

Ground roll attenuation with least-squares and robust inversion

by

Iliana Papathanasaki

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

Master of Science  
in  
Geophysics

Department of Physics  
University of Alberta

©Iliana Papathanasaki, 2020

# Abstract

This thesis proposes a framework for ground roll removal based on regularized inversion. Both ground roll and reflections are represented in the frequency-space ( $f - x$ ) domain as a linear function of unknown complex amplitudes. An inversion algorithm is developed to estimate the coefficients that model the ground roll and reflections separately. The latter allows for an independent synthesis of ground roll and reflections. The synthesized ground roll is subtracted from the seismic records to yield a seismic record with enhanced reflections.

In the first part of this thesis, a regularized least-squares inversion algorithm is proposed. I further propose to adopt two types of regularization to guarantee the stability of the inversion. To be more specific, I compare classical quadratic regulation with sparsity promoting regularization. Numerical experiments show that sparsity promoting regularization yields better results than quadratically regularized inversion. A robust regularized inversion algorithm is also proposed to cope with erratic noise often present in onshore data.

Through synthetic and field data examples, we prove the effectiveness and the limitations of these algorithms.

# Acknowledgements

First of all, I want to thank my supervisor, Dr. Mauricio Sacchi, for his endless support and his advice towards the development of this thesis.

I also appreciate my committee members for their support on this project.

I want to thank all my colleagues in the Signal Analysis and Imaging Group (SAIG) for our unforgettable moments and their help in my research.

Furthermore, I appreciate the sponsors of SAIG for their financial support and advice from our annual meetings.

I want to thank my colleague Felix Oghenekohwo for his help in understanding some necessary steps of this project.

Finally, I would like to thank my family back in Greece for their support all these years here at the University of Alberta.

# Contents

<b>1</b>	<b>Introduction</b>	<b>1</b>
1.1	Seismic exploration and wave propagation . . . . .	1
1.1.1	Seismic reflection and traveltimes . . . . .	3
1.2	Surface waves: Rayleigh waves . . . . .	5
1.3	Ground roll and attenuation methods . . . . .	7
1.4	Scope of this thesis . . . . .	11
1.4.1	Contributions . . . . .	11
1.4.2	Overview . . . . .	13
<b>2</b>	<b>The Forward Problem</b>	<b>14</b>
2.1	Introduction . . . . .	14
2.2	Ground roll modelling . . . . .	15
2.3	Method . . . . .	16
2.3.1	Modeling the signal . . . . .	17
2.3.2	Modeling the coherent noise . . . . .	18
2.4	Synthetic modeling . . . . .	18
2.5	Synthetic Data in FX domain . . . . .	19
2.6	Summary . . . . .	22

<b>3</b>	<b>Inversion Algorithms</b>	<b>23</b>
3.1	Introduction . . . . .	23
3.2	Least-Squares inversion algorithms . . . . .	23
3.3	Robust inversion algorithms . . . . .	25
3.4	Summary . . . . .	27
<b>4</b>	<b>The Inverse Problem</b>	<b>28</b>
4.1	Introduction . . . . .	28
4.2	Least-Squares inversion with quadratic regularization . . . . .	28
4.2.1	Parameters' selection . . . . .	29
4.3	Least-Squares inversion with sparse regularization . . . . .	31
4.3.1	Inversion through IRLS and FISTA algorithms . . . . .	31
4.3.2	Parameters' selection . . . . .	33
4.4	Robust inversion . . . . .	34
4.4.1	The ADMM algorithm . . . . .	34
4.4.2	Parameters' selection . . . . .	35
4.5	Comparison of convergence for IRLS, FISTA, and ADMM . . . . .	36
4.6	Numerical examples . . . . .	37
4.6.1	Synthetic examples . . . . .	37
4.6.2	Field data examples . . . . .	48
4.7	Summary . . . . .	71
<b>5</b>	<b>Conclusions</b>	<b>72</b>
5.1	Contributions . . . . .	72
5.2	Future developments . . . . .	74
	<b>Bibliography</b>	<b>75</b>

# List of Tables

4.1	Parameters used for inversion when $\tau, v$ pairs of the model are known-synthetic data . . . . .	41
4.2	Parameters used for inversion when $\tau, v$ pairs of the model are unknown-synthetic data. . . . .	46
4.3	Parameters used for inversion when $\tau, v$ pairs of the model are known-3D real data. . . . .	54
4.4	Parameters used for inversion when $\tau, v$ pairs of the model are unknown-3D real data . . . . .	58
4.5	Parameters used for inversion when $\tau, v$ pairs of the model are known-2D real data . . . . .	66
4.6	Parameters used for inversion when $\tau, v$ pairs of the model are unknown-2D real data . . . . .	70

# List of Figures

1.1	Schematic figure of a land seismic survey. A vibrator truck produces waves that propagate into the subsurface. Downgoing waves are reflected by interfaces and propagate upwards. Upward propagating waves are recorded by geophones (retrieved from <a href="http://geologylearn.blogspot.com/2015/06/marine-and-land-seismic-aquisition.html">http://geologylearn.blogspot.com/2015/06/marine-and-land-seismic-aquisition.html</a> ). . . . .	2
1.2	Schematic figure of a marine seismic survey. An air gun produces waves that propagate into the subsurface and are recorded by hydrophones (retrieved from <a href="https://www.liberaldictionary.com/seismic/">https://www.liberaldictionary.com/seismic/</a> ). . . . .	3
1.3	Sketch showing a P wave reflection and refraction to P and S waves in two media . . . . .	4
1.4	Sketch showing a P wave reflected to P and SV waves for a free surface on an elastic medium . . . . .	5
1.5	Sketch showing the particle motion for Rayleigh propagating from left to right .	6
1.6	Sketch showing group and phase velocities for a sum of waves. . . . .	7
1.7	Synthetic shot gather showing dispersive ground roll masking reflections. . . .	8
1.8	Not aliased signal in the $f - k$ domain. A filter can remove the frequency containing the signal we wish to attenuate. . . . .	9
1.9	The alias phenomenon is visible when displaying a record in the $f - k$ domain. For frequencies 25 to 125 HZ, the signal cannot be removed without affecting any other existing signal (reflections etc.) . . . . .	10
2.1	A simple definition for the forward and inverse problems in Geophysics for a model $m$ and the data $d$ . . . . .	15
2.2	An irregular offset 3D data set with very strong ground roll. Note that in traces 450-650, the coherent noise has lost its linearity. . . . .	16
2.3	Source and receivers for the 3D synthetic example adopted to test our signal and ground roll separation method. . . . .	20
2.4	a) 3D synthetic shot gather. b) 3D synthetic shot gather contaminated with random noise ( $SNR = 1$ ). . . . .	21

3.1	Comparison of non robust and robust regression when outliers are present. Non robust algorithms consider them as good observations (left image). Robust algorithms are not affected by them (right image). (Adapted by Sacchi, 2015.) . . . . .	24
4.1	An example of how to estimate the $\mu$ parameter using the L-curve method . . . . .	30
4.2	Separation of reflections and ground roll using least-squares inversion with quadratic regularization. The $\tau, v$ pairs describing hyperbolic events are assumed known for the inversion algorithm. a) 3D synthetic shot gather. b) 3D synthetic shot gather contaminated with random noise ( $SNR = 1$ ). c) Estimated signal. d) Estimate coherent noise (ground roll). e) Observations (b) minus estimated coherent noise (d). . . . .	37
4.3	Separation of reflections and ground roll using least-squares inversion with sparse ( $l_1$ ) regularization (IRLS). The $\tau, v$ pairs describing hyperbolic events are assumed known for the inversion. a) 3D synthetic shot gather. b) 3D synthetic shot gather contaminated with random noise ( $SNR = 1$ ). c) Estimated signal. d) Estimates coherent noise (ground roll). e) Observations (b) minus estimated coherent noise (d). . . . .	38
4.4	Separation of reflections and ground roll using least-squares inversion with sparse ( $l_1$ ) regularization (FISTA). The $\tau, v$ pairs describing hyperbolic events are assumed known for the inversion. a) 3D synthetic shot gather. b) 3D synthetic shot gather contaminated with random noise ( $SNR = 1$ ). c) Estimated signal. d) Estimates coherent noise (ground roll). e) Observations (b) minus estimated coherent noise (d). . . . .	38
4.5	Separation of reflections and ground roll using robust inversion (ADMM). a) 3D synthetic shot gather. b) 3D synthetic shot gather contaminated with random noise ( $SNR = 1$ ). c) Estimated signal. d) Estimate coherent noise (ground roll). e) Observations (b) minus estimated coherent noise (d). . . . .	39
4.6	Spectra of the ground roll $ m_c(f, p) $ . a) Estimation via least-squares inversion with quadratic ( $l_2$ ) regularization. b) Least-squares inversion with sparse ( $l_1$ ) regularization (IRLS). c) Least-squares inversion with sparse ( $l_1$ ) regularization (FISTA). d) Sparse inversion (ADMM). . . . .	40
4.7	Spectra of the ground roll $ m_c(f, p) $ for $f = 20 Hz$ . a) Estimation via least-squares inversion with quadratic ( $l_2$ ) regularization. b) Least-squares inversion with sparse ( $l_1$ ) regularization (IRLS). c) Least-squares inversion with sparse ( $l_1$ ) regularization (FISTA). d) Sparse inversion (ADMM). Dashed lines indicate the theoretical dispersion modes used to generate the synthetic data. . . . .	40
4.8	Separation of reflections and ground roll using least-squares inversion with quadratic regularization. The $\tau, v$ pairs describing hyperbolic events are unknown. Therefore, we assume a coarse discretization of $\tau, v$ . a) 3D synthetic shot gather. b) 3D synthetic shot gather contaminated with random noise ( $SNR = 1$ ). c) Estimated signal. d) Estimate coherent noise (ground roll). e) Synthetic observations (b) minus estimated coherent noise (d). . . . .	42

4.9	Separation of reflections and ground roll using least-squares inversion with sparse ( $l_1$ ) inversion using IRLS. The $\tau, v$ pairs describing hyperbolic events are unknown for the inversion process. Coarse discretization of the pairs $\tau, v$ is adopted. a) 3D synthetic shot gather. b) 3D synthetic shot gather contaminated with random noise ( $SNR = 1$ ). c) Estimated signal. d) Estimate coherent noise (ground roll). e) Synthetic observations (b) minus estimated coherent noise (d). . . . .	43
4.10	Separation of reflections and ground roll using least-squares inversion with sparse ( $l_1$ ) inversion using FISTA. The $\tau, v$ pairs describing hyperbolic events are unknown for the inversion process. Coarse discretization of the pairs $\tau, v$ is adopted. a) 3D synthetic shot gather. b) 3D synthetic shot gather contaminated with random noise ( $SNR = 1$ ). c) Estimated signal. d) Estimate coherent noise (ground roll). e) Synthetic observations (b) minus estimated coherent noise (d). . . . .	43
4.11	Separation of reflections and ground roll using robust inversion with ADMM. The $\tau, v$ pairs describing hyperbolic events are unknown for the inversion process. Coarse discretization of the pairs $\tau, v$ is adopted. a) 3D synthetic shot gather. b) 3D synthetic shot gather contaminated with random noise ( $SNR = 1$ ). c) Estimated signal. d) Estimate coherent noise (ground roll). e) Synthetic observations (b) minus estimated coherent noise (d). . . . .	44
4.12	Spectra of the ground roll $ m_c(f, p) $ . a) Estimation via least-squares inversion with quadratic ( $l_2$ ) regularization. b) Least-squares inversion with sparse ( $l_1$ ) regularization (IRLS). c) Least-squares inversion with sparse ( $l_1$ ) regularization (FISTA). d) Sparse inversion (ADMM). . . . .	45
4.13	Spectra of the ground roll $ m_c(f, p) $ for $f = 20 Hz$ . a) Estimation via least-squares inversion with quadratic ( $l_2$ ) regularization. b) Least-squares inversion with sparse ( $l_1$ ) regularization (IRLS). c) Least-squares inversion with sparse ( $l_1$ ) regularization (FISTA). d) Robust inversion (ADMM). Dashed lines indicate the theoretical dispersion modes used to generate the synthetic data. . . . .	45
4.14	Spectra of hyperbolic events coarsely discretized $\sum_f Re[m_s(f, \tau, v)]$ . a) Estimation via least-squares inversion with quadratic ( $l_2$ ) regularization. b) Least-squares inversion with sparse ( $l_1$ ) regularization. . . . .	47
4.15	Source and receivers for the irregular field data example adopted to test the effectiveness of our signal and ground roll separation methods. . . . .	49
4.16	Automatic Gain Control with window size 501. Reflections are stronger for the next processing steps. . . . .	50
4.17	Velocity Analysis applied on a shot section to estimate the intercept time-velocity pairs. The areas with higher energy show the possible pairs. . . . .	51
4.18	Separation of reflections and ground roll using least-squares inversion with quadratic regularization. a) 3D real shot gather. b) Estimated coherent noise (ground roll). c) Real observations (a) minus estimated coherent noise (b). . . . .	52

4.19	Separation of reflections and ground roll using least-squares inversion with sparse ( $l_1$ ) regularization using IRLS. a) 3D real shot gather. b) Estimated coherent noise (ground roll). c) Real observations (a) minus estimated coherent noise (b).	52
4.20	Separation of reflections and ground roll with least-squares inversion with sparse ( $l_1$ ) regularization using FISTA. a) 3D real shot gather. b) Estimated coherent noise (ground roll). c) Real observations (a) minus estimated coherent noise (b).	53
4.21	Separation of reflections and ground roll with robust inversion using ADMM. a) 3D real shot gather. b) Estimated coherent noise (ground roll). c) Real observations (a) minus estimated coherent noise (b).	53
4.22	a) Separation of reflections and ground roll using least-squares inversion with ( $l_2$ ) regularization, b) with ( $l_1$ ) regularization using IRLS, c) using FISTA and d) using ADMM	54
4.23	Separation of reflections and ground roll using least-squares inversion with quadratic regularization. a) 3D real shot gather. b) Estimated coherent noise (ground roll). c) Real observations (a) minus estimated coherent noise (b).	55
4.24	Separation of reflections and ground roll using least-squares inversion with sparse ( $l_1$ ) regularization using IRLS. a) 3D real shot gather. b) Estimated coherent noise (ground roll). c) Real observations (a) minus estimated coherent noise (b).	56
4.25	Separation of reflections and ground roll with least-squares inversion with sparse ( $l_1$ ) regularization using FISTA. a) 3D real shot gather. b) Estimated coherent noise (ground roll). c) Real observations (a) minus estimated coherent noise (b).	56
4.26	Separation of reflections and ground roll with robust inversion using ADMM. a) 3D real shot gather. b) Estimated coherent noise (ground roll). c) Real observations (a) minus estimated coherent noise (b).	57
4.27	a) Separation of reflections and ground roll using least-squares inversion with ( $l_2$ ) regularization, b) with ( $l_1$ ) regularization using IRLS, c) using FISTA and d) using ADMM	57
4.28	2D field data set with very dispersive ground roll	59
4.29	Source and receivers for the irregular 2D field data example adopted to test the effectiveness of our signal and ground roll separation methods.	60
4.30	Automatic Gain Control with window size 501. Reflections are stronger for next processing steps.	61
4.31	Velocity Analysis applied on a shot section to estimate the intercept time-velocity pairs. The areas with higher energy show the possible pairs.	62
4.32	Separation of reflections and ground roll using least-squares inversion with quadratic regularization. a) 3D real shot gather. b) Estimated coherent noise (ground roll). c) Real observations (a) minus estimated coherent noise (b).	63

4.33	Separation of reflections and ground roll using least-squares inversion with sparse ( $l_1$ ) regularization using IRLS. a) 3D real shot gather. b) Estimated coherent noise (ground roll). c) Real observations (a) minus estimated coherent noise (b).	63
4.34	Separation of reflections and ground roll with least-squares inversion with sparse ( $l_1$ ) regularization using FISTA. a) 3D real shot gather. b) Estimated coherent noise (ground roll). c) Real observations (a) minus estimated coherent noise (b).	64
4.35	Separation of reflections and ground roll with robust inversion using ADMM. a) 3D real shot gather. b) Estimated coherent noise (ground roll). c) Real observations (a) minus estimated coherent noise (b).	64
4.36	a) Zoomed area from raw data containing dispersive ground roll, b) Separation of reflections and ground roll using least-squares inversion with ( $l_2$ ) regularization, c) with ( $l_1$ ) regularization using IRLS, d) using FISTA and e) using ADMM	65
4.37	Separation of reflections and ground roll using least-squares inversion with quadratic regularization. a) 3D real shot gather. b) Estimated coherent noise (ground roll). c) Real observations (a) minus estimated coherent noise (b).	67
4.38	Separation of reflections and ground roll using least-squares inversion with sparse ( $l_1$ ) regularization using IRLS. a) 3D real shot gather. b) Estimated coherent noise (ground roll). c) Real observations (a) minus estimated coherent noise (b).	67
4.39	Separation of reflections and ground roll with least-squares inversion with sparse ( $l_1$ ) regularization using FISTA. a) 3D real shot gather. b) Estimated coherent noise (ground roll). c) Real observations (a) minus estimated coherent noise (b).	68
4.40	Separation of reflections and ground roll with robust inversion using ADMM. a) 3D real shot gather. b) Estimated coherent noise (ground roll). c) Real observations (a) minus estimated coherent noise (b).	68
4.41	a) Zoomed area from raw data containing dispersive ground roll. b) Separation of reflections and ground roll using least-squares inversion with ( $l_2$ ) regularization, c) with ( $l_1$ ) regularization using IRLS, d) using FISTA and e) using ADMM	69

---

---

# CHAPTER 1

---

## Introduction

### 1.1 Seismic exploration and wave propagation

Reflection seismology is a geophysical method often used to estimate images of the subsurface. Reflection seismology adopts the principles of wave propagation phenomena, and it heavily relies on data processing techniques to preconditioning seismic records before imaging. Reflection seismology is commonly used for commercial purposes, such as exploring and exploiting hydrocarbons and mineral deposits (Liu et al. (2018); Malehmir et al. (2012)).

The method can be summarized as follows. In an onshore or marine survey, a source generates waves, propagating through the different subsurface layers. These waves are reflected by geological interfaces and propagate back to Earth's surface where they are registered by arrays of receivers in the form of time series, often called seismic traces. In an onshore survey, dynamites and vibroseis are the most commonly used sources. Their use depends on surface topographic complexity and the accessibility of the survey area. In marine reflection seismology, the source consists of airguns (Dragoset (2000)).

Geophones for land data and hydrophones for marine acquisitions are used to record the seismic waves. These detectors are deployed as arrays specifically designed for noise attenuation and imaging purposes. Figures (1.1) and (1.2) show a land and marine seismic surveys.

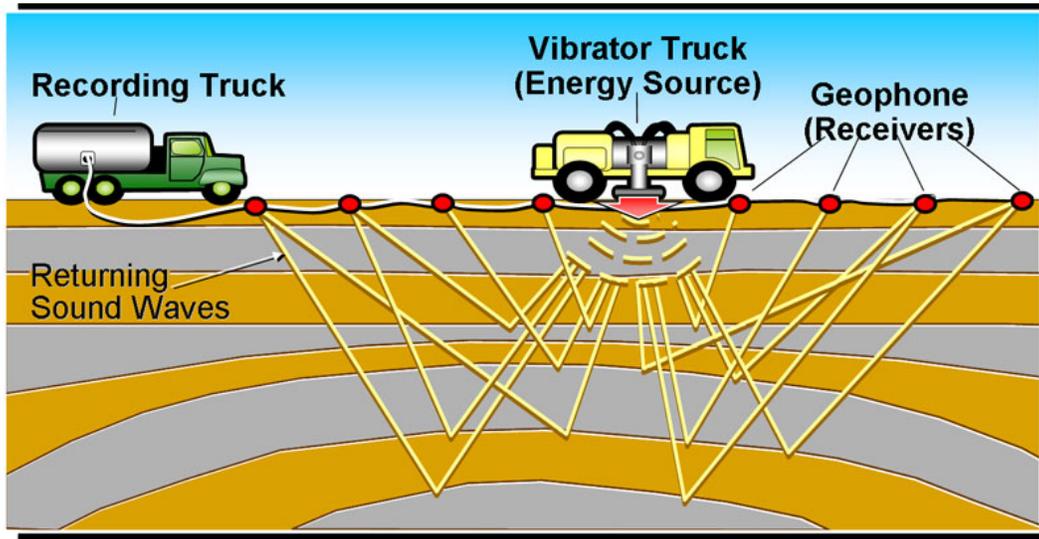


Figure 1.1: Schematic figure of a land seismic survey. A vibrator truck produces waves that propagate into the subsurface. Downgoing waves are reflected by interfaces and propagate upwards. Upward propagating waves are recorded by geophones (retrieved from <http://geologylearn.blogspot.com/2015/06/marine-and-land-seismic-aquisition.html>).

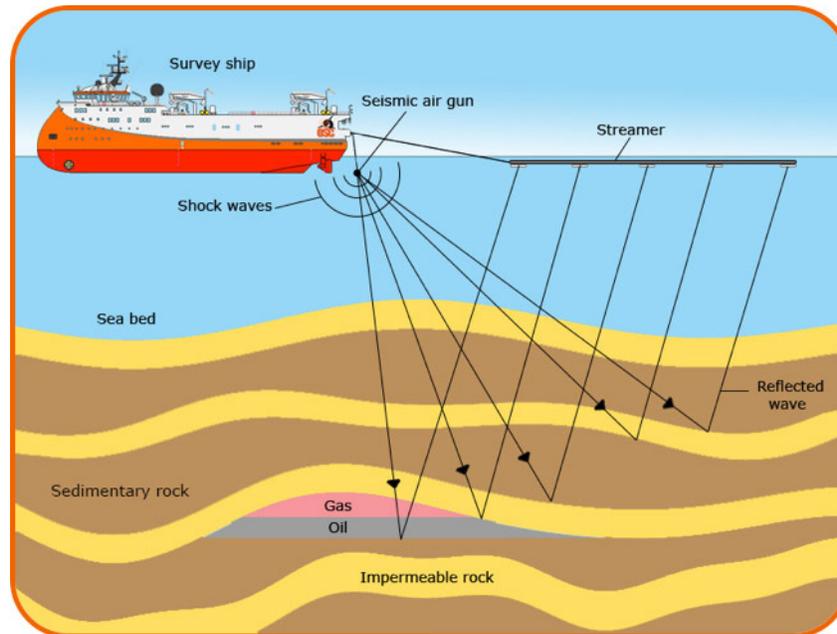


Figure 1.2: Schematic figure of a marine seismic survey. An air gun produces waves that propagate into the subsurface and are recorded by hydrophones (retrieved from <https://www.liberaldictionary.com/seismic/>).

Seismic waves generated by sources, as described above, propagate into the earth. This energy is refracted and reflected and propagates to the surface, and then it is recorded by receivers. The signal recorded by receivers contains the desired information for imaging the subsurface. However, before imaging, a series of processing steps are applied to eliminate unwanted noise and enhance interest signals.

### 1.1.1 Seismic reflection and traveltimes

Waves propagate in two ways: Body and Surface waves (Udias and Bufo (2017); Shearer (2019)). Body waves propagate in deep structures of the subsurface, forming P, SH, and SV waves. In contrast, surface waves travel along the surface, and they do not contain information about the subsurface structure. They form Love and Rayleigh waves, and they cause noise in seismic records.

Let us see what happens when waves propagate from one medium to another. The interior of the earth is not homogeneous. In other words, rocks have different elastic properties, which yields to stratified layers of materials of varying P-wave, S-wave, and density. When the waves travel propagate through these discontinuities, part of the energy is propagated into the next medium, and part is reflected back.

According to Fermat's principle, the waves follow the path of shortest duration (Udias and Bufo (2017); Shearer (2019); Bóna and Slawinski (2003)). For both reflection and refraction, these paths follow Snell's law, resulting from Fermat's principle (Udias and Bufo (2017); Shearer (2019); Bóna and Slawinski (2003); Yajima and Nagahama (2007)). Figure 1.3 shows two media where from an incident P wave, P and S waves are reflected and refracted from a medium M.

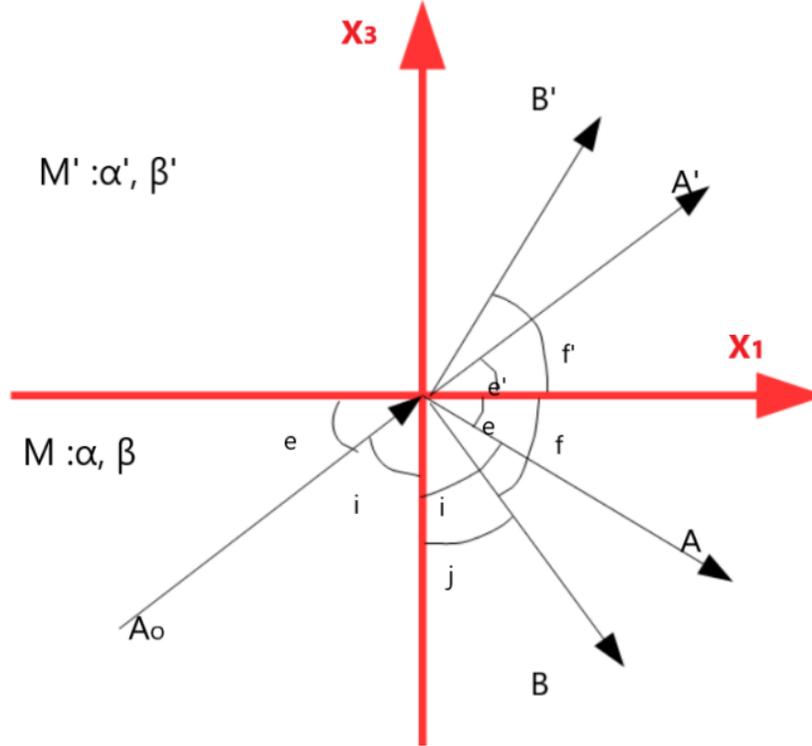


Figure 1.3: Sketch showing a P wave reflection and refraction to P and S waves in two media .

