Multichannel Semi-blind Sparse Deconvolution of Seismic Signals

Merabi Mirel

Multichannel Semi-blind Sparse Deconvolution of Seismic Signals

Research Thesis

As Partial Fulfilment of the Requirements for
the Degree of Master of Science in Electrical Engineering

Merabi Mirel

Submitted to the Senate of the Technion—Israel Institute of Technology

Iyar 5778 Haifa April 2018

Acknowledgement

The research thesis was done under the supervision of Professor Israel Cohen in the Department of Electrical Engineering. I would like to thank Prof. Cohen for his personal guidance and supervision throughout the process. His experience in this research field had significant influence and it was reflected in the professional comments and suggestions to the research.
I would also like to thank Anthony A. Vassiliou from GeoEnergy for his valuable inputs to the research.

I would like to express my big love and gratitude to my family and my wife, Shira, for their care, patience and support. Thank you for your constant love and encouragementt.

Contents

List of Figures

List of Tables

List of Publications

1. M. Mirel and I. Cohen, "Multichannel Semi-blind Deconvolution of Seismic Signals", Signal Processing, Vol. 135, June 2017, pp. 253-262.

2. M. Mirel, I. Cohen and A. A. Vassiliou, "Multichannel Semi-blind Sparse Deconvolution of Seismic Signals", Proc. Society of Exploration Geophysicist International Conference, Exposition and 87th Annual Meeting, Houston, Texas, 24-27 September 2017.

Abstract

Seismic waves are waves that can propagate in the subsurface and are reflected from underground layers of the earth. Reflection seismology involves emission of seismic waves by a seismic source on the earth's surface, and obtaining the reflected waves by an array of geophones. The reflected signals are measured on the earths' surface and the measurements combine the seismic data. Important features are extracted from this data in order to better understand the subsurface structure. The kernel defining the channel is called wavelet and the signal describing the reflected impulses is called reflectivity.
Convolution is a well-defined operator that is used to mathematically model the connection between the input and output signals of an LTI system given a third signal, called the kernel, that describes the system.
Deconvolution is a process where we hold the output signal and the kernel and try to find the input signal, this is called the non-blind deconvolution. Sometimes the kernel is also unknown and then the problem is called blind deconvolution.
Seismic deconvolution is a general problem associated with recovering the reflectivity series from the seismic data when the wavelet is known. Many methods and algorithms were presented for solving this kind of problem, using different kind of modelings and assuming different a-priori knowledge on the signal and on the system.
Many deconvolution methods were developed to solve the seismic deconvolution problem. When partial information is given on the wavelet, the blind deconvolution methods can ignore it and solve it as a blind problem. Alternatively, non-blind deconvolution methods can assume the partial information is the full one if possible or make an educated guess to fill the missing information and solve it as a non-blind problem.
In this research, we solve the problem of semi-blind seismic deconvolution, where the wavelet is known up to some error. The multichannel semi-blind deconvolution (MSBD)

model was developed for cases where there is some uncertainty in the assumed wavelet. We model the wavelet uncertainty as an additive noise to the wavelet and analyze that noise with our developed method. We present a novel, two-stage iterative algorithm that recovers both the reflectivity and the wavelet. While the reflectivity series is recovered using sparse modeling of the signal, the wavelet is recovered using L2 minimization, exploiting the fact that all channels share the same wavelet. The L2 minimization solution is revised to suit the multichannel case.

As mentioned before, our method assumes an additive noise to the wavelet. We test 2 use case. In the first one we test the straightforward case where an additive white Gaussian noise is added to the wavelet, and in the second we test the case where we are uncertain with a parameter defining the wavelet. In the second case the noise is not additive to the true wavelet but we show how we model it so we can apply our general method on this case also. We show that our algorithm outperforms both the non-blind and the blind methods.

Notation

$\mathbf{DW}_0$	Convolution Matrix of $w\left(t; f_0\right)$
$\mathbf{DW}_1$	Convolution Matrix of $\frac{\partial w(t;f)}{\partial f}$
$\mathbf{DW}_2$	Convolution Matrix of $\frac{\partial^2 w(t;f)}{\partial f^2}$
f	Frequency parameter of the Ricker wavelet
$h(t)$	General impulse response of an LTI system
i	Channel index
n	Discrete time index
N	Number of channels
N_r	Length of the reflectivity series
N_w	Length of the wavelet
$r[n]$	Reflectivity series
$r_i[n]$	Reflectivity series in channel number i
$\mathbf{r}_i$	Vector form of $r_i[n]$
$s[n]$	Seismic data
$s_i[n]$	Seismic data in channel number i
$\mathbf{s}_i$	Vector form of $s_i[n]$
t	Continuous time index
$v[n]$	Noise in the channel
$v_i[n]$	Noise in channel number i
$\mathbf{v}_i$	Vector form of $v_i[n]$
$\mathbf{v'}_i$	New noise term
$\bar{v}'_i$	First form of total noise
$w[n]$	Wavelet, earth's impulse response
$\mathbf{w}$	Vector form of $w[n]$
$\mathbf{W}$	Convolution matrix of $\mathbf{w}$
$\mathbf{W}_0$	Convolution matrix of the initial wavelet vector

$w(t; f_0)$	Wavelet in time domain
$\frac{\partial w(t;f)}{\partial f}$	First derivative of $w(t; f_0)$ with respect to f
$\frac{\partial^2 w(t;f)}{\partial f^2}$	Second derivative of $w(t; f_0)$ with respect to f
$\mathbf{w}'$	Vector form of initial wavelet
$\mathbf{w}_n$	Vector form of the additive noise to the wavelet
$\mathbf{W}'$	Convolution matrix of $\mathbf{w}'$
$\mathbf{W}_n$	Convolution matrix of the wavelet error vector $\mathbf{w}_n$
$\mathbf{w}^{\text{opt}}$	Recovered wavelet
$\mathbf{w_{init}}$	Initial wavelet
$\mathbf{w_{init}}$	Current recovered wavelet (At a specific iteration)
$x(t)$	General input signal to an LTI system
$\hat{\mathbf{x}}$	Recovered signal
$y(t)$	General output signal of an LTI system
ϵ	Trade-off parameter of second form of the optimization problem
λ	Trade-off parameter of first form of the optimization problem
σ_f	Standard deviation of the parameter defining the wavelet
σ_v	Noise standard deviation
$\sigma_{\bar{v}'_i}$	Standard deviation of the total noise
σ_w	Standard deviation of the additive noise to the wavelet
σ'_w	Approximated standard deviation of the additive noise
$\mathcal{N}(0, \sigma_v^2)$	Normal Distribution
$\|\mathbf{x}\|_0$	L0 norm of $\mathbf{x}$
$\|\mathbf{x}\|_1$	L1 norm of $\mathbf{x}$
$\|\mathbf{x}\|_2$	L2 norm of $\mathbf{x}$

Abbreviations

ARMA	Auto-Regressive Moving Average
AWGN	Additive white Gaussian Noise
BG	Bernoulli-Gaussian
BPD	Basis Pursuit Decomposition
BPDN	Basis Pursuit Denoising
BPI	Basis Pursuit Inversion
IID	Independent and Identically Distributed
IWM	Iterated Window Maximization
LTI	Linear and Time Invariant
MBG	Markov Bernoulli-Gaussian
MPD	Matching Pursuit Decomposition
MPM	Modified Posterior Mode
MSBD	Multichannel Semi-Blind Deconvolution
SMBD	Sparse Multichannel Blind Deconvolution
SMLR	Single Most Likely Replacement
SNR	Signal to Noise Ratio
SOOT	Smoothed One-Over-Two
SSI	Sparse Spike Inversion

Chapter 1

Introduction

1.1 Motivation and Goals

Convolution is a mathematical operator which has many applications in the fields of physics and engineering. It operates on two functions and creates a third function. The convolution between $f(t)$ and $g(t)$ is denoted by $f * g$ and is defined in the following way,

$$(f * g)(t) \stackrel{def}{=} \int_{-\infty}^{\infty} f(\tau)g(t-\tau)d\tau. \tag{1.1}$$

In the discrete time domain the convolution between $f[n]$ and $g[n]$ will be as follows,

$$(f * g)[n] \stackrel{def}{=} \sum_{-\infty}^{\infty} f[m]g(n-m). \tag{1.2}$$

Its most common application is to describe models of linear and time invariant(LTI) systems. As we can see in Figure 1.1, if we denote the impulse response of an LTI system as $h(t)$ and the input signal as $x(t)$ then the output signal of the system, $y(t)$, is the convolution between $x(t)$ and $h(t)$,

$$y(t) = x(t) * h(t). \tag{1.3}$$

Of course that the same holds for the discrete time domain. This description of LTI systems is the basis of signal processing. Any described LTI system in physics and engineering researched from the point of signal processing will utilize the above described formulation.

The ability to mathematically formulate systems can be used for many purposes. For

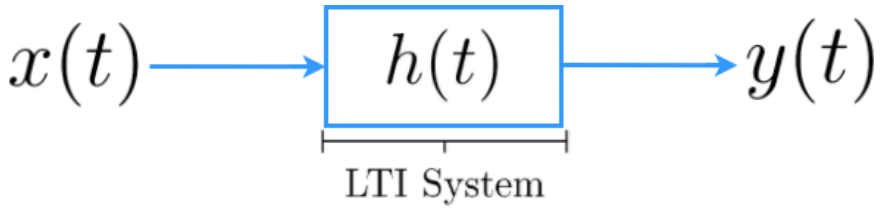

Figure 1.1: LTI system.

example, we can simulate a given system and analyze expected outcomes given certain inputs. The common to all these usages is the wish to analyze the output signal, $y(t)$, while knowing or assuming the system, described by $h(t)$, and/or the input signal, $x(t)$.

A very interesting follow-up question will be, can we reverse the system using knowledge on its outputs? Or in other words, can we learn anything about the system, $h(t)$, and/or the input signal, $x(t)$, while measuring or assuming the output signal, $y(t)$?
The process of reversing an LTI system is called Deconvolution. Deconvolution is not a specific operator nor has specific steps to apply, as opposed to the convolution operator. The deconvolution process needs to be defined per system and adjusted to its specific input signals and characteristics.
The field of deconvolution algorithms is of main interest because of its possible contributions. For example, it can help us remove affects of undesired convolutions applied from systems not under our control, help us better understand systems not developed by us, or mathematically describe and simulate nature systems.
The common to all deconvolution algorithms is that the output signal, $y(t)$, is known. If in addition we know the system, $h(t)$, and aim to find the input signal, $x(t)$, then this problem is called the non-blind deconvolution problem. If we know the input signal and aim to find the system, $h(t)$, then this problem is called the kernel estimation problem. It is very common to call the impulse response of an LTI system, $h(t)$, as the kernel of the system. If not the input signal, nor the system are known and we aim to find both of them then this is called the blind deconvolution problem. Of course that in the last case we will need some extra information. The most common type of extra information is derived from a multichannel setup. That means we have several input signals going through the same system and we hold measurements of all the corresponding outputs.

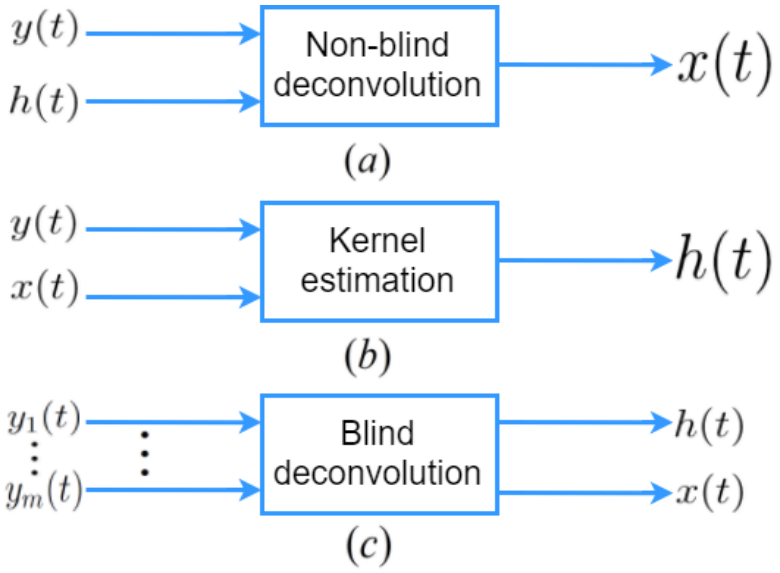

Figure 1.2: Deconvolution Problems: (a) Non-Blind Deconvolution, (b) Kernel Estimation, (c) Blind Deconvolution.

The different problems are described in Figure 1.2

Deconvolution problems have been widely researched over the last years in many fields in physics and engineering. Our research focuses on seismic signals from the geophysics field. An overview on different methods for seismic deconvolution can be found on Section 2.1. Also, we discuss with further details on seismic signals in the last part of this section.

There is a more generalized way to look on the deconvolution problem. According to Figure 1.2 the difference between blind and non-blind deconvolution is the knowledge of the kernel, $h(t)$. A semi-blind deconvolution defines a problem that has some knowledge on the kernel but not all of it. An example for additional knowledge on the kernel in a semi-blind deconvolution problem can be the mathematical model of the kernel but without the parameters defining it, or it can be the true kernel contaminated with some noise. Any additional information that can be mathematically formulated and imposed on the solution can fit into the semi-blind type of the deconvolution problem. A solution to a semi-blind problem will be even more specific to a certain problem and model because of the additional information on the kernel.

Seismic blind and non-blind deconvolution problems have a variety of solutions leaning on different assumptions, mathematical models and mathematical tools. In our research

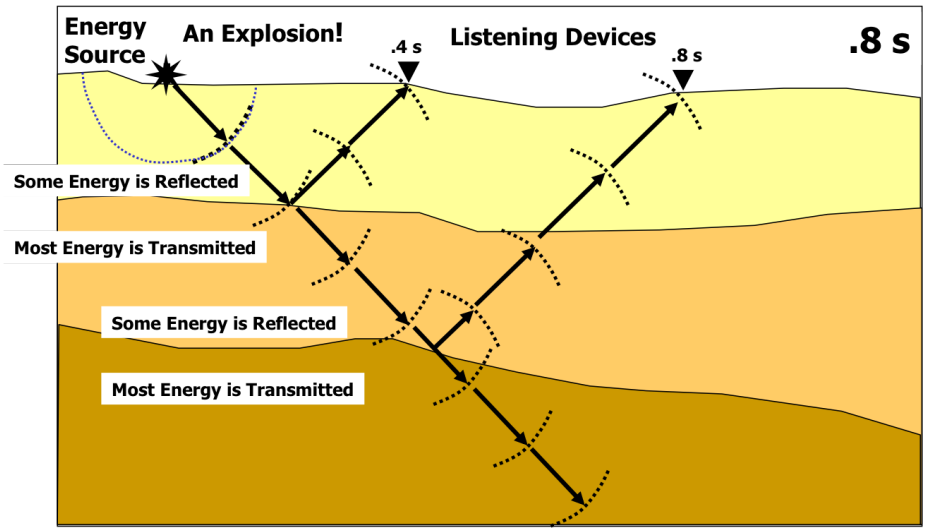

Figure 1.3: Seismic refraction and reflection. (from [1])

we would like to test and find an algorithm best suited to a general semi-blind seismic deconvolution problem.

To do that, we need to define a mathematical model for the signals in interest. As mentioned before, our interest is in seismic signals. Before we do that, we shortly review seismic waves and reflection seismology.

Seismic waves are a form of energy that travel through the earths layers. They can be a result of different natural phenomenons, such as earthquakes, volcanic eruptions, magma movement, large landslides, etc. They can also be a result of deliberated man-made explosions. These waves propagate in the underground layers of the earth with a velocity that depends on the density and elasticity of the medium. Those characters are mathematically translated into a parameter that is called acoustic impedance, which is a measure of the opposition that a system presents to the acoustic flow resulting of an acoustic pressure applied to the system. When a seismic wave traveling through the earth encounters an interface between two areas with different acoustic impedance, some of its energy is reflected and some of it is refracted as can be observed in Figure 1.3.

Reflection seismology is a field which uses the properties of seismic waves to explore the subsurface and reconstruct an image of it. Seismic waves are generated using an energy source such as dynamite explosion. The reflected signals are received and

measured by geophones, which in time converts ground motion into an analog electrical signal. Each recorded measure is called a trace. All traces are combined into what is called seismic data after a process of stack and migration. Seismic deconvolution is the process where we take the seismic data and remove the distortion created by the propagation in the medium. The distortion is modeled as a convolution between the desired reflectivity image and a signal defining the medium, called wavelet.

In the literature, $x(t)$ in Figure 1.1 is called the reflectivity signal, $h(t)$ is called the wavelet and $y(t)$ is called the seismic data. As said before, in order to best suit the deconvolution algorithm for our particular problem, we first need to mathematically formulate the reflectivity signal. In this research we model the reflectivity signal as a sparse signal. Sparse modeling of signals has been widely researched and novel algorithms of deconvolution have been presented relying on those models.

A sparse signal, or a sparse vector, is a vector in which most of the elements are zero. To be more precise, the sparsity of a vector is defined by the ratio of the number of zero elements to the total number of elements in it. Sparse deconvolution methods utilize the sparse modeling of the signals and rely on the sparsity when applying deconvolution methods. Theoretically, any type of deconvolution problem can retrieve the desired signal, but in the real world we hold noisy observation of the output signals of the LTI system and sometimes we observe systems that are non-invertible. The main problem of the legacy deconvolution algorithms is that in most cases, when not carefully applied, they can greatly amplify that noise. When using sparse deconvolution algorithms, as is detailed in later chapters, we can design a more general and generic algorithm that is less affected by the problem mentioned above. The sparse representation clearly best defines the reflectivity signal so sparse deconvolution methods have the most potential in recovering it.

When dealing with a blind or semi-blind deconvolution problem it is very common to use an iterative method. We have 2 sets of unknowns, the input signal and the kernel, or the reflectivity signal and the wavelet if talking in terms of seismic signals. To analytically

ease the solution at each iteration we usually have 2 steps, at each step we assume to know one unknown and use non-blind deconvolution methods to recover the other unknown. When designing this kind of solution to the blind or semi-blind deconvolution problem the important thing is to "stitch" those two steps together correctly with wise parameters calculations when moving from one step to another. Wise adaptions of the non-blind deconvolution methods combined with an iterative solution can compose a state-of the art solution to any blind or semi-blind deconvolution problem without going into deep analytical analysis of recovering both unknowns simultaneously.

1.2 Thesis Structure

The rest of the thesis is organized as follows: In Chapter 2 we review the different common seismic deconvolution methods. We start with the Non-Blind deconvolution, where we review two state-of-the-art methods for the Non-Blind case, also utilizing sparse modeling of the reflectivity signal. We go into deep mathematical analysis in those methods because they will affect one of the core steps in our proposed algorithm. In addition, we review different methods of the multichannel case, which is usually the setup when we meet a Blind deconvolution problem. In the chapter after that, Chapter 3, we present the specific semi-blind deconvolution problem discussed in this thesis. We formulate the problem and discuss our offered solution throughout all its steps.

In addition, we review two different use-cases of MSBD. We analyze both of them mathematically, fit them in the right formulation of the problem, as developed and presented in the previous section of the problem formulation and analyze the results in different levels of details. We discuss general performances of each case and compare the results to other seismic deconvolution methods. In the last chapter, Chapter 5, we summarize the work and research that was done in this thesis and review some future possible research questions we indicated throughout the process.

Chapter 2

Seismic Deconvolution

Seismic deconvolution is the field that is responsible for recovering the reflectivity signal and/or the wavelet out of the seismic data. Different approaches and models of the same problem can lead to different performances. When developing a solution for a problem in this field, one of the first things to determine is whether the solution is going to be based on a stochastic or deterministic model as well as defining the parameters. Our method focuses on deterministic modeling of the problem, but first we review different methods from both disciplines with greater emphasis on deterministic modeling methods.

In this chapter we review different approaches to seismic deconvolution, we shortly go over a variety of methods developed over the years. Then we formulate the problem mathematically and describe two of these methods with great details. These two methods describe a certain solution for the non-blind deconvolution problem and they are used in our algorithm which is presented later on. In the last part of this chapter we review different methods and approaches to the blind deconvolution problem. It is important to review methods from both non-blind and blind deconvolution disciplines because our research deals with the semi-blind deconvolution problem, so ideas, performances, advantages and disadvantages from both disciplines can give us quantitative and qualitative parameters to compare to.

