fibre-optic cable in a distributed acoustic sensor.

12.1 Future work

The work in this thesis answered many questions with regards to distributed acoustic sensing; however, it also produced many new questions which need to be answered. We include a few of those open questions below.

While we produced a model of data acquired using a distributed acoustic sensor, the next step is an analytic model of DAS-acquired data which shows reflections and transmissions. Our work on exact solutions for reflection and transmission coefficients in Chapter 2 provides a step in this direction, but an extension from plane wave reflection coefficients to spherical reflection coefficients is necessary.

In Chapter 7, we develop a homotopy which deforms helically wound fibre into straight fibre. We assumed that the helix kept its shape without the aid of any other materials. In reality, foam, aluminum, or some other material is essential for the helical fibre to maintain its helix. Investigating how the material maintaining the helix affects the deformation of the helical fibre into the straight fibre is of interest. In some instances, these materials could be considered a hole in the homotopy. A potential method for addressing this issue involves using fundamental groupoids to deal with the presence of a hole.

While independent component analysis showed potential for separating the vehicle's signal from the noise in Chapter 9, the Gaussian mixture model taught using a training set containing both the vehicle's signal and the noise statistically performed better than the vehicle detector trained on the independent component containing the vehicle in each of our examples. We only considered data which contained the vehicle's signal and noise. Given that independent component analysis provide good estimates of the vehicle's signal separated from the noise, it is worth investigating how well our methods from Chapter 9 work on data which contains more than one type of signal and noise. For data containing multiple vehicles or even vehicles and pedestrians, the independent component analysis would likely enable the Gaussian mixture model to detect a specific type of signal. It is of interest to test this hypothesis in future work.

Finally, applying the Hyperbola Method to real data acquired using a distributed acoustic sensing is necessary. Currently, the application of the Hyperbola Method to synthetic data shows its theoretical effectiveness for determining the distance between a distributed acoustic sensor and an event. Implementing the method on real data would show how beneficial it would be for monitoring projects as well as security. In order to achieve this application to real data acquired using distributed acoustic sensing, it is necessary to fully automate the process.

Bibliography

Adcock, B., and Dexter, N., 2019, Lecture I - Introduction to machine learning with neural networks, PIMS CRG Summer School: Deep Learning for Computational Mathematics.

Aki, K., and Richards, P., 2002, Quantitative Seismology: University Science Books.

Aktaş, M., Akgün, T., Demirçin, M., and Büyükaydın, D., 2017, Deep learning based multi-threat classification for phase-OTDR fiber optic distributed acoustic sensing applications.

Aminzadeh, F., Dasgupta, S. N., and Hill, D. G., 2003, Geophysics for Petroleum Engineers: Elsevier Science & Technology.

APPEA, 2019, Seismic surveys, APPEA's website, `https://www.appea.com.au/oil-gas-explained/operation/seismic-surveys/`.

Ashmead, J., 2012, Morlet wavelets in quantum mechanics: Quanta, **1**, No. 1.

Aster, R. C., Borchers, B., and Thurber, C. H., 2005, Parameter Estimation and Inverse Problems: Elsevier Academic Press, Inc.

Auger, F., and Flandrin, P., 1995, Improving the readability of time-frequency and time-scale representations by the reassignment method: IEEE Tansactions on Signal Processing, **43**, 1068–1089.

Benedetto, J. J., and Frazier, M., 1993, Wavelets: Mathematics and Applications: CRC Press.

Bentley, P., and McDonnell, J., 1994, Wavelet transforms: an introduction: Electronics and Communication Engineering Journal, **6**, 175–186.

Beylkin, G., Coifman, R., Daubechies, I., Mallat, S., Meyer, Y., Raphael, L., and Ruskai, M., 1991, Wavelets and their Applications: Jones and Bartlett.

Boer, J. D., Mateeva, A., Pearce, J., Mestayer, J., Birch, W., Lopez, J., Hornman, J., and Kuvshinov, B., 2013, Detecting broadside acoustic signals with a fiber optical distributed acoustic sensing (DAS) assembly, u.S. Patent Application 15/677640.

Bourbié, T., Coussy, O., and Zinszner, B., 1987, Acoustics of porous media: Editions Technip, Paris.