If  $e$  and  $f$  are the angles of incident and reflected P and S waves in medium M of velocities,  $\alpha$ ,  $\beta$  and  $e'$ ,  $f'$  are those of refracted waves in medium M' with velocities  $\alpha'$  and  $\beta'$ , then Snell's law is given by

$$\frac{\cos e}{\alpha} = \frac{\cos f}{\beta} = \frac{\cos e'}{\alpha'} = \frac{\cos f'}{\beta'} = \frac{1}{c} = p \quad (1.1)$$

where  $p$  is the ray parameter and  $c$  is the apparent velocity of propagation.

This thesis is focused on Rayleigh waves, which combine P and SV motions, and their displacements are in vertical planes (Udias and Bufo (2017); Shearer (2019)). Rayleigh

waves are idealized for what in exploration seismology is called ground roll and are considered unwanted signals. They are also called coherent noise, and they are dominant in land seismic records. They mask the body waves (reflections), making further processing steps difficult.

## 1.2 Surface waves: Rayleigh waves

The existence of a free surface in a medium leads to surface waves generation (Udias and Buforn (2017); Shearer (2019)). In a free surface, the boundary conditions are that components of stress across it are zero. This has great interest as the second medium is a vacuum (figure 1.4).

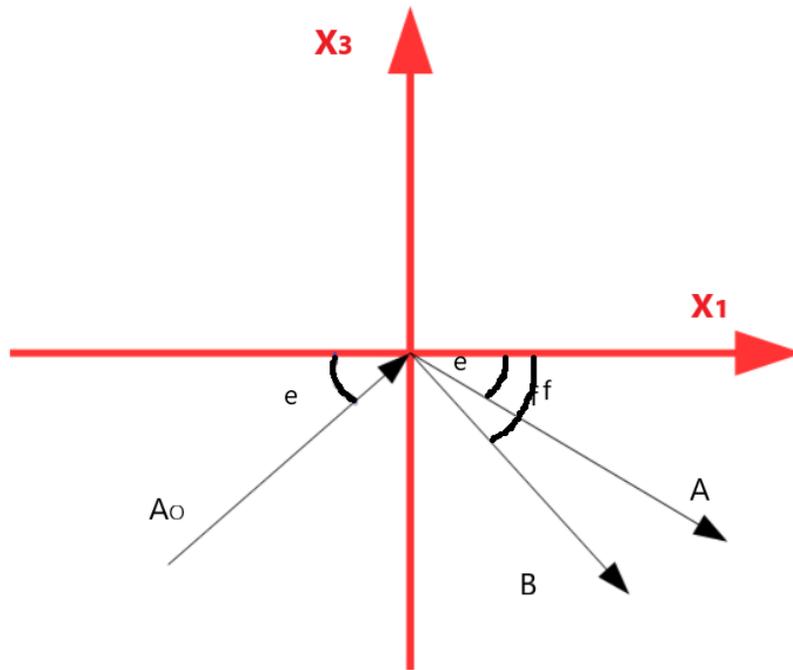


Figure 1.4: Sketch showing a P wave reflected to P and SV waves for a free surface on an elastic medium .

These waves travel with low velocity, strong amplitude, and the strongest arrivals in a record. Their velocities depend on the frequency, and they also have decreased amplitudes with depth. Rayleigh waves in a laterally homogeneous medium exist at any free surface. These waves are represented through a P/SV pair. In other words, the reflections on the surface contain both of them. When P and SV waves interact with a free surface, the horizontal slowness  $p$  only depends on  $S$ -wave velocity  $\beta$  and  $P$ -wave velocity  $\alpha$ .

Figure (1.5) shows for one wavelength the particle motion of Rayleigh waves for the fundamental mode (Udias and Bufo (2017); Shearer (2019)).

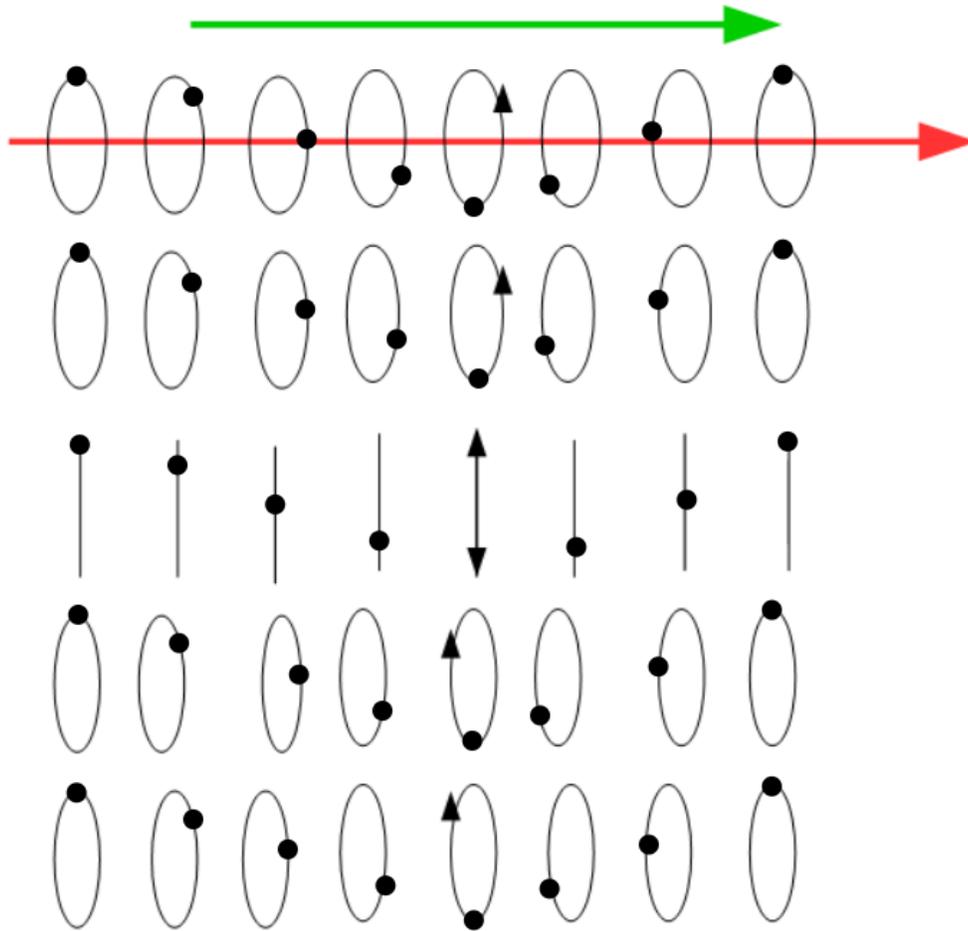


Figure 1.5: Sketch showing the particle motion for Rayleigh propagating from left to right .

The Rayleigh wave travels along the surface with phase velocity  $V_p = 1/p$  at a given frequency. However, when the phase velocity varies as a function of frequency, then group velocity  $V_g$  is used, with which the amplitude of the wave travels in space. In other words, for a sum of waves, phase velocity applies to the wave with the higher frequency and group velocity for the waves with the lowest frequency. This is a condition where the waves become dispersed, a common characteristic of ground roll.

Figure (1.6) show the definition of group and phase velocities.

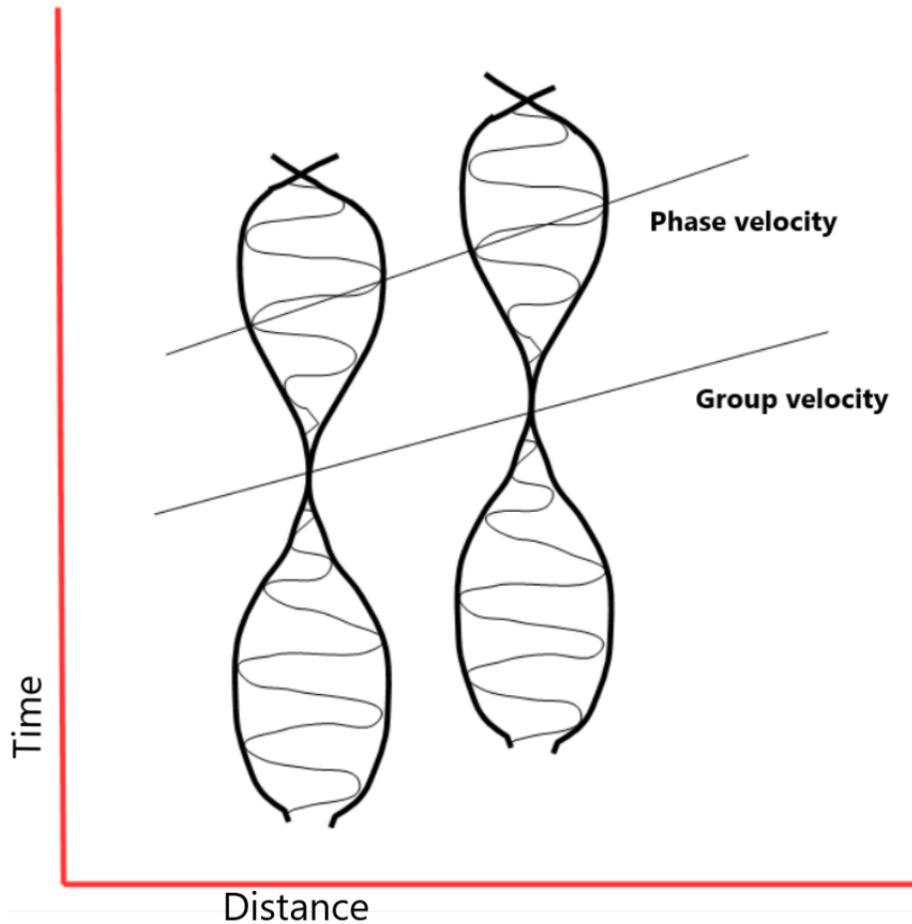


Figure 1.6: Sketch showing group and phase velocities for a sum of waves.

### 1.3 Ground roll and attenuation methods

Incoherent or random noise exists in almost every seismic record obtained from field measurements, and it can be generated by acquisition trucks, people working, wind, nearby roadways, etc. On the other hand, noise that appears with a consistent phase in seismic data is not random. It is called coherent and, in most cases, is source-generated noise, such as ground roll.

The ground roll attenuation is one of the first steps of seismic data processing of land data. In general, ground roll removal is a difficult task because it is dispersive and often aliased.

Alias is caused by insufficient receiver sampling (Duijndam et al. (2001)).

Figure (1.7) shows a synthetic shot gather where the ground roll is dispersive. The events seem linear, and their velocity change with frequency. This is one of the characteristics that makes ground roll attenuation difficult. Figure 1.8 shows in frequency-wavenumber ( $f - k$ ) domain a non-aliased signal with maximum frequency of 120 Hz. Figure 1.9 shows the same signal when it is aliased. The signal is in the frequency range of 25Hz and 75Hz shows alias. When this phenomenon happens to ground roll, it overlaps with the  $f - k$  spectrum of the reflections making frequency-wavenumber filtering a difficult task (Beresford-Smith and Rango (1989); Hosseini et al. (2015); Liu and Marfurt (2004); Liu (2010)).

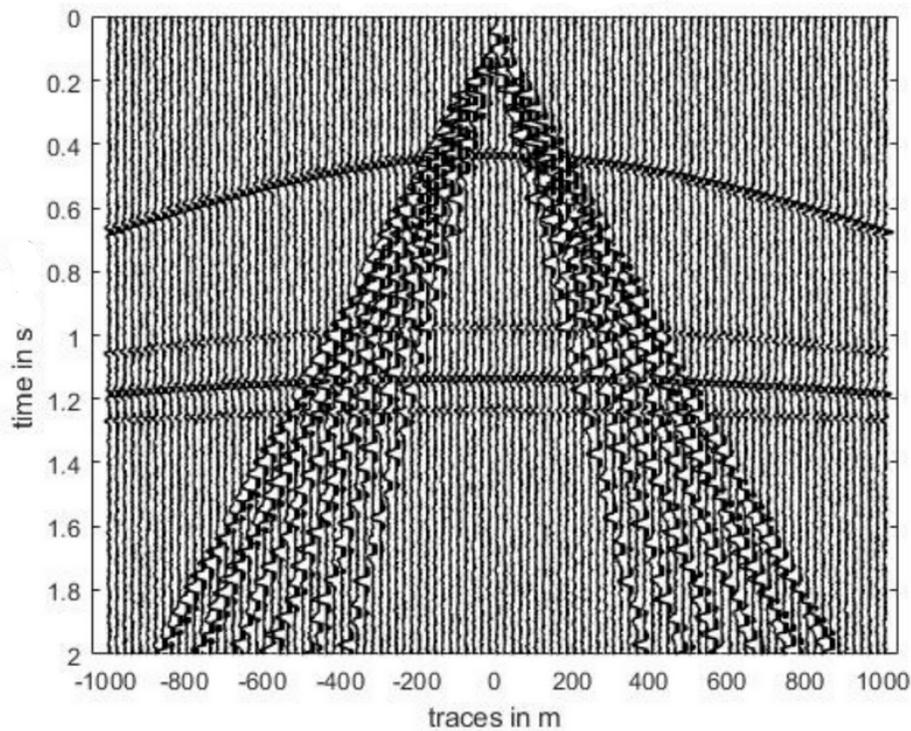


Figure 1.7: Synthetic shot gather showing dispersive ground roll masking reflections.

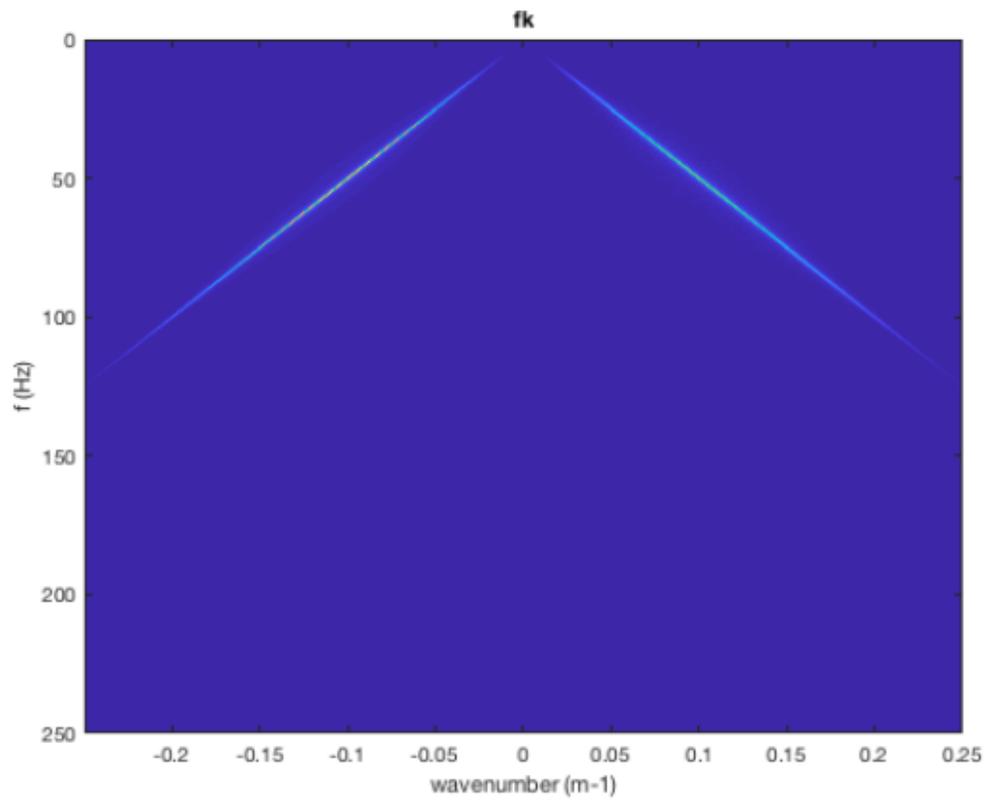


Figure 1.8: Not aliased signal in the  $f-k$  domain. A filter can remove the frequency containing the signal we wish to attenuate.

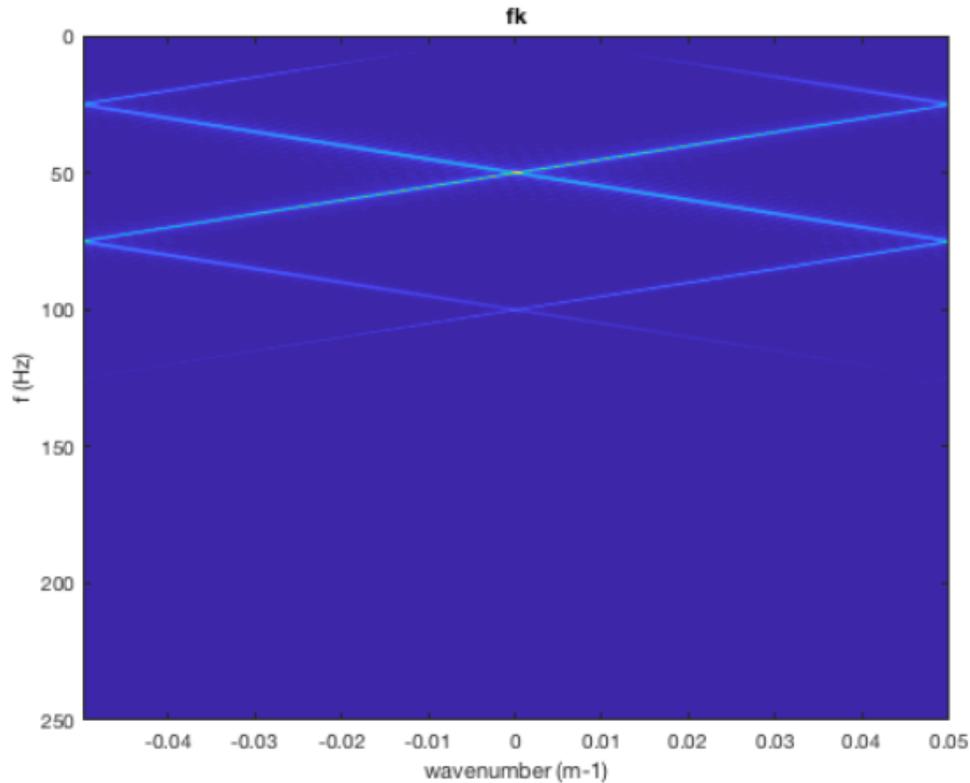


Figure 1.9: The alias phenomenon is visible when displaying a record in the  $f - k$  domain. For frequencies 25 to 125 HZ, the signal cannot be removed without affecting any other existing signal (reflections etc.)

Ground roll appears with higher energy than reflections. However, in certain spatial positions, such as for small offsets, coherent noise detection is difficult. Different methods have been developed for ground roll attenuation. Since the '80s, this task is one of the most important first steps in seismic data processing. The famous  $FK$  filtering introduced by Clement (1973), was one of the first steps for coherent noise attenuation. Later, Simaan and Love (1984) proposed an optimum filter array to remove linear noise from sections. However, the dispersive and aliased character of ground rolls demands more powerful tools. Radon Transform became popular in late '80s by Hampson (1987) (Discrete RT) and later by Fyfe and Kelamis (1992) (Linear RT) and Schonewille and Duijndam (1997) (Parabolic RT). Radon transforms in some occasions can cope with ground roll however, their parametrization to cope with dispersive signals is a difficult task and in general, Radon transform are most often adopted for multiple reflection removal Hampson (1987).

The Wavelet Transform is a well-known tool for ground roll suppression (Yu et al., 2002; de Matos and Osório, 2002; Goudarzi and Ali Riahi, 2012) because it can cope with the local dispersive character of ground roll. However, separating signals from the ground roll in the wavelet domain can become a complicated task because the wavelet coefficients that model these two signals might overlap for some wavelet scales. Other transforms, such as S and XFK (Askari and Siahkoochi, 2008), SVD (Porsani et al., 2009) have also been applied for ground removal. However, all previous methods work well if the ground roll is not aliased, and there is good separation of ground roll and reflections in the transform domain.

The Curvelet Transform has also been widely used for ground roll attenuation (Zhang et al., 2017; Naghizadeh and Sacchi, 2018, 2011). One advantage of the Curvelet transform is that it allows separating reflections from ground roll based on their local orientation (dip) and scale.

In this thesis, rather than adopting a transform-based framework for ground roll suppression, I propose an inversion approach based on regularized least-squares inversion and regularized robust inversion. To achieve this goal, I have adopted tools of sparsity promoting inversion and methodologies for the solution of linear inverse problems such as Iterative Re-weighted Least-Squares (Burrus et al. (1994); Daubechies et al. (2008); Chartrand and Yin (2008)) and the alternating direction method of multipliers (ADMM) (Boyd et al. (2011); Wang et al. (2019)).

The proposed method entails simultaneously modelling reflections and ground role via an operator in frequency-space  $f - x$  domain. The technique was inspired in the work of Perkins and Zwaan (2000), who also proposed modelling ground role and reflections in the  $f - x$  domain. However, the method I have developed has fewer restrictions than the technique of Perkins and Zwaan (2000) as I adopted a sparse regularization technique to isolate coefficients that model ground roll and reflections.

## 1.4 Scope of this thesis

### 1.4.1 Contributions

The main contributions of this thesis are summarized as follows:

- I propose an algorithm to concurrently model seismic reflections (hyperbolic events) and ground roll (dispersive linear noise) by estimating the coefficients that model hyperbolic events and coherent noise.

- The algorithm takes place in the  $f - x$  domain. Two operators are used to estimate coefficients that model hyperbolas and dispersive noise. The coefficients are calculated via inversion.
- I firstly apply least-squares inversion with quadratic and sparse regularization showing the superiority of the sparse inversion methods for this type of problem. Solutions with sparse regularization are obtained via the Iterative Reweighted Least-Squares (IRLS) method.
- I then use scaled ADMM to estimate the coefficients via robust inversion. This solver gives more flexibility than IRLS as it permits to incorporate a robust misfit measure in the inversion.
- Via synthetic examples, I show that one can separate the ground roll from reflections when an approximate model of the  $\tau, v$  pairs is provided. The couple  $\tau, v$  represents two-wave travel-time and NMO velocity of seismic reflections parametrized via classical hyperbolic assumption. These parameters are easy to obtain via velocity analysis.
- Through field data, we compare all previous methods and reach conclusions about each technique's effectiveness.

### 1.4.2 Overview

**Chapter 1** explains the basic Physics one needs to know to understand the content of this work and provides an introduction to ground roll attenuation methods reported in the literature.

**Chapter 2** explores an algorithm for generating 3D synthetic shot gathers in the  $f - x$  domain under the assumption of a layered Earth with minimal spatial structural variability, based on the work of Perkins and Zwaan (2000) and Le Meur et al. (2008). Then, I first investigate applying a  $l_1$  regularization constraint to retrieve the sparse coefficients that model signal and ground roll.

**Chapter 3** proposes three inversion algorithms to retrieve the coefficients that model reflections and ground roll. Rather than adopting a general mathematical transform inspired by the fields of harmonic analysis and image processing, we create modelling operators that resemble seismic reflections and dispersive ground roll. First, we investigate the application of a  $l_2$  regularization constraint. Then, we replace with an  $l_1$  regularization term to retrieve the sparse coefficients that model signal and ground roll by adopting the traditional approach of Iterative Reweighted Least-squares (IRLS) (Burrus et al., 1994; Daubechies et al., 2008; Chartrand and Yin, 2008) and FISTA algorithm (Beck and Teboulle (2009)). We also research the applicability of the proposed inversion algorithm to cases where we do not have a precise knowledge of the  $\tau, v$  pairs (intercept-time and RMS velocities) of the reflections. Besides, we test the efficiency of robust inversion in ground roll detection and separation from reflections using a robust algorithm using scaled ADMM (Alternating Direction Method of Multipliers). All the above algorithms are tested for their efficiency through synthetic and real data examples.

**Chapter 4** summarizes the contributions and limitations of the methods of this thesis and provides some recommendations for future steps.

---

---

## CHAPTER 2

---

### Forward Problem

#### 2.1 Introduction

In geophysics, it is a common and essential practice to reach conclusions about parameters from data. We use the laws of Physics and Geophysics to build a model based on parameters we insert. In this thesis, for example, we use the laws of wave propagation in the Earth's interior, we insert parameters such as velocities, densities of the different layers, and compute a seismic record model. This is called the forward problem.

The opposite is the inverse problem, and I will talk more about this in chapter 3. In that case, we try to reconstruct the model by using a set of measurements. We try to find the exact theory that describes how the measurements are combined to give the model that is almost similar to the data through this process. Figure 2.1 shows a simple definition of forward and inverse problems.

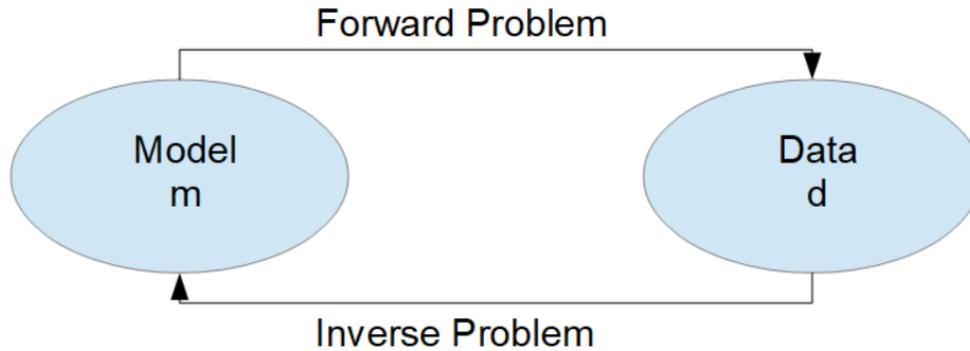


Figure 2.1: A simple definition for the forward and inverse problems in Geophysics for a model  $m$  and the data  $d$

The scope of chapter 2 is to show how the forward problem can be applied to make synthetic data, a seismic record which will contain reflections and coherent noise-ground roll.

## 2.2 Ground roll modelling

Ground roll modelling is a challenging problem. It is challenging to build a realistic model that will contain coherent noise with dispersion, with higher amplitude and lower velocity and frequency compared to the P waves.

It is tough to model ground roll and remove it from raw data by preserving the reflections' amplitudes (Perkins and Zwaan (2000); Le Meur et al. (2008)). The traditional filters where one suppresses frequencies, such as filtering in FK domain, do not work if there is aliased coherent noise. It is better to make a model that will be close enough to the original data and then subtract it from the seismic data.

One has to make a realistic near-surface Earth model using appropriate values for density,  $Q_p$ ,  $Q_s$  factors, velocities for P and S waves and thicknesses for different layers. Homogeneity or heterogeneity should also be taken into account for the basic model (Campbell et al., 2001).

