

Seismic ground-roll separation using sparsity promoting L1 minimization

by

Carson Edward Yarham

B.Sc., The University of British Columbia, 2004

A THESIS SUBMITTED IN PARTIAL FULFILMENT OF
THE REQUIREMENTS FOR THE DEGREE OF

Master of Science

in

The Faculty of Graduate Studies

(Geophysics)

THE UNIVERSITY OF BRITISH COLUMBIA

(Vancouver)

April, 2008

© Carson Edward Yarham 2008

Abstract

The removal of coherent noise generated by surface waves in land based seismic is a prerequisite to imaging the subsurface. These surface waves, termed as ground roll, overlay important reflector information in both the t-x and f-k domains. Standard techniques of ground roll removal commonly alter reflector information as a consequence of the ground roll removal. We propose the combined use of the curvelet domain as a sparsifying basis in which to perform signal separation techniques that can preserve reflector information while increasing ground roll removal. We examine two signal separation techniques, a block-coordinate relaxation method and a Bayesian separation method. The derivations and background for both methods are presented and the parameter sensitivity is examined. Both methods are shown to be effective in certain situations regarding synthetic data and erroneous surface wave predictions. The block-coordinate relaxation method is shown to have major weaknesses when dealing with seismic signal separation in the presence of noise and with the production of artifacts and reflector degradation. The Bayesian separation method is shown to improve overall separation for both seismic and real data. The Bayesian separation scheme is used on a real data set with a surface wave prediction containing reflector information. It is shown to improve the signal separation by recovering reflector information while improving the surface wave removal. The abstract contains a separate real data example where both the block-coordinate relaxation method and the Bayesian separation method are compared.

Table of Contents

Abstract	ii
Table of Contents	iii
List of Tables	iv
List of Figures	v
Preface	x
Acknowledgments	xi
Dedication	xii
Statement of Co-Authorship	xiii
1 Introduction	1
1.1 Theme	2
1.1.1 Sparse representations of seismic data	2
1.1.2 Signal separation background	6
1.1.3 Extension to redundant transforms	8
Bibliography	10
2 Ground roll prediction and separation in the sparse curvelet domain	13
2.1 Introduction	13
2.1.1 Sparse representations of seismic data	14
2.2 Signal separation	16
2.2.1 Block-coordinate relaxation method	18
2.2.2 Bayesian separation method	21
2.3 Application to synthetic data	24
2.3.1 Synthetic data generation	24

Table of Contents

2.3.2	Separation scheme comparison	25
2.3.3	Parameter sensitivity	27
2.4	Application to real field data	27
2.5	Discussion	38
2.6	Conclusion	39
2.7	Acknowledgments	40
Bibliography		41
3 Conclusions		44
3.1	Signal separation methods	45
3.2	Open and future research	47
Bibliography		49
 Appendices		
A Examples		55
A.1	Oz Yilmaz 25 dataset	55
A.1.1	Standard surface wave estimation	55
A.1.2	Curvelet domain estimation	62
Bibliography		73

List of Tables

2.1	The block-coordinate relaxation algorithm. Table from Herrmann et al. (2007b)	21
2.2	The Bayesian-separation algorithm.	24
2.3	Signal-to-noise ratios for Bayesian and block-coordinate relaxation (BCR) separation methods. The SNR for the synthetic data is -1.67 for all tests except for the third test which has added white noise decreasing the SNR to -1.92. These values were calculated with respect to data that has no ground roll or noise present.	26
2.4	Sensitivity of the parameters associated with the Bayesian signal separation scheme. The parameters λ_1^* , λ_2^* and η^* refer to 10, 2 and 3.5 respectively. These parameters are the same set used to generate the results of the second test in Table ?? and are chosen by the expected sparsity of the estimated signal and the fidelity of the ground roll prediction. Small adjustments were then made by parameter search to select the most effective values for the synthetic data.	33
2.5	Sensitivity of the parameters associated with block-coordinate relaxation. These parameters are associated with the aggressiveness of the soft thresholding operator defined in the outer loop and then applied in the inner loop. In this case, the values $C_1 = C_2 = 0.4$ are used.	33

List of Figures

1.1	Representations of curvelets in the spatial and Fourier domains. (a) Four separate curvelets at different scales and angles. (b) The associated Fourier domain for the displayed curvelets. The four curvelets represented display four different scales and angles with the finest scales associated with the highest frequencies in the Fourier domain.	6
1.2	Decay rates for synthetic seismic data. (a) shows the decay rate of the entire record with (b) and (c) showing the decay rates for the exact reflector and surface wave components respectively.	7
2.1	Descriptive plots showing the relation of the specific transforms to the α and β parameters. (a) is augmented with physical representations of the individual atoms and (b) contains representations of the Fourier spectrum. This figure is adapted from Demanet and Ying (2007).	15
2.2	Decay rates for real seismic waveforms. (a) shows the decay rate of the entire record with (b) and (c) showing the decay rates for the estimated reflector and surface wave components respectively. The curvelet domain shows the most rapid decay rates of all the transforms examined while the t-x and f-k domains have the slowest decay making them the least efficient for our methods.	17
2.3	Synthetic model for pressure wave velocity used in generation of the (a) reflectors and (b) ground roll. Shear wave velocity and density models are not shown but exhibit the same overall structure. Synthetic data was generated separately and combined to produce the data shown in Figure ??	25

List of Figures

2.4	Synthetic data used to test the signal separation schemes. The full synthetic data set (a), the reflector component (b) and the generated surface wave (c). The signal to noise ratio of the full synthetic data is -1.67.	28
2.5	Synthetic results for the estimated reflector signal using the Bayesian separation scheme. The generated results using predictions consisting of the exact noise (a), 5% error on modeling parameters (b), 5% error on modeling parameters with added white noise (c), Hilbert transform of the exact noise (d).	29
2.6	Synthetic results for the estimated surface wave signal using the Bayesian separation scheme. The generated results using an exact noise prediction (a), 5% error on modeling parameters (b), 5% error on modeling parameters with added white noise (c), Hilbert transform of the exact noise prediction (d).	30
2.7	Separation results using a prediction containing 5% error and the block-coordinate relaxation method (a) and the Bayesian separation method (b). The initial dataset has a SNR of -1.67 dB while the SNR of the separated results are 9.41 and 9.58 from the block-coordinate relaxation method and the Bayesian separation method respectively.	31
2.8	Estimated surface waves from the block-coordinate relaxation (a) and Bayesian (b) solutions to the 5% model error prediction for the synthetic data. Whereas Table ?? shows the SNR for the Bayesian solution is roughly equal to the block-coordinate relaxation method, we can see reflector information has been estimated as part of the surface wave where as the Bayesian solution has been more conservative and estimated less reflector information with the surface wave.	32
2.9	The real dataset to be improved with the Bayesian separation method. This dataset comes from a line that was located near the source thus causing a dominant ground roll signal in the center of the image. This data is also contaminated with several noisy traces that become increasingly obvious for later times in the signal as the data has had an applied time power gain added.	34
2.10	The predictions to be used with the Bayesian separation method for the (a) reflectors and (b) surface wave. The reflector prediction is a supplied denoised data set with the surface wave prediction as the difference between the total data and the reflector prediction.	35

List of Figures

2.11	The Bayesian (a) reflector and (b) surface wave estimates for the real data set after five iterations. These images show improved energy content for both the reflector and surface wave estimates. There is still a small amount of linear noise in the estimated reflectors but it is less than the original prediction.	36
2.12	The difference between the predicted surface wave and the Bayesian estimated surface wave. This image shows the energy transferred between predictions through the estimation scheme. As we can see, reflector information, particularly in the top left section of the image and ground roll energy has been transferred back to their appropriate estimates.	37
A.1	Oz25 data set from Yilmaz (2001). This data set exhibits high amplitude, aliased ground roll with varying reflector amplitude underneath.	56
A.2	Ground roll removal by F-K filtering. (a) A conservative estimate of the reflectors based on dip filtering in the Fourier domain. (b) The associated surface wave that is predicted during the F-K filtering process.	57
A.3	Results of signal block-coordinate relaxation method. (a) The estimated reflectors. (b) The associated estimated surface wave. We can see that some reflector information has been moved to the surface wave estimate while the reflector estimate still contains ground roll. There is a component of noise centered at one second and zero offset that has not been identified as part of the coherent noise either by the angle filtering or by the signal separation scheme and other methods need to be examined to address this signal. A large amount of curvelet like artifacts have also been generated during the separation process.	58
A.4	The difference between the predicted surface wave and the block-coordinate relaxation estimated surface wave. We can see a large amount of ground roll that has been identified as part of the surface wave estimate. Some higher amplitude reflector information has also been identified as ground roll and erroneously moved to the surface wave estimate.	59

List of Figures

A.5	Results of the Bayesian signal separation process. (a) The Bayesian estimated reflectors. (b) The associated estimated surface wave. In this case, very little to no reflector information has ended up in the estimated surface wave. Some ground roll still remains in the reflector estimate.	60
A.6	The difference between the predicted surface wave and the Bayesian estimated surface wave. Again, we can see a large amount of ground roll that has been identified as part of the surface wave estimate. In this case, all reflector information has been preserved where the block-coordinate relaxation method removed some high amplitude reflectors.	61
A.7	Difference between the block-coordinate method relaxation and Bayesian separation method surface wave estimations. Again, we can see the reflector information removed by the block-coordinate relaxation method and introduced artifacts. There is also some minor difference in the surface wave estimations.	62
A.8	The original data (a) and the data after a threshold defined by Equation ?? has been applied with $C_T = 1$ (b).	64
A.9	Curvelet scale threshold results for varying values of C_T . The C_T constant is equal to (a) 10, (b) 2, (c) 0.5, (d) 0.1.	65
A.10	A conservative estimate of the surface wave is created by performing a curvelet scale thresholding followed by application of a dip filter. We can see that the predicted reflectors (a) and the predicted surface wave (b) are not matched in amplitude but the prediction of the surface wave is enough to used as a prediction for the signal separation schemes.	66
A.11	The block-coordinate relaxation estimated reflectors (a) and estimated surface wave (b) generated using a curvelet scale thresholding prediction for the surface wave and reflectors. While the majority of the ground roll is removed, strong artifacts were generated by the signal separation scheme. As these artifacts are curvelet like, they begin to look like reflection events which is unacceptable.	67
A.12	The components of the surface wave that were not identified through the initial prediction but identified as part of the coherent noise signal through the block solver scheme. While all of the ground roll was removed, we see the generated artifacts and reflector information that has been removed from the data.	68

List of Figures

A.13	The Bayesian estimated reflectors (a) and estimated surface wave (b) generated by a curvelet scale thresholding with dip filter prediction. While there is more ground roll present in the reflector estimation, we can see that no reflector information is present in the surface wave estimation. Very few, if any artifacts have been generated with this method, especially when compared to the block coordinate relaxation method.	69
A.14	The difference between the curvelets based prediction and the estimated surface wave generated by the Bayesian separation scheme. There are no reflectors present in this difference indicating that no reflectors were estimated as ground roll.	70
A.15	The difference between the surface wave estimations of the two different prediction methods for the (a) block-coordinate relaxation and (b) Bayesian separation method. The shape of the removed data suggests that the larger curvelet coefficients that were not identified with the original f-k filtering technique have been removed by using a curvelet based thresholding scheme to predict the surface wave.	71

Preface

This thesis was prepared with Madagascar, a reproducible research software package available at rsf.sf.net. Many results shown in this thesis are directly linked to the codes that generated them. This reproducibility allows direct access to the experiments performed in this work allowing transparency and dissemination of knowledge for the Seismic Laboratory for Imaging and Modeling (SLIM), its sponsors, and the entire research community as a whole.

The programs required to reproduce the reported results are Madagascar programs written in C/C++, Matlab[®], or Python. The numerical algorithms and applications are mainly written in Python as part of SLIMpy (slim.eos.ubc.ca/SLIMpy) with a few exceptions written in Matlab[®] or Python as external files not part of the SLIMpy package.

Acknowledgments

I would first like to thank Felix Herrmann for both his leadership and support since his first days at UBC. His ability to spark curiosity in complex and engaging subjects through his seemingly endless enthusiasm is an inspiration that will leave a deep impression on me for many years to come. Without his vision and courage none of this (and much, much more) would have been possible.

I would also like to thank Henryk Modzelewski, Gilles Hennenfent, Darren Thompson, Cody Brown, Sean Ross Ross and the rest of the SLIM team past and present for their encouragement, commitment and spirit. It has been an honor to be part of this dynamic group and to witness the amazing process of development that it has undergone.

I am very fortunate to have Michael Bostock and Ronald Clowes as part of my supervisory committee. The knowledge, insight and inspiration that they have provided through my time at UBC has been invaluable.

I would also like to thank Dough Oldenburg for being part of my thesis defence and for providing some of the most informative and engaging courses of my university career, both undergraduate and graduate.

Without the support and love of my parents, Lance and Irene, my brother Bryan, sister Michelle and the rest of my family, I certainly would never have accomplished any of this. They have my unending thanks and love.

Finally, I would like to thank Crystal McConeghy. As a source of motivation, she has provided me with the strength and motivation to accomplish the seemingly impossible and my gratitude is unending.

To my family.

Statement of Co-Authorship

Chapter 2 was written with Felix J. Herrmann. The manuscript was jointly written with Felix making major contributions to the theoretical parts. I wrote the more applied parts as well as preformed the research and coding outside of the SLIMpy applications.

Chapter 1

Introduction

All processing of land based seismic faces the challenge of coherent noise contamination by surface waves generally termed as ground roll. The dominant component of ground roll consists of channeled Rayleigh waves within the near surface, low velocity layers (Olhovich, 1964; Saatcilar and Canitez, 1988). Ground roll can typically be identified due to its low velocity, low frequency content and high amplitude. The magnitude of this coherent noise often obscures underlying reflector information and thus it needs to be removed early in the processing stage.

One method of dealing with ground roll can be implemented at the stage of survey design. Arrangement of geophone and shot arrays can be designed to limit information from specific wave numbers and thus suppress the initial ground roll signature (McKay, 1954; Holzman, 1963). This approach can result in several complications. By suppressing certain wave numbers to exclude ground roll, any reflector information within those wave numbers will also be suppressed. The loss of information within certain wave numbers is generally something that is not desired. Even if noise is recorded, filtering techniques can often be applied post acquisition to remove the noise. Challenges are also faced when attempting to model survey design around wave number suppression since most survey design is ultimately dictated by the general site conditions.

The low frequency, low velocity components of ground roll have lead to the most common methods of ground roll suppression utilizing the Fourier domain. Methods such as band-pass filtering (Yilmaz, 2001), multichannel filtering (Galbraith and Wiggins, 1968), spectral balancing (Coruh and Costain, 1983), and various other flavors of frequency and velocity filtering (Embree et al., 1963; Fail and Grau, 1963; Treitel et al., 1967; Harlan et al., 1984; Gelişli and Karşlı, 1988) have been utilized in industry for many years due to their low cost and general simplicity. These methods rely on a good separation of the frequency components associated with ground roll and reflectors. Unfortunately, a good separation of reflector information and ground roll in the Fourier domain is generally uncommon. There are often other complications such as spatial aliasing of the ground roll itself

and overlapping ground roll and reflector information in the Fourier domain. While Fourier methods have been used extensively, several articles (McMechan and Sun, 1991; Liu, 1999; Karsli and Bayrak, 2004) have shown that when these methods are applied, distortion of reflector information can also be a consequence of denoising.

Other methods of ground roll removal have also been utilized. The use of the Karhunen-Loeve transform (Londono et al., 2005) generally causes less reflector distortion but is still limited in its ability to identify the ground roll. Radon transforms (Russel et al., 1990; Trad et al., 2003) have also been shown to work for ground roll removal but tend to be very data specific and could likely benefit from our signal separation methods to avoid reflector degradation. The Wiener-Levinson algorithm (Karsli and Bayrak, 2004) can also be applied to this problem but seems to lack the ability to recover information beneath the ground roll, thus leaving a definite shadow. One and two-dimensional wavelet transforms have also been applied to the problem of ground roll (Deighan and Watts, 1997; Zhang and Ulrych, 2003). While these show promise, wavelets are generally not as parsimonious as curvelets for representations of the same signals of ground roll and reflector information and care must also be taken to avoid reflector degradation.

1.1 Theme

The main theme of this thesis is twofold. First, to utilize the curvelet domain as a sparse representation of seismic data to simplify ground roll removal. Second is to make use of the sparsity provided by the curvelet domain to perform a subtraction of ground roll from reflector information through the use of ℓ_1 minimization schemes. The two main separation schemes that are examined are a block-coordinate relaxation method and a Bayesian separation scheme. Each scheme is examined, tested and utilized in this context.

1.1.1 Sparse representations of seismic data

To simplify the problem of ground roll signal separation, the data can be transformed into a domain that is sparse, that is, it generates a coefficient vector that contains a small number of large coefficients relative to its length. If our signal components were composed of well separated and distinct frequencies, we would expect that the Fourier domain would be the best choice for a sparse representation. While frequency information can be used to represent wavefronts, generally, many are required and overlap is common. The frequency spectrum of the response of the reflectors and the surface

wave are anything but sparse and tend to overlap in the Fourier domain. This situation can be made worse by undersampling in the spatial domain thus aliasing the surface wave number components.

As Figure 1.2(a) shows, the Fourier spectrum is not the most sparse domain for seismic signal representation. This is due to the fact that a seismic signal is produced by the convolution of the source function and the response of the earth, each of which contain a bandwidth limited but continuous range of frequencies. This is compounded by the addition of surface waves to this signal which generally dominate low frequencies and tend to overlap with the frequency content of the reflector information.

Naturally, a seismic source is a wavelet package that, while generally not known, is convolved with the Green's function of the Earth to produce a superposition of wavelets along a recorded trace. Thus, particularly for a wavelet transform with an appropriate source wavelet, a very sparse representation can be achieved. In practice, this is very difficult and lacks the ability to represent the continuity of wave fronts. For this, we look to two-dimensional wavelet transforms.

Two-dimensional waveforms

As the Fourier transform provides a sparse representation of frequency components, there is a collection of transforms that allow a sparse representation of two and three-dimensional wave components. As with the Fourier transform, their success lies in the similarity of the physical representation of the basic elements of the transform to the signals they are trying to approximate. This is accomplished by the use of prototype waveforms to generate a library of components based on parameters of scale, location and/or angle which allow them the variation to match the different wavefronts that exist in seismic signals. The corresponding transform coefficients are a product of the superposition of the waveform and the image at a specific location and orientation.

The extension of wavelets from one to multiple dimensions can be accomplished in several ways. To examine these, we will represent the multiscale nature and directionality as α and β in the same fashion as Demanet and Ying (2007) with relation to the dilation matrix

$$D_j = \begin{bmatrix} 2^{j\alpha} & 0 \\ 0 & 2^{j\beta} \end{bmatrix}, \quad (1.1)$$

which controls the general physical wavelet shape. A specific transform can be represented as multiscale ($\alpha = 1$) or not ($\alpha = 0$) with localized and

poorly directional elements ($\beta = 1$) or extended and fully directional elements ($\beta = 0$). The essential support of $\varphi(k)$ is approximately of size $2^{-\alpha*j}$ by $2^{-\beta*j}$ with j representing the scale of the wave packet with oscillations of an approximate wavelength 2^{-j} . In the Fourier domain, the two-dimensional essential support consists of two equal areas approximately equal to $2^{\alpha*j}$ by $2^{\beta*j}$ for scale j at opposing angles and equal distance ($\sim 2^j$) from the origin. The range of these properties lead to variety in the ability of the specific domain to efficiently represent seismic signals. A transform's ability to efficiently represent a seismic signal can be judged by examining the sorted and length normalized coefficient vector decay rates (Figure 1.2). A decay curve with a rapid decay to near zero will represent a more parsimonious representation of the seismic image. This is examined for a synthetic data set, as well as the exact reflectors and ground roll. A real data set is examined in the following chapter with similar results.

Two-dimensional wavelets (Daubechies, 1992) can be formulated as an extension of one-dimensional wavelets by applying the wavelet in the time, offset, and diagonally to the image. The multiscale nature of wavelets are preserved ($\alpha = 1$) but they are poorly directional ($\beta = 1$). This is to be expected since this basis function is isotropic. Since seismic wave fields contain wavefronts moving in all directions, we expect that it will require a larger amount of two-dimensional wavelet elements to properly represent them. This representation could be improved if anisotropic wave packets are used and this is what we find as shown in Figure 1.2(a).

The most obvious solution to the lack of directionality came in the form of the ridgelets (Candès, 1999). Like wavelets, they are a two-dimensional representation of a prototype wave form but unlike wavelets they do not scale and are fully directional elements that extend over the length of the image. In a more familiar sense, they are similar to a radon transform convolved with a wavelet across and along the radon line. With $\alpha = 0$ and $\beta = 1$, ridgelets are dyadic and highly directional. This has the potential to be a very efficient way of representing plane wave fronts but loses its efficiency when wavefronts are discontinuous, curved or otherwise more complicated than simple linear waves. Ground roll, even with its definitively linear structure, does not generally appear in a continuous wave front due to a lack of spatial sampling and cannot be represented efficiently with ridgelets as we see in Figure 1.2(c).

While similar in physical support to two-dimensional wavelets, Wave Atoms are anisotropic to allow representations of directionality. Specifically designed to represent oscillatory waves, Wave Atoms have abandoned the circular symmetry of standard wavelets in favor of harmonic oscillations in

one direction and smoothness along the wavefront. To describe directional waveforms, a third parameter l is needed to represent the angle of the wavelet package which in two-dimensions has now become $\varphi_{j,k,l}$. This transform sets $\beta = 1/2$ which corresponds to a parabolic scaling relation connecting the scale and size of the waveform with its frequency spectrum. It is this parabolic scaling relation that has been shown to efficiently represent wave propagation from a point source by imposing a second dyadic decomposition on the Fourier domain (Seeger et al., 1991). For Wave Atoms, this equates to $wavelength \simeq diameter^2$. While Wave Atoms are effective in displaying certain wave fields, generally, they are not the best choice for seismic data. Their ability to represent oscillatory fields can cause problems when data sets with large variations are encountered making curvelets a better choice. While not as efficient as curvelets, wave atoms generally outperform wavelets under the same circumstances.