Bruns, A., 2004, Fourier-, Hilbert- and wavelet-based signal analysis: are they really different approaches?: Journal of Neuroscience Methods, **137**, 321–332.

Bublin, M., 2019, Machine learning for distributed acoustic sensors, classic versus image and deep neural networks approach, `1904.11546`.

Byrd, R., Gilbert, J., and Nocedal, J., 2000, A trust region method based on interior point techniques for nonlinear programming: Mathematical Programming, **89**, No. 1, 149–185.

Byrd, R., Hribar, M. E., and Nocedal, J., 1999, An interior point algorithm for large-scale nonlinear programming: SIAM Journal on Optimization, **9**, No. 4, 877–900.

Carcione, J., Herman, G., and Kroode, P., 2002, Y2K review article: Seismic modeling: Geophysics, **67**, 1304–1325.

Chan, S. K., and Chen, C., 2017, Big data and monitoring the grid: The Power Grid, 253–285.

Chang, T., and Kuo, C.-C. J., 1993, Texture analysis and classification with tree-structured wavelet transforms: IEEE Transactions on Image Processing, **2**, 429–441.

Chen, S. S., Donoho, D. L., and Saunders, M. A., 1998, Atomic decomposition by basis pursuit: SIAM Journal on Scientific Computing, **20**, No. 1, 33–61, `https://doi.org/10.1137/S1064827596304010`.
URL `https://doi.org/10.1137/S1064827596304010`

Chopra, S., and Marfurt, K., 2006, Seismic attributes - a promising aid for geologic prediction: CSEG Recorder, **31**.

Chui, C. K., 1992a, Wavelet Analysis and its Applications: An Introduction to Wavelets, vol. 1: Academic Press, Inc.

Chui, C. K., 1992b, Wavelet Analysis and its Applications: Wavelets- A Tutorial in Theory and Applications, vol. 2: Academic Press, Inc.

CMC Research Institutes, 2015, CMC Containment and Monitoring Institute, `http://cmcghg.com/business-units/cami/`.

Cohen, M. X., 2019, A better way to define and describe Morlet wavelets for time-frequency analysis: NeuroImage, **199**, 81–86.

Couëslan, M. L., 2007, Processing and interpretation of time-lapse vertical seismic profile data from the penn west CO_2 monitoring project: M.Sc. thesis, University of Calgary, Calgary, AB.

Cova, R., Hardeman-Vooys, H. K., Li, D., and McDonald, M., 2018, DAS applications for near-surface characterization and traffic conditions assessment: CREWES Research Report, **30**, 1–11.

CREWES, 2008, Blackfoot 3D-3C seismic data, data set.

Culshaw, B., and Kersey, A., 2008, Fiber-optic sensing: A historical perspective: Journal of Lightwave Technology, **26**, No. 9, 1064–1078.

Daley, P. F., Krebes, E. S., and Lines, L. R., 2010, Reflection and transmission coefficient TI media: exact and linearized phase velocities, eikonals and polarization vectors: CREWES Research Report, **22**, 1–22.

Daley, T. M., Freifeld, B. M., Ajo-Franklin, J., Dou, S., Pevzner, R., Shulakova, V., Kashikar, S., Miller, D. E., Goetz, J., Henninges, J., and Lueth, S., 2013, Field testing of fibre-optic distributed acoustic sensing (DAS) for subsurface seismic monitoring: The Leading Edge.

Daubechies, I., 1988, Orthonormal bases of compactly supported wavelets: Communications on Pure and Applied Mathematics, **XLI**, 901–996.

Daubechies, I., 1992, Ten lectures on wavelets: SIAM, CBMS-NSF Regional Conference Series in Applied Mathematics.

Daubechies, I., and Maes, S., 1996, A nonlinear squeezing of the continuous wavelet transform based on auditory nerve models, *in* Aldroubi, A., and Unser, M., Eds., Wavelets in Medicine and Biology, CRC Press, 527–546.

Dou, S., Lindsey, N., Wagner, A. M., Daley, T. M., Freifeld, B., Robertson, M., Peterson, J., Ulrich, C., Martin, E. R., and Ajo-Franklin, J. B., 2017, Distributed acoustic sensing for seismic monitoring of the near surface: A traffic-noise interferometry case study: Scientific Reports, **7**, No. 11620.