However, this type of modelling, using finite differences, does not lead to a realistic and complex ground roll model, as any small change in the values of model parameters gives different data results.

Another issue is the acquisition geometry. Most of the filters work well in 2D data with a regular grid. This is because, in these cases, the ground roll is linear. However, when the

sampling is irregular, and the data set is 3D, the coherent noise loses its linearity. Figure 2.2 shows a 3D irregular data set.

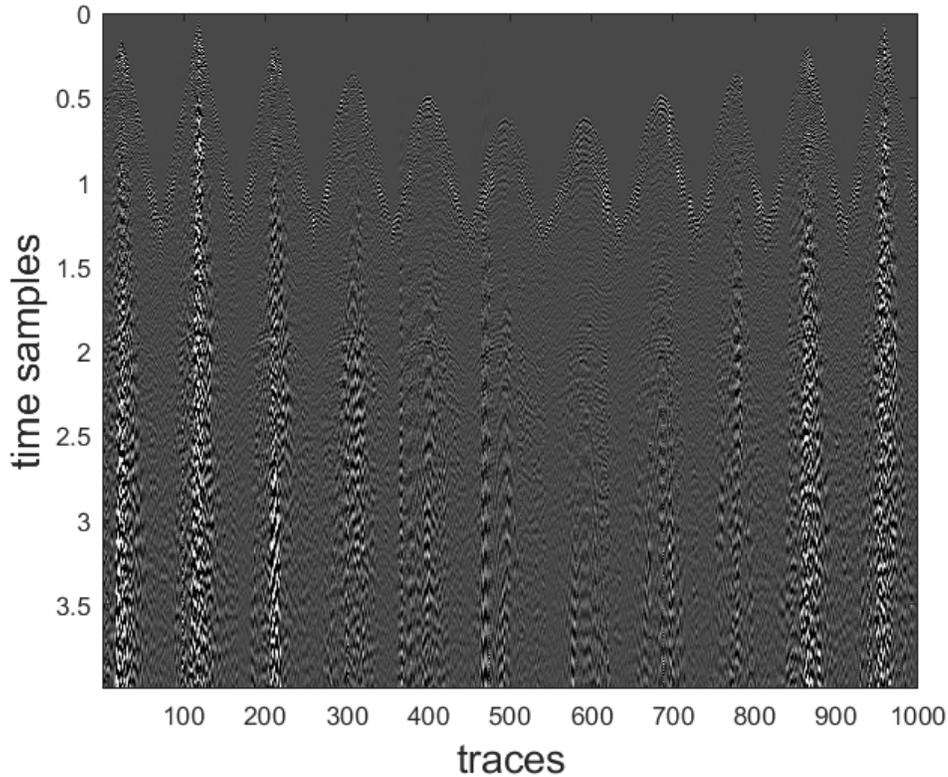


Figure 2.2: An irregular offset 3D data set with very strong ground roll. Note that in traces 450-650, the coherent noise has lost its linearity.

## 2.3 Method

To solve the forward problem, we first consider 3D irregular seismic data. We also decided to work on  $f-x$  domain, as it is easier to handle a range of frequencies each time, on which the dispersive ground roll depends on.

We assume a 3D shot gather composed of two signals, hyperbolic events and coherent noise. The coherent noise is the dispersive ground roll, and it is denoted by  $\mathbf{d}_c(t, h)$  where  $h$  is the source-receiver distance, and  $t$  is time. The signal (superposition of hyperbolic events) is denoted by  $\mathbf{d}_s(t, h)$ . The model in the time domain becomes

$$\mathbf{d}(t, h) = \mathbf{d}_s(t, h) + \mathbf{d}_c(t, h) \quad (2.1)$$

Equation 2.1 can be transformed to the frequency domain

$$\mathbf{d}(f, h) = \mathbf{d}_s(f, h) + \mathbf{d}_c(f, h) \quad (2.2)$$

where  $\mathbf{d}_s(f, h)$  is the signal and  $\mathbf{d}_c(f, h)$  is the coherent noise in the frequency-space ( $f - x$ ) domain. Clearly, we are simplifying notation by adopting  $\mathbf{d}(\cdot, h)$  for both describing data in time and frequency domain.

### 2.3.1 Modeling the signal

We create a model for the signal in the frequency domain. We consider a non-dipping layered model that consists of hyperbolas as the signal of interest. The following equation provides the travel time for the reflection hyperbola (Yilmaz (2001))

$$t(h) = \sqrt{\tau^2 + \frac{h^2}{v^2}} \quad (2.3)$$

where  $\tau$  is the zero offset travel time,  $h$  is the true shot-receiver distance, and  $v$  is the root mean square velocity. A superposition of hyperbolas can be represented in the  $f - x$  domain via the following expression

$$\begin{aligned} \mathbf{d}_s(f, h) &= \sum_{k=1}^{N_s} a(k) w_s(f) e^{-i2\pi f \sqrt{\tau_k^2 + \frac{h^2}{v_k^2}}} \\ &= \sum_{k=1}^{N_s} \mathbf{m}_s(f, k) e^{-i2\pi f \sqrt{\tau_k^2 + \frac{h^2}{v_k^2}}} \end{aligned} \quad (2.4)$$

where  $a_{s_k}$  is the amplitude of the  $k$ -th hyperbolic event,  $w_s(f)$  is the corresponding seismic wavelet in the frequency domain. Similarly to what Sacchi and Ulrych (1995) applied to their work in the frequency domain Radon transform, we can collapse the wavelet and the amplitude term into coefficients denoted by  $\mathbf{m}_s(f, k) = a(k)w_s(\omega)$ . The coefficients  $\mathbf{m}_s(k, f)$ ,  $k = 1 \dots N_s$ ,  $f \in [f_{min}, f_{max}]$  are the unknowns of our inverse problem.

We remind the reader that for a 3D acquisition, the source-receiver distance  $h$  is given by

$$h = \sqrt{(r_x - s_x)^2 + (r_y - s_y)^2} \quad (2.5)$$

where  $\vec{r} = (r_x, r_y)$  and  $\vec{s} = (s_x, s_y)$  are receiver and source coordinates, respectively. Equation 2.4 is the synthesis in  $f-x$  domain of  $N_s$  hyperbolas of parameters  $\tau_k, v_k, k = 1, \dots, N_s$ . We will analyze two cases. First, we assume known pairs  $\tau_k, v_k$ . Then, we assume a coarse regular grid of pairs  $\tau_k, v_k$  that might or might not contain an event.

### 2.3.2 Modeling the coherent noise

For simplicity, we assume ground roll modelled via the superposition of linear events via the following expression

$$\begin{aligned} \mathbf{d}_c(f, h) &= \sum_{k=1}^{N_c} a_c(k) w_c(f) e^{-i2\pi f p_k h} \\ &= \sum_{k=1}^{N_c} \mathbf{m}_c(f, k) e^{-i2\pi f p_k h} \end{aligned} \quad (2.6)$$

where  $a_c(k)$  is the amplitude of the  $k$ -th linear event,  $w_c(f)$  is the corresponding wavelet,  $p$  is the ray parameter. The coefficients  $\mathbf{m}_c(f, k), k = 1 \dots N_c, f \in [f_{min}, f_{max}]$  are also unknown.

## 2.4 Synthetic modeling

To produce a synthetic model, we need to run the forward operator using inserted physical values. In the end, we will be able to estimate the amplitude and the wavelet of the signals and build a model similar to a real one. For this step, we use the simple least-squares inversion.

For a single frequency  $f$ , we first define the observed data in vector form

$$\mathbf{d}(f) = \begin{pmatrix} d(f, h_1) \\ d(f, h_2) \\ d(f, h_3) \\ \vdots \\ d(f, h_N) \end{pmatrix} \quad (2.7)$$

where  $N$  is the number of receivers. Similarly, we define vectors of complex amplitude coefficients for the signal and coherent noise

$$\mathbf{m}_s(f) = \begin{pmatrix} m_s(f, 1) \\ m_s(f, 2) \\ m_s(f, 3) \\ \vdots \\ m_s(f, N_s) \end{pmatrix}, \mathbf{m}_c(f) = \begin{pmatrix} m_c(f, 1) \\ m_c(f, 2) \\ m_c(f, 3) \\ \vdots \\ m_c(f, N_c) \end{pmatrix}. \quad (2.8)$$

Equation 2.2 can be expressed as follows

$$\mathbf{d}(f) = \mathbf{A}_s(f) \mathbf{m}_s(f) + \mathbf{A}_c(f) \mathbf{m}_c(f) + \mathbf{e}(f), \quad (2.9)$$

where to make our model more realistic, we have included an additive noise term  $\mathbf{e}(f)$ . We can simplify notation by omitting the dependency on  $f$  and understand that the inversion and the synthesis of signals must be carried out for a predefined frequency range  $f \in [f_{min}, f_{max}]$ ,

$$\mathbf{d} = \mathbf{A}_s \mathbf{m}_s + \mathbf{A}_c \mathbf{m}_c + \mathbf{e}. \quad (2.10)$$

We now define the following augmented matrix and vector

$$\mathbf{A} = \begin{pmatrix} \mathbf{A}_s & \mathbf{A}_c \end{pmatrix}, \mathbf{m} = \begin{pmatrix} \mathbf{m}_s \\ \mathbf{m}_c \end{pmatrix} \quad (2.11)$$

which can be combined into a single expression

$$\mathbf{d} = \mathbf{A} \mathbf{m} + \mathbf{e}. \quad (2.12)$$

The form of equation 2.12 is to solve linear problems. Our goal is to adopt regularized inversion to estimate  $\mathbf{m}$ . We then split the augmented vector of coefficients into  $\mathbf{m}_s$  and  $\mathbf{m}_c$ . The coefficients are used to synthesize  $\mathbf{d}_s$  and  $\mathbf{d}_c$ . Finally, the process entails mapping the separated data from  $f - x$  to  $t - x$ .

## 2.5 Synthetic Data in FX domain

We model our data via a superposition of hyperbolic events and two dispersive modes with phase velocity given by

$$v(f) = v_{min} + \frac{(v_{max} - v_{min})}{\sqrt{1 + (\frac{f}{f_0})^4}}. \quad (2.13)$$

The following parameters have been adopted:

- 1st dispersive mode:  $v_{min} = 400m/s, v_{max} = 900m/s, f_0 = 10Hz$
- 2nd dispersive mode:  $v_{min} = 600m/s, v_{max} = 900m/s, f_0 = 20Hz$
- Central frequency of ground roll wavelet is  $10Hz$

The central frequency of the signal (hyperbolic events) is  $20Hz$ . We have used six hyperbolic events with intercept-velocity pairs  $\tau = [0.3, 0.39, 0.5, 0.6, 0.83, 1.2]s$  and  $v = [2.0, 2.4, 3.0, 3.4, 3.4, 4.0]km/s$ . In our first numerical experiment we will assume that the  $\tau, v$  pairs (intercept time-velocity pairs) are known.

Figure 2.3 shows the distribution of receivers for the 3D synthetic shot gather used to test our algorithms. Data consist of  $N_x = 50$  receivers per line and  $N_y = 5$  receivers lines, the length of each record in 400 samples and the sampling interval is  $\Delta t = 4ms$ . The additive noise in equation (2.10) is Gaussian, uncorrelated in space and band-limited in time. The variance of the noise is  $\sigma^2 = 0.07$  which corresponds to an  $SNR = 1$ .<sup>1</sup>

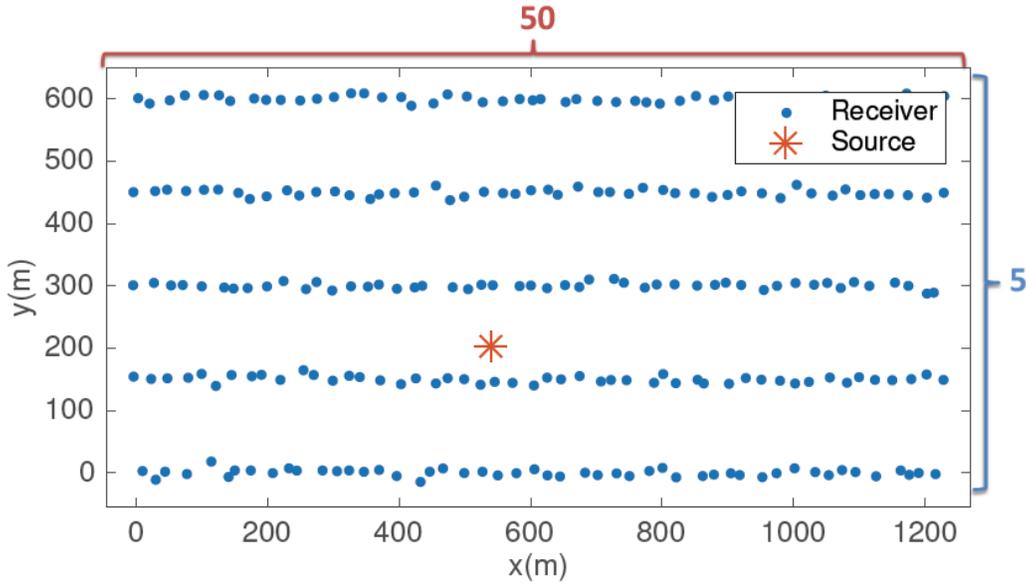


Figure 2.3: Source and receivers for the 3D synthetic example adopted to test our signal and ground roll separation method.

<sup>1</sup> $SNR = \frac{\text{Power of the clean data}}{\text{Power of the noise}}$

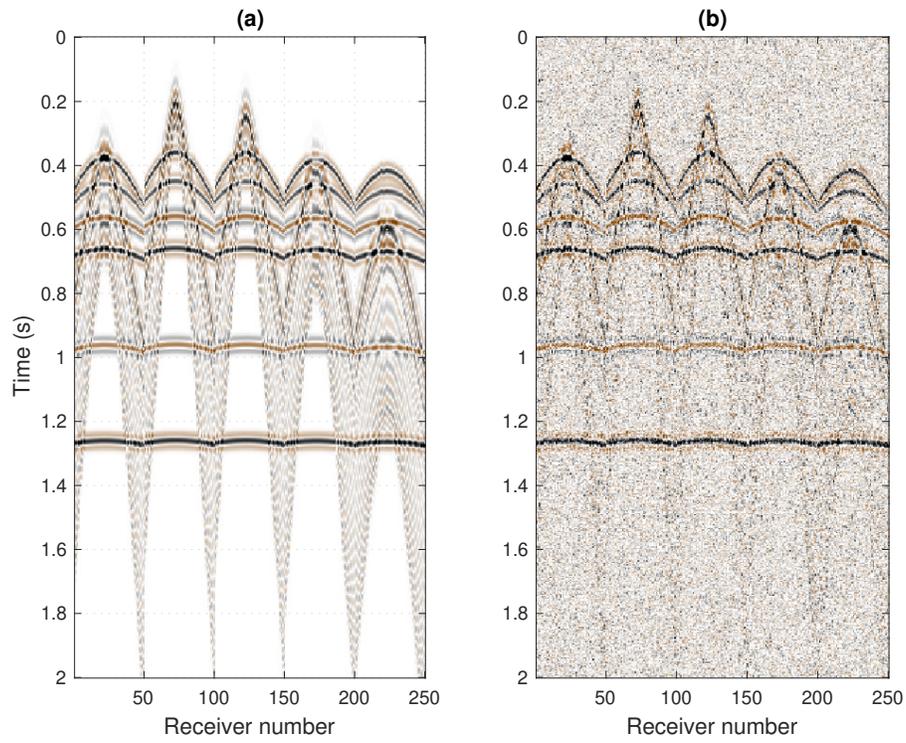


Figure 2.4: a) 3D synthetic shot gather. b) 3D synthetic shot gather contaminated with random noise ( $SNR = 1$ ).

## 2.6 Summary

In general, generating a realistic synthetic coherent noise model is very difficult. The near-surface (top 10-30 m) has very slow velocities, substantial heterogeneity, and increased velocity with depth, causing strong near-surface guided waves and strong attenuation.

We have developed a simple, realistic algorithm for simultaneously modelling reflections and linear dispersive ground roll. We built two operators for hyperbolas and dispersive noise, and we modelled a synthetic 3D shot gather in the  $f - x$  domain. This is a simple approach inspired by Perkins and Zwaan (2000) and Le Meur et al. (2008) that allows us to develop a simple, practical model to be used in the least-squares inversion.

---

---

## CHAPTER 3

---

### Inversion Algorithms

#### 3.1 Introduction

In the previous chapter, we discussed the forward and the inverse problem in Geophysics. There are many algorithms to solve linear problems through inversion.

In this thesis, we apply algorithms with simple mathematical procedures that we believe can give satisfying results.

In this chapter, we introduce the reader to Least-Squares and Robust inversion.

#### 3.2 Least-Squares inversion algorithms

Least-squares inversion is a simple mathematical procedure that gives solutions to linear problems. The cost function we minimize for the least-squares inversion only contains a misfit function

$$J = \min \|\mathbf{A}\mathbf{m} - \mathbf{d}\|_2^2. \quad (3.1)$$

We must add a quadratic regularization term for the stability of the solution (Tikhonov and Arsenin, 1977; Menke, 1989). Hence, the cost function we have to minimize given by

$$J = \min \|\mathbf{A}\mathbf{m} - \mathbf{d}\|_2^2 + \mu \|\mathbf{m}\|_2. \quad (3.2)$$

The minimum of the latter is often called the damped least-squares solution

$$\hat{\mathbf{m}} = (\mathbf{A}^* \mathbf{A} + \mu \mathbf{I})^{-1} \mathbf{A}^* \mathbf{d}, \quad (3.3)$$

where  $\mu$  is the trade-off parameter (more details in chapter 4).

The Least-squares solution is straightforward. It turns out to be statistically correct if data errors are normally distributed. However, this solution is negatively affected by outliers because they distort the distribution of errors (Russell (1988); Schuster (2017); Kijko (1994)). When we use least-squares fitting, the outliers are considered correct observations and the resulting regression is affected. To solve this problem, we can adopt a robust regression algorithm. Figure 3.1 shows a simple graph showing the behaviour of robust and non-robust algorithms in the presence of outliers.

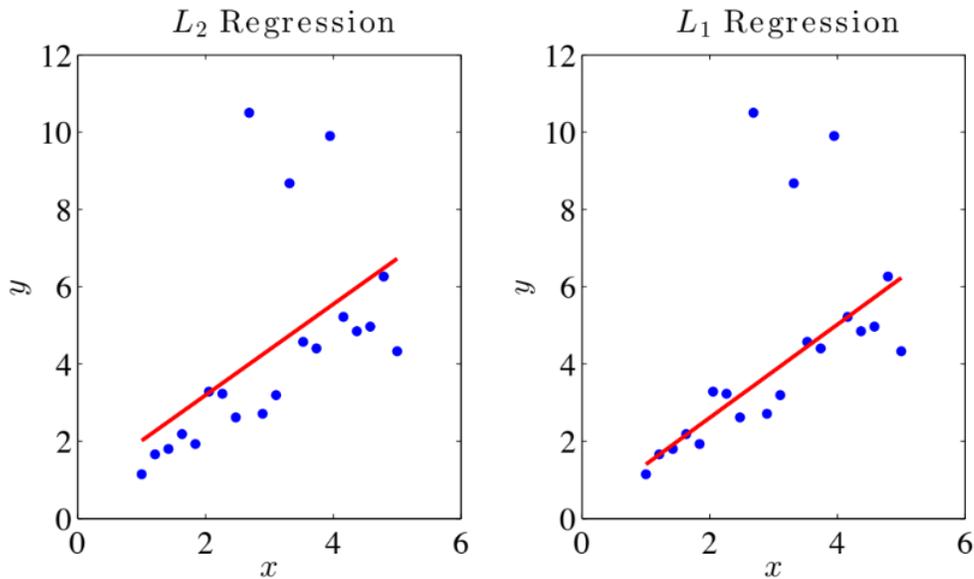


Figure 3.1: Comparison of non robust and robust regression when outliers are present. Non robust algorithms consider them as good observations (left image). Robust algorithms are not affected by them (right image). (Adapted by Sacchi, 2015.)

The solution of the least-squares fitting problem with an  $l_1$  regularization leads to sparse solutions. This technique has been applied for the computation of the so-called Sparse (High-Resolution) Radon Transform (Sacchi and Ulrych, 1995; Wang and Nimsaila, 2014), Fourier reconstruction methods with sparsity constraints (Sacchi and Ulrych, 1996; Zwartjes and Gisolf, 2007) and regional-residual separation of potential field data (Guspi and Introcaso, 2000).

In this case, an alternative solution is the  $\ell_1$  norm solution. That means, in our least-squares misfit, we add a sparse regularization norm which entails adding to the misfit the  $\ell_1$  penalty term

$$J = \mathbf{min} \|\mathbf{A}\mathbf{m} - \mathbf{d}\|_2^2 + \mu \|\mathbf{m}\|_1. \quad (3.4)$$

The solution of the latter is compute via iterative algorithms (Berman and Plemmons (1974); Shen et al. (2011); Bing et al. (1992)) such as IRLS

$$\mathbf{m} = (\mathbf{A}^* \mathbf{A} + \mu \mathbf{Q})^{-1} \mathbf{A}^* \mathbf{d}. \quad (3.5)$$

where  $\mathbf{Q}$  is also diagonal and given by

$$Q_{ii} = \frac{1}{|m_i| + \epsilon}. \quad (3.6)$$

All the above research has been done with the assumption that additive incoherent noise is Gaussian. However, it is highly probable that incoherent noise will be non-Gaussian erratic noise (Anstey (1986)). This leads to the requirement of a robust measure of misfit, which results in formulating a robust cost function in the next section.

### 3.3 Robust inversion algorithms

To completely ignore the influence of outliers in the solution of the linear regression problem, we propose to minimize the following cost function

$$J = \mathbf{min} \|\mathbf{A}\mathbf{m} - \mathbf{d}\|_1, \quad (3.7)$$

which is also called the robust misfit.

The minimum of  $J$  leads to a solution that is more resistant to outliers because it does not square each of the terms in the misfit measure. On the other hand, the solution, which minimizes the robust misfit, is more complicated to be computed. The functional  $f$  of the residuals  $r$  is given by

$$f(\mathbf{m}) = \|\mathbf{r}\|_1 = \sum |\mathbf{r}_i|. \quad (3.8)$$

Hence the the gradient of  $f$  is

$$\nabla f(\mathbf{m}) = \mathbf{A}^T \mathbf{R} \mathbf{r} = \mathbf{A}^T \mathbf{R} (\mathbf{d} - \mathbf{A} \mathbf{m}), \quad (3.9)$$

where  $R$  is a diagonal matrix with elements that are the absolute values of the residuals. To find the  $\ell_1$  solution, we solve  $\nabla f(\mathbf{m}) = \mathbf{0}$  and we have

$$\mathbf{A}^T \mathbf{R}(\mathbf{d} - \mathbf{A}\mathbf{m}) = \mathbf{0} \quad (3.10)$$

or

$$\mathbf{A}^T \mathbf{R}(\mathbf{d} - \mathbf{A}\mathbf{m}) = \mathbf{A}^T \mathbf{R} \mathbf{d} \quad (3.11)$$

Since  $\mathbf{R}$  depends on  $\mathbf{m}$ , 3.11 is a nonlinear system of equations and we need an iterative method to find the appropriate weights in  $\mathbf{R}$ .

Then for more stability, we add an  $l_1$  regularization term and we have

$$J = \min \|\mathbf{A}\mathbf{m} - \mathbf{d}\|_1 + \mu \|\mathbf{m}\|_1. \quad (3.12)$$

For our sparse inversion experiment, the cost function we have to minimize is

$$J = \operatorname{argmin}(\|\mathbf{e}\|_1 + \mu \|\mathbf{m}\|_1), \quad (3.13)$$

where  $\mathbf{e} = \mathbf{A}\mathbf{m} - \mathbf{d}$  and it contains erratic noise.

To solve equation 3.13, we need to use an iterative method. The steps are not so easy, as in the case of the least-squares inversion with a  $L - 1$  norm. In chapter 4, we adopt the algorithm ADMM to solve the problem above.

### 3.4 Summary

To solve linear problems in Geophysics, we use inversion algorithms.

Least squares inversion is a simple mathematical procedure which can give satisfying results.

To make the solution stable, one must add a penalty term.

However, we need to consider the behaviour of the algorithms in case the observations contain outliers. In this case, quadratic regression will consider them as useful observations and fitting the data will be wrong. Considering a robust regression will avoid fitting outliers.

---

---

## CHAPTER 4

---

### Inverse Problem

#### 4.1 Introduction

In this chapter, we describe different algorithms to separate ground roll from reflections. A comparison between least-squares inversion with penalty terms ( $L_2$  and  $L_1$ ) and robust inversion through synthetic and field data shows when a coherent noise attenuation is possible.

#### 4.2 Least-Squares inversion with quadratic regularization

Having read the work of Perkins and Zwaan (2000) and Le Meur et al. (2008) who used least-squares inversion for ground roll attenuation and seeing that their results were promising, I decided to expand their research by exploring regularization methods for the solution of inverse problems.

Our first algorithm is the simple least-squares inversion which now utilizes an  $l_2$  norm to measure the modeled and observed data's fidelity. The stability of the solution is guaranteed by adding the  $l_2$  model norm as a penalty constraint (Tikhonov and Arsenin, 1977; Menke, 1989)

$$J = \|\mathbf{d} - \mathbf{A}\mathbf{m}\|_2^2 + \mu \|\mathbf{m}\|_2^2. \quad (4.1)$$

To compute the coefficients that minimize  $J$ , we take the derivative of the cost function with respect to  $\mathbf{m}$  and then equate the result to zero. After some simple algebra, we arrive at the damped least-squares solution

$$\hat{\mathbf{m}} = (\mathbf{A}^* \mathbf{A} + \mu \mathbf{I})^{-1} \mathbf{A}^* \mathbf{d}. \quad (4.2)$$

The solution  $\hat{\mathbf{m}}$  can be separated into two parts  $\hat{\mathbf{m}} = (\hat{\mathbf{m}}_s, \hat{\mathbf{m}}_c)^T$ . The recovered amplitudes are finally used to synthesize signal and coherent noise

$$\hat{\mathbf{d}}_s = \mathbf{A}_s \hat{\mathbf{m}}_s \quad (4.3)$$

$$\hat{\mathbf{d}}_c = \mathbf{A}_c \hat{\mathbf{m}}_c. \quad (4.4)$$

In general, we will subtract the modeled coherent noise  $\hat{\mathbf{d}}_c$  from the original data. The latter is a common practice for multiple suppression.