Curvelets, as shown in Figure 1.1, represent the most sparse domain for seismic data (Candes et al., 2005). As seen in Figure 1.2, this is the case for our test synthetic data and will also be shown to be the case for real data in the following chapter. While it is not strictly sparse in the sense that the two-dimensional curvelet domain has a redundancy of over seven times, proportional to the length, it stores more energy on less coefficients than any of the other transforms. This is accomplished by the unique shape of curvelets, allowing local detection of wavefronts due to their high correlation values. As the curvelet domain is energy preserving, the concentration of the energy associated with the individual wave fronts correlated with the curvelets produces a very sparse representation of seismic data.

The curvelet wave packets are defined by parameters of $\alpha = 1$ and $\beta = 1/2$ to compare with the previous transforms. Similar to Wave Atoms, curvelets also obey a parabolic scaling principle with $length^2 \simeq width$, with the frequency of the oscillations across the curvelet increasing with increasing scales. Unsimilar to Wave Atoms, they have a physical representation closer to a ridgelet of a finite length than to a wavelet and become increasingly anisotropic and more needle like with finer and finer scales. The increasingly anisotropic nature of curvelets can be attributed to the tiling of the frequency plane into a collection of angular wedges positioned at different scales with the number of angular wedges doubling at every second scale. These wedges in the frequency plane are strictly localized while the curvelets have a rapid decay with an effective support defined by $width \simeq 2^{j/2}$, $length \simeq 2^j$, and an orientation defined by angle $\theta = 2\pi l 2^{j/2}$.

One challenge posed by the curvelet domain is the overcomplete representation. Applying the curvelet transform to its adjoint thus produces

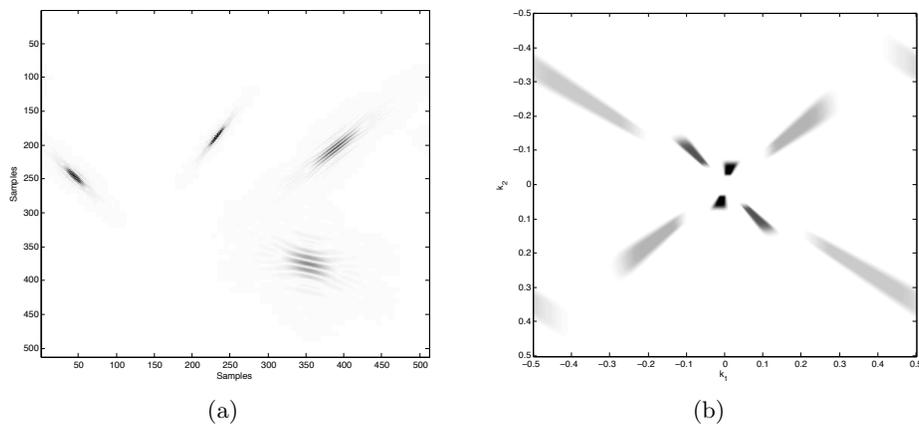


Figure 1.1: Representations of curvelets in the spatial and Fourier domains. (a) Four separate curvelets at different scales and angles. (b) The associated Fourier domain for the displayed curvelets. The four curvelets represented display four different scales and angles with the finest scales associated with the highest frequencies in the Fourier domain.

a projection. This is something that needs to be addressed when using this transform in certain algorithms. The signal separation schemes that will be used have been developed to incorporate this challenge providing that the curvelet vector is sufficiently sparse. Another challenge posed by the overcomplete representation is the curvelet vector length. While two-dimensional data generally is small enough that vector storage is not a problem, three dimensional data has a redundancy of nearly twenty four times. Problems could be encountered for large three dimensional data cubes and care needs to be taken in the application of algorithms to the coefficient vectors to avoid memory management issues.

The application of the curvelet transform to the problems associated with ground roll removal produces an effective domain in which to accomplish our goals. The properties of the curvelet transform are favorable to seismic signal separation problems and allow for the development of effective algorithms. The transform is also relatively fast as it is built on Fourier transforms, thus allowing for iterative methods to be applied to find solutions rapidly to the ℓ_1 minimization algorithms.

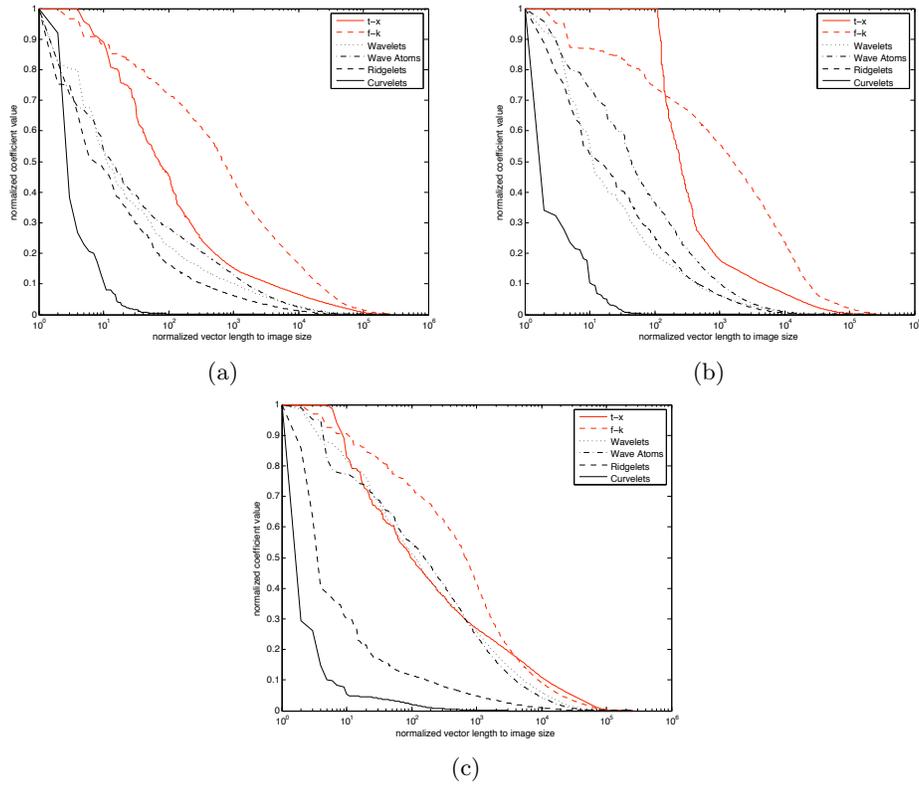


Figure 1.2: Decay rates for synthetic seismic data. (a) shows the decay rate of the entire record with (b) and (c) showing the decay rates for the exact reflector and surface wave components respectively.

1.1.2 Signal separation background

Considering our signal to be a combination of reflector signal (\mathbf{s}_1) and ground roll (\mathbf{s}_2), the aim of denoising is to recover \mathbf{s}_1 from $\mathbf{b} = \mathbf{s}_1 + \mathbf{s}_2 + \mathbf{n} \in \mathbb{R}^M$ with \mathbf{n} representing white noise. Recent developments in signal processing (Candès et al., 2005; Elad et al., 2005; Donoho et al., 2006) pertaining to the removal of white noise from overcomplete signal representations have utilized sparse representations of some data \mathbf{x} to formulate $\mathbf{b} = \mathbf{A}\mathbf{x} + \mathbf{n}$, with \mathbf{A} as a sparsity synthesis or composition matrix. The vector \mathbf{x} can then be recovered by solving the nonlinear optimization problem

$$P1 : \begin{cases} \min_{\mathbf{x}} \|\mathbf{x}\|_1 & \text{subject to } \|\mathbf{b} - \mathbf{A}\mathbf{x}\|_2 \leq \epsilon \\ \hat{\mathbf{s}} = \mathbf{A}\hat{\mathbf{x}}, \end{cases} \quad (1.2)$$

with $\hat{\cdot}$ associated with estimates found through optimization. If \mathbf{A} adheres to certain conditions, it has been shown that equation 1.2 is the constrained variation of the basis-pursuit denoising algorithm (Chen et al., 2001) and that the solution lies within the noise level (Candès et al., 2005; Elad et al., 2005).

The tolerance (ϵ) is dependant on the standard deviation of the noise vector \mathbf{n} . The noise vector $\mathbf{n}_{1\dots M} \in N(0, \sigma^2)$ has the probability of exceeding its mean by plus two standard deviations of less than 5%. Since $\|\mathbf{n}\|_2^2$ has a chi squared distribution with a mean $M \cdot \sigma^2$ and a variance of $\sqrt{2M} \cdot \sigma^2$, choosing $\epsilon^2 = \sigma^2(M + \nu\sqrt{2M})$ with $\nu = 2$ will give enough confidence that ϵ is adequate to represent the noise level.

1.1.3 Extension to redundant transforms

Under the condition that an orthonormal sparsity matrix can be used, it has been shown that a solution to P_1 can be found. Work by Donoho (1995), Mallat (1997) and Chen et al. (2001) define a soft thresholding procedure that permits an explicit solution of $\hat{\mathbf{s}} = \mathbf{S}_o^T T_\lambda(\mathbf{S}_o \mathbf{b})$ with \mathbf{S}_o as an orthonormal transform and T_λ representing a soft thresholding operator equivalent to $\text{sgn}(\mathbf{x}) \cdot \max(0, |\mathbf{x}| - |\lambda|)$. The thresholding operation solves

$$\hat{\mathbf{s}} = \arg \min_{\mathbf{s}} \|\mathbf{b} - \mathbf{s}\|_2^2 + \lambda \|\mathbf{S}_o \mathbf{s}\|_1 \quad (1.3)$$

that, with $\mathbf{u} := \mathbf{S}_o \mathbf{y}$, $\mathbf{v} := \mathbf{S}_o \mathbf{s}$ and $\mathbf{S}_o^T = \mathbf{S}_o^{-1}$ is equivalent to

$$\begin{cases} \min_{\mathbf{v}} \|\mathbf{u} - \mathbf{v}\|_2^2 + \lambda \|\mathbf{v}\|_1 \\ \hat{\mathbf{s}} = \mathbf{S}_o^T \hat{\mathbf{v}}. \end{cases} \quad (1.4)$$

The use of orthonormal transforms is not always the optimal choice. The work of Coifman and Donoho (1995) and Starck et al. (2004) has shown that the use of non-decimated wavelets gives superior results when compared to orthonormal wavelets. The choice of a sparse basis relative to a normalized length is also advantageous as long as we are dealing with an energy preserving tight frame. As seen in Figure 1.2, orthonormal transforms will not always provide the most sparse representation of seismic data and thus extension of this theory for use in overcomplete systems is needed.

In the context of overcomplete systems for non-decimated wavelets, we have $\mathbf{S} \in \mathfrak{R}^{N \times M}$ denoting a redundant, tight frame, transform. Due to the energy preservation in a tight frame, $\mathbf{S}^T \mathbf{S} = \mathbf{I}$ and, conversely, $\mathbf{S} \mathbf{S}^T$ is a projection due to the redundancy. As with all overcomplete systems, $\mathbf{S}^T \neq \mathbf{S}^{-1} = \mathbf{S}^\dagger$ with $\mathbf{S}^\dagger := (\mathbf{S}^T \mathbf{S})^{-1} \mathbf{S}^T$. In this context, the equivalence between equations 1.3 and 1.4 is no longer valid as recovery of $\mathbf{x} \in \mathfrak{R}^N$ becomes an under determined problem with $N = M \cdot \log M \gg M$ for $M \gg 1$.

Extending the problem to an overcomplete representation and following Elad et al. (2005), P_1 can be replaced by a series of simpler optimization problems in an effort to recover the sparse vector \mathbf{x}_0 .

$$P_\lambda : \begin{cases} \min_{\mathbf{x}} \|\mathbf{b} - \mathbf{S}\mathbf{x}\|_2^2 + \lambda \|\mathbf{x}\|_1 \\ \hat{\mathbf{s}} = \mathbf{S}\hat{\mathbf{x}} \end{cases} \quad (1.5)$$

where λ is an unknown Lagrangian multiplier, can be solved by use of a cooling method that slowly decreases λ from a large starting value until the result is within an error tolerance (ϵ). This occurs when λ is at its largest value for which $\|\mathbf{b} - \mathbf{S}\hat{\mathbf{x}}\|_2 \leq \epsilon$. To impose the cooling method, an iterative thresholding technique is preformed that derives from the Landweber descent method described in Daubechies et al. (2005); Elad et al. (2005); Figueiredo and Nowak (2003). Two iterative loops are imposed in solving the problem, first the outer loop where the Lagrangian multiplier is lowered, and second, the inner loop which estimates the coefficient vector for a fixed λ . The coefficient vector for the m^{th} outer loop is calculated as

$$\mathbf{x}^{m+1} = T_\lambda(\mathbf{x}^m + \mathbf{S}^\dagger(\mathbf{y} - \mathbf{S}\mathbf{x}^m)), \quad (1.6)$$

with $\lambda = \lambda_m$. For a fixed λ , it has been shown by Daubechies et al. (2005) that this iteration converges to the solution of the sub problem in equation 1.5 as long as m is large enough and $\|\mathbf{S}\| < 1$. During calculation of the solution, the under determined tight frame matrix \mathbf{S} is inverted by imposing a sparsity promoting l_1 norm. This regularizes the inverse problem (Daubechies et al., 2005) as well as meeting the recovery conditions as

described by Donoho et al. (2006); Tropp (2006) for equations 1.2 and 1.5. The algorithm is presented in detail by Elad et al. (2005) in reference to generic signal separation with an augmented system of sparsifying biases. The work of Herrmann et al. (2007) has extended this for use in seismic signal separation and is discussed in the next chapter.

Bibliography

- Candès, E., J. Romberg, and T. Tao, 2005, Stable signal recovery from incomplete and inaccurate measurements: *Comm. Pure Appl. Math.*, **59**, 1207–1223.
- Candès, E. J., 1999, Harmonic analysis of neural networks: *Appl. Comput. Harmon. Anal.*, **6**, 197–218.
- Candès, E. J., L. Demanet, D. L. Donoho, and L. Ying, 2005, Fast discrete curvelet transforms: *SIAM Multiscale Model. Simul.*, **5**, 861–899.
- Chen, S., D. Donoho, and M. Saunders, 2001, Atomic decomposition by basis pursuit: *SIAM J. Sci. Comp.*, **43**, 129–159.
- Coifman, R. and D. Donoho, 1995, Wavelets and statistics, chapter translation-invariant de-noising, 125–150. Springer-Verlag.
- Coruh, C. and J. K. Costain, 1983, Noise attenuation by vibroseis whitening (vsw) processing: *Geophysics*, **48**, 543–554.
- Daubechies, I., 1992, Ten lectures on wavelets: SIAM.
- Daubechies, I., M. Defrise, and C. de Mol, 2005, An iterative thresholding algorithm for linear inverse problems with a sparsity constraints: *CPAM*, 1413–1457.
- Deighan, A. and D. Watts, 1997, Ground-roll suppression using the wavelet transform: *Geophysics*, **62**, 1896–1903.
- Demanet, L. and L. Ying, 2007, Wave atoms and sparsity of oscillatory patterns: *Appl. Comput. harmon. Anal.*, **23-3**, 368–387.
- Donoho, D., 1995, De-noising by soft thresholding: **41**, 613–627.
- Donoho, D., M. Elad, and V. Temlyakov, 2006, Stable recovery of sparse overcomplete representations in the presence of noise: **52**, 6–18.
- Elad, M., J. L. Starck, P. Querre, and D. L. Donoho, 2005, Simultaneous cartoon and texture image inpainting using morphological component analysis (MCA): *Appl. Comput. Harmon. Anal.*, **19**, 340–358.

- Embree, P., J. P. Burg, and M. M. Backus, 1963, Wide band velocity filtering-the pie-slice process: *Geophysics*, **28**, 948–974.
- Fail, J. P. and G. Grau, 1963, Les filters on eventail: *Geophys. Prospect.*, **11**, 131–163.
- Figueiredo, M. and R. Nowak, 2003, An EM algorithm for wavelet-based image restoration: *IEEE Trans. Image Processing*, **12**, 906–916.
- Galbraith, J. N. and R. A. Wiggins, 1968, Characteristics of optimum multichannel stacking filters: *Geophysics*, **33**, 36–48.
- Gelişli, K. and H. Karşlı, 1988, F-k filtering using the hartley transform: *J. Seism. Explor.*, **7**, 101–108.
- Harlan, W. S., J. F. Claerbout, and F. Rocca, 1984, Signal/noise separation and velocity estimation: *Geophysics*, **49**, 1869–1880.
- Herrmann, F. J., U. Boeniger, and D.-J. E. Verschuur, 2007, Nonlinear primary-multiple separation with directional curvelet frames: *Geoph. J. Int.* (To appear.).
- Holzman, M., 1963, Chebysev optimized geophone arrays: *Geophysics*, **28**, 145–155.
- Karşlı, H. and Y. Bayrak, 2004, Using the Wiener-Levinson algorithm to suppress ground-roll: *Journal of Applied Geophysics*, **55**, 187–197.
- Liu, X., 1999, Ground roll spression using the Karhuren-Loeve transform: *Geophysics*, **64**, 564–566.
- Londono, E. G., L. C. Lòpez, and T. S. Kazmierczak, 2005, Using the karhunen-loève transform to suppress ground roll in seismic data: *Earth Sci. Res. J.*, **9**, 139–147.
- Mallat, S. G., 1997, *A wavelet tour of signal processing*. Academic Press.
- McKay, A. E., 1954, Review of pattern shooting: *Geophysics*, **19**, 420–437.
- McMechan, G. A. and R. Sun, 1991, Depth filtering of first breaks and ground roll: *Geophysics*, **56**, 390–396.
- Olhovich, V. A., 1964, The causes of noise in seismic reflection and refraction work: *Geophysics*, **29**, 1015–1030.
- Russel, B., D. Hampson, and J. Chun, 1990, Noise elimination and the radon transform; part 1: *The Leading Edge*, **9**, no. 10, 18–23.
- Saatçilar, R. and N. Canitez, 1988, A method of ground-roll elimination: *Geophysics*, **53**, 894–902.

Bibliography

- Seeger, A., C. Sogge, and E. Stein, 1991, Regularity properties of fourier integral operators: *Annals of Math*, **134**, 231–251.
- Starck, J. L., M. Elad, and D. Donoho, 2004, Redundant multiscale transforms and their application to morphological component separation: *Advances in Imaging and Electron Physics*, **132**.
- Trad, D., T. Ulrych, and M. Sacchi, 2003, Latest views of the sparse radon transform: *Geophysics*, **68**, 386–399.
- Treitel, S., J. L. Shanks, and C. W. Fraster, 1967, Some aspects of fan filtering: *Geophysics*, **32**, 789–800.
- Tropp, T., 2006, Just relax: convex programming methods for subset selection and sparse approximation: *IEEE Trans. Inform. Theory*. (To appear.).
- Yilmaz, O., 2001, *Seismic data analysis. Curves and Surfaces*. Society of Exploration Geophysicists.
- Zhang, R. and T. J. Ulrych, 2003, Physical wavelet frame denoising: *Geophysics*, **68**, 225–231.

Chapter 2

Ground roll prediction and separation in the sparse curvelet domain

2.1 Introduction

Ground roll removal is an essential processing step with all land based seismic studies. This large amplitude coherent noise component can swamp important reflector signals and needs to be removed in a manner that does not alter the underlying reflection information. Standard methods of surface wave removal make a trade off removing an adequate amount of ground roll with as little reflector information as possible but in most cases, reflector information is altered or even partially removed. Many methods attempt to utilize the low-velocity and low-frequency nature of ground roll through survey design (McKay, 1954; Holzman, 1963), band-pass filtering (Yilmaz, 2001), multichannel filtering (Galbraith and Wiggins, 1968), spectral balancing (Coruh and Costain, 1983), or other frequency-based methods (Embree et al., 1963; Fail and Grau, 1963; Treitel et al., 1967; Harlan et al., 1984; Gelişli and Karşli, 1988). It has also been shown that these methods can cause distortion due to alteration of the frequency components of the reflector signals (McMechan and Sun, 1991; Liu, 1999; Karşli and Bayrak, 2004). Beyond the Fourier domain, work has been done to identify and remove ground roll with the Karhunen-Loeve transform (Londono et al., 2005), Wiener-Levinson algorithm (Karşli and Bayrak, 2004), Radon transform (Russel et al., 1990; Trad et al., 2003) and with one and two dimensional wavelets (Deighan and Watts, 1997; Zhang and Ulrych, 2003) but challenges still remain, particularly when dealing with aliased ground roll with poor frequency separation.

We propose the use of the curvelet domain for ground-roll removal and

A version of this chapter will be submitted for publication. Ground-roll prediction and separation in the sparse curvelet domain written with Felix Herrmann

show that for the data of interest, it provides the most sparse representation. With the sparsifying curvelet domain, we examine two separation methods that solve sparsity promoting ℓ_1 minimization problems through iterative soft thresholding. Our first-generation signal separation algorithm (Herrmann et al., 2007b), the block-coordinate relaxation method, is described and tested against our second-generation method, the Bayesian separation algorithm (Saab et al., 2007; Wang et al., 2007). To examine the results of the two introduced separation methods, a synthetic dataset is constructed such that the signal-to-noise ratio with respect to the true reflector information can be calculated. We use signal-to-noise ratios relative to the exact reflectors to judge the two separation schemes and to judge the sensitivity of the parameters associated with each scheme. These methods are then put to use on a real dataset containing complex ground roll.

2.1.1 Sparse representations of seismic data

Seismic data are recorded in the source-receiver-time domain due to the limitations of the recording devices but this is not necessarily the most efficient domain in which to perform many seismic data processing operations. In certain situations it becomes advantageous to convert data to a different representation, preferably one designed to sparsely represent seismic wave fields such that the amplitude sorted coefficients have a rapid decay rate. This advantage was shown initially by Sacchi et al. (1998) and followed by Trad et al. (2003), Xu et al. (2005), Abma and Kabir (2006) and Zwartjes and Sacchi (2007) for Fourier and Radon transforms and more recently by Herrmann et al. (2007c,a) for curvelets. For example, if we are interested in the global frequency content of the seismic signal, the Fourier domain can be used allowing access to individual stationary frequency components of the signal. This is the basis of countless processing techniques and methods and is fundamental to seismic signal processing. Generally this approach works well when the frequency components of the recorded signal are stationary, well sampled with separated coefficients for the different wavefields but this cannot always be assumed. For instance, spatial aliasing, particularly with respect to ground roll, provides a challenge for f-k processing (McMechan and Sun, 1991; Liu, 1999; Karsli and Bayrak, 2004) and it becomes necessary to develop alternative methods based on sparsity promotion in another domain to separate these signals.