2.1 Background of Seismic Deconvolution

First we briefly describe the different stochastic modeling methods. Kormylo and Mendel (1982) estimate the wavelet using ARMA (Auto-Regressive Moving Average) and SMLR (Single Most Likely Replacement) for the reflectivity estimation while assuming a BG (Bernoulli-Gaussian) model [2]. Both algorithms use second-order statistics. The SMLR is an iterative search algorithm which is designed for a specific case but it can be easily extended to more general cases. One has to notice a significant drawback of this algorithm, the SMLR must be performed off-line. Kaaresen and Taxt (1998) also assume the BG model but they use the IWM (Iterated Window Maximization) algorithm for reflectivity estimation and the Least-Squares method for wavelet estimation [3]. They were one of the first to a-priori assume sparse reflectivity signal which corresponds to a layeredearth model. In order to not analytically complicate the solution they use a sub-optimal iterative procedure for the optimization problem. Heimer et al. (2009) proposed a blind multichannel deconvolution method based on the statistical properties of the signal [4]. Specifically, it is based on the Markov-Bernoulli random field modeling. Their method accounts for layer discontinuities resulting from splitting, merging, starting or terminating layers within the region of interest. The computational complexity of the proposed algorithm is generally higher than that of the IWM algorithm but under some configurations this complexity can be reduced. Ram et al. (2010) also propose a method based on the statistical properties of the signal [5], where the spatial dependency between neighboring traces is exploited by a priori assuming 2D reflectivity. The algorithm is based on the MBG (Markov Bernoulli-Gaussian) reflectivity model. In this research they present two multichannel blind deconvolution algorithms. The first algorithm uses in each step a modified posterior mode (MPM) algorithm which estimates a reflectivity signal from a specific channel using the estimation of the preceding channel and the current seismic data observed on the current channel. The second algorithm uses estimations of both the preceding and subsequent channels. The second algorithm performs better than the first one because it uses more information from neighboring channels.

Methods using higher order statistics are also very important to mention. For

example, Velis and Ulrych (1999) propose a fourth-order cumulant matching method [6]. They use the mean-square error as a measure for matching between the trace cumulant and the wavelet moment under the assumption of a non-Gaussian, stationary, and statistically independent reflectivity series. This leads to an optimization problem that is highly nonlinear and is usually solved with some kind of linearization. Alternatively, they suggest a hybrid strategy that uses simulated annealing algorithm to provide reliability of the numerical solutions by reducing the risk of being trapped in local minima. The method is robust enough to detect time-varying phase changes and then phase-only corrections can be made by means of a time-varying phase rotation. Van der Baan (2009) also uses a higher order statistics for wavelet estimation and deconvolution [7]. He maximizes the kurtosis measure for phase estimation in cases of phase mismatches. Menanno and Mazzotti (2012) use the vectorial nature of the wavefield [8] and exploit correlation between different channels in the multichannel case in a method of quaternion deconvolution. Their method is an extension of the Wiener filter and uses high order statistics. By acquiring the data by multicomponent sensors they decompose the vector signals in a different way from the standard scalar deconvolution methods.

There are also methods that do not utilize stochastic solutions to seismic inversion and wavelet estimation. Instead, they present a deterministic mathematical model to describe the seismic data and use methods that depend on the deterministic properties of the signal. One common method used in deconvolution problems is the Euclid Deconvolution in which a homogeneous system of equations is solved. The classic Euclid Deconvolution has stability problems in the presence of noise. Concerning solutions based on the deterministic properties of the signal, it has recently become very common to exploit the sparse representation of the reflectivity series. Methods using the sparse properties of the signal have different modeling mechanisms for the reflectivity series. Another major difference is the definition of the optimization problem and as a consequence, the algorithm that is used for solving the problem. Different approaches to this problem result in different performances with some advantages and disadvantages depending on a priori knowledge, sparsity depth, wavelet type, etc.

A modification of the Euclid Deconvolution is presented by Kazemi and Sacchi (2014), who developed the SMBD (Sparse Multichannel Blind Deconvolution) method [9]. They exploit the sparsity property in the reflectivity series to make the Euclid Deconvolution more robust to noise and generally improve the method. In practice, they define an L2 minimization problem with a regularization term and a requirement on the sparsity of the signal. Their proposed method can tolerate moderate levels of noise and does not require any a priori knowledge on the length of the wavelet. Because of the blind setup of the problem, the wavelet estimation is a by-product of the process. The homogenous system of equations are solved with sparsity constraints via a solver that uses the steepest descent method and to avoid the trivial solutions they constrain the reflectivity signal to have a unity norm.

Repetti et al. (2015) [10] present in a new study a different approach to the L1/L2 minimization problem for the blind deconvolution case. They propose a new penalty, Smoothed One-Over-Two (SOOT), based on a smooth approximation of the L1/L2 function. The SOOT penalty enables them to avoid the problems raised from solving non-convex and non-smooth optimization problems. They develop a proximal-based algorithm to solve the minimization problem and derive theoretical convergence. They show the algorithm is quite effective with blind deconvolution problems with the advantage of guaranteed convergence.

Furthermore, there are strong connection between blind deconvolution and blind compressed sensing, which presents the basic idea of recovering a sparse signal from a small number of linear measurements. Blind compressed sensing is widely studied and discussed by Gleichman and Eldar (2011) [11]. They suggest different constraints on the sparsity basis which allow them to guarantee a unique solution, while avoiding the need of prior knowledge of the sparsity basis, which is essential for the recovery process. They introduce a general sampling and reconstruction process which can be suited for all the signals that have a sparse representation and are under the conditions and restraints presented in their work. Another interesting work on blind compressed sensing is introduced by Rosenbaum and Tsybakov (2010) [12], which also account for matrix uncertainty. In their work they suggest new estimators, since they find previous estimators as unstable. Their main

conclusion of choosing smartly the regularization parameter as a key to success recovery supports our findings that are discussed later on regarding the regularization parameters .

Nguyen and Castagna (2010) presented a method that exploits the sparse properties of the reflectivity series using Matching Pursuit Decomposition (MPD) [13]. MPD involves a few steps and eventually decomposes the seismic data into a superposition of wavelet atoms generated from the locations, amplitudes and scaling (physically translated into different center frequencies) of a base wavelet form. The method correlates a wavelet dictionary with the data and marks the parameters iteratively, recording the best-fit wavelet in each iteration. It is very important to keep the wavelet dictionary orthogonal in this case. Wang also uses MPD in his first step [14] [15]. In the next steps he is exploiting lateral coherence as a constraint with hope to improve the uniqueness of the solution.

Zhang and Castagna (2011) presented a method based on Basis Pursuit Decomposition (BPD) for seismic inversion [16]. They used an algorithm presented by Chen et al. (2001) [17] [18] to solve the Basis Pursuit problem, which is an L1 optimization problem. They also utilized a special dictionary form. The special form results from the Dipole Decomposition process on the reflectivity series. Dipole Decomposition is a method that decomposes the reflectivity series into a summation of even and odd impulse pairs, weighted by different amplitudes. The BPD method is presented by Bork and Wood (2001) [19]. A brief summary of the method can be found in [16]. The basis pursuit algorithm is an L1 optimization problem that can be solved in different ways. The great advantage of BPD over MPD comes from the process used to solve the problems. MPD is an iterative process that extracts the best-fit wavelet atom, subtracts it from the data, finds the next best-fit atom, etc. The solution therefore depends on the order of the wavelet atoms in the wavelet dictionary, so different ordering of the same wavelet dictionary can lead to different solutions. Conversely, BPD is not affected by the ordering of the wavelet dictionary and obtains one solution for all the different combinations that eventually construct the same wavelet dictionary. BPD has more advantages over MPD, including interference handling, computational efficiency, and good stability even when

the wavelet dictionary is not orthogonal. More information and analytical developments about Basis Pursuit and Matching Pursuit are widely discussed in [20].

The inversion method that uses BPD is called Basis Pursuit Inversion (BPI) [16] [21] [22]. Before BPI was commonly used, another inversion method was in use, also exploiting sparse properties of the reflectivity series and solving the L1 optimization problem. This method is called Sparse Spike Inversion (SSI) [23] [24] [25]. SSI does not utilize any special dictionary, unlike BPI. There are some differences between BPI and SSI in the method used to solve the optimization problem, but the main difference is in the dictionary. While SSI uses a dictionary created from the direct formulation of the problem, BPI uses a dictionary created after a Dipole Decomposition process is performed on the reflectivity series.

There are two more important parameters used in the literature to classify the different kinds of problems. The first is the number of channels. Basically, a distinction is made between single-channel problems and multichannel problems. Single channel means that we have only one sensor that can sample the seismic trace. Multichannel means that there are multiple sensors spread over some area (close to one another) that simultaneously sample the seismic traces. The multichannel approach can give us extra information because of the correlation between the different traces. There are several potential causes of this correlation. For example, different channels share the same kernel or impulses generated from the same layer but sampled in two (or more) close points in the field. The second parameter used for classification is knowledge about the kernel. Basically, the literature deals with two main categories - problems that use full knowledge of the kernel, called non-blind deconvolution problems, and problems that assume nothing about the kernel, called blind deconvolution problems. Of course, there are problems that fall somewhere along the spectrum between not knowing anything or knowing everything about the kernel, but this classification of the problem is made by a more specific definition.

As is shown later on in greater detail, we define the problem as a Multichannel

Semi-Blind Deconvolution problem. This means that we assume to have an array of sensors located close enough to each other, to get some correlation between different seismic traces, and we also assume to know a noisy version of the kernel. We combine methods of kernel estimation and deconvolution to solve this problem.

Semi-blind deconvolution problem is not that common in seismic signals. We can find semi-blind problems in other disciplines of signal processing such as image processing, specifically on image denoising problems. That is why it will be very interesting to see how semi-blind setup can impact sparse signals and seismic deconvolution.

In the next sections we define the non-blind deconvolution problem and review with greater details some of the methods suggested to this problem. Later on we define the semi-blind problem and present our solution to it.

2.2 Problem formulation

We denote the earth's impulse response, the wavelet, by $w[n]$. The reflectivity series and the seismic data are denoted by $r[n]$ and $s[n]$, respectively. The input-output relation between the reflectivity series, the wavelet and the seismic data are given by

$$s[n] = r[n] * w[n] + v[n] \tag{2.1}$$

where $*$ is the well known convolution operator and $v[n]$ are Independent and Identically Distributed (IID) additive white Gaussian noise (AWGN), i.e., $v[n] \sim \mathcal{N}(0, \sigma_v^2)$.

We assume an array of N seismic sensors, where all channels share the same wavelet and the noise in the channels are statistically independent and identically distributed. Denoting i as the channel index we get the following set of input-output relations:

$$s_i[n] = r_i[n] * w[n] + v_i[n],\ 1 \leqslant i \leqslant N. \tag{2.2}$$

We can write (2.2) in the following vector-matrix form:

$$\mathbf{s}_i = \mathbf{W}\mathbf{r}_i + \mathbf{v}_i,\ 1 \leqslant i \leqslant N \tag{2.3}$$

where $\mathbf{s}_i \in \mathbb{R}^{N_r+N_w-1}$ is a vector representation of the seismic signal $s_i[n],\ 1 \leqslant i \leqslant N,\ 1 \leqslant n \leqslant N_r+N_w-1$, $\mathbf{W} \in \mathbb{R}^{(N_r+N_w-1)\times N_r}$ is the convolution matrix of $w[n]$, $\mathbf{r}_i \in \mathbb{R}^{N_r}$ is a vector representation of the reflectivity signal $r_i[n], 1 \leqslant i \leqslant N,\ 1 \leqslant n \leqslant N_r$ and $\mathbf{v}_i \in \mathbb{R}^{N_r+N_w-1}$ is a vector representation of the noise signal $v_i[n],\ 1 \leqslant i \leqslant N,\ 1 \leqslant n \leqslant N_r + N_w - 1$.

The goal of the basic problem is to recover $\mathbf{r}_i$ from $\mathbf{s}_i$ while assuming full knowledge of $\mathbf{W}$ and σ_v. This problem has been widely investigated and is called the non-blind deconvolution problem. A wide variety of solutions have been proposed to solve this problem, depending on the model of the signal $\mathbf{r}_i$.

2.3 Single Channel Deconvolution

In this section we discuss single channel seismic deconvolution. As was briefed in the previous chapter, there are a lot of methods for dealing with this kind of problem. We focus on two methods, SSI and BPI, as they exploit the sparse properties of the reflectivity signal and analyze the problem from a very interesting mathematical point of view.

In seismic deconvolution, assuming a sparse model for $\mathbf{r}_i$, sparse deconvolution methods have been proposed. One of them is SSI which is demonstrated in the next section. In the literature, when not talking on seismic signal this method is referred as Basis Pursuit Denoising (BPDN) [20] [26] [27] [28]. Referring to seismic signals we have BPI which is basically the same as SSI but it utilizes a different model for the optimization problem. This is detailed in next sections. BPDN, or SSI when referring to seismic signals, is an approach that solves the following optimization problem:

$$\hat{\mathbf{x}} = \underset{\mathbf{x}}{\operatorname{argmin}} \frac{1}{2} \|\mathbf{A}\mathbf{x} - \mathbf{y}\|_2^2 + \lambda \|\mathbf{x}\|_1 \tag{2.4}$$

when we know that $\mathbf{y} = \mathbf{A}\mathbf{x} + \mathbf{b}$.

Minimization of the term $\|\mathbf{A}\mathbf{x} - \mathbf{y}\|_2^2$ maintains fidelity to the observations, and minimization of the term $\|\mathbf{x}\|_1$ maintains sparsity of the recovered signal. The parameter λ controls the trade-off between them. The minimization problem can also be presented in

the following form:

$$\hat{\mathbf{x}} = \underset{\mathbf{x}}{\operatorname{argmin}} \ \|\mathbf{x}\|_1 \quad \text{s.t.} \quad \|\mathbf{A}\mathbf{x} - \mathbf{y}\|_2 < \epsilon \tag{2.5}$$

where ϵ controls the abovementioned trade-off.

Without getting into details one should notice that the above presented form of the minimization problem is a relaxation of the following form:

$$\hat{\mathbf{x}} = \underset{\mathbf{x}}{\operatorname{argmin}} \ \|\mathbf{x}\|_0 \quad \text{s.t.} \quad \|\mathbf{A}\mathbf{x} - \mathbf{y}\|_2 < \epsilon. \tag{2.6}$$

As we are talking on sparsity of the signal as the core of the mathematical modeling of the reflectivity signal, the $\|\mathbf{x}\|_0$ term best describes and takes care of the sparsity of the recovered signal. However, we are still using $\|\mathbf{x}\|_1$, and this is because the problem defined with $\|\mathbf{x}\|_0$ is non-convex, has stability issues and no guarantees even to the theoretical solution. In contrast, when using $\|\mathbf{x}\|_1$ we can cast it as a linear programming problem and as such, it can be solved using modern interior-point methods for example.

2.3.1 SSI - Sparse Spike Inversion

Here we present some results of SSI on general signals and on seismic signals as well. As mentioned before, SSI is about solving (2.5). We solve this optimization problem using the SPGL1 solver which is presented in `http://www.cs.ubc.ca/~mpf/spgl1/`, and is according to [29] [30].

First, we create some random data with no specific model as a reference. $\mathbf{A}$ is a matrix of size M $\times$ N with IID and normally distributed elements. i.e.,

$$\text{IID} \quad A_{m,n} \sim \mathcal{N}(0,1) \tag{2.7}$$

where m is in range of 1 to M and n is in range of 1 to N. The input signal, $\mathbf{x} \in \mathbb{R}^{N\times 1}$, is sparse where each element of $\mathbf{x}$ has p probability of being non-zero. If the element is non-zero, it is normally distributed. i.e., define a binary vector $a \in \mathbb{R}^{N\times 1}$,

$$P(a[n] = 1) = p \tag{2.8}$$

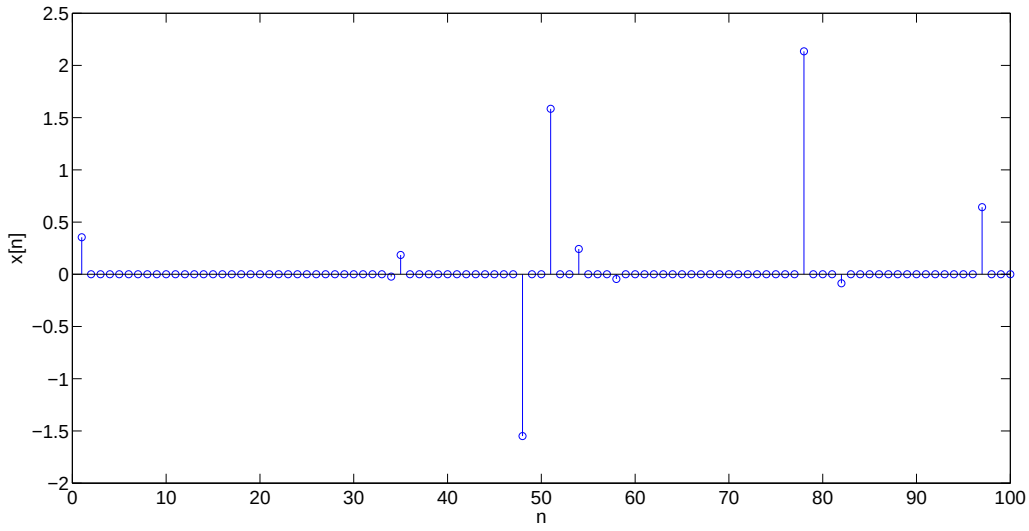

Figure 2.1: Example for input signal: $M = 50$, $N = 100$, $p = 0.1$.

$$P(a[n] = 0) = 1 - p. \tag{2.9}$$

Now define a vector $\mathbf{b}$ with IID and normally distributed elements,

$$b[n] \sim \mathcal{N}(0, 1) \tag{2.10}$$

and now define the elements of $\mathbf{x}$ as,

$$x[n] = a[n] \cdot b[n]. \tag{2.11}$$

First, we simulate $\mathbf{y} = \mathbf{A}\mathbf{x}$ and use SSI to recover $\mathbf{x}$. In our example we chose the parameters as follows, $M = 100, N = 50, p = 0.1$.

In Figure 2.1 we can see an example to the input signal and in Figure 2.2 we can see the output signal that was created from $\mathbf{x}$ and $\mathbf{A}$ according to $\mathbf{y} = \mathbf{A}\mathbf{x}$. We can clearly see the sparsity of $\mathbf{x}$, although not as expected. We can see 8 non-zeros elements although we expected 10 ($N \cdot p = 10$). This is because we are looking at only one realization of $\mathbf{x}$ and its length is not big enough.

Now we move to present the results of SSI. In this case, because we had no noise in

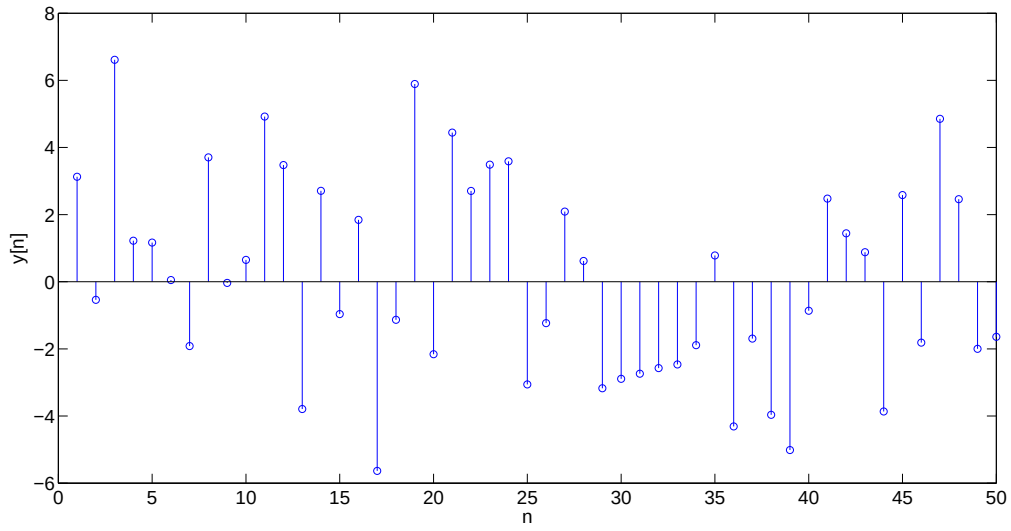

Figure 2.2: Example for output signal: $M = 50$, $N = 100$, $p = 0.1$.

the channel we chose $\epsilon = 0$ in (2.5). i.e., in this case we solve the problem:

$$\hat{\mathbf{x}} = \underset{\mathbf{x}}{\operatorname{argmin}} \ \|\mathbf{x}\|_1 \quad \text{s.t.} \ \ \mathbf{A}\mathbf{x} = \mathbf{y}. \tag{2.12}$$

In Figure 2.3 we can see the recovered signal obtained from solving (2.12). We can see a good recovery of the input signal. We now add a quantitative parameter for the recovery, the cross correlation. We calculate the cross correlation between the input signal and the recovered signal. For the tested case in the previous example we get correlation measure of 1, which means perfect recovery.

Now we test a case that is similar to the previous case but with noise added to the output signal. The new output signal is as follows,

$$\mathbf{y} = \mathbf{A}\mathbf{x} + \mathbf{v} \tag{2.13}$$

where $\mathbf{v}$ is a noise vector where its elements are IID Gaussian noise. We use the same input signal as before and add noise to the output signal, $v[n] \sim \mathcal{N}(0, \sigma^2)$, $\sigma = 0.3$. We can see the contaminated output signal in Figure 2.4 and the recovered signal using SSI with $\epsilon = 0$ in Figure 2.5.