The size of the matrix  $\mathbf{A}$  is  $N \times (N_s + N_c)$ . According to equation (4.2), the inversion of a matrix of size  $(N_s + N_c) \times (N_s + N_c)$  is required for each frequency  $f \in [f_{min}, f_{max}]$ . If the number of traces  $N < (N_s + N_c)$  (underdetermined system of equations) we prefer to replace equation (4.2) by

$$\hat{\mathbf{m}} = \mathbf{A}^* (\mathbf{A} \mathbf{A}^* + \mu \mathbf{I})^{-1} \mathbf{d}. \quad (4.5)$$

It is easy to show that equations (4.2) and (4.5) are equivalent (Sacchi and Ulrych, 1995). For the underdetermined problem, we prefer equation (4.5) because it requires the inversion of a matrix of size  $N \times N$ .

### 4.2.1 Parameters' selection

The trade-off parameter  $\mu$  in expressions (4.1) and later in (4.6), is required to avoid over-fitting or under-fitting the data.

A small value of  $\mu$  leads to a solution that fits the additive noise and, consequently, the solution becomes unstable. On the other hand, a large trade-off parameter  $\mu$  will lead to under-fitting. In this case, the separated signals  $\mathbf{d}_s$  and  $\mathbf{d}_c$  will not be adequately represented. In our synthetic examples, we have adopted the  $\chi^2$  criterion to find an optimum trade-off parameter  $\mu$ .

We try different values of trial trade-off parameters and evaluate the global misfit function in the  $t, h$  domain. For instance, if the original signal is denoted by  $d(t, h)$  and the predicted signal containing both reflections and ground roll is  $\hat{d}(t, h)$ , we compute  $\chi^2$  as follow

$$\chi^2 = \frac{\sum_{t,h} |d(t,h) - \hat{d}(h,t)|^2}{\sigma^2}$$

we choose  $\mu$  such that  $\chi^2 \approx N^o$  where  $N^o$  is the total number of observations  $N_o = N_t \times N_{traces}$ .

Figure 4.1 show an example of how we use the model norm and misfit to select the best  $\mu$  among a a given range.

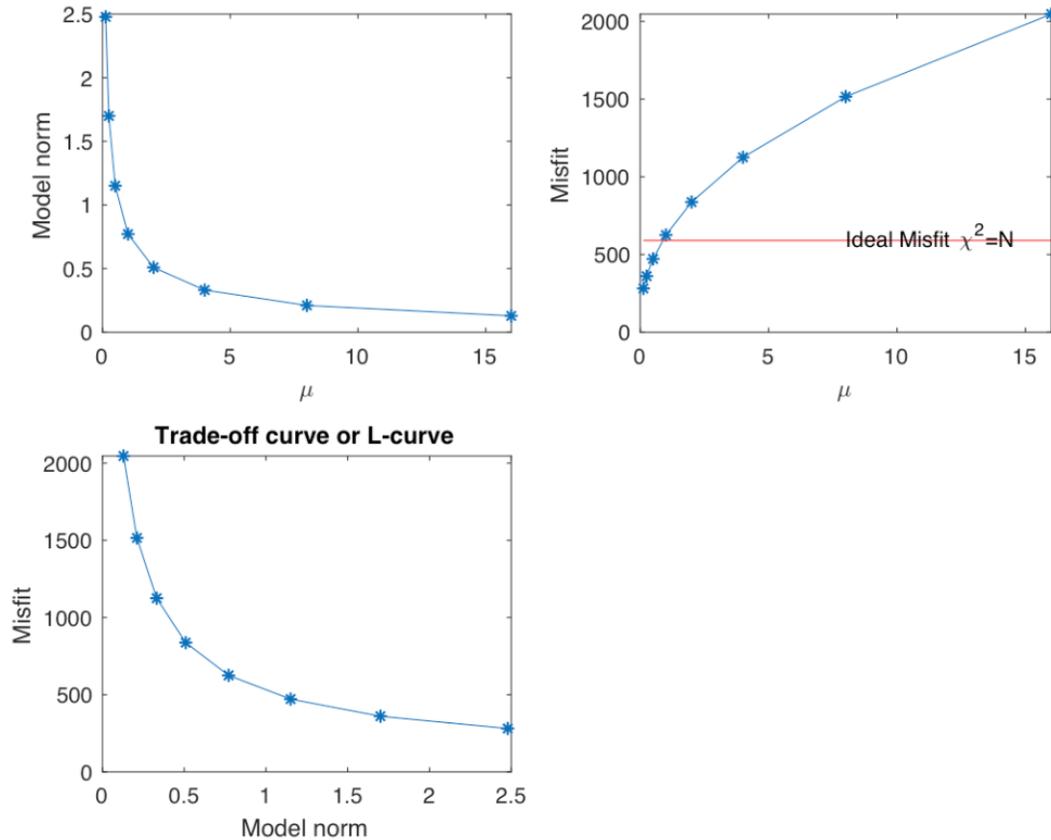


Figure 4.1: An example of how to estimate the  $\mu$  parameter using the L-curve method

We must mention that a more formal treatment of the parameter selection problem will require us to run the  $\chi^2$  test for each frequency  $f$  where the inversion is carried out. In essence, the parameter  $\mu$  should be frequency dependent. We have found that adopting a single parameter  $\mu$  for all frequencies is a reasonable and practical solution for our synthetic tests. The latter should not prevent defining a strategy to select  $\mu(f)$  to selectively damp

frequency components according to their *SNR*.

### 4.3 Least-Squares inversion with sparse regularization

In our second algorithm, we use least squares inversion, but we now use a sparse ( $L_1$ ) penalty constraint to avoid deal with outliers.

We then propose to minimize the following cost function

$$J = \|\mathbf{d} - \mathbf{A}\mathbf{m}\|_2^2 + \mu \|\mathbf{m}\|_1. \quad (4.6)$$

Different methods were proposed to minimize  $J$ . We will firstly adopt the traditional approach of Iterative Reweighed Least-squares (IRLS) (Burrus et al., 1994; Daubechies et al., 2008; Chartrand and Yin, 2008) and then we will try the Fast Iterative Shrinkage-Thresholding Algorithm for Linear Inverse problems (FISTA) (Beck and Teboulle, 2009) but bear in mind that other methods exist to minimize  $J$  (Tropp and Gilbert, 2007; Van den Berg and Friedlander, 2011).

#### 4.3.1 Inversion through IRLS and FISTA algorithms

##### IRLS

The derivative of the cost function (equation 4.6) can be approximated by

$$\nabla_{\mathbf{m}} J = \mathbf{A}^*(\mathbf{A}\mathbf{m} - \mathbf{d}) + \mathbf{D} = 0 \quad (4.7)$$

where  $\mathbf{D}$  is a diagonal matrix with elements given by  $D_{ii} = \text{sign}(m_i)$ . The sign function and, consequently,  $\mathbf{D}$ , can be approximated by the following expression

$$D_{ii} = \frac{m_i}{|m_i| + \epsilon} \quad (4.8)$$

where  $\epsilon$  is small number. In matrix form

$$\mathbf{D} = \mathbf{Q}\mathbf{m} \quad (4.9)$$

where  $\mathbf{Q}$  is also diagonal

$$Q_{ii} = \frac{1}{|m_i| + \epsilon}. \quad (4.10)$$

In essence, the condition for the minimum of  $J$  is given by

$$\nabla_{\mathbf{m}} J = \mathbf{A}^*(\mathbf{A}\mathbf{m} - \mathbf{d}) + \mathbf{Q}\mathbf{m} = 0, \quad (4.11)$$

from where one can write the following solution

$$\mathbf{m} = (\mathbf{A}^*\mathbf{A} + \mu\mathbf{Q})^{-1}\mathbf{A}^*\mathbf{d}. \quad (4.12)$$

Now, an important comment is in order. The diagonal matrix  $\mathbf{Q}$  is a function of the solution  $\mathbf{m}$ . Therefore, one must adopt an iterative algorithm to solve for  $\mathbf{m}$

$$\mathbf{m}^k = (\mathbf{A}^*\mathbf{A} + \mu\mathbf{Q}^{k-1})^{-1}\mathbf{A}^*\mathbf{d}$$

where  $k$  denotes iteration number. The IRLS algorithm often converges in about 10 iterations. IRLS is not restrictive to inversions with an  $l_1$  regularization term; it can also be adopted for least-squares inversion with the Cauchy, Huber and Geman-McClure sparsity criteria (Sacchi, 1997). Last, we mention that for the underdetermined problem (Sacchi and Ulrych, 1995) ( $N < (N_s \times N_c)$ ) we prefer to solve

$$\hat{\mathbf{m}} = \mathbf{P}\mathbf{A}^*(\mathbf{A}\mathbf{P}\mathbf{A}^* + \mu\mathbf{I})^{-1}\mathbf{d}, \quad (4.13)$$

where  $\mathbf{P} = \mathbf{Q}^{-1}$ . We also mention that equation (4.13) is also adopted by the re-weighted minimum norm signal reconstruction algorithm called FOCUSS (Gorodnitsky and Rao, 1997).

### FISTA

From the cost function (equation 4.6), if  $f(x) = \|Am - d\|^2$  and  $g(x) = \lambda\|m\|_1$ , then according to the basic iterative shrinkage method

$$m_{k+1} = \tau_{\lambda t}(m_k - 2tA^T(Am_k - d)) \quad (4.14)$$

where  $t$  is the appropriate stepsize and  $\tau$  is the shrinkage operator. Then, the stepsize can be defined by

$$t = \frac{1}{L(f)} \quad (4.15)$$

where  $L(f)$  is a Lipschitz constant of  $\nabla f$ .

For  $y_1 = m_o$  and  $t_1 = 1$  and for  $k$  iterations,

$$m_k = p_L(y_k) \quad (4.16)$$

where  $p_L$  is the iterative shrinkage operator. Also,

$$t_{k+1} = \frac{1 + \sqrt{1 + 4t_k^2}}{2} \quad (4.17)$$

and

$$y_{k+1} = m_k + \frac{t_{k-1}}{t_{k+1}}(m_k - m_{k-1}) \quad (4.18)$$

are needed for the FISTA algorithm. A more detailed description is provided by Beck and Teboulle (2009). The new aspect of the FISTA method is that the shrinkage operator in the solution showed in (4.16) is not based on the previous  $m_{k-1}$ , but on  $y_k$ .

### 4.3.2 Parameters' selection

To compute least-squares solutions with  $l_1$  penalty constraints using the iterative algorithms described before, a number of iterations should be provided in case it is difficult to define a number.

For the  $l_2$  norm regularization there is a closed-form solution. However, for IRLS and FISTA, we need iterative methods. IRLS is a simple algorithm that needs 10-20 iterations to converge depending on the data set to be inverted. In most cases, less than 20 iterations are enough. On the other hand, FISTA is more complex and needs more iterations. Stopping after a few iterations does not leave to solutions that are sparse and fit the data.

There are many definitions for a stopping criterion (Ascher and Roosta-Khorasani, 2016). Given a tolerance, a small number of  $10^{-2} - 10^{-5}$ , we force the algorithm to stop at such a point where the difference between two points of the cost function will be close to zero or, in other words, smaller than the given tolerance. Given a large tolerance, the iterative method will stop before convergence. For a small tolerance, the algorithm will reach a maximum iterations or stop at numerical convergence. If the cost function is  $\mathbf{J}$  and for  $k$  iterations we have  $J(k)$ , then we propose following stopping criterion

$$\frac{\|J_k - J_{k+1}\|}{\|J_{k+1}\|} < tol \quad (4.19)$$

where, for iteration  $k = 1$ ,  $\mathbf{m} = \mathbf{0}$  and  $J_{(1)} = \|\mathbf{A}\mathbf{m} - \mathbf{d}\|_2^2 = \|\mathbf{d}\|_2^2$ .

In addition, the trade-off parameter is still computed using the  $L$ -curve method, as previously described.

## 4.4 Robust inversion

Until now, we were assuming that the data noise is Gaussian. To further expand our research, we accept that land data are contaminated with erratic noise due to field operations, wind, debris produce by explosive sources, and in general, small near-surface heterogeneities. In this case, it is more logical to use robust inversion and rather than making the Gaussian assumption. For our sparse inversion experiment, the cost function we have to minimize is given by

$$J = \operatorname{argmin}(\|\mathbf{e}\|_1^1 + \mu\|\mathbf{m}\|_1) \text{ s.t. } \mathbf{e} = \mathbf{d} - \mathbf{A}\mathbf{m} \quad (4.20)$$

To solve this problem, we decided to use the Alternating Direction Method of Multipliers (ADMM), an Augmented Lagrangian method for solving constrained optimization problems. It is an iterative algorithm where the problem to be solved is distributed into smaller subproblems and at the same time is simple and powerful. Boyd et al. (2011) described the ADMM algorithm and its benefits as “an algorithm that is intended to blend the decomposability of dual ascent with the superior convergence properties of the method of multiplier.”

### 4.4.1 The ADMM algorithm

Much research has been done for algorithms that could handle large problems while they offer excellent optimization. It acts like a decomposition procedure, in which the solutions for the large problem is given by solving small subproblems.

ADMM combines the benefits of dual decomposition and augmented Lagrangian methods for constrained optimization. It has similar characteristics as many other algorithms, such as Douglas-Rachford splitting from numerical analysis (Eckstein and Bertsekas, 1992), (Boyd et al., 2011), Spingarn’s method of partial inverses (Spingarn, 1985), Dykstra’s alternating projections method (Boyle and Dykstra, 1986), and Bregman iterative algorithms for  $l_1$

problems in signal processing (Goldstein and Osher, 2009), (Yin et al., 2008). This algorithm is not new, and it has been used over the decades in different fields.

The algorithm for robust inversion is organized as follows (Wang et al., 2019):

$$(\mathbf{e}^{k+1}) = \operatorname{argmin}(\|\mathbf{e}\|_1 + \frac{\rho}{2}\|\mathbf{e} - \mathbf{A}\mathbf{m}^k + \mathbf{d} - \mathbf{y}^k\|_2^2) \quad (4.21)$$

$$(\mathbf{m}^{k+1}) = \operatorname{argmin}(\mu\|\mathbf{m}\|_1 + \frac{\rho}{2}\|\mathbf{e}^k - \mathbf{A}\mathbf{m} + \mathbf{d} - \mathbf{y}^k\|_2^2) \quad (4.22)$$

$$\mathbf{y}^{k+1} = \mathbf{y}^{k+1} - [\mathbf{e}^{k+1} - (\mathbf{d} - \mathbf{A}\mathbf{m}^{k+1})] \quad (4.23)$$

where  $k$  is the number of iterations and  $\rho$  is a positive number.

#### 4.4.2 Parameters' selection

To compute the solution with robust inversion using the iterative algorithms described before, maximum number of iterations should be provided. However it is hard to define this number to which the inversion will give the most accurate solution.

There are many definitions for a stopping criterion (Ascher and Roosta-Khorasani, 2016). Given a tolerance, a small number  $10^{-2} - 10^{-5}$ , we force the algorithm to stop in such a point where the difference between two points of the cost function will be close to zero or in other words smaller than the given tolerance. Similar to previous sections we define the following stopping criterion

$$\frac{\|J_k - J_{k+1}\|}{\|J_{k+1}\|} < tol \quad (4.24)$$

where for iteration  $k = 1$ ,  $\mathbf{m} = 0$  and  $J(1) = \|\mathbf{A}\mathbf{m} - \mathbf{d}\|_2^2 = \|\mathbf{d}\|_2^2$

It is important to mention here that the trade-off parameter  $\mu$  is not computed with the method of L-curve. We found that allowing the user to insert a constant value works very well for the inversion.

However, the algorithm is not too sensitive to the parameter  $\mu$ . Instead, one needs to be careful with the selection of the penalty parameter  $\rho$ .

One can say that the convergence can be improved by selecting a different parameter for each iteration, but this is hard to prove. Theory can still support the choice of a fixed value after a number of iterations. Boyd et al. (2011) propose a method to compute the value

of the penalty parameter. In our work, we insert a fixed  $\rho$ , which does not change while the algorithm iterates and the results we estimate are optimal enough for processing noise seismic data.

## 4.5 Comparison of convergence for IRLS, FISTA, and ADMM

In previous sections, we talked about the number of iterations through which the algorithm provides the best optimum solution. It is difficult to define a number of iterations for large problems because we are not sure if the solution will be optimal. As mentioned in previous sections, we use the stopping criterion, where we define a range of error by giving a tolerance. However, we need to be very careful when choosing the value as in an opposite case, the algorithm will stop earlier than the best result or later, where the solution does not dramatically change, and the computational cost is increased.

IRLS is a simple algorithm, and usually, about 10-20 iterations are needed to converge. However, FISTA and ADMM are more complex, and it is hard to define the maximum number of iterations.

According to Chambolle and Dossal (2015), FISTA algorithm is based on a simple step with a varying parameter, and many choices of parameters roughly lead to the same grade of convergence. It is not easy to experimentally define the convergence of the algorithm, and more research needs to be done to prove the convergence of the FISTA algorithm. The step  $\alpha$  seems to play a small role there, but it does not offer any dramatic changes to the process and a value greater than two works. In our work, we calculate the step alpha to be  $\geq \max(\text{eigenvalue}(\mathbf{A}^H \mathbf{A}))$  to guarantee the convergence of our algorithm. We then apply a stopping criterion to ensure that our result will be close to the optimal solution without increasing the computational cost.

Research has shown that convergence can be very slow when trying to reach high accuracy. In our work, we adopt the convergence criterion where when the difference of 2 points of the cost function reaches a value close to zero, the algorithm stops. We have also designed our algorithm based on FISTA, where we need to provide the step  $\alpha$ . Convergence is highly influenced by the value of the parameter  $\rho$ , as described earlier.

## 4.6 Numerical examples

### 4.6.1 Synthetic examples

#### Intercept time-velocity pairs are known

We have first assumed that the  $\tau, v$  pairs that model the reflections are known. In real datasets, this information can be obtained via velocity analysis. We have discretized the  $p$  axis by defining 240 parameters in the interval  $p \in [0.001, 0.0033]s/m$ . We first apply our least-squared inversion algorithm with quadratic regularization. Figure 4.2a portrays the noise-free data, Figure 4.2b shows the data contaminated by noise. Figures 4.2c and d portray the synthesis of reflections and ground roll obtained after inversion with  $l_2$  penalty. Finally, Figure 4.2e shows the observed data minus the modeled ground roll.

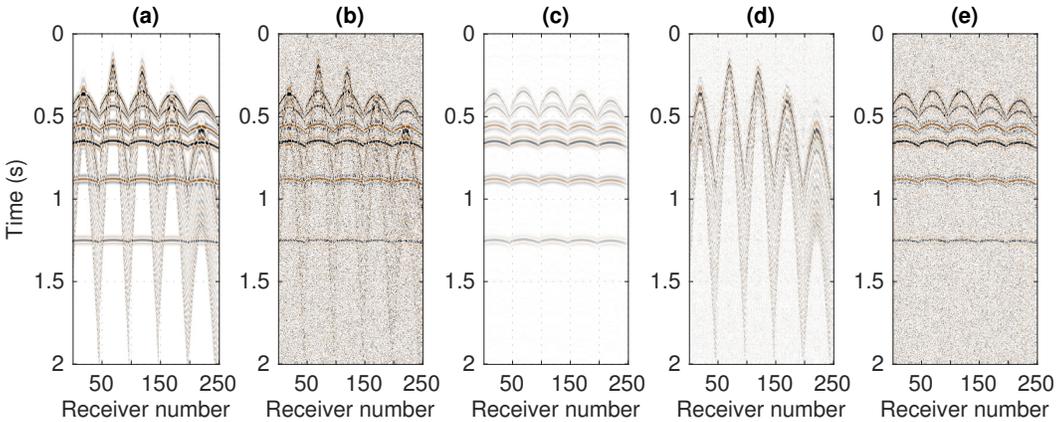


Figure 4.2: Separation of reflections and ground roll using least-squares inversion with quadratic regularization. The  $\tau, v$  pairs describing hyperbolic events are assumed known for the inversion algorithm. a) 3D synthetic shot gather. b) 3D synthetic shot gather contaminated with random noise ( $SNR = 1$ ). c) Estimated signal. d) Estimate coherent noise (ground roll). e) Observations (b) minus estimated coherent noise (d).

We repeat the example but now via least-squares inversion with the sparse ( $l_1$ ) regularization term. The results for this experiments solved by IRLS are shown in Figures 4.3a-e and by FISTA in 4.4a-e. We apply our robust inversion algorithm and figure 4.5a portrays the noise-free data, figure 4.5b shows the data contaminated by noise. Figures 4.5c and d portray the synthesis of reflections and ground roll obtained after inversion. Finally, figure 4.5e shows the observed data minus the modeled ground roll.

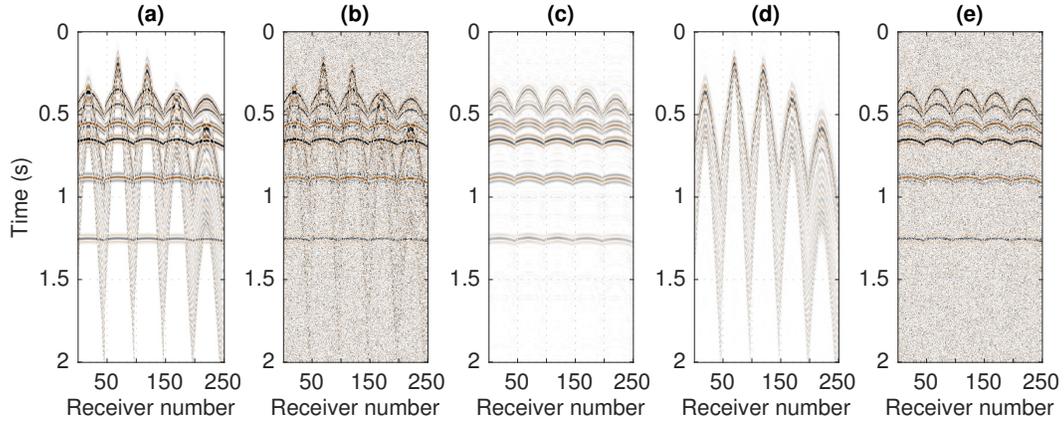


Figure 4.3: Separation of reflections and ground roll using least-squares inversion with sparse ( $l_1$ ) regularization (IRLS). The  $\tau, v$  pairs describing hyperbolic events are assumed known for the inversion. a) 3D synthetic shot gather. b) 3D synthetic shot gather contaminated with random noise ( $SNR = 1$ ). c) Estimated signal. d) Estimates coherent noise (ground roll). e) Observations (b) minus estimated coherent noise (d).

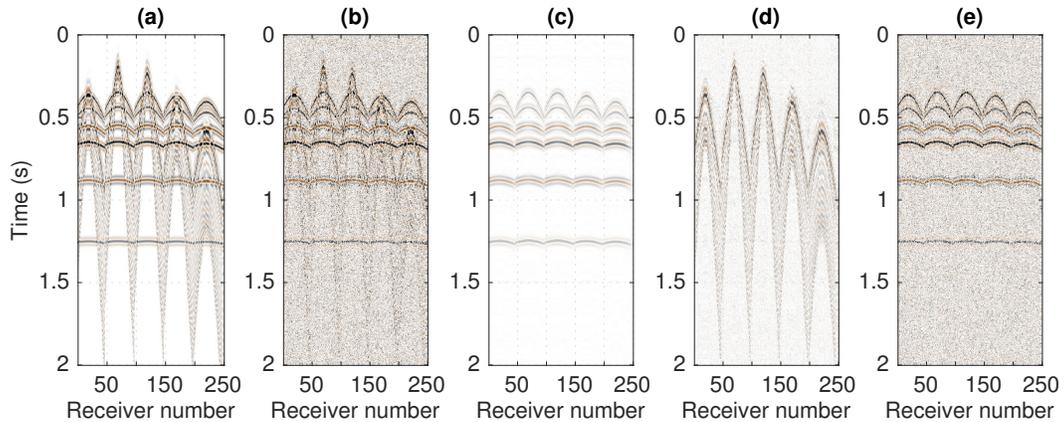


Figure 4.4: Separation of reflections and ground roll using least-squares inversion with sparse ( $l_1$ ) regularization (FISTA). The  $\tau, v$  pairs describing hyperbolic events are assumed known for the inversion. a) 3D synthetic shot gather. b) 3D synthetic shot gather contaminated with random noise ( $SNR = 1$ ). c) Estimated signal. d) Estimates coherent noise (ground roll). e) Observations (b) minus estimated coherent noise (d).

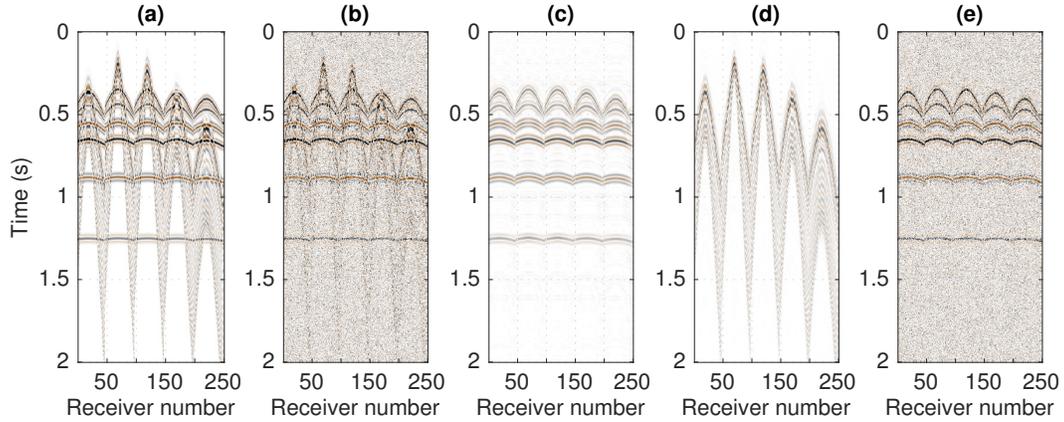


Figure 4.5: Separation of reflections and ground roll using robust inversion (ADMM). a) 3D synthetic shot gather. b) 3D synthetic shot gather contaminated with random noise ( $SNR = 1$ ). c) Estimated signal. d) Estimate coherent noise (ground roll). e) Observations (b) minus estimated coherent noise (d).