The global nature of Fourier transforms tends to limit their efficiency when representing time varying signals. For one-dimensional signals, multiscale wavelet transforms remedy this limitation by representing a seismic

trace in terms of discrete wavelets, $\varphi_{j,k}$, designed to expand the signal with respect to wavelets at different scales (j) and positions (k). Since seismic traces can be considered bandwidth limited, correlations of the reflection events with $\varphi_{j,k}$ will only be significant for wavelets that occupy roughly the same location and have a similar spectral content. As such, the wavelet vector will contain a limited number of significant coefficients. While wavelets are an effective way to represent one-dimensional signals, they do not generalize to higher dimensions. As we will show, the isotropic nature of wavelets limits their ability to represent continuous curved wavefronts as parsimoniously as other directional transforms can. Thus a different approach is needed for seismic images and data volumes.

To remedy the shortcoming of wavelets to sparsely represent data with arbitrarily curved wavefronts, different tilings of phase space (space spanned by physical space and Fourier space) have been proposed. These tilings generate families of directional two dimensional space domain atoms that differ in shape. This shape is controlled by the dilation matrix (Demanet and Ying, 2007)

$$D_j = \begin{bmatrix} 2^{j\alpha} & 0 \\ 0 & 2^{j\beta} \end{bmatrix} \quad (2.1)$$

where the index j refers to the scale and where α controls the multiscale nature ($\alpha = 0$ for single scale, $\alpha = 1$ for multiscale) and β the directional selectivity ($\beta = 0$ directionally selective, $\beta = 1$ directionally indiscriminate) of the atoms (See Figure 2.1 adapted from Demanet and Ying (2007)).

If we follow the α -axis in Figure 2.1 from the origin, the mono-scale and isotropic Gabor (Gabor, 1946), or windowed Fourier transform, changes into the directional, single aspect ratio ridgelet transform (Candès and Donoho, 1999). This choice of parameters ($\alpha = 1, \beta = 0$) makes ridgelets well adapted to represent unit length, single direction plane waves but on the same token, limits their ability to represent point scatterers or curved wavefronts. Leaving the origin toward wavelets ($\alpha = 1, \beta = 1$) in Figure 2.1 yields a transform that is multiscale but which lacks directional selectivity, a desired feature for our purpose of separating coherent wavefields. Adding this last requirement into the mix, leaves us with two alternatives, namely wave atoms ($\alpha = 1/2, \beta = 1/2$) and curvelets ($\alpha = 1, \beta = 1/2$). Both transforms involve atoms that exhibit a parabolic scaling relation essential for capturing wave phenomena (see Demanet and Ying (2007) for details). Wave atoms are tailored toward representing periodic "texture" while curvelets are designed to capture curved wave fronts. Wave atoms are continuous along the wave front with oscillations in the direction of the wavefront and obey the

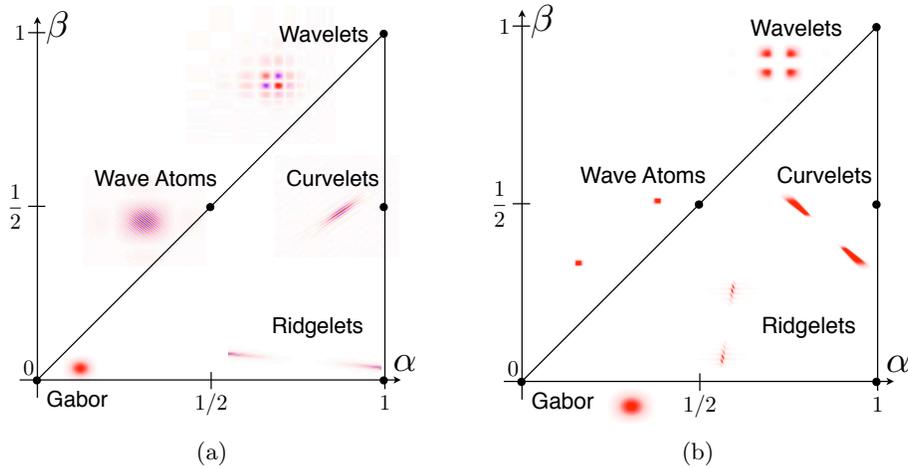


Figure 2.1: Descriptive plots showing the relation of the specific transforms to the α and β parameters. (a) is augmented with physical representations of the individual atoms and (b) contains representations of the Fourier spectrum. This figure is adapted from Demanet and Ying (2007).

parabolic scaling relation $wavelength \simeq diameter^2$. Curvelets (Candès and Donoho, 2000) have a physical shape closer to a ridgelet as seen in Figure 2.1(a) and obey the parabolic scaling principle $length^2 \simeq width$ with the frequency of a fixed number of oscillations across the curvelet increasing with higher scales. This property causes curvelets to become increasingly anisotropic and more needle like at finer scales.

Because of the complex nature of seismic data, it is not clear which of the two alternatives should be preferred. For that reason we, include in Figure 2.2, an empirical study of the decay rates for the aforementioned transforms in relation to our signals of interest. As shown by Candès et al. (2005) and Hennenfent and Herrmann (2007) with confirmation in Figure 2.2, curvelets present the most parsimonious domain for the chosen seismic data. While the curvelet transform is slightly less than eight times redundant, with respect to the total vector length it stores more energy on less coefficients than any of the other transforms we considered. This length normalized sparsity is what proves important when solving optimization problems by ℓ_1 minimization.

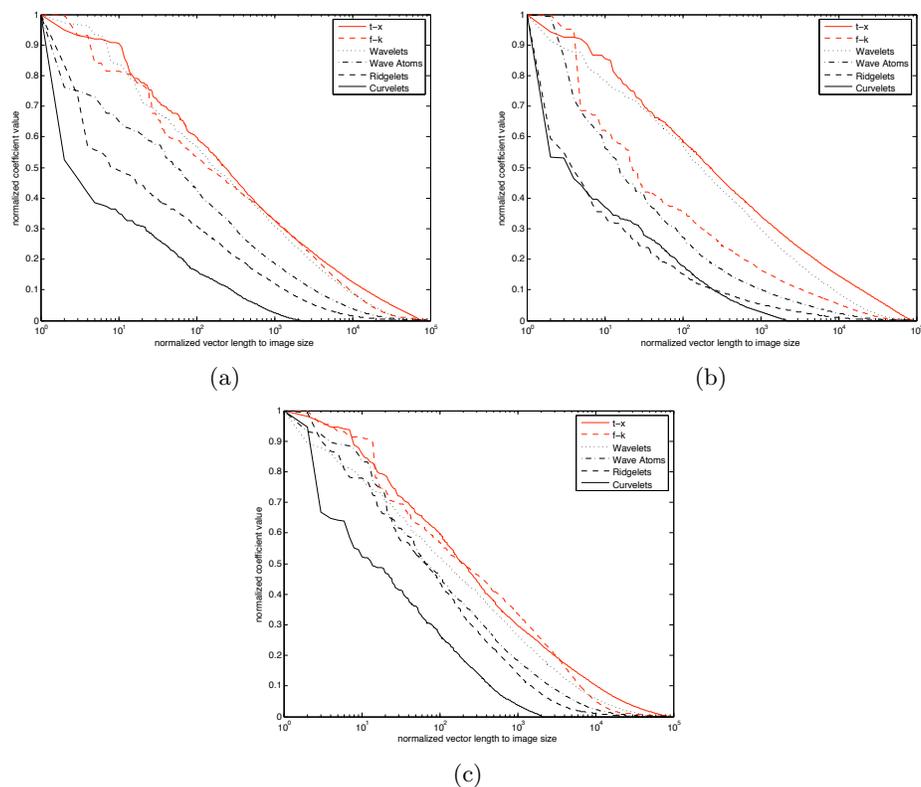


Figure 2.2: Decay rates for real seismic waveforms. (a) shows the decay rate of the entire record with (b) and (c) showing the decay rates for the estimated reflector and surface wave components respectively. The curvelet domain shows the most rapid decay rates of all the transforms examined while the t-x and f-k domains have the slowest decay making them the least efficient for our methods.

2.2 Signal separation

To be clear, we will refer to predictions as required to perform some signal separation and estimations as the products of the signal separation. The most simple separation scheme would be to calculate a surface wave estimate from a direct subtraction of some prediction from a total dataset. Even though this may work in cases where the prediction of the signal is nearly perfect, it has the potential to fail completely if the prediction contains translations, phase shifts or amplitude errors. It is therefore advantageous to use a scheme which improves on this basic method by utilizing the sparse representation of the signals in the curvelet domain to add robustness.

Developments in mathematical signal processing aimed at finding stable solutions to problems utilizing overcomplete representations (Starck et al., 2004; Elad et al., 2005) have led to signal separation schemes specifically designed for sparse signals. These new methods have complemented the list of subtraction schemes by utilizing ℓ_1 minimization on sparse vectors. This allows us to exploit the sparse representation of the curvelet signal which is based on the fundamental knowledge that we are separating wavefronts.

We apply these wavefield separation methods to our two signals consisting of reflector and surface waves even though these methods are not limited to only two signals (see Starck et al. (2004) for details). The formulation is based on the following representation for the total data \mathbf{b} defined as

$$\mathbf{b} = \mathbf{s}_1 + \mathbf{s}_2 + \mathbf{n}. \quad (2.2)$$

We define \mathbf{s}_1 and \mathbf{s}_2 as reflectors and ground roll respectively with \mathbf{n} as a zero-mean Gaussian noise vector with a standard deviation σ .

In our sparsity-promoting formulation of the wavefield separation problem, we aim to find a curvelet coefficient vector for each wavefield components \mathbf{x}_i , $i = 1 \dots 2$. These estimations will be generated by the separation scheme but predictions will be required to define the weighted thresholds that will be used to help drive the signals apart. The prediction for the reflector coefficients is defined as $\mathbf{b}_1 = \mathbf{b} - \mathbf{b}_2$ such that \mathbf{b}_1 and \mathbf{b}_2 are equal to the predicted reflector and ground roll components. We obtain the following system of equations

$$\mathbf{b}_i = \mathbf{A}\mathbf{x}_i + \mathbf{n}_i \quad (2.3)$$

with \mathbf{A} defined as the curvelet synthesis matrix with \mathbf{n}_i representing two independent Gaussian noise processes. With this forward model, we can now derive two alternative formulations for the wavefield separation problem. The first, based on a block-relaxation method, assumes that there is no

error in the prediction and sets $\mathbf{b}_1 = \mathbf{b} - \mathbf{b}_2 = \mathbf{A}\mathbf{x}_1$ and $\mathbf{b}_2 = \mathbf{A}\mathbf{x}_2$, yielding

$$\mathbf{b} = [\mathbf{A}_1 \quad \mathbf{A}_2][\mathbf{x}_1 \quad \mathbf{x}_2]^T + \mathbf{n}. \quad (2.4)$$

The second, more recent approach, derives from a Bayesian formulation of the separation problem where $\mathbf{b}_i = \mathbf{A}_i\mathbf{x}_i + \mathbf{n}_i$ with $i = 1 \dots 2$, and $\mathbf{n}_1 = \mathbf{n} - \mathbf{n}_2$. We will show that both sets of assumptions lead to distinctly different formations for our wavefield separation methods.

2.2.1 Block-coordinate relaxation method

The block-coordinate relaxation method builds on the concepts of morphological component analysis introduced by Starck et al. (2004) and Elad et al. (2005). Recognizing \mathbf{b} and \mathbf{x} are summations of the separate wavefields, an augmented system is constructed with $\mathbf{A} = [\mathbf{A}_1 \quad \mathbf{A}_2]$ where $\mathbf{A}_{1,2}$ are curvelet synthesis matrices and $\mathbf{x} = [\mathbf{x}_1 \quad \mathbf{x}_2]^T$ are the unknown curvelet coefficient vectors. This method is originally designed to separate two different signals, each sparse in some domain but incoherent in the sparsifying domain of the other signal. An example of this would be spikes and sinusoids with the Dirac and Fourier domain. By utilizing a sparse representation of one signal and an incoherent representation of the other, iterative soft thresholding is applied in the sparse domains to recover the signals in a fashion similar to solving a denoising problem (Elad et al., 2005). This is accomplished for the first signal by finding the minimization of an ℓ_1 denoising problem in the first sparse domain using soft thresholding of the sparse coefficients to remove the incoherent coefficients. The method is repeated for the other signals in their respective sparse domains and the entire algorithm is run iteratively to produce the signal separation. As shown in Starck et al. (2004), this method successfully separates cartoon-like features from texture components. Even though this method has been shown to be useful for separation of two incoherent basis (ie, Fourier and Dirac), use in the context of coherent seismic wavefield separation requires some modification.

With both reflectors and ground roll consisting of bandwidth limited wavefronts, it is a challenge to find two transforms that exhibit the properties discussed previously. Instead of using two sparse domains, Herrmann et al. (2007b) developed a method to adapt the solution to this problem to use a single sparse domain such that $\mathbf{A}_1 = \mathbf{A}_2$. To accomplish this, the ℓ_1 -

optimization problem is recast such that

$$\begin{cases} \min_{\mathbf{x}} \|\mathbf{x}\|_{\mathbf{w},1} & \text{subject to } \|\mathbf{b} - \mathbf{A}\mathbf{x}\|_2 \leq \epsilon \\ \widehat{\mathbf{b}}_i = \mathbf{A}\widehat{\mathbf{x}}_i, & i = 1, \dots, 2 \\ \text{given: } \mathbf{b}_1 & \text{and } \mathbf{w}(\mathbf{b}, \mathbf{b}_1) \end{cases} \quad (2.5)$$

with $\mathbf{A}\mathbf{x} = \mathbf{A}_1\mathbf{x}_1 + \mathbf{A}_2\mathbf{x}_2$ and a given reflector prediction \mathbf{b}_1 and $\widehat{\cdot}$ referring to estimates calculated via optimization. Built on the block-coordinate relaxation method, a weighting scheme was introduced that allows for incorporation of reflector, \mathbf{b}_1 , and surface wave, \mathbf{b}_2 , predictions such that $\mathbf{b} = \mathbf{b}_1 + \mathbf{b}_2$. With reasonable predictions of the reflectors and ground roll, which we provide, a separation can be preformed in the same sparse domain. The strictly positive weights are defined as

$$\begin{cases} \mathbf{w}_1 := \max(\sigma \cdot \sqrt{2 \log N}, C_1 |\mathbf{A}^T \mathbf{b}_1|) \\ \mathbf{w}_2 := \max(\sigma \cdot \sqrt{2 \log N}, C_2 |\mathbf{A}^T \mathbf{b}_2|) \end{cases} \quad (2.6)$$

and then combined such that $\mathbf{w} = [\mathbf{w}_1 \quad \mathbf{w}_2]^T$. The weights are then applied to the minimization of the 1-norm in the definition of the soft thresholding operator. For these definitions, σ refers to the standard deviation of the noise vector with the constants, C_1 and C_2 , used to normalize the ℓ_2 norms of the predictions.

The work of Elad et al. (2005) finds a solution by solving a series of simpler optimization problems and is implemented in Herrmann et al. (2007b) for the weighted problem. The solution to this problem is an algorithm that iteratively computes a series of outer and inner loops. The following optimization problem,

$$\widehat{\mathbf{x}}_j = \arg \min_{\mathbf{x}_j} \frac{1}{2} \|\mathbf{b} - \mathbf{A}\mathbf{x}_j - \sum_{i \neq j} \mathbf{A}\mathbf{x}_i\|_2^2 + \|\mathbf{x}_j\|_{1, \gamma \cdot \mathbf{w}_j} \quad j = 1, \dots, 2, \quad (2.7)$$

solves for the \mathbf{x}_j component in the outer-loop. Deriving the m^{th} iteration, the inner loop computes

$$\mathbf{x}_j^m = T_{\gamma \cdot \mathbf{w}_j}(\mathbf{x}_j^m + \mathbf{A}^T(\mathbf{b} - \mathbf{A}\mathbf{x}_j^m - \sum_{i \neq j} \mathbf{A}\mathbf{x}_i)), \quad (2.8)$$

with T_u as the element-wise soft thresholding operator defined as $T_u(x) = \text{sign}(x) \cdot \max(0, |x| - |u|)$. As part of the outer loop, the Lagrange multiplier γ^m , is used as a cooling parameter with $\gamma^m > \gamma^{m+1}$. This cooling parameter determines the threshold in the inner loop. As γ is lowered, more elements

are allowed to enter in to the solution to build up the signal estimates. The inner loop is solved L times unless a stopping criterion is met such that $R^m(\mathbf{z}_1, \mathbf{z}_2) > R^{m-1}$ with $\mathbf{z}_1 = \mathbf{A}(\mathbf{A}^T \mathbf{b} - \mathbf{x}_2^m)$, $\mathbf{z}_2 = \mathbf{A}(\mathbf{A}^T \mathbf{b} - \mathbf{x}_1^m)$ and

$$R^m(\mathbf{z}_1, \mathbf{z}_2) = \frac{\mathbf{z}_1^T \mathbf{z}_2}{\|\mathbf{z}_1\| \|\mathbf{z}_2\|}. \quad (2.9)$$

Table 2.1 summarizes the algorithm used to estimate the signal components $\hat{\mathbf{x}}_j$.

The block-coordinate relaxation method contains two parameters as constants that define the relative weights of the two signal components. Setting $C_1 > C_2$, the threshold that is applied to the reflector information will be more aggressive than the threshold for the ground roll estimate. These parameters are more sensitive than the parameters associated with the Bayesian solution. The estimates will be over thresholded if they are set too high causing reflector degradation. If they are set too low, convergence will take many iterations. If over thresholding occurs, signal mixing causing a poor separation or possibly no separation at all may occur. Proper balancing of these parameters is required to avoid this problem.

The block-coordinate relaxation scheme is most effective when applied to two signals that are relatively well separated in a certain domain, such as Fourier. Ground roll and reflectors are generally a good example of this. Near the apex where high amplitudes and curvature are encountered for all wavefronts, problems can be encountered. This large amplitude variation can cause the block-coordinate method to generate curvelet like artifacts in the denoised data.

2.2.2 Bayesian separation method

Recently Saab et al. (2007) and Wang et al. (2007) introduced a new separation scheme based on a Bayesian formulation, which we adapt to the problem of ground roll. This method computes separation estimates from two signal predictions making it similar to the block-coordinate relaxation method. This will introduce parameters that allow control based on expected sparsity and prediction fidelity making it more flexible than the previous method. This flexibility comes from the ability to define the ℓ_1 minimization problem to favor the more reliable prediction vs the less reliable one. The weights applied to the soft thresholding operator are defined by fixed Lagrangian parameters that are set to the expected sparsity of the estimated signals.

```

Initialize:  $\mathbf{x}_1^0 = 0, \mathbf{x}_2^0 = 0, m = 0, R^0 = 1;$ 
Choose:  $L, \lambda_1 \geq \lambda_2 \geq \dots$ 
while  $\|\mathbf{b} - \mathbf{A}\mathbf{x}^m\|_2 > \epsilon$  and  $R^m \leq R^{m-1}$  do
     $m = m + 1, l = 0;$ 
     $\mathbf{x}^m = \mathbf{x}^{m-1};$ 
     $\mathbf{w}_1 := \max(\lambda \cdot \sigma, C_{1,s}|\mathbf{A}^T \mathbf{b}_1|)$  {Set reflector weights},
     $\mathbf{w}_2 := \max(\lambda \cdot \sigma, C_{2,s}|\mathbf{A}^T \mathbf{b}_2|)$  {Set ground roll weights}
     $R^m(\mathbf{z}_1, \mathbf{z}_2) = \frac{\mathbf{z}_1^T \mathbf{z}_2}{\|\mathbf{z}_1\| \|\mathbf{z}_2\|};$  {Compute decorrelation}
    while  $l \leq L$  and  $\epsilon$  and  $R^m \leq R^{m-1}$  do
         $l = l + 1;$ 
         $\mathbf{b}_2 = \mathbf{A}\mathbf{x}_2^m;$  {Synthesize}
         $\mathbf{r}_1 = \mathbf{b} - \mathbf{b}_2;$  {Calculate residual}
         $\mathbf{x}_1^m = \mathbf{x}_1^m + \mathbf{A}^T(\mathbf{r}_1 - \mathbf{A}\mathbf{x}_1^m);$  {Descent update}
         $\mathbf{x}_1^m = \text{sign}(\mathbf{x}_1^m) \cdot \max(0, |\mathbf{x}_1^m| - |\lambda^m \cdot \mathbf{w}_1|);$  {Soft threshold}
         $\mathbf{b}_1 = \mathbf{A}\mathbf{x}_1^m;$  {Synthesize}
         $\mathbf{r}_2 = \mathbf{b} - \mathbf{b}_1;$  {Calculate residual}
         $\mathbf{x}_2^m = \mathbf{x}_2^m + \mathbf{A}^T(\mathbf{r}_2 - \mathbf{A}\mathbf{x}_2^m);$  {Descent update}
         $\mathbf{x}_2^m = \text{sign}(\mathbf{x}_2^m) \cdot \max(0, |\mathbf{x}_2^m| - |\lambda^m \cdot \mathbf{w}_2|);$  {Soft threshold}
    end while
end while

```

Table 2.1: The block-coordinate relaxation algorithm. Table from Herrmann et al. (2007b)

The noise term \mathbf{n}_i in the context of Equation 2.3 is defined as the noise associated with the reflector estimate with each component as $N(0, \sigma_i^2)$. This noise term is assumed independent of the recorded signal noise \mathbf{n} with the surface wave estimate noise $\mathbf{n}_2 = \mathbf{n} - \mathbf{n}_1$. The assumption of white noise is made in the absence of an accurate probabilistic model but is consistent with previous work regarding primary-multiple separation (Verschuur et al., 1992) and can be extended to colored Gaussian noise as long as the corresponding covariance matrices are known. This will lead to an ℓ_2 -norm penalty for the error of the optimization problem such that we assign a cost function that penalizes the misfit between the predicted and estimated signal components.