We can clearly see that in the presence of noise the recovery of the input signal is not perfect. The correlation measure in this case is 0.991. But it is clear that we can improve

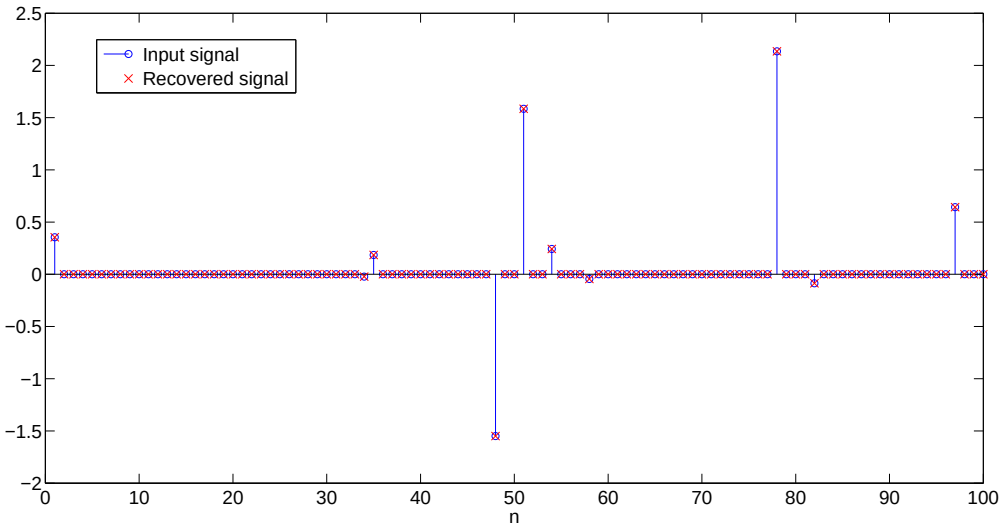

Figure 2.3: SSI. Example for recovered signal: $M = 50$, $N = 100$, $p = 0.1$.

this result because we did not change ϵ at all. As mentioned in previous sections, ϵ has a big connection to the total noise in the problem. We do not go into details on how to evaluate the noise in the problem because this is not clear from the literature and each algorithm suits it for itself. We still would like to see how the affect of ϵ improves our result so we use SSI on the same input signal with different values of ϵ and look for the maximum correlation measure. In Figure 2.6 we can see the correlation measure vs. ϵ for the case presented before. We can see that the maximum correlation is obtained when choosing $\epsilon = 2.5$. The recovered signal using $\epsilon = 2.5$ is shown in Figure 2.7. We can clearly see the improvement in the recovered signal.

2.3.2 BPI - Basis Pursuit Inversion

Here we present the extension of BPI over SSI. We mentioned that BPI utilizes a different mathematical model for the optimization problem. Let us present that model and then discuss some results.

As presented in [16], any reflectivity series can be decomposed into a summation of impulse pairs [19]. BPI utilizes this decomposition to represent the reflectivity series as a linear combination of two impulse pairs, even and odd. If we look at the layers composing

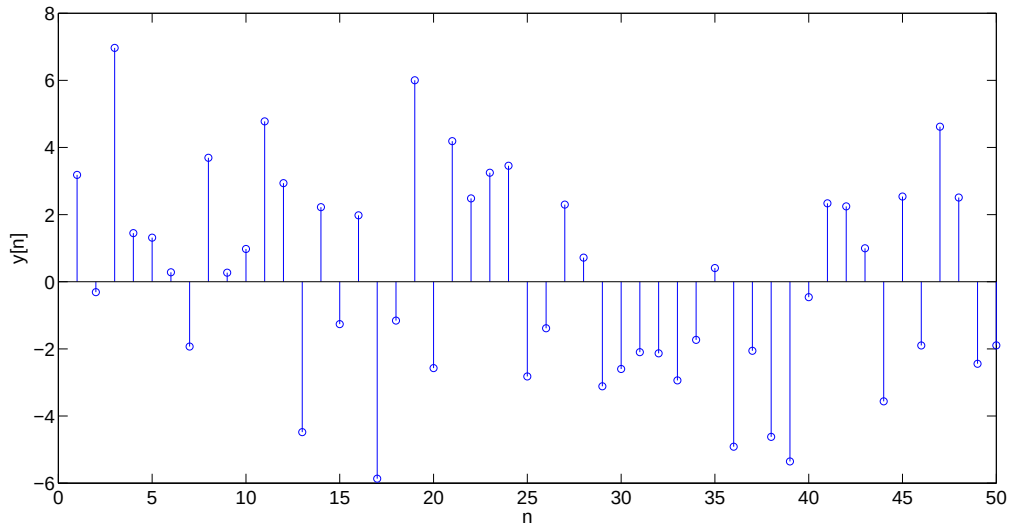

Figure 2.4: SSI. Example for output signal with noise: $M = 50$, $N = 100$, $p = 0.1$, $\sigma = 0.3$.

the earth, we can represent the top and base reflectors of a specific layer as two impulse functions $c\delta(t)$ and $d\delta(t + n\Delta t)$, where Δt represents the thickness of the layer, and c and d are two reflection coefficients. The dipole decomposition decomposes each reflector pair into one even pair r_e and one odd pair r_o with corresponding coefficients a and b in the following way:

$$r_e = \delta(t) + \delta(t + n\Delta t) \tag{2.14}$$

$$r_0 = \delta(t) - \delta(t + n\Delta t) \tag{2.15}$$

$$c\delta(t) + d\delta(t + n\Delta t) = ar_e + br_o. \tag{2.16}$$

Usually the thickness is unknown, so n varies from zero to the maximal assumed multiplication of Δt.

To construct the most general reflectivity signal using this model we assume:

$$r_e(t, m, n, \Delta t) = \delta(t - m\Delta t) + \delta(t - m\Delta t + n\Delta t) \tag{2.17}$$

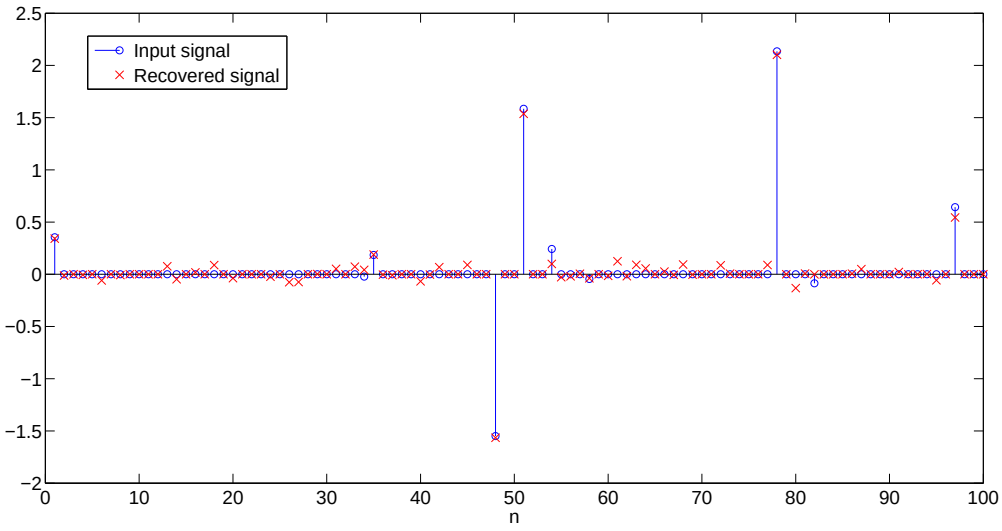

Figure 2.5: SSI. Example for Recovered signal with channel noise: $M = 50$, $N = 100$, $p = 0.1$, $\sigma = 0.3$.

$$r_o(t, m, n, \Delta t) = \delta(t - m\Delta t) - \delta(t - m\Delta t + n\Delta t) \tag{2.18}$$

where we use Δt as our sampling rate, so $m\Delta t$ represents the location of a specific layer and m ranges from 1 to the number of samples in the seismic trace. So we get the general reflectivity signal model:

$$r(t) = \sum_{n=1}^{N} \sum_{m=1}^{M} \left(a_{n,m} r_e(t, m, n, \Delta t) + b_{n,m} r_o(t, m, n, \Delta t)\right) \tag{2.19}$$

and the seismic signal will be:

$$s(t) = \sum_{n=1}^{N} \sum_{m=1}^{M} w(t) * \left(a_{n,m} r_e(t, m, n, \Delta t) + b_{n,m} r_o(t, m, n, \Delta t)\right). \tag{2.20}$$

We can see that basically we are staying with the same model from the point of view of the optimization problem but we are utilizing a slight different dictionary. Specifically, the new dictionary will be just a horizontal duplication of the old dictionary.

Now we continue to show some results using BPI. We use the same data used in the previous section to test SSI. The first example regarding Figure 2.1 with same A that was used in the SSI example, we can see the signal that was recovered by BPI in Figure 2.8.

The correlation measure in this case is also 1 as in SSI.

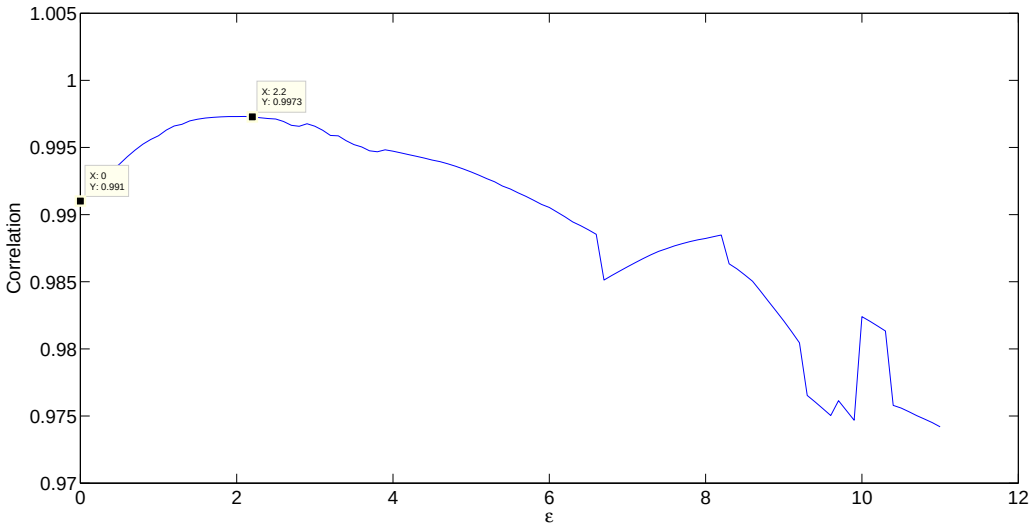

Figure 2.6: SSI. Example for correlation measure vs. ϵ.

Now we also move to test the case where there is noise in the channel. Again, we test the same signal as in the SSI section. In Figure 2.9 we can see the recovered signal from Figure 2.4 using BPI and choosing $\epsilon = 0$. The correlation measure in this case is 0.987. As done in the SSI case, we also try to find the value of ϵ that will maximize the correlation measure. It is important to say that choosing ϵ in SSI and BPI may have a lot in common but they do not have to be equal to each other, i.e., the value of ϵ that gives maximum correlation in SSI is not necessarily the value of ϵ that gives maximum correlation in BPI, so every one of them has to be chosen apart. In Figure 2.10 we can see the correlation measure vs. ϵ using BPI.

We can see there is no much difference from the SSI case. In our research we could not implement BPI to outperform SSI although Zhang and Castanga show in [16] how BPI has advantages over SSI. In the previous section where we presented SSI, although not tested for all cases we can see that SSI, when ϵ chosen correctly, can achieve very good results in recovering the input signal. This may be due to the fact that BPI is offered as a method to improve results where the top and bottom reflectors of an underground layer can be considered as different layers if not modeled correctly. BPI might have no advantages over SSI in cases where the abovementioned distinguish is not made.

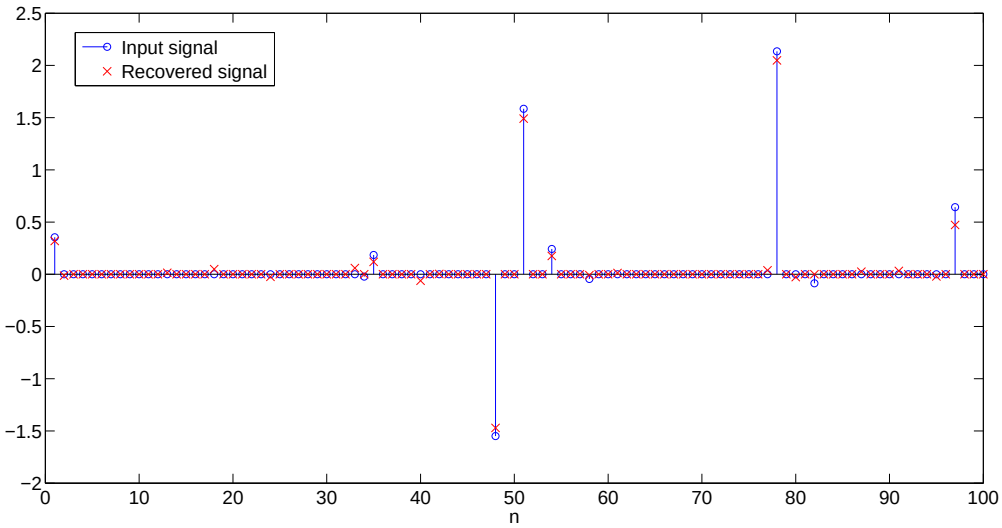

Figure 2.7: SSI. Recovered signal achieving maximum correlation.

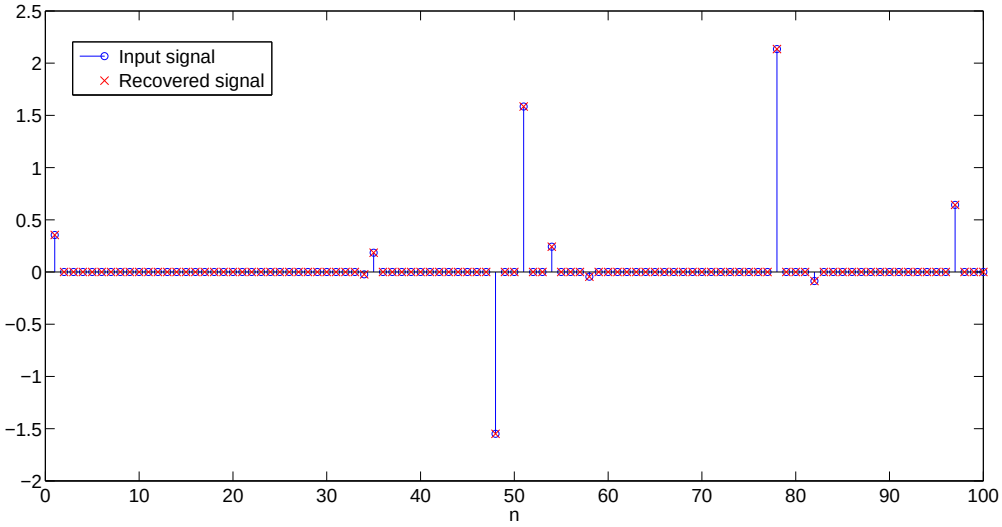

Figure 2.8: BPI. Example for recovered signal: $M = 50$, $N = 100$, $p = 0.1$.

2.4 Multichannel Blind Deconvolution

In Section 2.2 we formulated the non-blind deconvolution problem. If we go back to (2.3) we can define the goal of the blind deconvolution problem to recover $\mathbf{r}_i$ from $\mathbf{s}_i$ while assuming full knowledge only on σ_v. The immediate conclusion is that in the blind deconvolution case we have to be in the multichannel setup in order to have the ability to solve the problem. In this section we go into details of one of the most

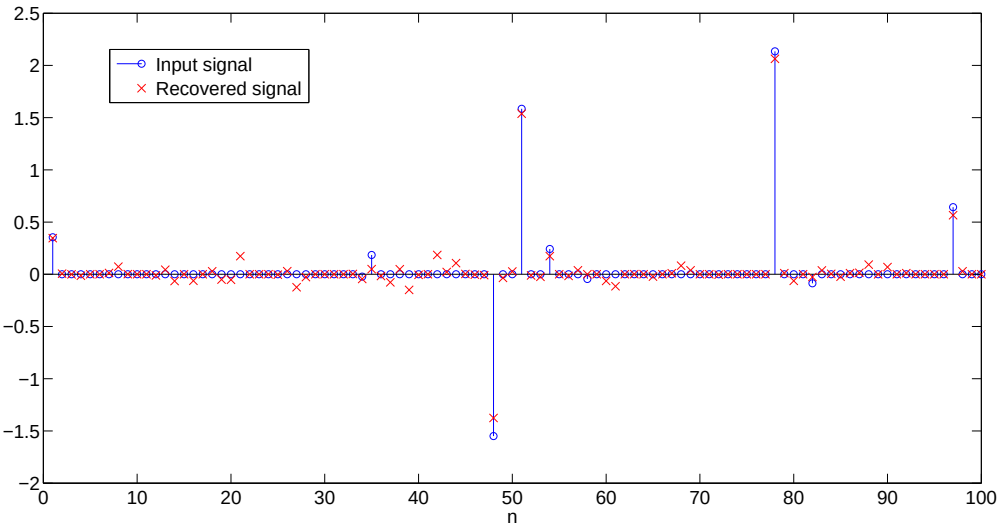

Figure 2.9: BPI. Example for Recovered signal with channel noise: $M = 50$, $N = 100$, $p = 0.1$, $\sigma = 0.3$.

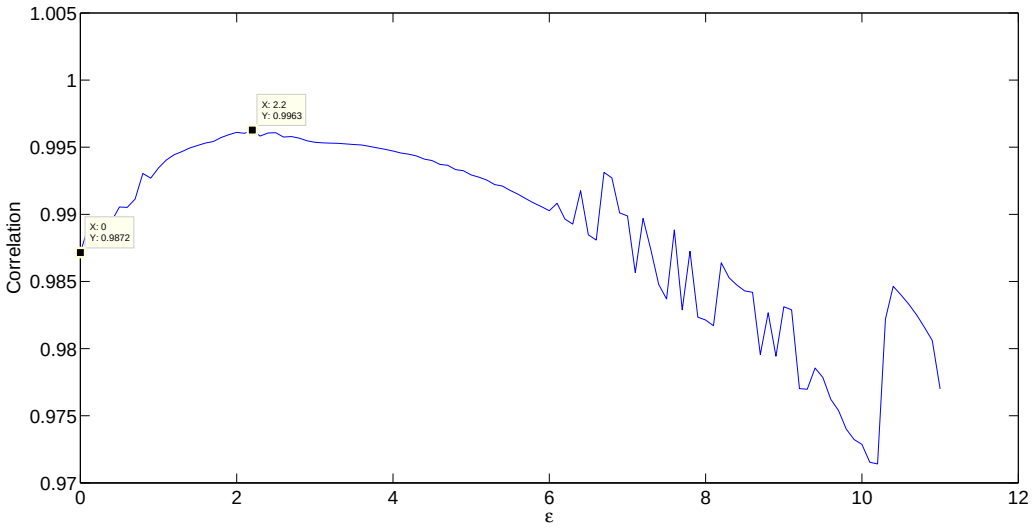

Figure 2.10: BPI. Example for correlation measure vs. ϵ.

common multichannel blind deconvolution methods for seismic signals, SMBD [9], as it is compared to our method later on.