Figures 4.6a-d compare the  $f-p$  spectra for the least-squares inversion with  $l_2$  and  $l_1$  norm regularization, respectively. Superimposed to these images, we have added the theoretical expressions for the reciprocal of the phase velocity for the two dispersive modes that were used to generate the ground roll. The  $l_1$  regularization provides more resolution of the dispersive modes than the  $l_2$  regularization. The enhanced resolution can be corroborated by Figures 4.7a-d where we compared the spectra of dispersive events at  $f = 20Hz$ . It is important to point out that Figures 4.6a-d portray the absolute value of coefficients of the coherent noise  $|m_c(f, p)|$ .

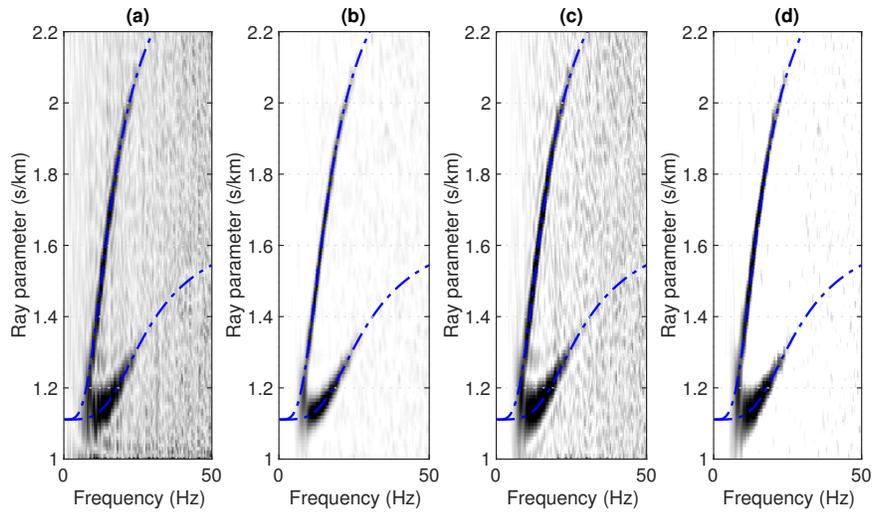


Figure 4.6: Spectra of the ground roll  $|m_c(f, p)|$ . a) Estimation via least-squares inversion with quadratic ( $l_2$ ) regularization. b) Least-squares inversion with sparse ( $l_1$ ) regularization (IRLS). c) Least-squares inversion with sparse ( $l_1$ ) regularization (FISTA). d) Sparse inversion (ADMM).

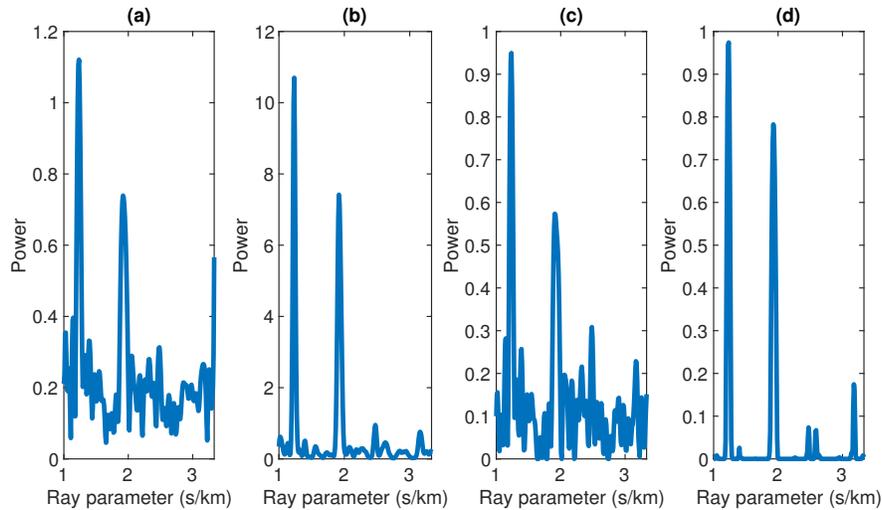


Figure 4.7: Spectra of the ground roll  $|m_c(f, p)|$  for  $f = 20 \text{ Hz}$ . a) Estimation via least-squares inversion with quadratic ( $l_2$ ) regularization. b) Least-squares inversion with sparse ( $l_1$ ) regularization (IRLS). c) Least-squares inversion with sparse ( $l_1$ ) regularization (FISTA). d) Sparse inversion (ADMM). Dashed lines indicate the theoretical dispersion modes used to generate the synthetic data.

It is important to mention that one could have used a slat-stack (linear Radon transform) to map data from  $t, h$  to  $\tau, p$ . However, given the dispersive nature of the ground roll, it is more convenient to compute coefficient in  $f, p$  rather than in  $\tau, p$ .

This example shows that both  $l_1$  (sparse) and the  $l_2$  (damped least-squares) regularization terms and robust inversion show identical results as per the separability of hyperbolas from dispersive coherent noise. We believe that this is because we know the  $\tau, v$  pairs that model the hyperbolas.

By examining the ground roll's spectra as presented in previous figures, one can notice a higher quality's separation for sparsity promotion of least squares and robust inversion. The estimated coefficients of coherent noise after inversion seem to match very well with the theoretical dispersive nodes. More detailed, spectra for a chosen frequency confirm the previous explanation. Sparsity and robust inversion clearly show the concentration of estimated coefficients around the theoretical nodes as two peaks. Quadratic least-squares contain noise, which makes the separation lower quality. Note that the appearance of FISTA, which solves the same problem as IRLS, has similar separability as damped least squares.

Table 4.1 shows the parameters and the results for each of the inversion algorithm used above:

<i>Method</i>	$\rho$	$\mu$	$\mu_1$	$\mu_2$	<i>Tolerance</i>	<i>Iterations</i>	$SNR_{out}dB$	<i>runtime(s)</i>
L2	-	-	0.1	0.1	-	1	9.69	6.39
IRLS	-	-	0.1	0.01	-	5	9.78	17.28
FISTA	-	16	-	-	$10^{-2}$	5	9.63	10.68
ADMM	1	50	-	-	$10^{-3}$	15	7.67	6.4

Table 4.1: Parameters used for inversion when  $\tau, v$  pairs of the model are known-synthetic data

where  $SNR_{out}$  for the signal only is calculated by

$$10 * \log \frac{\|d^{observed}\|_2^2}{\|d^{observed} - d_s\|_2^2} \quad (4.25)$$

### Intercept time-velocity pairs are unknown

We now run a synthetic experiment to test the proposed inversion algorithm in the realistic situation where one does not know the  $\tau, v$  pairs that model the hyperbolas. Trying a large number of intercept times ( $\tau$ ) and velocities ( $v$ ), as one does in different applications of the

Radon transform, will lead to a large operator  $\mathbf{A}_s$  and, consequently to a computationally expensive algorithm. For this reason, we have adopted a coarse discretization of  $\tau$  and  $v$ . For instance, the intercept axis  $\tau$  varies from  $0.2s$  to  $1.4s$  in 40 intervals sampled every  $0.0308s$  which is about seven times the sampling interval of the data  $\Delta t$ . Similarly, the velocity axis varies from  $1500m/s$  to  $5000m/s$  in 20 intervals. Therefore, for each parameter  $\tau$  we associate 20 trial velocities. The total number of coefficients for each frequency becomes  $N_s = 20 \times 40 = 800$ .

Figures 4.8a-e illustrate the least-squares solution with damping. Similarly, Figures 4.9a-e shows the least-squares solution with sparse ( $l_1$ ) regularization, using the IRLS algorithm while Figures 4.10 a-e shows the sparse solution using the FISTA algorithm and Figures ?? a-e shows the sparse solution using the ADMM algorithm.

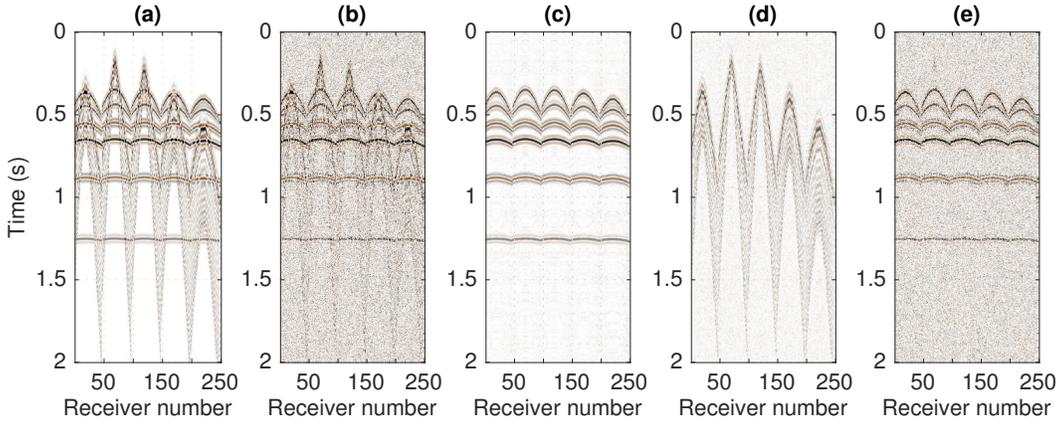


Figure 4.8: Separation of reflections and ground roll using least-squares inversion with quadratic regularization. The  $\tau, v$  pairs describing hyperbolic events are unknown. Therefore, we assume a coarse discretization of  $\tau, v$ . a) 3D synthetic shot gather. b) 3D synthetic shot gather contaminated with random noise ( $SNR = 1$ ). c) Estimated signal. d) Estimate coherent noise (ground roll). e) Synthetic observations (b) minus estimated coherent noise (d).

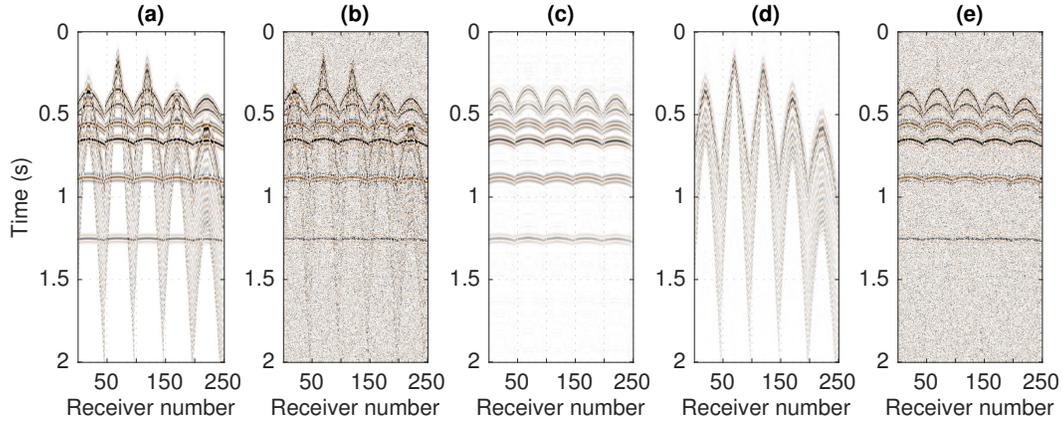


Figure 4.9: Separation of reflections and ground roll using least-squares inversion with sparse ( $l_1$ ) inversion using IRLS. The  $\tau, v$  pairs describing hyperbolic events are unknown for the inversion process. Coarse discretization of the pairs  $\tau, v$  is adopted. a) 3D synthetic shot gather. b) 3D synthetic shot gather contaminated with random noise ( $SNR = 1$ ). c) Estimated signal. d) Estimate coherent noise (ground roll). e) Synthetic observations (b) minus estimated coherent noise (d).

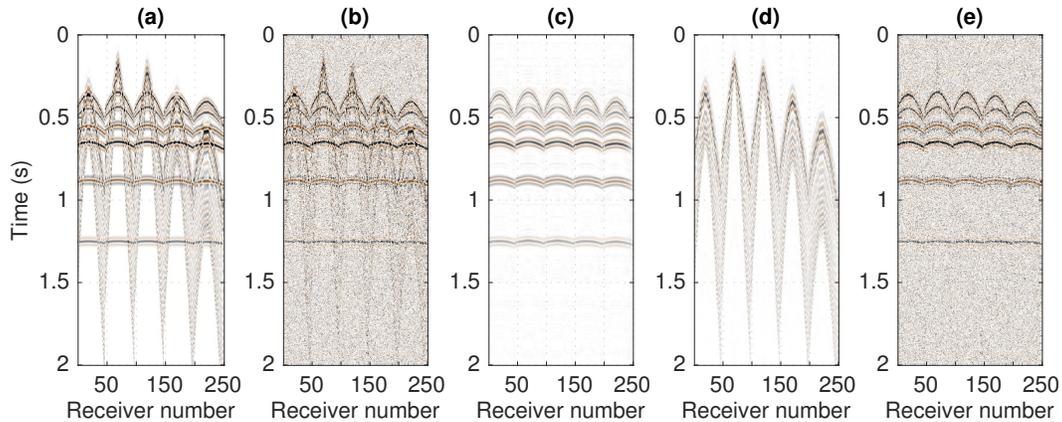


Figure 4.10: Separation of reflections and ground roll using least-squares inversion with sparse ( $l_1$ ) inversion using FISTA. The  $\tau, v$  pairs describing hyperbolic events are unknown for the inversion process. Coarse discretization of the pairs  $\tau, v$  is adopted. a) 3D synthetic shot gather. b) 3D synthetic shot gather contaminated with random noise ( $SNR = 1$ ). c) Estimated signal. d) Estimate coherent noise (ground roll). e) Synthetic observations (b) minus estimated coherent noise (d).

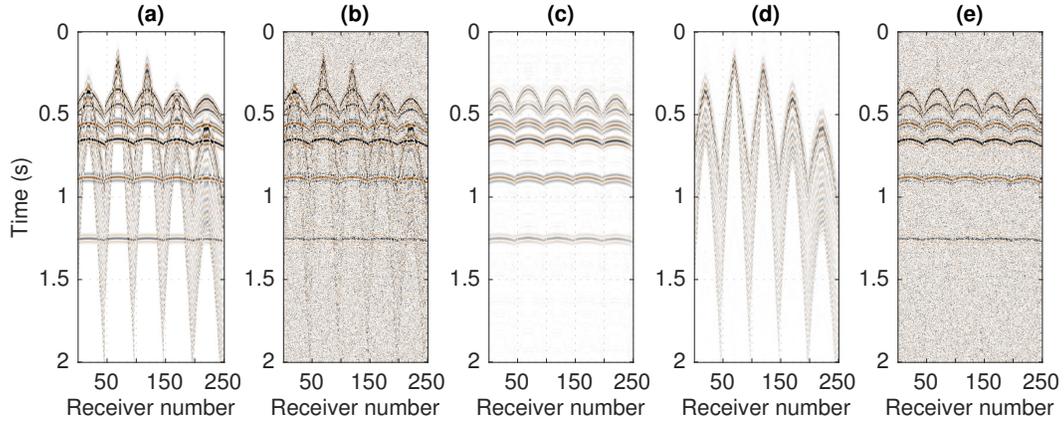


Figure 4.11: Separation of reflections and ground roll using robust inversion with ADMM. The  $\tau, v$  pairs describing hyperbolic events are unknown for the inversion process. Coarse discretization of the pairs  $\tau, v$  is adopted. a) 3D synthetic shot gather. b) 3D synthetic shot gather contaminated with random noise ( $SNR = 1$ ). c) Estimated signal. d) Estimate coherent noise (ground roll). e) Synthetic observations (b) minus estimated coherent noise (d).

Figures 4.12a-d illustrate the  $f, p$  spectra of the dispersive coherent noise for the least-squares solution with damping and sparsity constraints, respectively. Figures 4.13a-d show the spectra of the dispersive noise at  $f = 20Hz$ . Again, one can observe an important gain in resolution when the sparse prior is used to estimate the coefficients that model the ground roll.

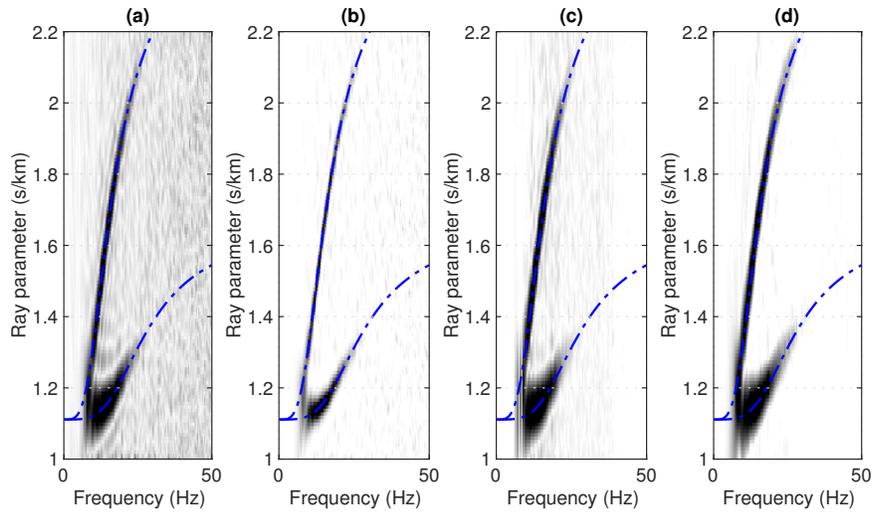


Figure 4.12: Spectra of the ground roll  $|m_c(f, p)|$ . a) Estimation via least-squares inversion with quadratic ( $l_2$ ) regularization. b) Least-squares inversion with sparse ( $l_1$ ) regularization (IRLS). c) Least-squares inversion with sparse ( $l_1$ ) regularization (FISTA). d) Sparse inversion (ADMM).

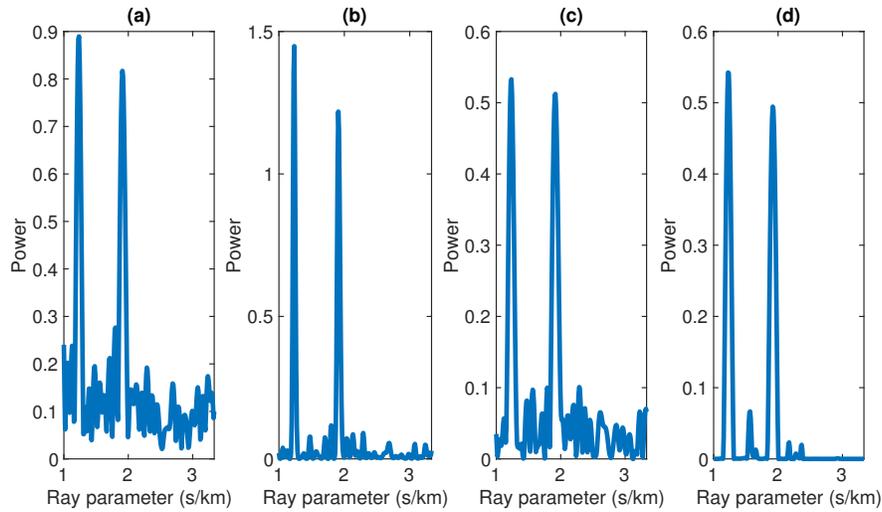


Figure 4.13: Spectra of the ground roll  $|m_c(f, p)|$  for  $f = 20 Hz$ . a) Estimation via least-squares inversion with quadratic ( $l_2$ ) regularization. b) Least-squares inversion with sparse ( $l_1$ ) regularization (IRLS). c) Least-squares inversion with sparse ( $l_1$ ) regularization (FISTA). d) Robust inversion (ADMM). Dashed lines indicate the theoretical dispersion modes used to generate the synthetic data.

Table 4.2 shows the parameters and the results for each of the inversion algorithms used above:

<i>Method</i>	$\rho$	$\mu$	$\mu 1$	$\mu 2$	<i>tolerance</i>	<i>Iterations</i>	<i>runtime(s)</i>
L2	-	-	0.1	0.01	-	1	40.4
IRLS	-		0.1	0.01	-	5	148
FISTA	-	0.125	-	-	$10^{-2}$	2	30.4
ADMM	1	50	-	-	$10^{-3}$	1	34.2

Table 4.2: Parameters used for inversion when  $\tau, v$  pairs of the model are unknown-synthetic data.

This example shows similar results as in the case where we know the intercept time-velocity pairs of reflections. In synthetic examples, all methods seem to provide similar results in the separation of two types of signals. However, as spectra of ground roll show, the accuracy of a robust inversion instead of a least-squares inversion with sparse or quadratic regularization is much higher.

This example is more challenging because we expect algorithms to detect the correct time-velocity pairs. By examining the ground roll's spectra, one can notice a higher quality's separation for sparsity promotion of least squares and robust inversion. The estimated coefficients of coherent noise after inversion seem to match very well with the theoretical dispersive nodes. More detailed, spectra for a chosen frequency confirm the previous explanation. Sparsity and robust inversion clearly show the concentration of estimated coefficients around the theoretical nodes as two peaks. Quadratic least-squares contain noise, which makes the separation lower quality. In the previous case, where we assumed known time-velocity pairs, FISTA had a poor separation result. However, in the current demanding case, both solvers of least-squares inversion with sparse penalty give results similar to the robust inversion.

Last, for completeness, we have also computed an image Figure 4.14 that represents the energy of the coefficients that model the hyperbolic events in the low-resolution  $\tau, v$  grid. The estimated coefficients  $m_s(f, k)$  were mapped to the cube  $m_s(f, \tau, p)$  and then the following velocity spectra was computed for plotting purposes  $\sum_f Re[m_s(f, \tau, p)]$  which it can be shown to be similar to the display one computes for the Hyperbolic Radon Transform. This experiment shows that one does not need to have precise knowledge of velocities and intercept times to separate ground roll from reflections properly. The analysis also indicates an improvement in the separability of coherent noise from reflections when the inverse problem is regularized via the  $l_1$  (sparse) penalty term.

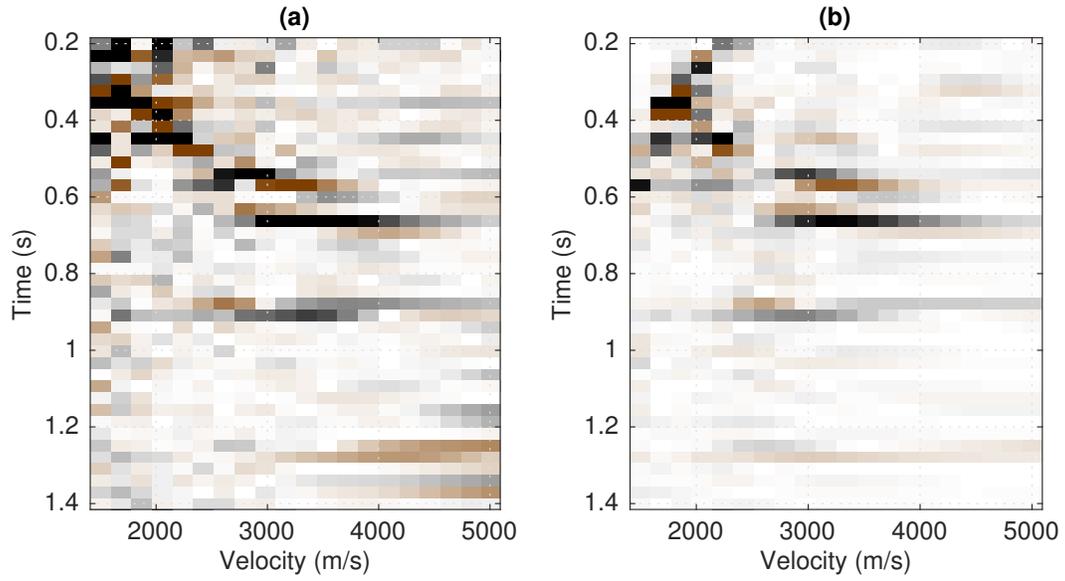


Figure 4.14: Spectra of hyperbolic events coarsely discretized  $\sum_f Re[m_s(f, \tau, v)]$ .  
a) Estimation via least-squares inversion with quadratic ( $l_2$ ) regularization. b) Least-squares inversion with sparse ( $l_1$ ) regularization.

## 4.6.2 Field data examples

We apply our algorithms to test their effectiveness in two irregular field data sets: a 3D and a 2D. As in synthetic examples, we examine two cases. In the first one, we estimate the intercept time-velocity pairs through a semblance map, and in the second case, we use a dense grid of intercept times and velocities and we expect through inversion to estimate the correct pairs.

### Example 1: 3D data set

Figure 4.15 shows the distribution of receivers for the first real 3D shot gather data set used to test the effectiveness of our algorithms. Data consist of  $N_y = 9$  receivers lines, the length of each record is 2000 samples; the sampling interval is  $\Delta t = 2ms$  and the time duration is  $4s$ .

We have discretized the  $p$  axis by defining 2000 parameters in the interval  $p \in [0.1, 0.002]s/m$ . For velocity analysis we have applied Automatic Gain Control (AGC) to strengthen the reflections (figure 4.16). We then tried the velocity analysis (figure 4.17) and we estimated  $\tau = [1.84, 1.862.22, 3.33]s$  and  $v = [2242, 2420, 3126, 4030]m/s$ .