Eaid, M., and Innanen, K., 2018, Modeling the sensitivity of shaped DAS fibre-optic cables to elastic wave data, Geoconvention Partnership, Geoconvention 2018.

Eaid, M., Li, J., and Innanen, K., 2017, A coupled DAS shaped-fibre and 3D elastic finite difference wave model: CREWES Research Report, **29**, 1–19.

Farge, M., Hunt, J. C. R., and Vassilicos, J. C., 1993, Wavelets, fractals, and Fourier transforms: Oxford Clarendon Press.
URL `http://openlibrary.org/books/OL1395827M`

Ferrebus, A. J. G., 2019, Processing of DAS and geophone VSP data from the CaMI Field Research Station: M.Sc. thesis, University of Calgary, Calgary, AB.

Fleet, P. J. V., 2008, Discrete Wavelet Transforms: An Elementary Approach with Applications: John Wiley and Sons, Inc.

Fryer, G., 1980, A slowness approach to the reflectivity method of seismogram synthesis: Geophysical Journal of the Royal Astronomical Society, **63**, 747–758.

Fuchs, K., and Müller, G., 1971, Computation of synthetic seismiograms with the reflectivity method and comparison with observations: Geophysical Journal of the Royal Astronomical Society, **23**, 417–433.

Gabai, H., and Eyal, A., 2016, On the sensitivity of distributed acoustic sensing: Optics Letters, **41**, No. 24, 5648–5651.

Gabai, H., and Eyal, A., 2017, How to specify and measure sensitivity in distributed acoustic sensing (das)?, *in* 2017 25th Optical Fiber Sensors Conference (OFS), 1–4.

Gävert, H., Hurri, J., Särelä, J., and Hyvärinen, A., 2005, FastICA 2.5, `https://www.mathworks.com/matlabcentral/mlc-downloads/downloads/submissions/49614/versions/2/previews/Thermal%20Pattern%20Separation/FastICA_25/fastica.m/index.html`.

Gordon, A., and Lawton, D. C., 2018, Walk away VSP processing of DAS and geophone data at CaMI Field Research Station Newell County, Alberta: CREWES Research Report, **30**, 1–22.

Goupillaud, P., Grossman, A., and Morlet, J., 1984, Cycle-octaves and related transforms in seismic signal analysis: Geoexploration, **23**, 85–102.

Grandi, S., Dean, M., and Tucker, O., 2017, Efficient containment monitoring with distributed acoustic sensing: Feasibility studies for the former Peterhead CCS project: Energy Procedia, **114**, 3889–3904.

Graps, A., 1995, A introduction to wavelets: IEEE Computational Science and Engineering, **2**, No. 2, 1–18.

Grattan, L., and Meggitt, B., 2000, Optic fiber sensor technology: Springer US.

Grossmann, A., and Morlet, J., 1984, Decomposition of hardy functions into square integrable wavelets of constant shape: SIAM Journal on Mathematical Analysis, **15**, No. 4, 723–736.

Gulati, J. S., Stewart, R. R., Peron, J., and Parkin, J. M., 1997, 3C-3D VSP: Normal moveout correction and VSPCDP transformation: CREWES Research Report, **9**, 1–30.

Guoshen, Y., 2012, Solving inverse problems with piecewise linear estimators: From Gaussian mixture models to structured sparsity: IEEE Transactions on Image Processing, **21**, No. 5, 2481–2499.

Hall, K. W., Lawton, D. C., Daley, T. M., Freifeld, B. M., and Cook, P., 2018, Effect of source effect and source distance on DAS data at CaMI.FRS, Newell County, Alberta, Geoconvention Partnership, Geoconvention 2018.

Han, J., Ciocanel, M.-V., Hardeman, H., Nasserden, D., Son, B., and Ye, S., 2015, Deducing rock properties from spectral seismic data - final report: Institute for Mathematics and its Applications Math Modeling XIX: Math Modeling in Industry Proceedings, 1–21.

Hardeman, H. K., and Lamoureux, M. P., 2016, Exact solutions for reflection coefficients, in 1D and 2D: CREWES Research Report, **28**, 1–14.

Hardeman, H. K., McDonald, M., Daley, T., Freifeld, B., Lamoureux, M. P., and Lawton, D. C., 2017, Vertical seismic profiling using distributed acoustic sensing: CREWES Research Report, **29**, 1–7.