Following Saab et al. (2007) and Wang et al. (2007) who cast the signal separation problem in terms of a maximization of a conditional probability of

$$\begin{aligned}
 P(\mathbf{x}_1, \mathbf{x}_2 | \mathbf{b}_1, \mathbf{b}_2) &= \frac{P(\mathbf{x}_1, \mathbf{x}_2)P(\mathbf{b}_1 | \mathbf{x}_1, \mathbf{x}_2)P(\mathbf{b}_2 | \mathbf{b}_1, \mathbf{x}_1, \mathbf{x}_2)}{P(\mathbf{b}_1, \mathbf{b}_2)} \\
 &= \max_{x_1, x_2} - \left(\lambda_1 \|\mathbf{x}_1\|_{1, \mathbf{w}_1} + \lambda_2 \|\mathbf{x}_2\|_{1, \mathbf{w}_2} + \frac{\|\mathbf{A}\mathbf{x}_2 - \mathbf{b}_2\|_2^2}{\sigma_2^2} \right. \\
 &\quad \left. + \frac{\|\mathbf{A}(\mathbf{x}_1 + \mathbf{x}_2) - (\mathbf{b}_1 + \mathbf{b}_2)\|_2^2}{\sigma^2} \right), \tag{2.10}
 \end{aligned}$$

for the two signals. Given the generated predictions of the reflector (\mathbf{b}_1) and ground roll (\mathbf{b}_2), this method attempts to maximize the probability that the actual reflector and ground roll coefficient vectors will be generated. This expression is reworked into a minimization of predicted ground roll and the total data in the following functional

$$\begin{aligned}
 f(\mathbf{x}_1, \mathbf{x}_2) &= \lambda_1 \|\mathbf{x}_1\|_{1, \mathbf{w}_1} + \lambda_2 \|\mathbf{x}_2\|_{1, \mathbf{w}_2} + \frac{\|\mathbf{A}\mathbf{x}_2 - \mathbf{b}_2\|_2^2}{\sigma_2^2} \\
 &\quad + \frac{\|\mathbf{A}(\mathbf{x}_1 - \mathbf{x}_2) - (\mathbf{b}_1 + \mathbf{b}_2)\|_2^2}{\sigma^2}, \tag{2.11}
 \end{aligned}$$

with \mathbf{A} represents the curvelet synthesis matrix, \mathbf{x}_1 , the reflector coefficient vector and \mathbf{x}_2 as the ground roll coefficient vector. The two signals will be driven apart by the weights applied to the ℓ_1 -norms.

The formulation of the conditional probability statement defined by Equation 2.10 to the minimization problem (equation 2.11) introduces the parameters λ_1 , λ_2 and η . These control parameters allow input of a priori information pertaining to the expected sparsity of the reflectors and ground roll estimates through λ_1 and λ_2 respectively as well as η to represent the fidelity of the two predictions. For example, if a more parsimonious reflector

signal relative to our ground roll signal is expected, then setting $\lambda_1 > \lambda_2$ will reflect this. This is the case of very disordered ground roll or regular and simple reflector signatures. If the opposite is true, these parameters are adjusted to reflect more sparse ground roll components. The parameter η represents the fidelity of the prediction. Reducing η will reduce the effect off the reflector prediction on the calculation of the ground roll estimation while increasing the aggressiveness of the thresholding operators. Increasing η will have the opposite outcome of reducing the thresholding aggressiveness while increasing the effect of the reflector prediction on the surface wave estimation. The effect of these parameters can be examined empirically with synthetic data as shown in Table 2.4.

The minimization of Equation 2.11 is accomplished by applying a soft thresholding operator T_u iteratively. This method of solving an ℓ_1 minimization is consistent with the work of Elad et al. (2005) and is similar to the previous block-coordinate relaxation method. From the initial starting conditions, the n^{th} iteration computes for the estimates

$$\mathbf{x}_1^{n+1} = T_{\frac{\lambda_1 \mathbf{w}_1}{2\eta}} [\mathbf{A}^T \mathbf{b}_2 - \mathbf{A}^T \mathbf{A} \mathbf{x}_2^n + \mathbf{A}^T \mathbf{b}_1 - \mathbf{A}^T \mathbf{A} \mathbf{x}_1^n + \mathbf{x}_1^n] \quad (2.12)$$

$$\mathbf{x}_2^{n+1} = T_{\frac{\lambda_2 \mathbf{w}_2}{2(1+\eta)}} [\mathbf{A}^T \mathbf{b}_2 - \mathbf{A}^T \mathbf{A} \mathbf{x}_2^n + \mathbf{x}_2^n + \frac{\eta}{1+\eta} (\mathbf{A}^T \mathbf{b}_1 - \mathbf{A}^T \mathbf{A} \mathbf{x}_1^n)]. \quad (2.13)$$

With the use of strictly positive weights, it is shown that this algorithm converges to the minimizer of $f(x_1, x_2)$ (Daubechies et al., 2005; Wang et al., 2008). The separation algorithm can be seen in Table 2.2 with further details described in Wang et al. (2008).

The solution is faster and more stable under more seismic signal separation problems than the previous block-coordinate relaxation methods used for ground-roll removal in Yarham et al. (2006). Results for both synthetic and real data generally require roughly ten iterations to reach convergence with the Bayesian separation method applied to a two dimensional slice of data. This generally equates to less than one minute of processing on the average desktop computer.

To maximize the preservation the of reflector information for these two separation schemes, it is advantageous to produce a conservative estimate of the surface wave excluding as much reflector information as possible. By providing a conservative estimate, the reflector information is located in one prediction instead of spread across both predictions. As no reflector information will be located in the surface wave prediction, no components of the reflectors will have to be identified through thresholding of the surface wave coefficient vector. This will minimize the chance that coefficients associated with the reflectors will end up in the surface wave estimation.

```

Initialize:  $\mathbf{x}_1^0 = 0, \mathbf{x}_2^0 = 0, m = 0,$ 
Choose:  $m_{\max}, \eta, \lambda_1, \lambda_2$ 
Define Threshold:  $\mathbf{w}_1 = (\lambda_1 |\mathbf{A}^T \mathbf{b}_2|) / (2\eta), \mathbf{w}_2 =$ 
 $(\lambda_2 |\mathbf{A}^T \mathbf{b}_1|) / (2(\eta + 1))$ 
while  $m < m_{\max}$  do
     $m = m + 1$ 
     $\tilde{\mathbf{x}}_1 = \mathbf{A}^T \mathbf{b}_2 - \mathbf{A}^T \mathbf{A} \mathbf{x}_2^m + \mathbf{A}^T \mathbf{b}_1 - \mathbf{A}^T \mathbf{A} \mathbf{x}_1^m + \mathbf{x}_1^m;$  {Coefficient
    update}
     $\tilde{\mathbf{x}}_2 = \mathbf{A}^T \mathbf{b}_2 - \mathbf{A}^T \mathbf{A} \mathbf{x}_2^m + \mathbf{x}_2^m + \frac{\eta}{\eta+1} (\mathbf{A}^T \mathbf{b}_1 - \mathbf{A}^T \mathbf{A} \mathbf{x}_1^m);$  {Coefficient
    update}
     $\mathbf{x}_1^{m+1} = \text{sign}(\tilde{\mathbf{x}}_1) \cdot \max(0, |\tilde{\mathbf{x}}_1| - |\mathbf{w}_1|);$  {Soft threshold}
     $\mathbf{x}_2^{m+1} = \text{sign}(\tilde{\mathbf{x}}_2) \cdot \max(0, |\tilde{\mathbf{x}}_2| - |\mathbf{w}_2|);$  {Soft threshold}
end while

```

Table 2.2: The Bayesian-separation algorithm.

2.3 Application to synthetic data

To judge the effectiveness of the block-coordinate relaxation method and the Bayesian separation method, synthetic datasets are used. By generating the reflector and ground roll information separately, we have exact knowledge of our reflector signal that can be used to generate signal-to-noise ratio measurements in which to judge our separation methods by. The same methods used to generate the ground roll can also be used to generate predictions for the separation scheme of varying and controlled fidelity. By examining the signal-to-noise ratios of the estimated reflectors, we can get an idea of how effective the separation methods are.

2.3.1 Synthetic data generation

To generate the synthetic data, we use a full elastic finite-difference code (Graves, 1996). The synthetic data is generated as two separate signals of ground roll and reflectors and then combined. This allows for a ground truth for the reflector information used in calculating the SNR, and to allow for the generation of the surface wave to occur on a finer grid making it more accurate. The first dataset is a simple reflector signal on a ten meter grid with a horizontal layer at 500m, followed by two dipping reflectors centered at 1500 and 2650m depth above a half space. The ground roll is generated on a 1m grid with linear increase in compressional and shear wave velocity and density from the surface to 25m depth over a half space. Within the

layers, a one percent normally distributed random perturbation is added to all three parameters varying it in two dimensions. These two models are shown in Figure 2.3. Before combination, all signals are convolved with a Ricker source wavelet with a center frequency of 20 Hz. This provides a dataset with an initial SNR of -1.67 dB with respect to ground-roll free data and is shown in Figure 2.4(a).

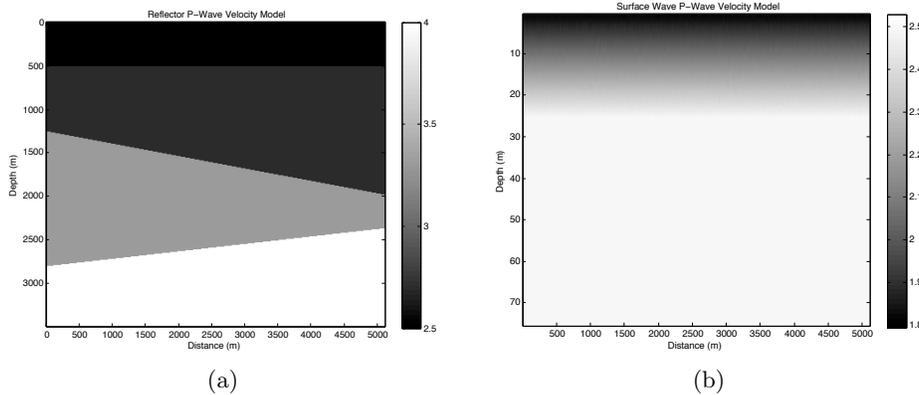


Figure 2.3: Synthetic model for pressure wave velocity used in generation of the (a) reflectors and (b) ground roll. Shear wave velocity and density models are not shown but exhibit the same overall structure. Synthetic data was generated separately and combined to produce the data shown in Figure 2.4.

2.3.2 Separation scheme comparison

Four different synthetic examples are used to test the separation schemes. The first test uses the exact ground roll as a prediction to our separation schemes. The second test simulates inaccurate modeling by using a surface wave prediction that contains modeling errors from the exact surface wave. This erroneous predictions is generated with 5% uniformly distributed random error applied to the total depth of the surface layers as well as the start and end points of the linearly increasing parameters. The third test uses a noisy dataset and noisy prediction consisting of the previous modeling error data. The added white noise exhibits a normal distribution and a maximum variation of 5% of the maximum of the signal and is applied to both the dataset and the prediction. The last test involves a Hilbert transform ($\pi/2$ degrees phase rotation) applied in the fast direction of the ground roll model to generate a erroneous prediction. These four examples are examined

for both the Bayesian signal separation and the block-coordinate relaxation scheme. Table 2.3 shows the signal-to-noise ratios of the experiments. These separations increase the signal-to-noise ratios of by over 10 dB when compared to direct subtraction except for the block-coordinate relaxation on the model error with added noise.

Noise prediction	Subtraction	Bayesian separation	BCR
1) Exact noise	147.96	20.58	13.80
2) 5% model error	-4.38	9.58	9.41
3) 5% model error with noise	-4.52	9.09	3.53
4) Hilbert transform of model	-4.67	14.93	13.33

Table 2.3: Signal-to-noise ratios for Bayesian and block-coordinate relaxation (BCR) separation methods. The SNR for the synthetic data is -1.67 for all tests except for the third test which has added white noise decreasing the SNR to -1.92. These values were calculated with respect to data that has no ground roll or noise present.

To solve these synthetic problems with the curvelet domain separation schemes, we need to look at how sparse the expected results will be and the fidelity of the predictions. For the Bayesian separation method, the expected relative sparsity of the signal estimates define the settings for λ_1 and λ_2 and the prediction fidelity defines η (Equation 2.11). As the synthetic reflector signal consists of three sparse reflectors, we expect that the \mathbf{x}_1 will be very sparse. With this expected sparsity, we set $\lambda_1 > \lambda_2$. The ground-roll prediction that will be used is subtracted from the total data to generate the reflector prediction. Since the ground-roll prediction consists of a erroneous surface wave, when subtracting from the total data, the result will introduce more error to the original signal. Using a more accurate ground roll prediction that contains no reflector information, we define $\eta > 1$. With the block-coordinate relaxation method, the best results are achieved by setting $C_1 = C_2$.

As shown in Table 2.3, both methods generate improved signal-to-noise ratios through the separation schemes. Except for the test using a noise prediction containing 5% model error and added white noise, the tests show similar results. The test with 5% model error limits the performance of the block-coordinate relaxation method due the increased noise level introduced into the thresholding. This method cannot estimate the reflector

information below the noise level causing loss of amplitude in the estimated reflectors. The values of the SNR seem roughly equal, but a look at Figure 2.8 shows the block-coordinate relaxation method generally effects the reflector information by removing some of the energy as estimated surface wave.

2.3.3 Parameter sensitivity

These methods require the selection of certain parameters before the separation is preformed and it is important we understand the stability and sensitivity of these parameters.

Table 2.4 shows variations of the chosen separation values $(\lambda_1^*, \lambda_2^*, \eta^*)$ and the effect on the signal-to-noise ratios as the Bayesian signal separation method is preformed with the prediction containing 5% model error. Parameters λ_1^*, λ_2^* and η^* are the values used to generate the results in Table 2.3. These results show stability for variations of the sparsity parameters with η representing the most sensitive parameter. Large variations of η without balancing of the sparsity parameters when defining the weights tend to produce the worst results as this is equivalent to much higher confidence in a specific prediction while disregarding sparsity for $\eta \gg (\lambda_1 \cdot \lambda_2)/2$ causing under thresholding or by forcing an unreasonably sparse signals for $\eta \ll (\lambda_1 \cdot \lambda_2)/2$ causing over thresholding.

SNR (dB)	$0.1 \cdot (\lambda_1^*, \lambda_2^*)$	$2 \cdot \lambda_1^*, \lambda_2^*$	λ_1^*, λ_2^*	$\lambda_1^*, 2 \cdot \lambda_2^*$	$10 \cdot (\lambda_1^*, \lambda_2^*)$
$0.1 \cdot \eta^*$	8.35	3.33	4.46	4.46	1.56
$\frac{1}{2} \cdot \eta^*$	1.86	5.83	8.88	9.00	3.33
η^*	1.76	6.93	9.59	9.02	4.45
$2 \cdot \eta^*$	-3.90	6.48	1.27	2.78	5.97
$10 \cdot \eta^*$	-4.28	-2.56	-3.38	-3.18	8.05

Table 2.4: Sensitivity of the parameters associated with the Bayesian signal separation scheme. The parameters λ_1^*, λ_2^* and η^* refer to 10, 2 and 3.5 respectively. These parameters are the same set used to generate the results of the second test in Table 2.3 and are chosen by the expected sparsity of the estimated signal and the fidelity of the ground roll prediction. Small adjustments were then made by parameter search to select the most effective values for the synthetic data.

The block-coordinate relaxation constants are more sensitive to variation with the best performance occurring when $C_1 = C_2$. This equal reduction

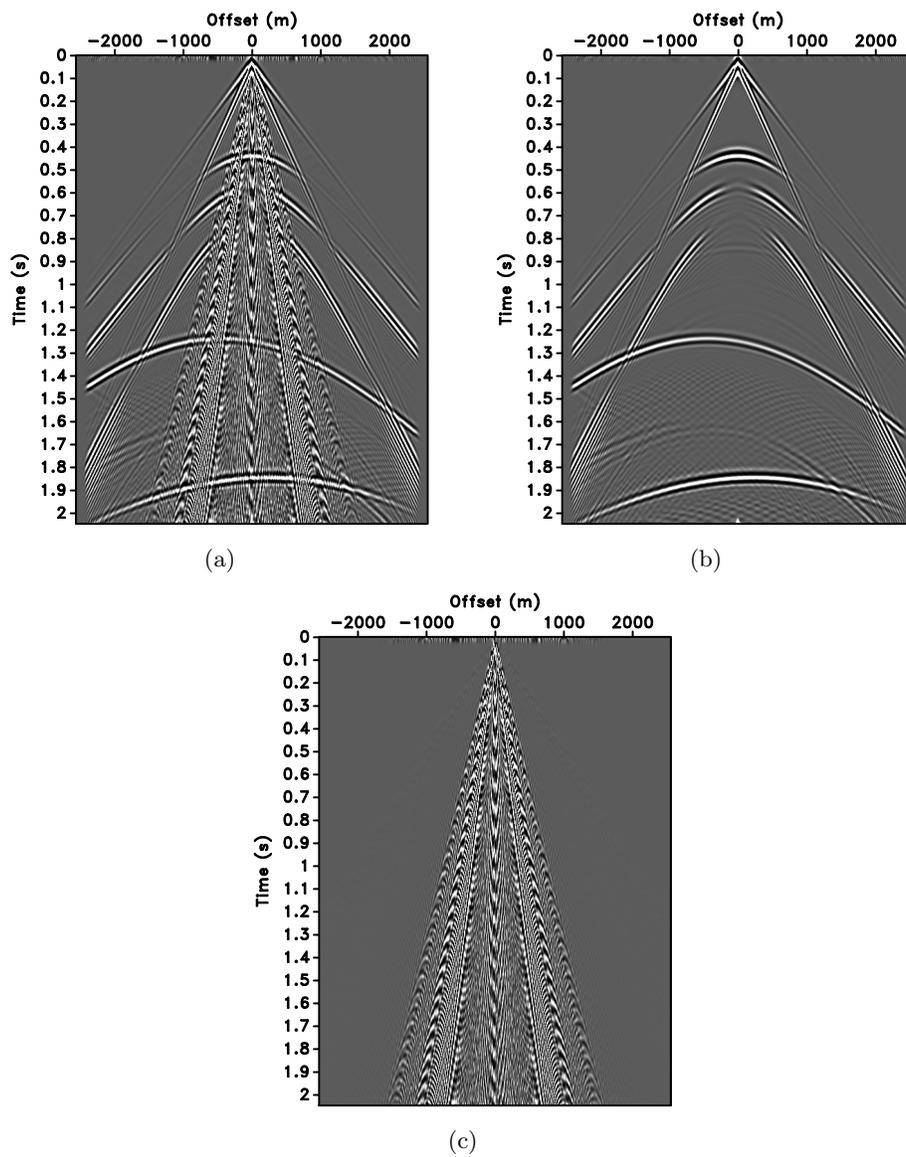


Figure 2.4: Synthetic data used to test the signal separation schemes. The full synthetic data set (a), the reflector component (b) and the generated surface wave (c). The signal to noise ratio of the full synthetic data is -1.67.

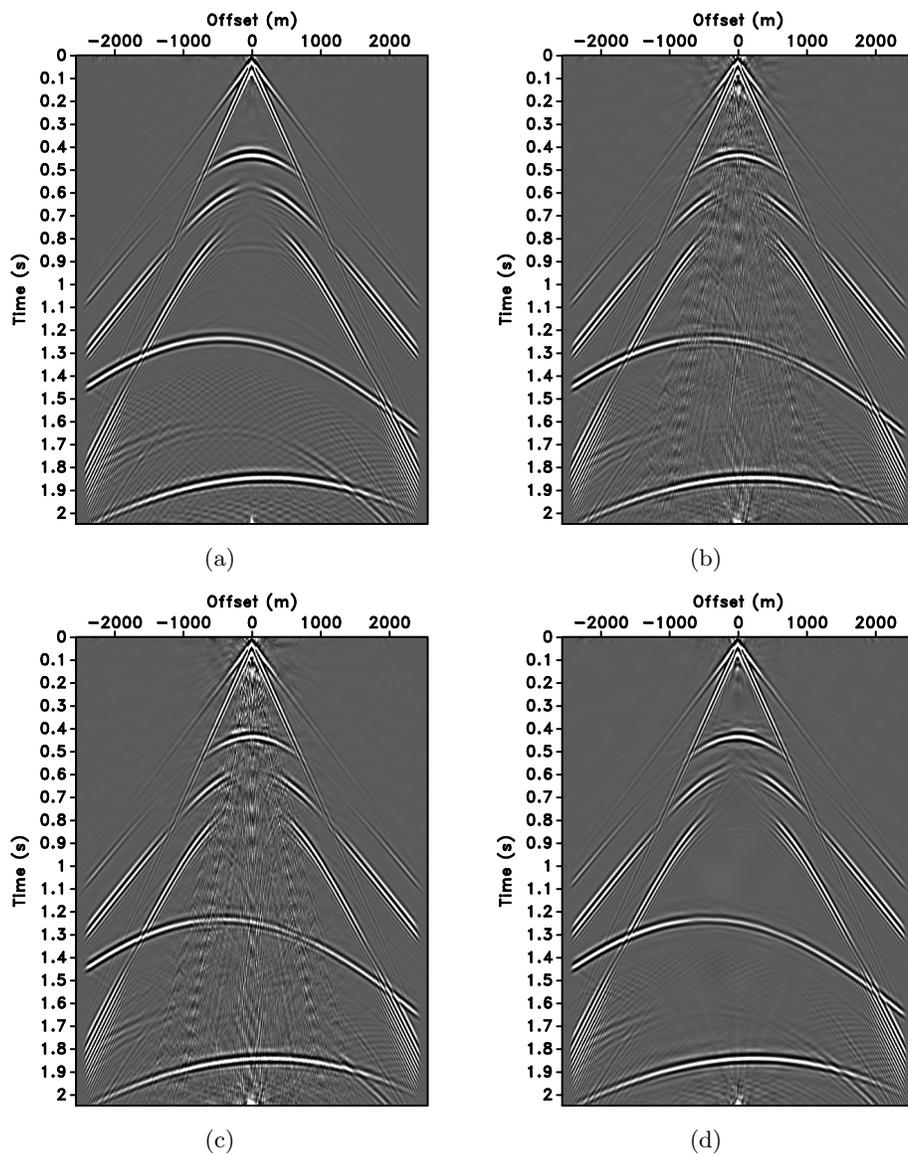


Figure 2.5: Synthetic results for the estimated reflector signal using the Bayesian separation scheme. The generated results using predictions consisting of the exact noise (a), 5% error on modeling parameters (b), 5% error on modeling parameters with added white noise (c), Hilbert transform of the exact noise (d).

