WAVELET APPROACHES TO SEISMIC DATA ANALYSIS

A Dissertation

Presented to

the Faculty of the Department of Mathematics

University of Houston

In Partial Fulfillment

of the Requirements for the Degree

Doctor of Philosophy

By

Qingqing Liao

December 2012

WAVELET APPROACHES TO SEISMIC DATA ANALYSIS

Qingqing Liao

APPROVED:

Dr. Donald J. Kouri, Chairman

Dr. Bernhard G. Bodmann

Dr. Manos Papadakis

Dr. Fuchun Gao
Total E&P Research and Technology USA

Dean, College of Natural Sciences and Mathematics

Acknowledgements

First and foremost, I must give my high, respectful gratitude to my advisor Dr. Donald J. Kouri for his continuous guidance, help, and encouragement throughout my study.

I would like to thank my committee members. Dr. Bernhard G. Bodmann gave me excellent feedback on my research over the years and pointed out importance references. Dr. Manors Parabasis helped to improve my dissertation. Dr. Fuchun Gao shared his knowledge on seismic data with me and encouraged me to think creatively.

I also want to thank my husband, Dr. Lue Du, for his support, love, and care during my preparation of this dissertation. I am grateful to my family for their love and support.

I extend my appreciation to my group members and fellow graduate students for their discussions and friendship.

I thank Ms. Pamela Hensley from UH International Friendship Program for her friendship which added a lot of fun to my study experience.

I appreciate the financial support from Petroleum Geo-Services (PGS) and Total E&P. I thank Total E&P for the opportunity of internship and its approval of including my internship research in this dissertation.

WAVELET APPROACHES TO SEISMIC DATA ANALYSIS

An Abstract of a Dissertation
Presented to
the Faculty of the Department of Mathematics
University of Houston

In Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy

By
Qingqing Liao
December 2012

Abstract

There have been extensive applications of wavelets to petroleum seismic data. In this dissertation, we focus on developing and testing new wavelets approaches to seismic data compression, microseismic first arrival picking, seismic event picking, and seismic reflectivity inversion.

First, we developed new methodologies for seismic data compression based on wavelets. We started with applying matching pursuit to obtain a sparse representation of seismic signals on a dictionary, so we only need to store and transmit the sparse representation. The dictionaries tested initially are Symlets. To improve the performance of compression further, we proposed the new idea of using subspace matching pursuit to obtain perfect reconstruction for a phase-rotated signal. We obtained better fidelity than matching pursuit, but the convergence is slowed down due to the incompleteness of the dictionary. Finally we proposed using matching pursuit with a combination of Symlets dictionary and subspace dictionary, thereby obtaining the best quality with the same compression ratio.

Second, we report a new method of automatic first break detection of P-waves and S-waves. Our method is based on a time-frequency analysis of the seismic trace using minimum uncertainty (μ-)wavelets, in particular in the minimum-phase form. We have tested our method on both lab data with various signal-to-noise ratio (S/N or SNR) and on field data.

Third, we explored methods of automatic seismic event picking. It is known that no single automatic seismic event indicator works for all data; therefore, we explored

two indicators based on the μ-wavelets and on an energy ratio. Thresholding was applied to pick seismic events. We have tested the methods with both synthetic data and offshore field data.

Finally, we proposed new seismic sparse inversion methods based on complex basis pursuit (CBP) and a modified complex basis pursuit (MCBP). In practice, constant phase wavelets are used for seismic inversion algorithms, for example, the basis pursuit (BP). If the phase of the estimated wavelet is wrong, this will surely cause an error in reflectivity. We can obtain more accurate reflectivity even though the estimated wavelet has biased phase by using CBP and MCBP rather than BP.

Contents

List of Figures

List of Tables

CHAPTER 1

Introduction

1.1 Background

There have been extensive applications of wavelets to petroleum seismic data [25, 27, 28, 26, 14]. Wavelet transforms were applied for downward continuation [18], shear-wave discrimination [31], and seismic data processing [24]. Dessing and Wapenaar used wavelets for wavefield extrapolation [9] and seismic migration with one-way operators in the wavelet transform domain [10]. Wang and Pann developed compressed Kirchhoff migration based on wavelets [49]. Li and Ulrych studied tomography via wavelet transform constraints [19] and Hilbert attribute analysis [20]. Wavelets have also been applied to the interpretational aspects [13], such as amplitude variation with offset [52], thin-bed analysis [59], direct hydrocarbon detection [41], and SPICE attribute [40].

In this dissertation, we focus on wavelets approaches to seismic data compression, microseismic first arrival picking, seismic event picking, and seismic reflectivity inversion.

1.2 Seismic data compression

In current 3-D seismic acquisitions, data volumes of a few terabytes are common. As wide azimuth and even rich azimuth acquisition become the new standards, seismic volumes are becoming even larger. An efficient seismic data compression method can speed up handling large data volumes and possibly speed up the imaging algorithm.

Generally, there are four types of data compression techniques: basic methods (e.g. scalar quantization, run length encoding), statistical methods (e.g. Huffman coding, arithmetic coding), transform methods (e.g. wavelet, Fourier transform), and dictionary based methods (e.g. GZIP, PNG)[39]. Among these, there are two groups: lossless and lossy techniques. If we apply lossless methods on SEG-Y format files, the compression is not significant. This is because there is hardly any absolute repetition of the entire 32-bit vector even if millions of different data traces are analyzed [51].

To obtain higher compression performance, we focus on wavelet-based methods which are lossy in Chapter 2. Wavelets have been applied as a popular tool for seismic data processing. An important goal of wavelet-based compression is to represent seismic data in a sparse way. Once we have a sparse representation of seismic data, we can store a relatively small number of coefficients and indices as compressed data. We can transmit the compressed data and reconstruct the full data for later processing. A sparser representation with good fidelity is preferred.

We applied matching pursuit [22] to compute the sparse representation of seismic signals on a dictionary. Matching pursuit is a greedy algorithm that iteratively selects a vector in a complete and redundant dictionary to obtain a sub-optimal approximation to a signal. The matching pursuit algorithm has been applied to seismic spectral decomposition [2], and it provides excellent spectral localization: the reflection, direct and surface waves, and artifact energy are clearly identifiable. In [35], the authors presented a new method to regularize irregularly sampled seismic data, based on a matching pursuit technique.

The dictionaries we have initially tested are Symlets. Symlets are a modified version of Daubechies wavelets with increased symmetry. We obtained good quality ($SNR > 40dB$) for a compression ratio of 10%, but obvious oscillating artifacts (mosaic phenomenon) appear at low level of compression.

To improve the performance of compression further, we explored ideas to leverage source wavelets or estimated source wavelets. We found that if we use the dictionary based on source wavelets, the results suffer from phase-rotation. Then we proposed a new idea of using subspace matching pursuit to obtain almost exact reconstructions for a phase-rotated signal. We have tested the subspace matching pursuit on a synthetic shot, and the results show better fidelity than matching pursuit with Symlets when the compression ratio is less than 3.3%. But its convergence slowed down due to the incompleteness of the dictionary.

Finally, we proposed using matching pursuit based on the combined Symlets and subspace dictionaries. We obtained the best quality with the same compression ratio. Only limited artifacts were observed for various compression ratios.

1.3 Microseismic first arrival picking

In refraction seismology, first break (or arrival time) detection has been applied to study the near surface low-velocity zone and determine static corrections. In recent years, with the advances in hydraulic fracturing techniques, first break detection of P-waves and S-waves has become crucial for locating microseismic events. Most of the first arrival picking methods are applied in the time domain. Examples include the multi-window algorithm [5] and the short-term average/long-term average method [54].

In Chapter 3, we proposed a first break detection method which is based on a time-frequency approach. It has been observed that the high-frequency component of microseismic traces increased dramatically at the onset of a microseismic event [42]. Taking into account the analysis of the frequency characteristics of the seismic trace, a continuous wavelet analysis with optimal temporal and frequency localization may yield accurate results and a method that can be automated easily. We propose to employ a minimum uncertainty (μ-) wavelets basis to perform spectral decomposition, since the μ-wavelets result from a constrained minimization of the

Heisenberg uncertainty product [15] and consequently can improve the temporal and frequency resolution [29]. Often, when the trace has been decomposed using zero-phase wavelets, the signal will be noncausal in the sense that a small amount of energy leaks into times before the true first arrival. Ricker (1953) noted this in the predigital analysis era. We circumvent this by transforming the original μ-wavelets (non-causal) into minimum-phase wavelets, which are causal.

The reference [36] took the maximum of the cross-correlation of the traces with a "model" trace as the indicator of the shift of the traveltime. This was based on the assumption that the waveforms in each trace are reasonably similar to those of the "model" trace which is obtained by hand-aligning two chosen traces and summing them. In a time domain approach, one could take the μ-wavelets as the "model" traces, and cross-correlate them with each trace and take the maximum as the indicator of the time shift. Since the μ-wavelet is initialized to time zero, the time of the maximum of the cross-correlation corresponds to the first arrival time. Unfortunately this doesn't work well, as our tests show. An alternative to a purely time-domain approach is to use time-frequency analysis.

In Chapter 3, we introduced a time-frequency decomposition of the signal using μ-wavelets. We employed the maximum of the power mean, computed from the time-frequency record as our first arrival indicator function. We applied a peak-picking algorithm to find the largest two peaks which are the indicators for the P-wave and S-wave first arrivals. We specified a potential region of the first arrivals for each of P-wave and S-wave. In each potential region, we neglected every other points and then applied Hermite distributed approximating functionals (HDAFs) to fill in the neglected points [56]. This yielded a less noisy indicator function at the potential region. Next we obtained the maximum of all the points in each of the potential arrival region as our indicator. Thus we obtained two indicators, the smaller one for the P-wave first arrival and the other one for the S-wave arrival. We have tested our method on both lab data with various signal-to-noise ratio (S/N or SNR) and on field data. Our method picks the first arrival time for lab data contaminated with both high and low S/N with an accuracy of 0.5 μs. We also compare our automatic detection of the first break in field data with manual detection. The difference is less than 1 ms. Our results indicate that our method is robust for automatically detecting the first arrival time for field data.

1.4 Seismic event picking

In seismic traveltime tomography, the automatic or manual picking of seismic events are used to give a correct velocity model. In this tomographic velocity model updating process, the manual picking of prestack events is a primary bottleneck. Besides, Laplace domain waveform inversion requires muting prestack seismic data before the first arrival, protecting refracted waves. We explored methods of automatic seismic event picking in Chapter 4.

It is known that no single automatic seismic event indicator works for all data; therefore, we explored two indicators based on μ-wavelets and an energy ratio method. Thresholding was applied to pick seismic events. We have tested with synthetic data and offshore field data.

We first applied a μ-wavelets based seismic event indicator as we introduced in Chapter 4. We applied the indicator on each trace and picked the first arrivals on synthetic shots of BP2004. We obtained good picking results but each shot took 2069s on a work station. This efficiency is not satisfactory and is due to the computational complexity of the spectral decomposition. We next explored an energy ratio method, which is less computationally complex.

We computed short-term average/long-term average (STA/LTA) ratios trace by trace and then applied thresholding to mute. We have added various levels of white noise to the synthetic shots and then picked seismic events using the energy ratio method. We obtained good results, as demonstrated by successfully muting the noise before the first arrival for the synthetic shots with various noise levels. With extremely high levels of noise (e.g SNR$\leq$2dB), the indicator does yield errors, as expected. We have tested the method on field data from Total E&P, but due to a confidentiality constraint we can not show the results.

1.5 Seismic reflectivity inversion

Seismic reflectivity inversion is one of the important digital signal processing methods in geophysical exploration. Traditional seismic deconvolution uses a wavelet inverse filter to yield a bandlimited reflectivity. Sparse seismic inversion methods can produce a significant increase in bandwidth content from band-limited seismic observations (e.g. [37]). This has become more important because the main task in today's seismic exploration is seeking to locate subtle hydrocarbon traps. Assume the seismic trace $s(t)$ is a simple convolution of a stationary seismic wavelet $w(t)$ and the reflectivity $r(t)$ with additive noise $n(t)$.

$$s(t) = w(t) * r(t) + n(t) \tag{1.1}$$

where t is the two-way traveltime and $*$ means convolution. This is the so called convolutional model and it assumes that a set of planar layers of constant impedance can sufficiently represent the earth structure. Seismic reflectivity inversion aims to obtain the reflectivity $r(t)$ given the data (measurements) $s(t)$, and with assumption of seismic wavelet $w(t)$. It is an ill-posed mathematical inversion problem for which the solution (reflectivity $r(t)$) is not unique, because the seismic wavelet is band limited and seismic data are finite and inaccurate. Among all of the possible solutions (reflectivity) that fit the data, constraints and tolerance are specified to pick the "optimal" reflectivity. It is widely known that applying valid constraints in seismic reflectivity inversion can produce a higher bandwidth of the solution [44]. Besides, it is a common practice to assume some prior information about the solution. For example, sparse-spike deconvolution supposes that the reflectivity consists of a sparse sequence of spikes as a prior [48]. With the assumption of sparseness, the seismic inversion objective is to determine the location of the reflection coefficients and their amplitudes. Various methods were proposed to optimize some norm that forces the solution to be sparse ([34],[38],[50], and [58]).

It is well known that the success of the seismic reflectivity inversion depends on the quality of the estimated wavelet $w(t)$ [34]. There are several methods to derive the estimated wavelet (e.g. [47]), but the quality varies with data. For simplicity, a constant phase wavelet is used for seismic inversion [57, 48]. If the phase of the estimated wavelet was wrong, this will surely cause an error in reflectivity [58, 48]. In this dissertation, I proposed new sparse inversion methods that do not require that the phase of the estimated wavelet is the same as the seismic data. We constructed a complex dictionary of the estimated wavelets and performed a complex basis pursuit to decompose the complex seismic traces to corresponding real and imaginary reflectivities. The complex dictionary consisted of the estimated wavelet with various shifts as the real part and the corresponding Hilbert transform of the wavelets as the imaginary part. In this way, we can obtain more accurate reflectivity even though the estimated wavelet has a biased phase. I also proposed a modified complex basis pursuit method to invert the real seismic trace (not the complex seismic trace) for reflectivity. In this case, we minimized the real part of the least square error instead of the complex least square error. We obtained results which are visually the same as the fully complex basis pursuit.

1.6 Outline of dissertation

In Chapter 2, we describe two new wavelet-based methods on seismic data compression. Test results for a synthetic shot are discussed.

In Chapter 3, we describe a new automatic first break detector for microseismic data using μ-wavelets. We have tested the approach on lab data and field microseismic data to pick both P-wave and S-wave arrivals.

In Chapter 4, we describe methods of automatic seismic event picking based on wavelets and energy ratio.

In Chapter 5, we describe new seismic sparse inversion methods based on complex basis pursuit and a modified basis pursuit.

In Chapter 6, we present our conclusions and discuss future work.

In the Appendix A, we derive the elastic wave equation in a pre-stressed medium. The pre-stress causes anisotropy of the elastic tensor and wave velocities. We obtain the Green-Christoffel equation by considering the harmonic plane wave solution. Also we analyze experimental data to obtain the elastic tensor under a uni-axial pre-stress.

CHAPTER 2

Wavelet-based seismic data compression

We describe two new methodologies for seismic data compression based on wavelets in this chapter. We started with applying matching pursuit to obtain a sparse representation of seismic signals on a dictionary, so we only need to store and transmit the sparse representation instead of the original signals. The dictionary tested initially is composed of Symlets. To improve the performance of compression further, we explore ideas to leverage source wavelets or estimated source wavelets. We propose a new idea of using subspace matching pursuit to obtain perfect reconstruction for phase-rotated signals. We have tested this subspace matching pursuit on a synthetic shot, and the results show better fidelity than matching pursuit for compression ratios less than 3.3%, but the convergence slowed down due to the incompleteness of the dictionary. Finally we propose using matching pursuit with a combination of the Symlets dictionary and subspace dictionary. We obtain the best quality result with the same compression ratio with this approach.

2.1 Matching pursuit with Symlets

In this section, we introduce a matching pursuit algorithm with Symlets to compress seismic data. Symlets are a modified version of Daubechies wavelets with increased symmetry. Matching pursuit is a greedy algorithm that iteratively selects a vector in a complete and redundant dictionary to obtain a sub-optimal approximation to a signal. We generate the synthetic shot data by forward modeling on a three layer velocity model to test the compression performance.

2.1.1 Algorithm

Let Γ be a index set and $\mathcal{D} = \{\omega_p\}_{p\in\Gamma}$ be a dictionary of P unit norm vectors $\|\omega_p\| = 1$ in a signal space $\mathbb{C}^N$. Finding an optimal M-term approximation $f_M \in \mathcal{D}$ of a signal $f \in \mathbb{C}^N$ with M vectors selected in a redundant dictionary $\mathcal{D}$ is NP-hard [21]. Thus it is necessary to rely on suboptimal approximations by computational algorithms. Several algorithms were investigated in [21]. A best-basis algorithm selects the orthogonal vectors in the basis but the rigidity of orthogonality limits the approximation. Matching pursuit has the freedom to incorporate more patterns using large and non-orthogonal dictionary. Matching pursuit iteratively obtains one vector from a redundant dictionary at a time [22]. Matching pursuit is nonlinear and it maintains an energy conservation which ensures its convergence. It is closely related to projection pursuit strategies used in statistics [11]. The two algorithms were developed independently in a very different context but the share similar underlying mathematics. The matching pursuit algorithm proceeds as follows:

Let $R^0 f = f$. Suppose that the mth-order residue $R^m f$ is already computed for $m \geq 0$. The next iteration chooses $\omega_{p_m} \in \mathcal{D}$ such that

$$| < R^m f, \omega_{p_0} > | = \max_{p \in \Gamma} | < R^m f, \omega_p > |. \tag{2.1}$$

and projects $R^m f$ on ω_{p_m}:

$$R^m f =< R^m f, \omega_{p_m} > \omega_{p_m} + R^{m+1} f. \tag{2.2}$$

The orthogonality of $R^m f$ and ω_{p_m} implies

$$\|R^m f\|^2 = | < R^m f, \omega_{p_m} > |^2 + \|R^{m+1} f\|^2, \tag{2.3}$$

which can be interpreted as the conservation of energy. Summing (2.2) on m between 0 and $M - 1$ yields

$$f = \sum_{m=0}^{M-1} < R^m f, \omega_{p_m} > \omega_{p_m} + R^M f. \tag{2.4}$$

Analogously, summing (2.3) on m between 0 and $M - 1$ yields

$$\|f\|^2 = \sum_{m=0}^{M-1} | < R^m f, \omega_{p_m} > |^2 + \|R^M f\|^2. \tag{2.5}$$

Next, we show that if the residual $\|R^m f\|$ has a minimum rate of decay, the matching pursuit has an exponential decay [22]. The decay of $\|R^m f\|$ depends on the correlation between the residues and the dictionary elements. Let $\mu(r, \mathcal{D})$ be the coherence of a vector r relative to the dictionary $\mathcal{D}$ and it is defined as

$$\mu(r, \mathcal{D}) = \max_{p \in \Gamma} | < \frac{r}{\|r\|}, \omega_p > | \leq 1. \tag{2.6}$$

Lemma 2.1.1. *Let $\mathcal{D}$ be a complete dictionary in a finite dimensional space $\mathbb{C}^N$,*

$$\mu_{min}(\mathcal{D}) = \inf_{r \in \mathbb{C}^N, r \neq 0} \mu(r, \mathcal{D}) > 0. \tag{2.7}$$

Proof. We prove the lemma by contradiction. Suppose $\mu_{min}(\mathcal{D}) = 0$. There exist $\{f_m\}_{m \in \mathbb{N}}$ with $\|f_m\| = 1$ such that

$$\lim_{m \to} \sup_{p \in \Gamma} | < f_m, \omega_p > | = 0. \tag{2.8}$$

Since the unit sphere of $\mathbb{C}^N$ is compact, there exists a subsequence $\{f_{m_k}\}_{k \in \mathbb{N}}$ that converges to a unit vector $f \in \mathbb{C}^N$. Then we have

$$\sup_{p \in \Gamma} | < f, \omega_p > | = \lim_{m \to +\infty} \sup_{p \in \Gamma} | < f_{m_k}, \omega_p > | = 0. \tag{2.9}$$

So $< f, \omega_p >= 0$ for all $\omega_p \in \mathcal{D}$. Since $\mathcal{D}$ contains a basis of $\mathbb{C}^N$, necessarily $f = 0$, which is contradict with $\|f\| = 1$. Therefore,

$$\mu_{min}(\mathcal{D}) > 0. \tag{2.10}$$

Theorem 2.1.2. *Let $\mathcal{D}$ be a complete dictionary in a finite dimensional space $\mathbb{C}^N$. The residual $R^m f$ computed by a matching pursuit satisfies*

$$\|R^m f\|^2 \leq (1 - \mu_{min}(\mathcal{D}))^m \|f\|^2, \tag{2.11}$$

where $1 \geq \mu_{min}(\mathcal{D}) > 0$. *As a consequence,*

$$f = \sum_{m=0}^{+\infty} < R^m f, \omega_{p_m} > \omega_{p_m}, \tag{2.12}$$

and

$$\|f\|^2 = \sum_{m=0}^{+\infty} | < R^m f, \omega_{p_m} > |^2. \tag{2.13}$$

Proof. The energy conservation equation 2.3 implies

$$\frac{\|R^{m+1} f\|^2}{\|R^m f\|^2} = 1 - | < \frac{R^m f}{\|R^m f\|^2}, \omega_{p_m} > |^2 \leq 1 - \mu^2(R^m f, \mathcal{D}) \tag{2.14}$$

We iterate on this equation and proves that

$$\|R^m f\|^2 \leq (1 - \mu_{min}(\mathcal{D}))^m \|f\|^2. \tag{2.15}$$

By Lemma 2.1.1, we have

$$1 - \mu^2_{min}(\mathcal{D}) < 1, \tag{2.16}$$

and it follows that

$$\lim_{m \to +\infty} \|R^m f\| = 0. \tag{2.17}$$

Inserting this into equation 2.4 and equation 2.5 proves that equation 2.12 and equation 2.13 hold respectively.

In the limit of infinite dimensional space, Jones' theorem shows that the matching pursuit still converges but the convergence is not exponential [16].

We have the algorithm written in Algorithm 1 in the form of pseudocode.

For seismic shot gathers, we applied matching pursuit trace by trace to obtain coefficients and indices on the dictionary. We first used well-known discrete wavelets

Algorithm 1 Matching pursuit

Input: signal(column vector of length N),dictionary(matrix of N by K),threshold, maxiteration
Output: coefficient (column vector of length Niter), index (column vector of length Niter)

```
Initialization: residue ← signal , m ← 0
while m ≤ maxiteration and norm(residue) ≤ threshold do
    m ← m + 1
    for i = 1 → K do
        innerprod(i) ← dot-product(dictionary(:, i), residue)
    end for
    index(m) ← the index of max(abs(innerprod))
    coefficient(m) ← innerprod(index(m))
    residue ← residue - coefficient(m)*dictionary(:,index(m))
end while
Niter ← m
```

to form the dictionary. We use Daubechies's least-asymmetric wavelets symJ (J is the number of vanishing moments) [8] as one example for demonstration. Symlets are a modified version of Daubechies wavelets with increased symmetry. SymJ have a minimum support $[-J+1, J]$ with J vanishing moments. So each symJ has $2J$ samples which have been stored in a separate data file, for example, sym4.mat in the MATLAB wavelet toolbox.

We generate a dictionary of Symlets using inverse discrete wavelet transform. We illustrate the procedure by generating a dictionary of Sym4 at level 2 for a signal of $N = 64$ samples. We set the approximation and detail coefficients in the way that they consist a unit diagonal matrix of N by N as shown in Figure 2.1. Then we apply 1-D inverse discrete wavelet transform using Sym4 column by column to obtain a dictionary as shown in Figure 2.2. Each column is a wavelet in the dictionary.

The 1-st to $\frac{N}{4}$-th components are plotted using different colors in Figure 2.3a. The $\frac{N}{4}+1$-th to $\frac{N}{2}+1$-th components are plotted in Figure 2.3b. The $\frac{N}{2}+1$-th to N-th components are plotted in Figure 2.3c. To look at one of the component closely, we plotted three components in Figure 2.4. The 3-rd component is in black as shown in Figure 2.4 and it has support of $4J$. The 25-th component in blue as shown in Figure 2.4 and it has support of $4J$. The 59-th component in red as shown in Figure 2.4 and it has support of $2J$.

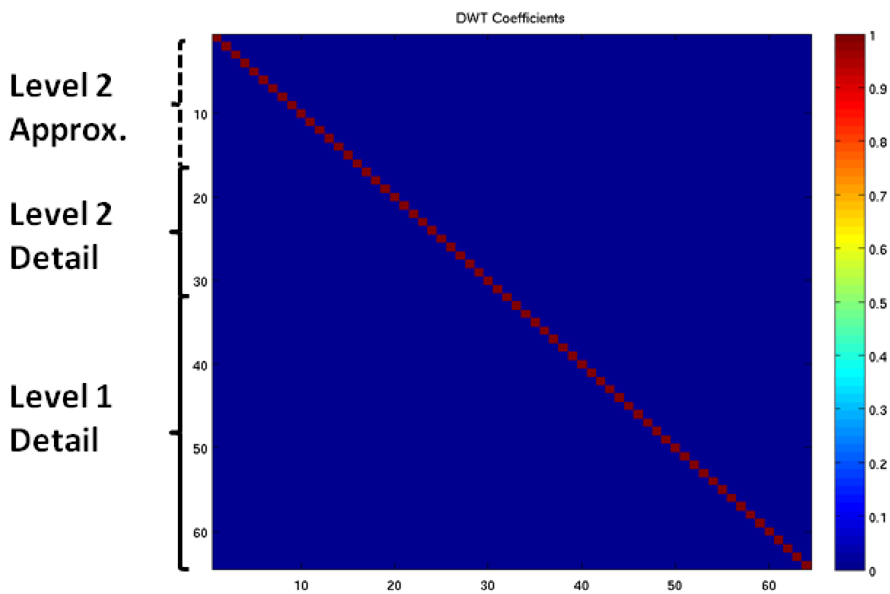


Figure 2.1: The approximation and detail coefficients.

Next we applied the matching pursuit algorithm on a synthetic trace. We generated a synthetic trace shown in Figure 2.5a by superposition of the 22-nd and the 58-th components of the dictionary.

We applied the matching pursuit algorithm using the previously generated dictionary. The algorithm stopped after two iterations and the coefficients and indices

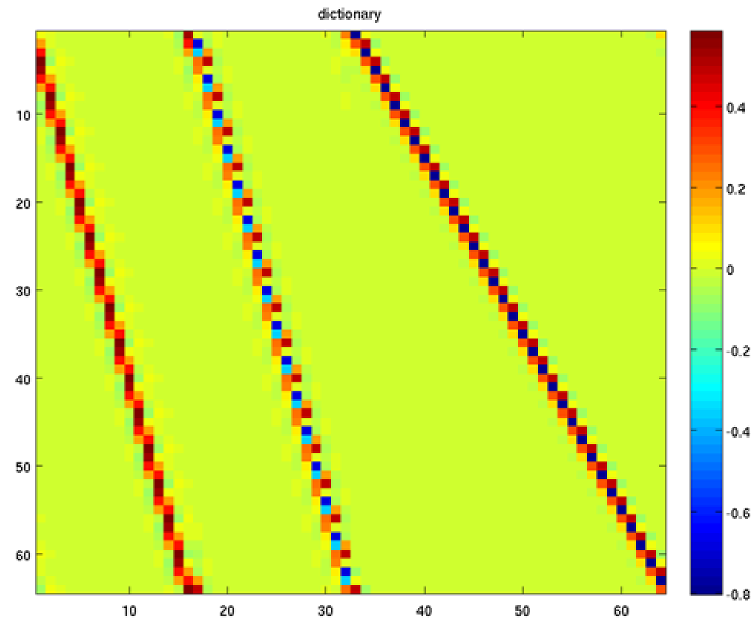
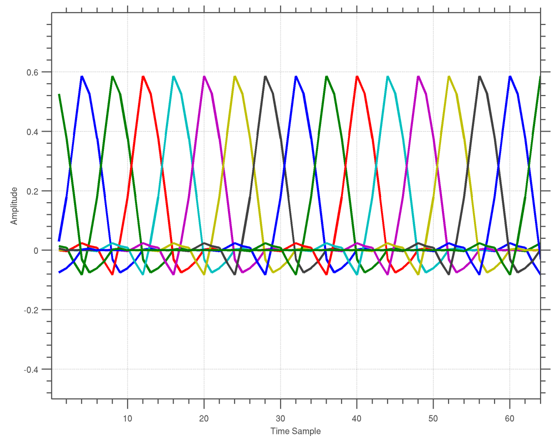
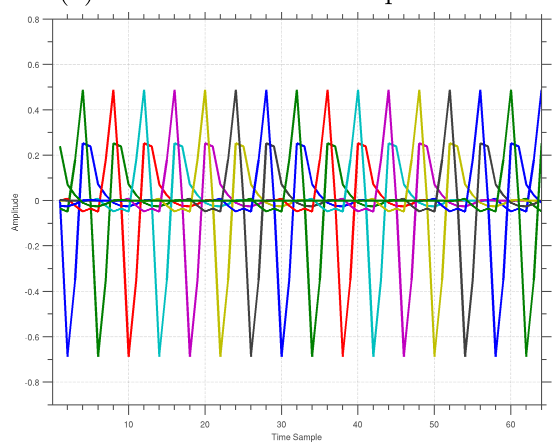


Figure 2.2: The dictionary generated by inverse discrete wavelet transform using Sym4.

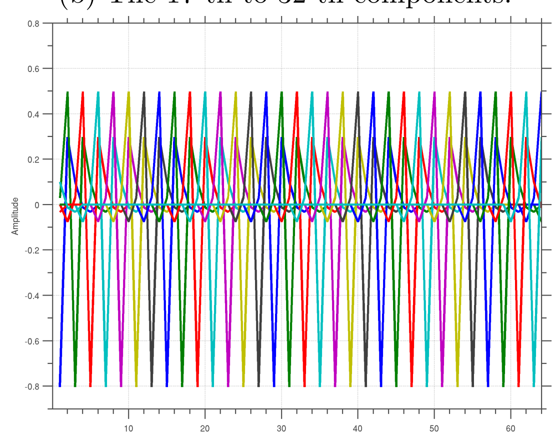
computed are shown in Fig 2.5b. The first index is 58 (black) and its coefficient 0.8. The second index is 22 (red) and its coefficient is 0.4. They were exactly the components that consisted the synthetic signal. For a real case, we don't know the original components of the signal. We must compare the reconstructed signal with the original one. We obtain the reconstructed signal in Figure 2.5c by summing the components with index 58 (black) and 22 (red) with their coefficients as weight. It is exactly the same signal as in Figure 2.5a. We obtained a compression ratio of $CR = 3\%$.



(a) The 1-st to 16-th components.



(b) The 17-th to 32-th components.



(c) The 33-th to 64-th components.

Figure 2.3: Dictionary of Symlets with $N = 64, J = 4$

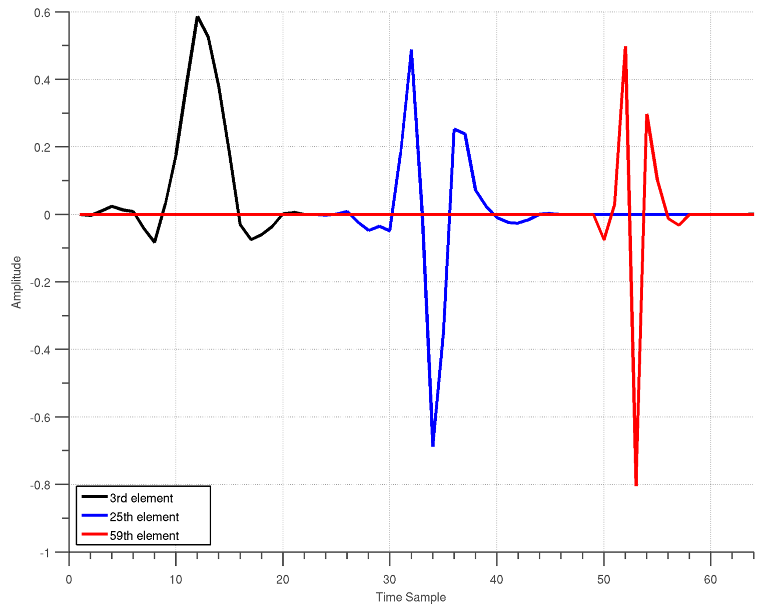


Figure 2.4: Three components of the dictionary with $N = 64, J = 4$

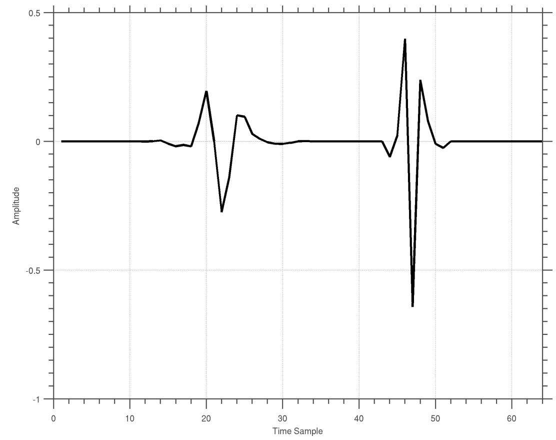
2.1.2 Numerical experiments

We generated a synthetic shot by forward modeling on a three layer velocity model. It is a common practice to use a Ricker wavelet as the source. The Ricker wavelet can be described by

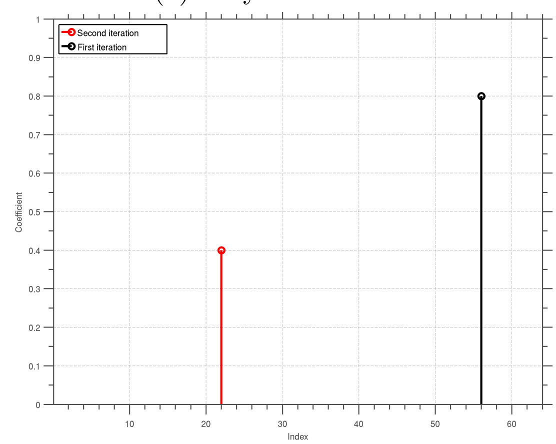
$$Ricker(t, f_0) = (1 - 2\pi^2 f_0^2 t^2) e^{-2\pi^2 f_0^2 t^2} \tag{2.18}$$

where f_0 is the dominant frequency. Here we used a Ricker wavelet with dominant frequency $f_0 = 10Hz$. The shot record has 220 traces and each has 1001 time samples with sample rate $8ms$. The right panel of Figure 2.7 shows the shot record in grey scale. We applied Algorithm 1 trace by trace with $maxiteration = \frac{N}{CR}$ where CR is compression ratio. In data compression literature, CR is defined as

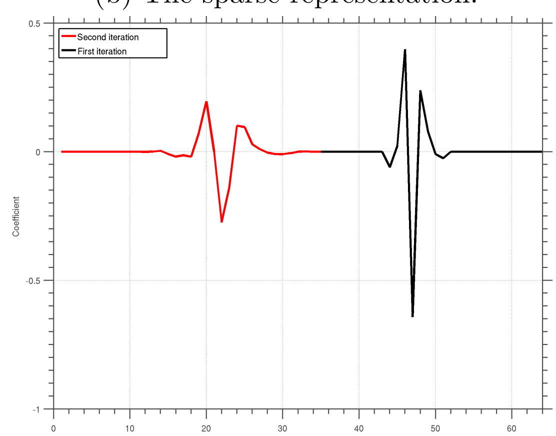
$$CR = 100 * \frac{Size\ of\ the\ compressed\ data}{Size\ of\ the\ original\ data} \% \tag{2.19}$$



(a) A synthetic trace



(b) The sparse representation.

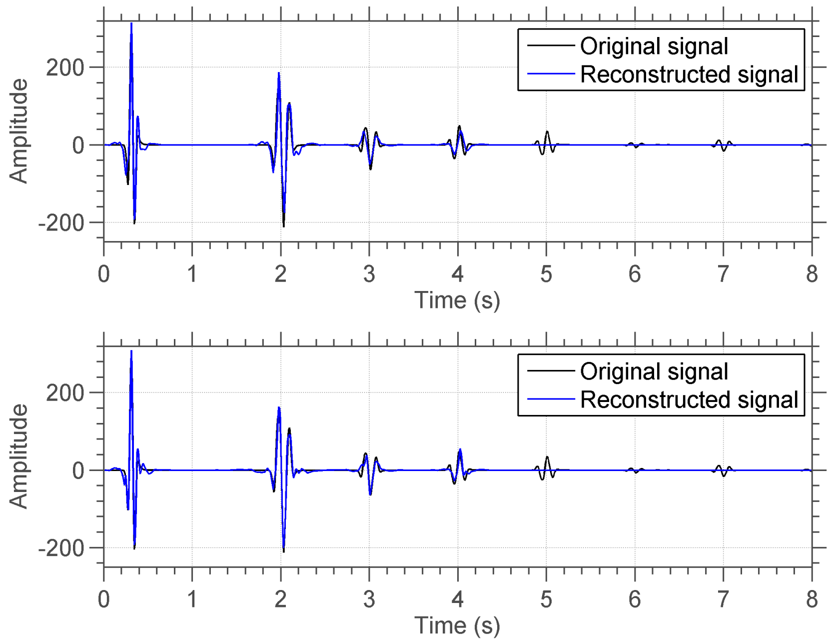


(c) The reconstructed signal.