Hardeman, H. K., McDonald, M., and Lamoureux, M. P., 2018, VSP using distributed acoustic sensing at the CaMI Field Research Station in Newell County, AB - August 2018: CREWES Research Report, **30**, 1–13.

Hartog, A., 2017, An introduction to distributed optical fibre sensors: CRC Press.

Herrera, R. H., Han, J., and van der Baan, M., 2014, Applications of the synchrosqueezing transform in seismic time-frequency analysis: Geophysics, 55–64.

Holschneider, M., 1995, Wavelets: An Analysis Tool: Oxford: Clarendon Press.

Hornman, J., 2016, Field trial of seismic recording using distributed acoustic sensing with broadside sensitive fibre-optic cables: Geophysical Prospecting.

Hornman, K., Kuvshinov, B., Zwartjes, P., and Franzen, A., 2013, Field trial of a broadside-sensitive distributed acoustic sensing cable for surface seismic, 75th EAGE Conference & Exhibition incorporating SPE EUROPEC 2013.

Huot, F., and Biondi, B., 2018, Machine learning algorithms for automated seismic ambient noise processing applied to das acquisition, 2018 SEG International Exposition and Annual Meeting, SEG.

Hyvärinen, A., 1999, Fast and robust fixed-point algorithms for independent component analysis: IEEE Transactions on Neural Networks, **10**, No. 3, 626 – 634.

Hyvärinen, A., and Oja, E., 2000, Independent component analysis: Algorithms and applications: Neural Networks, **13**, No. 4-5, 411–430.

Innanen, K., 2016, A geometrical model of DAS fibre response: CREWES Research Report, **28**.

Innanen, K., 2017, Determination of seismic tensor strain from HWC-DAS cable with arbitrary and nested-helix winds, SEG, SEG 2017.

Innanen, K., and Eaid, M., 2017, Design of DAS fibres for elastic wave mode discrimination: CREWES Research Report, **29**.

Innanen, K., Mahmoudian, F., and Eaid, M., 2017, Detection and characterization of microseismic sources with shaped DAS fibre: CREWES Research Report, **29**.

Islam, M., and Kabir, M., 2013, A new feature-based image registration algorithm: Computer Technology and Application, **4**, 79–84.

Jabbari, S., and Innanen, K., 2015, PS and SP converted wave reflection coefficients and their application to time-lapse difference AVO: CREWES Research Report, **27**, 1–14.

Jousset, P., Reinsch, T., Ryberg, T., Blanck, H., Clarke, A., Aghayev, R., Hersir, G. P., Henninges, J., Weber, M., and Krawczyk, C. M., 2018, Dynamic strain determination using fibre-optic cables allows imaging of seismological and structural: Nature Communications, **9**.

KaewTraKulPong, P., and Bowden, R., 2001, An improved adaptive background mixture model for real-time tracking with shadow detection: Springer US.

Karrenbach, M., Cole, S., Ridge, A., Boone, K., Kahn, D., Rich, J., Silver, K., and Langton, D., 2018, Fibre-optic distributed acoustic sensing of microseismicity, strain and temperature during hydraulic fracturing: Geophysics, **84**, No. 1, D11 – D23.

Kennett, B., 1979, Theoretical reflection seismograms for an elastic medium: Geophysical Prospecting, **27**, 301–321.

Kennett, B., 1983, Seismic wave propagation in stratified media: John Wiley and Sons, Inc.

Kimbell, J. F., 2013, History and analysis of distributed acoustic sensing (das) for oilfield applications: M.Sc. thesis, Texas A&M University.

Kind, R., 1976, Computation of reflection coefficients for layered media: Geophysics, **41**, 191–200.

Kroon, D.-J., 2010, fminlbfgs, `https://www.mathworks.com/matlabcentral/fileexchange/23245-fminlbfgs-fast-limited-memory-optimizer`.

Kumar, P., and Foufoula-Georgiou, E., 1994, Wavelets in geophysics: Academic Press, Inc.

Kumar, P., and Foufoula-Georgiou, E., 1997, Wavelet analysis for geophysical applications: Reviews of Geophysics, **35**, No. 4, 385–412.

Kuvshinov, B., 2015, Interaction of helically wound fibre-optic cables with plane seismic waves: Geophysical Prospecting, **64**.