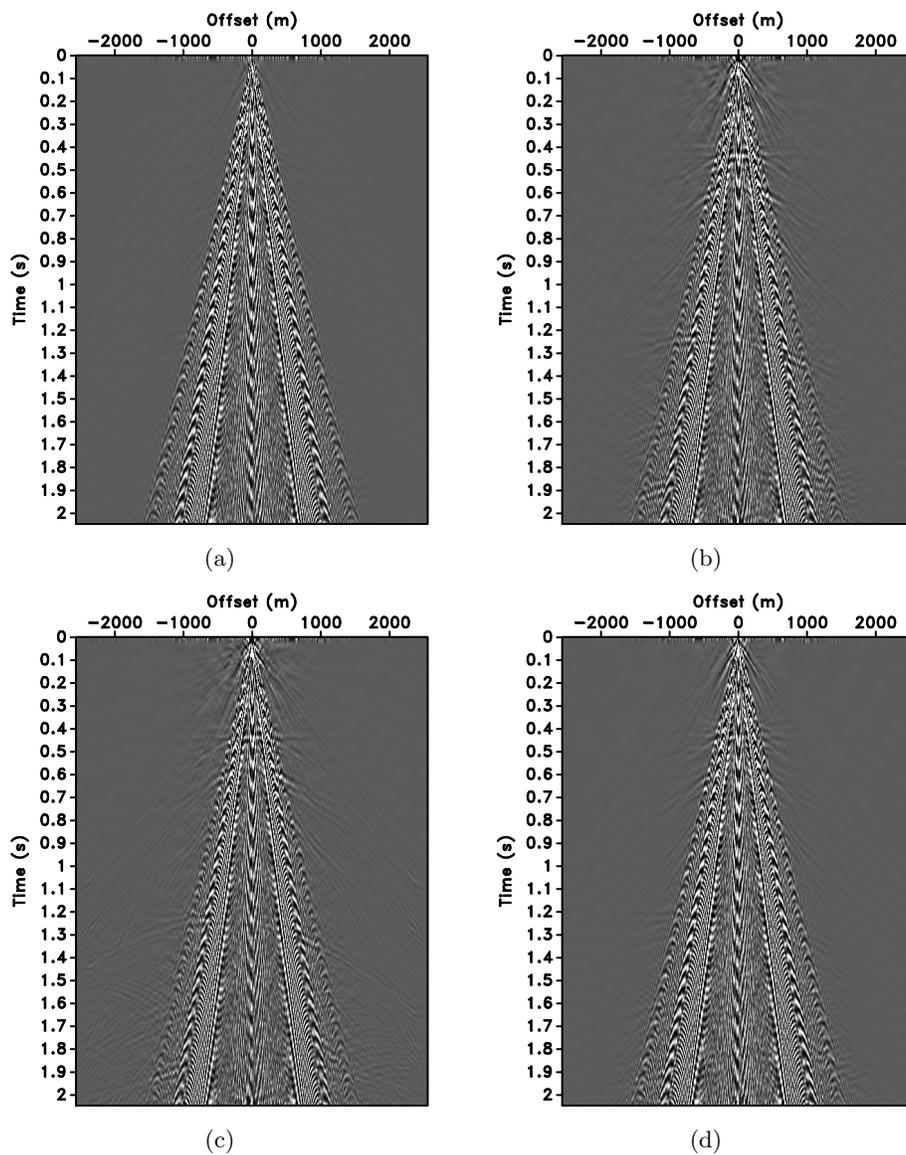


Figure 2.6: Synthetic results for the estimated surface wave signal using the Bayesian separation scheme. The generated results using an exact noise prediction (a), 5% error on modeling parameters (b), 5% error on modeling parameters with added white noise (c), Hilbert transform of the exact noise prediction (d).

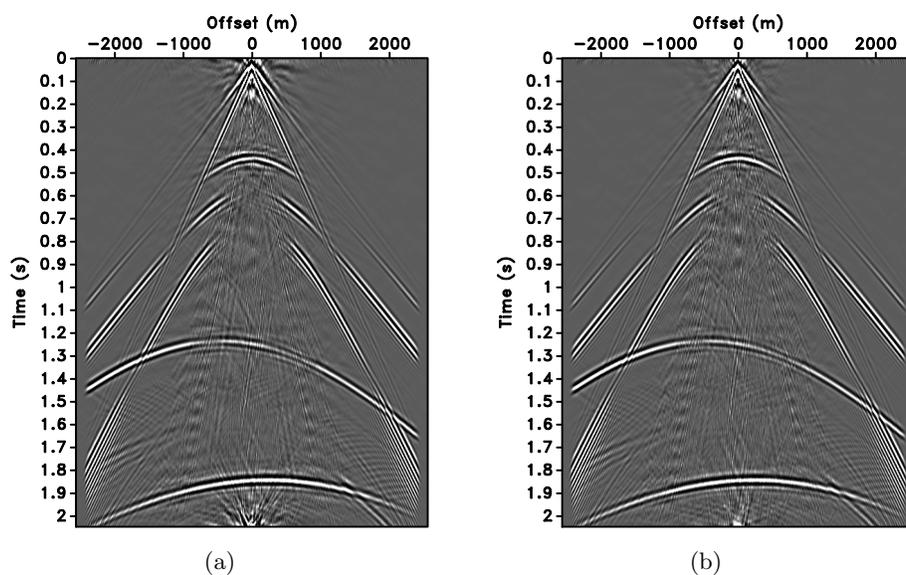


Figure 2.7: Separation results using a prediction containing 5% error and the block-coordinate relaxation method (a) and the Bayesian separation method (b). The initial dataset has a SNR of -1.67 dB while the SNR of the separated results are 9.41 and 9.58 from the block-coordinate relaxation method and the Bayesian separation method respectively.

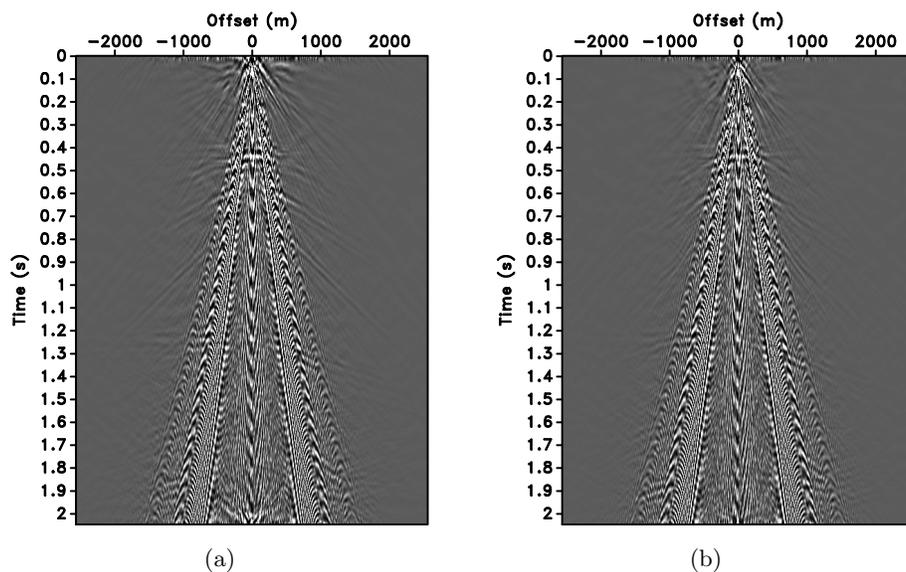


Figure 2.8: Estimated surface waves from the block-coordinate relaxation (a) and Bayesian (b) solutions to the 5% model error prediction for the synthetic data. Whereas Table 2.3 shows the SNR for the Bayesian solution is roughly equal to the block-coordinate relaxation method, we can see reflector information has been estimated as part of the surface wave where as the Bayesian solution has been more conservative and estimated less reflector information with the surface wave.

of the thresholds within the algorithm is expected to be the optimal choice for most situations as it prevents over-thresholding of each signal. This is expected with two signals of similar sparsity as over thresholding of reflector signal (for example) would start to select incoherent information associated with the ground roll signal.

SNR (dB)	$0.1 \cdot c_1^*$	$\frac{1}{2} \cdot c_1^*$	c_1^*	$2 \cdot c_1^*$	$10 \cdot c_1^*$
$0.1 \cdot c_2^*$	6.24	0.35	3.00	-1.57	-1.67
$\frac{1}{2} \cdot c_2^*$	4.09	8.83	0.62	-1.35	-1.66
c_2^*	3.00	5.01	9.41	-0.56	-1.24
$2 \cdot c_2^*$	1.66	2.47	4.02	7.85	8.42
$10 \cdot c_2^*$	0.00	0.01	0.12	0.15	0.21

Table 2.5: Sensitivity of the parameters associated with block-coordinate relaxation. These parameters are associated with the aggressiveness of the soft thresholding operator defined in the outer loop and then applied in the inner loop. In this case, the values $C_1 = C_2 = 0.4$ are used.

2.4 Application to real field data

This data was shot in the foothills of the Rocky Mountains and is likely to contain reflector information showing complex structure. This real data (shown in Figure 2.9) set faces common challenges such as unknown original signals, possible aliased surface waves and large amplitude variations. There is a strong ground roll signature for this line as the shot location was located midway along the line. It is our goal to improve the signal separation as much as possible while preserving the reflector information. As there was a time power gain applied to the data, noisy traces become increasingly prevalent at later times. This section of data is a two-dimensional line of a larger three-dimensional dataset.

As a prediction for the Bayesian separation method, a dataset with the coherent and incoherent noise removed by conventional methods, has been provided which will be used as a reflector prediction. This original denoised data contains not only ground roll, but also noisy traces and some reflector information that was incorrectly identified as unwanted noise. The goal of our separation algorithm is to improve the ground roll and reflector predictions (seen in Figure 2.10) by reducing the amount of ground roll in the reflector prediction and moving the incorrectly identified reflector informa-

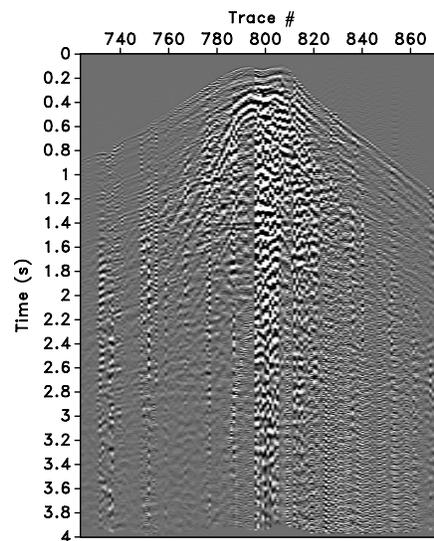


Figure 2.9: The real dataset to be improved with the Bayesian separation method. This dataset comes from a line that was located near the source thus causing a dominant ground roll signal in the center of the image. This data is also contaminated with several noisy traces that become increasingly obvious for later times in the signal as the data has had an applied time power gain added.

tion in the ground roll prediction back to the reflector estimation.

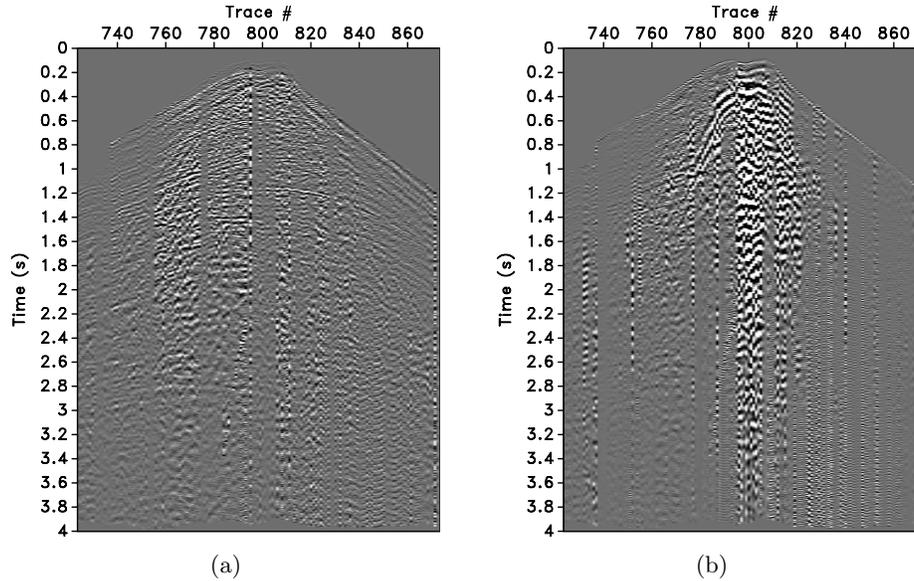


Figure 2.10: The predictions to be used with the Bayesian separation method for the (a) reflectors and (b) surface wave. The reflector prediction is a supplied denoised data set with the surface wave prediction as the difference between the total data and the reflector prediction.

The parameters of the Bayesian solver need to be set to generate an appropriate separation. The sparsity of the expected estimates will tend to lean to the ground roll estimate being more sparse than the reflectors. The reflector information is present in the majority of the signal and can potentially be complex, containing large and small scale structures. The ground roll will only occupy a region defined by the low velocity of the surface wave and will not be present in any other part of the signal. There is extra associated noise that will be part of the ground roll prediction but this is not expected to contribute much to the overall sparsity of the signal. For these reasons we set $\lambda_1 = 1$ and $\lambda_2 = 3$. The parameter η is set to 3 in this case as the reflector prediction was constructed with the expressed purpose of excluding as much surface wave information as possible. We will attempt to recover the data that has been removed through the original denoising processes while simultaneously reducing the noise in the reflector signal. As the reflector prediction scheme was similar for all the slices pertaining to this data set, these parameters can generally be applied to all of the two-

dimensional slices.

The calculated reflector and surface wave estimations after five iterations of the Bayesian scheme are shown in Figure 2.11. The difference between the surface wave prediction and estimation is shown in Figure 2.12. From this difference plot, we can see that the reflector information that is located in the upper left corner of the image have been identified as reflector information and this energy has been moved to the reflector estimate. We can also see much of the ground roll energy that was left behind through the original prediction generation has been identified as surface wave energy and thus moved to the surface wave estimate.

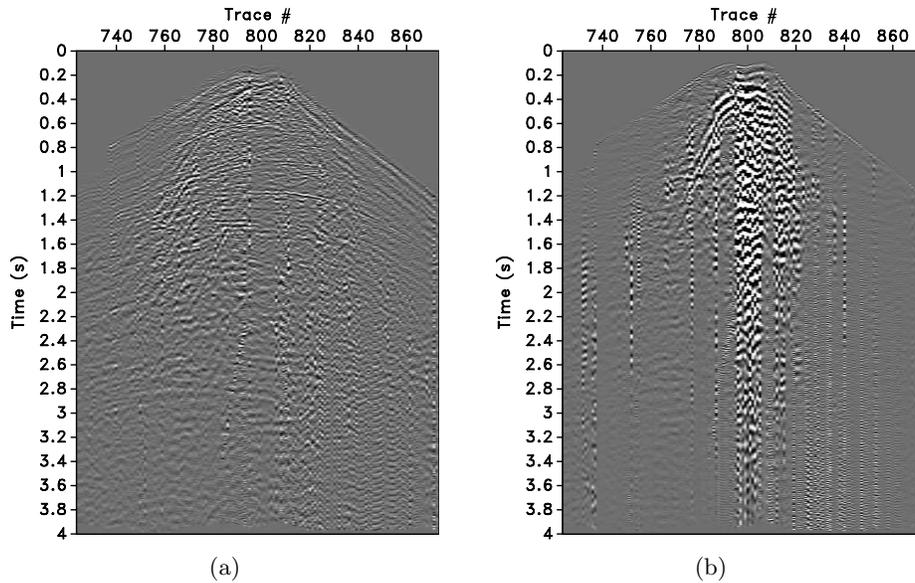


Figure 2.11: The Bayesian (a) reflector and (b) surface wave estimates for the real data set after five iterations. These images show improved energy content for both the reflector and surface wave estimates. There is still a small amount of linear noise in the estimated reflectors but it is less than the original prediction.

2.5 Discussion

Both signal separation methods performed well with the synthetic data suite with the exception of the block-coordinate relaxation method on the third test consisting of a erroneous prediction and added noise. This is

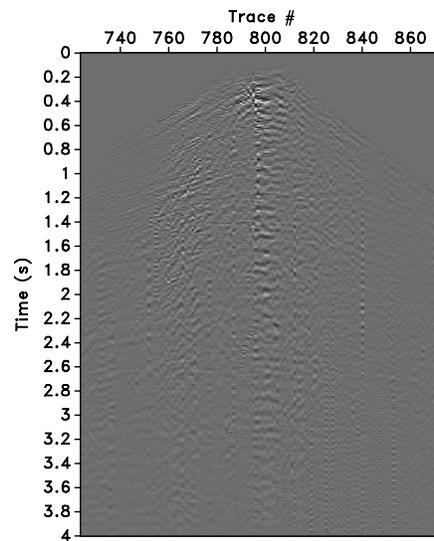


Figure 2.12: The difference between the predicted surface wave and the Bayesian estimated surface wave. This image shows the energy transferred between predictions through the estimation scheme. As we can see, reflector information, particularly in the top left section of the image and ground roll energy has been transferred back to their appropriate estimates.

due to the noise level interfering with the weighted thresholding scheme. The smaller curvelet coefficients that are associated with features below the noise level become over thresholded due to the settings of the noise level. To improve these results, the ϵ value in 2.5 can be lowered below the noise level while reducing the block-coordinate threshold coefficients C_1 and C_2 and increasing the number of outer loops that is preformed. In the case where $C_1 = C_2 = .1$ with $\epsilon = 5e^{-4}$ and 50 outer loops, the SNR can be increased to 4.30. While this is an improvement, this there is roughly ten times the iterations as the previous result with a result that is still marginal at best.

The block-coordinate relaxation method was not used on the real data set due to the quality of the prediction used. The predicted reflectors and ground roll did not constitute a conservative enough prediction to generate an effective estimate. This problem becomes apparent when the signal separation is attempted as artifacts due to over thresholding start to appear. Due to the construction of the curvelet domain, the sparse representation of a single reflector will require a few large curvelet coefficients but the smaller coefficients need to be respected and cannot be assumed to be noise when they lie below the noise level. This is the weakness of the block-coordinate relaxaton method. The Bayesian separation method takes into account the noise in each signal and attempts to recover the exact signals and there associated noise components. As we can see with the result shown in Table 2.3, since the random noise component of the same variance was part of the prediction, this was associated with the ground roll signal and removed while preserving the reflectors.

The lack of a conservative estimate can be compensated by proper setting of the fidelity and sparsity parameters resulting in an improved separation. This is important as the generation of new conservative estimates for datasets can be time consuming and costly so the addition of this flexibility allows this method to be directly implemented on existing problems to improve signal separations.

Multiple predictive steps, if required, will also work well with these separation methods and it may be advantageous to combine two or more methods to generate a prediction covering a wide range of characteristics, such as frequency, velocity, or amplitude identifiers. Since the sum of the generated estimates is equal to the total data, no information is lost if repeated iterations are required of the entire process or if another noise estimate generated of the estimated results of a previous separation.

2.6 Conclusion

The two separation methods discussed take two different approaches to signal separation. The block-coordinate relaxation method was designed originally to utilize two distinct separate signals. This method requires these signals have sparse representations but originally does not account for signals that are sparse in the same domain. The addition of the weighting scheme by Herrmann et al. (2007b) has made use of a sparsifying transform applied to similar signals. The Bayesian separation scheme was designed to be more flexible under noise and varying predictions. By maximizing the probability of recovering the desired signals given certain estimates, and the addition of parameters that increase the flexibility of the method, an effective separation algorithm has been developed that can be applied to ground roll separation.

These separation methods take advantage of the curvelet domain as the most sparse representation for the full dataset, as well as the reflector and the ground roll components (as shown in Figure 2.2). The specific design of curvelets (see Figure 2.1) allows for the multiscale and multidirectional nature of the transform that supplies this unique efficiency.

Both these methods have proven effective for certain synthetic examples with the Bayesian separation method out performing the block-coordinate relation method in most circumstances. The design of the Bayesian separation method provides a more stable and effective platform for signal separation through incorporation of the noise components into the signal and increased flexibility which is shown by the noisy synthetic example. The block-coordinate relaxation method generally performs well for most situations but when noise is added it requires too many iterations to be effective. The parameters for the block-coordinate relaxation method have also proved to be much more sensitive than the bayesian separation method.

We have shown how the Bayesian separation method can be utilized to improve previous denoised results by improving the overall signal separation. By taking data that was previously denoised, we have shown that with very simple parameter settings, a improvement can be calculated for both the reflector and the surface wave estimates. This improvement is simple to apply and helps preserve the original reflector amplitudes while improving the overall denoising.

2.7 Acknowledgments

The authors would like to thank the project sponsors for making this research possible. The authors would also like to thank the authors of the Fast Discrete Curvelet Transform for making this code available. The authors would also like to thank David Campden, Henning Kuehl and Tommy Kwan at Shell Canada for providing the field dataset and for their feedback and support. This work is in part supported by NSERC Discovery Grant 22R81254 of Felix J. Herrmann and was carried out as part of the SINBAD project with support from BG Group, BP, Chevron, ExxonMobil and Shell.