To be consistent with the authors we formulate the problem and solution as presented in [9]. We hold data from total J channels,

$$d_j[n] = \sum_k w[n-k] r_j[k], \quad j = 1, ...J \tag{2.21}$$

where the multichannel seismic data is given by $\mathbf{d_j} = (d_j[1], d_j[2], ..., d_j[N])^T$, the reflectivity signal (or the impulse response signal) is given by $\mathbf{r_j} = (r_j[1], r_j[2], ..., r_j[M])^T$ and the seismic source (the wavelet) is given by $\mathbf{w} = (w[1], w[2], ..., w[L])^T$ and is common to all channels. We hold N = M + L -1. The above convolution can be represented in the z-domain as follows,

$$D_j(z) = W(z)R_j(z), \quad j = 1, ...J. \tag{2.22}$$

Let us look at two specific channels, p and q,

$$\begin{cases} D_p(z) = W(z)R_p(z) \\ D_q(z) = W(z)R_q(z) \end{cases}, \tag{2.23}$$

we multiply the first equation by $R_q(z)$ and the second equation by $R_p(z)$,

$$\begin{cases} D_p(z)R_q(z) = W(z)R_p(z)R_q(z) \\ D_q(z)R_p(z) = W(z)R_q(z)R_p(z) \end{cases}. \tag{2.24}$$

Now we subtract both equations and get,

$$D_p(z)R_q(z) - D_q(z)_p(z) = 0, \quad \forall p, q. \tag{2.25}$$

This can be written in the following vector-matrix form,

$$\mathbf{D}_p\mathbf{r}_q - \mathbf{D}_q\mathbf{r}_p = 0 \tag{2.26}$$

where $\mathbf{D}_p$ and $\mathbf{D}_q$ represent the convolution matrices of channels p and q, respectively. If we combine all the equations for every p and q, we get the following homogeneous equations set,

$$\mathbf{Ax} = \mathbf{0} \tag{2.27}$$

where $\mathbf{A}$ and $\mathbf{x}$ are defined in the following way,

$$\mathbf{A} = \begin{pmatrix} \mathbf{D_2} & -\mathbf{D_1} & 0 & \cdots & \cdots & \cdots & 0 \\ \mathbf{D_3} & 0 & -\mathbf{D_1} & 0 & \cdots & \cdots & 0 \\ \mathbf{D_4} & 0 & 0 & -\mathbf{D_1} & 0 & \cdots & 0 \\ \vdots & 0 & \cdots & 0 & \ddots & \cdots & 0 \\ 0 & \mathbf{D_3} & -\mathbf{D_2} & 0 & \cdots & \cdots & 0 \\ 0 & \mathbf{D_4} & 0 & -\mathbf{D_2} & 0 & \cdots & 0 \\ 0 & \vdots & 0 & 0 & \ddots & \cdots & 0 \\ 0 & \cdots & \cdots & 0 & \mathbf{D_j} & 0 & \mathbf{D_{j-2}} \\ 0 & \cdots & \cdots & \cdots & 0 & \mathbf{D_j} & \mathbf{D_{j-1}} \end{pmatrix} \tag{2.28}$$

and,

$$\mathbf{x} = [\mathbf{r_1}, \mathbf{r_2}, \cdots, \mathbf{r_J}]^T. \tag{2.29}$$

If we consider noise in the channel we can rewrite the above equations as,

$$\mathbf{Ax} = \mathbf{e} \tag{2.30}$$

where in [9] the authors consider $\mathbf{e}$ as a white Gaussian noise just to be able to develop an algorithm. Therefore, they propose to recover $\mathbf{x}$ with minimizing the $l2$ norm of the error $\mathbf{e}$ with a requirement that $\mathbf{x}$ is sparse. In addition, to avoid the trivial solution they add the constraint $\mathbf{x}^T\mathbf{x} = 1$. Finally, they propose to find $\mathbf{x}$ in the following way,

$$\hat{\mathbf{x}} = \underset{\mathbf{x}}{argmin}\ \mathbf{J}(\mathbf{x}), \quad s.t. \quad \mathbf{x}^T\mathbf{x} = 1 \tag{2.31}$$

where

$$\mathbf{J}(\mathbf{x}) = \frac{1}{2}\,||\mathbf{Ax}||_2^2 + \lambda\mathcal{R}_\epsilon(\mathbf{x}) \tag{2.32}$$

and

$$\mathcal{R}_\epsilon(\mathbf{x}) = \sum_i \left(\sqrt{x_i^2 + \epsilon^2} - \epsilon\right). \tag{2.33}$$

$\mathcal{R}_\epsilon(\mathbf{x})$ promotes sparse solutions and $||\mathbf{Ax}||_2^2$ maintains data fitting. λ is the trade-off parameter controlling the balance between the mentioned properties.

We do not go into details of presenting results of this method because it is not used in our developed algorithm. For experimental results of SMBD one can refer to [9].

2.5 Conclusions

In this chapter we went over the state-of-the-art algorithms and methods for seismic deconvolution, blind and non-blind, we briefly described different disciplines in seismic deconvolution and different considerations.
We formulated the problem mathematically and presented in details specific interesting algorithms for the blind and non-blind deconvolution problems which serves us for our algorithm and for comparison. In the next chapter we discuss our proposed method for semi-blind deconvolution, MSBD.

Chapter 3

Multichannel Semi-blind Deconvolution - MSBD

In this chapter we discuss MSBD, which is our proposed method for the semi-blind deconvolution problem. In Section 3.1 we mathematically formulate the semi-blind problem and discuss the differences from the non-blind deconvolution problem. In Section 3.2 we present the general method of MSBD, explain the principles behind it and discuss the iterated two steps method.

3.1 Problem Formulation

The semi-blind deconvolution is a very general problem where we hold some information on the wavelet. We mathematically presented the non-blind problem in Section 2.2 and in this section we add the mathematical formulation of the semi-blind problem.

In the non-blind deconvolution problem, $\mathbf{A}$ from (2.5) is fully known and in the blind problem $\mathbf{A}$ is unknown. In our case, we do not have full knowledge of $\mathbf{A}$. We have $\mathbf{A}'$, which is a noisy version of $\mathbf{A}$ and holds the relation $\mathbf{A} = \mathbf{A}' - \mathbf{A}_n$. $\mathbf{A}_n$ represents the uncertainty in $\mathbf{A}$ and in later chapters we address the problem with a specific definition of $\mathbf{A}_n$.

3.2 General Method

Our purpose is to establish a general method for the Semi-Blind Deconvolution problem and to specifically analyze two different cases of wavelet uncertainty as shown later. First, we introduce the general method. The method relies on the different modeling of each of the recovered signals. We know nothing about the reflectivity signal besides the fact that it is sparse, however we assume to know the wavelet up to some level of noise. The noise can be additive to the wavelet signal or intrinsic to one of the parameters that form the wavelet model, or any other noise that can be mathematically formulated. We assume to know the statistical parameters defining the noise. In addition, We assume non-sparse representation of the wavelet signal. This means that we use the standard deconvolution methods for the wavelet. In our case, we choose to work with the L2 minimization method. This method was adapted to best fit our problem. The adaption that was made and the full form of the deconvolution are shown later. The method we propose is an iterative method, with two steps in each iteration, as follows:

1. Assume to know the wavelet and use the sparse deconvolution method to recover the reflectivity signal.

2. Assume to know the reflectivity and use the L2 minimization method to recover the wavelet.

3.2.1 Reflectivity Recovery

In the first step we choose to work with the BPDN method for recovering sparse signals. As mentioned before regarding this method, the most important thing is to choose the trade-off parameter wisely. Our method relies on the fact that the uncertainty in the wavelet is represented as an additive noise to the true wavelet. Later we show that if this is not the case, then we can approximate the non-additive noise with an additive estimation.

We now introduce the general method of choosing the trade-off parameter when assuming additive noise to the wavelet and later on we demonstrate this method for two specific cases. Denote by $\mathbf{w} \in \mathbb{R}^{N_w}$ the vector representation of the signal $w[n], 1 \leqslant n \leqslant N_w$. The

additive noise to the wavelet is denoted by $\mathbf{w}_n \in \mathbb{R}^{N_w}$ and the corresponding convolution matrix is $\mathbf{W}_n \in \mathbb{R}^{(N_r+N_w-1)\times N_r}$. In the same way we denote the initial wavelet we are given and its corresponding convolution matrix as $\mathbf{w'} \in \mathbb{R}^{N_w}$ and $\mathbf{W'} \in \mathbb{R}^{(N_r+N_w-1)\times N_r}$ so we get the relation,

$$\mathbf{W} = \mathbf{W}' - \mathbf{W}_n. \tag{3.1}$$

Substituting (3.1) into (2.3) we get,

$$\mathbf{s}_i = \mathbf{W}\mathbf{r}_i + \mathbf{v}_i = (\mathbf{W}' - \mathbf{W}_n)\mathbf{r}_i + \mathbf{v}_i = \mathbf{W}'\mathbf{r}_i + \mathbf{v}_i - \mathbf{W}_n\mathbf{r}_i. \tag{3.2}$$

Looking at this relation,

$$\mathbf{s}_i = \mathbf{W}'\mathbf{r}_i + \mathbf{v}_i - \mathbf{W}_n\mathbf{r}_i \tag{3.3}$$

we can identify that $\mathbf{W'}$ is our "known" wavelet and it is treated as one, and $\mathbf{v}_i - \mathbf{W}_n\mathbf{r}_i$ is the term that represents the noise, or uncertainty, in the problem. For a wise choice of the trade-off parameter, variance analysis must be performed for that term. A major issue we have identified is that in each iteration the variance of the uncertainty term can be changed and a wise adaptation to that trade-off parameter needs to be made.
We denote the new noise term as,

$$\mathbf{v'}_i = \mathbf{v}_i - \mathbf{W}_n\mathbf{r}_i. \tag{3.4}$$

For the first step we assume to know the wavelet and recover the reflectivity series. As mentioned before, this is done by applying the BPDN solver to our problem. As we see in (2.5), we have to choose the trade-off parameter, ϵ, wisely. The literature does not prove, nor imply, a generic analysis for choosing this parameter, but as mentioned in previous sections, ϵ has a strong relation to the total standard deviation of the noise in the problem. In our case we decided to define the total standard deviation of the noise in two forms. Each form has its advantages and disadvantages.
The first form is as follows:

$$\bar{v}'_i = \sum_{n=1}^{N_r+N_w-1} v'_i[n] \tag{3.5}$$

and now ϵ is the standard deviation of $\bar{v}'_i$:

$$\epsilon_i = \sigma_{\bar{v}'_i} = \sqrt{E(\bar{v}'_i - E\bar{v}'_i)^2} \tag{3.6}$$

where E is the well known expectation operator.

The second form is:

$$\epsilon_i = \sqrt{\sum_{j=1}^{N_r} \sigma^2_{\mathbf{v}_j}}. \tag{3.7}$$

The main difference is that in the first form we look at different elements of the noise vector with common elements constructing them, hence there is a strong correlation between the sources of noise from different elements. In the second form we look at the total noise as the sum of independent noise sources in each element.

The selection of a specific form comes more from intuition and empirical processes and less from analytical proof that the chosen form is the only correct one. Different forms can be suggested, the important thing is to maintain a logical connection to the model of the problem and to take into account and quantify all the noise sources in the problem.

3.2.2 Wavelet Recovery

For the second step, we assume to know the reflectivity series and aim to recover the wavelet. Unlike the first step, here we cannot apply BPDN, or any other sparse deconvolution method for that matter. The simple reason is that the wavelet is not a sparse signal. We look at this problem from another point of view, dictionary learning. Dictionary learning is a broad field that can provide many insights on how to update the wavelet. Several ideas were tested according to [20] [31] [32] [33].

The seismic data can be considered a linear combination of the columns of $\mathbf{W}$, the dictionary, where the reflectivity series can be treated as the coefficients. This makes sense because the columns of $\mathbf{W}$ are shifted versions of $\mathbf{w}$. With that in mind, finding the wavelet when the reflectivity series is known can be treated by methods from the field of dictionary learning, as the purpose of this stage is to update and learn $\mathbf{W}$ (defined directly by $\mathbf{w}$). We use a method of dictionary update based on the Signature Dictionary as described in [20].

Specifically, we would like to minimize the l_2 expression $\sum_{i=1}^{N} \|\mathbf{s}_i - \mathbf{W}\mathbf{r}_i\|_2^2$, where $\{\mathbf{s}_i\}_{i=1}^N, \{\mathbf{r}_i\}_{i=1}^N$ are known and $\mathbf{W}$ has the special form of $\mathbf{w}_k = \mathbf{R}_k\mathbf{w}$, where $\mathbf{w}_k$ is the k-th

column of $\mathbf{W}$ and $\mathbf{R}_k$ is a matrix that fits $\mathbf{w}$ into a zero-vector from its k-th element, i.e., $\mathbf{R}_k$ looks like this:

$$\mathbf{R}_k = \begin{pmatrix} \mathbf{0}_{k-1\times N_w} \\ \mathbf{I}_{N_w} \\ \mathbf{0}_{N_r-k\times N_w} \end{pmatrix} \tag{3.8}$$

where $\mathbf{I}_{N_w}$ is the identity matrix os size N_w and $\mathbf{0}_{M\times N}$ is an $M \times N$ matrix of zeros.
This minimization problem was solved in [20] to obtain the optimal $\mathbf{w}$, although solved for different $\mathbf{R}_k$ matrices. We can represent the seismic data, $\mathbf{s}_i$, as a linear combination of the columns of $\mathbf{W}$:

$$\mathbf{s}_i = \mathbf{W}\mathbf{r}_i = \sum_{k=1}^{m} \mathbf{w}_k r_i[k] = \sum_{k=1}^{m} \mathbf{R}_k \mathbf{w} r_i[k]. \tag{3.9}$$

We recall that our purpose is to minimize $\sum_{i=1}^{N} \|\mathbf{s}_i - \mathbf{W}\mathbf{r}_i\|_2^2$. Substituting (3.9) into this we get the following minimization problem:

$$\mathbf{w}^{\text{opt}} = \underset{\mathbf{w}}{\text{argmin}} \quad \sum_{i=1}^{N} \left\| \mathbf{s}_i - \sum_{k=1}^{m} r_i[k]\mathbf{R}_k\mathbf{w} \right\|_2^2. \tag{3.10}$$

Thus, we obtain an update to $\mathbf{w}$ by optimizing the above l_2 term with respect to $\mathbf{w}$. We obtain:

$$\sum_{i=1}^{N} \left(\sum_{k=1}^{m} r_i[k]\mathbf{R}_k \right)^T \left(\mathbf{s}_i - \sum_{k=1}^{m} r_i[k]\mathbf{R}_k\mathbf{w} \right) = 0. \tag{3.11}$$

Solving this equation we get:

$$\mathbf{w}^{\text{opt}} = \left(\sum_{k=1}^{N_r}\sum_{j=1}^{N_r} \left[\sum_{i=1}^{N} \mathbf{r}_i[k]\mathbf{r}_i[j] \right] \mathbf{R}_k^T\mathbf{R}_j \right)^{-1} \sum_{i=1}^{N}\sum_{k=1}^{N_r} \mathbf{r}_i[k]\mathbf{R}_k^T\mathbf{s}_i \tag{3.12}$$

and now we can update the matrix $\mathbf{W}$.
This step is common to all different kinds of uncertainties in the wavelet since the true wavelet model and its connection to the seismic data are not affected by the wavelet uncertainty.

3.3 Conclusions

In this chapter we presented MSBD as a general method for semi-blind deconvolution problems with two steps where one of them is fixed from one use case to another. In the next chapter we go with further mathematical development to test MSBD with two different use cases.

Now we continue on to analyze two different cases of wavelet uncertainty. The first is AWGN contamination of the wavelet and the second is a parametric change in the wavelet model.

Chapter 4

Specific Use Cases

In this chapter we discuss two specific use cases for the general method we presented in the previous chapter. We intentionally choose two different use cases which differ in a few manners.

The first use case, wavelet AWGN, discusses a case where the wavelet is contaminated by some AWGN which we know its STD. This case is presented mainly to intuitively go over the different steps of MSBD and discuss some of its properties. As mentioned in previous chapters, our model deals with additive noise and so in this case we start with that as given in order to focus on the process of MSBD itself. In the second case, wavelet parametric change, we present a more realistic problem that can occur in reflection seismology. We talk about a case where we have a mathematical model of the wavelet but the parameters defining it are known with some kind of uncertainty factor. Although this is not straightforward modeled as an additive noise, we present a simple process to approximately fit it into an additive noise model. Then we discuss some results and affects of MSBD regarding this case.

4.1 Wavelet AWGN

4.1.1 MSBD solver Development

The first case of wavelet uncertainty is where the wavelet is contaminated with AWGN. The model we are assuming is as follows:

$$\mathbf{w}' = \mathbf{w} + \mathbf{w}_n \tag{4.1}$$

where the elements of $\mathbf{w}_n$ are IID and normally distributed with a known variance, i.e.,

$$\text{IID} \quad \mathbf{w}_n[k] \sim \mathcal{N}(0, \sigma_w^2), \quad 1 \leqslant k \leqslant N_w. \tag{4.2}$$

In addition we assume that $\{\mathbf{w}_n[k]\}_{k=1}^{N_w}$ and $\{\{\mathbf{v}_i[k]\}_{k=1}^{N_w+N_r-1}\}_{i=1}^{N}$ are statistically independent.

We recall that our purpose is to choose ϵ_i according to (3.5) and (3.6). In this case we have the exact form as in (3.1) so no further adaptations need to be made to fit the proposed model and method.
First let us examine the general element in $\mathbf{v}'_i$ as defined in (3.4). To do this we recall that

$$\mathbf{w}_n = [w_n[1], w_n[2], ..., w_n[N_w]]^T \text{ and } \mathbf{W}_n = \begin{pmatrix} w_n[1] & 0 & \cdots & 0 \\ w_n[2] & w_n[1] & & 0 \\ \vdots & w_n[2] & \ddots & \vdots \\ w_n[N_w] & \vdots & & 0 \\ 0 & w_n[N_w] & & w_n[1] \\ 0 & 0 & \ddots & w_n[2] \\ \vdots & 0 & \vdots & \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & w_n[N_w] \end{pmatrix}.$$

Substituting this into (3.4) we get,

$$\mathbf{v'}_i = \mathbf{v}_i - \mathbf{W}_n\mathbf{r}_i =$$
$$= \begin{bmatrix} \mathbf{v}_i[1] \\ \mathbf{v}_i[2] \\ \vdots \\ \mathbf{v}_i[N_r+N_w-1] \end{bmatrix} - \begin{pmatrix} w_n[1] & 0 & \cdots & 0 \\ w_n[2] & w_n[1] & & 0 \\ \vdots & w_n[2] & \ddots & \vdots \\ w_n[N_w] & \vdots & & 0 \\ 0 & w_n[N_w] & & w_n[1] \\ 0 & 0 & \ddots & w_n[2] \\ \vdots & 0 & \vdots & \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & w_n[N_w] \end{pmatrix} \cdot \begin{bmatrix} \mathbf{r}_i[1] \\ \mathbf{r}_i[2] \\ \vdots \\ \mathbf{r}_i[N_r] \end{bmatrix}. \tag{4.3}$$

So the general element in $\mathbf{v'}_i$ can be written as:

$$\mathbf{v'}_i[k] = \mathbf{v}_i[k] - \sum_{j=1}^{k} w_{v;k-j+1}\mathbf{r}_i[j], \quad 1 \leqslant k \leqslant N_r + N_w - 1. \tag{4.4}$$

In this case, we chose to work with the first form of the trade-off parameter selection. It is pretty clear that different noise vector elements have a strong correlation in their noise sources. Substituting this into (3.5) we get,

$$\begin{aligned} \bar{v}'_i &= \sum_{l=1}^{N_r+N_w-1} \mathbf{v'}_i[l] = \sum_{l=1}^{N_r+N_w-1} \left(\mathbf{v}_i[l] - \sum_{j=1}^{l} w_{v;l-j+1}\mathbf{r}_i[j] \right) = \\ &= \sum_{l=1}^{N_r+N_w-1} \mathbf{v}_i[l] - \sum_{l=1}^{N_r+N_w-1} \sum_{j=1}^{l} w_{v;l-j+1}\mathbf{r}_i[j] = \\ &= \sum_{l=1}^{N_r+N_w-1} \mathbf{v}_i[l] - \sum_{j=1}^{N_w} \left(\sum_{k=1}^{N_r} \mathbf{r}_i[k] \right) w_{v;j} \end{aligned} \tag{4.5}$$

We can see that $\bar{v}'_i$ is a linear combination of independent normally distributed random variables, so we can directly obtain the variance and the standard deviation of $\bar{v}'_i$:

$$\begin{aligned} \sigma^2_{\bar{v}'_i} &= \sum_{l=1}^{N_r+N_w-1} \sigma^2_{\mathbf{v}_i[l]} + \sum_{j=1}^{N_w} \left(\sum_{k=1}^{N_r} \mathbf{r}_i[k] \right)^2 \sigma_w^2 = \\ &= (N_r + N_w - 1)\sigma_v^2 + N_w \left(\sum_{k=1}^{N_r} \mathbf{r}_i[k] \right)^2 \sigma_w^2 \end{aligned} \tag{4.6}$$

and now we can obtain ϵ_i from (3.6):

$$\begin{aligned}\epsilon_i &= \sigma_{\bar{v}_i'} = \\ &= \sqrt{(N_r + N_w - 1)\sigma_v^2 + N_w \left(\sum_{k=1}^{N_r} \mathbf{r}_i[k]\right)^2 \sigma_w^2}\end{aligned} . \tag{4.7}$$

Notice the dependence on σ_v and σ_w, which changes from iteration to iteration. We show an easy way to update σ_w in each iteration to best fit ϵ to the current iteration. Updating σ_v is not trivial and has no analytical solution to date, so in our system we assume σ_v stays constant from one iteration to the next. In addition, although mathematically updating σ_v could improve results, this step has no logical justification because σ_v represents the noise in the received seismograph and we do not update the received signal throughout the process. In contrast, σ_w is part of the given wavelet we hold and we change the assumed wavelet from iteration to iteration so there is a strong justification to evaluate σ_w in each iteration.