Lagarias, J., Reeds, J., Wright, M., and Wright, P., 1998, Convergence properties of the Nelder-Mead simplex method in low dimensions: SIAM Journal of Optimization, **9**, No. 1, 112–147.

Lamoureux, M. P., Gibson, P. C., and Margrave, G. F., 2012, Reflection coefficients through a linear velocity ramp, in 1D: CREWES Research Report, **24**, 1–12.

Lamoureux, M. P., Gibson, P. C., and Margrave, G. F., 2013, Can we see a velocity ramp, in reflection seismic data: CREWES Research Report, **25**, 1–15.

Lapidus, L., and Pinder, G. F., 1982, Numerical Solutions of Partial Differential Equations in Science and Engineering: Wiley Press, Inc.

Li, C., and Liang, M., 2012, A generalized synchrosqueezing transform for enhancing signal time-frequency representation: Signal procesing, **92**, 2264–2274.

Lira, J., Weglein, A., Bird, C., and Innnanen, K., 2011, Determination of reflection coefficients by comparison of direct and reflected vsp events: CREWES Research Report, **23**, 1–13.

Ma, Y., Loures, L., and Margrave, G. F., 2004, Seismic modelling with the reflectivity method: CREWES Research Report, **16**, 1–7.

Mallat, S., 1989a, Multifrequency channel decomposition of images and wavelet models: IEEE Transactions on Acoustic Speech and Signal Processing, **37**, No. 12, 2091–2110.

Mallat, S., 1989b, A theory for multiresolution signal decomposition: The wavelet representation: IEEE Transactions on Pattern Analysis and Machine Intelligence, **11**, No. 7, 674–693.

Mallat, S., 2009, A Tour of Signal Processing: The Sparse Way: Elsevier, Inc.

Mann, S., and Haykin, S., 1991, The chirplet transform: A generalization of Gabor's logon transform: Vision Interface '91, 205–212, iSSN 0843-803X.

Margrave, G., 2006, Methods in seismic data processing, Geophysics 557/657, course Lecture Notes.

Margrave, G., 2018, CREWES toolbox version: 1713, `https://www.crewes.org/ResearchLinks/FreeSoftware/`.

Martin, E., Biondi, B., Cole, S., and Karrenbach, M., 2017, Continuous subsurface monitoring by passive seismic data recorded with distributed acoustic sensors: The "Stanford DAS array" experiment, EAGE, Expanded EAGE Abstracts.

Martin, J., Donno, D., Papp, B., and Hartog, A., 2014, Fibre optic distributed vibration sensing with directional sensitivity, wO2014201313.

Martins, W., Campos, M., Chaves, R., Lordelo, C., Ellmauthaler, A., Nunes, L., and Barfoot, D., 2017, Communication models for distributed acoustic sensing for telemetry: A tutorial: IEEE Sensors Journal, **PP**, 1–1.

Masoudi, A., and Newson, T. P., 2017, Analysis of distributed optical fibre acoustic sensors through numerical modelling: Opt. Express, **25**, No. 25, 32,021–32,040.
URL http://www.opticsexpress.org/abstract.cfm?URI=oe-25-25-32021

Mateeva, A., Lopez, J., Potters, H., Mestayer, J., Cox, B., Kiyashchenko, D., Willis, P., Grandi, S., Hornman, K., Kuvshinov, B., Bertang, W., Yang, Z., and Detomo, R., 2014, Distributed acoustic sensing for reservoir monitoring with vertical seismic profiling: Geophysical Prospecting, **62**.

Mateeva, A., Mestayer, J., Cox, B., Kiyashchenko, D., Willis, P., Lopez, J., Grandi, S., Hornman, K., Hornman, K., Lumens, P., Franzen, A., Hill, D., and Roy, J., 2012, Advances in distributed acoustic sensing (DAS) for VSP, SEG, SEG Technical Program Expanded Abstracts.

McDonald, M., 2015, Microseismic hypocentre location, via distributed acoustic, 2015 PIMS Industrial Problem Solving Workshop, problem proposed at workshop.

McDonald, M., 2016a, Chirp signal.

McDonald, M., 2016b, Distributed acoustic sensing, an optional note.