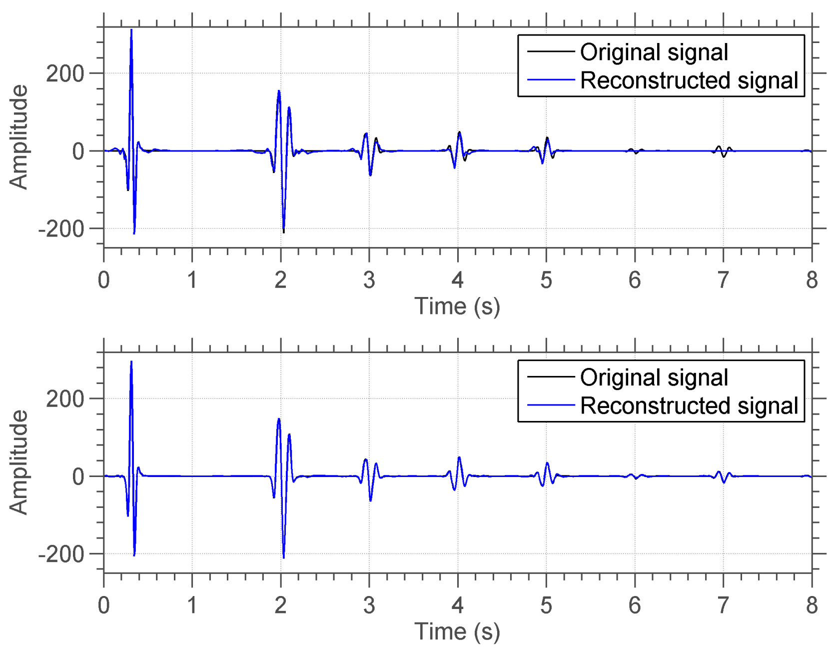
Figure 2.5: A example of synthetic trace

Specifically, our compression ratio can be computed by the ratio of the number of coefficients in the sparse representation and the number of the samples in the traces. We plotted the 20-th trace and its reconstructions in Figure2.6a with $CR = 1\%$ (top) and $CR = 2\%$ (bottom). The relative l_∞ errors of the reconstruction are 0.26 and 0.15 respectively. In Figure 2.6b, the original trace was compared with its reconstruction with $CR = 3.3\%$ (top) and $CR = 10\%$ (bottom). The relative l_∞ errors of the reconstruction are 0.07 and 0.01 respectively.

With $CR = 1\%$ and $CR = 2\%$ we obtained the reconstructed shot record in the left and middle panel of Figure 2.7. We observe serious oscillating artifacts (mosaic phenomenon) along the edges. The artifacts were reduced dramatically when we increase the size of compressed data as shown in Figure 2.8 when $CR = 3.3\%$ and $CR = 10\%$. The residues of the reconstructed shot record were concatenated in Figure 2.9 and they decreased as the size of compressed data increased.



(a) $CR = 1\%$ at the top, $CR = 2\%$ at the bottom.



(b) $CR = 3.3\%$ at the top, $CR = 10\%$ at the bottom.

Figure 2.6: Matching pursuit with Symlets reconstructed trace

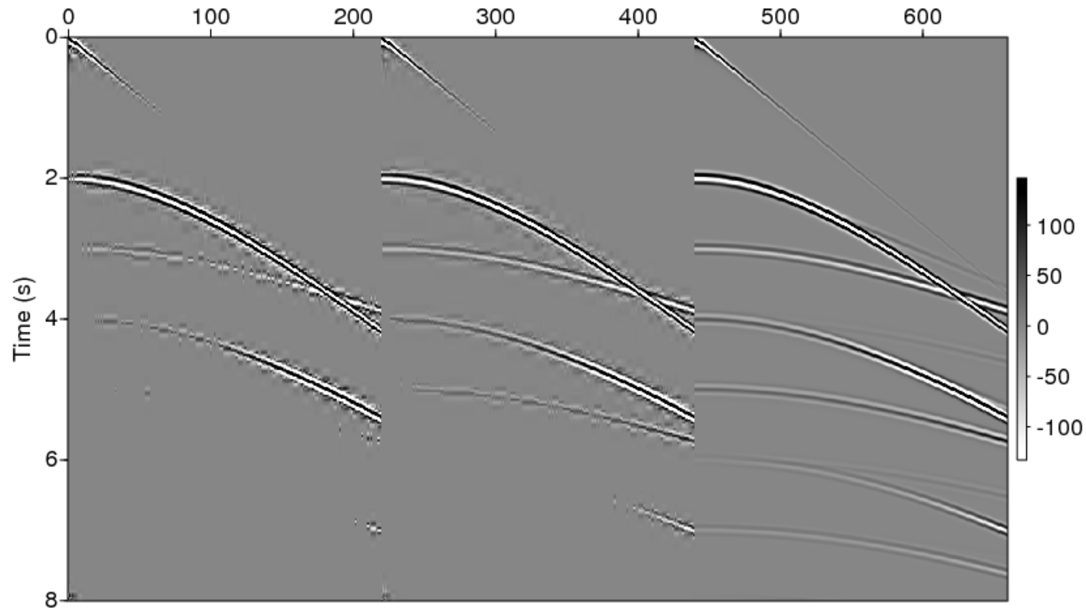


Figure 2.7: Matching pursuit with Symlets reconstructed data with $CR = 1\%$ on the left, $CR = 2\%$ in the middle, and original data on the right.

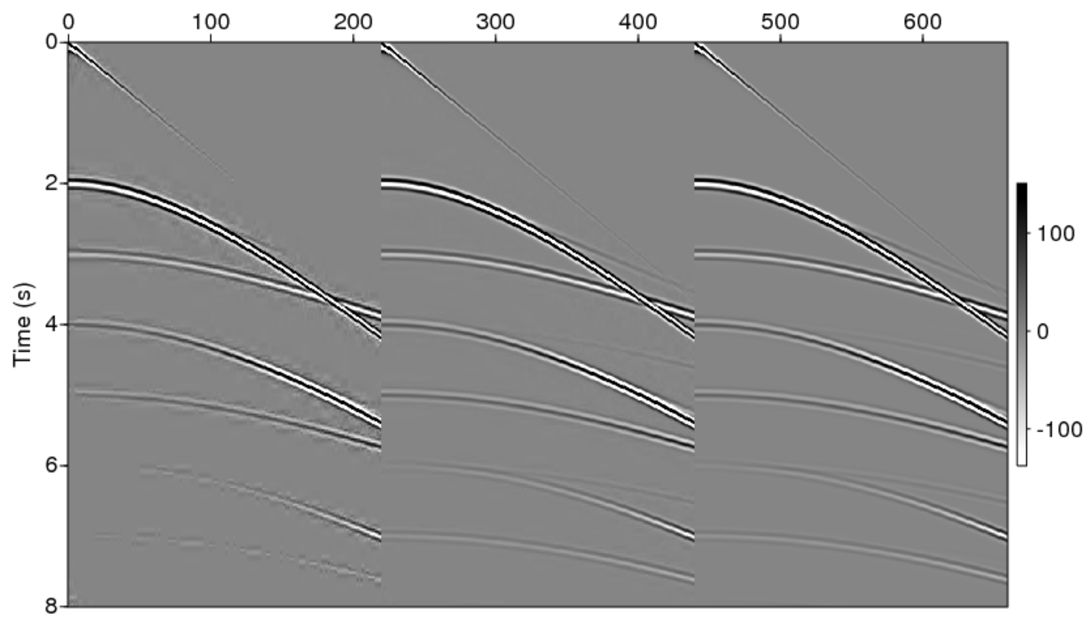


Figure 2.8: Matching pursuit with Symlets reconstructed data with $CR = 3.3\%$ on the left, $CR = 10\%$ in the middle, and original data on the right.

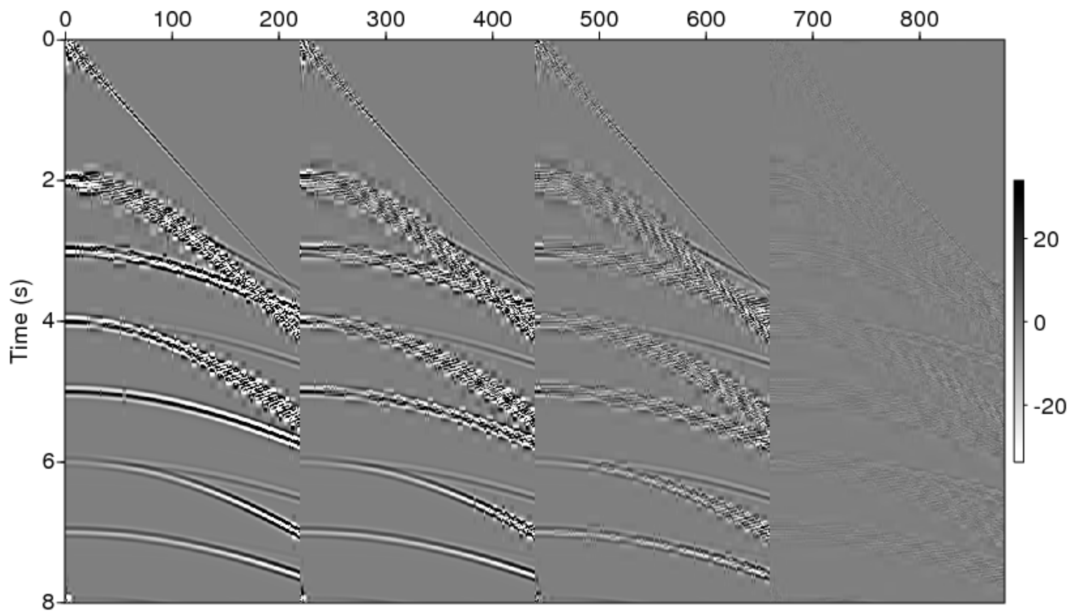


Figure 2.9: Matching pursuit with Symlets residue with $CR = 1\%, 2\%, 3.3\%, 10\%$ from left to right.

2.2 Subspace matching pursuit

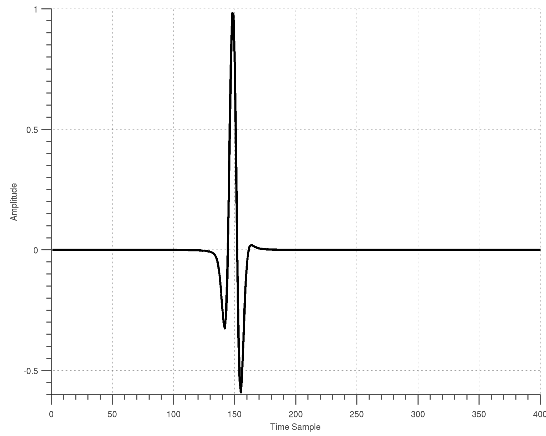
To improve the performance of compression further, we explored an idea to leverage source wavelets or estimated source wavelets. We found that if we use the dictionary based on a source wavelet, the results suffer from phase-rotation. Then we proposed a new idea of using subspace matching pursuit to obtain almost exact reconstructions for a phase-rotated signal. We have tested the subspace matching pursuit on the same synthetic shot as previous section.

2.2.1 Algorithm

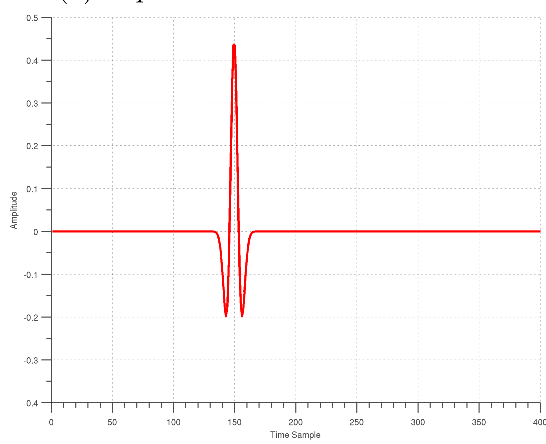
As is known, seismic signals constantly have phase rotations [43]. A phase-rotated signal can be decomposed into a linear combination of a zero-phase signal and the -90-degree phase-rotated signal. The zero-phase wavelet is symmetrical with a maximum at time zero (non-causal). We plotted a phase-rotated Ricker wavelet in Figure 2.10a. It can be decomposed into the zero-phase Ricker wavelet shown in Figure 2.10b and a -90-degree phase Ricker wavelet shown in Figure 2.10c.

We can construct a dictionary as a series of subspaces using zero-phase wavelets and their corresponding -90-degree phase-rotated wavelets. Then we can seek the correct subspace to reconstruct any phase-rotated signal from the subspace dictionary. We describe below our new matching pursuit strategy, called subspace matching pursuit, to obtain approximations to phase-rotated signals.

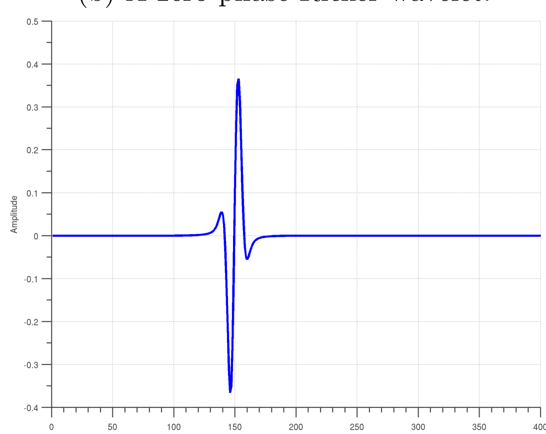
Let Γ be an index set and $\mathcal{D}_s = \{(\phi_p, \psi_p\}_{p\in\Gamma})$ be a subspace dictionary with ϕ_p



(a) A phase-rotated Ricker wavelet.



(b) A zero-phase Ricker wavelet.



(c) A -90 degree phase Ricker wavelet.

Figure 2.10: A phase-rotated Ricker wavelet in (a) can be decomposed into the zero-phase Ricker wavelet (b) and a -90-degree phase Ricker wavelet (c).

as a shifted zero-phase source wavelet and ψ_p as a shifted -90 degree phase wavelet.

Let $R^0 f = f$. Suppose that the mth-order residue $R^m f$ is already computed from $m \geq 0$. The next iteration chooses

$$(\phi_{p_m}, \psi_{p_m}) \in \mathcal{D}_s$$

such that

$$|< R^m f, \phi_{p_0} >| + |< R^m f, \psi_{p_0} >| \quad = \tag{2.20}$$

$$\max_{p \in \Gamma}(|< R^m f, \phi_p >| + |< R^m f, \psi_p >|). \tag{2.21}$$

and projects $R^m f$ on (ϕ_{p_m}, ψ_{p_m}):

$$R^m f =< R^m f, \phi_{p_m} > \phi_{p_m} + < R^m f, \psi_{p_m} > \phi_{p_m} + R^{m+1} f. \tag{2.22}$$

We stop at the Mth iteration when $\|R^M f\| < \epsilon$ or the maximum number of iterations is reached. Summing (2.22) on m between 0 and $M-1$ gets

$$f = \sum_{m=0}^{M-1} (< R^m f, \phi_{p_m} > \phi_{p_m} + < R^m f, \psi_{p_m} > \psi_{p_m}) + R^M f. \tag{2.23}$$

We summarize the subspace matching pursuit in Algorithm 2 and the generation of the subspace dictionary in Algorithm 3.

We generated a dictionary of size $N = 400$ by $2N = 800$ for the signal in Figure 2.10a with $N = 400$ time samples. The 1-st to the 400-th components are the zero-phase Ricker wavelets and we plotted 10% of them in Figure 2.11a. The 401-st to 800-th components are the -90 degree phase Ricker wavelets and we plotted 10% of them in Figure 2.11b.

Algorithm 2 Subspace matching pursuit

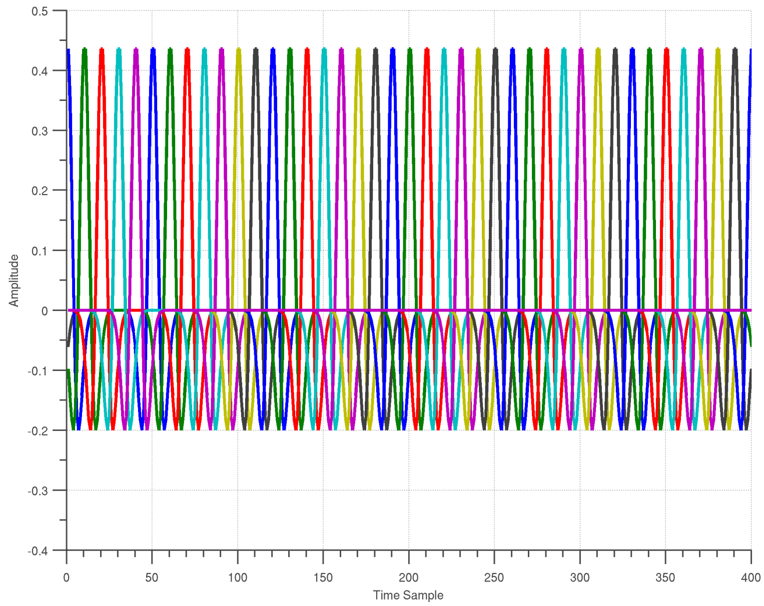
Input: signal(column vector of length N),dictionary(matrix of N by 2N),threshold, maxiteration
Output: coefficient (column vector of length 2Niter), index (column vector of length 2Niter)

```
Initialization: residue ← signal , m ← 0
while m ≤ maxiteration and norm(residue) ≤ threshold do
    m ← m + 1
    for i = 1 → N do
        innerprod0(i) ← dot-product(dictionary(:, i), residue)
        innerprod90(i)← dot-product(dictionary(:, i+N), residue)
    end for
    index(2m-1) ← the index of max(abs(innerprod0)+abs(innerprod90))
    index(2m) ← index(2m-1)+N
    coefficient(2m-1) ← innerprod0(index(m))
    coefficient(2m) ← innerprod90(index(m))
    residue ← residue - coefficient(2m-1)*dictionary(:, index(2m-1))
    residue ← residue - coefficient(2m)*dictionary(:, index(2m))
end while
Niter ← m
```

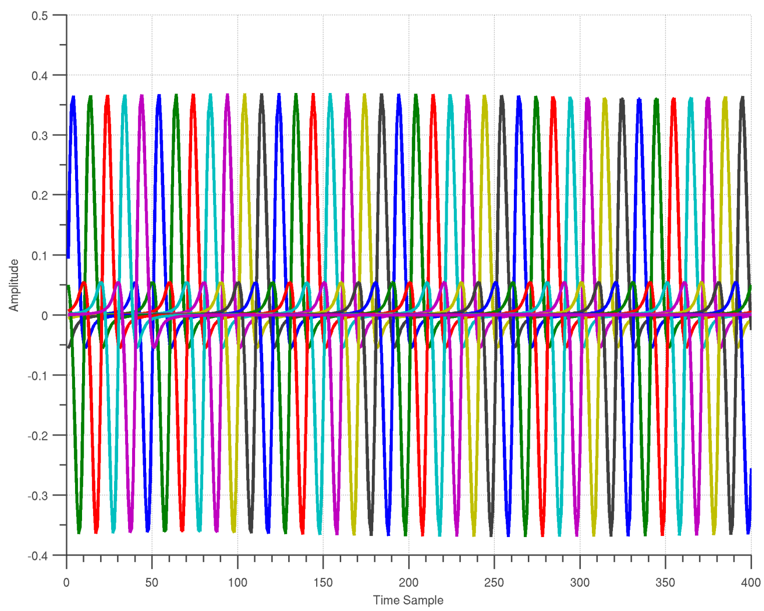
Algorithm 3 Generate Subspace Dictionary

Input: sourcewavelet(vector of length W), N(size of the signal)
Output: dictionary(matrix of N by 2N)

```
temp0 ← pad(sourcewavelet, N)
▷ Pad zeros to the end of array to be of length N
temp90 ← −90 degree phase-rotation of temp0
for i = 0 → N − 1 do
    dictionary(:, i) ← circle-shift(temp0, i)
    dictionary(:, i+N) ← circle-shift(temp90, i)
end for
```



(a) 10% of the zero-phase Ricker wavelets in the dictionary.



(b) 10% of the −90 degree phase Ricker wavelets in the dictionary.

Figure 2.11: Subspace dictionary.

After one iteration of subspace matching pursuit algorithm, we obtained two coefficients as shown in Figure 2.12. We obtain the reconstructed signal by summing the zero-phase Ricker wavelet (black) and −90-degree phase Ricker wavelet (red) with their coefficients as weights. It is exactly the same as the original signal.

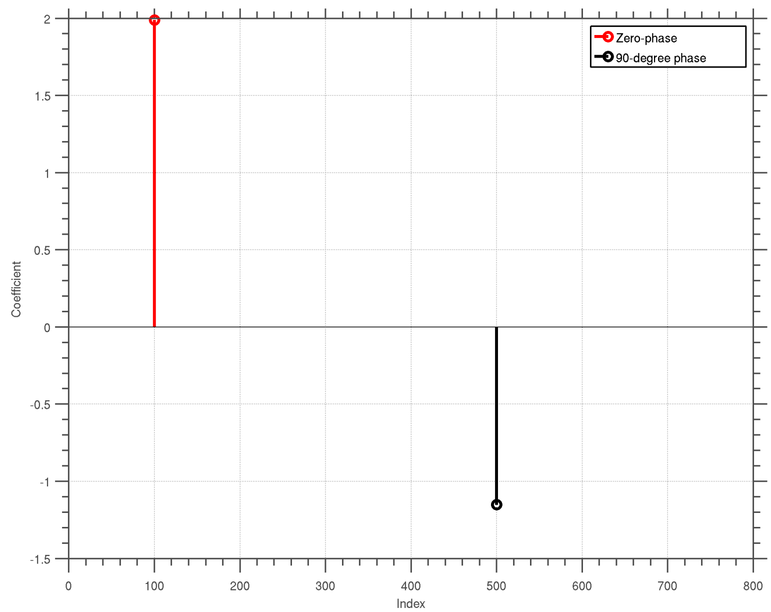


Figure 2.12: A sparse representation of a phase-rotated Ricker wavelet.

Next, we consider another example. We generated a synthetic trace by summing a 30-degree and 60-degree rotated Ricker wavelet as shown in Figure 2.13. We satisfied the threshold after 2 iterations; therefore, we only need to save 4 coefficients and 2 indices to reconstruct the signal. As can be seen from Figure 2.13, our reconstructed signal overlaps exactly with the original one. Our compression ratio is 1% in this case.

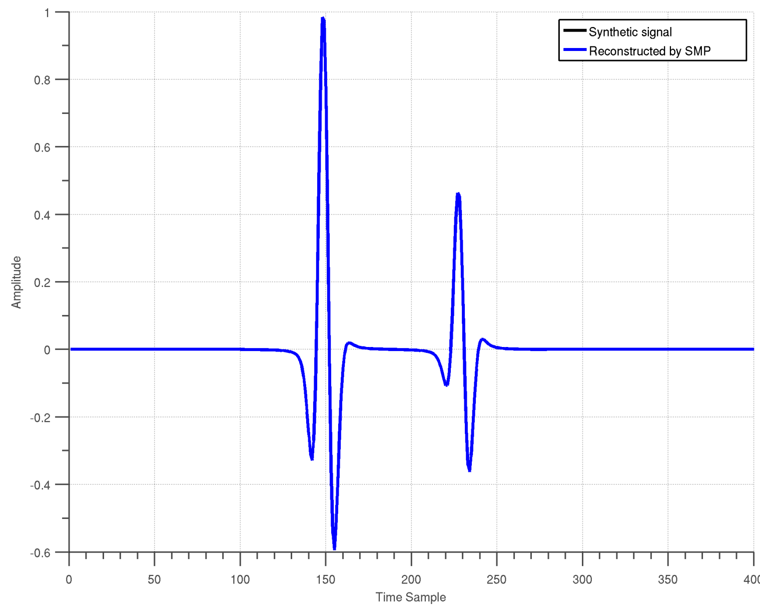
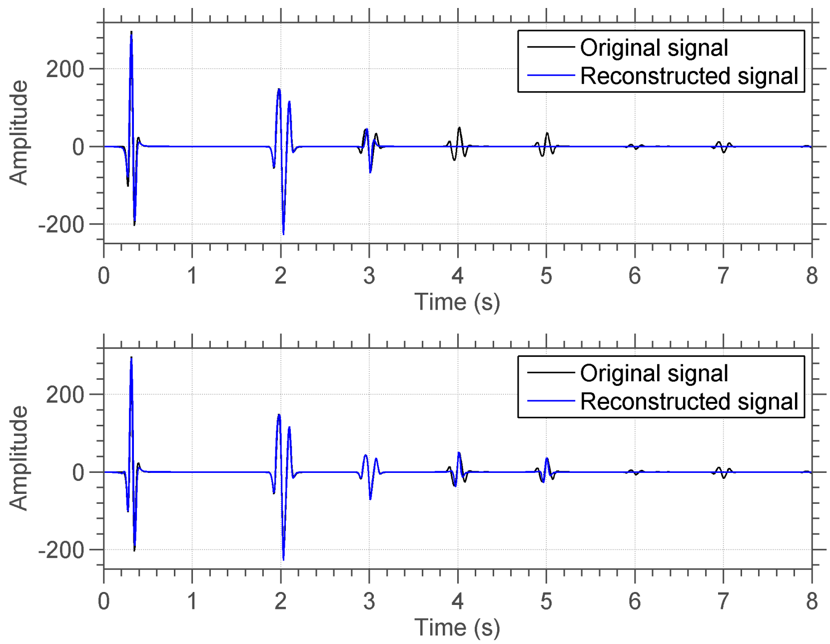


Figure 2.13: Subspace matching pursuit: the reconstructed and original signal agree perfectly.

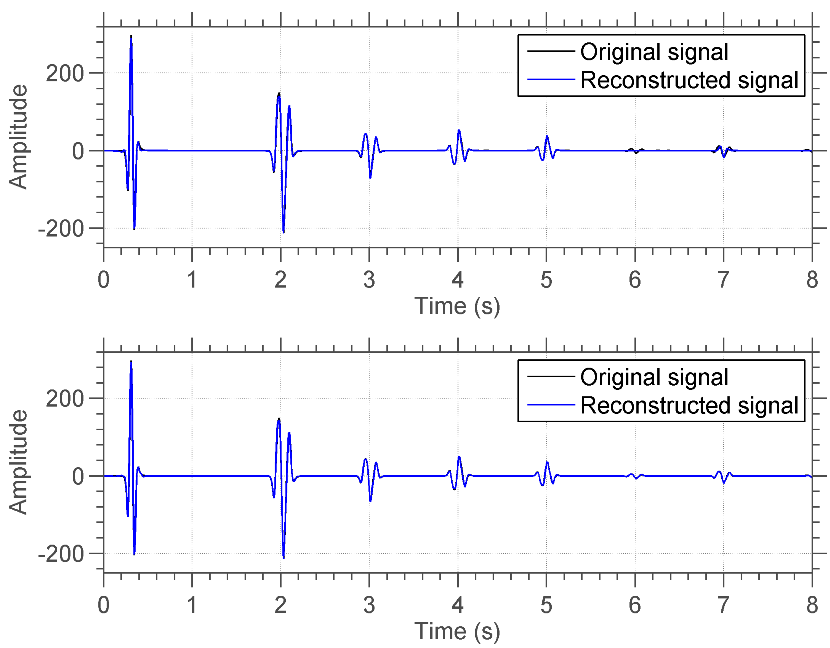
2.2.2 Numerical experiments

We have tested this approach on the same synthetic shot as previous chapter. We plotted the 20-th trace and its reconstructions in Figure2.14a with $CR = 1\%$ (top) and $CR = 2\%$ (bottom). The relative l_∞ errors of the reconstruction are 0.17 and 0.11 respectively. In Figure 2.14b, the original trace was compared with its reconstruction with $CR = 3.3\%$ (top) and $CR = 10\%$ (bottom). The relative l_∞ errors of the reconstruction are 0.05 and 0.03 respectively.

With $CR = 1\%$ and $CR = 2\%$ we obtained the reconstructed shot record in the left and middle panel of Figure 2.15. No oscillating artifacts (mosaic phenomenon) exist. We have more accurate reconstructed data as shown in Figure 2.16 when $CR = 3.3\%$ and $CR = 10\%$. The residues of the reconstructed shot record were



(a) $CR = 1\%$ at the top, $CR = 2\%$ at the bottom.



(b) $CR = 1\%$ at the top, $CR = 2\%$ at the bottom.

Figure 2.14: Subspace matching pursuit reconstructed trace

concatenated in Figure 2.17 and they decreased as the size of compressed data increased. We define common metrics to measure the quality of the reconstructed signal compared with the original. Let P be a signal (vector) or a image (matrix) and Q be the reconstructed one. Root mean square error is defined as

$$RMSE = \frac{norm(P-Q,2)}{\sqrt{length(P)}}$$

signal-to-noise ratio is

$$SNR = 20log_{10}\frac{norm(P,2)}{norm(P-Q,2)}$$

We have plotted the error and SNR comparison of Matching pursuit with Symlets and Subspace matching pursuit in Figure 2.18a and 2.18b. Subspace matching pursuit converges faster in the beginning, but then become slower than the matching pursuit with Symlets. Because the incompleteness of the subspace dictionary, it was hard to reconstruct the higher frequency residue. In the next section, we will solve this problem by using matching pursuit with combined dictionary.

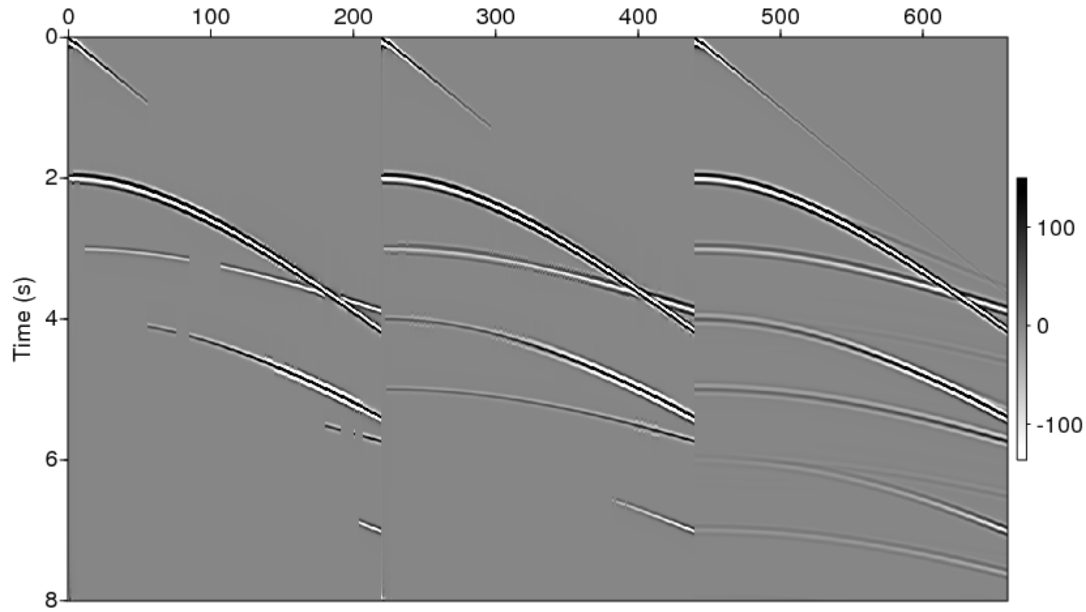


Figure 2.15: Subspace matching pursuit reconstructed data with $CR = 1\%$ on the left, $CR = 2\%$ in the middle, and original data on the right.

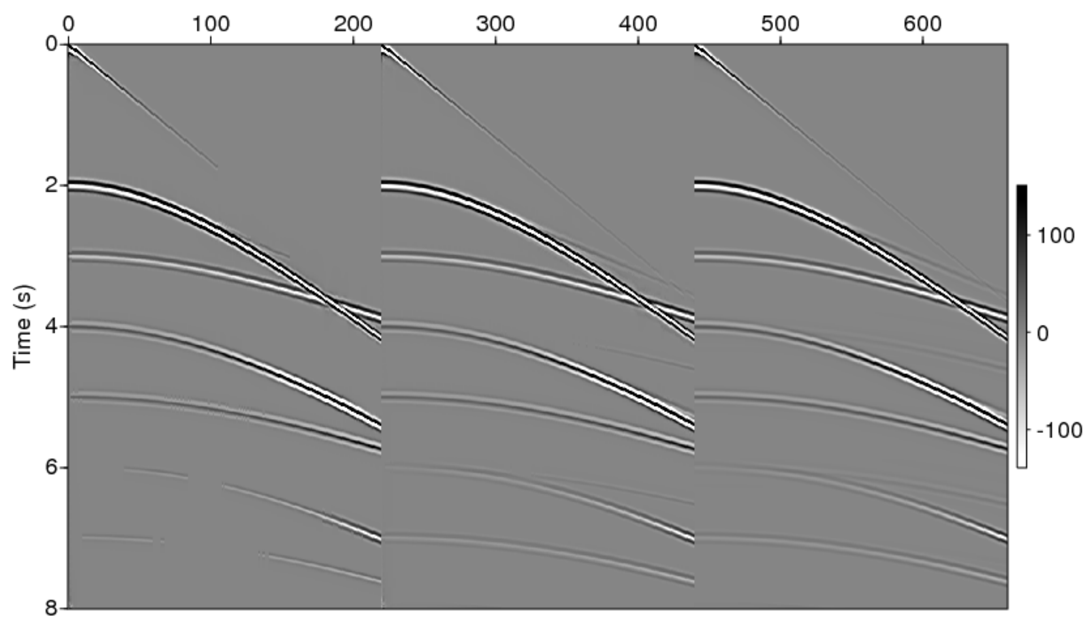


Figure 2.16: Subspace matching pursuit reconstructed data with $CR = 3.3\%$ on the left, $CR = 10\%$ in the middle, and original data on the right.

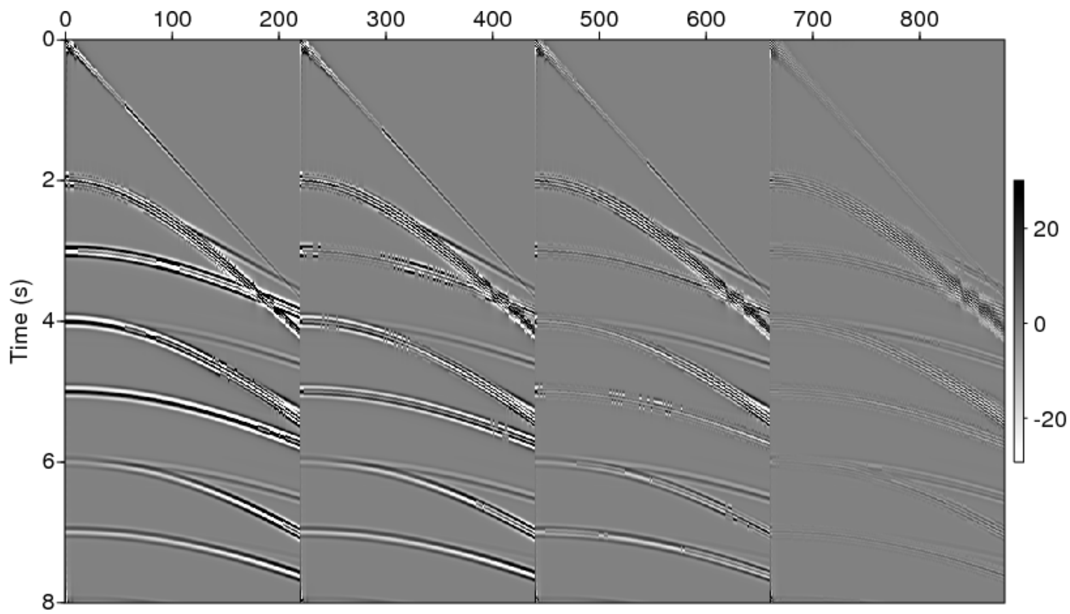
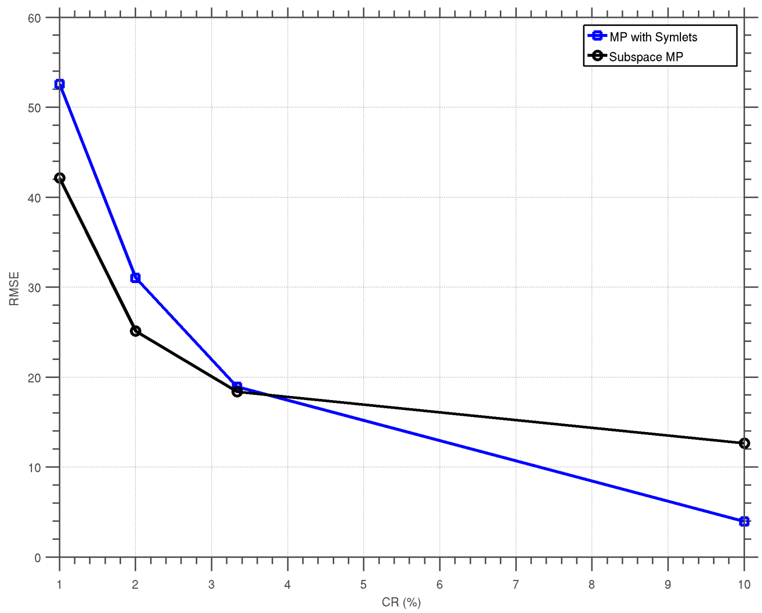
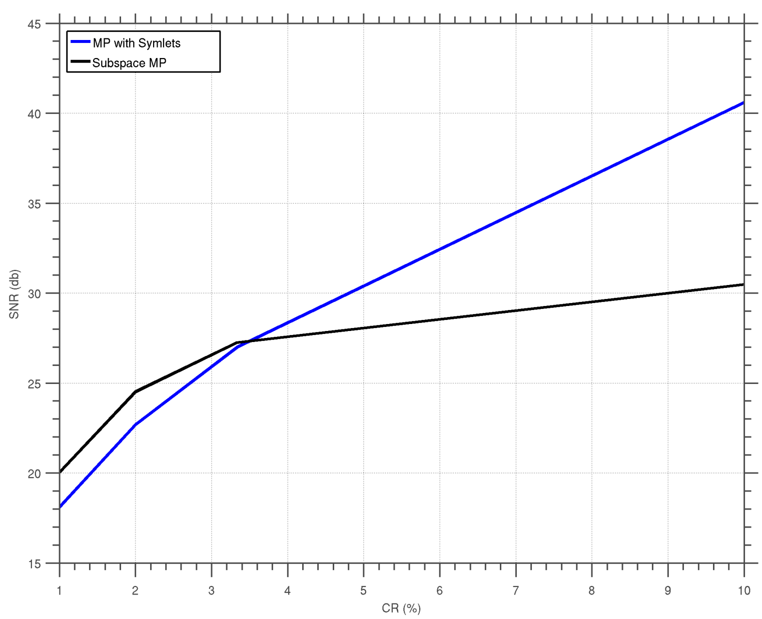


Figure 2.17: Subspace matching pursuit residue with $CR = 1\%, 2\%, 3.3\%, 10\%$ from left to right.