To update σ_w we analyze the current wavelet that was recovered from the last iteration and the initial wavelet that was given to us. The initial wavelet, $\mathbf{w_{init}}$ and the current wavelet, $\mathbf{w_{curr}}$, can be modeled as,

$$\begin{aligned}\mathbf{w_{init}} &= \mathbf{w} + \mathbf{w_n} \\ \mathbf{w_{curr}} &= \mathbf{w} + \mathbf{w}'_{\mathbf{n}}\end{aligned} \tag{4.8}$$

where we assume that $\mathbf{w}_n[k] \sim \mathcal{N}(0, \sigma_w^2), 1 \leqslant k \leqslant N_w$ and $\mathbf{w'}_n[k] \sim \mathcal{N}\left(0, \sigma_w'^2\right), 1 \leqslant k \leqslant N_w$. We know σ_w and aim to find σ_w'. We recall that we hold $\mathbf{w_{init}}$ and $\mathbf{w_{curr}}$ fixed, and analyze the following term,

$$\begin{aligned}&\|\mathbf{w_{init}} - \mathbf{w_{curr}}\|_2^2 = \|\mathbf{w} + \mathbf{w}_n - (\mathbf{w} + \mathbf{w'}_n)\|_2^2 = \\ &= N_w \left[\sigma_w^2 + (\sigma_w')^2\right]\end{aligned} . \tag{4.9}$$

It is easy to see that we can extract σ_w',

$$\sigma_w' = \sqrt{\frac{1}{N_w} \|\mathbf{w_{init}} - \mathbf{w_{curr}}\|_2^2 - \sigma_w^2}. \tag{4.10}$$

Now we can update σ_w at the beginning of each iteration.

4.1.2 Results and Discussion

In this part we present experimental results obtained from testing the performances of MSBD. We focus on synthetic data. The reflectivity sequences were created using the model presented in [34] with SNR varied in the range 0-20 dB. Fifty channels were used and the wavelet was created using the Ricker wavelet. The standard deviation of the wavelet AWGN, σ_w,was varied in the range 0.05 - 0.2. Twenty iterations were performed. Each combination of the above parameters was tested 1000 times. The presented correlation measure is the mean of those 1000 results.

We present the results in terms of the correlation between the recovered reflectivity and the original reflectivity, and compare this to the case where we assume a fixed wavelet. The fixed wavelet is tested with the SSI and SMBD algorithms that were presented in Section 2.1. The SSI representes the non-blind directive and the SMBD representes the blind directive in our comparison. All the results of those algorithms are presented, although when we dive into figures and detailed explanation of the results, we compare the MSBD only to SSI.

In the next figures we present some graphs and results related to an example of SNR = 10dB and $\sigma_w = 0.15$. First we examine in Figure 4.1 the true (original) reflectivity and the seismic data created from it using the Ricker wavelet as presented in Figure 4.2. The recovered reflectivity series from both (MSBD and SSI) methods are presented in Figure 4.3.
A visual inspection of the recovered reflectivity graphs indicates a major improvement with MSBD. Even without a quantitative measure we can see that the signal outlines are recovered nicely using MSBD. At certain points we can see that MSBD has created discontinuities, for example near channel number 25, at time 85 and near channel number 30 at time 160. These discontinuities are due to the fact that each channel is recovered irrespective of its neighbors; the neighbors are taken into account only at the stage of wavelet estimation. The correlation between recovered and true reflectivity for each channel is presented in Figure 4.4.

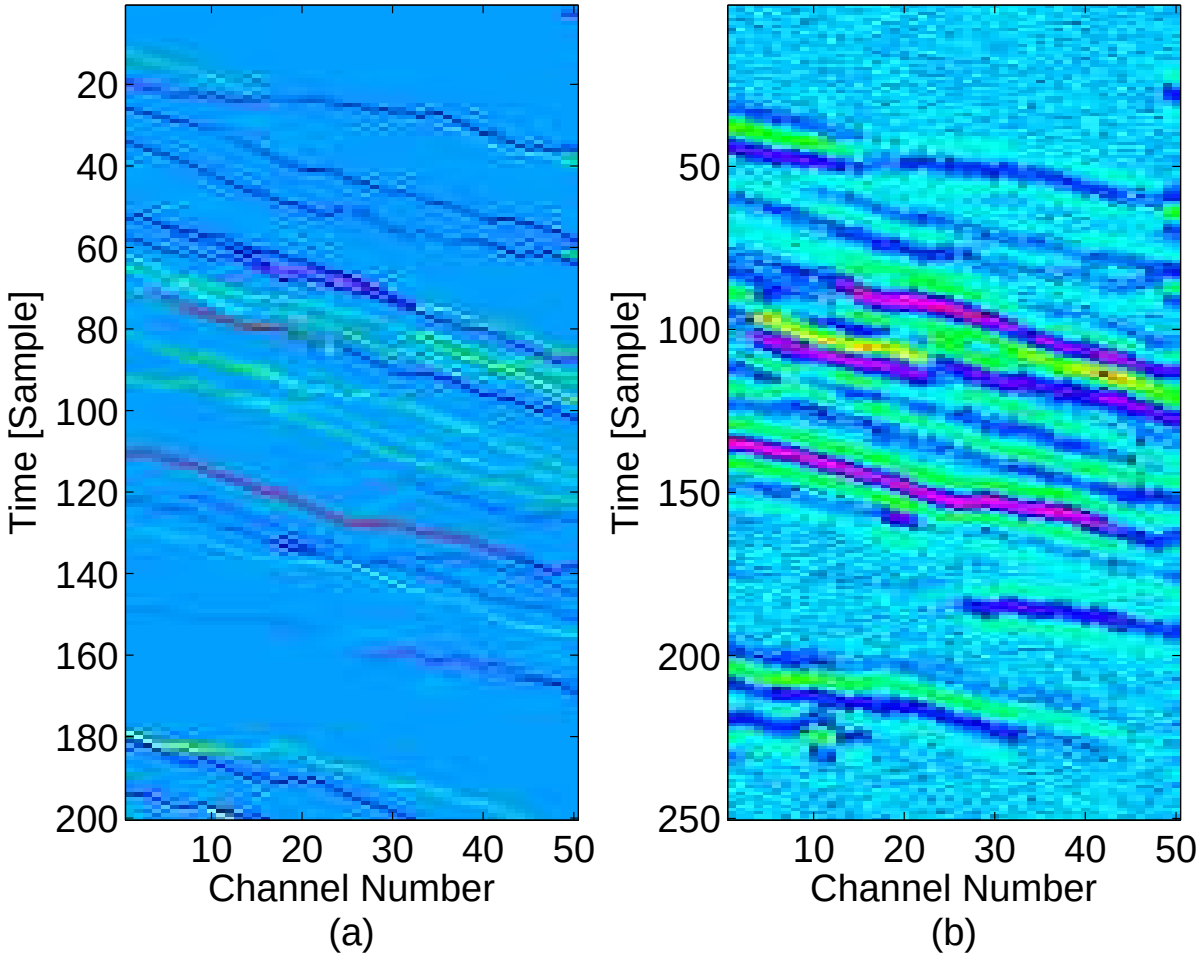

Figure 4.1: Wavelet AWGN, $SNR = 10dB$, $\sigma_w = 0.15$, True reflectivity and the seismic data: (a) true (original) reflectivity, (b) seismic data.

This provides a quantitative measure for testing the quality of the reflectivity recovery. It is clear that in most channels MSBD outperforms the SSI method.

Now we examine more "low-level" effects of the recovery by MSBD by focusing on recovery within a specific channel, in this case channel 15 (Figure 4.5).
When we examine the recovered and true reflectivity series (Figure 4.5(a)), the first effect we notice is the difficulty in recovering adjacent impulses, such as samples 88 to 90. This is because the wavelet is wide (in the time domain) so it is difficult to distinguish between two close impulses. If we look at sample 190 and 199 which are far from the previous and next impulses we note almost perfect recovery by MSBD. Now we examine the correlation between recovered and true reflectivity in this same channel as a function of iteration (Figure 4.5(b)). In general, the correlation tends to increase but there are

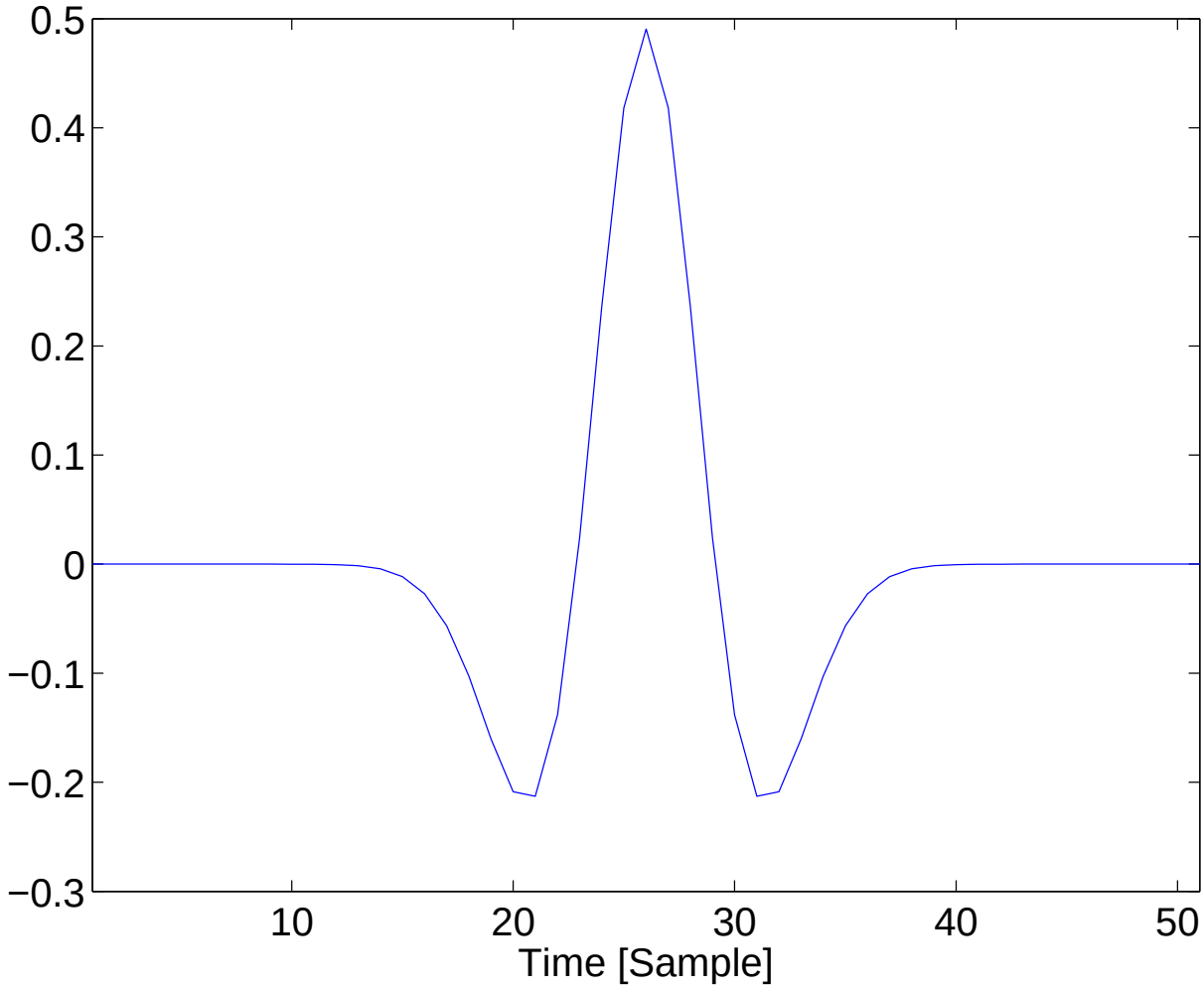

Figure 4.2: Wavelet AWGN: Ricker wavelet.

parts where it decreases. Also, we can clearly see that the final correlation is not the highest among all iterations. It is interesting to check an algorithm that finds the best moment to stop the iterations. Of course this kind of algorithm will have to take into account all the channels and not only a specific one.

Finally, we examine the estimation of the wavelet (Figure 4.6). We recall that MSBD is a two-stage algorithm, where the second stage is wavelet recovery. Here we can see the initial wavelet that was provided at the beginning, and the estimated wavelet at the last iteration, both compared to the true wavelet from which the seismic data were created. We can see that the estimation of the wavelet is almost perfect.

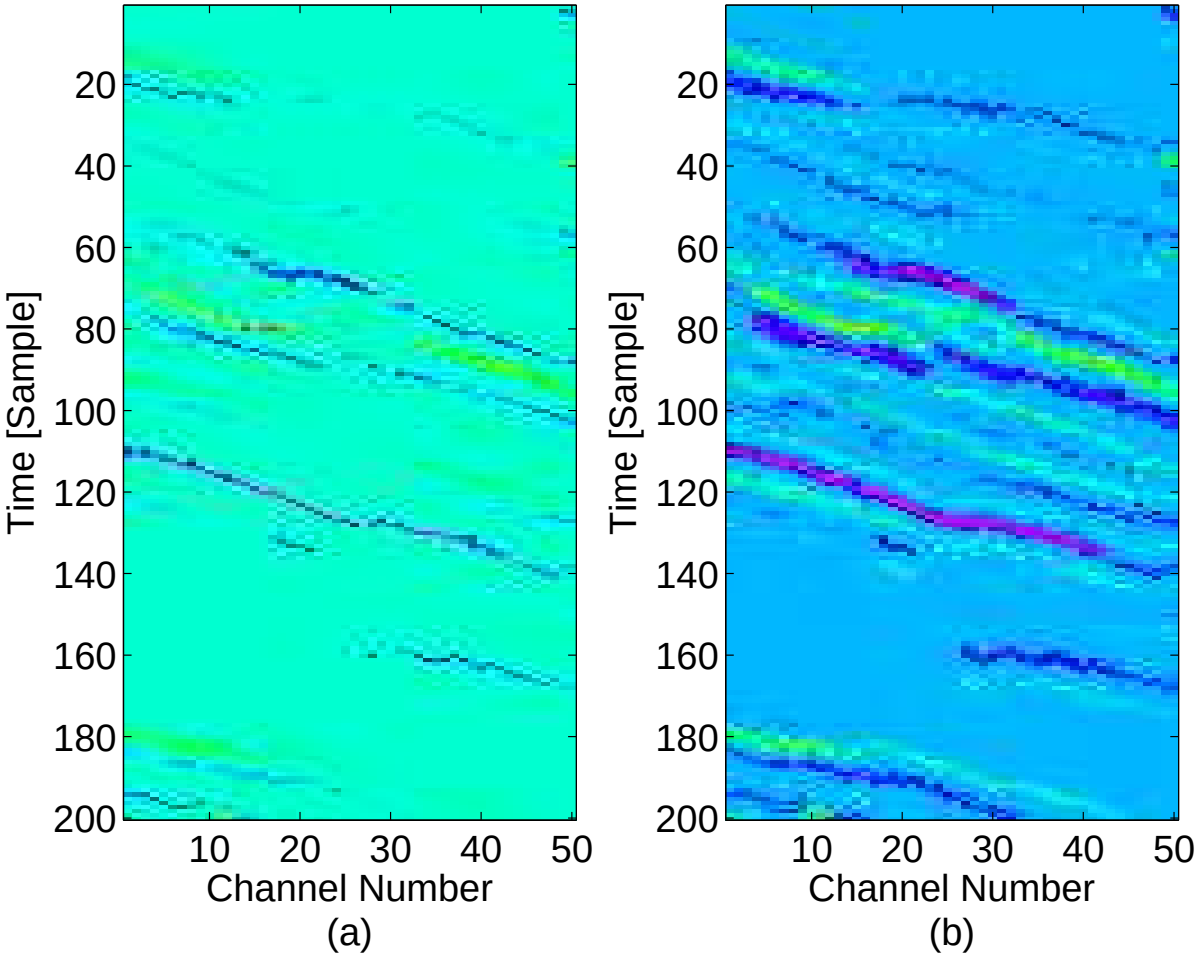

Figure 4.3: Wavelet AWGN, $SNR = 10dB$, $\sigma_w = 0.15$: (a) MSBD recovered reflectivity (b) SSI Recovered Reflectivity.

The mean correlations between the recovered and original reflectivity measured across all channels, for all tested cases, i.e. SNR varies from 0 to 20dB and σ_w varies from 0.05 to 0.2, are presented below in Table 4.1. Separate correlations for the SSI, SMBD and MSBD are presented for each value of SNR and σ_w.

Please notice that because SMBD is a method for the blind deconvolution problem it has nothing to do with the initial information on the wavelet so its performances are the same when looking on different values of σ_w.

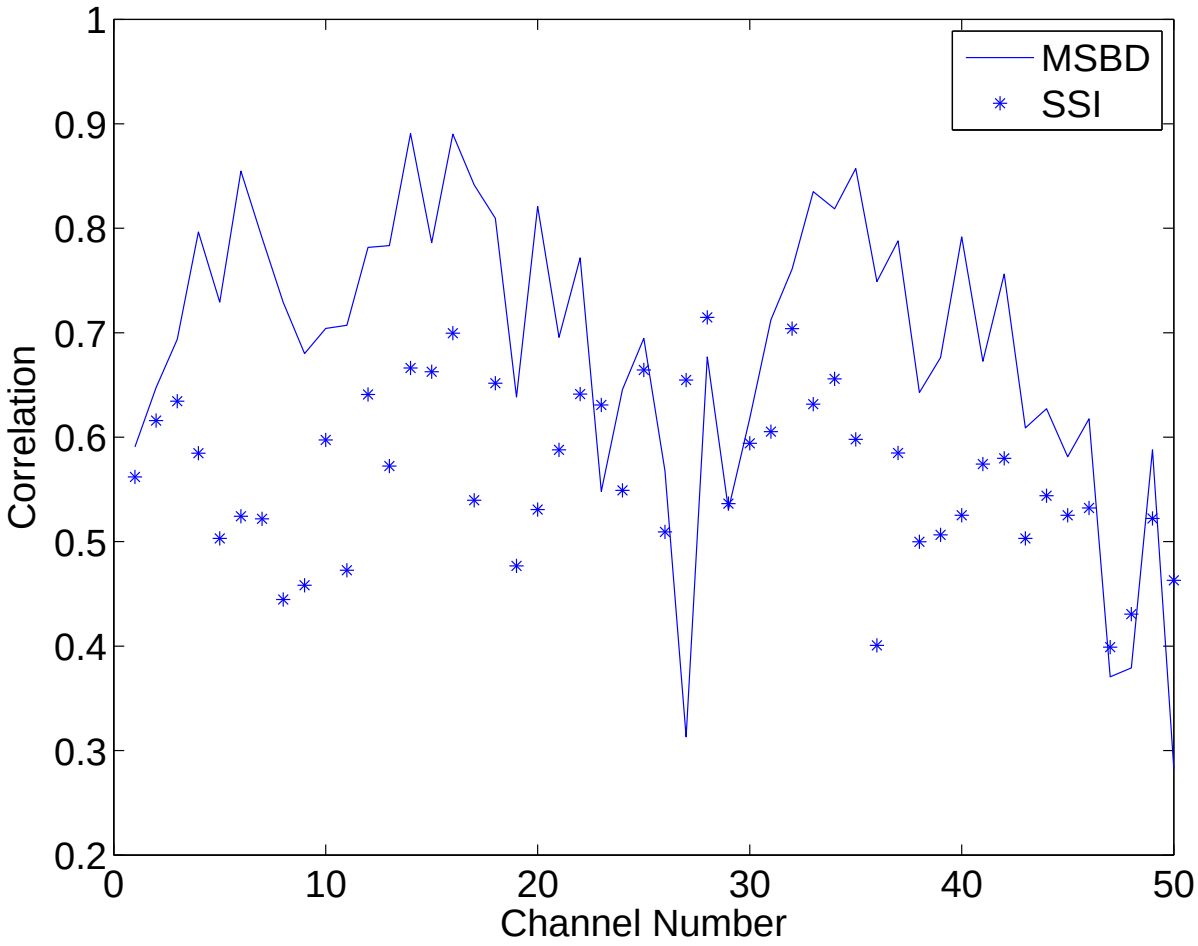

Figure 4.4: Wavelet AWGN, $SNR = 10dB$, $\sigma_w = 0.15$: Correlation measure vs. channel number.

Table 4.1: Correlations between recovered and original reflectivity for wavelet AWGN.