Mestayer, J., Cox, B., Wills, P., Kiyashchenko, D., Lopez, J., Costello, M., Bourne, S., Ugueto, G., Lupton, R., Solano, G., Hill, D., and Lewis, A., 2011, Field trials of distributed acoustic sensing for geophysical monitoring, SEG, SEG Technical Program Expanded Abstracts.

Mestayer, J., Karam, S., Cox, B., Willis, P., Mateeva, A., Lopez, J., Hill, D., and Lewis, A., 2016, DAS microseismic: CSEG Recorder Focus Article, 38–39.

Meyer, Y., 1992, Wavelets and Operators: Cambridge University Press.

Meyer, Y., 1993, Wavelets: Algorithms and Applications: Society of Industrial and Applied Mathematics.

Meyer, Y., and Roques, S., 1993, Progress in Wavelet Analysis and Applications: Frontiéres.

Morlet, J., Arens, G., Fourgeau, E., and Giard, D., 1982a, Wave propagation and sampling theory, 1, complex signal and scattering in multilayered media: Geophysics, **47**, No. 2, 203–221.

Morlet, J., Arens, G., Fourgeau, E., and Giard, D., 1982b, Wave propagation and sampling theory, 2, sampling theory and complex waves: Geophysics, **47**, No. 2, 222–236.

Muanenda, Y., 2018, Review article: Recent advances in distributed acoustic sensing based on phase-sensitive optical time domain reflectometry: Journal of Sensors, **2018**, 1–16.

Mueller, G., 1985, The reflectivity method: a tutorial: Zeitschrift für Geophysik, **58**, No. 1-3, 153–174.

Ng, A., Ngiam, J., Foo, C. Y., Mai, Y., Suen, C., Coates, A., Maas, A., Hannun, A., Huval, B., Wang, T., and Tandon, S., 2013, UFLDL: Welcome to the deep learning tutorial, `http://ufldl.stanford.edu/tutorial/`.

Nizkous, I., Gerritsen, S., Kiyashchenko, D., and Joinson, D., 2015, Distributed acoustic sensing (DAS) VSP for imaging and velocity model building, International Petroleum Technology Conference.

Porwik, P., and Lisowska, A., 2004, The Haar-wavelet transform in digital image processing: Its status and achievements: Machine Graphics and Vision, **13**, No. 1/2, 79–98.

Pujol, J., 2003, Elastic Wave Propagation and Generation in Seismology: Cambridge Press, Inc.

Ricker, N., 1953, The form and laws of propagation of seismic wavelets: Geophysics, **18**, No. 1, 10–40.

Rotman, J. J., 1988, An Introduction to Algebraic Topology: Springer-Verlag New York Inc.

Ruskai, M. B., Beylkin, G., Coifman, R., Daubechies, I., Mallat, S., Meyers, Y., and Raphael, L., 1992, Wavelets and Their Applications: Jones and Bartlett.

Schmidt, M., 2005, minfunc, `https://www.cs.ubc.ca/~schmidtm/Software/minFunc.html`.

Shamsa, A., and Paydayesh, M., 2019, Applications of independent component analysis and Gaussian mixture models in micro-seismic signal detection, Geoconvention Partnership, Geoconvention 2019.

Sharma, R. K., and Ferguson, R. J., 2009, Reflection and transmission coefficients for SH wave in plane wave domain: CREWES Research Report, **21**, 1–23.

Sharma, R. K., and Ferguson, R. J., 2010, Reflection and transmission coefficient for VTI media: CREWES Research Report, **22**, 1–22.

Sherman, C., Mellors, R., Morris, J., and Ryerson, F., 2019, Geomechanical modeling of distributed fiber-optic sensor measurements: Interpretation, **7**, SA21–SA27.

Shi, Y., Wang, Y., Zhao, L., and Fan, Z., 2019, An event recognition method for ϕ-OTDR sensing system based on deep learning: Sensors, **19**, 1–9.

Shiloh, L., Eyal, A., and Giryes, R., 2018, Deep learning approach for processing fiber-optic DAS seismic data, *in* 26th International Conference on Optical Fiber Sensors, Optical Society of America, ThE22.
URL `http://www.osapublishing.org/abstract.cfm?URI=OFS-2018-ThE22`

Shiloh, L., Eyal, A., and Giryes, R., 2019, Efficient processing of distributed acoustic sensing data using a deep learning approach: Journal of Lightwave Technology, **37**, No. 18, 4755–4762.