(a) Error.



(b) SNR.

Figure 2.18: Comparison of matching pursuit with Symlets and subspace matching pursuit.

2.3 Matching pursuit with combined dictionaries

In this section, we propose using matching pursuit with the combined dictionaries of Symlets and subspace. We obtain the best quality with the same compression ratio. Only limited artifacts were observed for various compression ratios for our test synthetic shot.

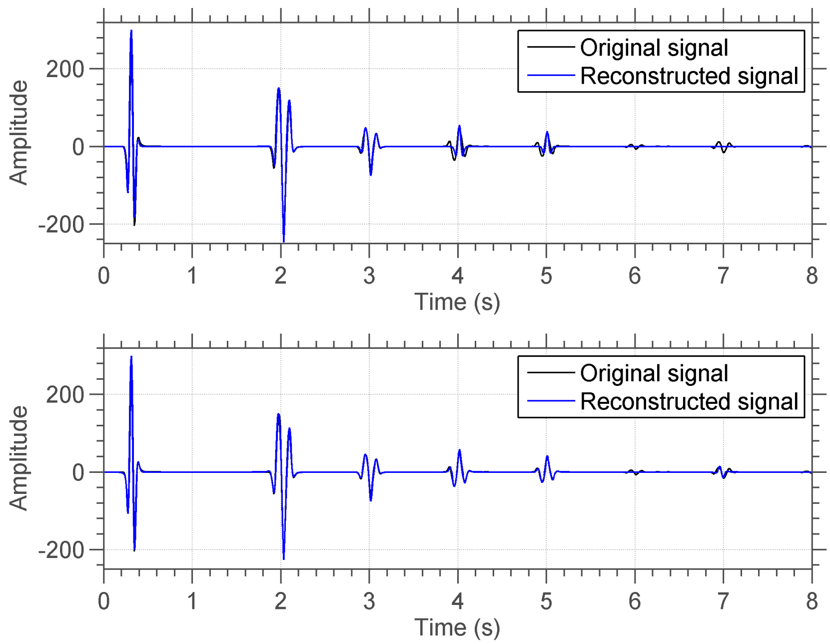
2.3.1 Algorithm

We combined the Symlets dictionary and subspace dictionary to form a dictionary $\mathcal{D}_c = \{\omega_p, \phi_p, \psi_p\}$. Then we applied matching pursuit on $\mathcal{D}_c$.

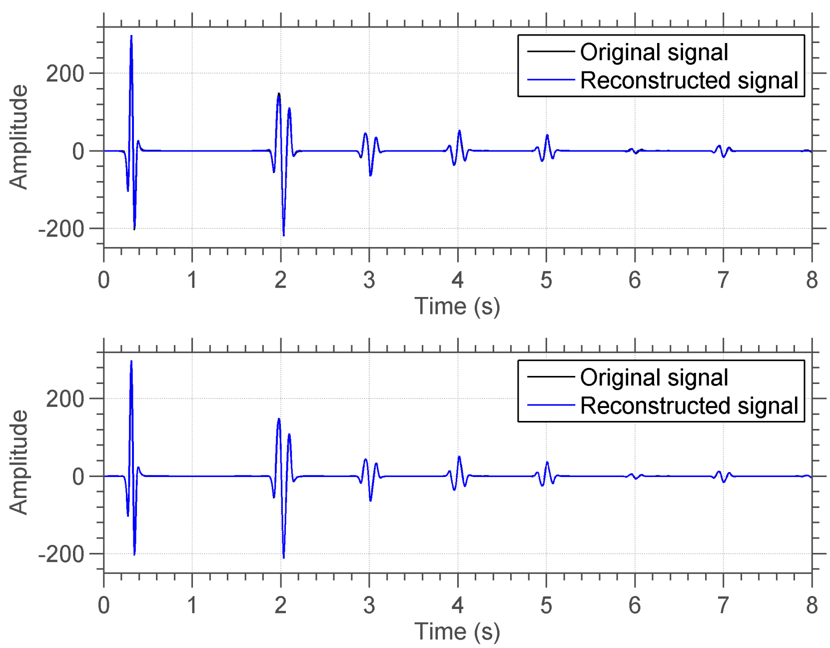
2.3.2 Numerical results

We have applied our combined Symlets dictionary and subspace dictionary matching pursuit. We plot the 20-th trace and its reconstructions in Figure 2.19a with $CR = 1\%$ (top) and $CR = 2\%$ (bottom). The relative l_∞ errors of the reconstruction are 0.12 and 0.07 respectively. In Figure 2.19b, the original trace was compared with its reconstruction with $CR = 3.3\%$ (top) and $CR = 10\%$ (bottom). The relative l_∞ errors of the reconstruction are 0.04 and 0.01 respectively.

With $CR = 1\%$ and $CR = 2\%$ we obtain the reconstructed shot record in the left and middle panel of Figure 2.20. Only limited artifacts were observed compared to Figure 2.7 and more seismic events were reconstructed compared to Figure 2.15. With $CR = 3.3\%$ and $CR = 10\%$ we obtained reconstructed shot record of better



(a) $CR = 1\%$ at the top, $CR = 2\%$ at the bottom.



(b) $CR = 1\%$ at the top, $CR = 2\%$ at the bottom.

Figure 2.19: Matching pursuit with combined dictionaries reconstructed trace.

quality in the left and middle panel of Figure 2.21.

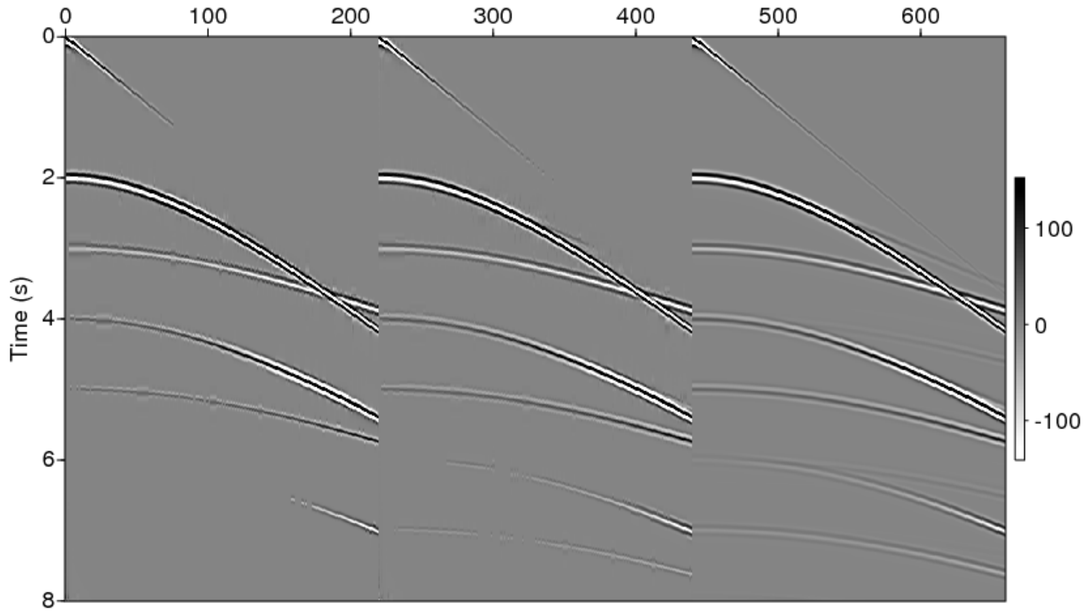


Figure 2.20: Matching pursuit with combined dictionary reconstructed data with $CR = 1\%$ on the left, $CR = 2\%$ in the middle, and original data on the right.

Next, we validated matching pursuit with combined dictionaries approach with noisy synthetic data. We added white noise at the level of $SNR = 33dB$ to the synthetic shot and then applied matching pursuit with combined dictionary.

We plotted the 20-th trace and its reconstructions in Figure2.23a with $CR = 1\%$ (top) and $CR = 2\%$ (bottom).In Figure 2.23b, the original trace was compared with its reconstruction with $CR = 3.3\%$ (top) and $CR = 10\%$ (bottom).

With $CR = 1\%$ and $CR = 2\%$ we obtained the reconstructed shot record in the left and middle panel of Figure 2.24. With $CR = 3.3\%$ and $CR = 10\%$ we obtained the reconstructed shot record in the left and middle panel of Figure 2.25. We have recovered as many seismic events as with clean synthetic data as shown in Figures

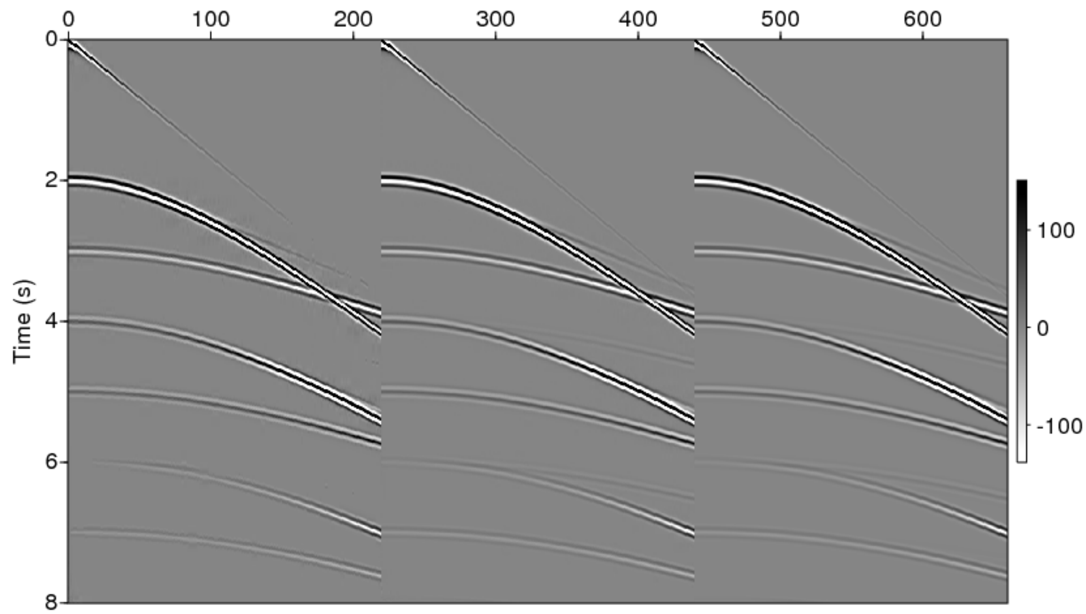


Figure 2.21: Matching pursuit with combined dictionary reconstructed data with $CR = 3.3\%$ on the left, $CR = 10\%$ in the middle, and original data on the right.

2.20 and 2.21. The residue for the reconstructed shot record were concatenated in Figure 2.26. The lateral correlation in the residue decreases when the size of the compressed data increases. For a compression ratio 10% we barely found lateral correlation which indicates good quality of reconstruction of seismic events, but not for the noise.

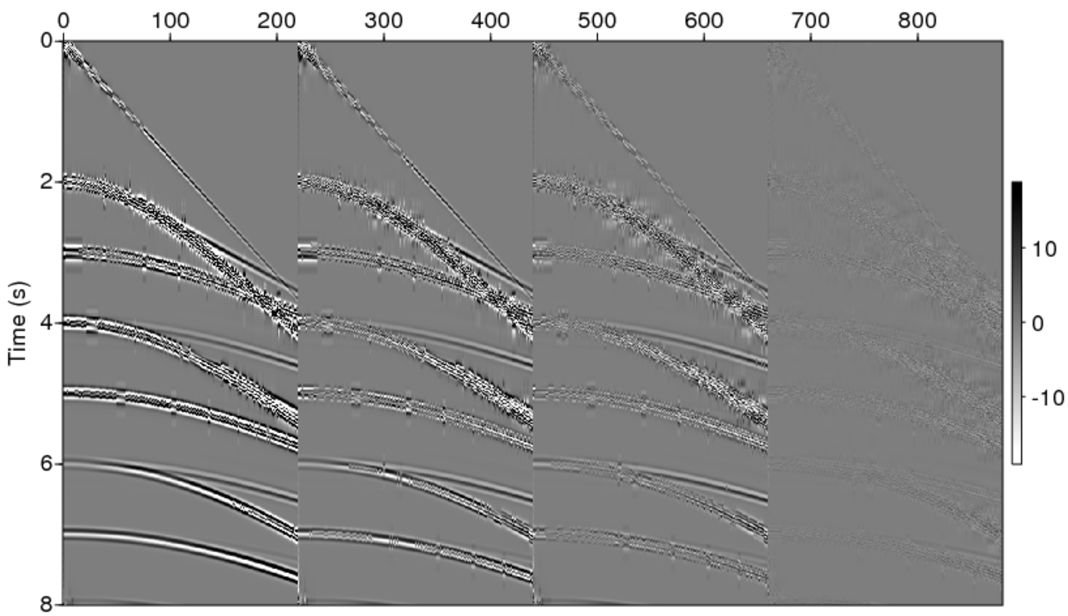
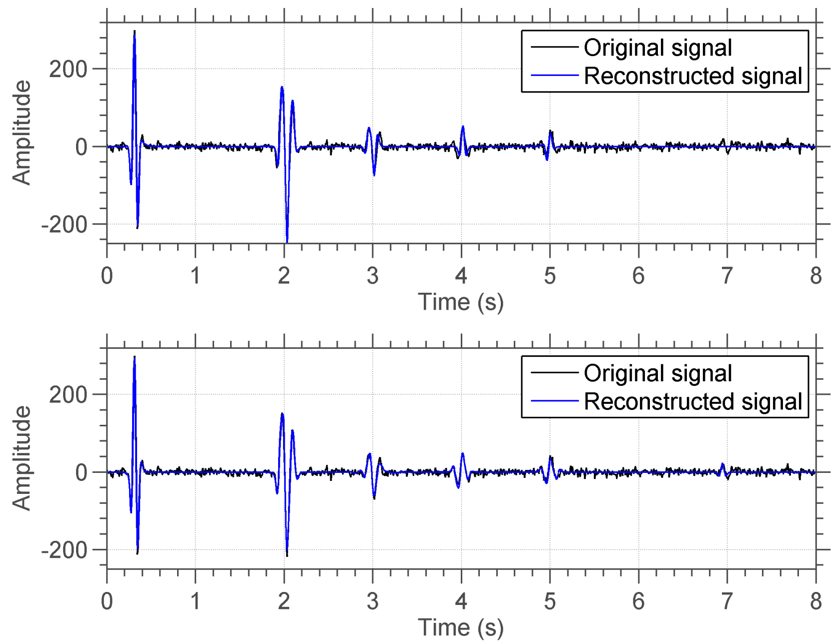
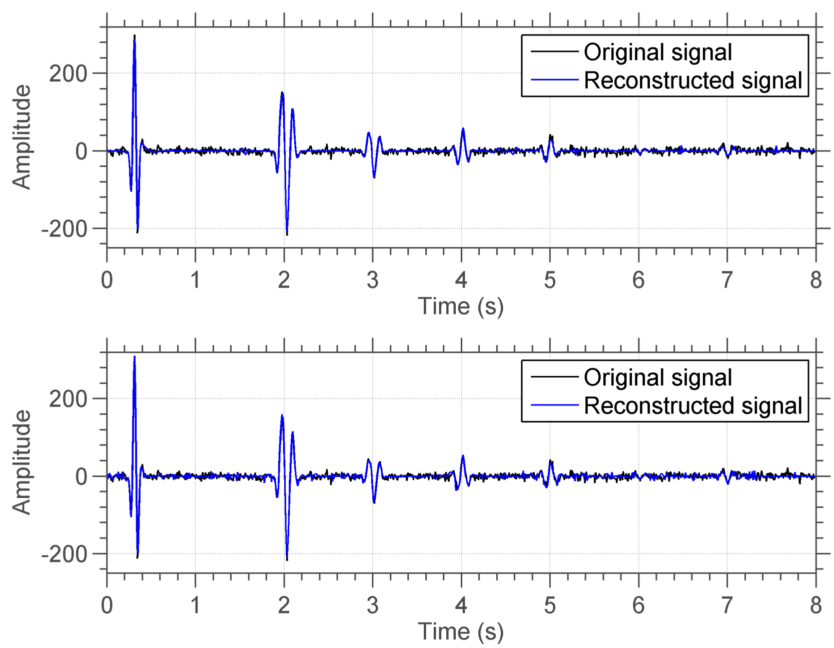


Figure 2.22: Matching pursuit with combined dictionary residue with $CR = 1\%$, 2%, 3.3%, 10% from left to right.



(a) $CR = 1\%$ at the top, $CR = 2\%$ at the bottom.



(b) $CR = 1\%$ at the top, $CR = 2\%$ at the bottom.

Figure 2.23: Matching pursuit with combined dictionaries reconstructed trace.

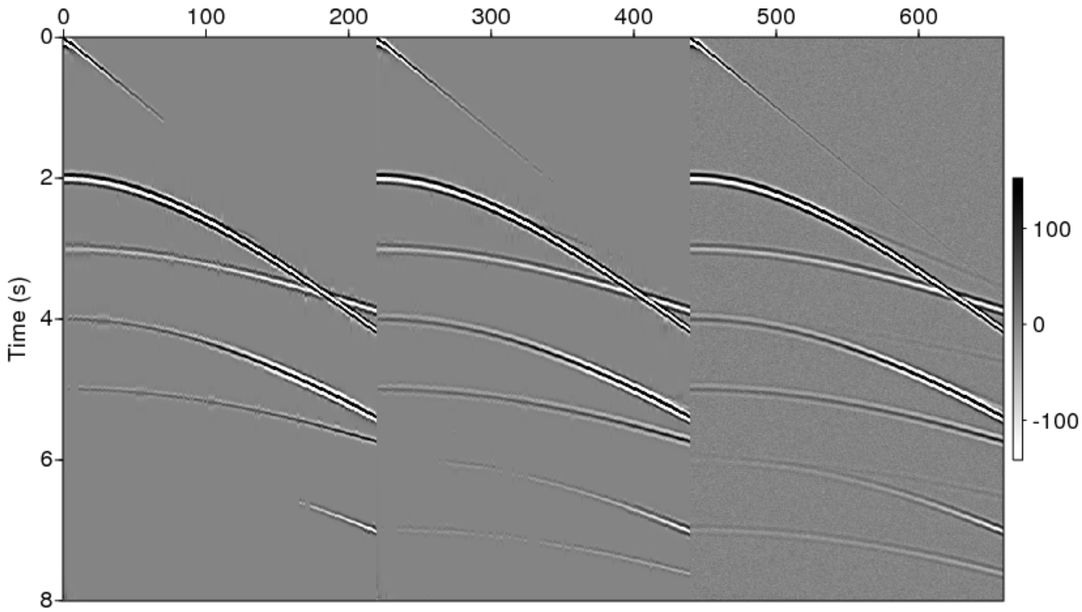


Figure 2.24: Matching pursuit with combined dictionary reconstructed data with $CR = 1\%$ on the left, $CR = 2\%$ in the middle, and original data on the right.

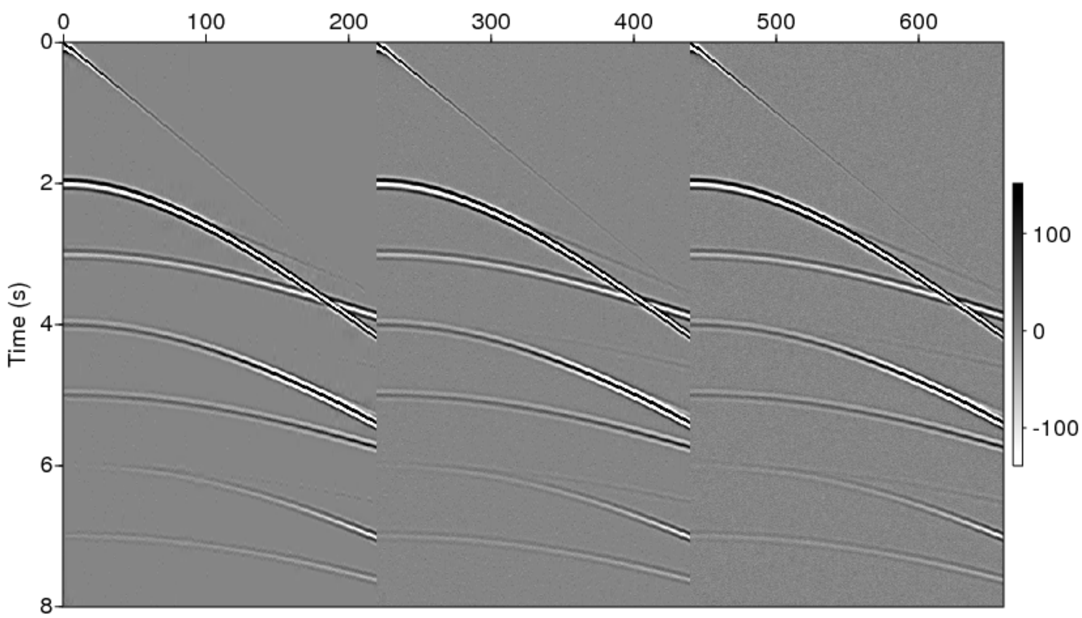


Figure 2.25: Matching pursuit with combined dictionary reconstructed data with $CR = 3.3\%$ on the left, $CR = 10\%$ in the middle, and original data on the right.

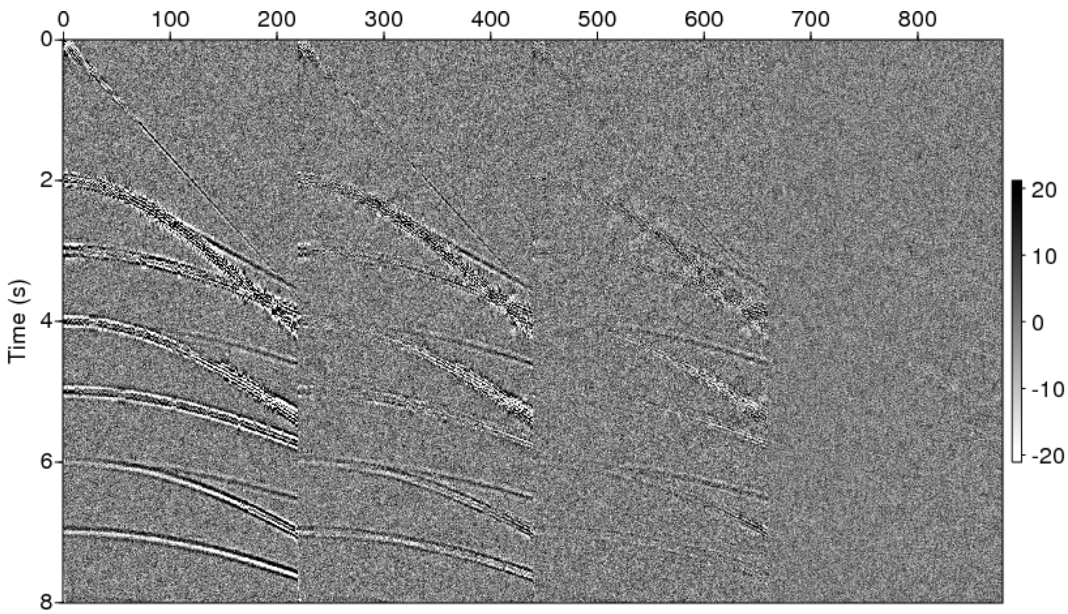


Figure 2.26: Matching pursuit with combined dictionary residue with $CR = 1\%$, 2%, 3.3%, 10% from left to right.

2.4 Comparison

We applied the three methods trace by trace on the same synthetic shot. For the 20-th trace, we have plotted the relative l_∞ error for all the three methods with various compression ratios in Figure 2.27. With $CR = 2\%$ we obtained the reconstructed

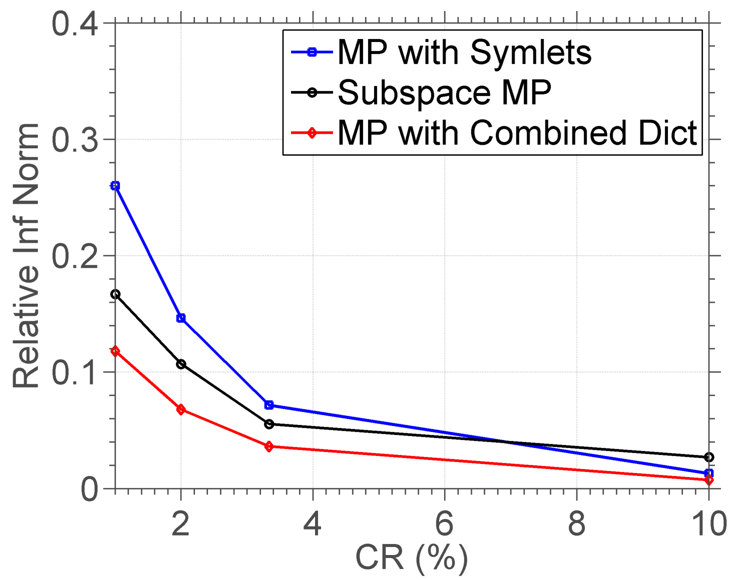


Figure 2.27: Relative infinity error comparison of three methods for the 20-th trace.

shot records and concatenated from left to right in Figures 2.28. The rightmost panel of Figure 2.28 shows the shot record in grey scale. From left to right, we show matching pursuit with Symlets, subspace matching pursuit, and matching pursuit with combined dictionaries. We also concatenated the residue of the three methods in the same order in Figures 2.29.

With $CR = 10\%$ we obtained the reconstructed shot records and concatenated from left to right in Figures 2.30. The rightmost panel of Figure 2.30 shows the shot

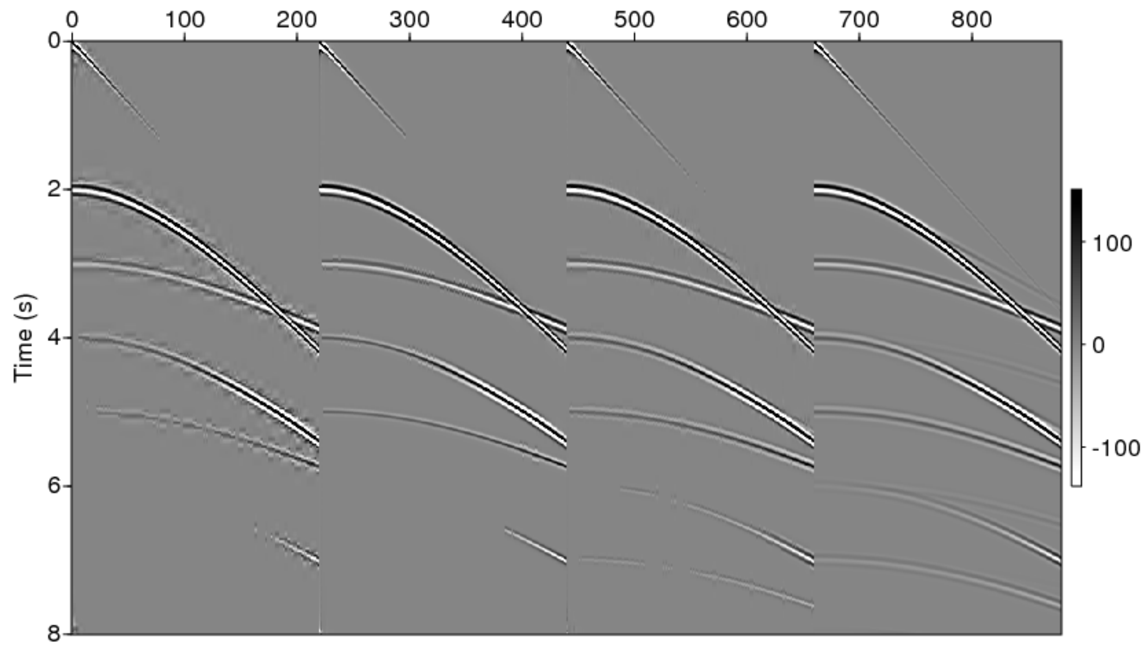


Figure 2.28: Original and reconstructed shot records for $CR = 2\%$

record in grey scale. From left to right, we show matching pursuit with Symlets, subspace matching pursuit, and matching pursuit with combined dictionaries. We also concatenated the residue of the three methods in the same order in Figures 2.31.

We observed serious oscillating artifacts (mosaic phenomenon) along the edges for matching pursuit with symlets in the first panel from the left in Figure 2.28. The artifacts were reduced dramatically when we increase the size of compressed data as shown in Figure 2.30. No oscillating artifacts (mosaic phenomenon) exist for subspace matching pursuit as seen in the second panel in both Figure 2.28 and 2.30. Only limited artifacts were observed compared for matching pursuit with combined dictionaries as shown in the third panel in both Figure 2.28 and 2.30. More seismic events were reconstructed compared to the other two.

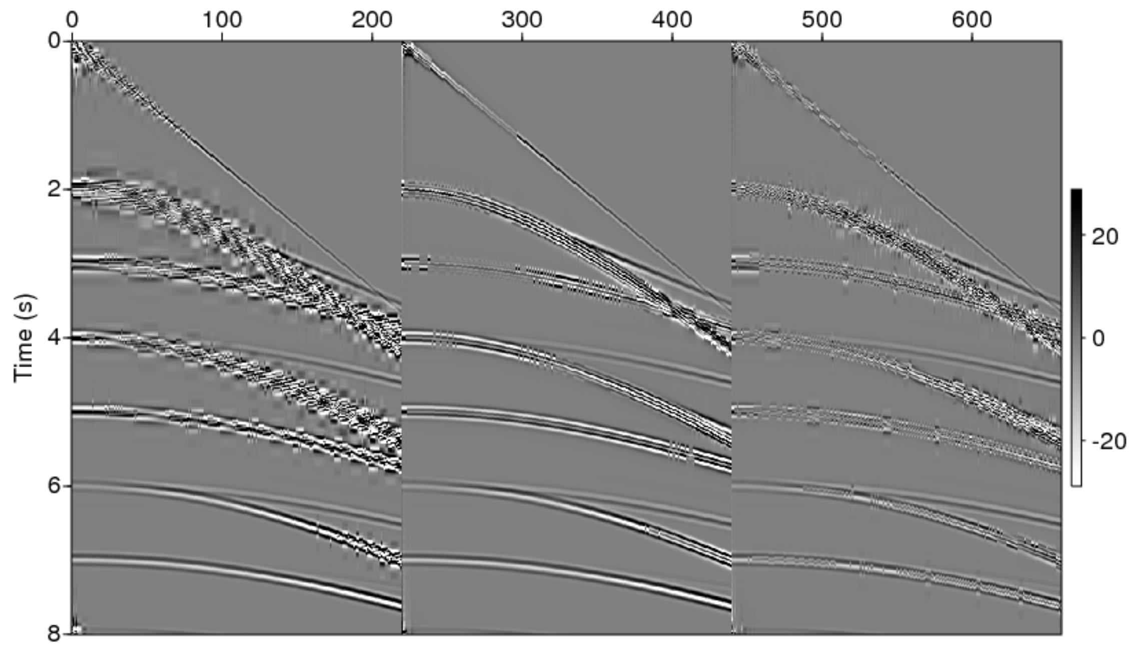


Figure 2.29: Residue of the three methods $CR = 2\%$

We compared the error and SNR for the three methods. As one can see from Figure 2.32a, errors for all methods decreased when the size of compressed data increases. Matching pursuit with combined dictionaries had the smallest error among the three methods. The subspace matching pursuit started with a smaller error than matching pursuit with Symlets, but the convergence slowed down due to incompleteness of its dictionary. SNR for all three methods increased as the size of compressed data increased. Matching pursuit with combined dictionaries obtained the highest SNR in Figure 2.32b.

We list the characteristics of the three methods below:

1. Matching pursuit with Symlets

    - worked without knowledge of source wavelet

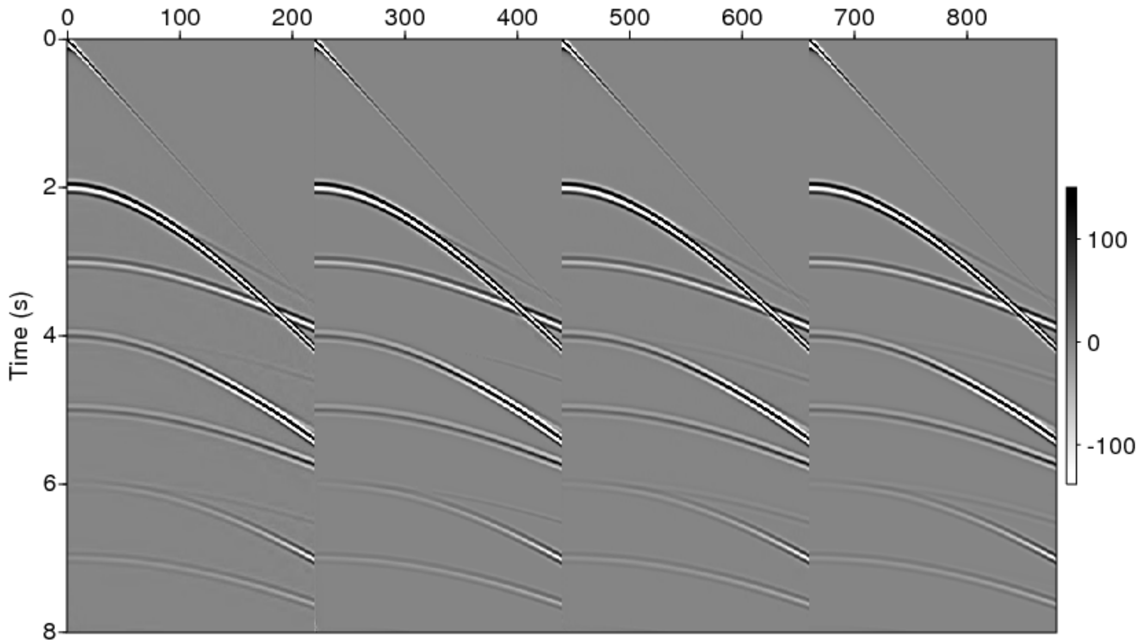


Figure 2.30: Original and reconstructed shot records for $CR = 10\%$

- obtained better quality than Discrete Wavelet Transform
- had obvious artifacts for too much compression

2. Subspace matching pursuit

- needed source wavelet or estimated source wavelet
- had no artifacts existing for various compression ratio
- had its convergence slowing down due to incompleteness of the dictionary

3. Matching pursuit with combined dictionaries (Symlets and Subspace)

- needed source wavelet or estimated source wavelet
- had limited artifacts existing for various compression ratio
- achieved best quality with fastest convergence among the three methods

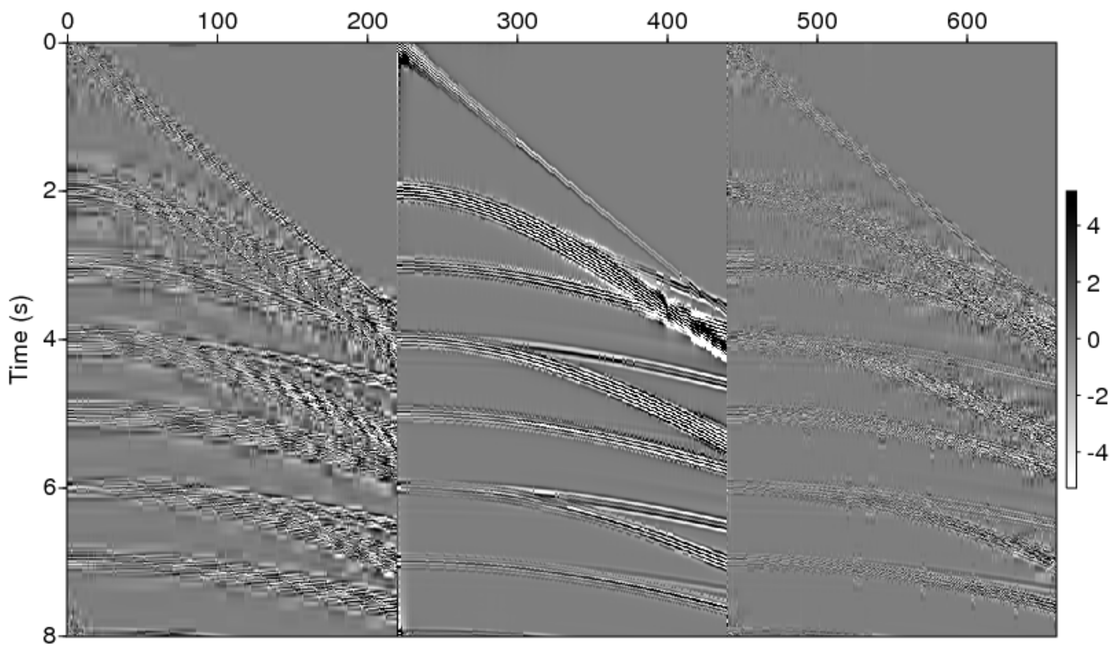
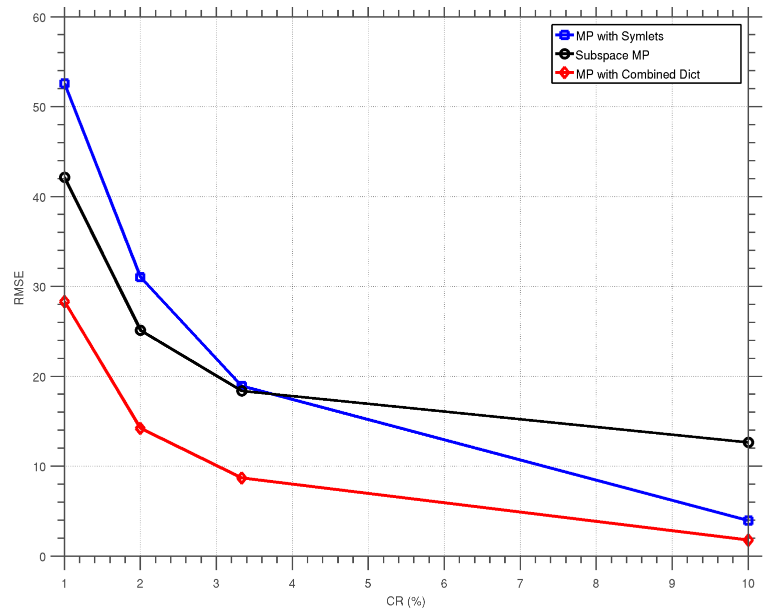
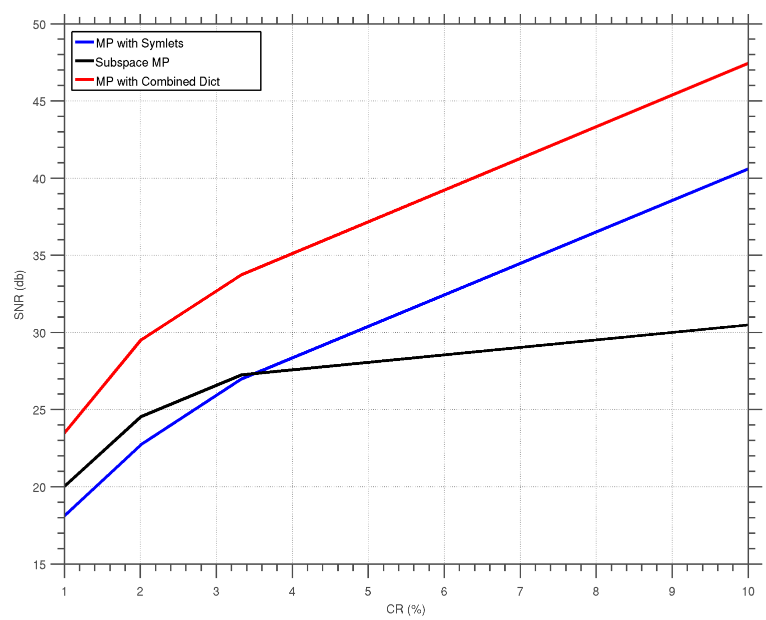


Figure 2.31: Residue of the three methods $CR = 10\%$



(a) Error.