Bibliography

- Abma, R. and N. Kabir, 2006, 3D interpolation of irregular data with a POCS algorithm: *Geophysics*, **71**, no. 6, E91 – E97.
- Candès, E. J., L. Demanet, D. L. Donoho, and L. Ying, 2005, Fast discrete curvelet transforms: *SIAM Multiscale Model. Simul.*, **5**, 861–899.
- Candès, E. J. and D. Donoho, 1999, Ridgelets: the key to high-dimensional intermittency: *Phil. Trans. R. Soc. Lond. A*, **357**, 2495–2509.
- Candès, E. J. and D. L. Donoho, 2000, Curvelets – a surprisingly effective nonadaptive representation for objects with edges. *Curves and Surfaces*. Vanderbilt University Press.
- Coruh, C. and J. K. Costain, 1983, Noise attenuation by vibroseis whitening (vsw) processing: *Geophysics*, **48**, 543–554.
- Daubechies, I., M. Defrise, and C. de Mol, 2005, An iterative thresholding algorithm for linear inverse problems with a sparsity constraint: *CPAM*, 1413–1457.
- Deighan, A. and D. Watts, 1997, Ground-roll suppression using the wavelet transform: *Geophysics*, **62**, 1896–1903.
- Demanet, L. and L. Ying, 2007, Wave atoms and sparsity of oscillatory patterns: *Appl. Comput. Harmon. Anal.*, **23-3**, 368–387.
- Elad, M., J. L. Starck, P. Querre, and D. L. Donoho, 2005, Simultaneous cartoon and texture image inpainting using morphological component analysis (MCA): *Appl. Comput. Harmon. Anal.*, **19**, 340–358.
- Embree, P., J. P. Burg, and M. M. Backus, 1963, Wide band velocity filtering-the pie-slice process: *Geophysics*, **28**, 948–974.
- Fail, J. P. and G. Grau, 1963, Les filters on eventail: *Geophys. Prospect.*, **11**, 131–163.
- Gabor, D., 1946, Theory of communication: *J. Inst. Elect. Eng.*, **93**, 429–457.
- Galbraith, J. N. and R. A. Wiggins, 1968, Characteristics of optimum multichannel stacking filters: *Geophysics*, **33**, 36–48.

- Gelişli, K. and H. Karşlı, 1988, F-k filtering using the hartley transform: *J. Seism. Explor.*, **7**, 101–108.
- Graves, R., 1996, Simulating seismic wave propagation in 3d elastic media using staggered-grid finite differences: *Bulletin of the Seismological Society of America*, **86**, 1091–1106.
- Harlan, W. S., J. F. Claerbout, and F. Rocca, 1984, Signal/noise separation and velocity estimation: *Geophysics*, **49**, 1869–1880.
- Hennenfent, G. and F. Herrmann, 2007, Irregular sampling: from aliasing to noise: Presented at the EAGE 69th Conference & Exhibition.
- Herrmann, F., D. Wang, G. Hennenfent, and P. Moghaddam, 2007a, Curvelet-based seismic data processing: a multiscale and nonlinear approach: *Geophysics*, **73**, A1–A5.
- Herrmann, F. J., U. Boeniger, and D.-J. E. Verschuur, 2007b, Nonlinear primary-multiple separation with directional curvelet frames: *Geoph. J. Int.* (To appear.).
- Herrmann, F. J., P. P. Moghaddam, and C. Stolk, 2007c, Sparsity- and continuity-promoting seismic imaging with curvelet frames. (Accepted for publication in the *Journal of Applied and Computational Harmonic Analysis*).
- Holzman, M., 1963, Chebysev optimized geophone arrays: *Geophysics*, **28**, 145–155.
- Karşlı, H. and Y. Bayrak, 2004, Using the Wiener-Levinson algorithm to suppress ground-roll: *Journal of Applied Geophysics*, **55**, 187–197.
- Liu, X., 1999, Ground roll spression using the Karhuren-Loeve transform: *Geophysics*, **64**, 564–566.
- Londono, E. G., L. C. Lòpez, and T. S. Kazmierczak, 2005, Using the karhunen-loève transform to suppress ground roll in seismic data: *Earth Sci. Res. J.*, **9**, 139–147.
- McKay, A. E., 1954, Review of pattern shooting: *Geophysics*, **19**, 420–437.
- McMechan, G. A. and R. Sun, 1991, Depth filtering of first breaks and ground roll: *Geophysics*, **56**, 390–396.
- Russel, B., D. Hampson, and J. Chun, 1990, Noise elimination and the radon transform; part 1: *The Leading Edge*, **9**, no. 10, 18–23.
- Saab, R., D. Wang, O. Yilmaz, and F. Herrmann, 2007, Curvelet-based primary-multiple separation from a bayesian perspective: Presented at the SEG International Exposition and 77th Annual Meeting.

- Sacchi, M. D., T. J. Ulrych, and C. Walker, 1998, Interpolation and extrapolation using a high-resolution discrete Fourier transform: *IEEE Trans. Signal Process.*, **46**, no. 1, 31–38.
- Starck, J. L., M. Elad, and D. Donoho, 2004, Redundant multiscale transforms and their application to morphological component separation: *Advances in Imaging and Electron Physics*, **132**.
- Trad, D., T. Ulrych, and M. Sacchi, 2003, Latest views of the sparse radon transform: *Geophysics*, **68**, 386–399.
- Treitel, S., J. L. Shanks, and C. W. Fraster, 1967, Some aspects of fan filtering: *Geophysics*, **32**, 789–800.
- Verschuur, D. J., A. J. Berkhout, and C. P. A. Wapenaar, 1992, Adaptive surface-related multiple elimination: *Geophysics*, **57**, 1166–1177.
- Wang, D., R. Saab, O. Yilmaz, and F. Herrmann, 2007, Recent results in curvelet-based primary-multiple separation: application to real data: Presented at the SEG International Exposition and 77th Annual Meeting.
- , 2008, Bayesian wavefield separation by transform-domain sparsity promotion: *Geophysics*. (Accepted for publication).
- Xu, S., Y. Zhang, D. Pham, and G. Lambare, 2005, Antileakage Fourier transform for seismic data regularization: *Geophysics*, **70**, no. 4, V87 – V95.
- Yarham, C., U. Boeniger, and F. Herrmann, 2006, Curvelet-based ground roll removal: Presented at the SEG International Exposition and 76th Annual Meeting.
- Yilmaz, O., 2001, *Seismic data analysis. Curves and Surfaces*. Society of Exploration Geophysicists.
- Zhang, R. and T. J. Ulrych, 2003, Physical wavelet frame denoising: *Geophysics*, **68**, 225–231.
- Zwartjes, P. M. and M. D. Sacchi, 2007, Fourier reconstruction of nonuniformly sampled, aliased seismic data: *Geophysics*, **72**, no. 1, V21–V32.

Chapter 3

Conclusions

The purpose of this thesis is to explore the ability of signal separation methods applied to sparse seismic data, focusing on the problem of ground roll removal. Traditionally, the ground roll problem has been approached through methods focused on the frequency and velocity properties of the ground roll. While these methods have been generally successful for the last few decades, it is our ambition to improve these techniques and update these methods.

The fundamental concept behind the methods discussed in this thesis are based on sparse seismic signal representation. While this is not a new concept (Sacchi et al., 1998; Trad et al., 2003; Xu et al., 2005; Abma and Kabir, 2006; Zwartjes and Sacchi, 2007; Herrmann et al., 2007a), the development of the many classes of two-dimensional transforms designed to represent wavefields is relatively new. The ridgelet, curvelet, and Wave Atom transforms have significantly added to the list of sparse transforms. The curvelet transform, in particular, has provided the most sparse representation of seismic data when normalized by length. This length normalized sparsity opens up many new methods that can be utilized and applied to seismic data such as the previous separation algorithms.

Two signal separation methods have been introduced in this thesis for application to ground roll removal. The block-coordinate relaxation method was adapted by previous methods of Starck et al. (2004) and Elad et al. (2005) by Herrmann et al. (2007b) for use with a single sparse transform and seismic problems pertaining to multiple removal. This first generation method was adapted to reflect factors such as expected sparsity and prediction fidelity by Saab et al. (2007) to produce the second generation Bayesian separation algorithm. Both methods have their benefits and limitations with the second generation Bayesian algorithm generally outperforming the block-coordinate relaxation method. The increased flexibility and stability of the algorithm provides a separation scheme that is effective under noise and real data. The Bayesian solution also reduces reflector degradation and generates few artifacts compared to the block-coordinate relaxation method.

One of the main benefits of this research is that it shows how the sig-

nal separation methods, particularly the Bayesian separation method, can be utilized with previously existing processing flows. The flexibility of the Bayesian separation method allows it to adapt to different predictive methods with little difficulty. The addition of the signal separation methods to ground roll removal can improve separations with better reflector preservation than was previously available. While the use of a signal separation method will require extra computational resources, the cost of preserving reflector information is small when considering the initial costs of recording this information. The simplicity of the Bayesian separation method in its ability to accept and improve previously denoised data, as shown with the real data set shown in Chapter 2, makes this method available to improve previously unacceptable ground roll separation results.

3.1 Signal separation methods

The signal separation methods discussed cannot simply be utilized without some thought to the properties of the data to be separated. The block-coordinate relaxation method requires the setting of two parameters pertaining to the aggressiveness of the soft thresholding operator that is applied to the two signals. These parameters are always set to values between zero and one but have been shown to be sensitive to variations. The simplest rules for these coefficients seem to be to set them to an equal value due to the sensitivity. Variation in the parameters with respect to each other generally over thresholds one coefficient vector in relation to the other, causing poor results.

The parameters associated with the Bayesian method tend to add flexibility without compromising the stability or efficiency of the method. While the λ parameters pertaining to expected sparsity are seemingly subjective, this tends only to be the case when dealing with synthetic or abnormally regular data. When dealing with real seismic data in the context of surface wave separation, it can generally be assumed that these parameters will be roughly equal. The η parameter allows the algorithm to take full advantage of conservative estimates and perform more efficiently. This parameter sets a preference for one prediction versus another such that an estimate containing only one signal can be recognised by the algorithm. This allows the effect of the coefficients associated with the opposing signal to be reduced.

The block-coordinate relaxation method, while effective under certain circumstances, is generally less effective than the Bayesian separation method. The major faults of this method are generally due to reflector degradation

and artifact generation. These are problems associated with over thresholding in the curvelet domain under noisy conditions and can generally be reduced by small modifications to the thresholding parameters. The reduction of the threshold coefficients to improve the results is followed by a marked increase in the number of outer loops that must be performed to facilitate the separation. This increase in iterations can make this method computationally expensive to the point where it is unfeasible for large datasets. The block-coordinate relaxation method can be used to successfully separate two signals as shown in the previous chapter, if it is used with certain predictive methods. It is generally better than no signal separation at all (ie, direct subtraction) but care needs to be taken to preserve the reflector information.

The Bayesian separation method has been shown to typically outperform the block-coordinate relaxation method in the presence of noise and for most other situations. The ability to add a priori information regarding expected sparsity leads to advances in the denoising of sparse synthetic data. While this information is generally not well known for real data, relative values can be estimated. The application to real problems benefits from the addition of the η parameter allowing information regarding the fidelity of the predictions to be taken into account and utilized. While this parameter is usually the most sensitive parameter of the Bayesian separation method, it is generally less sensitive than the parameters associated with the block-coordinate relaxation method. This parameter is critical when considering how to add these methods to a preexisting denoising flow as it allows the separation method to adapt to different prediction methods.

The methods discussed in this thesis have been focused on the problem of surface wave separation built from the generic signal separation problem. The only restriction that has been placed on these methods in relation to the type of data used is that data is required to be sufficiently sparse under the sparsifying transform. Typically, with seismic data and the curvelet transform, this will always hold. There are no limitations on the type of seismic separation problems that may be performed with these methods regardless of the signals to be separated. These methods may be adapted to a variety of signal separation problems as long as the signals are sparse under some basis, be it the curvelet, wavelet, Fourier or any other basis. The methods developed have focused on a single sparse basis by utilizing a weighting scheme thus allowing separation of similar signals, such as wavefronts. If the most parsimonious basis for two signals is something other than the curvelet domain, particularly if it is redundant, little modification needs to be done to utilize these methods. In the event that an orthonormal basis is found for a particular problem, modifications could be made to the algorithms to

accommodate the increased simplicity.

3.2 Open and future research

The use of the curvelet domain and the introduction of sparse signal separation relating to ground roll removal have opened up many new avenues of research. The introduction of both the block-coordinate relaxation method and the Bayesian separation method add two new tools into the kits of seismic data processors which may be advantageous for certain situations but are flexible enough to be incorporated into many new and old problems.

The use of the curvelet domain for seismic signals is still in its infancy and will likely continue to expand as new techniques and methods are developed to take advantage of its length normalized sparsity. The curvelet domain also allows access to specific scales, angles and locations which have yet to be fully utilized. Not only can the curvelet domain be utilized in the development of location specific velocity and frequency filters, but it may be possible to improve certain signal separation methods by utilizing location based information through the use of restricted curvelet vectors. If the signal estimation schemes are restricted to removal of ground roll only in areas where ground roll is expected, effects on reflector information outside the ground roll cone can be removed. This may also prove to be possible for the angles and frequencies of the reflectors but will likely run the risk of altering reflector information and thus restrictions should only be placed on surface wave components.

Another avenue for future work is the influx of new transforms. While the curvelet transform is the most parsimonious of the transforms discussed, many other similar systems are being developed. Shearlets (Labate et al., 2005) are a good example of this. While not developed to the extent that the curvelet transform has been, shearlets show promise as a competitive transform with the curvelet domain. New developments in sparse signal representation may lead to a transform that is more efficient than the methods currently used in this thesis and will warrant exploration.

Pertaining to the block-coordinate relaxation method, initially research was done with the purpose to find two efficient transforms for both ground roll and reflector information as was done by Elad et al. (2005). While the curvelet domain was used to represent reflectors, with Wave Atoms and ridgelets to described ground roll, it was determined that the ability of curvelets to sparsely represent both signals was a barrier to these original techniques. With the development of new transforms, it is possible that a

transform will emerge that meets the required conditions for this method and reevaluation would be required.

One of the more challenging parts of the discussed methods centers on parameter selection. There currently is no automated way of selecting optimal parameters for the separation methods. While fully automated parameter selection may be a bit ambitious, schemes that automatically adjust the parameters during the separation scheme may be possible. One possible method of this would inversion of the parameters using reflector information outside of the ground roll cone to minimize reflector degradation. For large scale applications, this may require significant computational resources and may not be feasible but with transforms of increasing speed, this may be something of interest in the future.

Bibliography

2008, Bayesian-signal separation by sparsity promotion: application to primary-multiple separation.

Abma, R. and N. Kabir, 2006, 3D interpolation of irregular data with a POCS algorithm: *Geophysics*, **71**, no. 6, E91 – E97.

Berkhout, A. J. and D. J. Verschuur, 2006, Focal transformation, an imaging concept for signal restoration and noise removal: *Geophysics*, **71**.

Black, M., G. Sapiro, and D. Marimont, 1998, Robust anisotropic diffusion: *IEEE Trans. Signal Process.*, **7**, 421–432.

Bregman, L., 1965, The method of successive projection for finding a common point of convex sets: *Soviet Math. Dokl.*, **6**, 688–692.

Bruce, A. G., 1998, Block coordinate relaxation methods for nonparametric signal de-noising: Presented at the .

Candès, E. and D. Donoho, 2005, Continuous Curvelet Transform I: Resolution of the Wavefront Set: *Appl. Comput. Harmon. Anal.*, **19**, 162–197.

Candès, E., J. Romberg, and T. Tao, 2005, Stable signal recovery from incomplete and inaccurate measurements: *Comm. Pure Appl. Math.*, **59**, 1207–1223.

Candès, E. J., 1999, Harmonic analysis of neural networks: *Appl. Comput. Harmon. Anal.*, **6**, 197–218.

Candès, E. J. and L. Demanet, 2004, The curvelet representation of wave propagators is optimally sparse: *Comm. Pure Appl. Math.*, **58**, 1472–1528.

Candès, E. J., L. Demanet, D. L. Donoho, and L. Ying, 2005, Fast discrete curvelet transforms: *SIAM Multiscale Model. Simul.*, **5**, 861–899.

Candès, E. J. and D. Donoho, 1999, Ridgelets: the key to high-dimensional intermittency: *Phil. Trans. R. Soc. Lond. A*, **357**, 2495–2509.

Candès, E. J. and D. L. Donoho, 2000, Curvelets – a surprisingly effective nonadaptive representation for objects with edges. *Curves and Surfaces*. Vanderbilt University Press.

- Candés, E. J. and J. Romberg, 2004, Practical signal recovery from random projections: Presented at the Wavelet Applications in Signal and Image Processing XI.
- Chauris, H., 2006, Seismic imaging in the curvelet domain and its implications for the curvelet design: Presented at the 76th Ann. Internat. Mtg., SEG, Soc. Expl. Geophys., Expanded abstracts.
- Chen, S., D. Donoho, and M. Saunders, 2001, Atomic decomposition by basis pursuit: *SIAM J. Sci. Comp.*, **43**, 129–159.
- Claerbout, J. and F. Muir, 1973, Robust modeling with erratic data: *Geophysics*, **38**, 826–844.
- Coifman, R. and D. Donoho, 1995, Wavelets and statistics, chapter translation-invariant de-noising, 125–150. Springer-Verlag.
- Coruh, C. and J. K. Costain, 1983, Noise attenuation by vibroseis whitening (vsw) processing: *Geophysics*, **48**, 543–554.
- Daubechies, I., 1992, Ten lectures on wavelets: SIAM.
- Daubechies, I., M. Defrise, and C. de Mol, 2005, An iterative thresholding algorithm for linear inverse problems with a sparsity constraints: CPAM, 1413–1457.
- Deighan, A. and D. Watts, 1997, Ground-roll suppression using the wavelet transform: *Geophysics*, **62**, 1896–1903.
- Deighan, A. J. and D. Watts, 1996, 2d filtering using the wavelet packet transform: Presented at the EAGE 58th Conference and Technical Exhibition.
- Demagnet, L. and L. Ying, 2007, Wave atoms and sparsity of oscillatory patterns: *Appl. Comput. Harmon. Anal.*, **23-3**, 368–387.
- Do, M. and M. Vetterli, 2001, Pyramidal directional filter banks and curvelets.: *Image Processing, 2001, Proceedings. 2001, International Conference on*, 158–161.
- Donoho, D., 1995, De-noising by soft thresholding: **41**, 613–627.
- Donoho, D., M. Elad, and V. Temlyakov, 2006, Stable recovery of sparse overcomplete representations in the presence of noise: **52**, 6–18.
- Donoho, D. L., 2006, Compressed sensing: *IEEE Trans. Inform. Theory*, **52**, 1289–1306.
- Douma, H. and M. de Hoop, 2006, Leading-order seismic imaging using curvelets: Presented at the 76th Ann. Internat. Mtg., SEG, Soc. Expl. Geophys., Expanded abstracts.

- Duijndan, A. and M. Schonmeville, 1999, Nonuniform fast Fourier transform: *Geophysics*, **64**, 539–551.
- Elad, M., J. L. Starck, P. Querre, and D. L. Donoho, 2005, Simultaneous cartoon and texture image inpainting using morphological component analysis (MCA): *Appl. Comput. Harmon. Anal.*, **19**, 340–358.
- Embree, P., J. P. Burg, and M. M. Backus, 1963, Wide band velocity filtering—the pie-slice process: *Geophysics*, **28**, 948–974.
- Fail, J. P. and G. Grau, 1963, Les filters on eventail: *Geophys. Prospect.*, **11**, 131–163.
- Figueiredo, M. and R. Nowak, 2003, An EM algorithm for wavelet-based image restoration: *IEEE Trans. Image Processing*, **12**, 906–916.
- Fomel, S. and A. Guitton, 2006, Regularizing seismic inverse problems by model reparameterization using plane-wave construction: *Geophysics*, **71**, A43–A47.
- Gabor, D., 1946, Theory of communication: *J. Inst. Elect. Eng.*, **93**, 429–457.
- Galbraith, J. N. and R. A. Wiggins, 1968, Characteristics of optimum multichannel stacking filters: *Geophysics*, **33**, 36–48.
- Gelişli, K. and H. Karsh, 1988, F-k filtering using the hartley transform: *J. Seism. Explor.*, **7**, 101–108.
- Graves, R., 1996, Simulating seismic wave propagation in 3d elastic media using staggered-grid finite differences: *Bulletin of the Seismological Society of America*, **86**, 1091–1106.
- Harlan, W. S., J. F. Claerbout, and F. Rocca, 1984, Signal/noise separation and velocity estimation: *Geophysics*, **49**, 1869–1880.
- Hennenfent, G. and F. Herrmann, 2006a, Application of stable signal recovery to seismic interpolation: Presented at the SEG International Exposition and 76th Annual Meeting.
- , 2007a, Irregular sampling: from aliasing to noise: Presented at the EAGE 69th Conference & Exhibition.
- Hennenfent, G. and F. J. Herrmann, 2006b, Seismic denoising with non-uniformly sampled curvelets: *IEEE Comp. in Sci. and Eng.*, **8**, 16–25.
- , 2007b, Simply denoise: wavefield reconstruction via jittered under-sampling: Technical report, UBC Earth & Ocean Sciences Department. (TR-2007-5).