σ_w	SNR														
	0 dB			5 dB			10 dB			15 dB			20 dB		
	SSI	SMBD	MSBD	SSI	SMBD	MSBD	SSI	SMBD	MSBD	SSI	SMBD	MSBD	SSI	SMBD	MSBD
0.05	0.41	0.437	**0.46**	0.57	0.58	**0.67**	0.62	0.56	**0.69**	0.57	0.62	**0.82**	0.67	0.69	**0.85**
0.1	0.4		**0.45**	0.52		**0.78**	0.41		**0.52**	0.45		**0.77**	0.29		**0.88**
0.15	0.39		**0.52**	0.46		**0.59**	0.38		**0.81**	0.38		**0.77**	0.32		**0.77**
0.2	0.3		**0.48**	0.42		**0.57**	0.31		**0.65**	0.45		**0.75**	0.30		**0.72**

A few interesting points arise from these results. First, we see that the mean correlation of MSBD outperforms the mean correlation of SSI and SMBD methods. This demonstrates the great improvement potential MSBD has to offer for seismic deconvolution. Second, we would have expected to see a fixed trend of improvement in the correlation with increasing of SNR and decreasing of σ_w . For SNR this assumption held true, besides three exceptions: $\sigma_w = 0.1$ and $SNR = 5dB$, $\sigma_w = 0.15$ and $SNR = 10dB$,

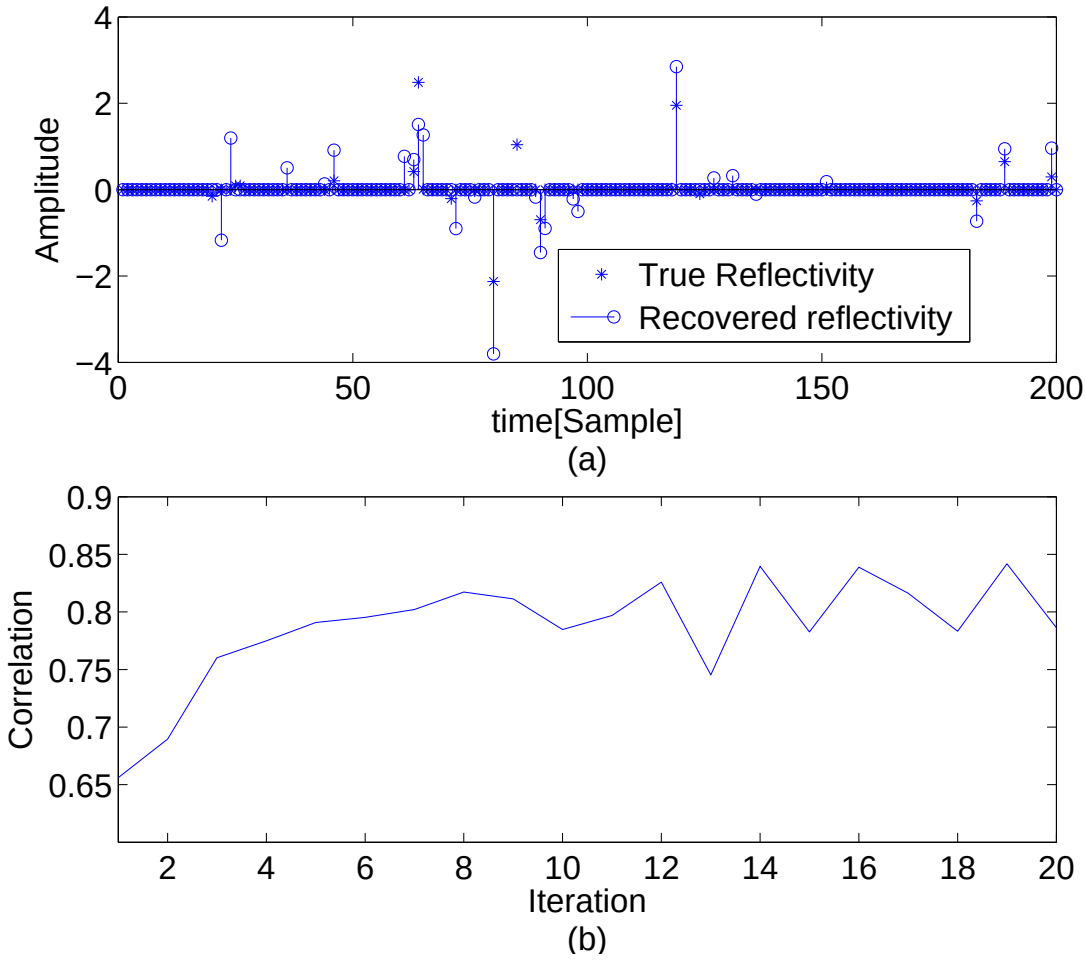

Figure 4.5: Wavelet AWGN, $SNR = 5dB$, $\sigma_w = 0.15$: (a) MSBD recovered reflectivity compared to the original reflectivity, within a specific channel, (b) The correlation between recovered and original reflectivity within a specific channel, as a function of iteration.

$\sigma_w = 0.2$ and $SNR = 15dB$, however for σ_w this assumption was invalid. This is due to the fact that recovery of the reflectivity is more related to the relative changes in adjacent samples of the wavelet than to the standard deviation of the noise in the signal.

4.2 Wavelet parametric change

4.2.1 MSBD solver Development

The second case of wavelet uncertainty is the one that involves a change in one of the parameters that define the wavelet. In this case we assume a certain model for the wavelet and analyze an uncertainty in one of its parameters.

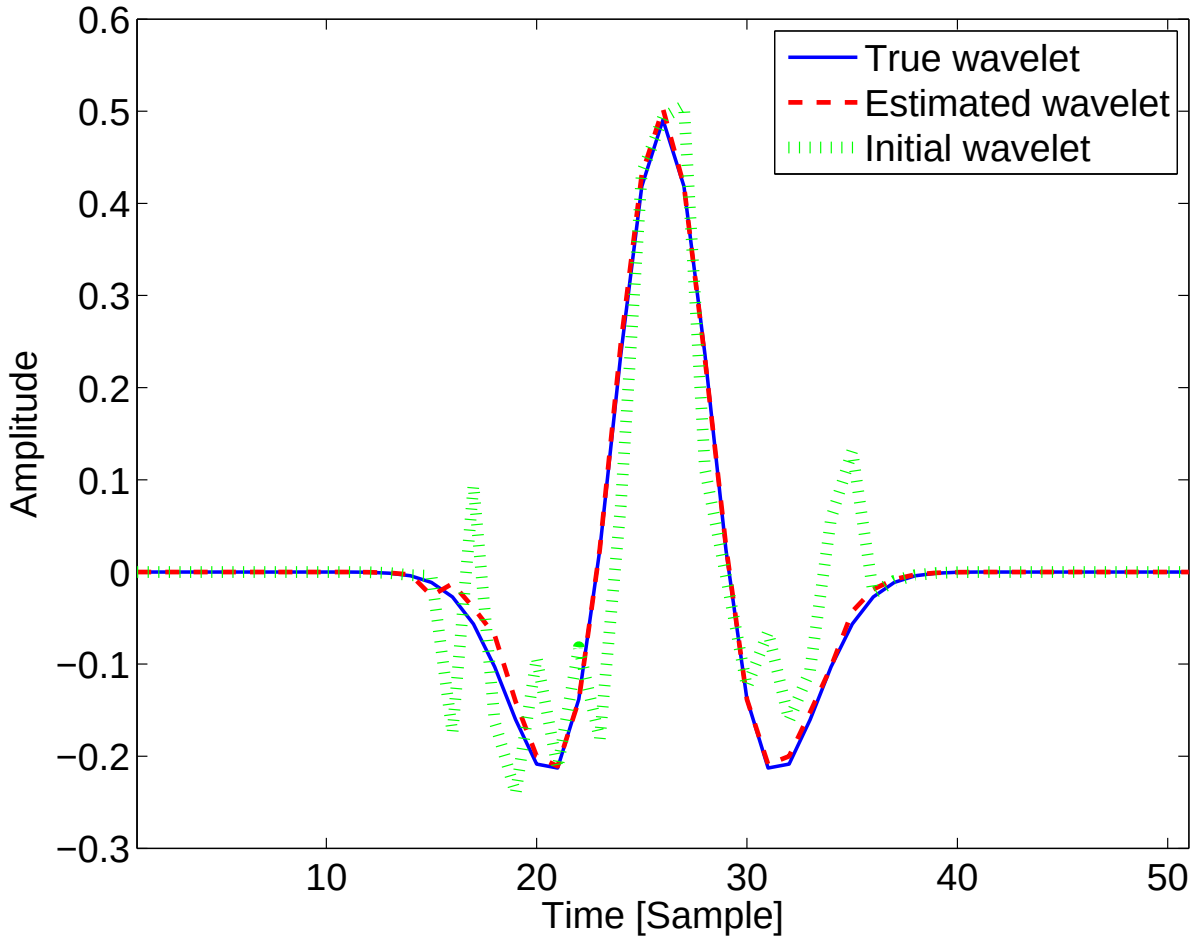

Figure 4.6: Wavelet AWGN: Wavelet estimation, $SNR = 10dB$, $\sigma_w = 0.15$.

A very common model for a seismic wavelet is the Ricker wavelet,

$$w(t; f) = \left(1 - 2\pi^2 f^2 t^2\right) e^{-\pi^2 f^2 t^2} \tag{4.11}$$

where f is a parameter that represents the frequency of the wavelet. In our case, we assume that the seismic data results from a wavelet defined by a frequency f_0, $w(t; f_0)$, but the wavelet we are initially given is defined by the parameter f, where $f \neq f_0$.

In order to simplify the later analysis of ϵ, we present the change in f as an additive term to the true wavelet. To do this, we need to represent $w(t; f)$ as a Taylor series. We

develop the series up to three elements. The first two derivatives are:

$$\begin{aligned}
&\tfrac{\partial w(t;f)}{\partial f} = -4\pi^2 f t^2 e^{-\pi^2 f^2 t^2} - 2\pi^2 f t^2 \left(1 - 2\pi^2 f^2 t^2\right) e^{-\pi^2 f^2 t^2} = \\
&= 2\pi^2 f t^2 \left(2\pi^2 f^2 t^2 - 3\right) e^{-\pi^2 f^2 t^2} = \left(4\pi^4 f^3 t^4 - 6\pi^2 f t^2\right) e^{-\pi^2 f^2 t^2} \\
&\tfrac{\partial^2 w(t;f)}{\partial f^2} = \\
&\left(12\pi^4 f^2 t^4 - 6\pi^2 t^2\right) e^{-\pi^2 f^2 t^2} - \\
&-2\pi^2 f t^2 \left(4\pi^4 f^3 t^4 - 6\pi^2 f t^2\right) e^{-\pi^2 f^2 t^2} = \\
&= \left(24\pi^4 f^2 t^4 - 8\pi^6 f^4 t^6 - 6\pi^2 t^2\right) e^{-\pi^2 f^2 t^2}
\end{aligned} \quad . \tag{4.12}$$

So we can approximately write that:

$$w\left(t;f\right) \approx w\left(t;f_0\right) + \left.\frac{\partial w\left(t;f\right)}{\partial f}\right|_{f=f_0} \left(f - f_0\right) + \frac{1}{2} \cdot \left.\frac{\partial^2 w\left(t;f\right)}{\partial f^2}\right|_{f=f_0} \left(f - f_0\right)^2 . \tag{4.13}$$

Putting this into a vector-matrix form in order to maintain the relation in (2.1) we get,

$$\mathbf{W} = \mathbf{W}_0 + \mathbf{DW}_1|_{f=f_0} \Delta f + \frac{1}{2} \mathbf{DW}_2|_{f=f_0} \Delta f^2 \tag{4.14}$$

where $\mathbf{W}_0, \mathbf{DW}_1$ and $\mathbf{DW}_2$ are the convolution matrices of $w\left(t;f_0\right)$, $\frac{\partial w(t;f)}{\partial f}$ and $\frac{\partial^2 w(t;f)}{\partial f^2}$, respectively, and $\Delta f = f - f_0$. Applying this notation to (3.1) we can see that $\mathbf{W}' = \mathbf{W}_0$ and $\mathbf{W}_n = -\left(\mathbf{DW}_1|_{f=f_0} \Delta f + \frac{1}{2} \mathbf{DW}_2|_{f=f_0} \Delta f^2\right)$. Following this we can write (3.4) as,

$$\mathbf{v'}_i = \mathbf{v}_i + \left(\mathbf{DW}_1|_{f=f_0} \Delta f + \frac{1}{2} \mathbf{DW}_2|_{f=f_0} \Delta f^2\right) \mathbf{r}_i . \tag{4.15}$$

Now we assume that $\Delta f \sim \mathcal{N}(0, \sigma_f^2)$ and analyze (3.4):

$\mathbf{v'}_i[k] = \mathbf{v}_i[k] + \left(\mathbf{DW}_1|_{f=f_0} \mathbf{r}_i\right)[k] \Delta f + \frac{1}{2} \left(\mathbf{DW}_2|_{f=f_0} \mathbf{r}_i\right)[k] \Delta f^2, \ 1 \leqslant k \leqslant N_r + N_w - 1.$

In this case, we choose to work with the second form of the trade-off parameter selection. Although it seems that different elements are also correlated with respect to the noise sources, the second form is more suitable here because of the independence of each element with the others.

We would like to use (3.7); to do so we must first examine the variance of $\mathbf{v'}_i[k]$:

$$\begin{aligned}
&E\left[\mathbf{v'}_i[k]\right] = \tfrac{1}{2} \left(\mathbf{DW}_2|_{f=f_0} \mathbf{r}_i\right)[k] \sigma_f^2 \\
&E\left[\left(\mathbf{v'}_i[k]\right)^2\right] = \sigma_v^2 + \left(\mathbf{DW}_1|_{f=f_0} \mathbf{r}_i\right)^2 [k] \sigma_f^2 + \\
&+ \tfrac{3}{4} \left(\mathbf{DW}_2|_{f=f_0} \mathbf{r}_i\right)^2 [k] \sigma_f^4 \\
&\sigma^2_{\mathbf{v'}_i[k]} = E\left[\left(\mathbf{v'}_i[k]\right)^2\right] - \left(E\left[\mathbf{v'}_i[k]\right]\right)^2 = \\
&= \sigma_v^2 + \left(\mathbf{DW}_1|_{f=f_0} \mathbf{r}_i\right)^2 [k] \sigma_f^2 + \tfrac{1}{2} \left(\mathbf{DW}_2|_{f=f_0} \mathbf{r}_i\right)^2 [k] \sigma_f^4
\end{aligned} \tag{4.16}$$

and we conclude:

$$\begin{aligned}\epsilon_i &= \sqrt{\sum_{n=1}^{N_r} \sigma^2_{\mathbf{v}_i}} = \\ &= \sqrt{\sum_{k=1}^{N_r} \left(\sigma_v^2 + \left(\mathbf{DW}_1|_{f=f_0} \mathbf{r}_i \right)^2 [k] \sigma_f^2 + \tfrac{1}{2} \left(\mathbf{DW}_2|_{f=f_0} \mathbf{r}_i \right)^2 [k] \sigma_f^4 \right)} = \\ &= \sqrt{N_r \sigma_v^2 + \sigma_f^2 \sum_{k=1}^{N_r} \left(\mathbf{DW}_1|_{f=f_0} \mathbf{r}_i \right)^2 [k] + \tfrac{1}{2} \sigma_f^4 \sum_{k=1}^{N_r} \left(\mathbf{DW}_2|_{f=f_0} \mathbf{r}_i \right)^2 [k]}\end{aligned} \quad . \tag{4.17}$$

We note here that the derivatives must be recalculated at each iteration.

4.2.2 Results and Discussion

In this part we present the results of the above discussed case. The reflectivity series was created in the same way as for the wavelet AWGN, according to the same element and with the same parameters. Likewise, the SNR of the seismic data varied from 0 to 20 dB, the wavelet was the Ricker Wavelet with $f_0 = 2$ and σ_f varied from 0.5 to 1.2. As in the former use case, each combination of the above parameters was tested 1000 times. The presented correlation measure is the mean of those 1000 results.

We illustrate the main results using an example with SNR = 10dB and $\sigma_f = 1$, again using 20 iterations. First we examine the true reflectivity series and the seismic data (Figure 4.7) created using the Ricker wavelet in Figure 4.8. The recovered reflectivity series after applying MSBD is presented in (Figure 4.9(a)) and the recovered reflectivity series using the SSI method is presented in (Figure 4.9(b)) where we assume a fix wavelet throughout the process.

Once again, we can visually see the superiority of MSBD and the improvement it offers, even without a quantitative measure. Using the SSI recovery, the outlines of the reflectivity signals are visible, but very unclear and contain many discontinuities. In contrast, the MSBD recovery series shows clear reflectivity outlines recovered with high correlation to the true reflectivity. Surprisingly, there are far fewer discontinuities here compared to the wavelet AWGN, even though neighboring channels were not taken into account when recovering a specific channel. Visual inspection of the correlation

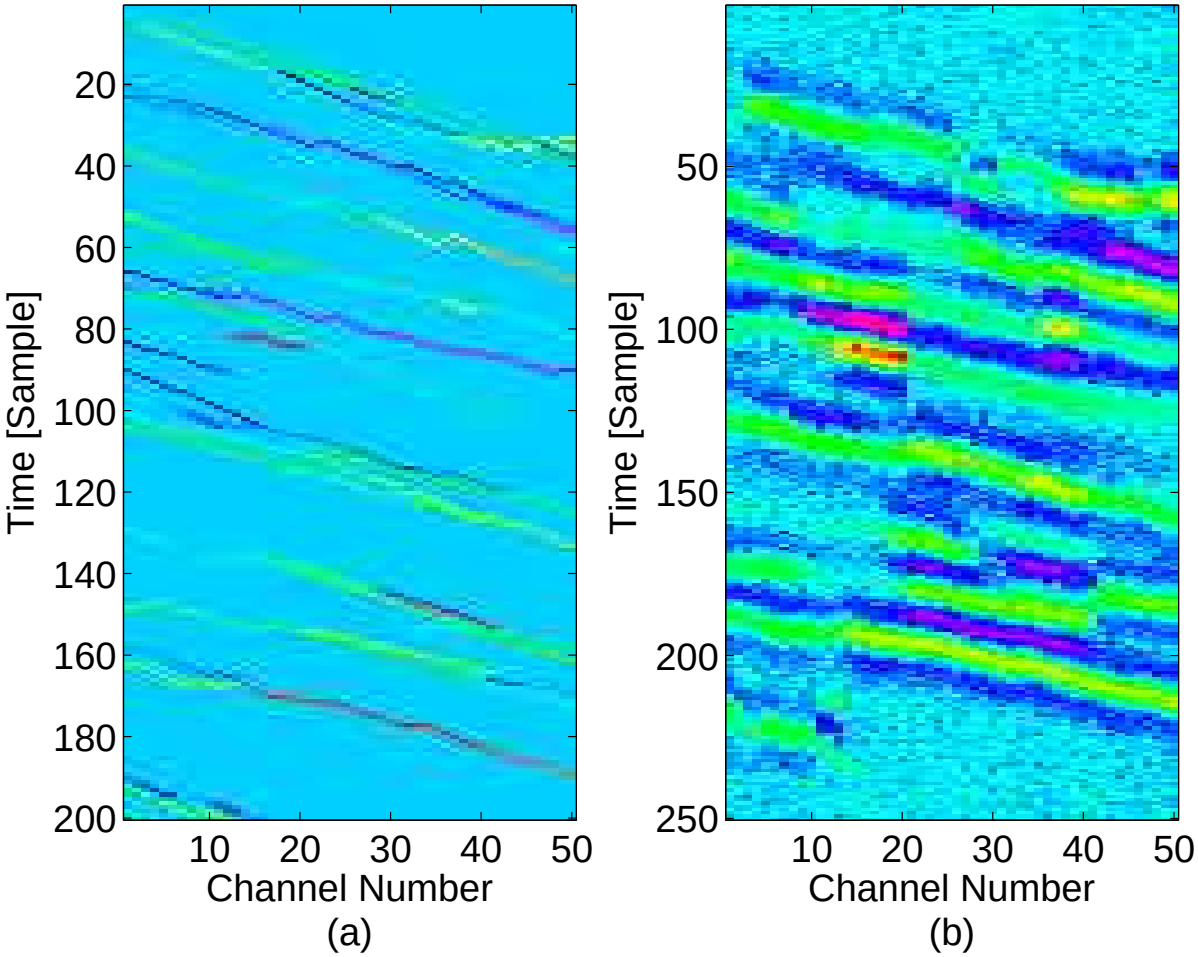

Figure 4.7: Wavelet parametric change, $SNR = 10dB$: True reflectivity and the seismic data: (a) true (original) reflectivity, (b) seismic data.

between recovered and true reflectivity for each channel (Figure 4.10) indicates that MSBD clearly outperforms the SSI method. The sharp increases and decreases in the correlation measure are due to the correlation calculation. Even if a certain recovery method recovers the amplitudes of the trace perfectly, a shift of one sample in the impulse times can produce a very small correlation measure at certain points in time (if we assume no adjacent impulses then the correlation will be 0). A shift of one sample in the recovery is quite common, since the wavelet itself is not an impulse and has a certain width in the time domain. This is highly affected by the wavelet sample rate.