(b) SNR.

Figure 2.32: Comparison of the three methods.

CHAPTER 3

Wavelet-based microseismic first arrival detection

We proposed a new method of automatic first break detection of P-waves and S-waves. Our method is based on a time-frequency analysis of the seismic trace using minimum uncertainty (μ-)wavelets, in particular in the minimum-phase form. We have tested our method on both lab data with various S/N and on field data.

3.1 Method

In this section, we propose a new automatic first arrival picking method based on μ-wavelets. We introduced the μ-wavelets for spectral decomposition of microseismic traces to generate a indicator function. We developed a technique to yield a less noisy indicator function using Hermite distributed approximating functionals (HDAFs).

3.1.1 μ-Wavelets and Heisenberg's uncertainty principle

It is well-known that the bandwidth-duration product of any physical signal has to be greater than a minimum universal value. Essentially, it is a rigorous mathematical bound which makes it impossible to create signals with arbitrarily narrow width simultaneously in both time and frequency. This is known as Heisenberg's uncertainty principle, and we will derive the μ-wavelets from it.

We define Heisenberg's uncertainty product as

$$\Delta = \Delta_t \Delta_\omega = \int_{-\infty}^{\infty} t^2 |\phi(t)|^2 dt \int_{-\infty}^{\infty} \omega^2 |\phi(\omega)|^2 d\omega \tag{3.1}$$

where $\phi(t)$ is the time domain signal and $\phi(\omega)$ is the Fourier transform of it:

$$\phi(\omega) = \int_{-\infty}^{+\infty} \phi(t) e^{-iwt} dt \tag{3.2}$$

Δ_t and Δ_ω are the standard deviation in time domain and frequency domain, respectively. Here we assumed that the expected values of $\phi(t)$ and $\phi(\omega)$ are zero. It is well-known that the absolute minimum of the uncertainty product is $\frac{1}{2}$ when $\phi(t)$ is a Gaussian signal. The uncertainty is associated with the standard deviation in time and frequency ($\Delta_t \Delta_\omega = constant$).

To generate a family of wavelets, we start with an arbitrary function $\phi_0(t)$ which we shall choose to be a Gaussian. We modify $\phi_0(t)$ to create a new function

$$\phi_1(t) = \phi_0(t) + \psi(t), \tag{3.3}$$

By Parseval's Theorem

$$\int_{-\infty}^{\infty} \omega^2 |\phi(\omega)|^2 d\omega = \int_{-\infty}^{\infty} |IFT(\omega\phi(\omega))|^2 dt = \int_{-\infty}^{\infty} |i\frac{\partial\phi(t)}{\partial t}|^2 dt \tag{3.4}$$

Thus the uncertainty product of ϕ_1 can be written as

$$\Delta = \int_{-\infty}^{\infty} |t\phi_1(t)|^2 dt \int_{-\infty}^{\infty} |\frac{\partial\phi_1(t)}{\partial t}|^2 dt \tag{3.5}$$

We substitute Equation (3.3) into the second integral in Equation (3.5),

$$\begin{aligned}
\Delta &= \int_{-\infty}^{\infty} |t\phi_1(t)|^2 dt \int_{-\infty}^{\infty} |\frac{\partial(\phi_0(t)+\psi(t))}{\partial t}|^2 dt \\
&= \int_{-\infty}^{\infty} |t\phi_1(t)|^2 dt \int_{-\infty}^{\infty} [(\frac{\partial\phi_0(t)}{\partial t})^2 + 2\frac{\partial\phi_0(t)}{\partial t}\frac{\partial\psi(t)}{\partial t} + (\frac{\partial\psi(t)}{\partial t})^2] dt \\
&= \int_{-\infty}^{\infty} |t\phi_1(t)|^2 dt \int_{-\infty}^{\infty} [(\frac{\partial\phi_0(t)}{\partial t})^2 + 2\frac{\partial\phi_0(t)}{\partial t}\frac{\partial\psi(t)}{\partial t}] dt + \int_{-\infty}^{\infty} |t\phi_1(t)|^2 dt \int_{-\infty}^{\infty} (\frac{\partial\psi(t)}{\partial t})^2 dt
\end{aligned}$$

It is convenient to separate Δ into 2 separate contributions, Δ_f and Δ_v, defined by

$$\begin{aligned}
\Delta_f &= \int_{-\infty}^{\infty} |t\phi_1(t)|^2 dt \int_{-\infty}^{\infty} [(\frac{\partial\phi_0(t)}{\partial t})^2 + 2\frac{\partial\phi_0(t)}{\partial t}\frac{\partial\psi(t)}{\partial t}] dt \\
\Delta_v &= \int_{-\infty}^{\infty} |t\phi_1(t)|^2 dt \int_{-\infty}^{\infty} (\frac{\partial\psi(t)}{\partial t})^2 dt
\end{aligned}$$

We see that Δ_v is the product of the full uncertainty in t times the uncertainty in ω that is strictly due to the modification function, $\psi(t)$. So

$$\Delta = \Delta_f + \Delta_v.$$

We minimize Δ_v to get the minimum of the full time uncertainty times the $\Delta\omega$ due solely to $\psi(t)$. Thus, by applying the Schwarz inequality, we get

$$\Delta_v \geq |\int_{-\infty}^{\infty} (t\phi_1(t))(\frac{\partial\psi(t)}{\partial t}) dt|^2 \tag{3.6}$$

The equality holds when $t\phi_1(t)$ and $\frac{\partial\psi(t)}{\partial t}$ are proportional

$$\frac{1}{\sigma^2}t\phi_1(t) = \frac{\partial\psi(t)}{\partial t} = \frac{\partial(\phi_1(t) - \phi_0(t))}{\partial t} \tag{3.7}$$

If we take the Fourier transform on both sides, we get

$$i\frac{1}{\sigma^2}\frac{\partial\phi_1(\omega)}{\partial\omega} = -i\omega(\phi_1(\omega) - \phi_0(\omega)) \tag{3.8}$$

We can rewrite it as

$$-\frac{1}{\sigma^2\omega}\frac{\partial\phi_1(\omega)}{\partial\omega} = \phi_1(\omega) - \phi_0(\omega) \tag{3.9}$$

If we define $\xi = \omega^2\sigma^2/2$, the Equation (3.9) is equivalent to

$$-\frac{\partial\phi_1(\xi)}{\partial\xi} = \phi_1(\xi) - \phi_0(\xi), \tag{3.10}$$

which has the solution

$$\phi_1(\xi) = \int_0^\xi d\xi' e^{-(\xi-\xi')}\xi_0(\xi')\phi_0(\xi'), \tag{3.11}$$

obeying the wavelet admissibility requirement that $\phi_1 = 0$ when $\omega = 0$ [15]. We can generate an infinite family of wavelets as

$$\phi_n(\xi) = \int_0^\xi d\xi' e^{-(\xi-\xi')}\xi_{n-1}(\xi')\phi_{n-1}(\xi'), \tag{3.12}$$

Thus, we obtain the μ-wavelets as

$$\phi_n(\xi) = e^{-\xi}\frac{\xi^n}{n!} = e^{-(\sigma^2\omega^2)/2}\frac{(\frac{\sigma^2\omega^2}{2})^n}{n!}, \tag{3.13}$$

For large n,

$$\phi_n(\xi) \to \frac{1}{\sqrt{2\pi n}}e^{-(\xi-n)^2/2n}, \tag{3.14}$$

which shows that the μ-wavelets are generalized Gaussians. The μ-wavelet can also be written in the time domain as

$$\mu_n^{\lambda,\sigma}(x) = \frac{1}{\sigma(\sqrt{2^n n!\sqrt{\pi}})} H_n(x) e^{-x^2}, \tag{3.15}$$

where $x = \frac{\sqrt{\lambda}(t-t_0)}{\sigma}$, λ and σ are parameters that control the bandwidth of the wavelets, t_0 is the center-time of the wavelets, and $H_n(x)$ is the n^{th} order Hermite polynomial which is defined by

$$H_n(x) = (-1)^n e^{x^2} \frac{d^n}{dx^n} e^{-x^2}, \tag{3.16}$$

The μ-wavelets are functions that are localized in both time and frequency. We have plotted a family of μ-wavelets with $n = 2, \cdots, 30$ in Figure 3.1.

It is useful to make a few comments about the μ-wavelets. First, we stress that ϕ_0, which is chosen to be a Gaussian, is therefore not a wavelet. In wavelet theory, it corresponds rather to the basic "low pass" function, since it is equal to the Gaussian. Indeed, the work of [15] showed that sums of the μ-wavelets, plus ϕ_0, give rise to infinitely smooth approximations to the ideal window. Second, we make the observation that the ϕ_n are, in fact, "generalized Gaussians ", in the following sense. It is well-known that the Gaussian is a solution of the diffusion equation with a Dirac delta function source. It is also true that ϕ_n, $n > 0$, is a solution of the diffusion equation with the n^{th} derivative of the Dirac delta function as the source. It is this fundamental connection of the ϕ_0 and μ-wavelets, ϕ_n, $n > 0$, that motivated us to investigate them for time-frequency analysis of seismic signals.

The original μ-wavelets are zero-phase wavelets. In Figure 3.2 we plotted the zero-phase μ-wavelet of degree 10. Throughout this article we choose $\sigma = 0.005$ and

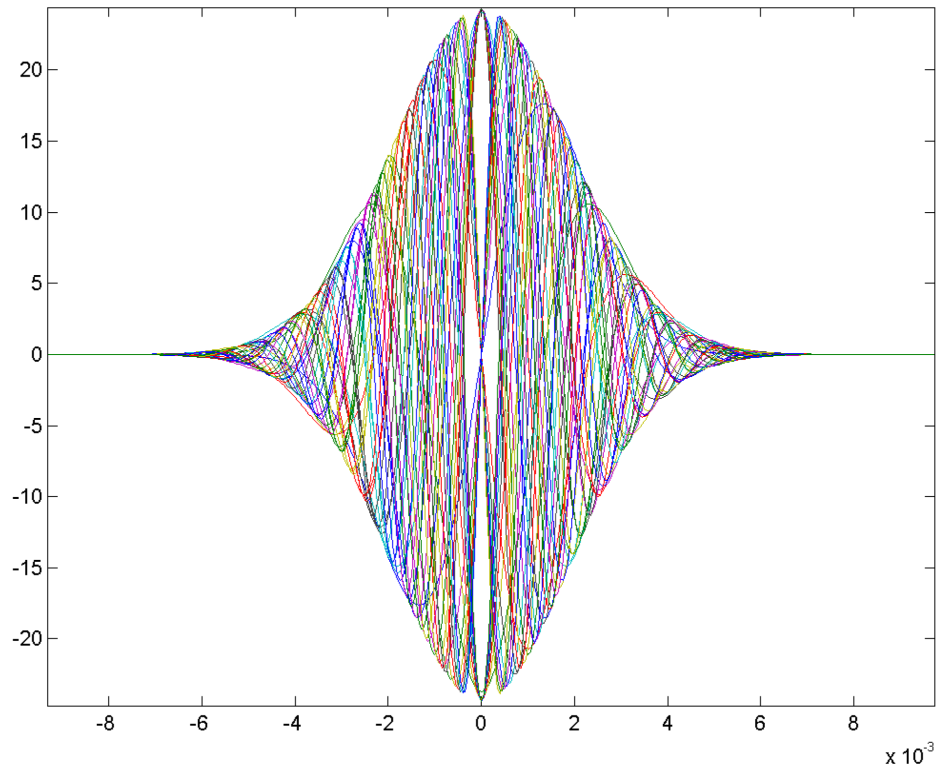


Figure 3.1: Wavelets family with $n = 2, \cdots, 30$.

$\lambda = 7$ after exploring the μ-wavelets dependence on its parameters.

As we know, the zero-phase wavelet is symmetrical with a maximum at time zero (non-causal). Often, when the trace has been decomposed using zero-phase wavelets, then there will be a small amount of energy leaks into times before the true first arrival. The minimum phase wavelet is the most front-loaded wavelet possible that is zero before time zero (causal) and has the given amplitude spectrum. If the phase is made smaller, then the wavelet becomes non causal. Of all causal wavelets with given amplitude spectrum, the phase spectrum given by the Hilbert transform

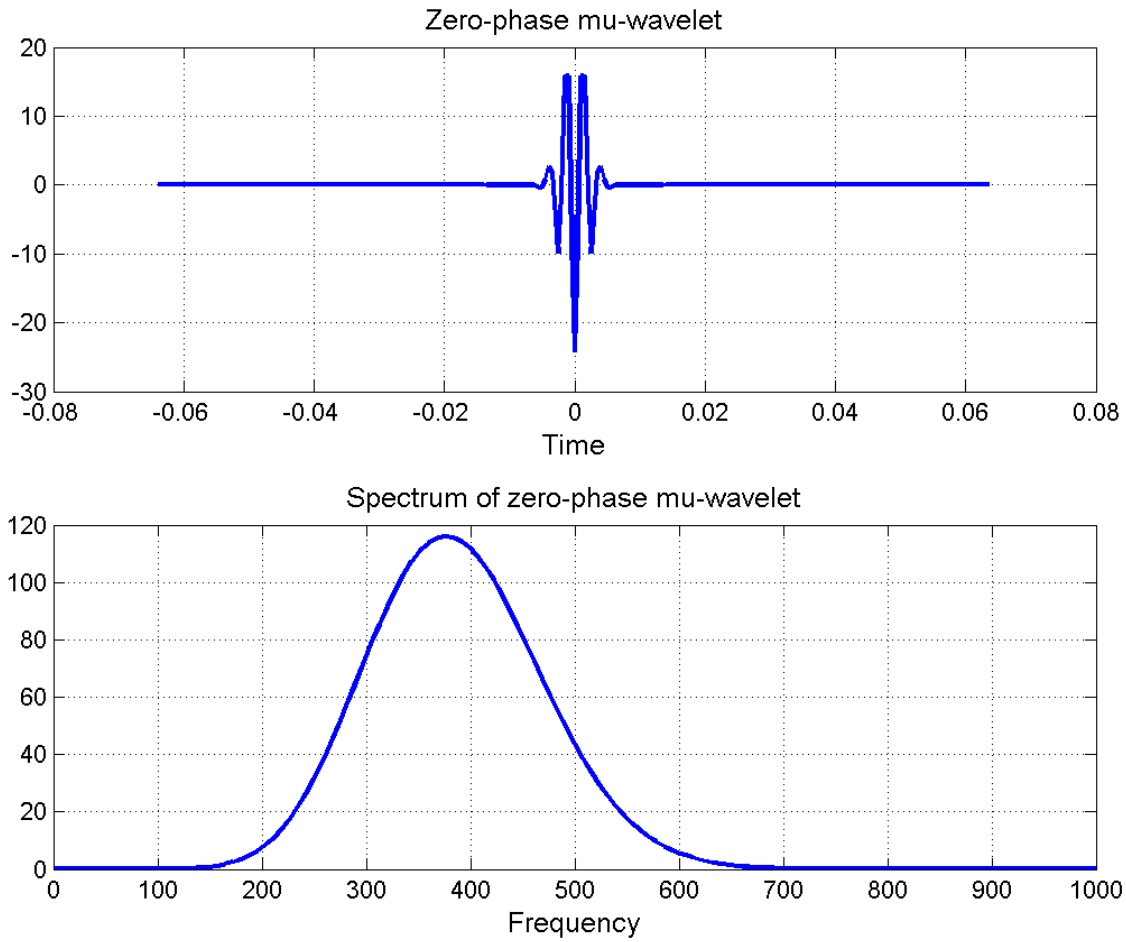


Figure 3.2: The zero-phase μ-wavelet of degree $n = 10$ (upper) and its amplitude spectrum (lower).

of the logarithm of the amplitude spectrum is the smallest in absolute value at any frequency. Therefore, we generate the minimum phase μ-wavelet with the phase spectrum as the Hilbert transform of the logarithm of the amplitude spectrum of the zero-phase μ-wavelets. In this way, we transform the zero-phase μ-wavelets into minimum-phase wavelets (see Figure 3.3 for example) and employ them to do spectral decomposition by the algorithm in section 3.1.3.

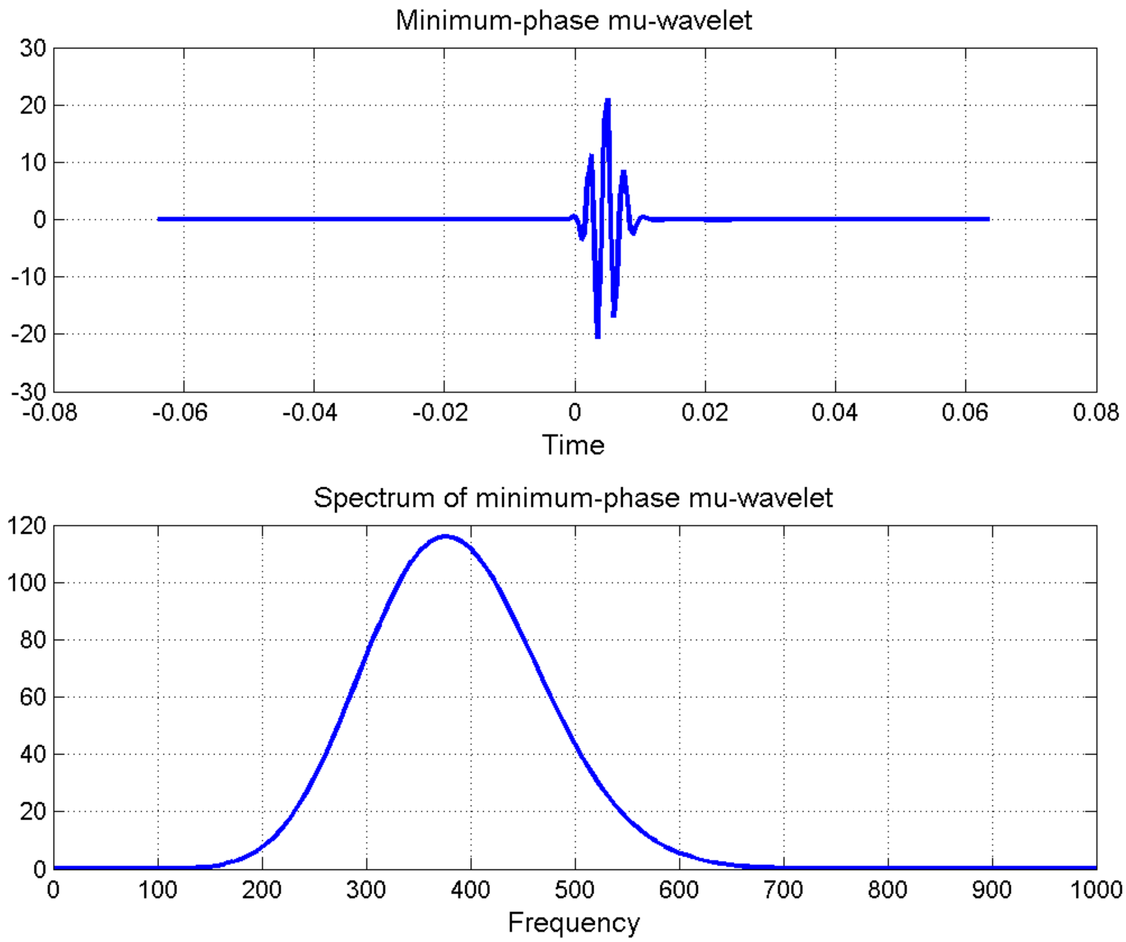


Figure 3.3: The minimum-phase μ-wavelet of degree $n = 10$ (upper) and its amplitude spectrum (lower).

3.1.2 Hermite distributed approximating functionals

In broad terms, approximating functions and their derivatives can be cast in terms of the Dirac delta function, having the following property:

$$f(x) = \int_{-\infty}^{\infty} \delta(x - x')f(x')dx', \tag{3.17}$$

We introduce the Hermite distributed approximating functionals or HDAFs.

$$\delta_M(x-x',\sigma)=\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(x-x')^2}{2\sigma^2}}\sum_{n=0}^{M/2}(-\frac{1}{4})^n\frac{1}{n!}H_{2n}(\frac{x-x'}{\sqrt{2}\sigma}), \tag{3.18}$$

It has been proved that the HDAFs yield unity approximations [3] [53]. Therefore

$$f_M(x)=\int_{-\infty}^{\infty}\delta_M(x-x')f(x')dx' \tag{3.19}$$

is a delta sequence of functions which approximates $f(x)$.

Let us consider a discrete uniform grid (x_i) of size h. By carrying out a numerical integration of (3.19), considering that the HDAF has a limited numerical support, the HDAFs approximation for $f(x)$ is:

$$f(x_i)\approx f_M(x_i)\approx\sum_{j=0}^{N}h\delta_M(x_i-x_j,\sigma)f(x_j)\approx\sum_{j=i-W/2}^{i+W/2}h\delta_M(x_i-x_j,\sigma)f(x_j), \tag{3.20}$$

where W is the HDAFs bandwidth. Next, we can formulate filling a gap in data using HDAFs [56]. Suppose we have a set of uniformly spaced grid points on the infinite line. $f(x)$ is a continuous function that is known on all grid points except for the set $x_J,\cdots,x_K$. We can estimate the unknown values by minimizing the cost function,

$$C=\sum_{p=-\infty}^{\infty}W_p(f(x_p)-f_M(x_p))^2 \tag{3.21}$$

where W_p is a weight assigned to the point x_p. In this dissertation it is chosen to be 1 on a finite grid and 0 elsewhere. We substitute the HDAFs approximation $f_M(x_p)$ in equation 3.20 into equation 3.21, we obtain

$$C=\sum_{p=-\infty}^{\infty}W_p(f(x_p)-\sum_{j=p-W/2}^{p+W/2}h\delta_M(x_p-x_t,\sigma)f(x_t))^2 \tag{3.22}$$

We minimize the cost function with respect to the unknown values, $f(x_J), \cdots, f(x_K)$ by specifying

$$\frac{\partial C}{\partial f(x_l)} = 0, J \leq l \leq K \tag{3.23}$$

to obtain the system of linear algebraic equations,

$$\sum_{p=-\infty}^{\infty} 2W_p(f(x_p) - \sum_{j=p-W/2}^{p+W/2} h\delta_M(x_p - x_t, \sigma)f(x_t))(\delta_{pl} - \delta_M(x_p - x_l, \sigma)) = 0, J \leq l \leq K \tag{3.24}$$

where the unknowns are $f(x_p)$ and $f(x_t)$ for $p = l$ or $t = l$. Here δ_{pl} is the Kronecker delta. We can obtain the predicted values of $f(x)$ on the grid points in the gap.

3.1.3 Algorithm

We next assume the signal can be expressed as a linear combination of the μ-wavelets:

$$S(t) = \sum_{n=0}^{N} C_n \mu_n(t) \tag{3.25}$$

We carry out the following steps to obtain the first arrival detection based on a time-frequency analysis.

Step 1: cross-correlate the analysis μ-wavelet $\mu_m(t)$ with both sides of equation 3.25 and obtain

$$\mu_m(t) \otimes S(t) = \mu_m(t) \otimes \sum_n C_n \mu_n(t), \tag{3.26}$$

$$\mu_m(t) \otimes S(t) = \sum_n C_n \mu_m(t) \otimes \mu_n(t), \tag{3.27}$$

where $g(t) \otimes f(t) = \int_{-\infty}^{\infty} g(t+\tau)f(t)dt$, and define

$$d_m(\tau) = \int_{-\infty}^{\infty} \mu_m(t+\tau)S(t)dt, \tag{3.28}$$

$$X_{mn}(\tau) = \int_{-\infty}^{\infty} \mu_m(t+\tau)\mu_n(t)dt, \tag{3.29}$$

where we use the Einstein summation convention for repeated indices. Thus equation 3.27 can be written in the matrix vector form as

$$d_m(\tau) = X_{mn}(\tau)C_n, \tag{3.30}$$

Thus we can get the coefficients C_n by

$$C_n(\tau) = X^+_{mn}(\tau)d_m(\tau), \tag{3.31}$$

where X^+_{mn} is the pseudo-inverse of X_{mn} which can be obtained by singular value decomposition.

Step 2: We obtain the time-frequency representation of the signal by multiplying the coefficients with the Fourier transform of the μ-wavelets

$$S(\tau,\omega) = \sum_n C_n(\tau)\mu_n(\omega), \tag{3.32}$$

where $\mu_n(\omega) = FT(\mu_n(t))$.

Step 3: We define the power spectrum $P(\tau,\omega) = S^2(\tau,\omega)$, of the time-frequency form of the signal and then integrate it over frequency at various times τ to define the first break indicator function, $f(\tau)$:

$$f(\tau) = \int S^2(\tau,\omega)d\omega. \tag{3.33}$$

Step 4: We apply a peak-picking algorithm to find the largest two peaks, and denote them as identifiers for the P-wave and S-wave first arrivals. We specify a potential region of the first arrivals for each of P-wave and S-wave. In each potential region, we neglect every other point and then apply Hermite distributed approximating functionals (HDAFs) to fill the neglected points [56]. This yields a less noisy indicator function at the potential region. Next we find the maximum of all the points in each potential region as our indicator. Thus we obtain two indicators, the smaller one for P-wave first arrival and the other one for S-wave arrival.

3.2 Examples

We validated our method using both lab data with various S/N and on field micro-seismic data.

3.2.1 Test on lab data

The experimental data was obtained for an S-wave source. The original sampling frequency was 50 MHz. We re-sampled the data with sampling frequency of 5 MHz for our test. The seismic trace was gathered by stacking 64 signals in Figures 3.4 and 3.5. But the seismic trace in Figure 3.6 was not stacked. The signals in Figures 3.5 and 3.6 are noisier than that of Figure 3.4, as we can see from both the original signals and the indicator functions. Our automatic time arrival detections for the three signals are very close to each other and to the actual arrival time of S-wave which is 0.22238 ms. Our automatic time arrival detection has error of less than 0.5 μs as shown in the figures. This illustrates the accuracy of our method.

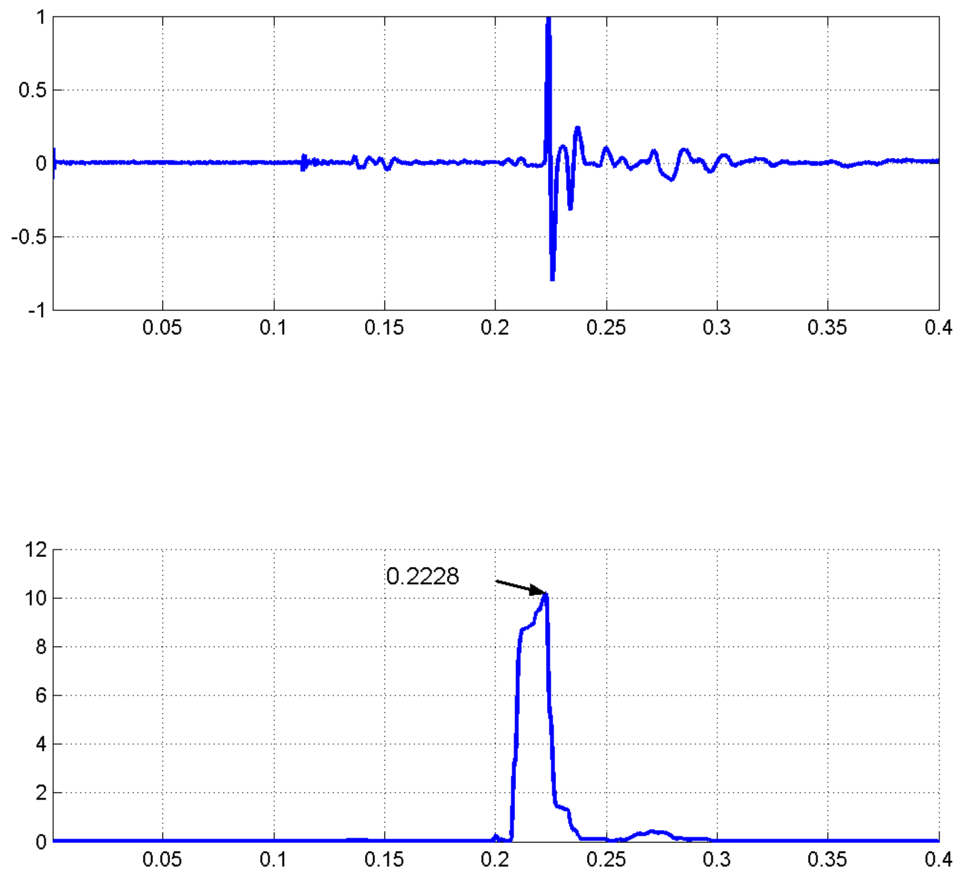


Figure 3.4: The zero-phase μ-wavelet of degree $n = 10$ (upper) and its amplitude spectrum (lower).The upper picture shows the stacked seismic trace of gain +30 dB, the lower one shows the first arrival indicator function. Our method detects the S-wave arrival time $t_s = 0.2228$ ms.

3.2.2 Test on field data

The field data we tested on have a sampling frequency of 2 KHz. Our data are for three components and was recorded at 8 stations. As shown in Figures 3.7 to 3.9, we detected both P-wave and S-wave arrival times since we have two obvious peaks in the indicator function. The first peak detects the P-wave arrival and the second peak detects the S-wave arrival respectively. The time arrivals we detected by each of the

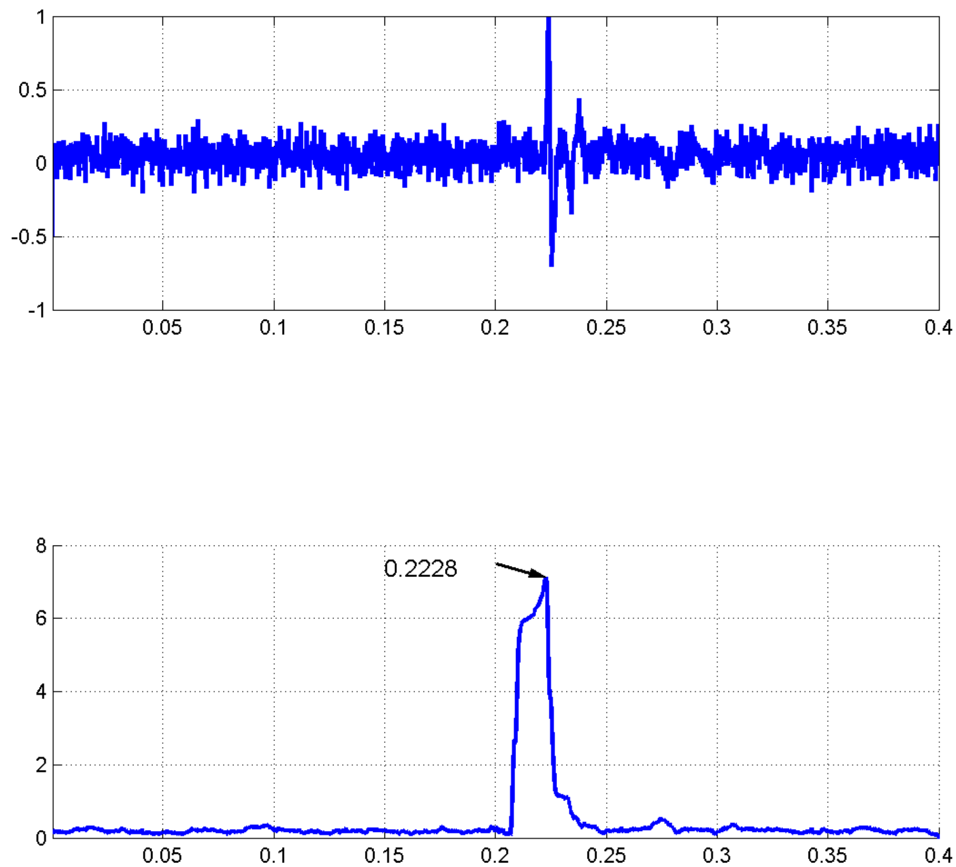


Figure 3.5: The upper picture shows the stacked seismic trace of gain 0 dB, the lower one shows the first arrival indicator function. Our method detects the S-wave arrival time $t_s = 0.2228$ ms.

three components are slightly different, so we calculated the average to obtain the arrival time for the station. We found that the Z-component and NS-component are better behaved to detect P-wave arrival, while the SW-component is better behaved to detect S-wave arrival.

We compare the result of our method with the manual time arrival picking in Table 3.1. Our detection of the arrival time is generally consistent with the manual detection. The differences are within 1 ms.

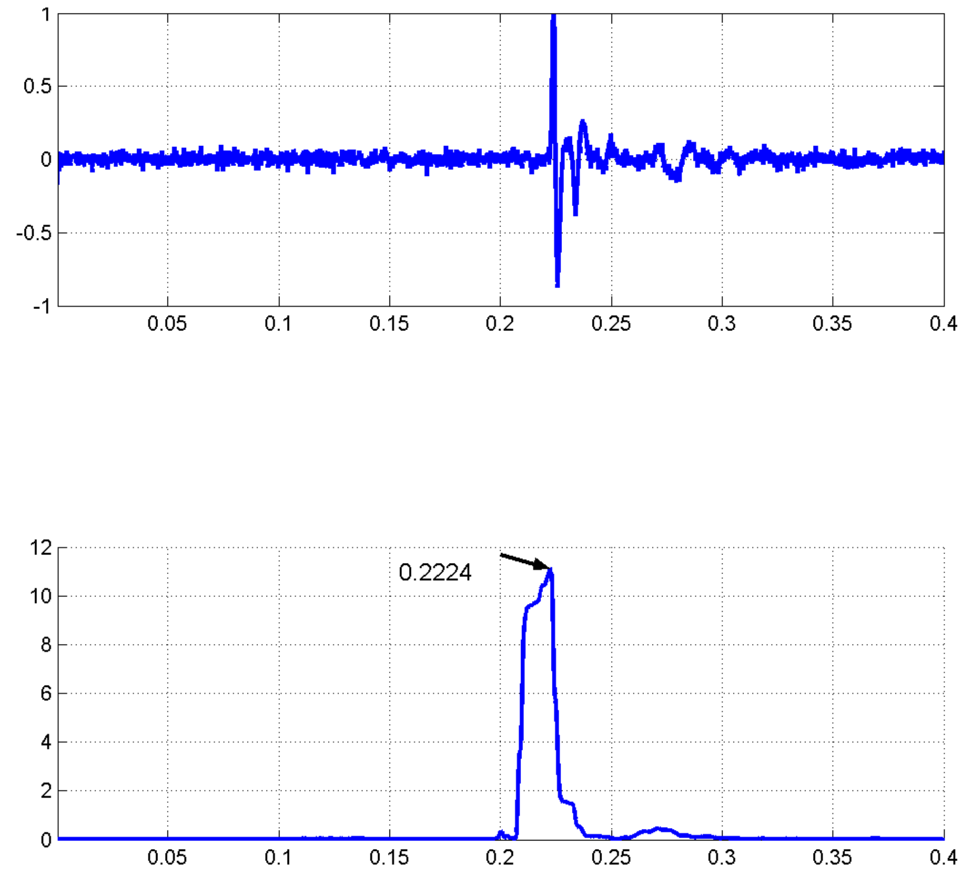


Figure 3.6: The upper picture shows one seismic trace of gain +30 dB, the lower one shows the first arrival indicator function. Our method detects the S-wave arrival time $t_s = 0.2224$ ms.