- Herrmann, F., D. Wang, G. Hennenfent, and P. Moghaddam, 2007a, Curvelet-based seismic data processing: a multiscale and nonlinear approach: *Geophysics*, **73**, A1–A5.
- Herrmann, F. J., 2003, Multifractional splines: application to seismic imaging: *Proceedings of SPIE Technical Conference on Wavelets: Applications in Signal and Image Processing X*, 240–258.
- Herrmann, F. J., U. Boeniger, and D.-J. E. Verschuur, 2007b, Nonlinear primary-multiple separation with directional curvelet frames: *Geoph. J. Int.* (To appear.).
- Herrmann, F. J. and G. Hennenfent, 2007a, Non-parametric seismic data recovery with curvelet frames: Technical report, UBC Earth & Ocean Sciences Department. (doi:10.1111/j.1365-246X.2007.03698.x.).
- , 2007b, Non-parametric seismic data recovery with curvelet frames: *Geophysical Journal International*, **170**, 781–799.
- Herrmann, F. J., P. P. Moghaddam, and R. Kirilin, 2005, Optimization strategies for sparseness- and continuity- enhanced imaging: Theory: Presented at the EAGE 67th Conference & Exhibition Proceedings.
- Herrmann, F. J., P. P. Moghaddam, and C. Stolk, 2007c, Sparsity- and continuity-promoting seismic imaging with curvelet frames. (Accepted for publication in the *Journal of Applied and Computational Harmonic Analysis*).
- Holzman, M., 1963, Chebyshev optimized geophone arrays: *Geophysics*, **28**, 145–155.
- Karsli, H. and Y. Bayrak, 2004, Using the Wiener-Levinson algorithm to suppress ground-roll: *Journal of Applied Geophysics*, **55**, 187–197.
- Labate, D., W.-Q. Lim, J. Kutyniok, and G. Weiss, 2005, Sparse multi-dimensional representations using shearlets: *Proc. SPIE Wavelets XI*, 254–262.
- Lee, N.-Y. and B. J. Lucier, 2001, Wavelet methods for inverting the Radon transform with noisy data: **10**, 79–94.
- Liu, X., 1999, Ground roll suppression using the Karhunen-Loeve transform: *Geophysics*, **64**, 564–566.
- Londono, E. G., L. C. Lòpez, and T. S. Kazmierczak, 2005, Using the karhunen-loève transform to suppress ground roll in seismic data: *Earth Sci. Res. J.*, **9**, 139–147.
- Mallat, S. G., 1997, *A wavelet tour of signal processing*. Academic Press.

- McKay, A. E., 1954, Review of pattern shooting: *Geophysics*, **19**, 420–437.
- McMechan, G. A. and R. Sun, 1991, Depth filtering of first breaks and ground roll: *Geophysics*, **56**, 390–396.
- McMechan, G. A. and M. J. Yedlin, 1981, Analysis of dispersive waves by wavefield transformation: *Geophysics*, **49**, 1169–1179.
- Olhovich, V. A., 1964, The causes of noise in seismic reflection and refraction work: *Geophysics*, **29**, 1015–1030.
- Rickett, J. E., 2003, Illumination-based normalization for wave-equation depth migration: *Geophysics*, **68**, no. 4.
- Russel, B., D. Hampson, and J. Chun, 1990, Noise elimination and the radon transform; part 1: The Leading Edge, **9**, no. 10, 18–23.
- Saab, R., D. Wang, O. Yilmaz, and F. Herrmann, 2007, Curvelet-based primary-multiple separation from a bayesian perspective: Presented at the SEG International Exposition and 77th Annual Meeting.
- Saatcilar, R. and N. Canitez, 1988, A method of ground-roll elimination: *Geophysics*, **53**, 894–902.
- Sacchi, M. D., T. J. Ulrych, and C. Walker, 1998, Interpolation and extrapolation using a high-resolution discrete Fourier transform: *IEEE Trans. Signal Process.*, **46**, no. 1, 31–38.
- Seeger, A., C. Sogge, and E. Stein, 1991, Regularity properties of fourier integral operators: *Annals of Math*, **134**, 231–251.
- Starck, J. L., M. Elad, and D. Donoho, 2004, Redundant multiscale transforms and their application to morphological component separation: *Advances in Imaging and Electron Physics*, **132**.
- Trad, D., T. Ulrych, and M. Sacchi, 2003, Latest views of the sparse radon transform: *Geophysics*, **68**, 386–399.
- Treitel, S., J. L. Shanks, and C. W. Fraster, 1967, Some aspects of fan filtering: *Geophysics*, **32**, 789–800.
- Tropp, T., 2006, Just relax: convex programming methods for subset selection and sparse approximation: *IEEE Trans. Inform. Theory*. (To appear.).
- Verschuur, D. J. and A. J. Berkhout, 1997, Estimation of multiple scattering by iterative inversion, part II: practical aspects and examples: *Geophysics*, **62**, 1596–1611.
- Verschuur, D. J., A. J. Berkhout, and C. P. A. Wapenaar, 1992, Adaptive surface-related multiple elimination: *Geophysics*, **57**, 1166–1177.

Bibliography

- Wang, D., R. Saab, O. Yilmaz, and F. Herrmann, 2007, Recent results in curvelet-based primary-multiple separation: application to real data: Presented at the SEG International Exposition and 77th Annual Meeting.
- , 2008, Bayesian wavefield separation by transform-domain sparsity promotion: *Geophysics*. (Accepted for publication).
- Xu, S., Y. Zhang, D. Pham, and G. Lambare, 2005, Antileakage Fourier transform for seismic data regularization: *Geophysics*, **70**, no. 4, V87 – V95.
- Yarham, C., U. Boeniger, and F. Herrmann, 2006, Curvelet-based ground roll removal: Presented at the SEG International Exposition and 76th Annual Meeting.
- Yilmaz, O., 2001, *Seismic data analysis. Curves and Surfaces*. Society of Exploration Geophysicists.
- Ying, L., L. Demanet, and E. J. Candés, 2005, 3D discrete curvelet transform: *Wavelets XI, Expanded Abstracts*, 591413, SPIE.
- Zhang, R. and T. J. Ulrych, 2003, Physical wavelet frame denoising: *Geophysics*, **68**, 225–231.
- Zwartjes, P. M. and M. D. Sacchi, 2007, Fourier reconstruction of nonuniformly sampled, aliased seismic data: *Geophysics*, **72**, no. 1, V21–V32.

Appendix A

Examples

A.1 Oz Yilmaz 25 dataset

This is a classic example of a spatially aliased surface wave signal taken from Oz Yilmaz's text: *Seismic Data Analysis* Yilmaz (2001) and is displayed in figure A.1. This example provides the challenge of real data with its associated problems on which we can display the methods and techniques that we have developed. This data set also examines the use of a curvelet based prediction for the ground roll and compares it with standard F-K techniques.

A.1.1 Standard surface wave estimation

The generation of a prediction can be done in any number of ways. It is not limited to a single method and it may be advantageous to combine two or more methods to generate a prediction that covers a wide band of characteristics of the signal that is to be represented. When using the block solver method, it is more important to be conservative in the estimates of the surface wave to avoiding mixing the characteristics of the signals than it is to predict as much of the surface wave as possible. One of the benefits of using a sparse domain such as the curvelet domain is that the curvelet coefficients associated with the surface wave and the reflectors will have some common characteristics, such as scale, angle and location. By using conservative threshold methods, a prediction can be found using the curvelet domain for the separation scheme. As an iterative ℓ_1 minimization is solved through weighted thresholding in a sparse domain, the predicted coefficients will be used to define the weights, effectively identifying the associated coefficients with the particular wave fronts.

Standard F-K techniques such as band-pass filtering (Yilmaz (2001)), multichannel filtering (Galbraith and Wiggins (1968)), spectral balancing (Coruh and Costain (1983)) and other various F-K techniques (Embree et al. (1963); Treitel et al. (1967); Gelişli and Karsh (1988)) are advantageous as they allow a conservative prediction to be calculated. The conservative

prediction ensures that all reflector information will be excluded from the surface wave prediction. Used alone to remove surface waves, the risk to altering the frequency content of the reflector information can be high and difficulty may be faced in dealing with aliased information. If a conservative setting is used, for example, the upper limit of a high pass filter to a value significantly below where the reflector information is expected or used in combination with sparse domain techniques, the frequency content of the reflector information can be preserved.

Using predictions generated though other methods such as the Karhunen-Loeve (KL) Transform (Londono et al. (2005)), the Wiener-Levinson algorithm (Karsli and Bayrak (2004)), more standard wavelet based methods (Deighan and Watts (1997)), or any locally preferred method will also work either alone or in combination with other methods. Since the separation that is preformed by the block solver method is complete, no information will be lost if repeated iterations are required of the entire process.

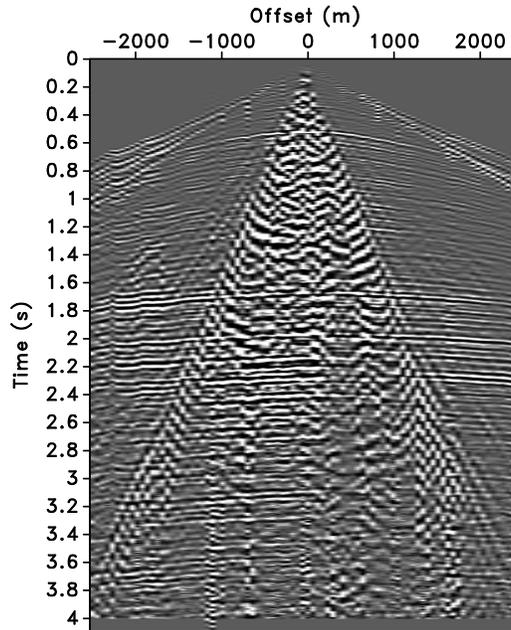


Figure A.1: Oz25 data set from Yilmaz (2001). This data set exhibits high amplitude, aliased ground roll with varying reflector amplitude underneath.

First, the data (shown in Figure A.1) is filtered using a classical F-K dip filter. This provides an estimate based on the velocities of the surface

wave and the reflectors. This can be seen in Figure A.2. This result is also the result of a standard subtraction. As we can see by the large amount of ground roll still present, this is not an acceptable solution. This may be slightly improved by applying more aggressive filter parameters but this will likely lead to increased reflector degradation and would not be an optimal choice for defining the weights used in the separation schemes.

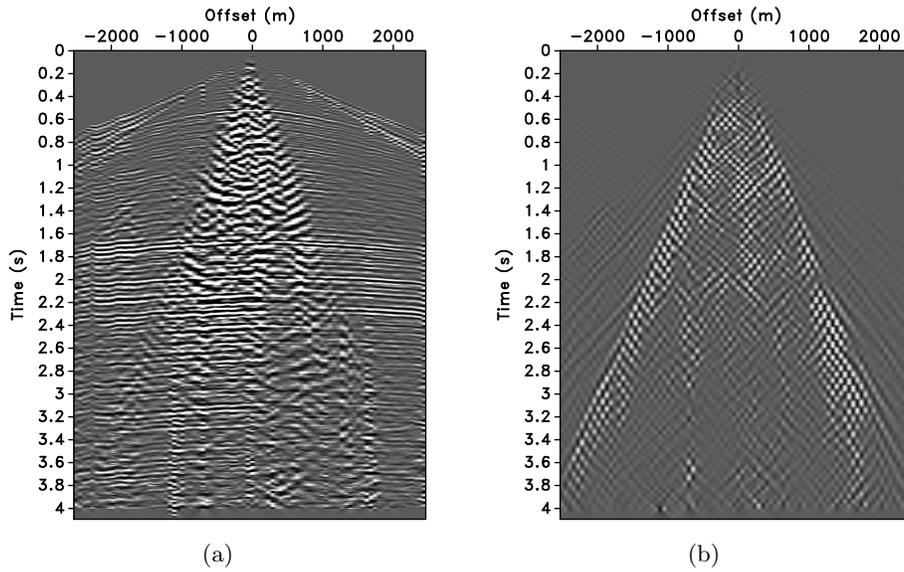


Figure A.2: Ground roll removal by F-K filtering. (a) A conservative estimate of the reflectors based on dip filtering in the Fourier domain. (b) The associated surface wave that is predicted during the F-K filtering process.

To improve the results, we can use these predictions of the surface wave and the reflectors with our separation methods to facilitate the signal separation. This technique takes advantage of the curvelet domain as a sparsity promoting transform. To conduct this separation, 3 inner and 3 outer loops were performed within the block-coordinate relaxation using a curvelet transform set to 6 scales with 16 angles at the second coarsest scale. A pad length of 200 was used and a noise level of 10^{-6} was assumed for the signal separation, which is roughly three orders of magnitude below our signal. The coefficients that define the threshold decay for the reflectors and the surface wave were both set to 0.2. We can see the results of this separation in figure A.3.

Figure A.4 highlights the effect of using the block-coordinate relaxation

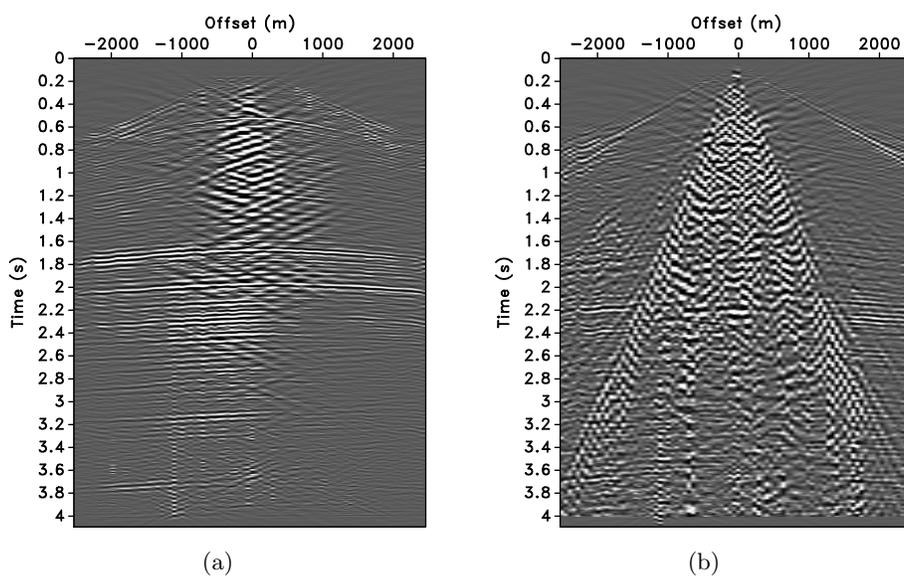


Figure A.3: Results of signal block-coordinate relaxation method. (a) The estimated reflectors. (b) The associated estimated surface wave. We can see that some reflector information has been moved to the surface wave estimate while the reflector estimate still contains ground roll. There is a component of noise centered at one second and zero offset that has not been identified as part of the coherent noise either by the angle filtering or by the signal separation scheme and other methods need to be examined to address this signal. A large amount of curvelet like artifacts have also been generated during the separation process.

separation algorithm as opposed to a simple subtraction by imaging at the difference between the two. Here, the extra information that has been assumed by the signal separation scheme to be part of the surface wave is shown. This signal is associated with the surface wave prediction and the algorithm removes it and makes it part of the reflector estimate. This extra information would have been associated with the reflectors by standard F-K filtering but is now with the associated reflectors in Figure A.3(a). We can see that some reflector information has been removed as well. There is also artifacts near the signal source that we wish to avoid.

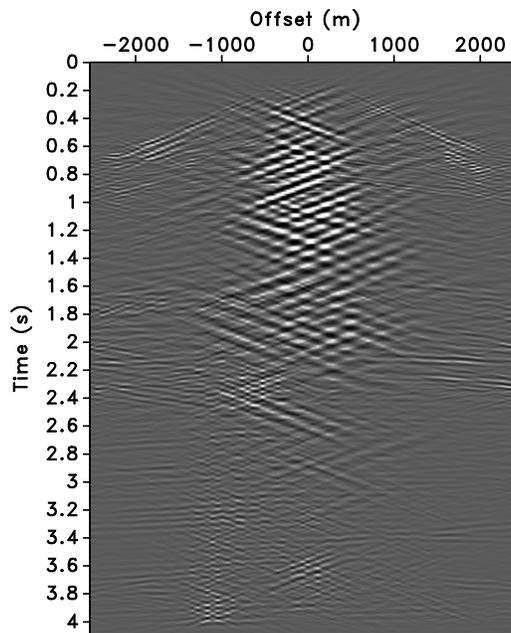


Figure A.4: The difference between the predicted surface wave and the block-coordinate relaxation estimated surface wave. We can see a large amount of ground roll that has been identified as part of the surface wave estimate. Some higher amplitude reflector information has also been identified as ground roll and erroneously moved to the surface wave estimate.

To compare the results, this prediction was also used to define the weights of the Bayesian separation method. For this separation, the parameters for defining the sparsity of the expected solution were set to 5 for the reflectors and .1 for the surface wave. The prediction confidence parameter η is set to .1 to reduce the effect of the reflector prediction on the surface wave

estimation. The results of the separation are shown in Figure A.5 with the effect of this separation process shown in Figure A.6.

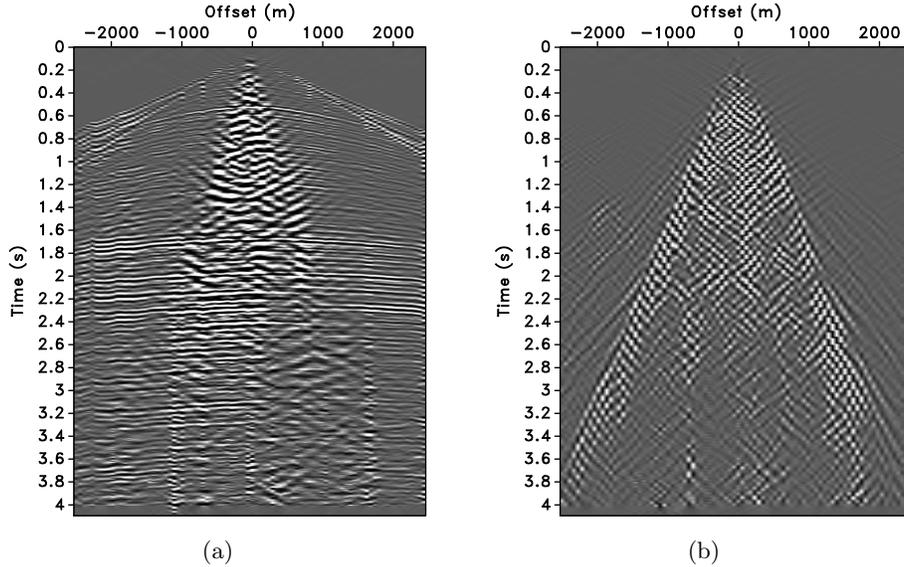


Figure A.5: Results of the Bayesian signal separation process. (a) The Bayesian estimated reflectors. (b) The associated estimated surface wave. In this case, very little to no reflector information has ended up in the estimated surface wave. Some ground roll still remains in the reflector estimate.

The difference between the results of these two methods using the same F-K surface wave prediction is shown in Figure A.7. As we can see, this is mostly reflector information that has been removed by the block-coordinate relaxation scheme. There is some ground roll difference but it is minor in comparison to the reflector information. This amount of reflector degradation caused by the block-coordinate relaxation method makes it essential to avoid. As we can see in Figure A.5(b), the Bayesian method has preserved the reflectors making this the method of choice.

A.1.2 Curvelet domain estimation

As we have seen, the curvelet domain is built by using two dimensional wave packets that vary in scale, angle and position. There is also an associated amplitude spectrum in the curvelet domain that relates to the amplitude variations seen in the transformed signal. Any differences in the characteristics of amplitude, angle and scale variations of the surface wave

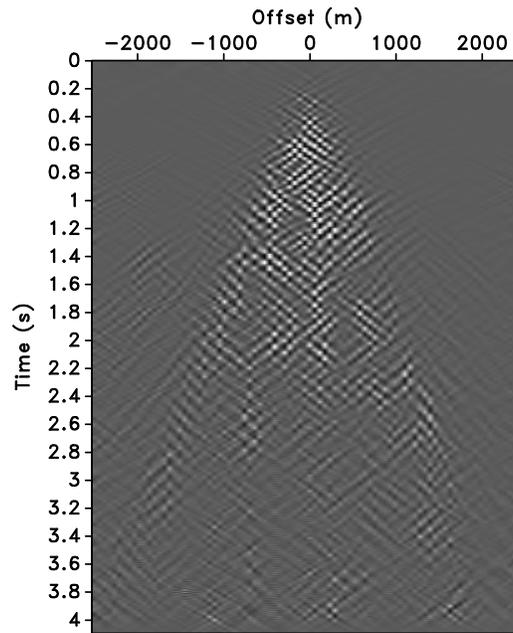


Figure A.6: The difference between the predicted surface wave and the Bayesian estimated surface wave. Again, we can see a large amount of ground roll that has been identified as part of the surface wave estimate. In this case, all reflector information has been preserved where the block-coordinate relaxation method removed some high amplitude reflectors.

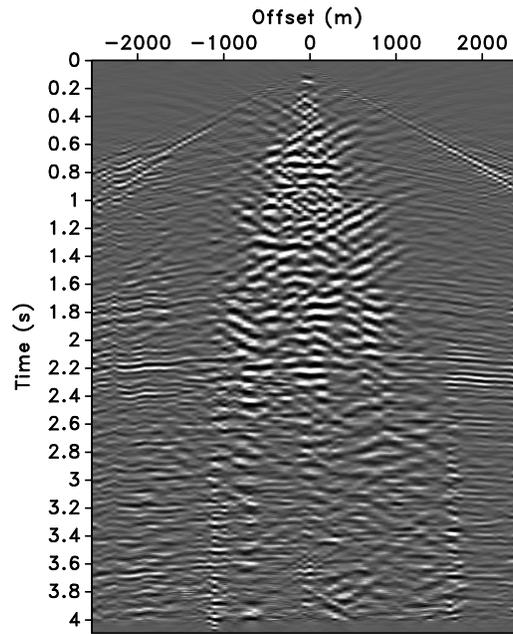


Figure A.7: Difference between the block-coordinate method relaxation and Bayesian separation method surface wave estimations. Again, we can see the reflector information removed by the block-coordinate relaxation method and introduced artifacts. There is also some minor difference in the surface wave estimations.

from the reflector signal can be identified in the curvelet domain. By applying a soft or hard thresholding operator to the signal in the curvelet domain, a simple prediction can be formulated for the reflectors and ground roll.