We now examine the recovery of a specific channel, in this case, channel 20 (Figure 4.11). Comparison between the reflectivity series recovered by MSBD and the true reflectivity series (Figure 4.11(a)) demonstrates the same effects observed for the

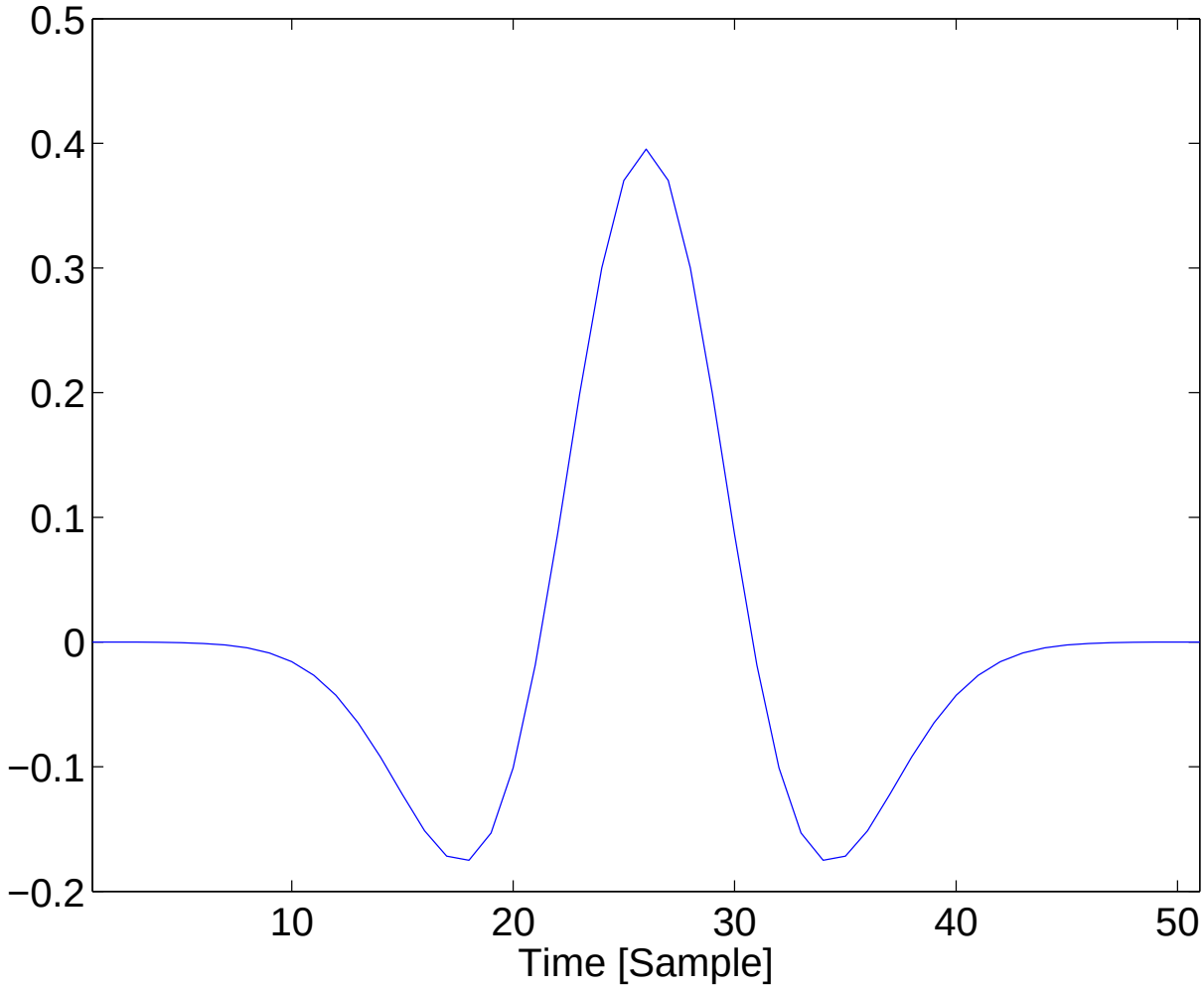

Figure 4.8: Wavelet parametric change: Ricker Wavelet, $f = 2$.

wavelet AWGN in a specific channel (Figure 4.5(a)). Now we examine the correlation between recovered and true reflectivity for the same channel as a function of iteration (Figure 4.11(b)). Again, the correlation demonstrates a generally increasing trend, however it also decreases at some points along the process. As in the case of wavelet AWGN, the correlation at the final iteration is not the highest among all iterations.

Finally, we examine the estimated wavelet at the last iteration compared to the true wavelet and the initial wavelet that was provided to us (Figure 4.12). We can see a very good estimation of the wavelet, even better than for the wavelet AWGN in Figure 4.6.

The mean correlations between the recovered and original reflectivity measured across

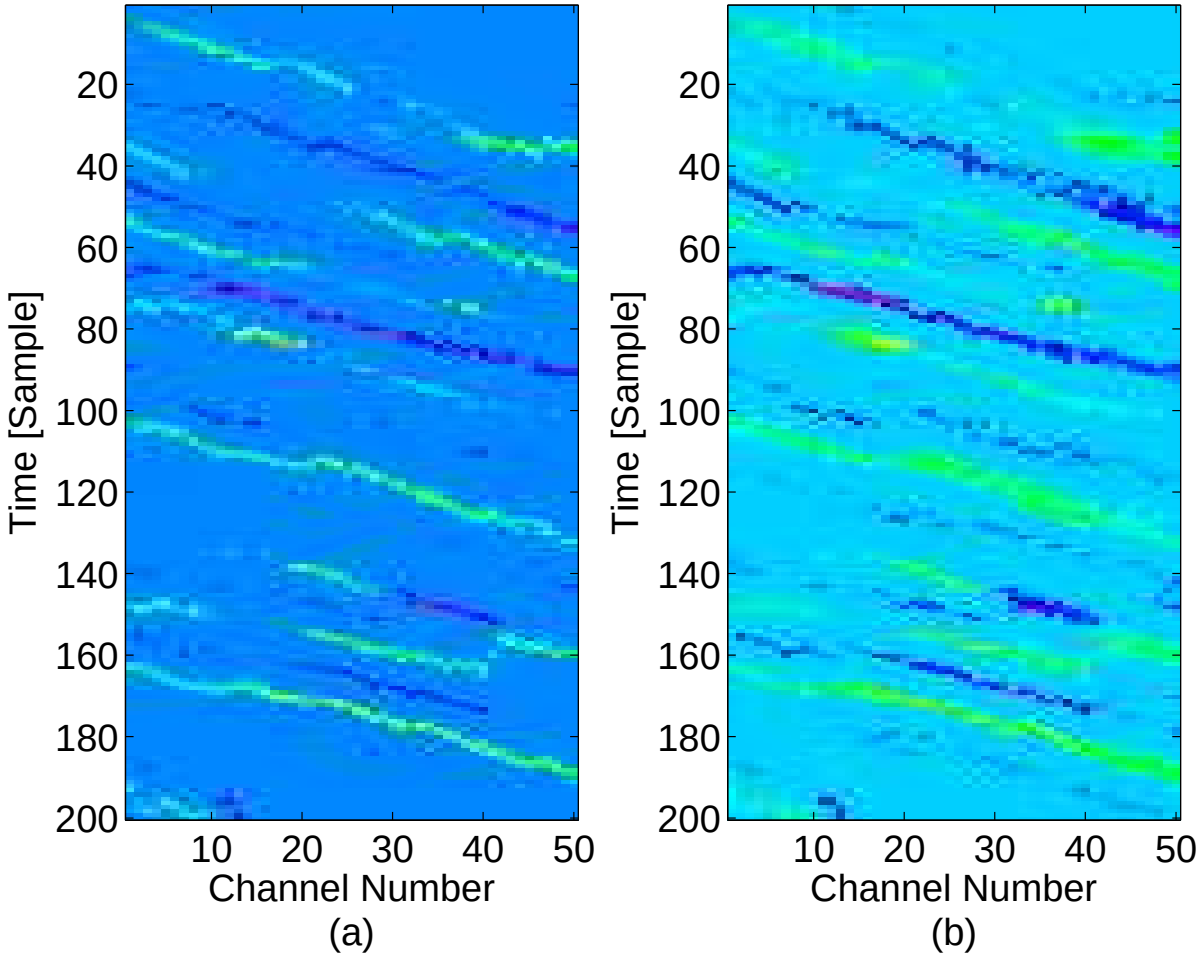

Figure 4.9: Wavelet parametric change, $SNR = 10dB$, $\sigma_f = 1$: (a) MSBD recovered reflectivity (b) SSI Recovered Reflectivity.

all channels, for all tested cases, i.e. SNR varies from 0 to 20 dB and σ_f varies from 0.5 to 1.2, are presented below in Table 4.2. Separate correlations for the SSI, SMBD and MSBD are presented for each value of SNR and σ_f. Please notice that because SMBD is a method for the blind deconvolution problem it has nothing to do with the initial information on the wavelet so its performances are the same when looking on different values of σ_f.

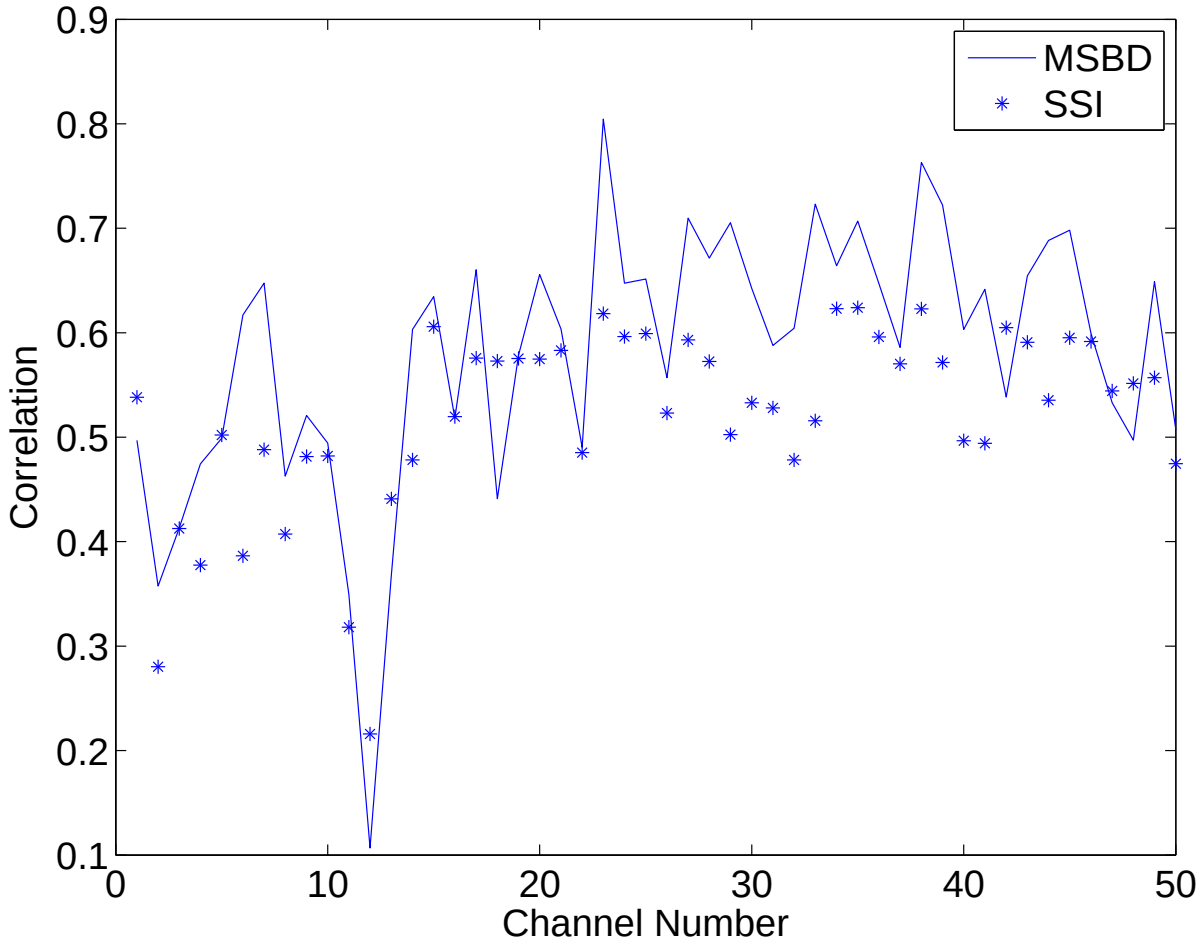

Figure 4.10: Wavelet parametric change, $SNR = 10dB$, $\sigma_f = 1$: Correlation measure vs. channel number.

Table 4.2: Correlations between recovered and original reflectivity for wavelet parametric change.

σ_f	SNR														
	0 dB			5 dB			10 dB			15 dB			20 dB		
	SSI	SMBD	MSBD	SSI	SMBD	MSBD	SSI	SMBD	MSBD	SSI	SMBD	MSBD	SSI	SMBD	MSBD
0.5	0.39	0.35	**0.53**	0.25	0.39	**0.61**	0.32	0.41	**0.92**	0.7	0.72	**0.77**	0.89	0.8	**0.89**
0.8	0.54		**0.6**	0.66		**0.74**	0.45		**0.69**	0.7		**0.82**	0.52		**0.81**
1	0.42		**0.56**	0.33		**0.69**	0.35		**0.69**	0.51		**0.76**	0.66		**0.73**
1.2	0.49		**0.53**	0.42		**0.65**	0.57		**0.67**	0.72		**0.78**	0.57		**0.80**

In general, the correlation between the recovered and original reflectivity increases as SNR increases , as was found for wavelet AWGN (Table 4.1). However, it is not clear whether a decrease in σ_f leads to better recovery. This is because under wavelet parametric change the relative amplitude of adjacent samples of the wavelet is maintained while only the width of the wavelet and the absolute amplitudes change. In other words,

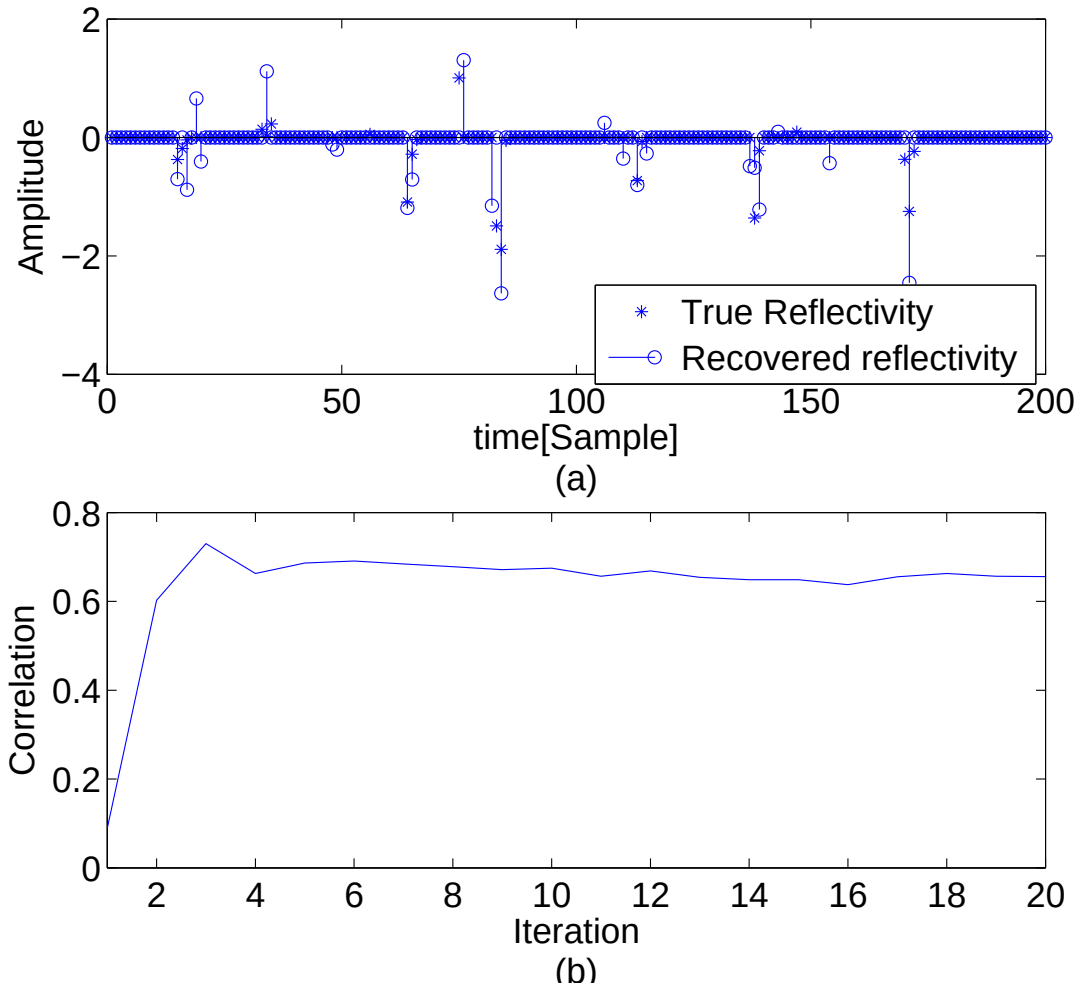

Figure 4.11: Wavelet parametric change, SNR = 10dB and $\sigma_f = 1$, specific channel: (a) MSBD recovered reflectivity compared to the original reflectivity, within a specific channel, (b) The correlation between recovered and true reflectivity in a specific channel as a function of iteration.

the maximum amplitude of the wavelet will be maintained at the center of the wavelet, so the quality of the recovery is not necessarily related to σ_f.

A very interesting point we can easily see in both simulations is that we can clearly see that if we have some initial knowledge about the wavelet, the semi-blind (MSBD) method has more potential in recovering the reflectivity series with better performance than the blind (SMBD) or non-blind (SSI) directives. This makes sense because non-blind methods assume the wavelet is given and does not change - which is not the case, and blind methods assume nothing about the initial wavelet so they are not exploiting all the information given in the problem. This clearly states the importance of the method we propose.

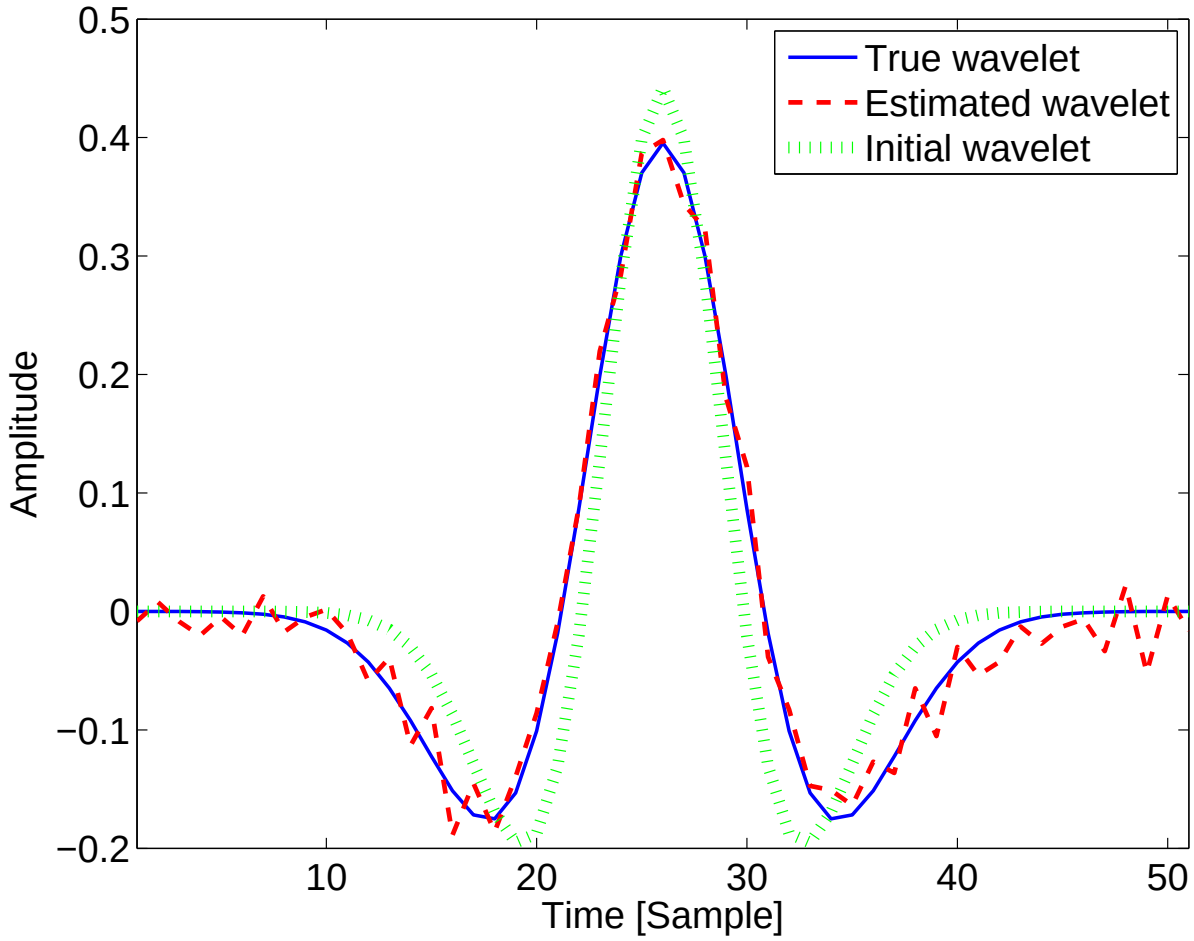

Figure 4.12: Wavelet parametric change: Wavelet estimation, SNR = 10dB and $\sigma_f = 1$.

4.3 Real Data Example

In the next figures we see a real data example, courtesy of GeoEnergy Inc., Texas. In this example we did not have a priori information about the wavelet, so we assumed a Ricker wavelet with $f = 5, \sigma_f = 0.5$. The SNR in this example is 2 dB. In Figure 4.13 we can see the seismic data that was given to us and the recovered reflectivity using MSBD. In Figure 4.14 we can see the initial wavelet that was given to the algorithm and the optimal wavelet found by the algorithm.