We tested on 440 microseismic traces and plot the difference of the manually detected S-wave arrivals and our automated picked S-wave arrivals using a histogram in Figure 3.10. The width of each bin is $0.5ms$ which is the sample rate. The mean of the difference is $0.16ms$.

We have proposed a new method of automatic first break detection of P-waves and S-waves. Our method is based on a time-frequency analysis of the seismic trace using minimum uncertainty (μ-) wavelets. The performance of our method has been

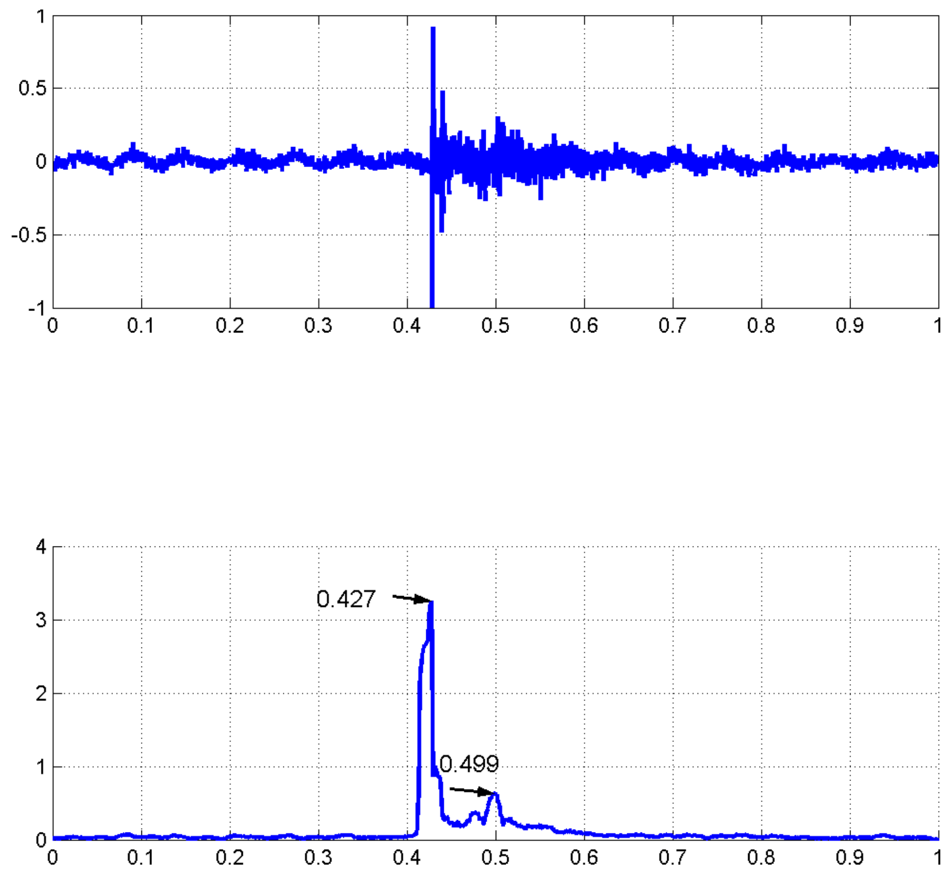
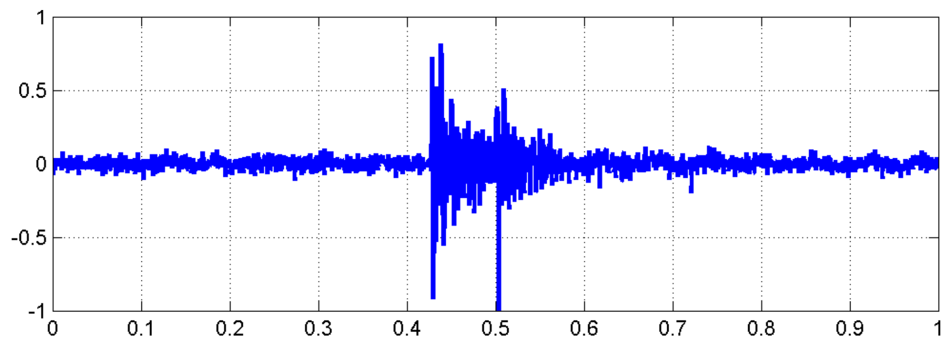


Figure 3.7: The upper picture shows the real passive seismic trace Z-component recorded at station 1, the lower one shows the first arrival indicator function. Our method detects that P-wave arrives at $t_p = 0.427$ s and the S-wave arrives at $t_s = 0.499$ s.

tested on both lab data and field data. The results suggest that our method is robust.



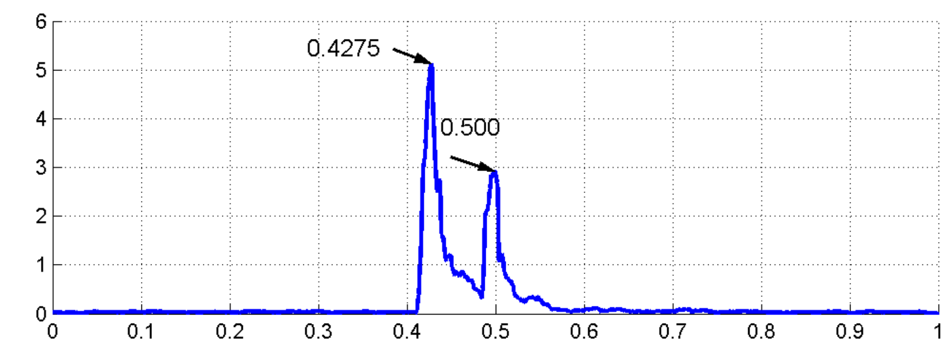


Figure 3.8: The upper picture shows the real passive seismic trace in NS-component recorded at station 1, the lower one shows the first arrival indicator function. Our method detects that P-wave arrives at $t_p = 0.4275$ s and the S-wave arrives at $t_s = 0.500$ s.

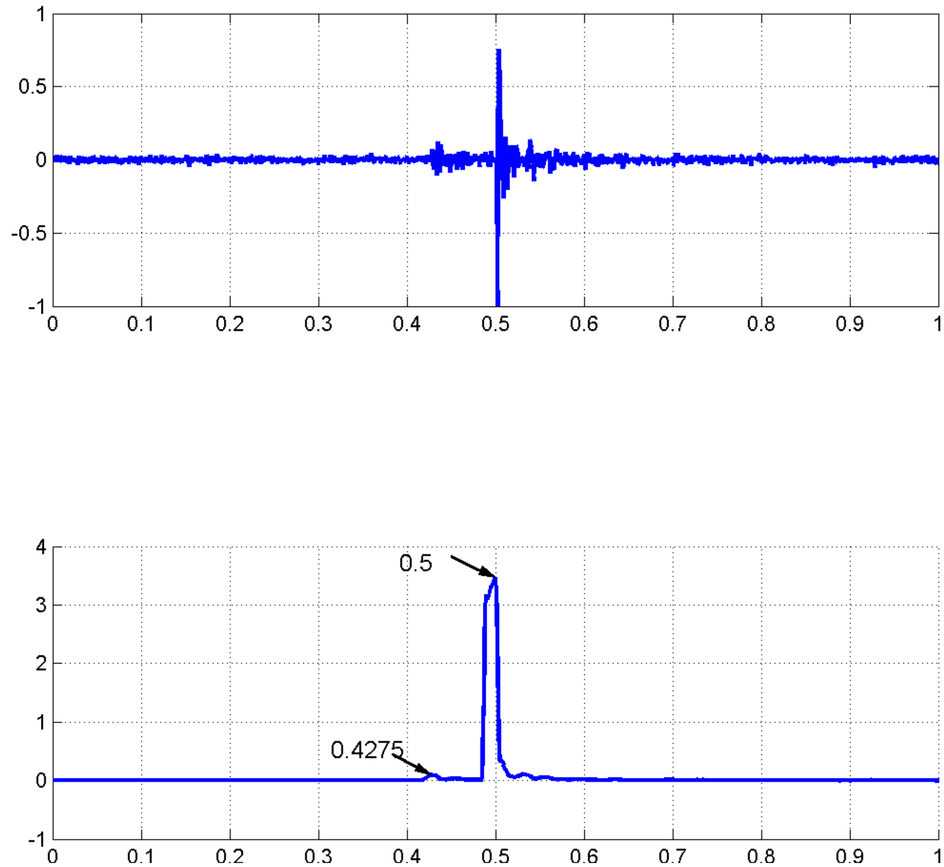


Figure 3.9: The upper picture shows the real passive seismic trace in EW-component recorded at station 8, the lower one shows the first arrival indicator function. Our method detects that P-wave arrives at $t_p = 0.4275$ s and the S-wave arrives at $t_s = 0.500$ s.

	P-wave arrival time (s)		S-wave arrival time (s)	
Station	Our Method	Manual	Our Method	Manual
1	0.4273	0.427	0.4997	0.5005
1	0.421	0.4215	0.490	0.4905
3	0.417	0.417	0.482	0.4825
4	0.4145	0.415	0.477	0.4775
5	0.412	0.4125	0.471	0.4715
6	0.4075	0.4085	0.463	0.464
7	0.4055	0.405	0.458	0.458
8	0.400	0.4005	0.452	0.452

Table 3.1: Comparison our automatic method and manually detection.

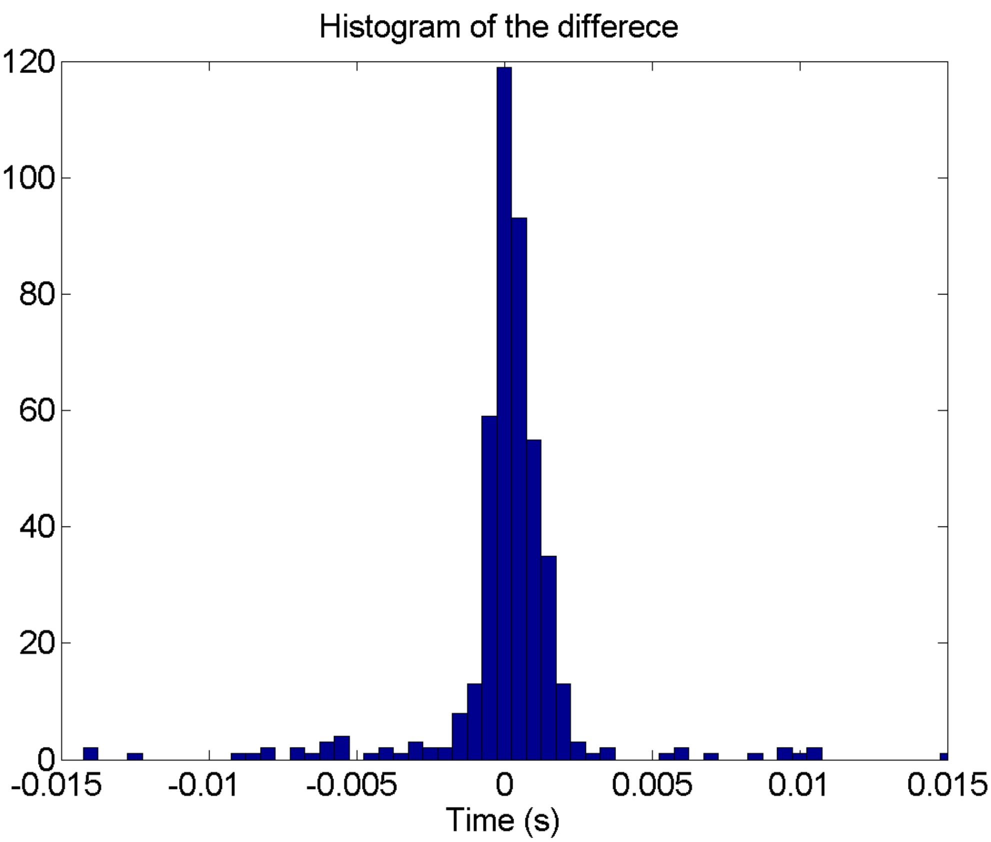


Figure 3.10: Histogram of the difference.

CHAPTER 4

Seismic event picking

We explored methods of automatic seismic event picking. It is known that no single automatic seismic event indicator works for all data; therefore, we explored two indicators based on μ-wavelets and on an energy ratio. Thresholding was applied to pick seismic events. We have tested with synthetic data and offshore field data.

4.1 μ-Wavelets-based seismic event picking

We applied μ-wavelets-based seismic event indicator as we introduced in the previous chapter on the synthetic data set BP2004. The dataset was generated using a 2D time-domain, acoustic, finite difference modeling algorithm. It was modeled with a streamer configuration, using a 15km streamer with 12.5m group interval and a 50m shot interval. Minimum offset is 0m, and sampling interval is 6ms. A total of 1340 shots were generated each with 1201 receivers. Shot 1 is shown in figure 4.1. We applied the indicator on each trace and picked the first arrivals which were shown in black in Figure 4.2. Shot 15 is shown in figure 4.3. We applied the indicator on each trace and picked the first arrivals which were shown in black in Figure 4.4. We have good picking results but each shot took $2069s$ on a work station. Its efficiency is not satisfactory because of the computational complexity of the spectral decomposition. We also explored an energy ratio method which is less computationally complex.

4.2 Energy ratio-based seismic event picking

In this section, we introduced a short-term average/long-term average (STA/LTA) ratio [54], applied trace by trace, and followed by thresholding to mute. We have added various levels of white noise to the synthetic shots and then picked seismic events using the energy ratio method for testing.

Let x_i be the time series representing a seismic trace. We define the energy ratio as the short-term average/long-term average ratio as in [54]. Let the length of a short-term window be N_s, and the length of the long-term window be N_l ($N_l > N_s$). For each time index $i (1 < i < N)$, we define short-term average of the energy as

$$STA_i = \frac{1}{N_s} \sum_{j=i-N_s-1}^{i} x_j^2$$

And long-term average of the energy is defined as

$$LTA_i = \frac{1}{N_l} \sum_{j=i-N_l-1}^{i} x_j^2$$

If $j \leq 0$,set $x_j = x_1$. Then we define STA/LTA ratio as

$$r_i = \frac{STA_i}{LTA_i}$$

When we apply this to shot records, we compute STA/LTA ratio trace by trace and then apply thresholding to mute. We generated the shot record by forward modeling using finite differences. The sampling interval is 8ms with a total of 1876 samples. One shot has 801 traces. We have added various levels of white noise to the shot and then pick seismic events using the energy ratio method. We have plotted a near offset trace (#600) with low noise added at the top of Figure 4.5a. We accurately picked the first arrival using the energy ratio indicator and muted the noise before the first arrival as shown at the bottom of Figure 4.5a. We plotted a far offset trace (#200) with low noise added at the top of Figure 4.5b and the corresponding accurately muted trace is shown at the bottom of Figure 4.5a. For the entire shot record with low noise, as shown in Figure 4.6, we have correctly muted the noise before the first arrival by the energy ratio method in Figure 4.7.

We plotted a near offset trace (#600) with medium noise added at the top of Figure 4.8a. We accurately picked the first arrival using energy ratio indicator and muted the noise before the first arrival as shown at the bottom of Figure 4.8a. We plotted a far offset trace (#200) with medium noise added at the top of Figure 4.8b and the corresponding accurately muted trace is shown at the bottom of Figure 4.8a. For the entire shot record with medium noise, as shown in Figure 4.9, we have correctly muted the noise before the first arrival by the energy ratio method in Figure 4.10.

We plotted a near offset trace (#600) with high noise added at the top of Figure 4.11a. We accurately picked the first arrival using energy ratio indicator and muted the noise before the first arrival as shown at the bottom of Figure 4.11a. We plotted a far offset trace (#100) with high noise added at the top of Figure 4.11b and the corresponding accurately muted trace is shown at the bottom of Figure 4.11a. For the entire shot record with high noise, as shown in Figure 4.12, and we have correctly muted the noise before the first arrival by the energy ratio method in Figure 4.13.

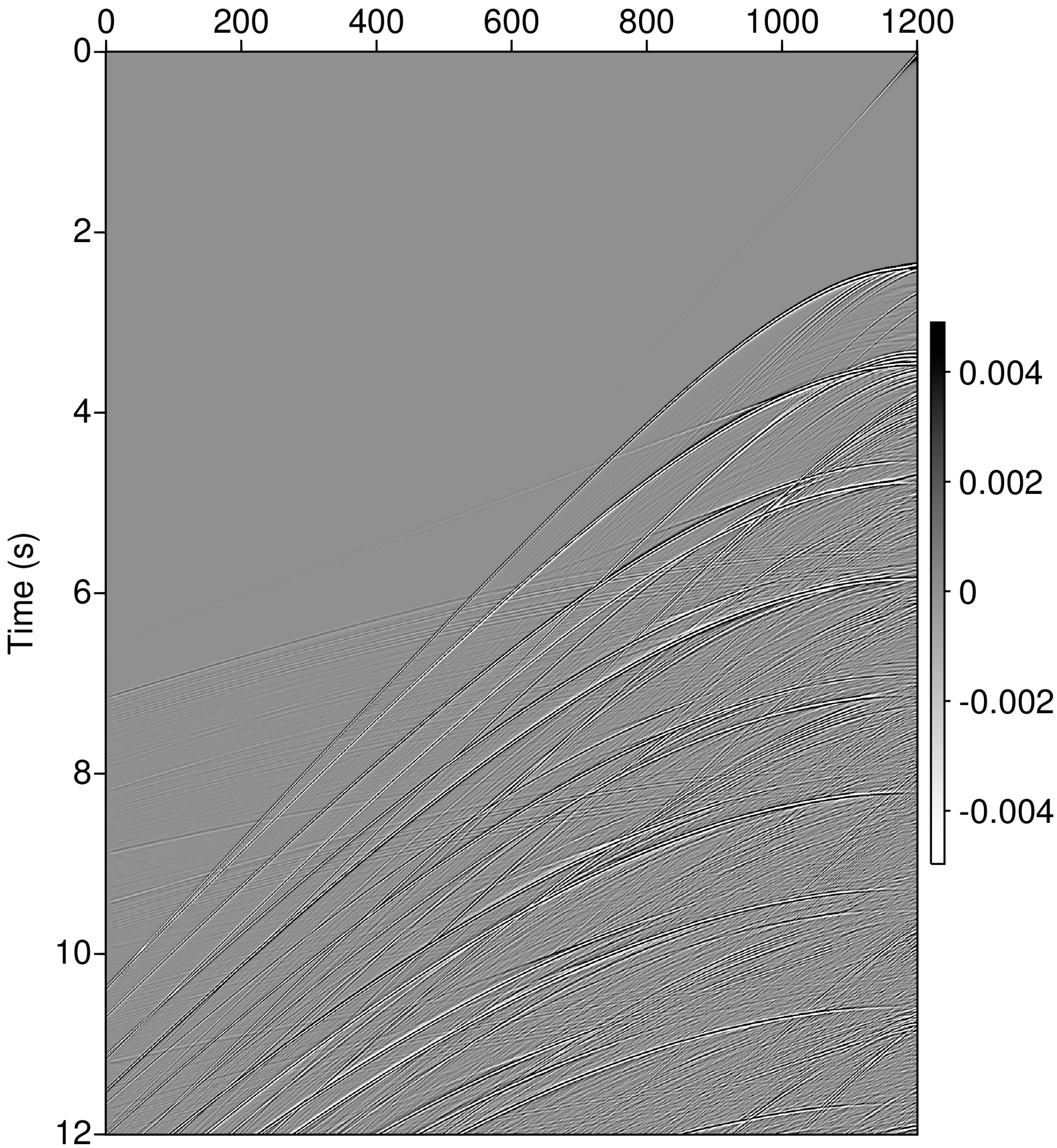


Figure 4.1: Shot 1.

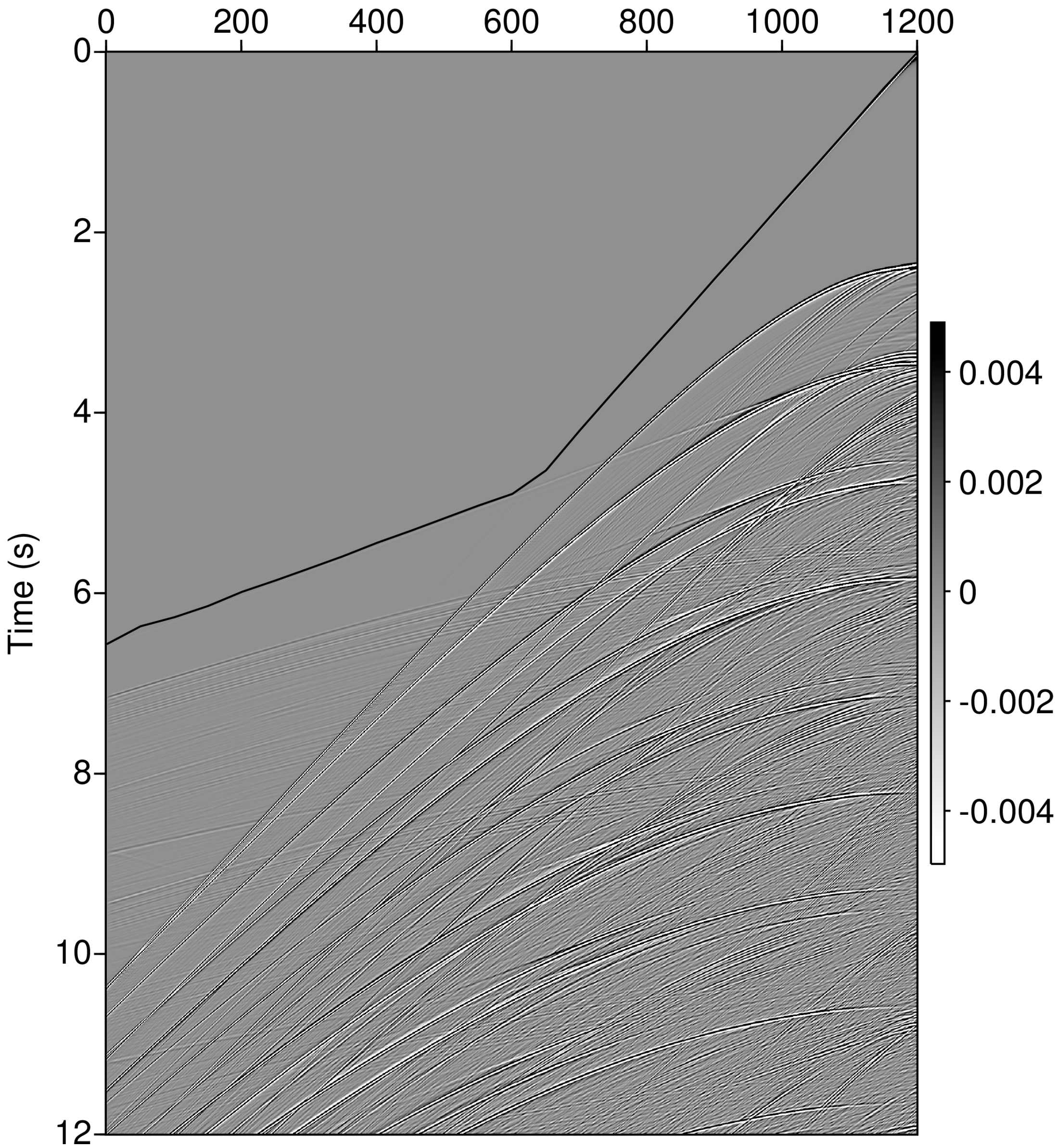


Figure 4.2: Shot 1 automatic picking.

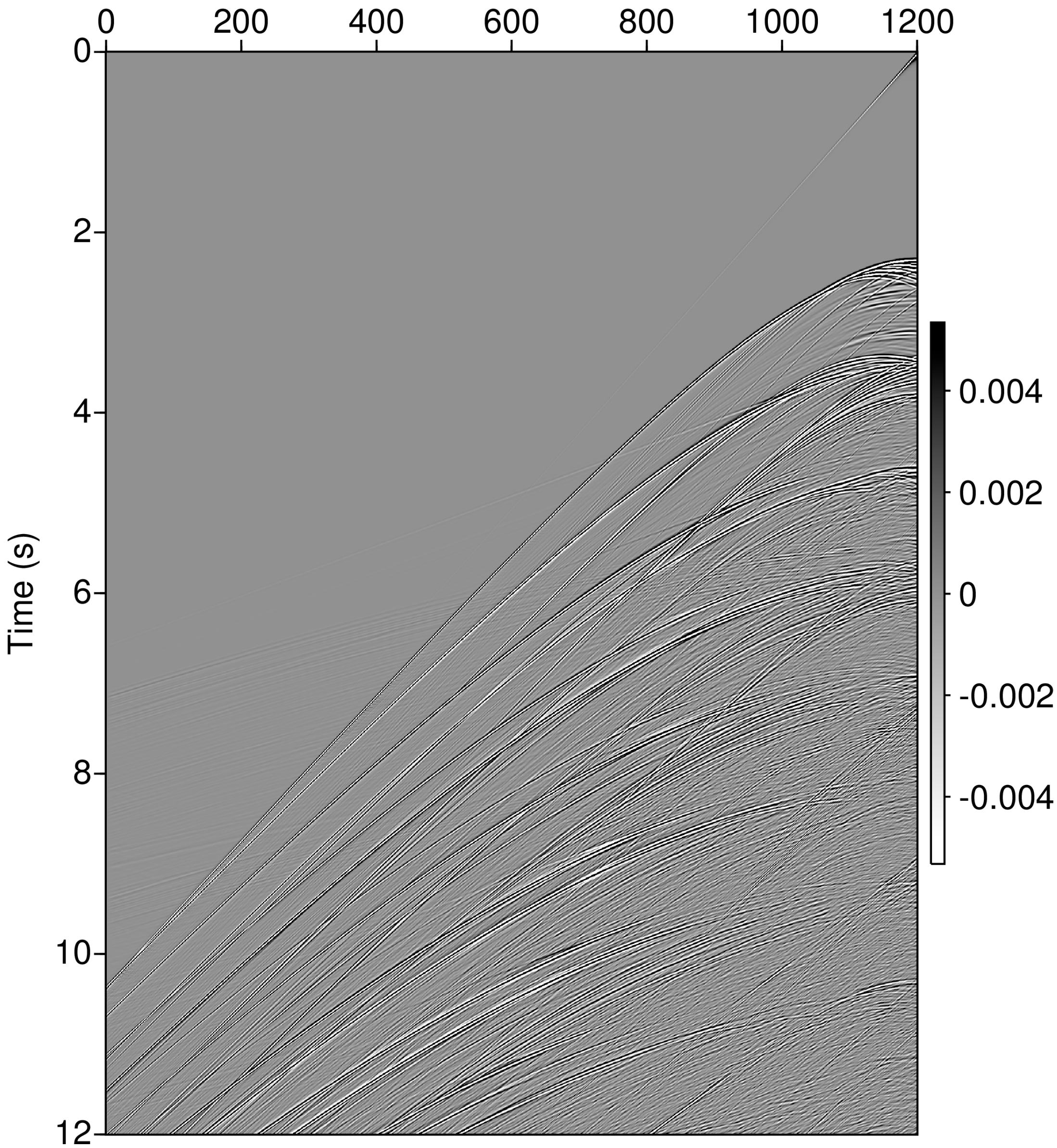


Figure 4.3: Shot 15.

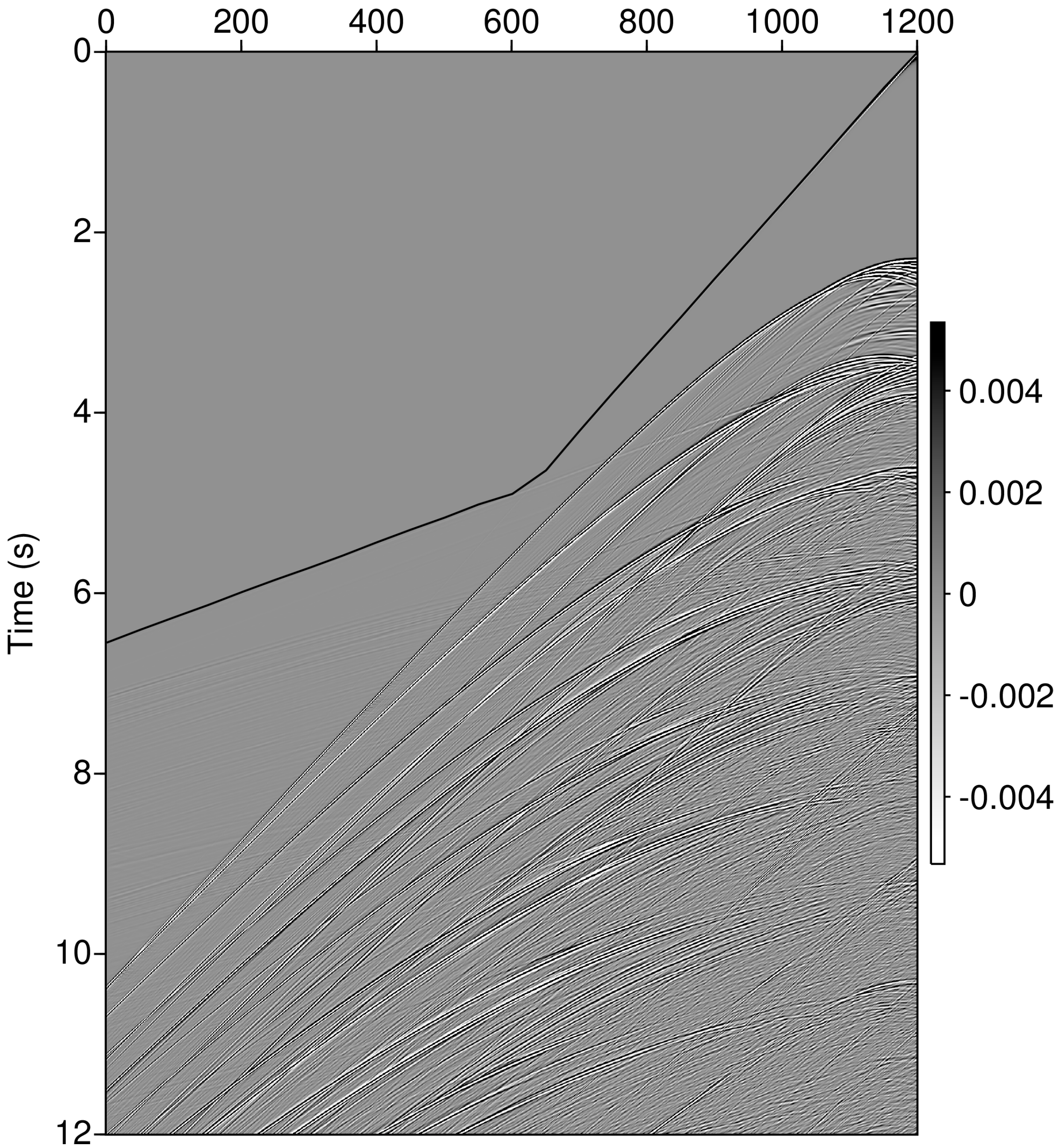
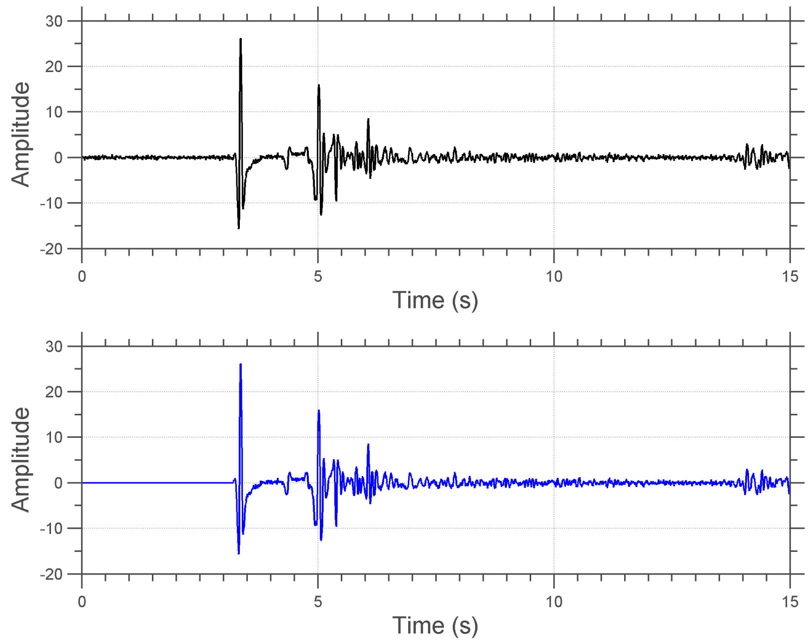
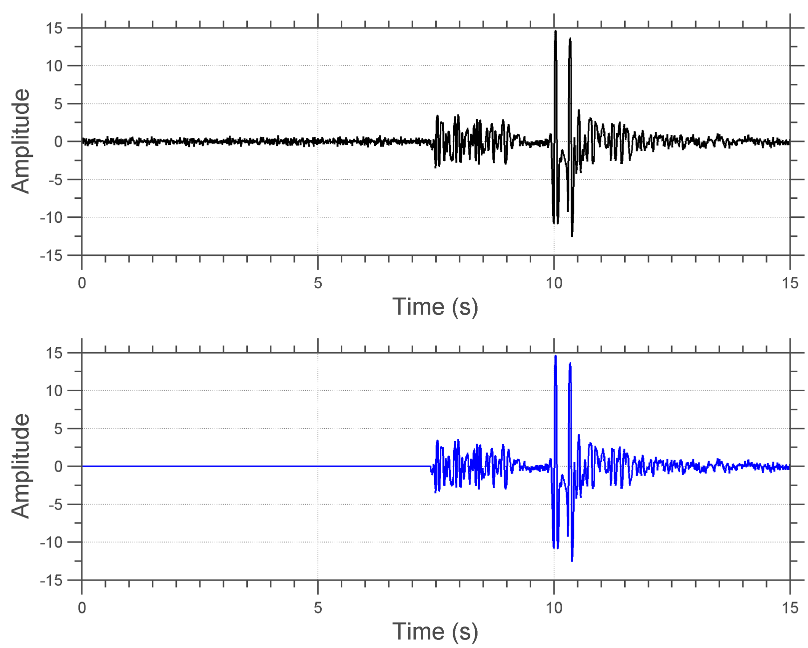


Figure 4.4: Shot 15 automatic picking.



(a) Synthetic trace #600 is at the top and its corresponding accurately muted trace at the bottom.



(b) Synthetic trace #200 is at the top and its corresponding accurately muted trace at the bottom.

Figure 4.5: Synthetic traces with low noise.

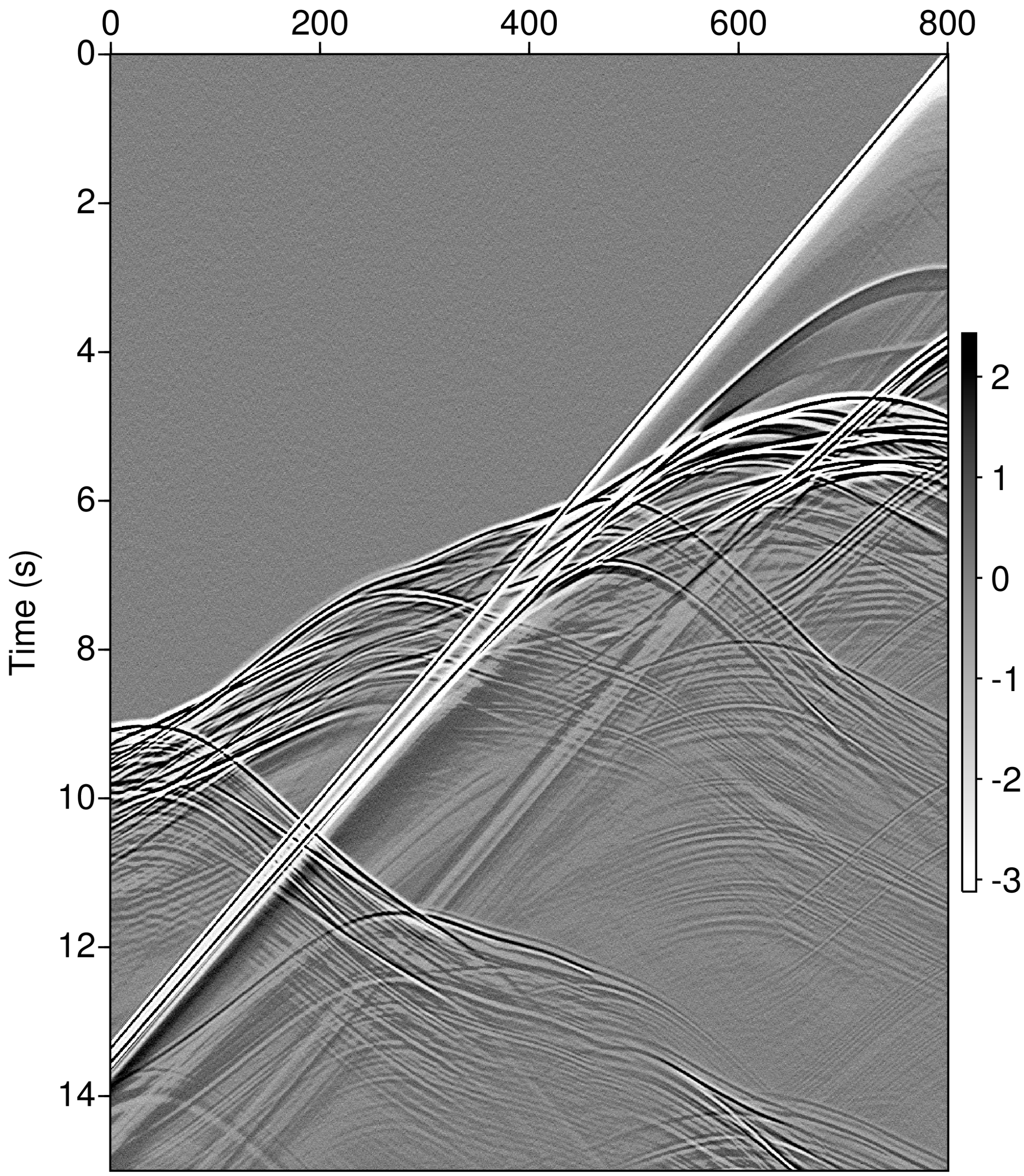


Figure 4.6: Synthetic shot record with low noise.