One useful method when dealing with reflector signals of lower amplitude and surface waves that have a high amplitude is to select a threshold defined by the root mean squared value of the individual scales of the curvelet domain. If we define a curvelet coefficient as $\mathbf{c}_{j,l,k}$ at scale j , angle l and position k , we can define the threshold t_c as,

$$\mathbf{t}_{\mathbf{c}_{j,l\dots L(j),k\dots K(j)}} = C_T \sqrt{\frac{1}{L(j)K(j)} \sum_{m=1}^{L(j)} \sum_{n=1}^{K(j)} \mathbf{c}_{j,m,n}^2}, \quad (\text{A.1})$$

with $L(j)$ and $K(j)$ representing the maximum number of scale dependent angles and locations respectively. This threshold also contains a constant C_T , which allows for the overall level of the threshold to be adjusted depending on the signal strength difference between the surface wave and the reflectors. For instance, if the surface wave and reflectors do not have a very large amplitude difference, this constant could be raised to ensure that there is no measure of reflector signal above the threshold. This introduces the flexibility needed to generate a conservative prediction that will only contain ground roll and no reflector information.

The generation of an effective prediction is a crucial step in the signal separation process. To generate a prediction by thresholding the curvelet vector as in Equation A.1 the C_T parameter needs to be set. Figure A.8(b) shows the effect of $C_T = 1$ applied to the original data. Figure A.9 shows a variety of C_T values. We can see that as the constant increases, less information is thresholded out. Examining Figure A.9(b) we can see that the first information to be thresholded out is the higher amplitudes, and most coherent wave fields. In this case, the ground roll is substantially reduced and the center reflectors have dimmed slightly. As C_T decreases, the rest of the signal is thresholded out.

To generate a second prediction of the surface wave and the reflectors in contrast to the the T-X domain prediction, a soft thresholding scheme is preformed in the curvelet domain. This takes advantage of the separation of scales in combination with the amplitude range of the the two signals. In this case, the threshold function is automatically set to a scaling of (4 times) the root mean squared value associated with each scale of the curvelet domain. On average, the reflector amplitudes are less than the surface wave and applying a soft threshold in the curvelet domain can be used to generate reflector and surface wave predictions. To supplement this, as some of

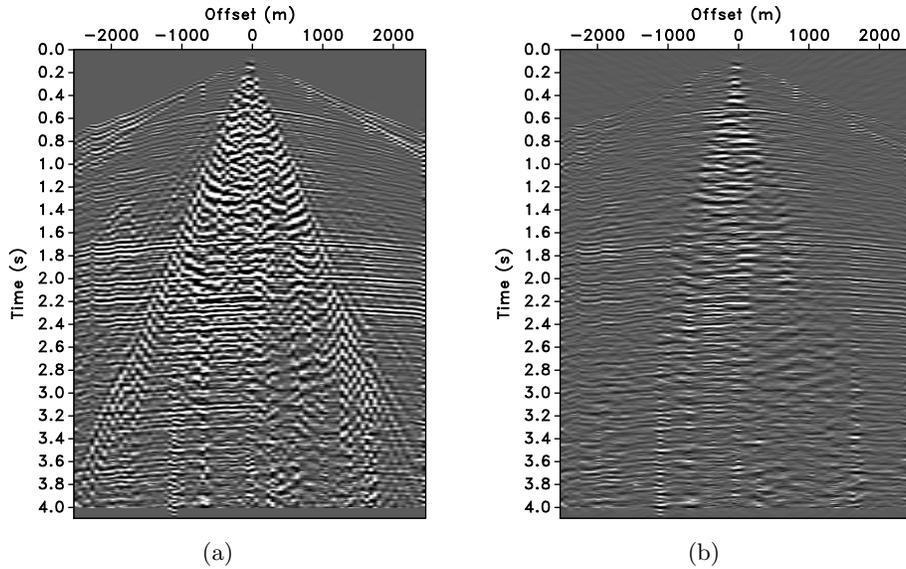


Figure A.8: The original data (a) and the data after a threshold defined by Equation A.1 has been applied with $C_T = 1$ (b).

the high amplitude reflectors in the center of the image are unintentionally thresholded, a F-K filter is applied to remove horizontal events from the surface wave and they are added to the reflector prediction. A second F-K filter is applied to the reflectors to remove any low amplitude ground roll that is then added to the surface wave prediction. Figure A.10 show the results of this process and the predictions for the signal separation algorithms. We can see that this is a more effective way to predict the complete surface wave and reveal the data that is underneath.

Using these predictions, again, we generate estimations for the reflectivity and the surface wave using the block-coordinate relaxation method. The parameter settings for both the block-coordinate relaxation and the Bayesian separation method are the set to the same values as the previous example with the F-K prediction. The results are shown in figure A.11. As before, we can also look at the difference between the curvelet defined prediction and the generated estimation, shown in Figure A.12.

Figure A.11 shows the results of the block-coordinate relaxation method applied to the data with the new curvelet thresholded prediction. The difference between the curvelet based prediction of the surface wave and the estimated surface wave is shown in Figure A.12. As we can see, the surface

Appendix A. Examples

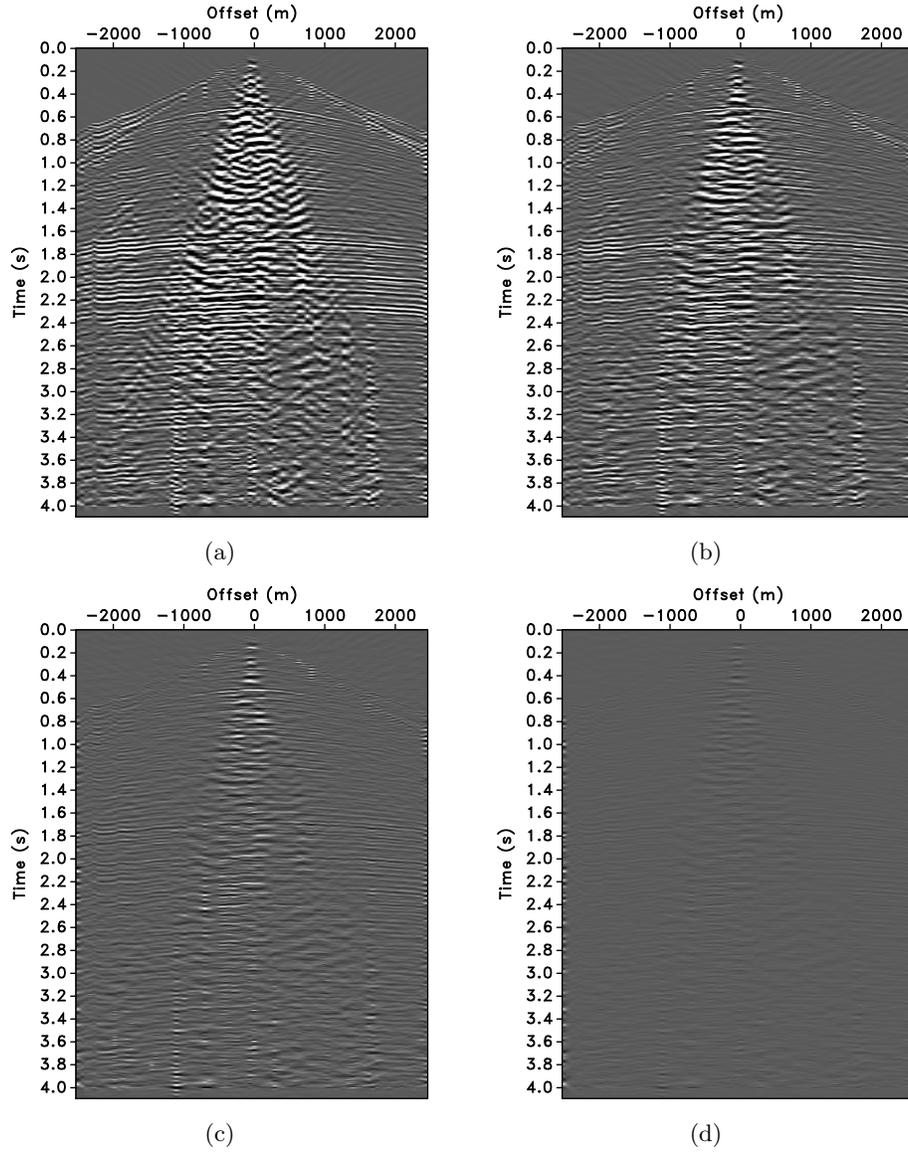


Figure A.9: Curvelet scale threshold results for varying values of C_T . The C_T constant is equal to (a) 10, (b) 2, (c) 0.5, (d) 0.1.

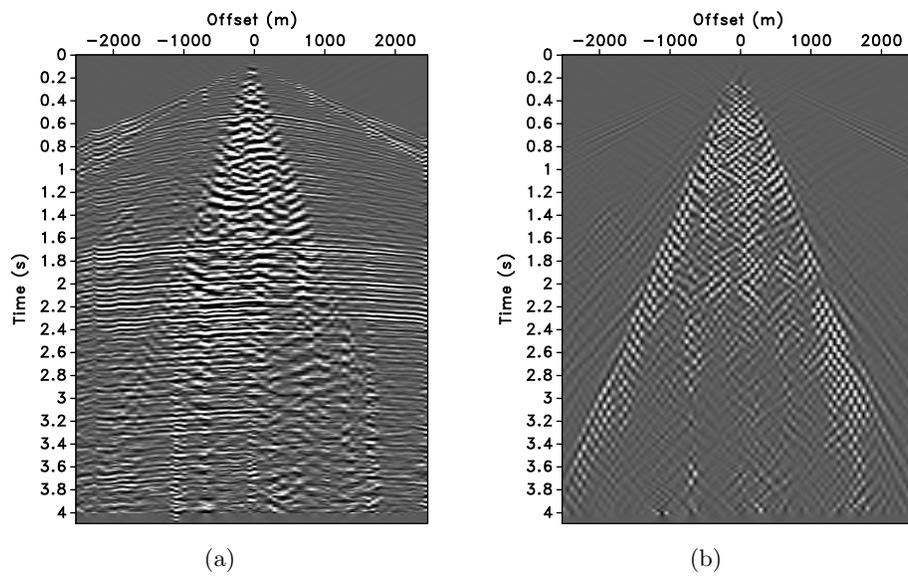


Figure A.10: A conservative estimate of the surface wave is created by performing a curvelet scale thresholding followed by application of a dip filter. We can see that the predicted reflectors (a) and the predicted surface wave (b) are not matched in amplitude but the prediction of the surface wave is enough to used as a prediction for the signal separation schemes.

wave has been removed but the reflectors have lost a lot of information and the curvelet like artifacts that have been introduced will cause a serious problem in previous processing steps.

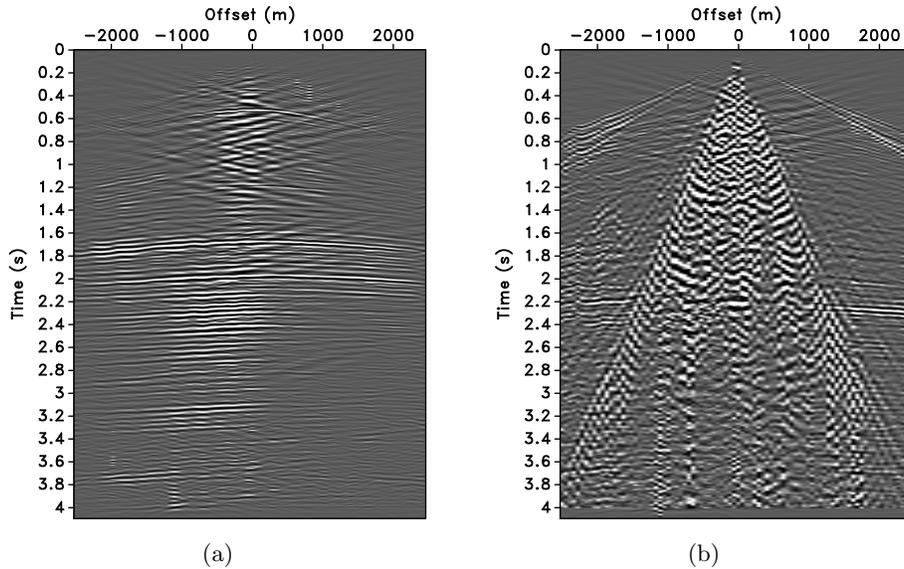


Figure A.11: The block-coordinate relaxation estimated reflectors (a) and estimated surface wave (b) generated using a curvelet scale thresholding prediction for the surface wave and reflectors. While the majority of the ground roll is removed, strong artifacts were generated by the signal separation scheme. As these artifacts are curvelet like, they begin to look like reflection events which is unacceptable.

This prediction was also used in conjunction with the Bayesian separation method with the estimated wave fields shown in Figure A.13. These results show a trade off with the block-coordinate relaxation method which removes much more information than is necessary. Here, the Bayesian method is more conservative, removing no reflector information but also leaving more ground roll energy. This method is the more valuable of the two as the reflector energy is protected and no artifacts have been left behind to contaminate the signal. Again, it is also advantageous to look at the difference between the separation scheme and the original curvelet estimate. This is shown in Figure A.14. As we can see, a significant amount of ground roll has been identified and moved to the estimate through the separation process. We should also note that no reflector information has been identified.

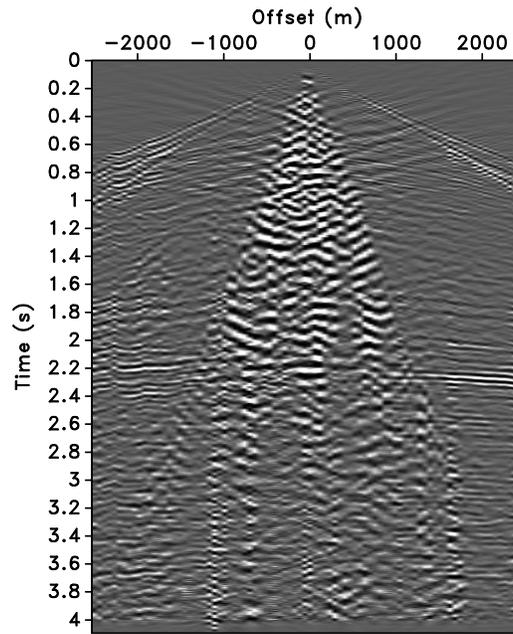


Figure A.12: The components of the surface wave that were not identified through the initial prediction but identified as part of the coherent noise signal through the block solver scheme. While all of the ground roll was removed, we see the generated artifacts and reflector information that has been removed from the data.

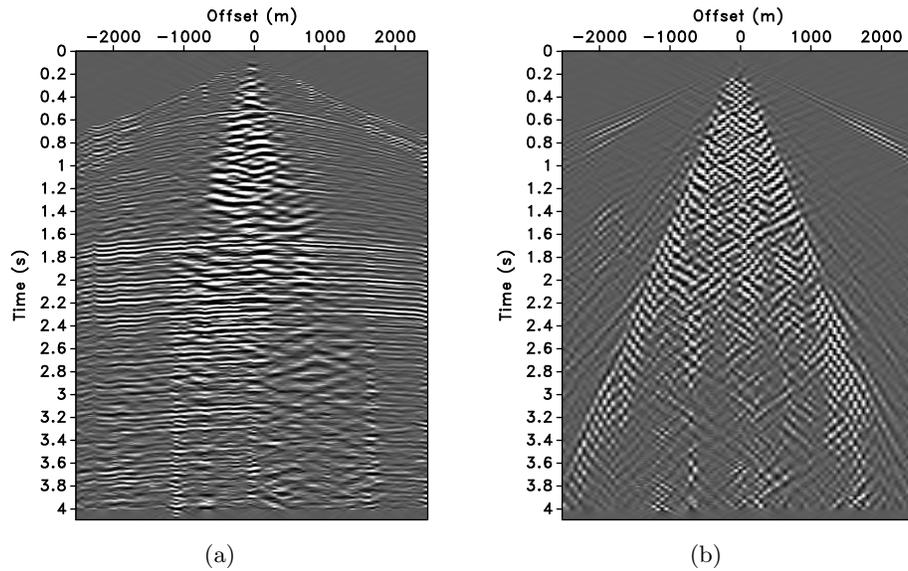


Figure A.13: The Bayesian estimated reflectors (a) and estimated surface wave (b) generated by a curvelet scale thresholding with dip filter prediction. While there is more ground roll present in the reflector estimation, we can see that no reflector information is present in the surface wave estimation. Very few, if any artifacts have been generated with this method, especially when compared to the block coordinate relaxation method.

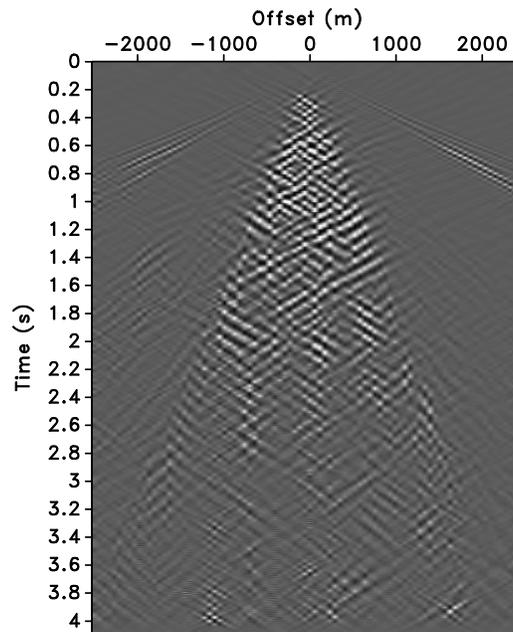


Figure A.14: The difference between the curvelets based prediction and the estimated surface wave generated by the Bayesian separation scheme. There are no reflectors present in this difference indicating that no reflectors were estimated as ground roll.

Appendix A. Examples

While the two separation methods have been compared, the methods of prediction have not. The difference between the separation results for the different predictions is shown in A.15. We can see that the primary difference between the predictive methods has been increased ground roll removal. The use of curvelet based predictions to estimate the surface wave has caused more surface wave to be predicted correctly. The curvelet like shape of the ground roll that has been removed with the curvelet based prediction method is due to the thresholding in the curvelet domain as the higher amplitude coefficients that are used to represent the ground roll have been identified regardless of their frequency content which may not have been identified by traditional F-K filtering.

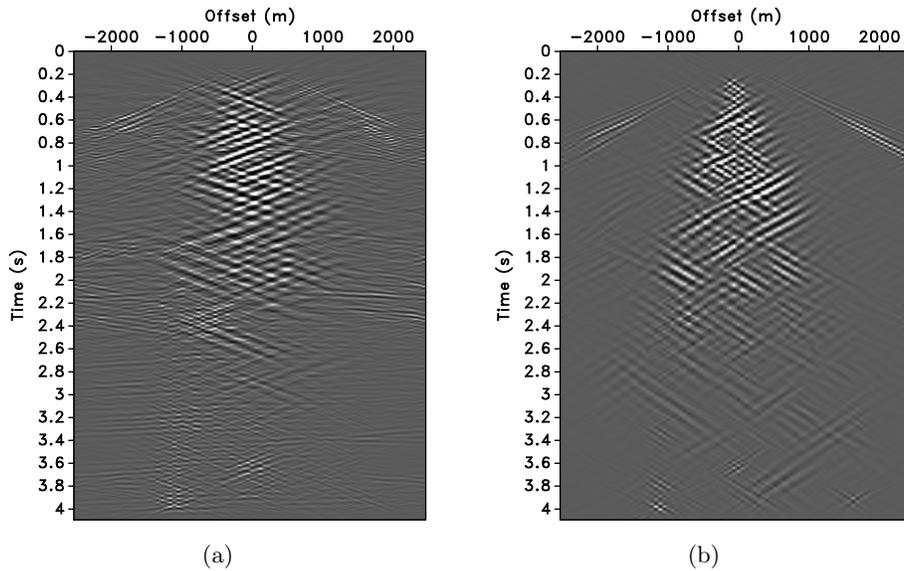


Figure A.15: The difference between the surface wave estimations of the two different prediction methods for the (a) block-coordinate relaxation and (b) Bayesian separation method. The shape of the removed data suggests that the larger curvelet coefficients that were not identified with the original f-k filtering technique have been removed by using a curvelet based thresholding scheme to predict the surface wave.

This data set highlights two important points. The first is the advantages and limitations of both methods. The block-coordinate relaxation method has a tendency to over threshold curvelet coefficients leading to artifacts, and reflector removal. The Bayesian separation method is more conserva-

tive, producing little or not artifacts but having a tendency to remove less ground roll from the reflector signal. The primary advantage of the Bayesian solver is its ability to preserve the reflector information from degradation. The second point is that the curvelet domain can be used to generate predictions for these separation methods. By applying a thresholding scheme to the curvelet vector, augmented with Fourier techniques, a more effective prediction can be produced than by simply using Fourier techniques alone. Direct thresholding of the curvelet vector to remove ground roll will have the same limitations as using Fourier techniques without signal separation but in combination with an effective separation method, these become viable results.

Bibliography

Coruh, C. and J. K. Costain, 1983, Noise attenuation by vibroseis whitening (vsw) processing: *Geophysics*, **48**, 543–554.

Deighan, A. and D. Watts, 1997, Ground-roll suppression using the wavelet transform: *Geophysics*, **62**, 1896–1903.

Embree, P., J. P. Burg, and M. M. Backus, 1963, Wide band velocity filtering-the pie-slice process: *Geophysics*, **28**, 948–974.

Galbraith, J. N. and R. A. Wiggins, 1968, Characteristics of optimum multichannel stacking filters: *Geophysics*, **33**, 36–48.

Gelişli, K. and H. Karşlı, 1988, F-k filtering using the hartley transform: *J. Seism. Explor.*, **7**, 101–108.

Karşlı, H. and Y. Bayrak, 2004, Using the Wiener-Levinson algorithm to suppress ground-roll: *Journal of Applied Geophysics*, **55**, 187–197.

Londono, E. G., L. C. Lòpez, and T. S. Kazmierczak, 2005, Using the karhunen-loève transform to suppress ground roll in seismic data: *Earth Sci. Res. J.*, **9**, 139–147.

Treitel, S., J. L. Shanks, and C. W. Fraster, 1967, Some aspects of fan filtering: *Geophysics*, **32**, 789–800.

Yilmaz, O., 2001, *Seismic data analysis. Curves and Surfaces*. Society of Exploration Geophysicists.