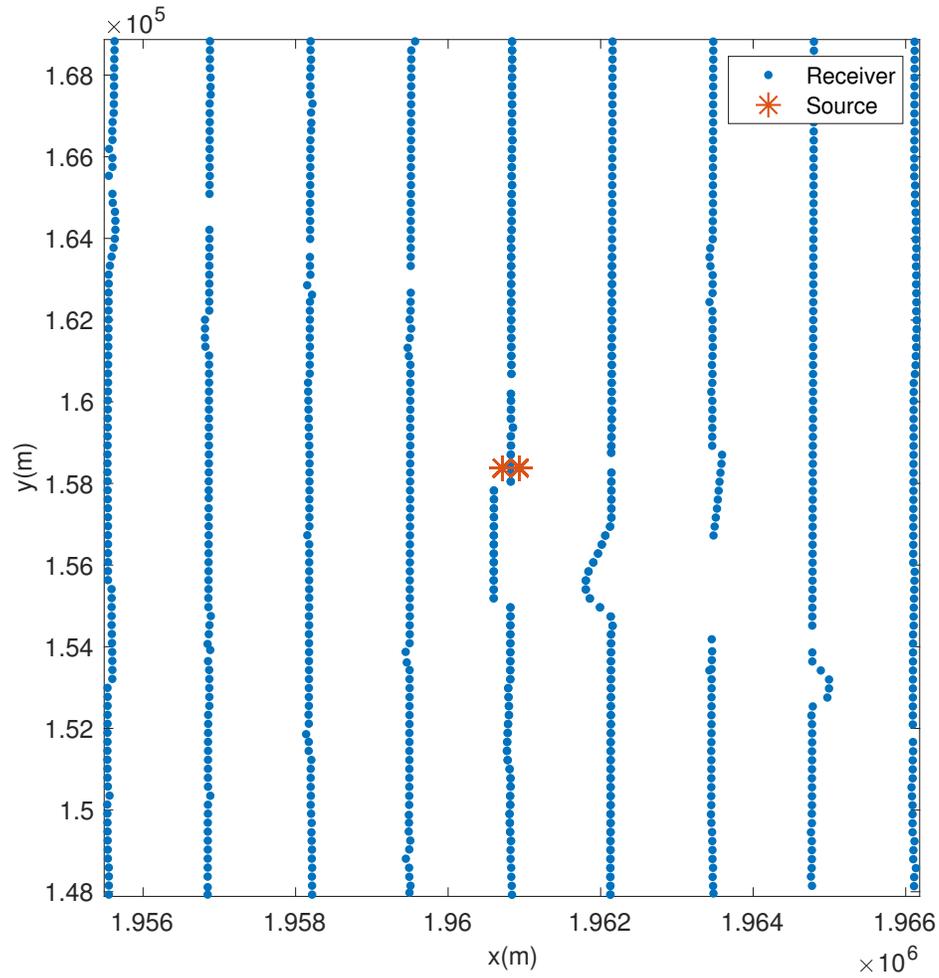


Figure 4.15: Source and receivers for the irregular field data example adopted to test the effectiveness of our signal and ground roll separation methods.

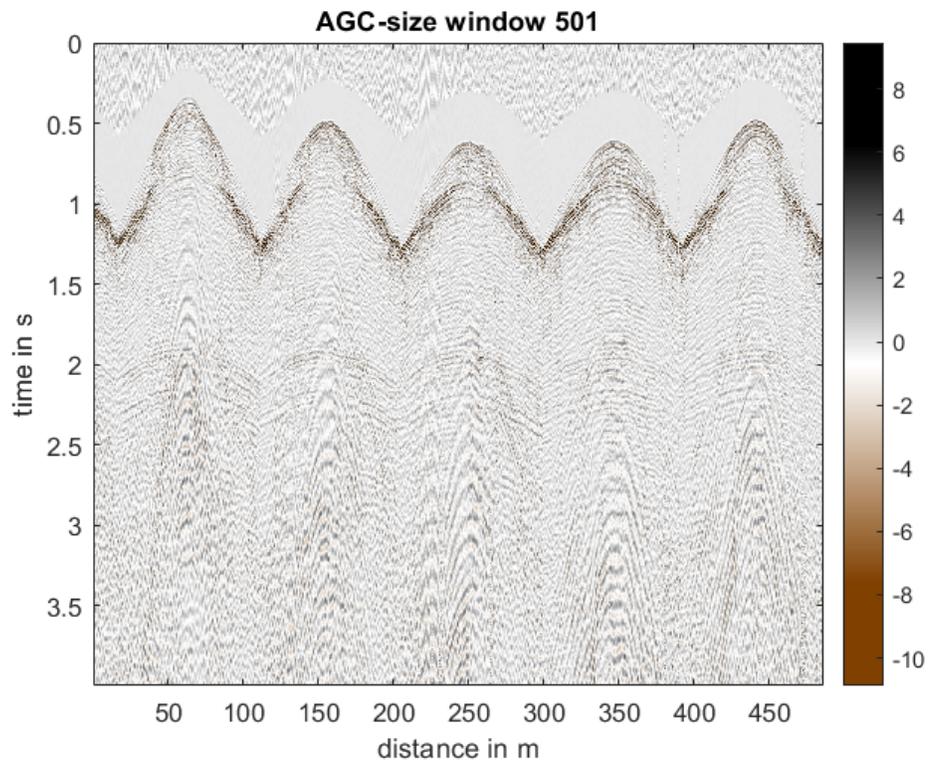


Figure 4.16: Automatic Gain Control with window size 501. Reflections are stronger for the next processing steps.

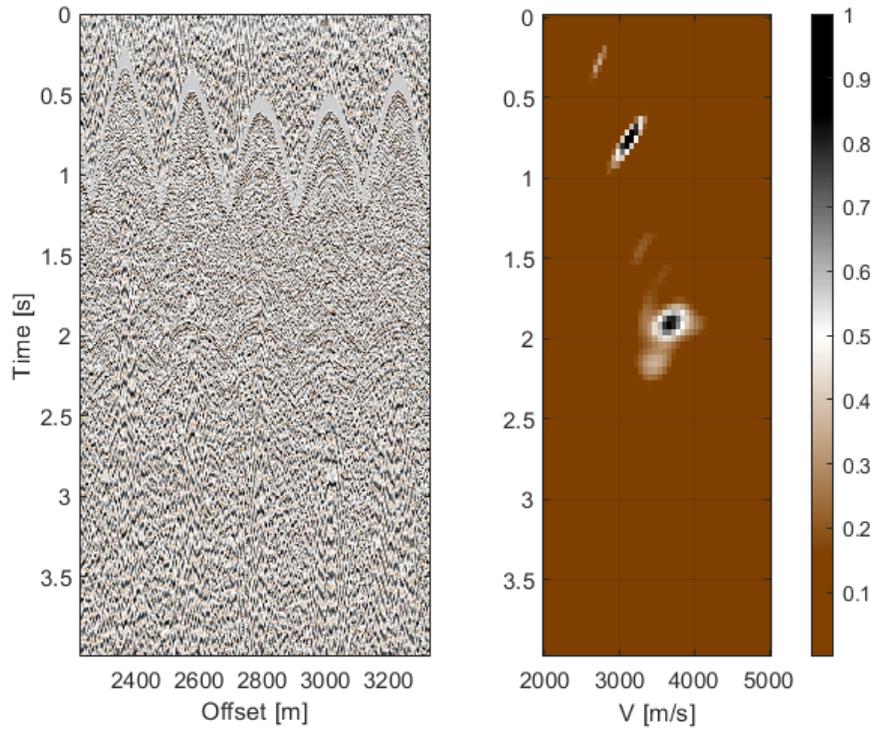


Figure 4.17: Velocity Analysis applied on a shot section to estimate the intercept time-velocity pairs. The areas with higher energy show the possible pairs.

We are now ready to examine all the previously proposed methods.

Figures 4.18a-c illustrate the least-squares solution with damping. Similarly, Figures 4.19a-c show the least-squares solution with sparse ( $l_1$ ) regularization using IRLS and Figures 4.20a-c shows the  $l_1$  regularization with FISTA algorithm. Finally, Figures 4.21 show denoising with robust inversion.

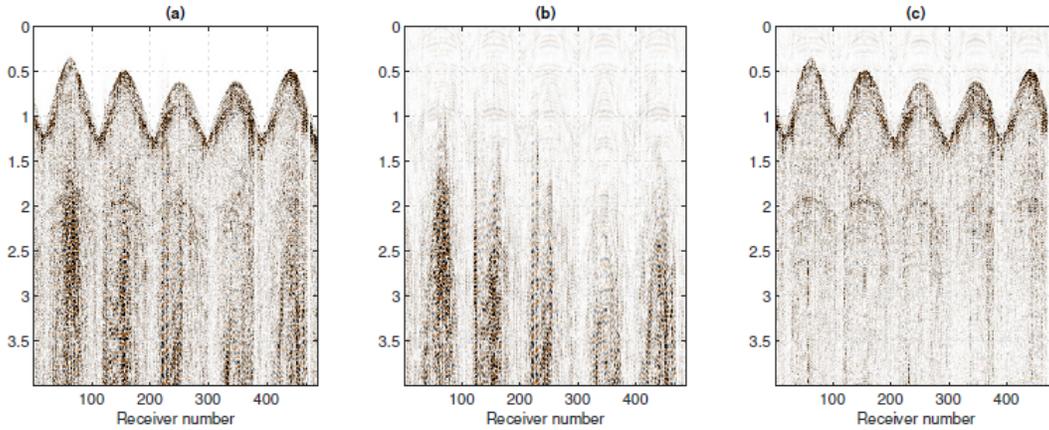


Figure 4.18: Separation of reflections and ground roll using least-squares inversion with quadratic regularization. a) 3D real shot gather. b) Estimated coherent noise (ground roll). c) Real observations (a) minus estimated coherent noise (b).

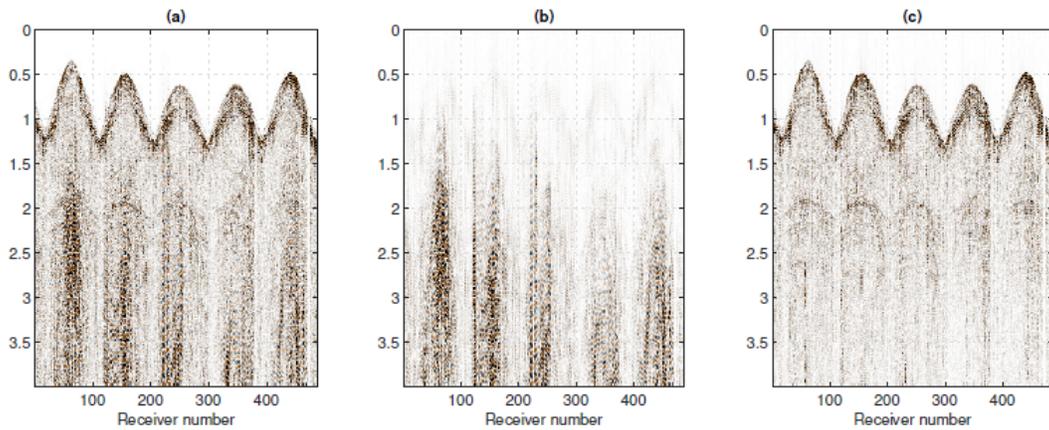


Figure 4.19: Separation of reflections and ground roll using least-squares inversion with sparse ( $l_1$ ) regularization using IRLS. a) 3D real shot gather. b) Estimated coherent noise (ground roll). c) Real observations (a) minus estimated coherent noise (b).

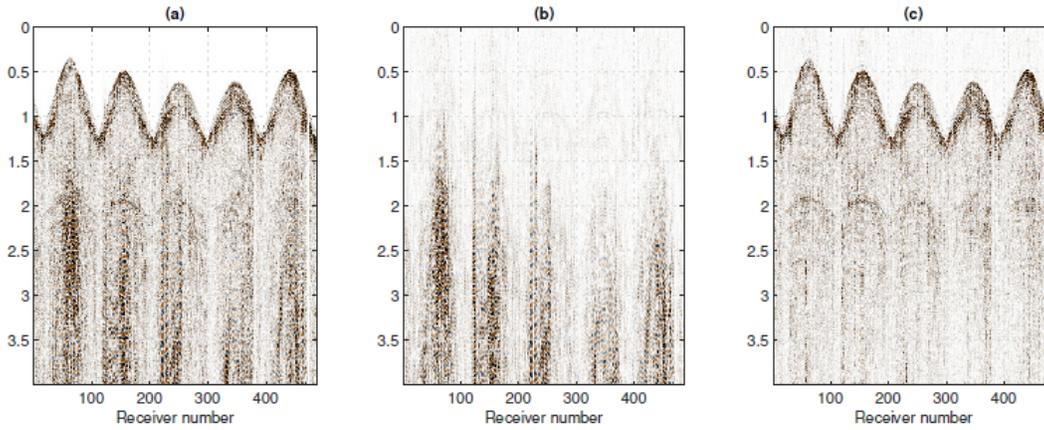


Figure 4.20: Separation of reflections and ground roll with least-squares inversion with sparse ( $l_1$ ) regularization using FISTA. a) 3D real shot gather. b) Estimated coherent noise (ground roll). c) Real observations (a) minus estimated coherent noise (b).

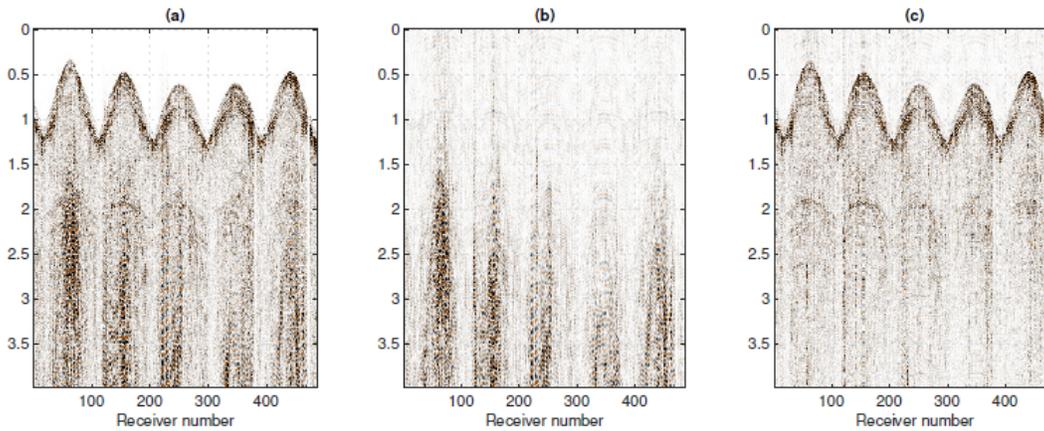


Figure 4.21: Separation of reflections and ground roll with robust inversion using ADMM. a) 3D real shot gather. b) Estimated coherent noise (ground roll). c) Real observations (a) minus estimated coherent noise (b).

Figure 4.22 shows a comparison for the 4 algorithms on the same zoomed part for the data, L2, IRLS, FISTA and ADMM and table 4.3 shows the parameters used for the above results. Both  $x$  and  $y$  axis show which part from main data is selected.

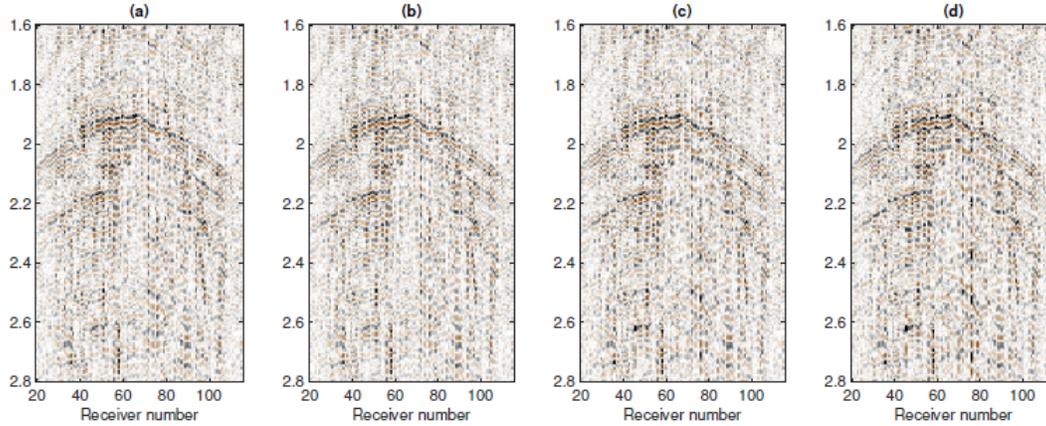


Figure 4.22: a) Separation of reflections and ground roll using least-squares inversion with  $(l_2)$  regularization, b) with  $(l_1)$  regularization using IRLS, c) using FISTA and d) using ADMM

Table 4.3 shows the parameters and the results for each of the inversion algorithms used above:

<i>Method</i>	$\rho$	$\mu$	$\mu_1$	$\mu_2$	<i>tolerance</i>	<i>Iterations</i>	<i>runtime(s)</i>
L2	-	-	0.1	0.1	-	1	92.3
IRLS	-	-	0.1	100	-	2	144.43
FISTA	-	0.125	-	-	$10^{-2}$	30	164.8
ADMM	0.0001	50	-	-	$10^{-3}$	311	268.9

Table 4.3: Parameters used for inversion when  $\tau, v$  pairs of the model are known-3D real data.

Denosing results appear similar after using all separation methods. Let's focus on the details. Separation with quadratic regularization is faster, but stronger artifacts are visible in the section which may cause problems for further processing steps (visible in 2.6-2.8 s). Separation with IRLS and FISTA need almost the same processing time, with IRLS giving a section with less noise (traces close to 60 and 80). Finally, ADMM algorithm needs the longest processing time, and the results seem to be better than FISTA and similar to IRLS. A comparison of all four methods show that with IRLS we can have a high quality denosing result with less processing time and a few iterations, which leads to a reduced computational cost.

We now try to apply the separation methods in case ground roll is very strong and it masks most of the seismic record and it is hard to estimate the time-velocity pairs.

The intercept axis  $\tau$  varies from  $1.5s$  to  $3.5s$  in 20 intervals. Similarly, the velocity axis varies from  $2000m/s$  to  $4000m/s$  in 10 intervals. Therefore, for each parameter  $\tau$  we associate 10 trial velocities. The total number of coefficients for each frequency becomes  $N_s = 20 \times 10 = 200$ .

Figures 4.23a-c illustrate the least-squares solution with damping. Similarly, Figures 4.24a-c show the least-squares solution with sparse ( $l_1$ ) regularization using IRLS and Figures 4.25a-c shows the  $l_1$  regularization with FISTA algorithm. Finally, Figures 4.26 show denoising with robust inversion.

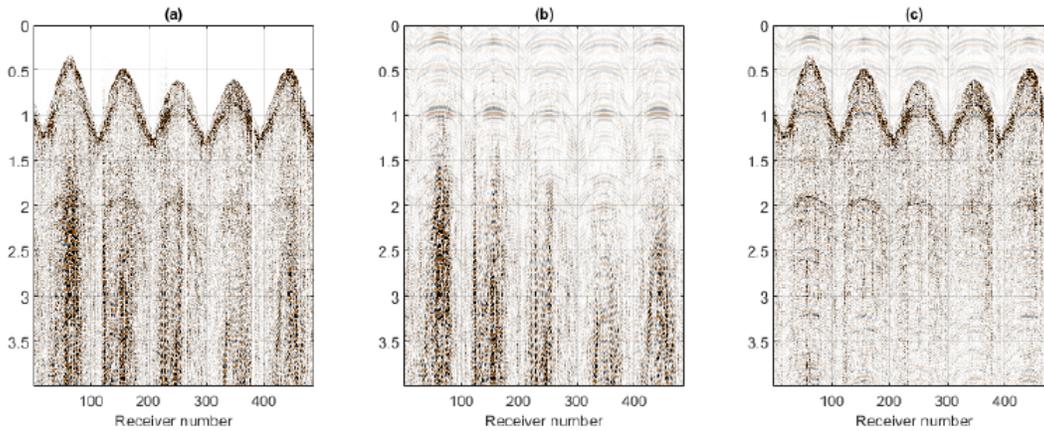


Figure 4.23: Separation of reflections and ground roll using least-squares inversion with quadratic regularization. a) 3D real shot gather. b) Estimated coherent noise (ground roll). c) Real observations (a) minus estimated coherent noise (b).

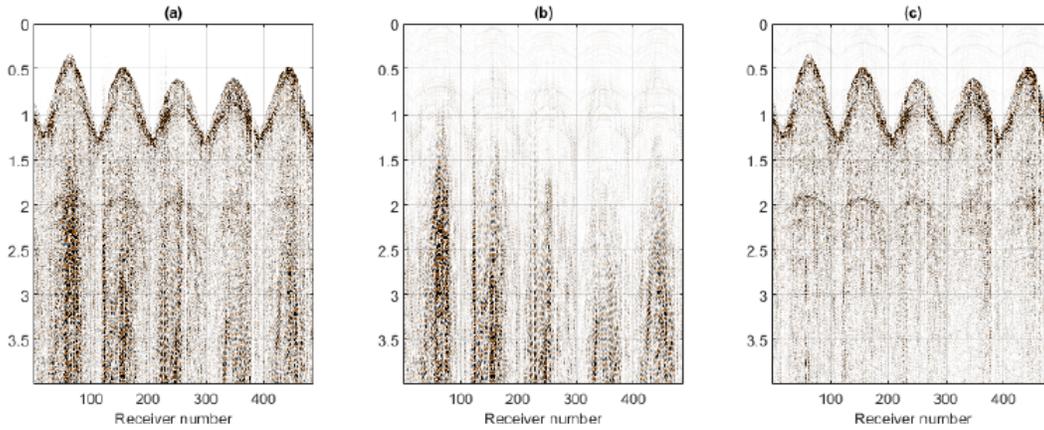


Figure 4.24: Separation of reflections and ground roll using least-squares inversion with sparse ( $l_1$ ) regularization using IRLS. a) 3D real shot gather. b) Estimated coherent noise (ground roll). c) Real observations (a) minus estimated coherent noise (b).

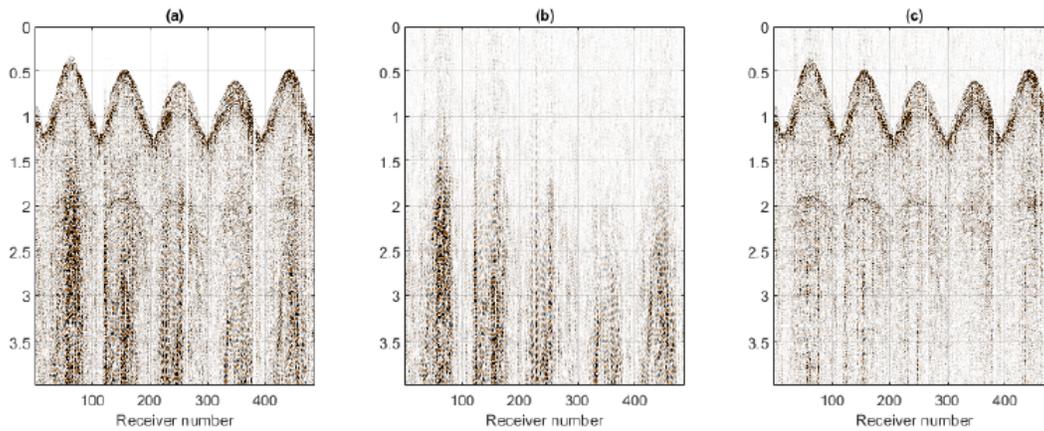


Figure 4.25: Separation of reflections and ground roll with least-squares inversion with sparse ( $l_1$ ) regularization using FISTA. a) 3D real shot gather. b) Estimated coherent noise (ground roll). c) Real observations (a) minus estimated coherent noise (b).

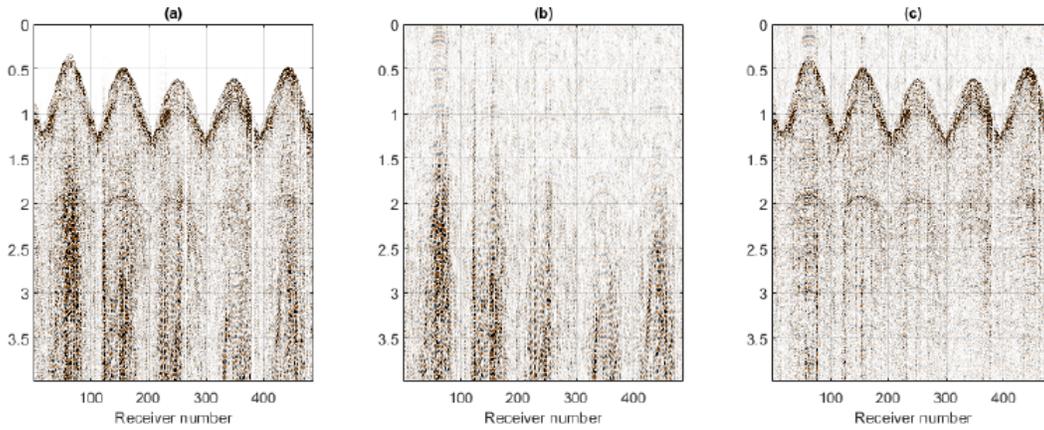


Figure 4.26: Separation of reflections and ground roll with robust inversion using ADMM. a) 3D real shot gather. b) Estimated coherent noise (ground roll). c) Real observations (a) minus estimated coherent noise (b).

Figure 4.27 shows a comparison for the 4 algorithms on the same zoomed part for the data,  $l_2$ , IRLS, FISTA and ADMM and table 4.4 shows the parameters used for the above results. Both  $x$  and  $y$  axis show which part from main data is selected.

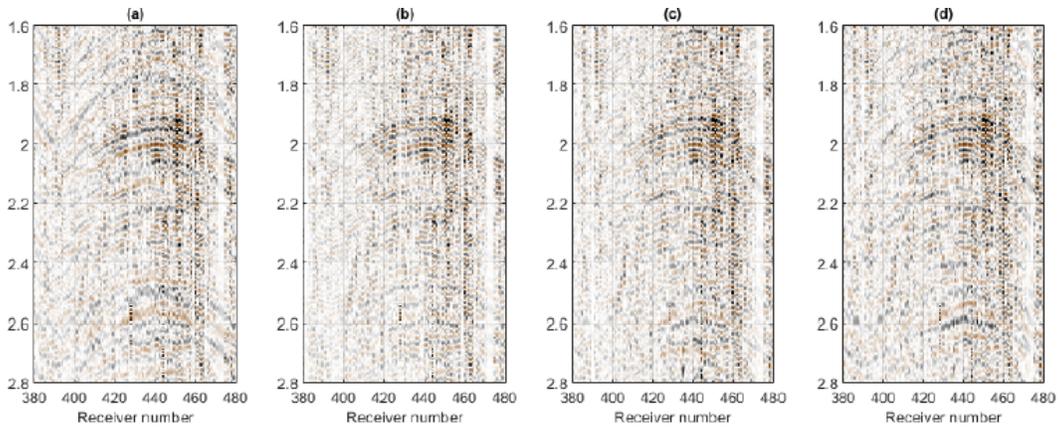


Figure 4.27: a) Separation of reflections and ground roll using least-squares inversion with  $(l_2)$  regularization, b) with  $(l_1)$  regularization using IRLS, c) using FISTA and d) using ADMM

Table 4.4 shows the parameters and the results for each of the inversion algorithms used above:

<i>Method</i>	$\rho$	$\mu$	$\mu_1$	$\mu_2$	<i>tolerance</i>	<i>Iterations</i>	<i>runtime(s)</i>
L2	-	-	0.01	10	-	1	44.5
IRLS	-	-	0.01	10	-	2	70.3
FISTA	-	0.125	-	-	$10^{-2}$	30	122.6
ADMM	0.00001	250	-	-	$10^{-3}$	39	89.7

Table 4.4: Parameters used for inversion when  $\tau, v$  pairs of the model are unknown-3D real data

In this case, when we cannot use a semblance map to estimate the  $\tau - v$  pairs, the results are very interesting. Compared to the case of known pairs, processing times are much shorter. Although we study a different area of the same data set than the previous case, one can have conclusions for the separation methods. As previously, least squares inversion with quadratic regularization norm gives poor quality results. The section contains many artifacts that we cannot surely say which of the reflections are real.