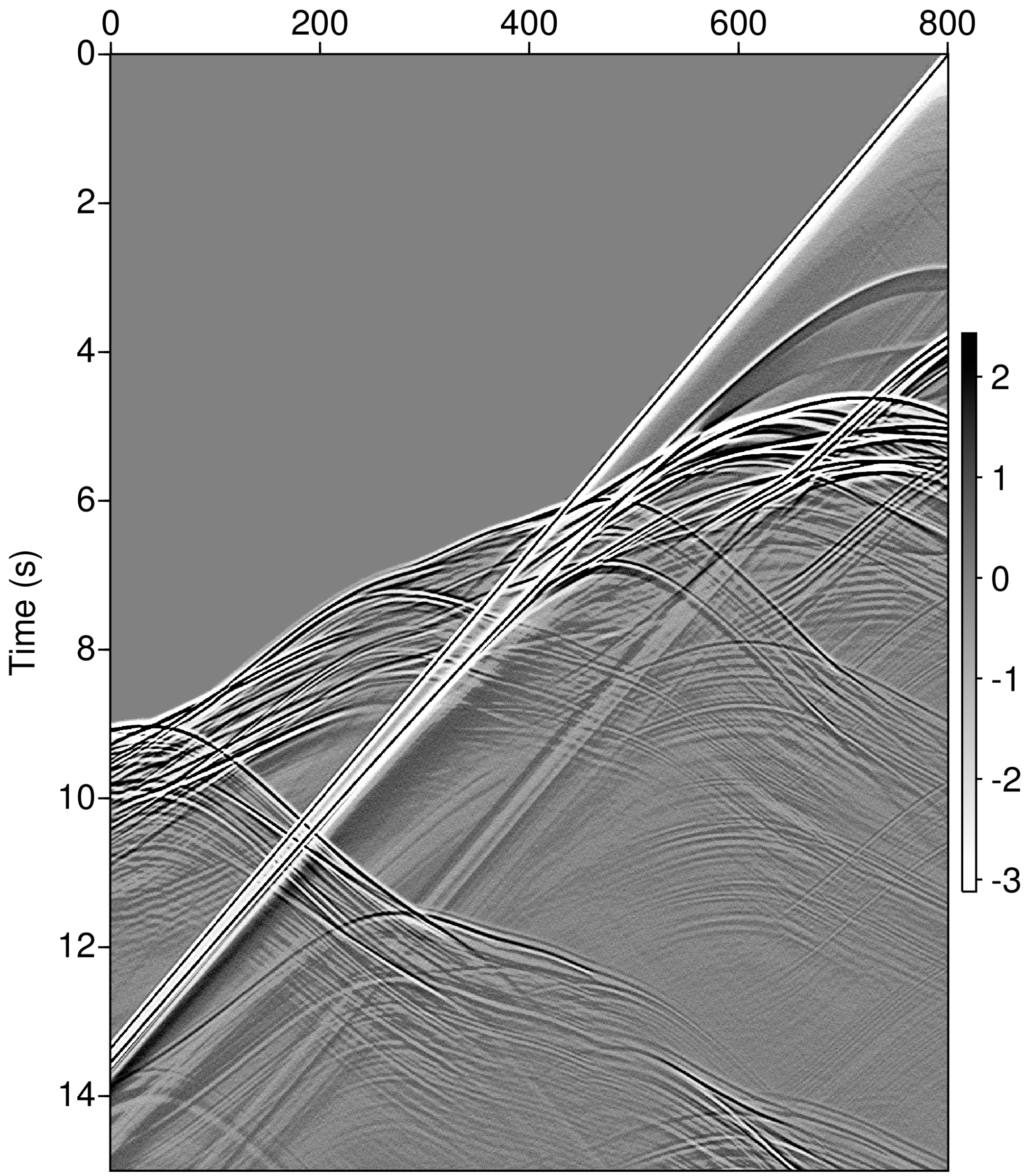
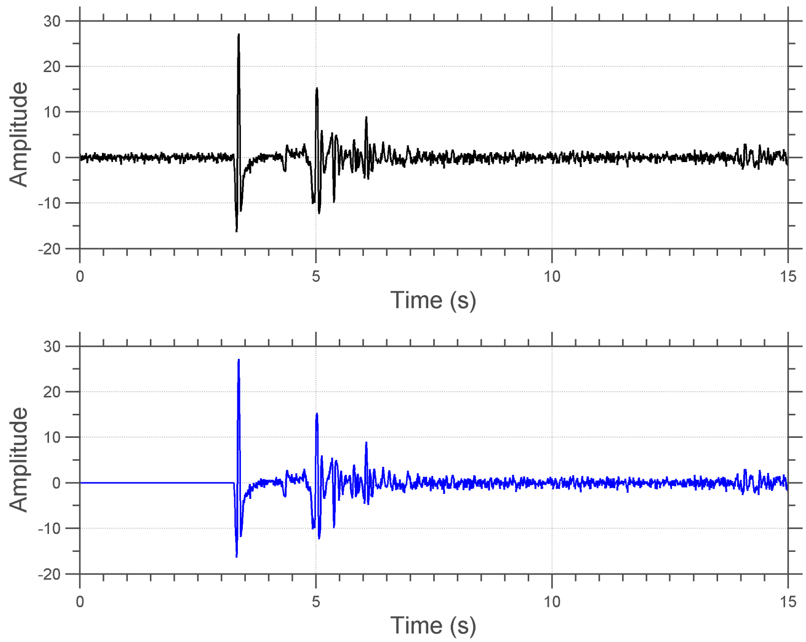
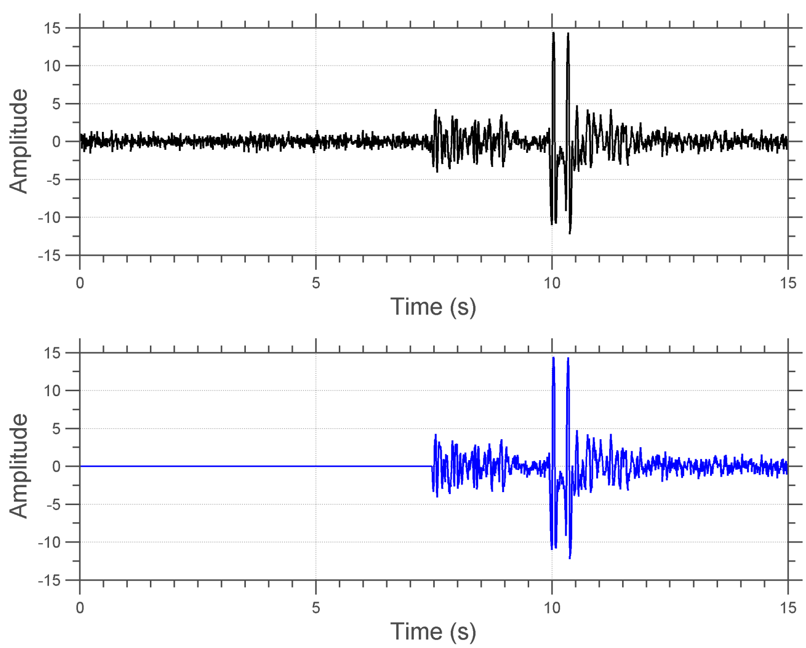


Figure 4.7: Accurately muted synthetic shot record with low noise by the energy ratio method.



(a) Synthetic trace #600 is at the top and its corresponding accurately muted trace at the bottom.



(b) Synthetic trace #200 is at the top and its corresponding accurately muted trace at the bottom.

Figure 4.8: Synthetic traces with medium noise.

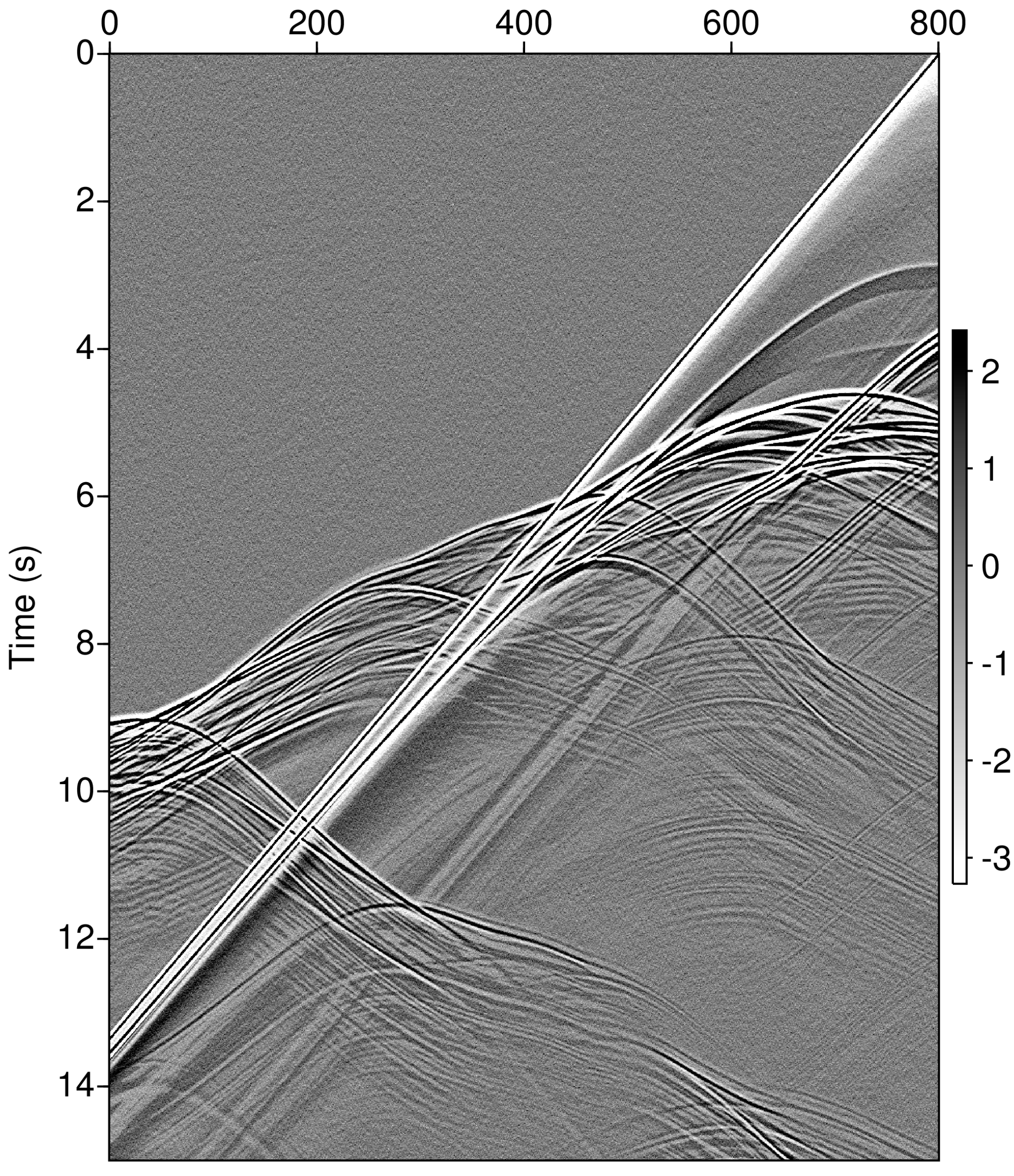


Figure 4.9: Synthetic shot record with medium noise.

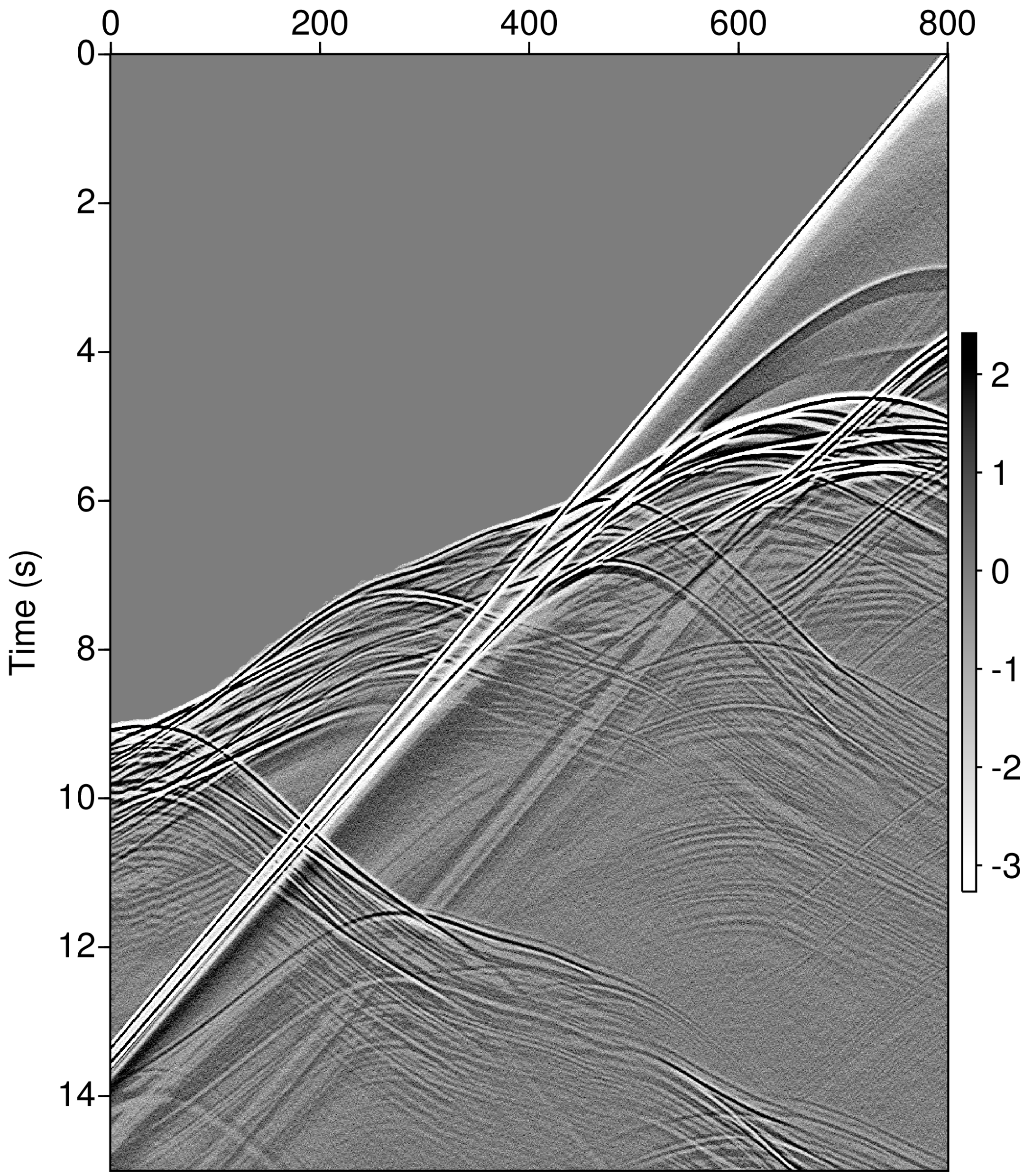
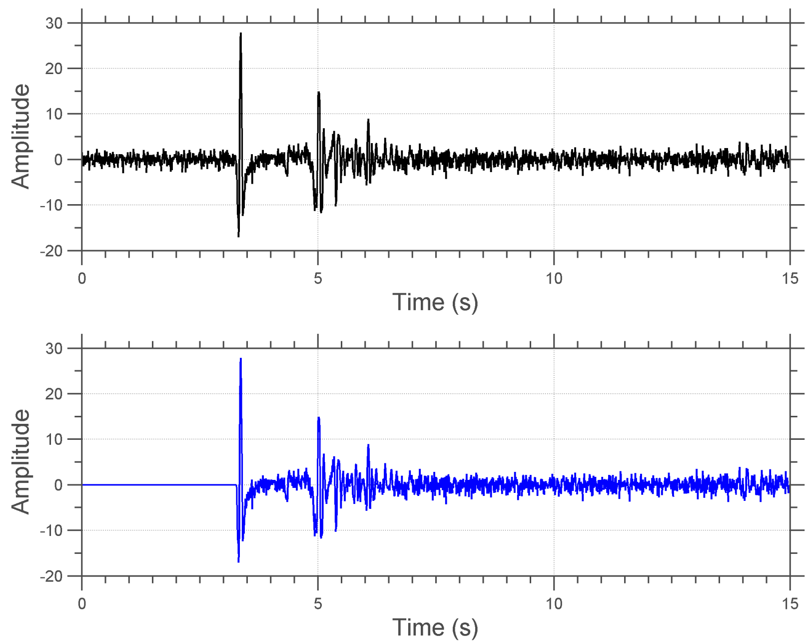
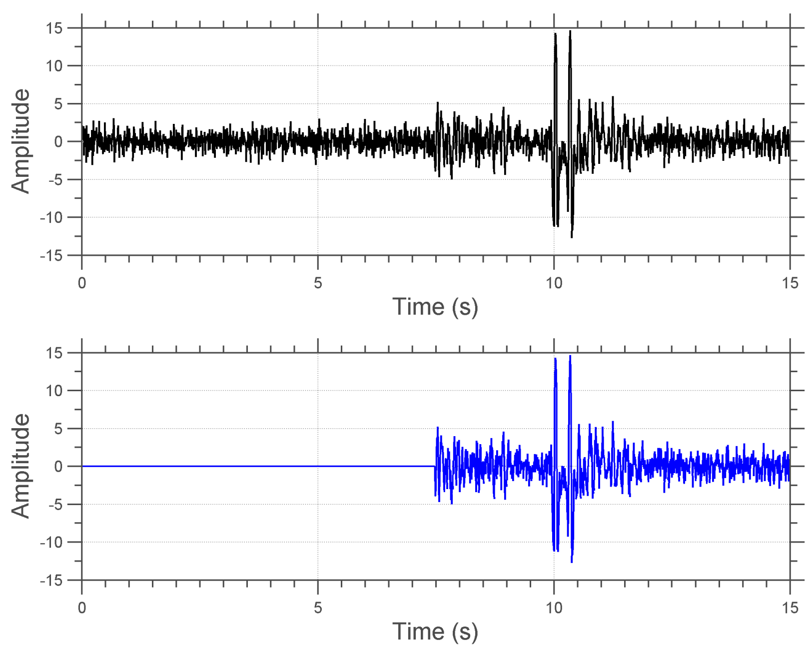


Figure 4.10: Accurately muted synthetic shot record with medium noise by the energy ratio method.



(a) Synthetic trace #600 is at the top and its corresponding accurately muted trace at the bottom.



(b) Synthetic trace #200 is at the top and its corresponding accurately muted trace at the bottom.

Figure 4.11: Synthetic traces with high noise.

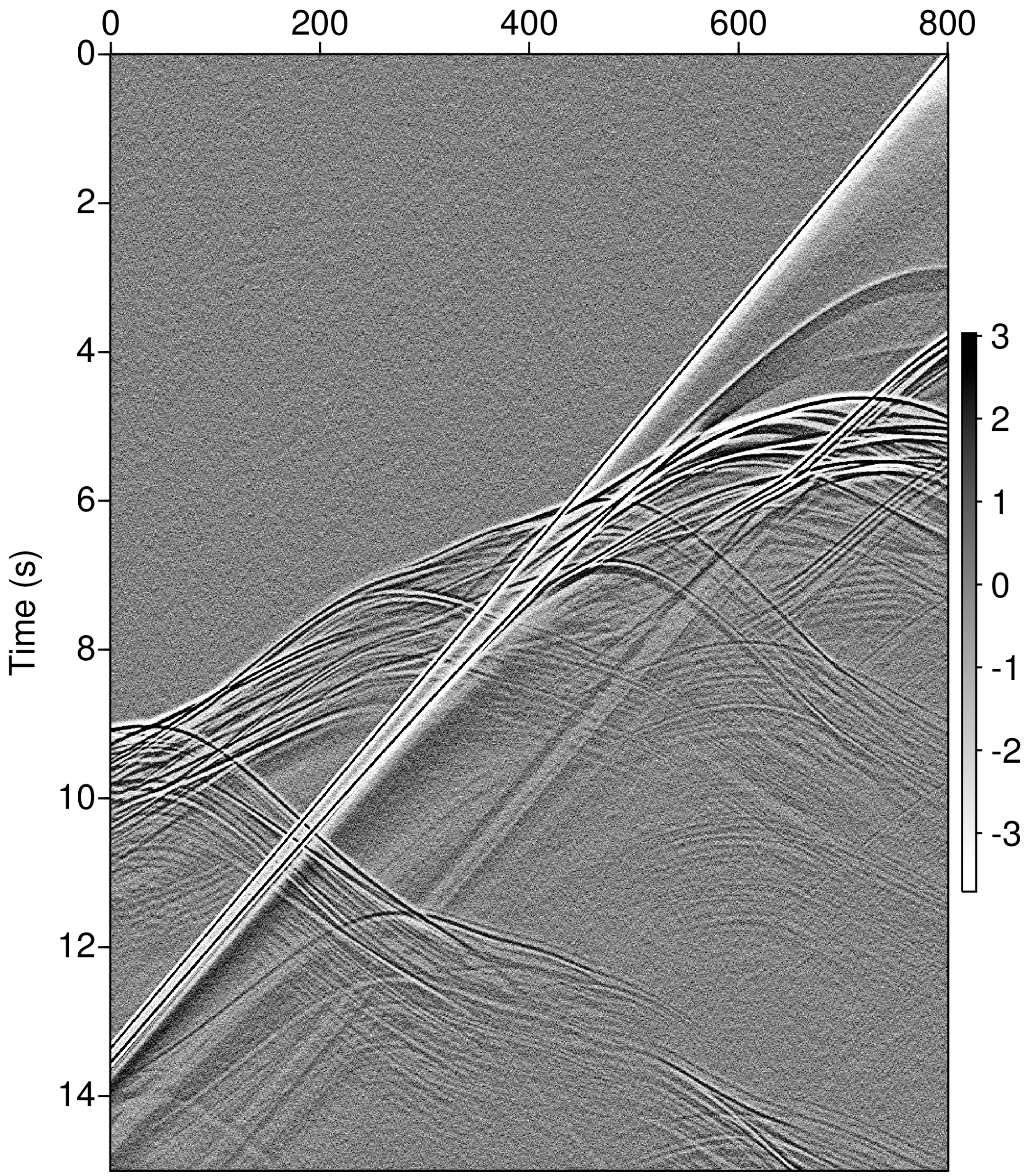


Figure 4.12: Synthetic shot record with high noise.

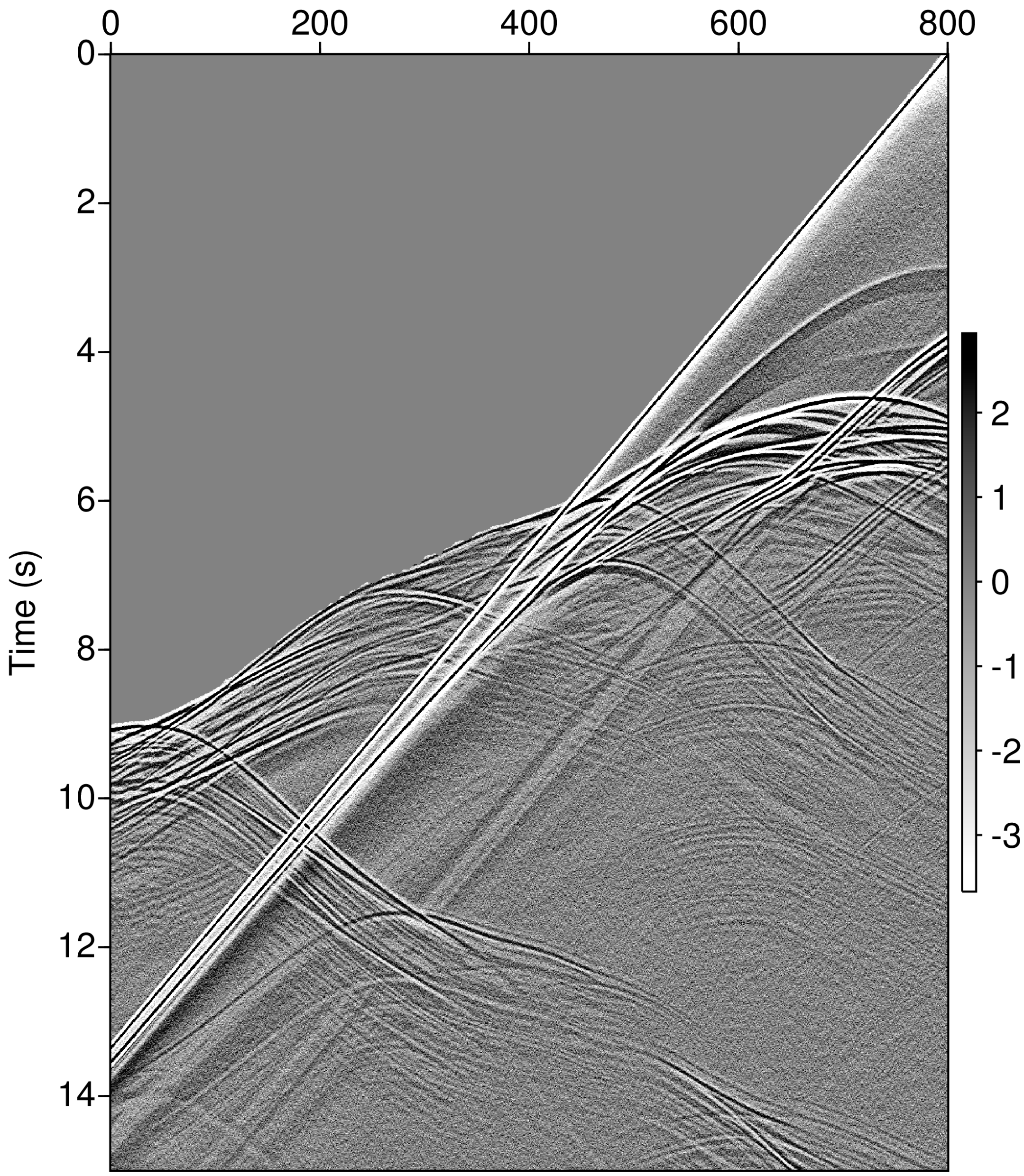


Figure 4.13: Accurately muted synthetic shot record with high noise by the energy ratio method.

CHAPTER 5

Seismic reflectivity inversion

We propose two new seismic sparse inversion methods based on complex basis pursuit and a modified basis pursuit. In practice, constant phase wavelets are used for seismic inversion algorithms (e.g the basis pursuit method). If the phase of the estimated wavelet is wrong, this can cause an error in reflectivity. We can obtain more accurate reflectivity even though the estimated wavelet has biased phase using complex basis pursuit and modified complex basis pursuit. We have tested the new approaches on a wedge model and the results are much more accurate than the standard basis pursuit method.

5.1 Basis pursuit

In this section, we introduce a seismic reflectivity inversion method using the basis pursuit method. A wedge model was generated for demonstration. We discuss the characteristics and limitations of the basis pursuit method for inverting reflectivity.

In the case of a discretized time series, the convolutional model 1.1 can be written as [57]

$$\mathbf{A}\mathbf{x} + \mathbf{e} = \mathbf{y} \tag{5.1}$$

where $\mathbf{A} \in \mathbb{R}^{N \times N}$ is the diagonal wavelet matrix whose columns consist of the time-shifted wavelets, $\mathbf{x} \in \mathbb{R}^N$ is the reflectivity series column vector with $P(P < N)$ nonzero elements., $\mathbf{y} \in \mathbb{R}^N$ is a column vector representing the seismogram, and $\mathbf{e} \in \mathbb{R}^N$ is a noise vector.

Let the wavelet vector have length of $K = 2M + 1$ and denote $w_k (1 \leq k \leq K)$ as the k-th element. Let $\mathbf{A}_{ij}$ be the element in the i-th row and j-th column of $\mathbf{A}$, then we have

$$\mathbf{A}_{ij} = \begin{cases} w_{i-(j-M)}, & \text{if } -M \leq i - j + 1 \leq M \\ 0, & \text{otherwise} \end{cases} \tag{5.2}$$

The columns of $\mathbf{A}$ are not linearly independent so that $rank(A) < N$. Equation 5.1 is an underdetermined linear system. Minimizing a quadratic loss

$$\|\mathbf{e}\|_2 = \|\mathbf{A}\mathbf{x} - \mathbf{y}\|_2 \tag{5.3}$$

can lead to over-fit. Here $\|\mathbf{x}\|_2 = (\sum_i x_i^2)^{\frac{1}{2}}$ denotes the l_2 norm of $\mathbf{x}$. One of the techniques to prevent over-fitting is Tikhonov regularization, which can be written as [30]

$$\min_{\mathbf{x}} \|\mathbf{A}\mathbf{x} - \mathbf{y}\|_2 + \lambda\|\mathbf{x}\|_2 \tag{5.4}$$

where $\lambda > 0$ is a regularization parameter. The Tikhonov regularization solutions $\mathbf{x}$ typically have all coefficients nonzero which is not desired for sparse inversion. The basis pursuit algorithm formulates an optimization problem that simultaneously minimizes both the l_2 norm of the error and l_1 norm of the solution [4]:

$$\min_{\mathbf{x}} \|\mathbf{A}\mathbf{x} - \mathbf{y}\|_2 + \lambda\|\mathbf{x}\|_1 \tag{5.5}$$

where $\|\mathbf{x}\|_1 = \sum_i |x_i|$ denotes the l_1 norm of $\mathbf{x}$, and $\lambda > 0$ is the regularization parameter. Throughout this study, we have used $\lambda = 0.01$ for our computations. Basis pursuit typically yields a sparse vector $\mathbf{x}$, that is an $\mathbf{x}$ that has relatively few nonzero coefficients.

We test the basis pursuit by convolving a $40-Hz$ zero-phase Ricker wavelet with reflection-coefficient pairs with varying ratios. We generated wedge models with $1ms$ sampling for a reflection-coefficient pair $r_1 = 1$ and $r_2 = 2$. The values were taken for better visualization. It is well known as the tuning effect that describes the phenomenon of constructive or destructive interference of waves from closely spaced reflections. At a spacing of less than $\frac{1}{4}$ of the wavelength, reflections have constructive interference and generate a single event of high amplitude. At spacing greater than that, the event begins to be resolvable as two separate events. The tuning thickness is the bed thickness at which two seismic events become indistinguishable in time, and it is important for seismic interpreters to know the tuning thickness to study thin reservoirs.

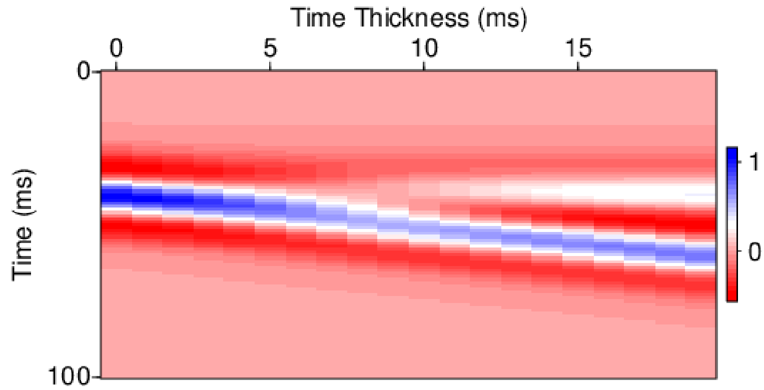
The tuning thickness of a thin-bed model with a Ricker wavelet can be calculated by [6]

$$t_R = \frac{\sqrt{6}}{2\pi f_0} \tag{5.6}$$

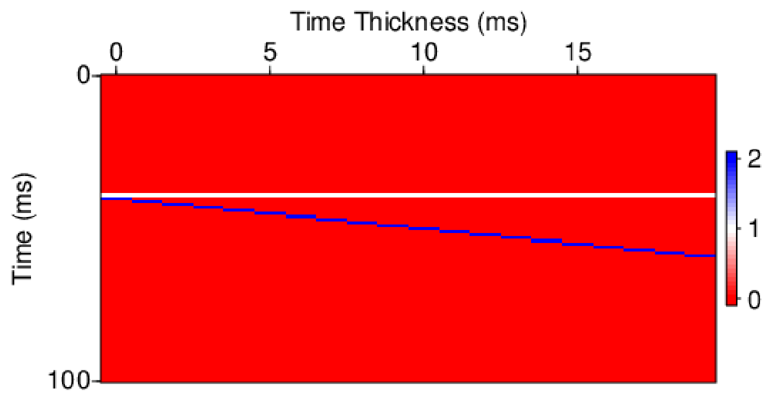
where f_0 is the dominant wavelet frequency. The equation assumes that the interfering wavelets are identical in frequency content and are zero-phase. In our case, the wedge model we produced satisfies the conditions to use the formula, for which $f_0 = 40Hz$, we have tuning thickness $t_R = 10ms$.

We plotted the wedge model as shown in Figure 5.1a and we notice that the event begins to be unresolvable as two separate events when time-thickness is less than the tuning thickness $t_R = 10ms$. When we invert with zero-phase wavelets, we obtain inverted reflectivity exactly by basis pursuit in Figure 5.1b. The top of the layer lies at $40ms$ and the base of the layer increased from $41ms$ to $61ms$. We notice that we have distinctly inverted the top and bottom of the wedge under the spacing of the tuning thickness.

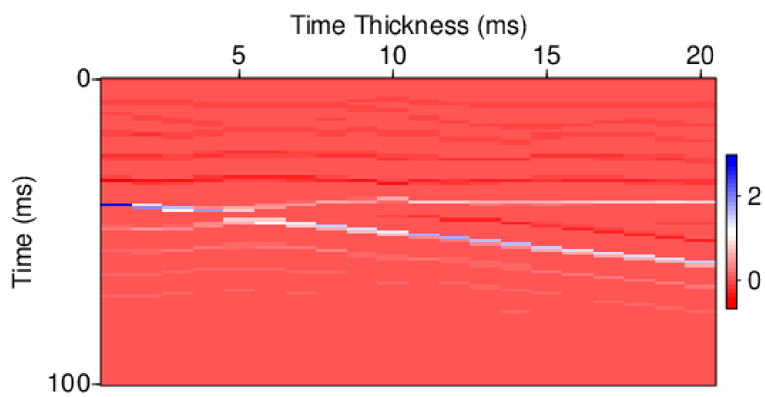
In the real world, given a seismogram, we couldn't know the wavelet with exact phase. Suppose we include a biased phase in the estimated wavelet. Then the reflectivity inversion will not be exact. This is demonstrated by the result of using a 30 degree phase wavelet to invert the same synthetic seismograms. The inverted reflectivity has artifacts as shown in Figure 5.1c for the wedge model. Therefore, we propose two new methods to invert reflectivity based on complex basis pursuit and modified complex basis pursuit to decompose the biased phase in the estimated wavelet.



(a) Wedge model



(b) Inverted reflectivity by basis pursuit with zero-phase seismic wavelet.



(c) Inverted reflectivity by basis pursuit with 30-degree seismic wavelet.

Figure 5.1: Wedge model and inverted reflectivity by basis pursuit.

5.2 Complex basis pursuit

In this section, we introduce a new method for seismic reflectivity inversion based on complex basis pursuit. We generated a complex seismic trace and a complex diagonal wavelet matrix and then applied complex basis pursuit to obtain real part and imaginary part of reflectivity. We have tested using our wedge model and the results are much more accurate than the basis pursuit method.