Sinha, S., Routh, P. S., Anno, P. D., and Castagna, J. P., 2005, Spectral decomposition of seismic data with continuous-wavelet transform: Geophysics, **70**, 19–25.

Spielberg, S., 1993, Jurassic Park.

Stauffer, C., and Grimson, W. E. L., 1999, Adaptive background mixture models for real-time tracking, *in* 1999 Conference on Computer Vision and Pattern Recognition (CVPR '99), 23-25 June 1999, Ft. Collins, CO, USA, 2246–2252.
URL `https://doi.org/10.1109/CVPR.1999.784637`

Stewart, R. R., 2001, VSP: An in-depth seismic understanding: CSEG Recorder, **26**.

Tary, J. B., Herrera, R. H., Han, J., and van der Baan, M., 2014, Spectral estimation–what is new? what is next?: Reviews of Geophysics, **52**, 1–27.

Taylor, M., 1996, Partial Differential Equations: Springer, Inc.

Matlab, 2019, Detecting cars using Gaussian mixture models, `https://www.mathworks.com/help/vision/examples/detecting-cars-using-gaussian-mixture-models.html`.

Thakur, G., and Wu, H.-T., 2011, Synchrosqueezing-based recovery of instantaneous frequency from nonuniform samples: SIAM Journal on Mathematical Analysis, **43**, No. 5, 2078–2095.
URL `http://dx.doi.org/10.1137/100798818`

Thomson, W., 1950, Transmission of elastic waves through a stratified solid medium: Journal of Applied Physics, **21**, 89–93.

Torrence, C., and Compo, G. P., 1998, A practical guide to wavelet analysis: Bulletin of the American Meteorological Society, **79**, No. 1, 61–78.

Ursenbach, C. P., and Hasse, A. B., 2004, An efficient method for calculating spherical-wave reflection coefficients: CREWES Research Report, **16**, 1–14.

Vooys, G., 2019, Homotopy discussion, Personal communication.

Waltz, R., Morales, J. L., Nocedal, J., and Orban, D., 2006, An interior algorithm for nonlinear optimization that combines line search and trust region steps: Mathematical Programming, **107**, No. 3, 391–408.

Wang, J. Z., 2001, Wavelets and imaging informatics: A review of the literature: Journal of Biomedical Informatics, **34**, 129–141.

Wang, Y., 2015, Generalized seismic wavelets: Geophysical Journal International, **203**, No. 2, 1172–1178.

Wang, Y., Yuan, H., Liu, X., Bai, Q., Zhang, H., Gao, Y., and Jin, B., 2019, A comprehensive study of optical fiber acoustic sensing: IEEE Access, **PP**, 1–1.

Webster, P., Molenaar, M., and Perkins, C., 2016, DAS microseismic: CSEG Recorder, **41**, No. 6.

Webster, P., Wall, J., Perkins, C., and Molenaar, M., 2013, Micro-seismic detection using distributed acoustic sensing, 2459–2463.

Williams, A., Kendall, J. M., Clarke, A., and Verdon, J., 2017, Challenges in Locating Microseismic Events Using Distributed Acoustic Sensors, *in* AGU Fall Meeting Abstracts, vol. 2017, S33B–0873.

Wornell, G., 1995, Signal Processing With Fractals: A Wavelet-Based Approach: Prentice-Hall.

Wu, B., 1992, An introduction to neural networks and their applications in manufacturing: Journal of Intelligent Manufacturing, **3**, 391–403.

Wu, H.-T., Flandrin, P., and Daubechies, I., 2011, One or two frequencies? The synchrosqueezing answers: Advances in Adaptive Data Analysis, **3(2)**, 29–39.

Wuestefeld, A., and Wilks, M., 2019, How to twist and turn a fiber: Performance modeling for optimal DAS acquisitions: The Leading Edge, **38**.

Yang, B., 2014, Bin Yang, https://github.com/byangderek.

Yavuz, S., Freifeld, B., Pevzner, R., Tertyshnikov, K., Dzunic, A., Ziramov, S., Shulakova, V., Robertson, M., Daley, T., Kepic, A., Urosevic, M., and Gurevich, B., 2016, Subsurface imaging using buried das and geophone arrays - preliminary results from co2crc otway project, EAGE, Extended EAGE Abstracts.