However, among IRLS, FISTA and ADMM, removal of ground roll is in the same level. But, IRLS and FISTA methods may create less artifacts.

**Example 2: 2D data set (CYNTHIA-TGS)**

Our algorithms are designed for denoising for 3D data sets. However, we applied them in 2D field (figure 4.28) set with irregular distribution, because the reflections are masked with a very dispersive and aliased ground roll and we would like to test our methods in a demanding environment.

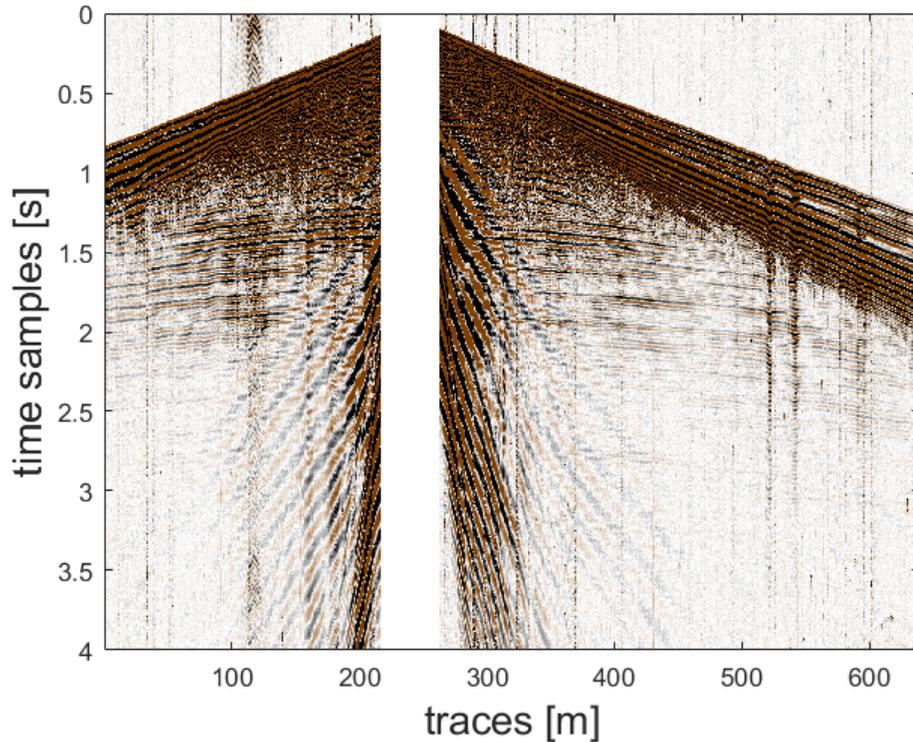


Figure 4.28: 2D field data set with very dispersive ground roll

Figure 4.29 shows the distribution of receivers for the first real 3D shot gather data set used to test the effectiveness of our algorithms. The sampling interval is  $\Delta t = 1ms$  and the time duration is  $4s$ .

We have discretized the  $p$  axis by defining 2000 parameters in the interval  $p \in [0.04, 0.002]s/m$ . For velocity analysis we have applied Automatic Gain Control (AGC) to strengthen the reflections (figure 4.30). We then tried the velocity analysis (figure 4.31) and we estimated  $\tau = [1.33, 1.48, 1.6, 1.74, 1.92]s$  and  $v = [3490, 3816, 4020, 4143, 4340]m/s$ .

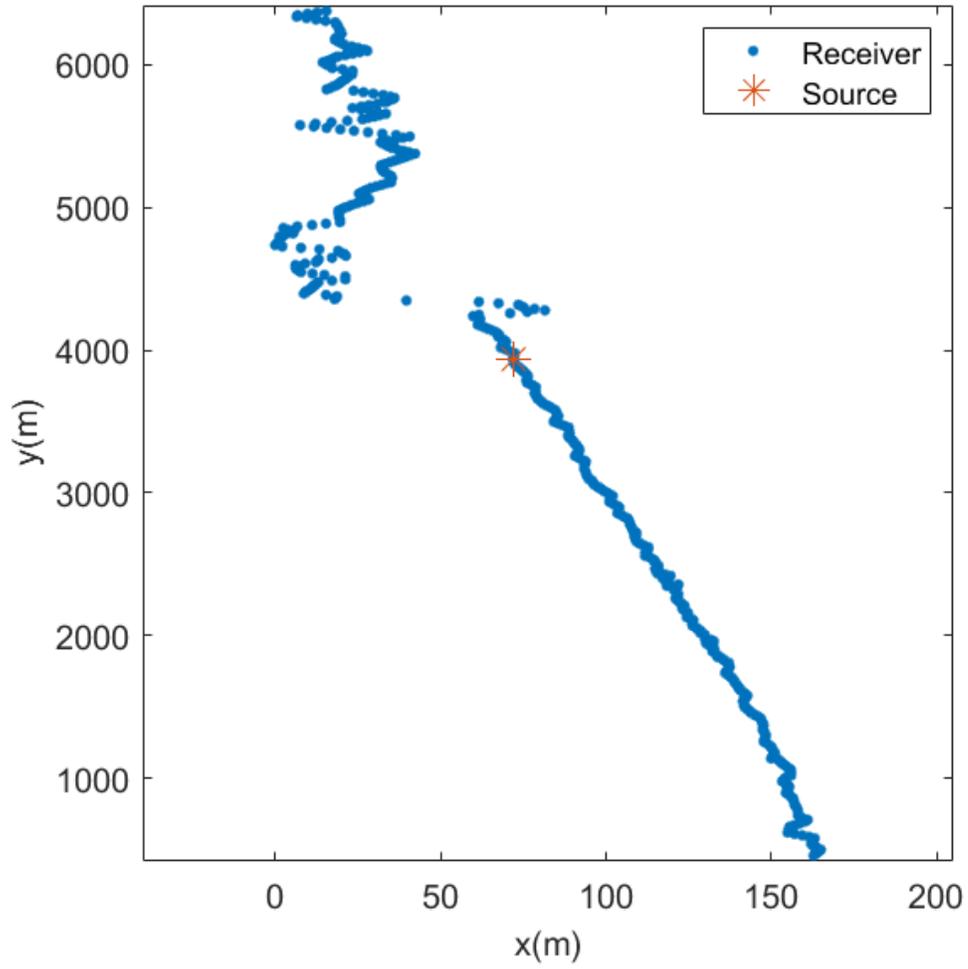


Figure 4.29: Source and receivers for the irregular 2D field data example adopted to test the effectiveness of our signal and ground roll separation methods.

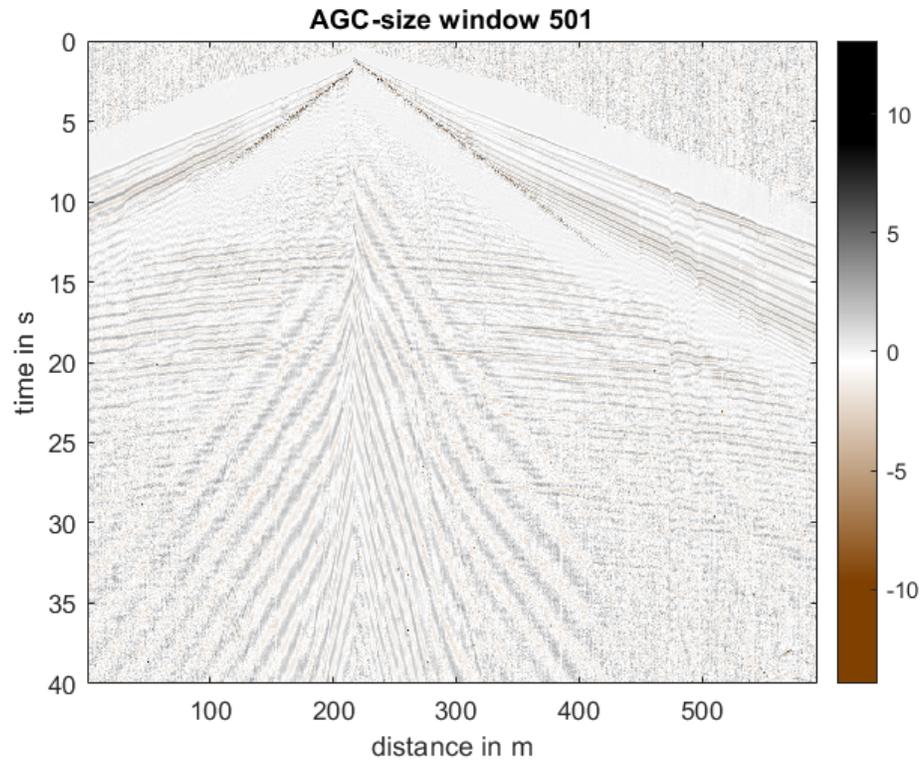


Figure 4.30: Automatic Gain Control with window size 501. Reflections are stronger for next processing steps.

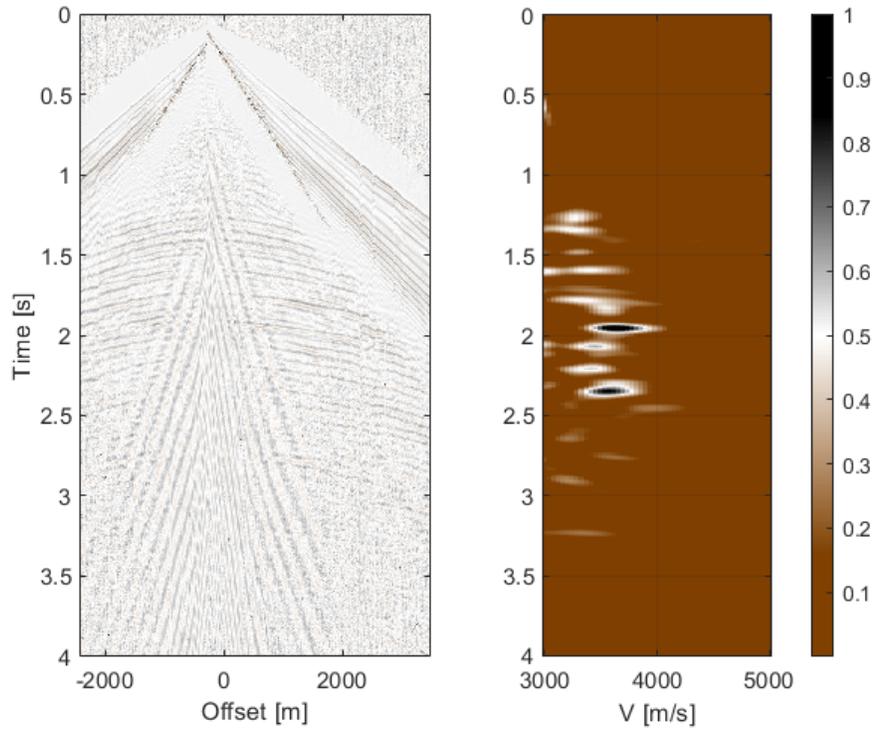


Figure 4.31: Velocity Analysis applied on a shot section to estimate the intercept time-velocity pairs. The areas with higher energy show the possible pairs.

We now examine all the previously proposed methods to test their effectiveness in a seismic record where ground roll is highly dispersive and aliased.

Figures 4.32a-c illustrate the least-squares solution with damping. Similarly, Figures 4.33a-c show the least-squares solution with sparse ( $l_1$ ) regularization using IRLS and Figures 4.34a-c shows the  $l_1$  regularization with FISTA algorithm. Finally, Figures 4.35 show denoising with robust inversion.

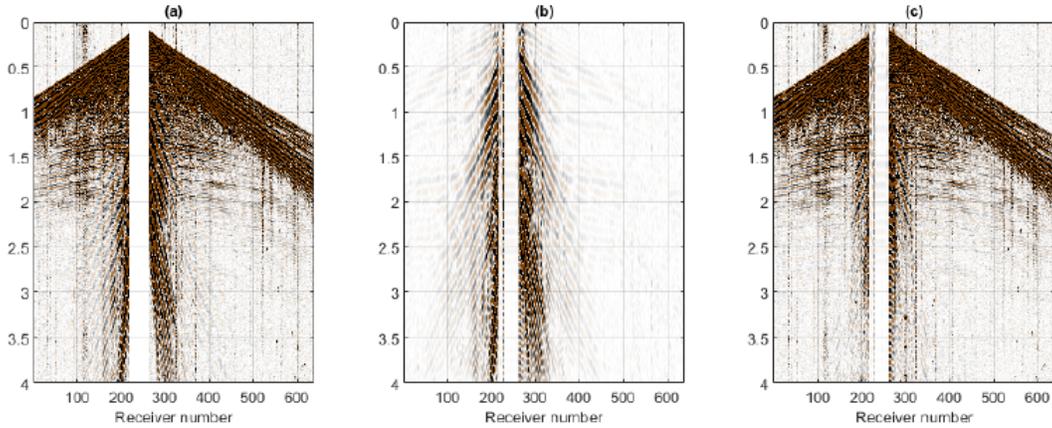


Figure 4.32: Separation of reflections and ground roll using least-squares inversion with quadratic regularization. a) 3D real shot gather. b) Estimated coherent noise (ground roll). c) Real observations (a) minus estimated coherent noise (b).

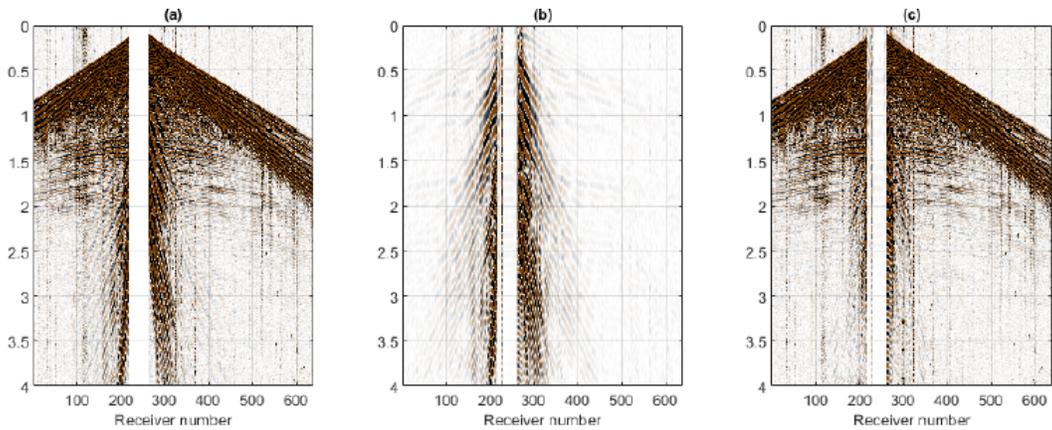


Figure 4.33: Separation of reflections and ground roll using least-squares inversion with sparse ( $l_1$ ) regularization using IRLS. a) 3D real shot gather. b) Estimated coherent noise (ground roll). c) Real observations (a) minus estimated coherent noise (b).

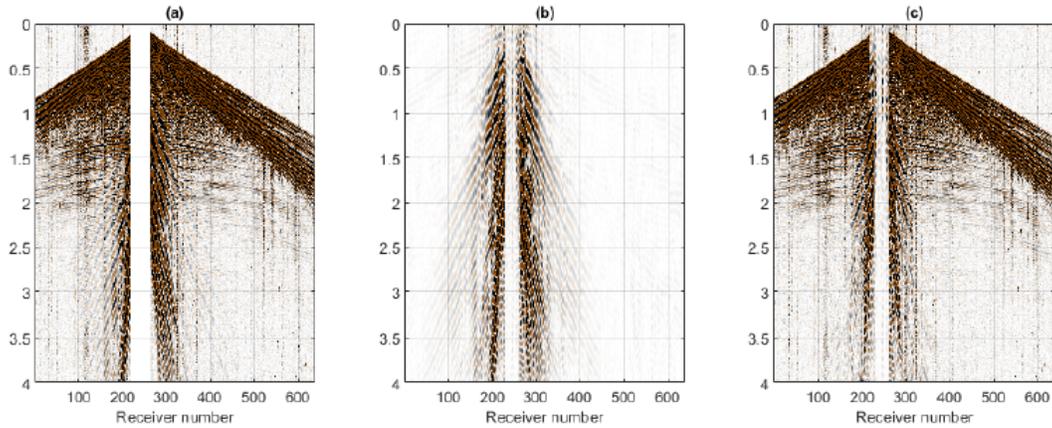


Figure 4.34: Separation of reflections and ground roll with least-squares inversion with sparse ( $l_1$ ) regularization using FISTA. a) 3D real shot gather. b) Estimated coherent noise (ground roll). c) Real observations (a) minus estimated coherent noise (b).

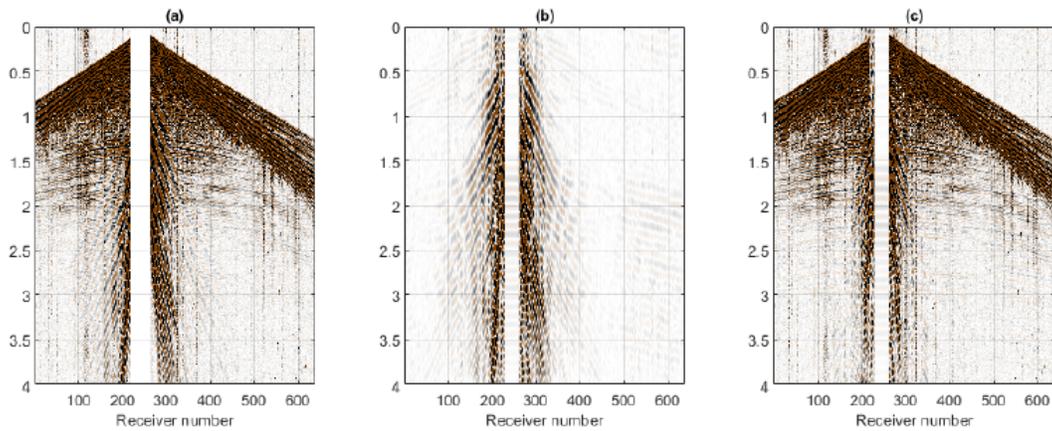


Figure 4.35: Separation of reflections and ground roll with robust inversion using ADMM. a) 3D real shot gather. b) Estimated coherent noise (ground roll). c) Real observations (a) minus estimated coherent noise (b).

Figure 4.36 shows a comparison for the 4 algorithms on the same zoomed part for the data, L2, IRLS, FISTA and ADMM and table 4.5 shows the parameters used for the above results. Both  $x$  and  $y$  axis show which part from main data is selected.

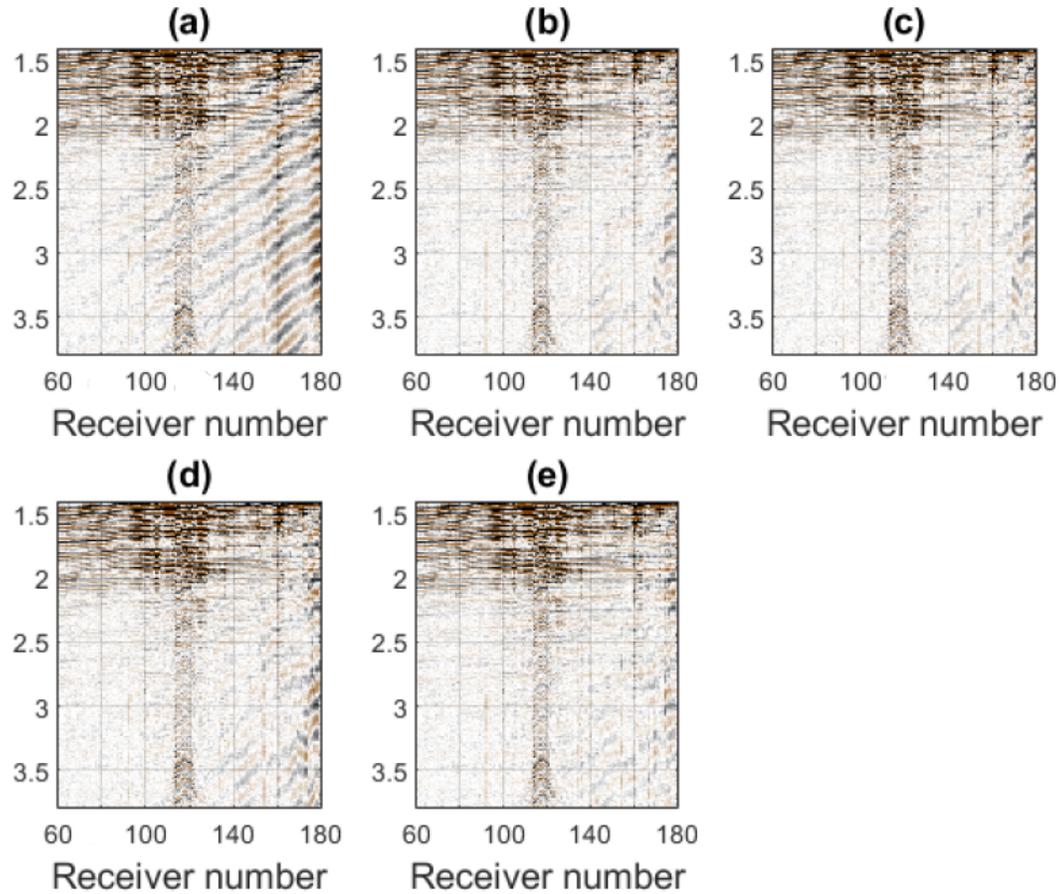


Figure 4.36: a) Zoomed area from raw data containing dispersive ground roll, b) Separation of reflections and ground roll using least-squares inversion with ( $l_2$ ) regularization, c) with ( $l_1$ ) regularization using IRLS, d) using FISTA and e) using ADMM

Table 4.5 shows the parameters and the results for each of the inversion algorithms used above:

<i>Method</i>	$\rho$	$\mu$	$\mu_1$	$\mu_2$	<i>tolerance</i>	<i>Iterations</i>	<i>runtime(s)</i>
L2	-	-	0.01	1	-	1	24.5
IRLS	-	-	0.01	1	-	2	38.7
FISTA	-	1024	-	-	$10^{-4}$	50	149.6
ADMM	1	50	-	-	$10^{-4}$	236	42.8

Table 4.5: Parameters used for inversion when  $\tau, v$  pairs of the model are known-2D real data

In a highly dispersive and aliased ground roll section, surprisingly all methods can successfully attenuate the coherent noise. However, if we consider the processing time, FISTA is more expensive. All other methods need similar time to give the same quality of results.

We now apply our algorithms in the case we cannot estimate the time-velocity pairs.

The intercept axis  $\tau$  varies from 1.0s to 2.5s in 20 intervals. Similarly, the velocity axis varies from 3000m/s to 5000m/s in 10 intervals. Therefore, for each parameter  $\tau$  we associate 10 trial velocities. The total number of coefficients for each frequency becomes  $N_s = 20 \times 10 = 200$ .

Figures 4.37a-c illustrate the least-squares solution with damping. Similarly, Figures 4.38a-c show the least-squares solution with sparse ( $l_1$ ) regularization using IRLS and Figures 4.39a-c shows the  $l_1$  regularization with FISTA algorithm. Finally, Figures 4.40 show denoising with robust inversion.

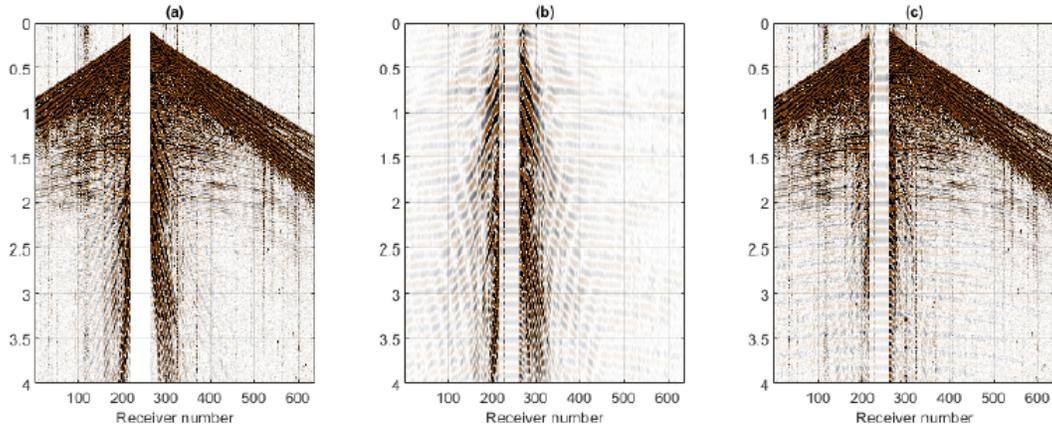


Figure 4.37: Separation of reflections and ground roll using least-squares inversion with quadratic regularization. a) 3D real shot gather. b) Estimated coherent noise (ground roll). c) Real observations (a) minus estimated coherent noise (b).