We define a complex diagonal wavelet matrix $\mathbf{A}^c$. Let $\mathbf{A}_j^c$ be a complex column vector of $\mathbf{A}^c$ which was defined by

$$\mathbf{A}_j^c = \mathbf{A}_j + iH(\mathbf{A}_j) \tag{5.7}$$

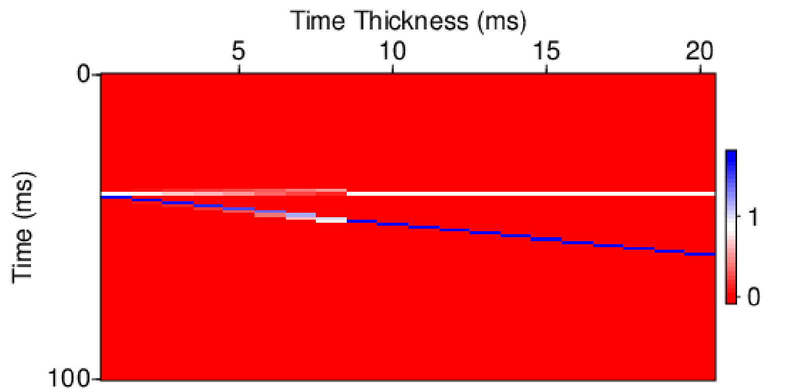
where $H(\cdot)$ is the Hilbert transform. The Hilbert transform has the effect of shifting the phase of the negative frequency components by +90 degrees and the phase of the positive frequency components by −90 degrees. Let $\mathbf{x}^c = \mathbf{x}^r + i\mathbf{x}^i, \mathbf{y}^c = \mathbf{y} + iH(\mathbf{y})$. We formulate the complex basis pursuit as

$$\min_{\mathbf{x}^c} \|\mathbf{A}^c\mathbf{x}^c - \mathbf{y}^c\|_2 + \lambda\|\mathbf{x}^c\|_1^c \tag{5.8}$$

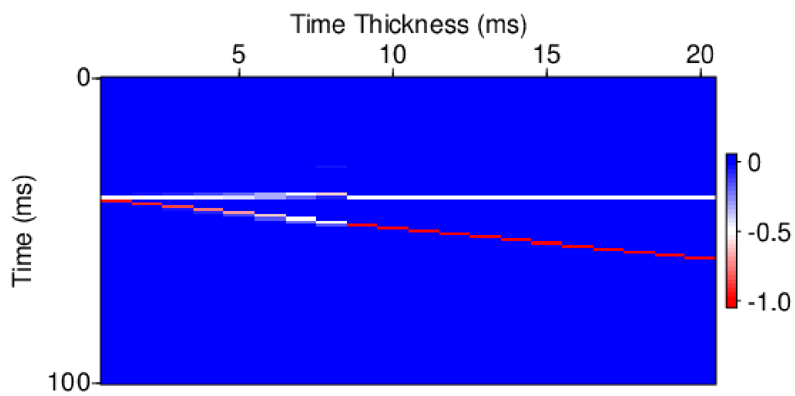
where the norm $\|\cdot\|_1^c$ was defined

$$\|\mathbf{x}^c\|_1^c = \sum_j |x_j^c| = \sum_j ((x_j^r)^2 + (x_j^i))^{\frac{1}{2}} \tag{5.9}$$

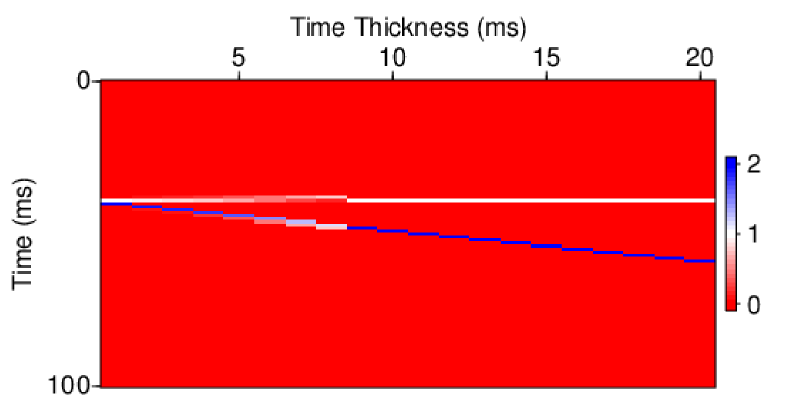
First we tested this approach on the same wedge models as shown in Figure 5.1a. We inverted the wedge model with complex basis pursuit. We obtained a real reflectivity in Figure 5.2a and imaginary reflectivity in Figure 5.2b with 30-degree phase seismic wavelet. We notice that we have inverted the top and bottom of the wedge under the spacing of tuning thickness $t_R = 10ms$. The inverted reflectivities are clean in Figure 5.2c, while artifacts existed in the basis pursuit result shown in Figure 5.1c.



(a) Inverted real reflectivity.



(b) Inverted imaginary reflectivity.



(c) Inverted reflectivity.

Figure 5.2: Complex basis pursuit with 30-degree phase seismic wavelet for wedge model.

5.3 Modified complex basis pursuit

In this section, we introduced a second method for seismic reflectivity inversion based on modified complex basis pursuit. We employ the complex diagonal wavelet matrix as the previous section, but we then apply a modified complex basis pursuit for the real seismic trace (not the complex seismic trace) to obtain the real part and imaginary part of the reflectivity. We have tested using our wedge model and the results are visually the same as the full complex pursuit method.

We formulate the modified complex basis pursuit as

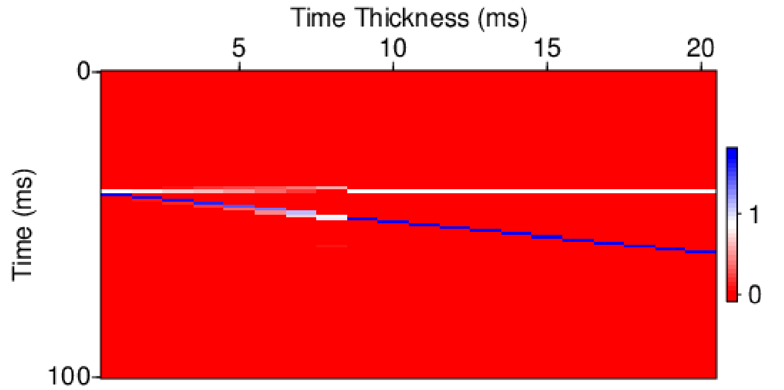
$$\min_{\mathbf{x}^c} \|\mathbf{A}^c \cdot \mathbf{x}^c - \mathbf{y}\|_2 + \lambda \|\mathbf{x}^c\|_1^c \tag{5.10}$$

where

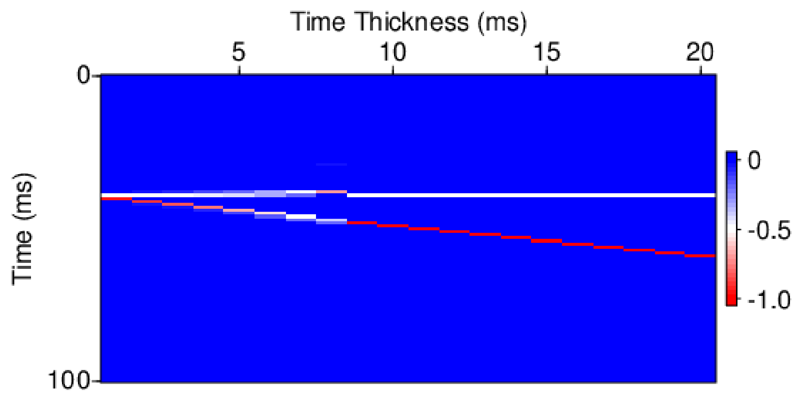
$$\mathbf{A}^c \cdot \mathbf{x}^c = \sum_j \mathbf{A}_j \mathbf{x}^r - H(\mathbf{A}_j)\mathbf{x}^i \in \mathbb{R}^N \tag{5.11}$$

We notice that here we minimize the real part of the complex least square error instead of the complex least square error in complex basis pursuit. First we tested on the same wedge models as shown in Figure 5.1a. We inverted the wedge model with modified complex basis pursuit. We obtained a real reflectivity in Figure 5.3a and imaginary reflectivity in Figure 5.3b with 30-degree phase seismic wavelet. We notice that we have inverted the top and bottom of the wedge under the spacing of the tuning thickness $t_R = 10ms$. The results are visually the same as the complex basis pursuit as shown in Figures 5.2a and 5.2b.

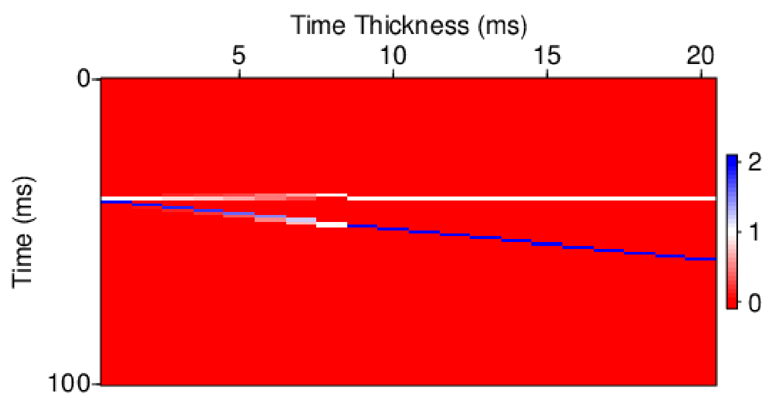
We compared the inverted reflectivity with the actual reflectivity as follows. We vertically concatenated the inverted reflectivity with 30 degree phase wavelets by



(a) Inverted real reflectivity.



(b) Inverted imaginary reflectivity.



(c) Inverted reflectivity.

Figure 5.3: Modified complex basis pursuit with 30-degree phase seismic wavelet for wedge model.

basis pursuit, complex basis pursuit, modified complex basis pursuit and the actual reflectivity from top to bottom as shown in Figure 5.4. We noticed that the inverted reflectivity by basis pursuit showed obvious artifacts (at the top), while the complex basis pursuit (top middle) and modified complex basis pursuit (bottom middle) resulted in much more accurate reflectivity. This is illustrated by the residue of the inverted relectivities as shown in Figure 5.5 where the residue by basis pursuit, complex basis pursuit and modified complex basis pursuit were vertically concatenated.

We tested using 60 degree phase wavelets. We noticed that the inverted reflectivity by basis pursuit showed obvious artifacts (at the top), while the complex basis pursuit (top middle) and modified complex basis pursuit (bottom middle) resulted in similar reflectivity as the actual reflectivity (bottom) in Figure 5.6. The residues were plotted in Figure 5.7 and it showed that the complex basis pursuit (middle) and modified complex basis pursuit (bottom) were much more accurate than the basis pursuit (top).

Besides, we tested using 90 degree phase wavelets. We noticed that the inverted reflectivity by basis pursuit showed obvious artifacts (at the top), while the complex basis pursuit (top middle) and modified complex basis pursuit (bottom middle) resulted in similar reflectivity as the actual reflectivity (bottom) in Figure 5.8. The residue were plotted in Figure 5.9 and it showed that the complex basis pursuit (middle) and modified complex basis pursuit (bottom) were much more accurate than the basis pursuit (top).

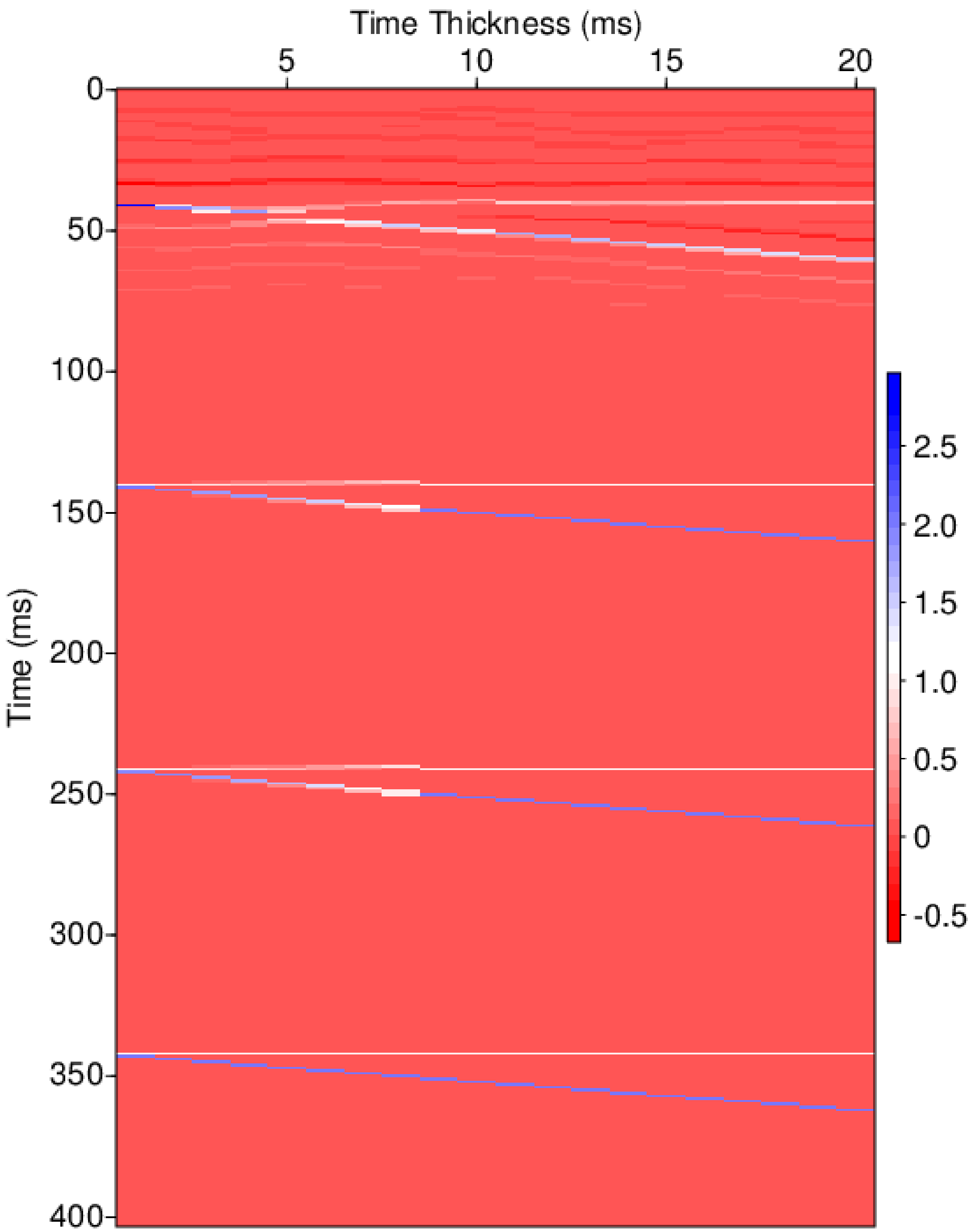


Figure 5.4: Inverted reflectivity comparison with 30-degree phase seismic wavelet for wedge model.

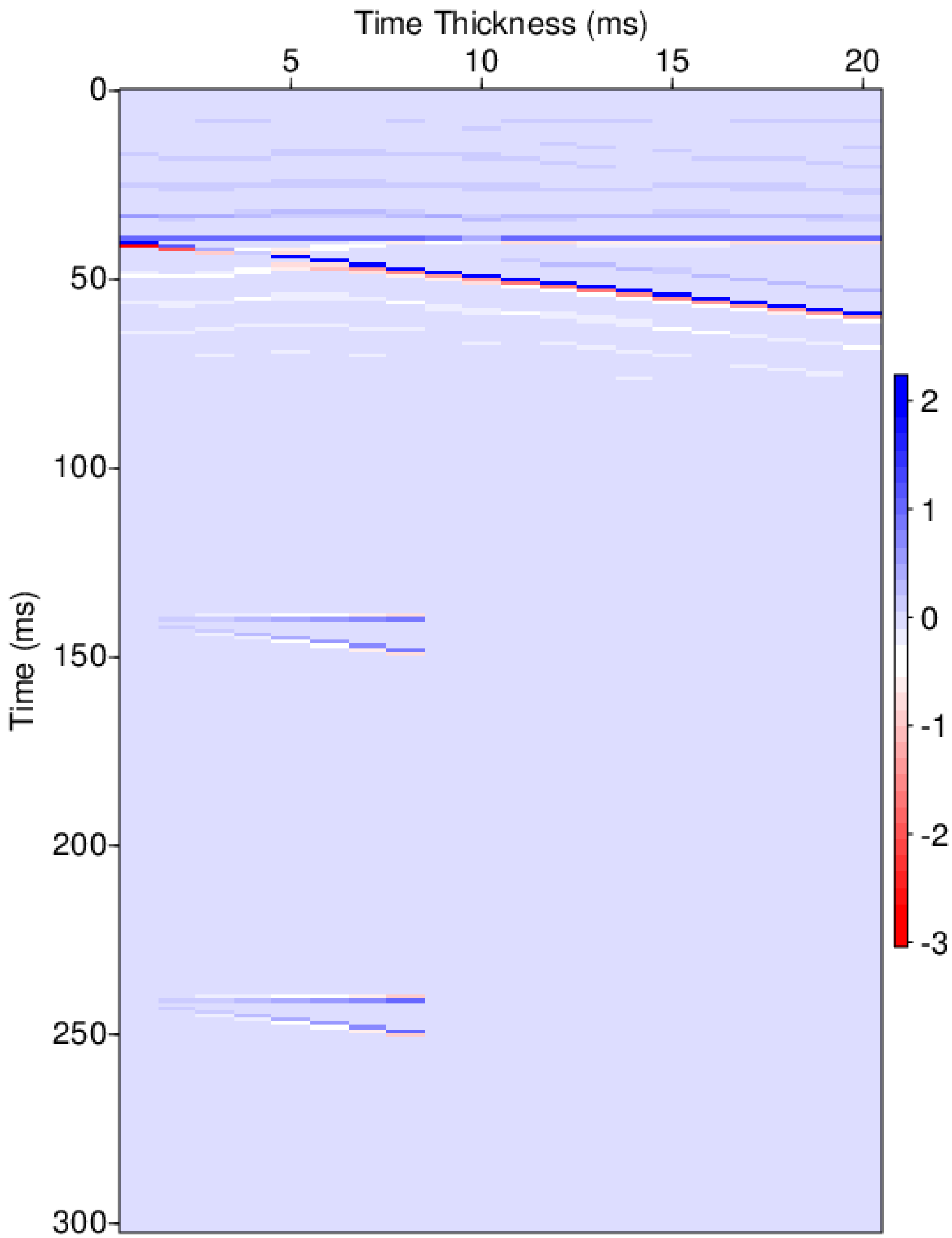


Figure 5.5: Inverted reflectivity residue comparison with 30-degree phase seismic wavelet for wedge model.

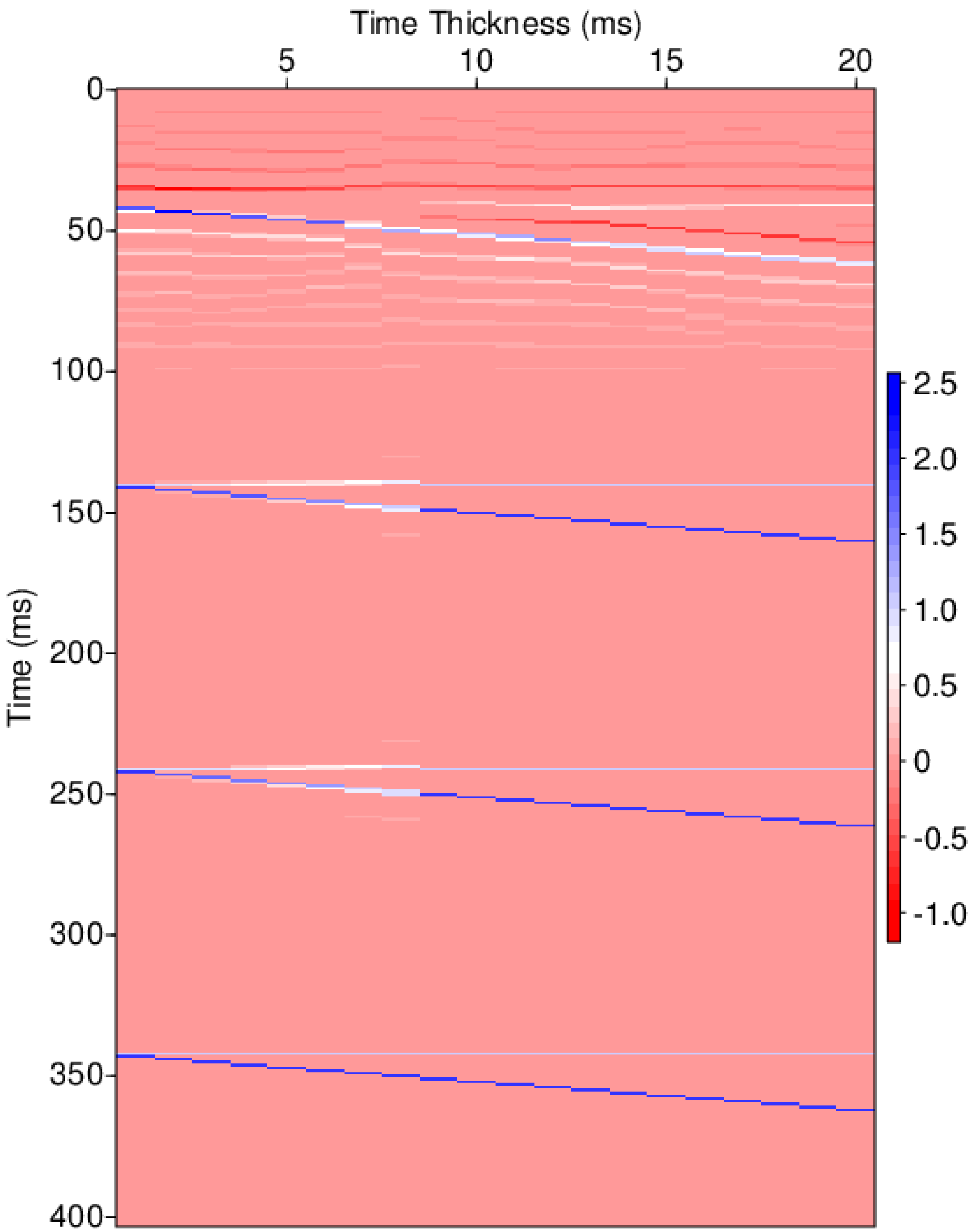


Figure 5.6: Inverted reflectivity comparison with 60-degree phase seismic wavelet for wedge model.

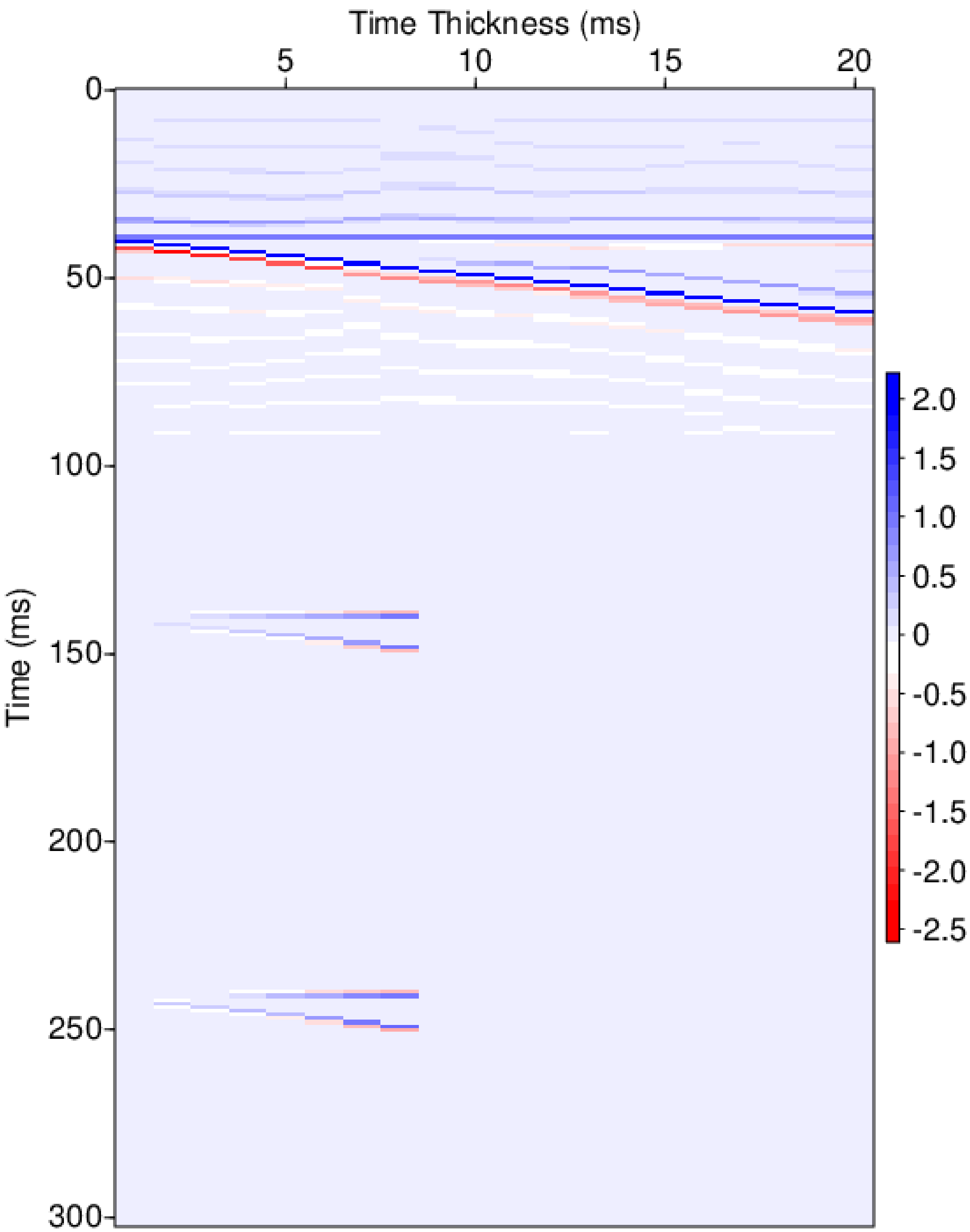


Figure 5.7: Inverted reflectivity residue comparison with 60-degree phase seismic wavelet for wedge model.

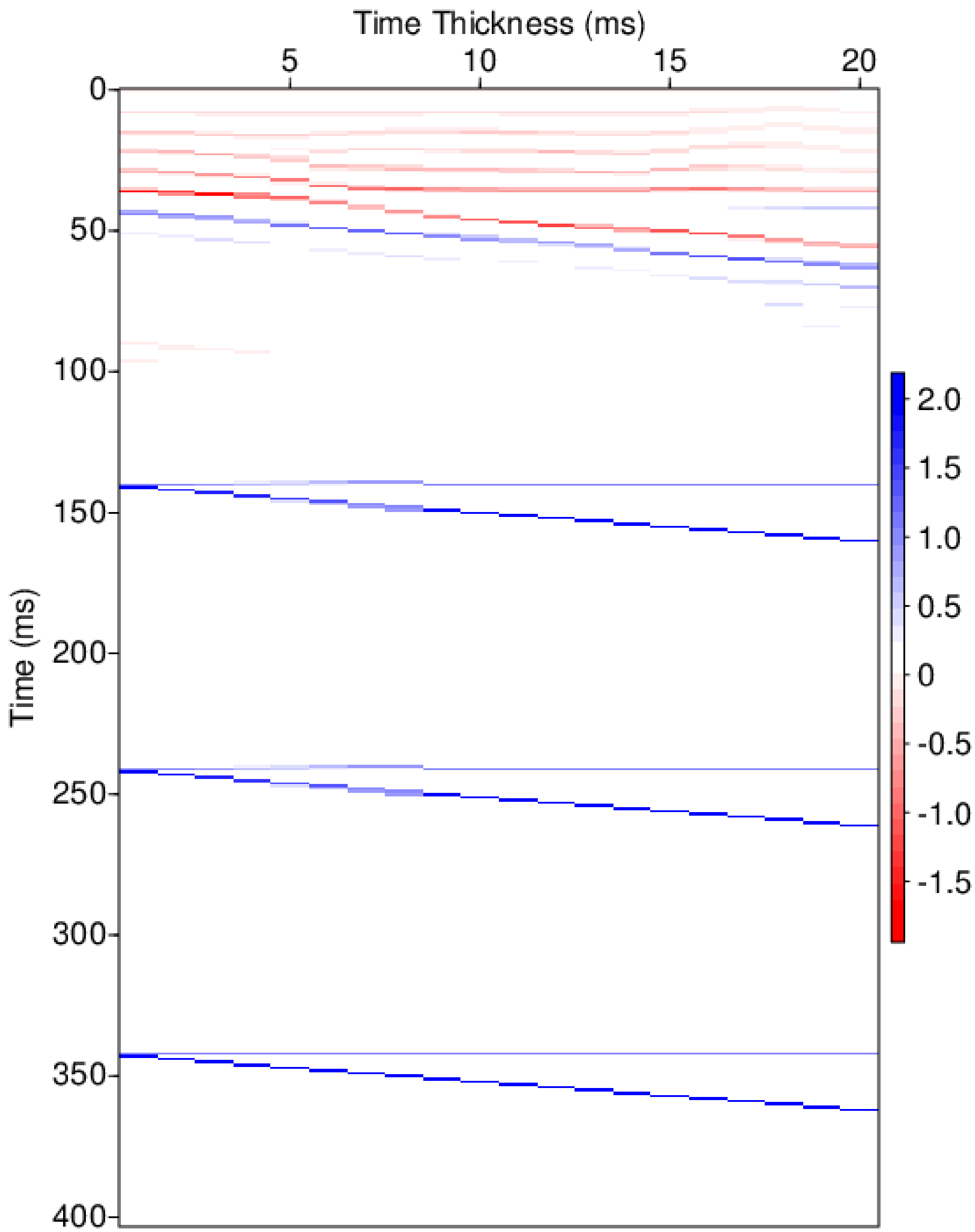


Figure 5.8: Inverted reflectivity comparison with 90-degree phase seismic wavelet for wedge model.

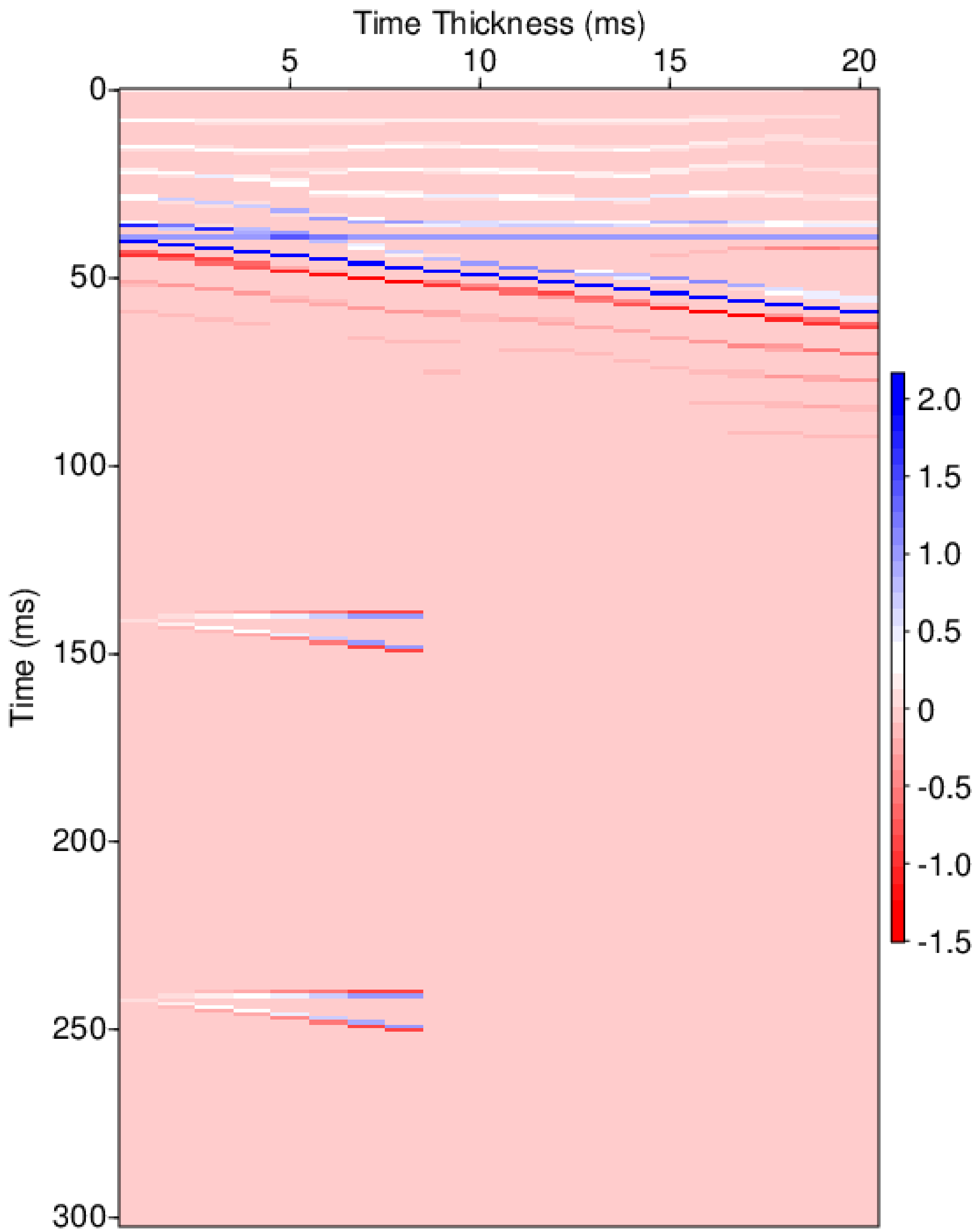


Figure 5.9: Inverted reflectivity residue comparison with 90-degree phase seismic wavelet for wedge model.

5.4 Implementation

To solve problems in equation 5.5 for basis pursuit, equation 5.8 for complex basis pursuit and equation 5.10 for modified complex basis pursuit we used CVX, a package for specifying and solving convex programs [7, 12]. More specifically, the solver we applied for both basis pursuit, complex basis pursuit and modified complex basis pursuit is SDPT3 [45, 46]. The optimization algorithm implemented in SDPT3 is a primal-dual interior-point algorithm that uses the path-following paradigm. In each iteration, a predictor search direction was computed to decrease the duality gap as much as possible. After that, the algorithm generated a Mehrotra-type corrector step [23] to keep the iterate near the central path. Neighborhood restrictions were not imposed on the iterates. Initial iterates need not be feasible -the algorithm tries to achieve feasibility and optimality of its iterates simultaneously.

CHAPTER 6

Conclusions and future work

An efficient seismic data compression method can speed up handling large data volumes and possibly speed up the imaging algorithm. In Chapter 2, we developed new methodologies using wavelet dictionary as new representation space so that seismic data can be represented in smaller size. We have tested our methodologies on a synthetic shot record, and the compressed data is 10% of the original with great fidelity ($SNR > 40dB$). We list the characteristics of the three methods below:

1. Matching pursuit with Symlets

    - worked without knowledge of the source wavelet

- obtained better quality than Discrete Wavelet Transform
- had obvious artifacts for too much compression

2. Subspace matching pursuit

- needed source wavelet or estimated source wavelet
- eliminated artifacts for various compression ratio
- decreased convergence rate due to incompleteness of the dictionary

3. Matching pursuit with combined dictionaries (Symlets and Subspace)

- needed source wavelet or estimated source wavelet
- had limited artifacts existing for various compression ratio
- achieved best quality with fastest convergence among the three methods

All three methods can have adaptive error control by specifying the stopping criteria for the iterations.

In refraction seismology, first break (or arrival time) detection has been applied to study the near surface low-velocity zone and determine the static corrections. In recent years, with the advances in hydraulic fracturing techniques, first break detection of P-waves and S-waves has become crucial for locating microseismic events. In Chapter 3, we proposed a new method of automatic first break detection of P-waves and S-waves. Our method is based on a time-frequency analysis of the seismic trace using minimum uncertainty (μ-) wavelets. The performance of our method has been tested on both lab data and field data. The results suggest that our method is robust to pick P-wave and S-wave arrivals for microseismic data, automatically.

In seismic traveltime tomography, the automatic or manual picking of seismic events are inverted to give a correct velocity model. In this tomographic velocity model updating process, the manual picking of prestack events is a primary bottleneck. In addition, Laplace domain waveform inversion requires muting prestack seismic data before the first arrival protecting refracted waves. We explored methods of automatic seismic event picking in Chapter 4. It is known that no single automatic seismic event indicator works for all data; therefore, we explored two indicators based on μ-wavelets and an energy ratio. Thresholding was applied to pick seismic events. We have tested with synthetic data with various level of noise.

For seismic reflectivity inversion, a constant phase wavelet is usually used in practice for simplification. If the phase of the estimated wavelet is wrong, this can cause an error in reflectivity. I proposed a new sparse inversion method that doesn't require that the phase of the estimated wavelet be the same as the seismic data. We constructed a complex dictionary of the estimated wavelets and performed a complex basis pursuit to decompose the complex seismic traces to corresponding real and imaginary reflectivity. The complex dictionary consisted of the estimated wavelet with various shifts of the real wavelet using the corresponding Hilbert transform of the wavelets for the imaginary part. In this way, we can obtain more accurate reflectivity even though the estimated wavelet has biased phase. We also proposed a modified complex basis pursuit method to invert the seismic trace for reflectivity. The difference is that we only minimize the real part of the least square error instead of the complex least square error. We obtained results which are visually the same as the complex basis pursuit. Application to field data is left for future work.

APPENDIX A

Elastic wave equation in pre-stressed medium

We derived an elastic wave equation in a pre-stressed medium. The pre-stress causes anisotropy of the elastic tensor and wave velocities. We obtained the Green-Christoffel equation by considering the harmonic plane wave solution. Also we analyzed an experimental study to get the elastic tensor under uni-axial pre-stress.

A.1 Derivation of equation of motion in Lagrangian

In the continuum mechanics, if the medium is such that the original relative positions of the particles have little or no effect on the internal forces throughout the body, the it will be convenient to work in the current coordinates of each particle. Otherwise, if the original relative positions affect the forces within the body, then it is appropriate to work in the original coordinates. The two formulations are called the Eulerian and the Lagrangian respectively. In elasticity, the force between any pair of particles depends on the difference between the current and original mutual distances, therefore, we will derive the elastic wave equation in Lagrangian from the equation of motion in Eulerian.