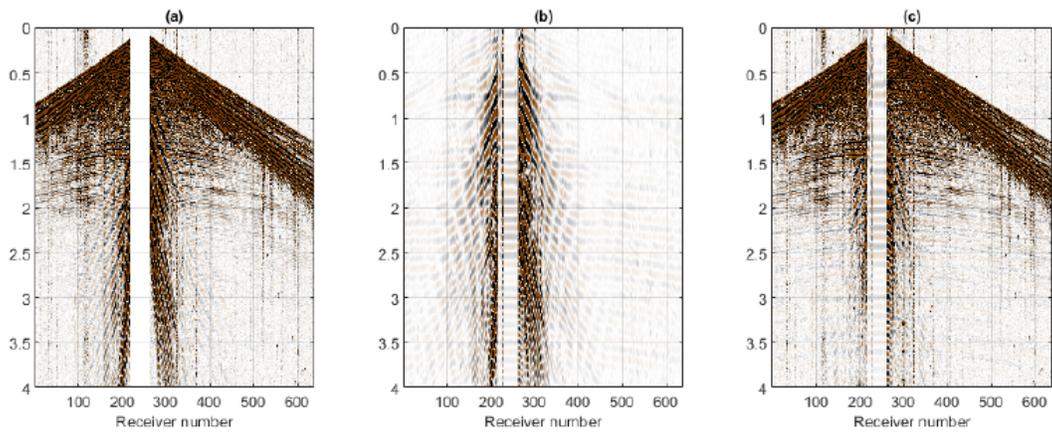


Figure 4.38: Separation of reflections and ground roll using least-squares inversion with sparse ( $l_1$ ) regularization using IRLS. a) 3D real shot gather. b) Estimated coherent noise (ground roll). c) Real observations (a) minus estimated coherent noise (b).

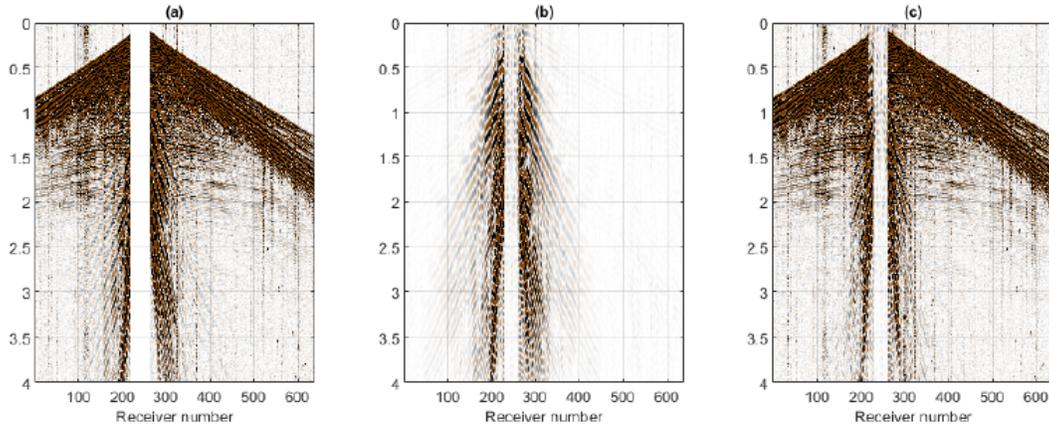


Figure 4.39: Separation of reflections and ground roll with least-squares inversion with sparse ( $l_1$ ) regularization using FISTA. a) 3D real shot gather. b) Estimated coherent noise (ground roll). c) Real observations (a) minus estimated coherent noise (b).

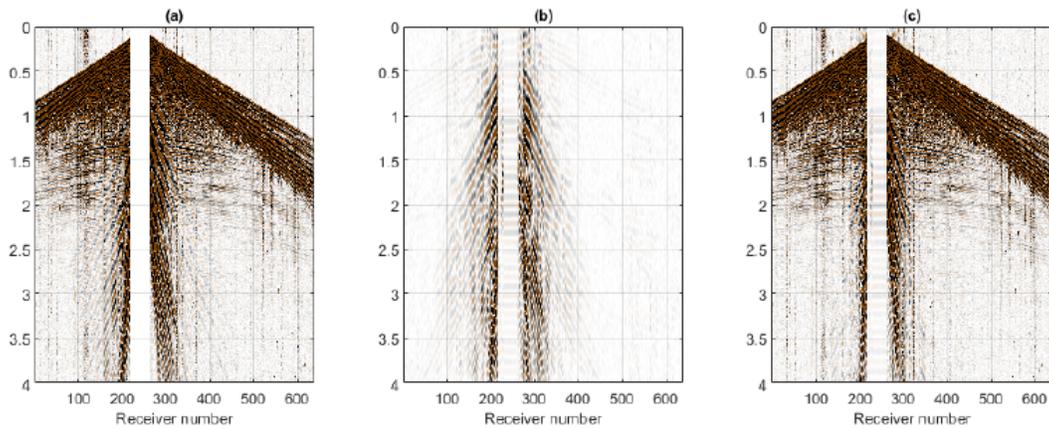


Figure 4.40: Separation of reflections and ground roll with robust inversion using ADMM. a) 3D real shot gather. b) Estimated coherent noise (ground roll). c) Real observations (a) minus estimated coherent noise (b).

Figure 4.41 shows a comparison for the 4 algorithms on the same zoomed part for the data, L2, IRLS, FISTA and ADMM and table 4.6 shows the parameters used for the above results. Both  $x$  and  $y$  axis show which part from main data is selected.

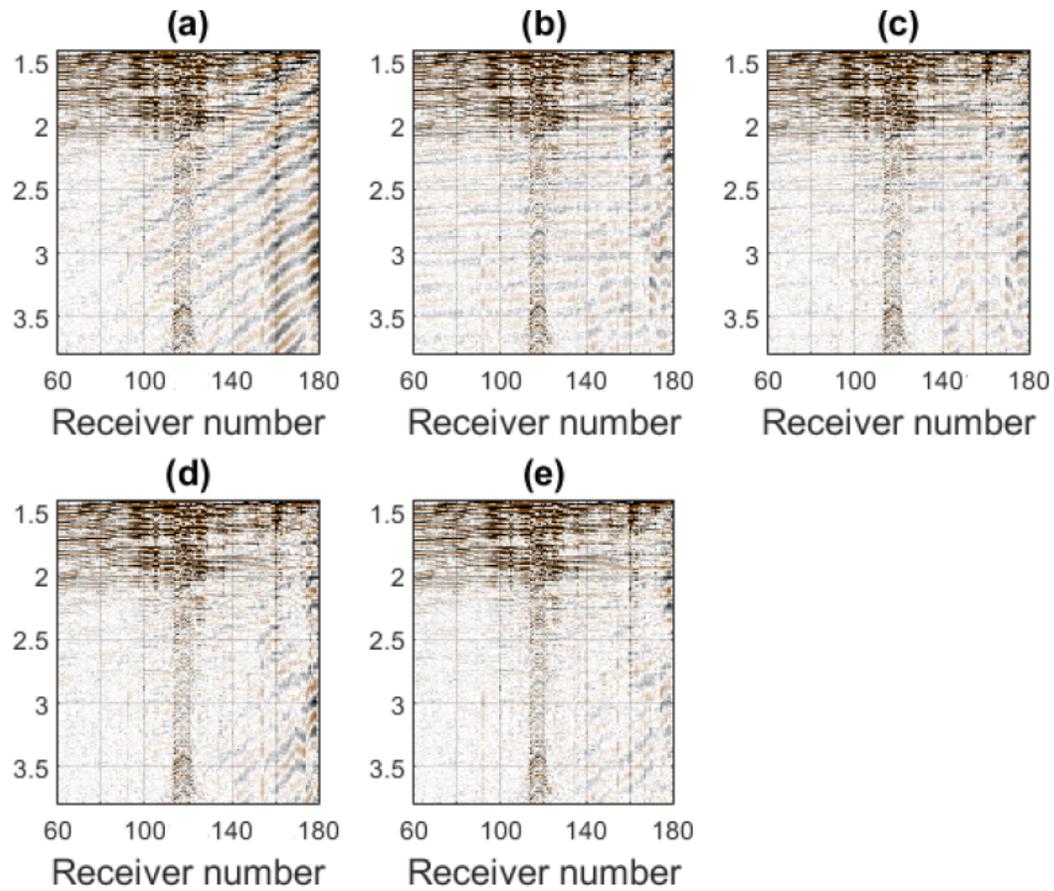


Figure 4.41: a) Zoomed area from raw data containing dispersive ground roll. b) Separation of reflections and ground roll using least-squares inversion with ( $l_2$ ) regularization, c) with ( $l_1$ ) regularization using IRLS, d) using FISTA and e) using ADMM

Table 4.6 shows the parameters and the results for each of the inversion algorithms used above:

<i>Method</i>	$\rho$	$\mu$	$\mu_1$	$\mu_2$	<i>tolerance</i>	<i>Iterations</i>	<i>runtime(s)</i>
L2	-	-	0.01	01	-	1	24.3
IRLS	-	-	0.01	1	-	10	144.43
FISTA	-	1024	-	-	$10^{-3}$	50	154.9
ADMM	1	0.001	-	-	$10^{-4}$	244	70.55

Table 4.6: Parameters used for inversion when  $\tau, v$  pairs of the model are unknown-2D real data

In the case where estimation of intercept time-velocity pairs is impossible, we expect our algorithms, or some of them to be able to find the correct answer. This is more challenging when ground roll has a complex character: strong dispersion and alias.

The above results (figures and table of parameters), wherever robust algorithms (penalty term or inversion) is used, the separation is successful. The ADMM algorithm needs the shortest processing time and no artifacts are visible.

FISTA has an increased computational time and some ground roll is left behind.

Finally, in the case of the quadratic penalty term, too many artifacts are present in the solution which show a false number of reflections.

## 4.7 Summary

Coherent linear noise exists in land data in the form of dispersive Rayleigh waves, known as ground roll. In this chapter, we explored several methods to attenuate it while preserving the signal's characteristics. We introduced the reader to least-squares inversion with penalty constraints as well as to robust inversion. We solved the  $L_2 - L_1$  problem with the well-know IRLS and FISTA methods. We then used the ADMM algorithm for robust inversion which seems to be a more logical approach for land data corrupted with erratic noise.

We applied these algorithms to both synthetic and real data and we explored two cases; one case where we know the intercept time and velocity pairs for our reflections and another case where we assume that it is difficult to define the pairs because ground roll is very strong and it masks most of seismic record's area. We use a dense grid in which we believe the pairs are located.

For synthetic examples, in the separation of reflections and ground roll, all algorithms lead to similar results. When examining the spectra of estimated coherent noise, it seems that least-squares inversion with sparse penalty norm and robust inversion have better outcomes, where the robust inversion has the best quality.

When coming to field data, we estimate the intercept time-velocity pairs through velocity analysis after Automatic Gain Control, because the reflections are too weak. Applying all our algorithms, the separation seems to be better when robust inversion or least squares inversion with robust regularization norm are used. These algorithms also seem to need the shortest processing time. That means the computational cost decreases.

We also explore the case when a semblance map is hard to be used for velocity estimation. Results show again that robust inversion or LS inversion with  $L_1$  term lead to a successful coherent noise attenuation within a short period of time. On the other hand, LS inversion with quadratic norm produces many artifacts in the final section which confuses us about the position of the actual reflections.

All the above algorithms show that separation of a highly dispersive and aliased ground roll is possible, while the amplitude of the signal is preserved.

---

---

## CHAPTER 5

---

### Conclusions

#### 5.1 Contributions

Ground roll can be aliased and dispersive, and these make its attenuation difficult. Many traditional methods, such as frequency filters, fail when the noise is aliased. This happens because coherent noise masks the reflections. If we try to remove it altogether, the signal will also be removed. In the opposite case, we try to preserve the signal of interest, a significant amount of surface waves will remain, and further processing will be complicated.

Based on this, we firstly explored an algorithm for simultaneously modelling reflections and dispersive ground roll in the  $f - x$  domain. Through two operators, coefficients are modeled into hyperbolas and dispersive noise. We compute synthetic 3D shot gathers with irregular distribution of source and receivers. In general, one should consider that modelling the ground roll is difficult due to the complexity of the real surface. So rather than using an operator that entails solving the wave equation for given subsurface parameters, we adopted simple operators that require only NMO velocities and a dense axis of dips (ray parameters) to capture complex ground roll.

I then tried to separate the ground roll from reflections based on a methodology inspired on the work of (Perkins and Zwaan, 2000 and Le Meur et al., 2008, who used the least-squares inversion method. This is a simple approach which preserves the signal amplitudes, and it can work well even when there is dispersion and alias.

For our solutions' stability, we added quadratic and sparse regularization norms to estimate the coefficients of linear noise in the FX domain, where we can easily handle one inverse

problem per temporal frequency.

In synthetic and real data, we conducted experiments with two cases. We use a semblance map to estimate the intercept time-velocity pairs and assume that ground roll is strong and masks most of the seismic section. We proved that ground roll could be separated from reflections when an approximate model of intercept times-velocities is provided.

Separation of dispersive and aliased ground roll from reflections using damped LS inversion leads to poor results as it creates many artifacts. We then tested the IRLS and the FISTA algorithm for the  $l_2 - l_1$  problem in synthetic and real data. The results were satisfying under the assumption that the data were contaminated by additive Gaussian noise. However, the FISTA algorithm needs more processing time, and this increases the computational cost.

Continuing from the previous step, I then applied robust inversion. I solve the robust and sparse inversion via ADMM. We assumed that additive noise can be modelled as a non-gaussian erratic noise. Such an assumption stems from considering typical land data surveys where we often observe anomalous erratic traces. These traces arise as a product of field operations, wind, debris produce by explosive sources, and in general, small near-surface heterogeneities non-modelable by our experiment.

The ground roll can have strong amplitudes in specific areas, such as for geophones close to the source. After separation reflections from ground roll, some noise leakage might still exists, which results in low quality near offset traces. This may cause problems for further seismic processing required for Amplitude versus Offset inversion (AVO).

## 5.2 Future developments

To further expand our work, one should consider including to the algorithms a term to handle dipping layers. The latter involves using a hyperbolic traveltime that considers an apex term (Schleicher et al. (1993); Larner et al. (1981)).

A physical transform for representing seismic data should be built with operators like  $\mathbf{A}_s$  and  $\mathbf{A}_c$ , which define data with simple kinematics. Additionally, localization via the windowing in t-x will limit the number of dips (coefficients) to be inverted and help the sparsity assumption. In other words, expanding the current operator to a Curvelet-like transform (Ma and Plonka (2010); Starck et al. (2002); Candes et al. (2006); Ying et al. (2005); Neelamani et al. (2008)) might be a challenging yet exciting problem for future research. The benefit is to add localization and scale to the proposed algorithms for making them more reliable for processing data with complex ground roll signatures.

To remove the leakage where the ground roll was strong, an additional filter that would capture these traces could improve the denoising after inversion steps. For this purpous, one might require to use a combination of methods applied in a cascade manner rather than a single method that attempts to correctly separate ground roll from reflections at all offsets.

# Bibliography

- Anstey, N. A., 1986, Part 1: Whatever happened to ground roll?: The Leading Edge, **5**, 40–45.
- Ascher, U., and F. Roosta-Khorasani, 2016, Algorithms that satisfy a stopping criterion, probably: Vietnam Journal of Mathematics, **44**, 49–69.
- Askari, R., and H. R. Siahkoochi, 2008, Ground roll attenuation using the s and xfk transforms: Geophysical Prospecting, **56**, 105–114.
- Beck, A., and M. Teboulle, 2009, A fast iterative shrinkage-thresholding algorithm for linear inverse problems: SIAM Journal on Imaging Sciences, **2**, 183–202.
- Beresford-Smith, G., and R. Rango, 1989, Suppression of ground roll by windowing in two domains: First Break, **7**.
- Berman, A., and R. J. Plemmons, 1974, Cones and iterative methods for best least squares solutions of linear systems: SIAM Journal on Numerical Analysis, **11**, 145–154.
- Bing, Z., S. Greenhalgh, and C. Sinadinovski, 1992, Iterative algorithm for the damped minimum norm, least-squares and constrained problem in seismic tomography: Exploration Geophysics, **23**, 497–505.
- Bóna, A., and M. A. Slawinski, 2003, Fermat’s principle for seismic rays in elastic media: Journal of applied geophysics, **54**, 445–451.
- Boyd, S., N. Parikh, E. Chu, B. Peleato, J. Eckstein, et al., 2011, Distributed optimization and statistical learning via the alternating direction method of multipliers: Foundations and Trends® in Machine learning, **3**, 1–122.
- Boyle, J. P., and R. L. Dykstra, 1986, A method for finding projections onto the intersection of convex sets in hilbert spaces, *in* Advances in order restricted statistical inference: Springer, 28–47.
- Burrus, C. S., J. A. Barreto, and I. W. Selesnick, 1994, Iterative reweighted least-squares design of fir filters: IEEE Transactions on Signal Processing, **42**, 2926–2936.
- Campbell, S., H. Wang, and J. Hoffmann, 2001, Combined nmo response and stack array analysis of model-based data in an environment of strong ground roll, *in* SEG Technical Program Expanded Abstracts 2001: Society of Exploration Geophysicists, 1969–1972.

- Candes, E., L. Demanet, D. Donoho, and L. Ying, 2006, Fast discrete curvelet transforms: *Multiscale Modeling & Simulation*, **5**, 861–899.
- Chambolle, A., and C. Dossal, 2015, On the convergence of the iterates of "fista".
- Chartrand, R., and W. Yin, 2008, Iteratively reweighted algorithms for compressive sensing: 2008 IEEE International Conference on Acoustics, Speech and Signal Processing, 3869–3872.
- Clement, W., 1973, Basic principles of two-dimensional digital filtering: *Geophysical prospecting*, **21**, 125–145.
- Daubechies, I., R. DeVore, M. Fornasier, and C. S. Güntürk, 2008, Iteratively reweighted least squares minimization for sparse recovery: *Communications on Pure and Applied Mathematics*, **63**, 1–38.
- de Matos, M. C., and P. L. M. Osório, 2002, Wavelet transform filtering in the 1d and 2d for ground roll suppression, *in* SEG Technical Program Expanded Abstracts 2002: Society of Exploration Geophysicists, 2245–2248.
- Dragoset, B., 2000, Introduction to air guns and air-gun arrays: *The Leading Edge*, **19**, 892–897.
- Duijndam, A., M. Schonewille, and C. Hindriks, 2001, Irregular and sparse sampling in exploration seismology, *in* *Nonuniform Sampling*: Springer, 479–518.
- Eckstein, J., and D. P. Bertsekas, 1992, On the douglas-rachford splitting method and the proximal point algorithm for maximal monotone operators: *Mathematical Programming*, **55**, 293–318.
- Fyfe, D., and P. Kelamis, 1992, Removing coherent noise using the linear radon transform: Presented at the 54th EAEG Meeting.
- Goldstein, T., and S. Osher, 2009, The split bregman method for l1-regularized problems: *SIAM journal on imaging sciences*, **2**, 323–343.
- Gorodnitsky, I. F., and B. D. Rao, 1997, Sparse signal reconstruction from limited data using focuss: a re-weighted minimum norm algorithm: *IEEE Transactions on Signal Processing*, **45**, 600–616.
- Goudarzi, A. R., and M. Ali Riahi, 2012, Adaptive seismic ground roll attenuation using the double density dual tree discrete wavelet transform (dwt) method: *Earth Sciences Research Journal*, **16**, 113–120.
- Guspi, F., and B. Introcaso, 2000, A sparse spectrum technique for gridding and separating potential field anomalies: *Geophysics*, **65**, 1154–1161.
- Hampson, D., 1987, The discrete radon transform: A new tool for image enhancement and noise suppression, *in* SEG Technical Program Expanded Abstracts 1987: Society of Exploration Geophysicists, 141–143.
- Hosseini, S. A., A. Javaherian, H. Hassani, S. Torabi, and M. Sadri, 2015, Shearlet transform in aliased ground roll attenuation and its comparison with fk filtering and curvelet

- transform: *Journal of Geophysics and Engineering*, **12**, 351–364.
- Kijko, A., 1994, Seismological outliers: L 1 or adaptive lp norm application: *Bulletin of the Seismological Society of America*, **84**, 473–477.
- Larner, K. L., L. Hatton, B. S. Gibson, and I.-C. Hsu, 1981, Depth migration of imaged time sections: *Geophysics*, **46**, 734–750.
- Le Meur, D., N. Benjamin, R. Cole, and M. Al Harthy, 2008, Adaptive groundroll filtering: 70th EAGE Conference and Exhibition incorporating SPE EUROPEC 2008.
- Liu, G., X. Meng, J. Ni, Z. Chen, and D. Zhang, 2018, Evaluation of the 2d reflection seismic method toward the exploration of thrust-controlled mineral deposits in southwestern fujian province, china: *Geophysics*, **83**, B209–B220.
- Liu, J., and K. J. Marfurt, 2004, 3-d high resolution radon transforms applied to ground roll suppression in orthogonal seismic surveys, *in* SEG Technical Program Expanded Abstracts 2004: Society of Exploration Geophysicists, 2144–2147.
- Liu, Y., 2010, Local time-frequency transform and its application to ground-roll noise attenuation, *in* SEG Technical Program Expanded Abstracts 2010: Society of Exploration Geophysicists, 3711–3716.
- Ma, J., and G. Plonka, 2010, The curvelet transform: *IEEE signal processing magazine*, **27**, 118–133.
- Malehmir, A., R. Durrheim, G. Bellefleur, M. Urosevic, C. Juhlin, D. J. White, B. Milkereit, and G. Campbell, 2012, Seismic methods in mineral exploration and mine planning: A general overview of past and present case histories and a look into the future: *Geophysics*, **77**, WC173–WC190.
- Menke, W., 1989, *Geophysical data analysis: Discrete inverse theory*: Elsevier Science. International Geophysics.
- Naghizadeh, M., and M. Sacchi, 2011, Ground-roll elimination by scale and direction guided curvelet transform: Presented at the 73rd EAGE Conference and Exhibition incorporating SPE EUROPEC 2011.
- , 2018, Ground-roll attenuation using curvelet downscaling: *Geophysics*, **83**, V185–V195.
- Neelamani, R., A. I. Baumstein, D. G. Gillard, M. T. Hadidi, and W. L. Soroka, 2008, Coherent and random noise attenuation using the curvelet transform: *The Leading Edge*, **27**, 240–248.
- Perkins, C., and M. Zwaan, 2000, Ground roll attenuation: 62nd EAGE Conference & Exhibition.
- Porsani, M. J., M. G. Silva, P. E. Melo, and B. Ursin, 2009, Ground-roll attenuation based on svd filtering: *SEG Technical Program Expanded Abstracts 2009*, 3381–3385.
- Russell, B. H., 1988, *Introduction to seismic inversion methods*: Society of Exploration Geophysicists.

- Sacchi, M. D., 1997, Reweighting strategies in seismic deconvolution: *Geophysical Journal International*, **129**, 651–656.
- Sacchi, M. D., and T. J. Ulrych, 1995, High-resolution velocity gathers and offset space reconstruction: *Geophysics*, **60**, 1169–1177.
- , 1996, Estimation of the discrete Fourier transform, a linear inversion approach: *Geophysics*, **61**, 1128–1136.
- Schleicher, J., M. Tygel, and P. Hubral, 1993, Parabolic and hyperbolic paraxial two-point traveltimes in 3d media: *Geophysical Prospecting*, **41**, 495–513.
- Schonewille, M., and A. Duijndam, 1997, Parabolic radon transform and z-squared fk-transform? aliasing and efficiency, *in* SEG Technical Program Expanded Abstracts 1997: Society of Exploration Geophysicists, 1410–1413.
- Schuster, G. T., 2017, *Seismic inversion*: Society of Exploration Geophysicists.
- Shearer, P. M., 2019, *Introduction to seismology*: Cambridge university press.
- Shen, Y., B. Li, and Y. Chen, 2011, An iterative solution of weighted total least-squares adjustment: *Journal of geodesy*, **85**, 229–238.
- Simaan, M., and P. Love, 1984, Optimum suppression of coherent signals with linear moveout in seismic data: *Geophysics*, **49**, 215–226.
- Spingarn, J. E., 1985, Applications of the method of partial inverses to convex programming: decomposition: *Mathematical Programming*, **32**, 199–223.
- Starck, J.-L., E. J. Candès, and D. L. Donoho, 2002, The curvelet transform for image denoising: *IEEE Transactions on image processing*, **11**, 670–684.
- Tikhonov, A. N., and V. Y. Arsenin, 1977, *Solution of ill-posed problems*: V.H. Winston, Washington DC.
- Tropp, J. A., and A. C. Gilbert, 2007, Signal recovery from random measurements via orthogonal matching pursuit: *IEEE Transactions on information theory*, **53**, 4655–4666.
- Udias, A., and E. Buforn, 2017, *Principles of seismology*: Cambridge University Press.
- Van den Berg, E., and M. P. Friedlander, 2011, Sparse optimization with least-squares constraints: *SIAM Journal on Optimization*, **21**, 1201–1229.
- Wang, B., Y. Zhang, W. Lu, and J. Geng, 2019, A robust and efficient sparse time-invariant radon transform in the mixed time-frequency domain: *IEEE Transactions on Geoscience and Remote Sensing*.
- Wang, P., and K. Nimsaila, 2014, Fast progressive sparse Tau-P transform for regularization of spatially aliased seismic data: 84th Annual International Meeting, SEG, Expanded Abstracts, 3589–3593.
- Yajima, T., and H. Nagahama, 2007, Kawaguchi space, zermelo’s condition and seismic ray path: *Nonlinear Analysis: Real World Applications*, **8**, 130–135.
- Yilmaz, Ö., 2001, *Seismic data analysis: Processing, inversion, and interpretation of seismic data*: Society of exploration geophysicists.

- Yin, W., S. Osher, D. Goldfarb, and J. Darbon, 2008, Bregman iterative algorithms for  $\ell_1$ -minimization with applications to compressed sensing: *SIAM Journal on Imaging sciences*, **1**, 143–168.
- Ying, L., L. Demanet, and E. Candes, 2005, 3d discrete curvelet transform: *Wavelets XI*, International Society for Optics and Photonics, 591413.
- Yu, Z., G. A. McMechan, J. F. Ferguson, and P. D. Anno, 2002, Adaptive wavelet filtering of seismic data in the wavelet transform domain: *Journal of Seismic Exploration*, **11**, 223–246.
- Zhang, H., X.-H. Chen, and L.-Y. Zhang, 2017, 3d simultaneous seismic data reconstruction and noise suppression based on the curvelet transform: *Applied Geophysics*, **14**, 87–95.
- Zwartjes, P., and A. Gisolf, 2007, Fourier reconstruction with sparse inversion: *Geophysical Prospecting*, **55**, 199–221.