The displacement of a particle is defined by

$$u_i = x_i - a_i \tag{A.1}$$

where a_i are the original coordinates (Lagrangian variables) and x_i are the current coordinates (Eulerian variables). Let $\hat{\sigma}_{ij}$ be the Eulerian stress tensor, and the force f_i exerted across the parallelogram with sides dx_i and δx_i (with normal n_i) can be given as

$$f_i = \hat{\sigma}_{ij}\varepsilon_{jkl}n_j dx_k \delta x_l$$

where ε_{ijk} is the permutation symbol:

$$\varepsilon_{ijk} = \begin{cases} +1 & \text{if } (i,j,k) \text{ is } (1,2,3),(3,1,2) \text{ or } (2,3,1), \\ -1 & \text{if } (i,j,k) \text{ is } (1,3,2),(3,2,1) \text{ or } (2,1,3), \\ 0 & \text{if } i=j \text{ or } j=k \text{ or } k=i \end{cases}$$

We define τ_{ij} as Lagrangian (Piola-Kirchhoff) stress tensor which satisfies that

$$f_i = \tau_{ij}\varepsilon_{jkl}n_j da_k \delta a_l$$

In this case, the force f_i is expressed in terms of the sides of the parallelogram in its original position da_i and δa_i. It is necessary and sufficient condition the relation of the Eulerian stress tensor $\hat{\sigma}_{ij}$ and the Lagrangian stress tensor τ_{ij} holds [1]:

$$\hat{\sigma}_{ij} = J(\tbinom{x}{a})\frac{\partial x_j}{\partial a_k}\tau_{ik} \tag{A.2}$$

where $J(\tbinom{x}{a})$ is the Jacobian $J(x_1, x_2, x_3/a_1, a_2, a_3)$, and

$$J(\tbinom{x}{a}) = \frac{\rho_0}{\rho}, J(\tbinom{a}{x}) = \frac{\rho}{\rho_0} \tag{A.3}$$

where ρ_0 and ρ are the initial and current densities.

The equation of motion in Eulerian form is

$$\frac{\partial \hat{\sigma}_{ij}}{\partial x_j} + \rho F_i = \rho \frac{D^2 u_i}{Dt^2} \tag{A.4}$$

where F_i is the body force per unit mass. To convert this equation into Lagragian form we need the following result, namely

$$\frac{\partial}{\partial x_j}(J(\tbinom{a}{x})\frac{\partial x_j}{\partial a_k}) = 0 \tag{A.5}$$

Proof. We have
$$\begin{aligned}
\varepsilon_{ijk}\frac{\partial a_l}{\partial x_i}\frac{\partial a_m}{\partial x_j}\frac{\partial a_n}{\partial x_k} &= \varepsilon_{123}\varepsilon_{ijk}\frac{\partial a_l}{\partial x_i}\frac{\partial a_m}{\partial x_j}\frac{\partial a_n}{\partial x_k} \\
&= \varepsilon_{lmn}\delta_{1l}\delta_{2m}\delta_{3n}\varepsilon_{ijk}\frac{\partial a_l}{\partial x_i}\frac{\partial a_m}{\partial x_j}\frac{\partial a_n}{\partial x_k} \\
&= \varepsilon_{lmn}\frac{\partial a_1}{\partial a_l}\frac{\partial a_2}{\partial a_m}\frac{\partial a_3}{\partial a_n}\varepsilon_{ijk}\frac{\partial a_l}{\partial x_i}\frac{\partial a_m}{\partial x_j}\frac{\partial a_n}{\partial x_k} \\
&= \varepsilon_{lmn}\varepsilon_{ijk}\frac{\partial a_1}{\partial x_i}\frac{\partial a_2}{\partial x_j}\frac{\partial a_3}{\partial x_k} \\
&= \varepsilon_{lmn}J(\tbinom{a}{x})
\end{aligned}$$

We multiply both sides by $\dfrac{\partial x_p}{\partial a_l}$,

$$\varepsilon_{ijk}\frac{\partial x_p}{\partial x_i}\frac{\partial a_m}{\partial x_j}\frac{\partial a_n}{\partial x_k} = \varepsilon_{lmn}J(\tfrac{a}{x})\frac{\partial x_p}{\partial a_l} \tag{A.6}$$

That is,

$$\varepsilon_{pjk}\frac{\partial a_m}{\partial x_j}\frac{\partial a_n}{\partial x_k} = \varepsilon_{lmn}J(\tfrac{a}{x})\frac{\partial x_p}{\partial a_l} \tag{A.7}$$

Then we differentiate both sides by $\dfrac{\partial}{\partial x_p}$,

$$\frac{\partial}{\partial x_p}(\varepsilon_{pjk}\frac{\partial a_m}{\partial x_j}\frac{\partial a_n}{\partial x_k}) = \frac{\partial}{\partial x_p}(\varepsilon_{lmn}J(\tfrac{a}{x})\frac{\partial x_p}{\partial a_l}) \tag{A.8}$$

$$\varepsilon_{pjk}\frac{\partial}{\partial x_p}(\frac{\partial a_m}{\partial x_j})\frac{\partial a_n}{\partial x_k} + \varepsilon_{pjk}\frac{\partial}{\partial x_p}(\frac{\partial a_n}{\partial x_k}\frac{\partial a_m}{\partial x_j}) = \varepsilon_{lmn}\frac{\partial}{\partial x_p}(J(\tfrac{a}{x})\frac{\partial x_p}{\partial a_l}) \tag{A.9}$$

Since

$$\frac{\partial}{\partial x_p}\frac{\partial}{\partial x_j} = \frac{\partial}{\partial x_j}\frac{\partial}{\partial x_p}$$

We have

$$\varepsilon_{pjk}\frac{\partial}{\partial x_p}(\frac{\partial a_m}{\partial x_j}) = \varepsilon_{jpk}\frac{\partial}{\partial x_j}(\frac{\partial a_m}{\partial x_p}) = \varepsilon_{jpk}\frac{\partial}{\partial x_p}(\frac{\partial a_m}{\partial x_j}) = -\varepsilon_{pjk}\frac{\partial}{\partial x_p}(\frac{\partial a_m}{\partial x_j})$$

Therefore, we obtain

$$\varepsilon_{pjk}\frac{\partial}{\partial x_p}(\frac{\partial a_m}{\partial x_j}) = 0$$

Analogously,

$$\varepsilon_{pjk}\frac{\partial}{\partial x_p}(\frac{\partial a_n}{\partial x_k}) = 0$$

Thus, we get

$$0 = \varepsilon_{lmn}\frac{\partial}{\partial x_p}(J(\tfrac{a}{x})\frac{\partial x_p}{\partial a_l}) \tag{A.10}$$

We multiply both sides by ε_{jmn}, since $\varepsilon_{jmn}\varepsilon_{lmn} = 2\delta_{jl}$

$$0 = 2\delta_{jl}\frac{\partial}{\partial x_p}(J(\tfrac{a}{x})\frac{\partial x_p}{\partial a_l}) \tag{A.11}$$

That is

$$0 = \frac{\partial}{\partial x_p}(J(\tbinom{a}{x})\frac{\partial x_p}{\partial a_j}) \tag{A.12}$$

We introduce the Kirchhoff stress tensor K_{ij} and the Eulerian stress tensor can be written as

$$\hat{\sigma}_{ij} = J(\tbinom{x}{a})\frac{\partial x_j}{\partial a_l}\frac{\partial x_i}{\partial a_k}K_{kl} \tag{A.13}$$

Comparing equation (A.2) and (A.13), we have the relationship of the two Lagrangian stress tensors:

$$\tau_{ik} = \frac{\partial x_i}{\partial a_l}K_{kl} \tag{A.14}$$

and

$$\begin{aligned}
\frac{\partial \hat{\sigma}_{ij}}{\partial x_j} &= \frac{\partial}{\partial x_j}(J(\tbinom{x}{a})\frac{\partial x_j}{\partial a_l}\frac{\partial x_i}{\partial a_k}K_{kl}), \text{by } (A.13) \\
&= \frac{\partial}{\partial x_j}(J(\tbinom{x}{a})\frac{\partial x_j}{\partial a_l})(\frac{\partial x_i}{\partial a_k}K_{kl}) + J(\tbinom{x}{a})\frac{\partial}{\partial x_j}(\frac{\partial x_i}{\partial a_k}K_{kl})\frac{\partial x_j}{\partial a_l}, \\
&= J(\tbinom{x}{a})\frac{\partial}{\partial a_l}(\frac{\partial x_i}{\partial a_k}K_{kl})), \text{by } (A.5) \\
&= J(\tbinom{x}{a})\frac{\partial}{\partial a_l}(\frac{\partial(u_i + a_i)}{\partial a_k}K_{kl}), \\
&= J(\tbinom{x}{a})(u_{i,k}K_{kl} + \delta_{ik}K_{kl})_{,l} = J(\tbinom{x}{a})(u_{i,k}K_{kj} + K_{ij})_{,j}, \\
&= J(\tbinom{x}{a})(u_{i,kj}K_{kj} + u_{i,k}K_{kj,j} + K_{ij,j}).
\end{aligned}$$

Let us represent the total stress by

$$K_{ij} = \tau^0_{ij} + \sigma_{ij} \tag{A.15}$$

where τ^0_{ij} is the initial stress and σ_{ij} is the stress produced by propagating waves.

Then we have

$$\begin{aligned}
\frac{\partial \hat{\sigma}_{ij}}{\partial x_j} &= J(\tbinom{x}{a})(u_{i,kj}(\tau^0_{kj}+\sigma_{kj}) + u_{i,k}(\tau^0_{kj,j}+\sigma_{kj,j}) + (\tau^0_{ij,j}+\sigma_{ij,j})) \\
&= J(\tbinom{x}{a})(\tau^0_{kj}u_{i,kj} + u_{i,kj}\sigma_{kj} + u_{i,k}\sigma_{kj,j} + \sigma_{ij,j}), \text{ since } \tau^0_{kj,j} = 0 \\
&\approx J(\tbinom{x}{a})(\tau^0_{kj}u_{i,kj} + \sigma_{ij,j}) \text{ ignoring the higher order terms}
\end{aligned}$$

We then substitute this into the equation of motion in Eulerian form A.4. By the equivalence of the material time derivative and $\frac{\partial}{\partial t}$ in the Lagrangian coordinate frame, we get the equation of motion in the Lagrangian form:

$$\frac{\rho}{\rho_0}(\tau^0_{kj}u_{i,kj} + \sigma_{ij,j}) + \rho F_i = \rho\frac{\partial^2 u_i}{\partial t^2} \tag{A.16}$$

or equivalently,

$$\tau^0_{kj}u_{i,kj} + \sigma_{ij,j} + \rho_0 F_i = \rho_0\frac{\partial^2 u_i}{\partial t^2} \tag{A.17}$$

We next introduce the linear infinitesimal strain tensor:

$$\varepsilon_{kl} = \frac{1}{2}(\frac{\partial u_k}{\partial a_l} + \frac{\partial u_l}{\partial a_k}) \tag{A.18}$$

The second term in equation $(A.17)$ can be written as

$$\sigma_{ij,j} = \frac{\partial \sigma_{ij}}{\partial a_j} = \frac{\partial \sigma_{ij}}{\partial \varepsilon_{kl}}\frac{\partial \varepsilon_{kl}}{\partial a_j} \tag{A.19}$$

We then have

$$\frac{\partial \sigma_{ij}}{\partial \varepsilon_{kl}} = C_{ijkl}(\tau^0_{mn}) \tag{A.20}$$

where the $C_{ijkl}(\tau^0_{mn})$ are the elastic constants in a pre-stressed medium. We substitute equation $(A.18)$ and $(A.20)$ into $(A.19)$ to obtain

$$\frac{\partial \sigma_{ij}}{\partial a_j} = C_{ijkl}(\tau^0_{mn})\frac{1}{2}\frac{\partial}{\partial a_j}(\frac{\partial u_k}{\partial a_l} + \frac{\partial u_l}{\partial a_k}) = C_{ijkl}(\tau^0_{mn})\frac{\partial^2 u_k}{\partial a_j \partial a_l} \tag{A.21}$$

The equation of motion $(A.17)$ can be rewritten finally as

$$C_{ijkl}(\tau^0_{mn})\frac{\partial^2 u_k}{\partial a_j \partial a_l} + \tau^0_{kj}\frac{\partial^2 u_i}{\partial a_j \partial a_k} + \rho_0 F_i = \rho_0 \frac{\partial^2 u_i}{\partial t^2} \tag{A.22}$$

A.2 The elastic tensor in anisotropic medium

Let us investigate the equation of elastic moduli in an arbitrarily, anisotropic initially stressed medium. Let us represent stress

$$\tau_{ij} = \tau^0_{ij} + \sigma_{ij} \tag{A.23}$$

where τ^0_{ij} is the initial stress and σ_{ij} is the stress produced by propagating waves. We represent the elastic potential as an expansion of the Green strain tensor ε_{ij}

$$W = W_0 + \tau^0_{ij}\varepsilon_{ij} + C_{ijkl}\varepsilon_{ij}\varepsilon_{kl} + C_{ijklmn}\varepsilon_{ij}\varepsilon_{kl}\varepsilon_{mn} + ... \tag{A.24}$$

where C_{ijkl} and C_{ijklmn} are the second order and third order elastic moduli. The initial stress can be written as

$$\tau^0_{ij} = -P\delta_{ij} + t^0_{ij} \tag{A.25}$$

where $P = -\frac{1}{3}\tau^0_{ii}$ is the initial pressure, and t^0_{ij} is the deviating part. Then we can rewrite the potential as

$$W = W_0 + \tau^0_{ij}\varepsilon_{ij} + C^0_{ijkl}(P)\varepsilon_{ij}\varepsilon_{kl} + B_{ijklmn}\varepsilon_{ij}\varepsilon_{kl}t^0_{mn} + ... \tag{A.26}$$

where C^0_{ijkl} is the second-order elastic tensor dependent on pressure and determining the type of symmetry in the medium without pre-stress; B_{ijklmn} is the fourth-order tensor that characterizes the anisotropy of the medium generated by the pre-stress. Comparing equation (A.24) and (A.26), we notice that

$$C_{ijkl} = C^0_{ijkl} + B_{ijklmn}t^0_{mn} \tag{A.27}$$

We can get the stress from the potential by

$$\tau_{ij} = \frac{1}{2}(\frac{\partial W}{\partial \varepsilon_{ij}} + \frac{\partial W}{\partial \varepsilon_{ji}}) \tag{A.28}$$

Then substituting equation(A.26) into (A.28), we obtain

$$\begin{aligned} \tau_{pq} &= \tau^0_{pq} + \lambda\varepsilon_{mm}\delta_{pq} + 2\mu\varepsilon_{pq} + \nu_1(t^0_{mn}\varepsilon_{mn}\delta_{pq} + \varepsilon_{jj}t^0_{pq}) \\ &+ \nu_2(\varepsilon_{pm}t^0_{mq} + t^0_{pm}\varepsilon_{mq}) \end{aligned} \tag{A.29}$$

where λ and μ are Lame constants which are dependent on the initial pressure P, and ν_1 and ν_2 are non-zero components of B_{ijklmn} [32]. From equation (A.24) and (A.28), we have

$$C_{ijkl} = \frac{\partial^2 W}{\partial \varepsilon_{ij}\varepsilon_{kl}} = \frac{1}{2}(\frac{\partial \tau_{ij}}{\partial \varepsilon_{kl}} + \frac{\partial \tau_{ij}}{\partial \varepsilon_{lk}}) \tag{A.30}$$

We substitute equation (A.29) into (A.30), to obtain

$$\begin{aligned} C_{pqkl} &= \lambda\delta_{pq}\delta_{kl} + \mu(\delta_{pk}\delta_{ql} + \delta_{pl}\delta_{qk}) + \nu_1(\delta_{pq}t^0_{kl} + t^0_{pq}\delta_{kl}) \\ &+ \frac{1}{2}\nu_2(\delta_{pk}t^0_{ql} + \delta_{pl}t^0_{qk}) + (t^0_{pk}\delta_{ql} + t^0_{pl}\delta_{qk}) \end{aligned} \tag{A.31}$$

The elastic modulus reduces to the isotropic homogeneous case when $\tau^0_{ij} = 0$:

$$C^0_{pqkl} = \lambda\delta_{pq}\delta_{kl} + \mu(\delta_{pk}\delta_{ql} + \delta_{pl}\delta_{qk}) \tag{A.32}$$

There is symmetry of the elastic tensor :

$$C_{ijkl} = C_{jikl} = C_{ijlk} = C_{klij}$$

If we write the tensor in the form of matrix, according to the rule 11-1,22-2,33-3,23-4,13-5,12-6, we have $C_{\alpha\beta} =$

$$\begin{pmatrix} C_{11} & \lambda - \nu_1 t^0_{33} & \lambda - \nu_1 t^0_{22} & \nu_1 t^0_{23} & (\nu_1+\nu_2)t^0_{13} & (\nu_1+\nu_2)t^0_{12} \\ C_{21} & C_{22} & \lambda - \nu_1 t^0_{11} & (\nu_1+\nu_2)t^0_{23} & \nu_1 t^0_{13} & (\nu_1+\nu_2)t^0_{12} \\ C_{31} & C_{32} & C_{33} & (\nu_1+\nu_2)t^0_{23} & (\nu_1+\nu_2)t^0_{13} & \nu_1 t^0_{12} \\ C_{41} & C_{42} & C_{43} & \mu - \frac{1}{2}\nu_2 t^0_{11} & \frac{1}{2}\nu_2 t^0_{12} & \frac{1}{2}\nu_2 t^0_{13} \\ C_{51} & C_{52} & C_{53} & C_{54} & \mu - \frac{1}{2}\nu_2 t^0_{22} & \frac{1}{2}\nu_2 t^0_{23} \\ C_{61} & C_{62} & C_{63} & C_{64} & C_{65} & \mu - \frac{1}{2}\nu_2 t^0_{33} \end{pmatrix}$$

where $C_{\alpha\alpha} = \lambda + 2\mu + 2(\nu_1+\nu_2)t^0_{\alpha\alpha}, \alpha = 1, 2, 3$, and $C_{\alpha\beta} = C_{\beta\alpha}$.

A.3 Green Christoffel equation (Plane wave solution)

Let us consider plane wave propagation of the displacement:

$$u_k = A_k e^{ik(n_p a_p - Vt)} \tag{A.33}$$

Therefore,

$$\frac{\partial^2}{\partial a_j \partial a_l}[A_k e^{ik(n_p a_p - Vt)}] = -k^2 n_j n_l A_k e^{ik(n_p a_p - Vt)} \tag{A.34}$$

$$\frac{\partial^2}{\partial a_j \partial a_k}[A_i e^{ik(n_p a_p - Vt)}] = -k^2 n_j n_k A_i e^{ik(n_p a_p - Vt)} \tag{A.35}$$

$$\frac{\partial^2}{\partial t^2}[A_i e^{ik(n_p a_p - Vt)}] = -V^2 k^2 A_i e^{ik(n_p a_p - Vt)} \tag{A.36}$$

If we substitute the terms above into equation ($A.22$) without a body force ($F_i = 0$) and cancel out the exponential term:

$$-C_{ijkl}(\tau^0_{mn})k^2 n_j n_l A_k - \tau^0_{kj} k^2 n_j n_k A_i = -\rho_0 V^2 k^2 A_i \tag{A.37}$$

Since $A_i = \delta_{ik} A_k$, and we cancel out k^2:

$$C_{ijkl}(\tau^0_{mn}) n_j n_l A_k + \tau^0_{kj} n_j n_k \delta_{ik} A_k = \rho_0 V^2 \delta_{ik} A_k \tag{A.38}$$

We group the terms:

$$(C_{ijkl}(\tau^0_{mn}) n_j n_l + \tau^0_{kj} n_j n_k \delta_{ik} - \rho_0 V^2 \delta_{ik}) A_k = 0 \tag{A.39}$$

We denote $\Gamma^*_{ik} = C_{ijkl}(\tau^0_{mn}) n_j n_l + \tau^0_{kj} n_j n_k \delta_{ik}$ as the Green-Christoffel tensor, and multiplying ($A.39$) by the exponential term we have

$$(\Gamma^*_{ik} - \rho_0 V^2 \delta_{ik}) u_k = 0 \tag{A.40}$$

We denote $\Gamma_{ik} = C_{ijkl} n_j n_l$ and its components are

$$\begin{aligned}
\Gamma_{11} &= C_{11}n_1^2 + C_{66}n_2^2 + C_{55}n_3^2 + 2C_{16}n_1n_2 + 2C_{15}n_1n_3 + 2C_{56}n_2n_3 \\
\Gamma_{22} &= C_{66}n_1^2 + C_{22}n_2^2 + C_{44}n_3^2 + 2C_{26}n_1n_2 + 2C_{46}n_1n_3 + 2C_{24}n_2n_3 \\
\Gamma_{33} &= C_{55}n_1^2 + C_{44}n_2^2 + C_{33}n_3^2 + 2C_{45}n_1n_2 + 2C_{35}n_1n_3 + 2C_{34}n_2n_3 \\
\Gamma_{12} &= C_{16}n_1^2 + C_{26}n_2^2 + C_{45}n_3^2 + (C_{12} + C_{66})n_1n_2 + (C_{14} + C_{56})n_1n_3 \\
&\quad + (C_{46} + C_{25})n_2n_3 \\
\Gamma_{13} &= C_{15}n_1^2 + C_{46}n_2^2 + C_{35}n_3^2 + (C_{14} + C_{46})n_1n_2 + (C_{13} + C_{55})n_1n_3 \\
&\quad + (C_{36} + C_{45})n_2n_3 \\
\Gamma_{23} &= C_{56}n_1^2 + C_{24}n_2^2 + C_{34}n_3^2 + (C_{46} + C_{25})n_1n_2 + (C_{36} + C_{45})n_1n_3 \\
&\quad + (C_{23} + C_{44})n_2n_3
\end{aligned}$$

The tensor has the symmetry $\Gamma_{12} = \Gamma_{21}$, $\Gamma_{13} = \Gamma_{31}$, $\Gamma_{23} = \Gamma_{32}$. Let $B = \tau^0_{kj} n_j n_k$, then the Green-Christoffel tensor is $\Gamma^*_{ik} = \Gamma_{ik} + B\delta_{ik}$, i.e.,

$$\Gamma^*_{ik} = \begin{cases} \Gamma_{ii} + B, & \text{i=k;} \\ \Gamma_{ik}, & \text{otherwise.} \end{cases}$$

Therefore, to obtain its eigenvalues we need to know ν_1, ν_2 and τ^0_{kj}.

Now we consider the velocities in an orthorhombic medium. (1) $n_1 = 1, n_2 = 0, n_3 = 0$, then

$$\begin{aligned} \Gamma^*_{11} &= C_{11} + \tau^0_{11} \\ \Gamma^*_{22} &= C_{66} + \tau^0_{11} \\ \Gamma^*_{33} &= C_{55} + \tau^0_{11} \\ \Gamma^*_{ij} &= 0, otherwise \end{aligned}$$

Thus the Green Christoffel tensor in such a medium is diagonal, and we have

$$\rho_0 V^2_{p1} = C_{11} + \tau^0_{11} = C_{11} + t^0_{11} - P \tag{A.41}$$

$$\rho_0 V^2_{sh1} = C_{66} + \tau^0_{11} = C_{66} + t^0_{11} - P \tag{A.42}$$

$$\rho_0 V^2_{sv1} = C_{55} + \tau^0_{11} = C_{55} + t^0_{11} - P \tag{A.43}$$

Analogously, (2)$n_1 = 0, n_2 = 1, n_3 = 0$, we have

$$\rho_0 V^2_{p2} = C_{22} + t^0_{22} - P \tag{A.44}$$

$$\rho_0 V^2_{sh2} = C_{66} + t^0_{22} - P \tag{A.45}$$

$$\rho_0 V^2_{sv2} = C_{44} + t^0_{22} - P \tag{A.46}$$

(3)$n_1 = 0, n_2 = 0, n_3 = 1$, we have

$$\rho_0 V_{p3}^2 = C_{33} + t_{33}^0 - P \tag{A.47}$$

$$\rho_0 V_{sh3}^2 = C_{55} + t_{33}^0 - P \tag{A.48}$$

$$\rho_0 V_{sv3}^2 = C_{44} + t_{33}^0 - P \tag{A.49}$$

From the components of C_{mn}, we have

$$C_{11} + C_{22} + C_{33} = 3\lambda + 6\mu \tag{A.50}$$

$$C_{33} + C_{44} + C_{55} = 3\mu \tag{A.51}$$

If we add equations ($A.41$,$A.44$,$A.47$), we get

$$\begin{aligned} \rho_0(V_{p1}^2 + V_{p2}^2 + V_{p3}^2) &= C_{11} + C_{22} + C_{33} + t_{11} + t_{22} + t_{33} - 3P \\ &= C_{11} + C_{22} + C_{33} - 3P \end{aligned} \tag{A.52}$$

If we substitute equation($A.50$) into equation($A.52$), we find

$$\lambda + 2\mu - P = \frac{1}{3}\rho_0(V_{p1}^2 + V_{p2}^2 + V_{p3}^2) \tag{A.53}$$

If we add equations ($A.42$,$A.46$,$A.48$), we obtain

$$\rho_0(V_{sh1}^2 + V_{sv2}^2 + V_{sh3}^2) = C_{44} + C_{55} + C_{66} - 3P \tag{A.54}$$

If we substitute equation($A.51$) into equation($A.54$), we get

$$\mu - P = \frac{1}{3}\rho_0(V_{sh1}^2 + V_{sv2}^2 + V_{sh3}^2) \tag{A.55}$$

We substitute $C_{11} = \lambda + 2\mu + 2(\nu_1 + \nu_2)t_{11}^0$ into equation ($A.41$) to obtain

$$\rho_0 V_{p1}^2 = \lambda + 2\mu - P + (2(\nu_1 + \nu_2) + 1)t_{11}^0 \tag{A.56}$$

We substitute $C_{33} = \lambda + 2\mu + 2(\nu_1 + \nu_2)t^0_{33}$ into equation $(A.47)$ to obtain

$$\rho_0 V^2_{p3} = \lambda + 2\mu - P + (2(\nu_1 + \nu_2) + 1)t^0_{33} \tag{A.57}$$

We substitute $C_{66} = \mu - \frac{1}{2}\nu_1 t^0_{33}$ into equation $(A.45)$ to obtain

$$\rho_0 V^2_{sh1} = \mu - P - \frac{1}{2}\nu_1 t^0_{33} + t^0_{11} \tag{A.58}$$

We substitute $C_{44} = \mu - \frac{1}{2}\nu_1 t^0_{11}$ into equation $(A.49)$ to obtain

$$\rho_0 V^3_{shv} = \mu - P - \frac{1}{2}\nu_1 t^0_{11} + t^0_{33} \tag{A.59}$$

From the experimental data in [33], we have $V_{p1} = 4.22$, $V_{p2} = 4.22$, $V_{p3} = 3.85$, $V_{sh1} = 2.68$, $V_{sh2} = 2.74$,$V_{sv2} = 2.73$, $V_{sv3} = 2.76$ while $\tau^0_{33} = 0.1kbar$. Then we solve equations $(A.56 - A.59)$, obtaining $\nu_1 = 11.8697$,$\nu_2 = 1.1685$,$t^0_{11} = -0.2427$,$t^0_{33} = 0.4853$. This yields $P = \tau^0_{33} + t_{33} = 0.5853$. Then we solve equations $(A.53, A.55)$ to obtain $\lambda = 6.4056$ and $\mu = 19.2835$ and finally construct the elastic tensor in a medium under initial uniaxial stress:

$$C_{\alpha\beta} = \begin{pmatrix} 38.3 & 0.16 & 7.9 & 0 & 0 & 0 \\ 0.16 & 38.3 & 7.9 & 0 & 0 & 0 \\ 7.9 & 7.9 & 57.3 & 0 & 0 & 0 \\ 0 & 0 & 0 & 19.9 & 0 & 0 \\ 0 & 0 & 0 & 0 & 19.9 & 0 \\ 0 & 0 & 0 & 0 & 0 & 19.0 \end{pmatrix}$$

Using this result, we can obtain the Green-Christoffel tensor Γ_{ik} which is a real symmetric matrix. Its eigenvalues are $\rho_0 V^2_p$,$\rho_0 V^2_{sh}$,$\rho_0 V^2_{sv}$. Therefore, we can calculate the velocities of the P-wave, S_h-wave and S_v-wave.

Bibliography

[1] D. R. Bland. *Nonlinear dynamic elasticity.* Blaisdell publishing company, 1969.

[2] A. Chakraborty and D. Okaya. Frequency-time decomposition of seismic data using wavelet-based methods. *Geophysics*, 60(6):1906–1916, 1995.

[3] C. Chandler and A. G. Gibson. Uniform approximation of functions with discrete appromixation functional. *Journal of Approximation Theory*, 100:233–250, 1999.

[4] S. Chen, D. Donoho, and M. Saunders. Atomic decomposition by basis pursuit. *SIAM Rev.*, 43(1):129–159, 2001.

[5] Z. Chen and R. Stewart. Multi-window algorithm for detecting seismic first arrivals. *CSEG National Convention*, pages 355–358, 2005.

[6] H. Chung and D. C. Lauwton. Frequency characteristics of sesmic reflections from thin beds. *Canadian Journal of Exploration Geophysics*, 31:32–37, 1995.

[7] Inc. CVX Research. CVX: Matlab software for disciplined convex programming, version 2.0 beta. `http://cvxr.com/cvx`, September 2012.

[8] I. Daubechies. *Ten lectures on wavelets.* Society for Industrial and Applied Mathematics, 1992.

[9] F. J. Dessing and C.P.A. Wapenaar. Wavefield extrapolation using the wavelet transform. *64th Annual International Meeting, SEG, Expanded Abstracts*, pages 1355–1358, 1994.

[10] F. J. Dessing and C.P.A. Wapenaar. Efficient migration with one-way operators in the wavelet transform domain. *65th Annual International Meeting, SEG, Expanded Abstracts*, pages 1240–1243, 1995.

[11] J. H. Friedman and W. Stuetzle. Projection pursuit regression. *Journal of the American Statistical Association*, 76:817–823, 1981.

[12] M. Grant and S. Boyd. Graph implementations for nonsmooth convex programs. In V. Blondel, S. Boyd, and H. Kimura, editors, *Recent Advances in Learning and Control*, Lecture Notes in Control and Information Sciences, pages 95–110. Springer-Verlag Limited, 2008. `http://stanford.edu/~boyd/graph_dcp.html`.

[13] J. Gridley and G. Partyka. Processing and interpretational aspects of spectral decomposition. *67th Annual International Meeting, SEG, Expanded Abstracts*, pages 1055–1058, 1997.

[14] A. Grossmann and J. Morlet. Decompositions of hardy functions into square-integrable wavelets of constant shape. *SIAM Journal on Mathematical Analysis*, 15:723–736, 1984.

[15] D. Hoffman and D. Kouri. Hierarchy of local minimum solutions of Heisenberg's uncertainty principle. *Physical Review Letters*, 85(25):5263–5267, 2000.

[16] L. K. Jones. On a conjecture of huber concerning the convergence of projection pursuit regression. *The Annals of Statistics*, 15(2):880–882, 1987.

[17] S. Kim, K. Koh, M. Lustig, Boyd S., and Gorinevsky D. An interior-point method for large-scale l_1-regularized least squares. *IEEE Journal of Selected Topics In Signal Processing*, 1(4):606–617, 2007.

[18] R. LeBras, G. Mellman, and M. Peters. A wavelet transform for downward continuation. *62nd Annual International Meeting, SEG, Expanded Abstracts*, pages 889–892, 1992.

[19] X.-G. Li and T. J. Ulrych. Tomography via wavelet transform constraints. *65th Annual International Meeting, SEG, Expanded Abstracts*, pages 1070–1073, 1995.

[20] X.-G. Li and T. J. Ulrych. Multi-scale attribute analysis and trace decomposition. *66th Annual International Meeting, SEG, Expanded Abstracts*, pages 1634–1637, 1996.

[21] S. Mallat. *A Wavelet Tour of Signal Processing*. Academic Press, 2009.

[22] S. Mallat and Z. Zhang. Matching pursuits with time-frequency dictionaries. *IEEE Trans. Signal Process.*, 41(12):3397–3415, 1993.

[23] S. Mehrotra. On the implementation of a primal-dual interior point method. *SIAM J. Optimization*, 2:575–601, 1992.

[24] X. G. Miao and Moon W. M. Application of the wavelet transform in seismic data processing. *64th Annual International Meeting, SEG, Expanded Abstracts*, pages 1461–1464, 1994.

[25] J. Morlet. Sampling theory and wave propagation. *51st Annual International Meeting, SEG, Expanded Abstracts*, 1981.

[26] J. Morlet. *Sampling theory and wave propogation, NATO ASI Series, Vol F1, Isuues in acoustic signal/image processing and recognition*. Springer Verlag, 1983.

[27] J. Morlet, G. Arens, E. Fourgeau, and D. Glard. Wave propagation and sampling theory-part I: Complex signal and scattering in multilayered media. *Geophysics*, 47(2):203–221, 1982.

[28] J. Morlet, G. Arens, E. Fourgeau, and D. Glard. Wave propagation and sampling theory-part II: Sampling theory and complex waves. *Geophysics*, 47(2):222–236, 1982.

[29] D. Nanda. Spectral decomposition of seismic data using conventional μ-wavelets and μ-pseudo-wavelets. Master's thesis, University of Houston, Houston, Texas, August 2010.

[30] A. Neumaier. Solving ill-conditioned and singular linear systems: A tutorial on regularization. *SIAM Review*, 40(3):636–666, 1998.

[31] H. Niitsuma, K. Tsuyuki, and H. Asanuma. Discrimination of split shear waves by wavelet transform. *Journal of the Canadian Society of Exploration Geophysicists*, 29(1):106–113, 1993.

[32] L. V. Nikitin and E. M. Chesnokov. Influence of a stressed condition on the anisotropy of elastic properties in a medium. *Izvestiya, Earth Physics*, 17(3):174–183, 1981.

[33] A Nur and G. Simmons. Stress-induced velocity anisotropy in rock: an experiment study. *Journal of Geophysical Research*, 74(27):6667–6674, 1969.

[34] D. W. Oldenburg, T. Scheuer, and S. Levy. Recovery of the acoustic impedance from reflection seismograms. *Geophysics*, 48:1318–1337, 1983.

[35] A. Ozbek, A. K. Ozdemir, and M. Vassallo. Interpolation by matching pursuit. *79th Annual International Meeting, SEG, Expanded Abstracts*, pages 3254–3258, 2009.

[36] R. Peraldi and A. Clement. Digital processing of refraction data study of first arrivals. *Geophysical Prospecting*, 20:529–548, 1972.

[37] P. V. Riel and A. J. Berkhout. Resolution in seismic trace inversion by parameter estimation. *Geophysics*, 50:1440–1455, 1985.

[38] M. D. Sacchi, D. R. Velis, and A. H. Cominguez. Minimum entropy deconvolution with frequency-domain constraints. *Geophysics*, 59:938–945, 1994.

[39] D. Salomon. *Data compression the complete reference*. Springer, 2004.

[40] J. Smythe, A. Gersztenkorn, B. Radovich, C.-F. Li, and C. Liner. Gulf of mexico shelf framework interpretation using a bed-form attribute from spectral imaging. *The Leading Edge*, 23(9):921–926, 2004.

[41] S. Sun, R. Siegfried, and J. Castagna. Examples of wavelet transform time-frequency analysis in direct hydrocarbon detection. *72nd Annual International Meeting, SEG, Expanded Abstracts*, pages 457–460, 199.

[42] J. Tan. *Classification of microseismic events from bitumen production at Cold Lake, Alberta*. PhD thesis, University of Calgary, Alberta, Canada, 2007.

[43] M. T. Taner. Complex seismic trace analysis. *Geophysics*, 44(6):1041–1063, 1979.

[44] A. Tarantola. *Inverse problem theory and methods for model parameter estimation*. Society for Industrial and Applied Mathematics, 2004.

[45] K. C. Toh and R. H. Tutuncu. Sdpt3 — a matlab software package for semidefinite programming. *Optimization Methods and Software*, 11:545–581, 1999.

[46] R. H. Tutuncu, K. C. Toh, and M. J. Todd. Solving semidefinite-quadratic-linear programs using sdpt3. *Mathematical Programming Ser. B*, 95:189–217, 2003.

[47] T. J. Ulrych, D. R. Velis, and M. D. Sacchi. Wavelet estimation revisited. *The Leading Edge*, 14:1139–1143, 1995.

[48] D. Velis. Stochastic sparse-spike deconvolution. *Geophysics*, 73(1):R1–R9, 2008.

[49] B. Wang and K. Pann. Kirchhoff migration of seismic data compressed by matching pursuit decomposition. *66th Annual International Meeting, SEG, Expanded Abstracts*, pages 1642–1645, 1996.

[50] J. Wang, X. Wang, and M. Perz. Structure preserving regularization for sparse deconvolution. *76th Annual International Meeting, SEG, Expanded Abstracts*, pages 2072–2076, 2006.

[51] W. Wang and P. Mishra. A partitioned bitmask-based technique for lossless seismic data compression. Technical report, University of Florida, FL, 2008.

[52] K. Wapenaar. Seismic reflection and transmission coefficients of a self-similar interface. *Geophysical Journal International*, 135(2):585–594, 1998.

[53] G. Wei, H. Wang, D. Kouri, M. Papadakis, I. Kakadiaris, and D. Hoffman. On the mathematical properties of distributed approximating functionals. *Journal of Mathematical Chemistry*, 30(1):83–107, 2001.

[54] J. Wong, L. Han, J. Bancroft, and R. Stewart. Automatic time-picking of first arrivals on noisy microseismic data. *CSEG National Convention*, 2009.

[55] Li X.-G., Wang B., Anderson J., and Deng L. Fast migration using a matching pursuit algorithm. *68th Annual International Meeting, SEG, Expanded Abstracts*, pages 1732–1735, 1998.

[56] D. Zhang, D. Kouri, D. Hoffman, and G. Gunaratne. Distributed approximating functional treatment of noisy signals. *Computer Physics Communications*, 120:1–12, 1999.

[57] R. Zhang. *Basis pursuit seismic inversion*. PhD thesis.

[58] R. Zhang and J. Castagna. Seismic sparse-layer reflectivity inversion using basis pursuit decomposition. *Geophysics*, 76(6):147–158, 2011.

[59] G. Zhu and Q. Li. The wavelet transform and thin bed analysis. *61st Conference and Exhibition, EAGE, Extended Abstracts*, page P002, 1999.