We can see the good recovery of MSBD. There are still discontinuities, for example in channel 32, around sample number 75. This discontinuities can be due to non-accurate assumptions on the noise in the setup. Still, we can clearly observe the deblurring effect

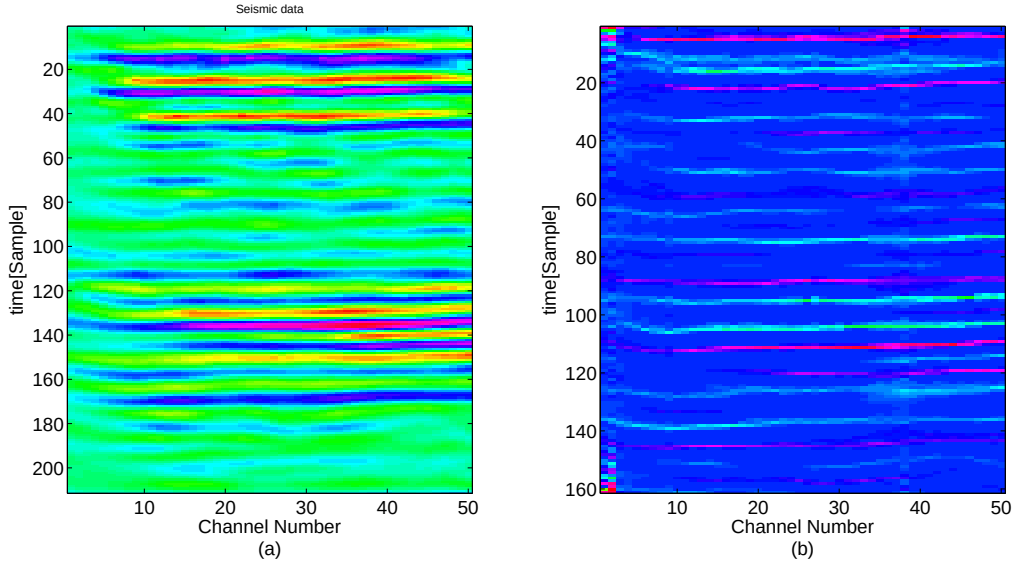

Figure 4.13: Real data example : (a) seismic data, (b) recovered reflectivity using MSBD.

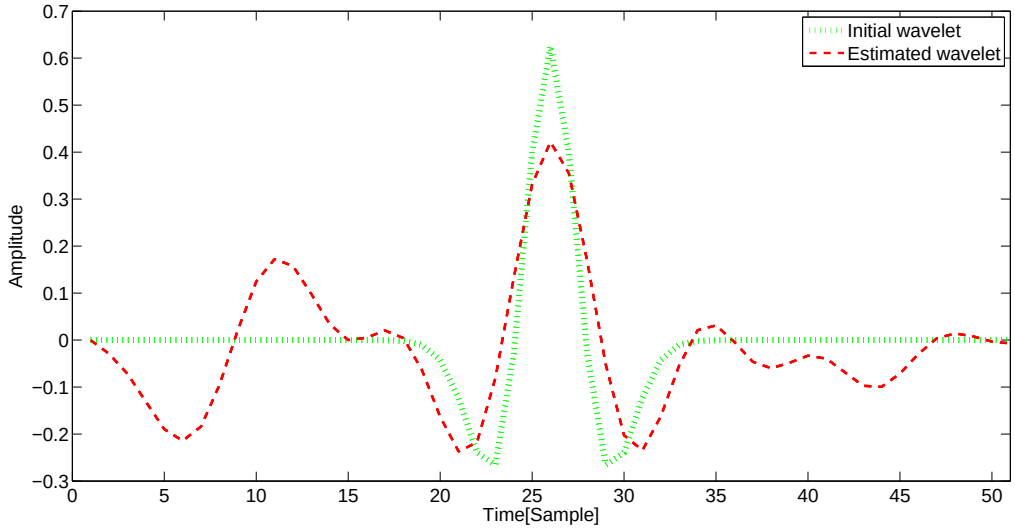

Figure 4.14: Real data example : Initial and estimated wavelet.

of the MSBD that has managed to distinguish between different spikes from near channels and from near samples in time.

4.4 Conclusions

In this chapter we tested MSBD with two different use case for the semi-blind deconvolution problem. For each case we mathematically developed its suitable solver with

emphasis on its unique way of total noise estimation in the channel. We compared the results of MSBD in each use case to SSI and SMBD algorithms and showed that MSBD outperforms them.

Chapter 5

Summary and conclusions

5.1 Summary

The deconvolution problem is a wide problem that is relevant in almost every physical system. Defining and solving deconvolution problem can help us learn physical systems, overcome undesired artifacts in a developed system, etc. Unlike the convolution operation, the deconvolution operation is not well defined. It depends on the mathematical model of the physical system under study, the mathematical model of the signals in that system, the a-priori information we hold about the system or the setup and so on. In this study we researched the deconvolution methods for seismic signals and presented a new method based on a two-stage iterative process that recovers the reflectivity series from the seismic data given a wavelet containing some kind of an uncertainty. We presented a general two-stage method, where one of the steps is fixed at the wavelet recovery stage, and the other changes from case to case at the reflectivity recovery stage. The reflectivity recovery stage does not fully change from case to case. All cases apply the BPDN solution for reflectivity recovery. The part in this stage that does change is the way we choose the trade-off parameter in the BPDN solution.

In this study we have presented two different cases in which we analytically calculated the trade-off parameter. For each case we presented the results of our proposed method and compared it to blind and non-blind methods. The results clearly show the advantage of MSBD. The immediate conclusion is that a stage of wavelet update is necessary and that the performance of our proposed method for both wavelet and reflectivity series

recovery is very promising.

5.2 Future Research

As was shown in the results section of MSBD, we currently choose the number of iterations with no analytic evaluation. We also see that the correlation measure does not always increase, nor the correlation measure of the last iteration is the highest among all iterations. A very helpful thing to research for will be the best point to stop the iterative process. This could potentially improve the results and open even more follow-up research questions.

A different point of research can be finding a more general form of the trade-off parameter. In the current research we found two forms to be helpful for the two different cases under test. As was mentioned before, those forms were chosen with no proof that they are the best ones. One of the most important things in the deconvolution process is to find the balance between sparsity and data fitting. So a wise choice of the trade-off parameter can greatly improve results.

A third point of follow-up research can be in the part of the wavelet recovery stage. As we can see, in the current algorithm, the wavelet recovery stage does not use the initial information given on the wavelet. The current method uses only the reflectivity signal to recover the wavelet. We believe that recovering the wavelet when taking into account the a-priori information on the wavelet could significantly improve the results.

Bibliography

[1] F. W. Schroeder, "The seismic method," *American Association Of Petroleum Geologists*.

[2] J. Kormylo and J. Mendel, "Maximum likelihood detection and estimation of Bernoulli-Gaussian processes," *IEEE, Trans. Inf. Theory*, vol. 28, pp. 482–488, 1982.

[3] K. F. Kaaresen and T. Taxt, "Multichannel blind deconvolution of seismic signals," *Geophysics*, vol. 63, pp. 2093–2107, 1998.

[4] A. Heimer, I. Cohen, and A. A. Vassiliou, "Multichannel seismic modeling and inversion based on Markov-Bernoulli random field," *IEEE trans. on Geosci. and Remote Sens.*, vol. 47, pp. 2047–2058, 2009.

[5] I. Ram, I. Cohen, and S. Raz, "Multichannel deconvolution of seismic signals using statistical MCMC methods," *IEEE Trans. on Signal Process.*, vol. 58, pp. 2757–2770, 2010.

[6] D. Velis and T. Ulrych, "Simulated annealing wavelet estimation via fourth-order cumulant matching," *Geophysics*, pp. 1939 – 1948, 1999.

[7] M. V. der Baan, "Time-varying wavelet estimation and deconvolution by kurtosis maximization," *Geophysics*, pp. V11 – V18, 2008.

[8] G. Menanno and A. Mazzotti, "Deconvolution of multicomponent seismic data by means of quaternions: Theory and preliminary results," *Geophys. Prospect.*, pp. 217 – 238, 2012.

[9] N. Kazemi and D. Sacchi, "Sparse multichannel blind deconvolution," *Geophysics*, vol. 79, pp. 143–152, 2014.

[10] A. Repetti, "Euclid in a taxicab: sparse blind deconvolution with smoothed l1/l2 regularization," *IEEE Signal Process. Lett.*, pp. 539–543, 2015.

[11] S. Gleichman and Y. Eldar, "Blind compressed sensing," *IEEE Trans. Inf. Theory*, pp. 6598 – 6975, 2011.

[12] M. Rosenbaum and A. Tsybakov, "Sparse recovery under matrix uncertainty," *The Annals of Statistics*, pp. 2620 – 2651, 2010.

[13] T. Nguyen and J. Castagna, "High resolution reflectivity inversion," *J. Seism. Explor.*, vol. 19, pp. 303–320, 2010.

[14] Y. Wang, "Seismic time-frequency spectral decomposition by matching pursuit," *Geophysics*, vol. 72, pp. V13–V20, 2007.

[15] ——, "Multichannel matching pursuit for seismic trace decomposition," *Geophysics*, vol. 75, pp. V61–V66, 2010.

[16] R. Zhang and J. Castagna, "Seismic sparse-layer reflectivity inversion using basic pursuit decomposition," *Geophysics*, vol. 76, pp. 147–158, 2011.

[17] S. S. Chen, D. L. Donoho, and M. A. Saunders, "Atomic decomposition by basis pursuit," *SIAM rev.*, vol. 43, pp. 129–159, 2001.

[18] S. S. Chen and D. L. Donoho, "Basis pursuit," *Signals Syst. Comput.*, vol. 1, pp. 41–44, 1994.

[19] J. Bork and L. Wood, "Seismic interpretation of sonic logs," *in: 71st Annual International Meeting, SEG, Expanded Abstracts*, pp. 510–513, 2001.

[20] M. Elad, *Sparse and Redundant Representations, 2010.* Springer, 2010.

[21] R. Zhang, M. K. Sen, and S. Srinivasan, "Multi-trace basis pursuit inversion with spatial regularization," *J. Geophys. Eng.*, vol. 10, 2013.

[22] ——, "A prestack basis pursuit seismic inversion," *Geophysics*, vol. 78, pp. R1–R11, 2013.

[23] C. Dossal and S. Mallat, "Sparse spike deconvolution with minimum scale," *in: Signal processing with adaptive sparse structured representations*, pp. 123–126, 2005.

[24] H. L. Taylor, S. C. Banks, and J. F. MccCoy, "Deconvolution with the L1 norm," *Geophysics*, vol. 44, pp. 39–52, 1979.

[25] D. W. Oldenburg, S. Levy, and K. J. Stinson, "Inversion of band-limited reflection seismograms: Theory and practice," *Proc. IEEE*, vol. 74, pp. 487 – 497, 1986.

[26] P. R. Gill, A. Wang, and A. Molnar, "The In-crowd algorithm for fast basis pursuit denoising," *IEEE Trans. Signal Process.*, vol. 59, pp. 4595–4605, 2011.

[27] W. Lu and N. Vaswani, "Modified basis pursuit denoising (Modified-BPDN) for noisy compressive sensing with partially known support," *IEEE International Conference on Acoustics, Speech and Signal Processing*, pp. 3926–3929, 2010.

[28] L. Dai and K. Pelckmans, "An ellipsoid based, two-stage screening test for BPDN," *in: European Signal Processing Conference*, pp. 654 – 658, 2012.

[29] E. V. den Berg and M. Friedlander, "Probing the Pareto Frontier for Basis Pursuit Solutions," *SIAM Journal on Scientific Computing*, vol. 31, pp. 890 – 912, 2011.

[30] ——, "Sparse optimization with least-squares constraints," *SIAM Journal on Scientific Computing*, vol. 21, pp. 1201 – 1209, 2008.

[31] I. Tosic and P. Frossard, "Dictionary learning - What is the right representation for my signal?" *IEEE Signal Process. Mag.*, pp. 27 – 38, 2011.

[32] K. Skretting and K. Engan, "Recursive least squares dictionary learning algorithm," *IEEE Trans. Signal Process.*, vol. 58, pp. 2121 – 2130, 2010.

[33] S. P. Lloyd, "Least squares quantization in PCM," *IEEE Trans. Inf. Theory*, vol. IT-28, pp. 129 – 137, 1982.

[34] J. Idier and Y. Goussard, "Multichannel seismic deconvolution," *IEEE Geosci. Remote Sens.*, vol. 31, pp. 961–979, 1993.

דה-קונבולוציה דלילה חצי עיוורת ורב ערוצית לאותות סייסמים

חיבור על מחקר

לשם מילוי חלקי של הדרישות לקבלת התואר מגיסטר למדעים בהנדסת חשמל

מרבי מירל

הוגש לסנט הטכניון – מכון טכנולוגי לישראל
אייר תשע"ח חיפה אפריל 2018

תודות

המחקר נעשה בהנחייתו של פרופ' ישראל כהן מהפקולטה להנדסת חשמל.

ברצוני להודות לפרופ' ישראל כהן על הנחייתו המסורה ועל התמיכה וההדרכה במהלך השלבים השונים במחקר. הידע והניסיון שלו בתחום מחקר זה הועילו לתהליך המחקר בצורה משמעותית ובאו לידי ביטוי בהערות ובהצעות המקצועיות למחקר.
בנוסף, ברצוני להודות לאנטוני ואסיליו מחברת GeoEnergy על תרומתו למחקר זה.

ברצוני להודות למשפחתי היקרה ולאשתי האהובה והיקרה, שירה, על העזרה, התמיכה וההבנה לאורך הדרך. תודה על האהבה והעידוד המתמידים.

תקציר

גלים סייסמים הם צורה של אנרגיה שמתפשטת בשכבות השונות אשר מתחת לפני הקרקע של כדור הארץ. מהירות ההתקדמות של גלים אלו תלויה בצפיפות ובאלסטיות של התווך. אלסטיות התווך נמדדת ע"י פרמטר כמותי שנקרא אימפדנס אקוסטי. כאשר גל סייסמי נע בתווך ופוגע בשכבה שמפרידה בין שני אזורים בעלי אימפדנסים אקוסטים שונים, חלק מהאנרגיה של הגל מוחזרת לכיוון פני הקרקע וחלק מהאנרגיה ממשיכה לשכבות עמוקות של התת-קרקע.

שיקוף סייסמולוגי (Reflection seismology) הוא תחום שמנצל את התכונות הנ"ל של גלים סייסמים כדי לחקור את התת-קרקע ולנסות לייצר תמונה אמינה שלו. במחקרים מסוג זה, מייצרים גלים סייסמים באופן מלאכותי ודוגמים את ההחזרים שלהם על פני הקרקע בעזרת מכשירים הנקראים סייסמוגרפים. לאחר מכן משתמשים בדגימות אלו כדי לשחזר את המסלולים שעברו הגלים הסייסמים וכך לצייר תמונה של התת-קרקע. אולם דגימות אלה לא תמיד מייצגות בצורה אמינה את התת-קרקע, שכן כל אות עובר תהליך של עיוות בזמן תנועתו בתווך.

דה-קונבולוציה הוא תהליך בו אנו יודעים או מודדים את יציאת המערכת וגרעין המערכת ומנסים לשערך או לחשב את כניסת המערכת. כאשר הגרעין ידוע, תהליך זה נקרא דה-קונבולוציה לא עיוורת. במקרים מסוימים איננו יודעים גם את גרעין המערכת ואז המטרה היא לשערך או לחשב את כניסת המערכת ובנוסף גם את גרעין המערכת. תהליך זה נקרא דה-קובנולוציה עיוורת.

תהליך הדה-קונבולוציה מתבצע לצורך מטרות שונות, כגון שערוך אותות ופרמטרים בהינתן מודל מתמטי של מערכת או הסרת אפקטים לא רצויים כתוצאה ממעבר של אותות במערכות אשר לא בשליטתנו. בניגוד לפעולת הקונבולוציה, פעולת הדה-קונבולוציה איננה מוגדרת היטב. קיימות גישות ושיטות שונות לביצוע הפעולה, כל אחת עם יתרונותיה וחסרונותיה.

דה-קונבולוציה סייסמית היא התהליך שבו לוקחים את הדגימות הגולמיות של הסייסמוגרפים ומסירים מהם את העיוותים שנוצרו כתוצאה מהמעבר בתווך. האות הרצוי נקרא אות השיקוף (reflectivity signal), האות המתאר את התווך נקרא אדוה (wavelet) והדגימות הסייסמיות נקראות אותות סייסמים (Seismic signal).

עד כה עסקו רבות בשיטות דה-קונבולציה סייסמיות אשר מניחות ידיעה מלאה של התווך או לחילופין, חוסר ידיעה מוחלט שלו. פותחו שיטות דה-קונבולוציה שונות, תחת הנחות של מודלים מתמטיים שונים, במטרה להתאים את הפתרון לבעיה הנתונה על הצד הטוב ביותר. כאשר קיים מידע חלקי על גרעין הקונבולוציה, שיטות של דה-קונבולוציה עיוורת יכולות להניח שהן לא יודעות את הגרעין כלל וכך לפתור את הבעיה. לחילופין, שיטות דה-קונבולוציה לא עיוורת יכולות להניח שהמידע החלקי הוא המידע המלא או לבצע ניחוש מושכל לצורך השלמת המידע החסר וכך לפתור את הבעיה.

מחקר זה עוסק בדה-קונבולוציה של אותות סייסמים תחת הנחת מודלים דלילים, כאשר קיים מידע חלקי בלבד על גרעין הקונבולוציה. מטרת המחקר הינה לפתח אלגוריתם חדש המנצל מידע חלקי בלבד כדי לשפר את ביצועי הדה-קונבולוציה. מידע חלקי זה יכול להיות נתון בצורות שונות, כגון מודל מתמטי של הגרעין אך ללא ערכי הפרמטרים המתארים אותו או גרעין מעוות תחת מודל עיוות ידוע. מעתה נתייחס לבעיה זו בשם "דה-קונבולוציה חצי עיוורת".

בדומה לדה-קונבולוציה העיוורת, גם בדה-קונבולוציה החצי עיוורת המטרה היא לשחזר את גרעין המערכת ואת אות הכניסה למערכת. כאמור, עבור אותות סייסמים אותות אלו נקראים אדוה ואות השיקוף. שני האותות הללו הינם בעלי מאפיינים ומודלים מתמטיים שונים, ולכן אופן השחזור של כל אחד מהם יהיה בהתאם לתכונותיו ומאפייניו.

במחקר זה נתמקד במודל רב-ערוצי של אותות סייסמים. המשמעות היא שיש מספר סייסמוגרפים אשר פזורים במרחב, וכל אחד מהם דוגם סימולטנית על פני הקרקע את החזרי הגלים הסייסמים. בנוסף, אנו מניחים מודל דליל עבור אותות השיקוף.

אנו מציגים אלגוריתם איטרטיבי חדשני, בעל שני שלבים בכל איטרציה. בשלב הראשון אנו מניחים שהאדוה ידועה לנו ומשתמשים בשיטות דה-קונבולוציה של אותות דלילים כדי לשחזר את אות השיקוף. אחד מהאלגוריתמים הנפוצים ביותר בדה-קונבולוציה סייסמית לא עיוורת הוא אלגוריתם SSI (Sparse Spike Inversion). אלגוריתם זה הינו בעצם מימוש של אלגוריתם דה-קונבולוציה ידוע הנקרא BPDN (Basis Pursuit Denoising). אלגוריתם זה נפוץ מאוד בפתרון בעיות דה-קונבולוציה עבור אותות דלילים, ולכן מקובל מאוד להשתמש בו עבור אותות סייסמים. SSI יכול לפעול על ערוץ סייסמי בודד, ובהינתן אות האדוה ניתן לשחזר את אות השיקוף.

עבור SSI (או BPDN) קיים פרמטר עיקרי אחד אשר נדרש לקבוע אותו טרם הפעלת האלגוריתם. הפרמטר מאזן בין דלילות האות המשוחזר לבין התאמה לאות הסייסמי הדגום. חשוב להדגיש כי הספרות אינה מגדירה איך נכון לבחור את הפרמטר הנ"ל. לכל סוג אות ניתן לחשב ולהתאים את הפרמטר בהתאם לאופיו ולתכונותיו. אחד מהמודלים המתמטיים אשר מייצגים את אלגוריתם ה-SSI טומן בחובו יכולת פירוש כמותית לפרמטר זה. עבור המודל המתמטי שבחרנו ל-SSI אנו בוחרים את הפרמטר המאזן להיות סטיית התקן של כמות הרעש הכוללת במערכת. אנו מציגים שיטות שונות לחישוב כמות הרעש הכוללת ומראים תחת אילו תנאים מומלץ לבחור כל אחת מהשיטות השונות. הרעש הכולל באלגוריתם המוצע מבוסס על הנחה של רעש גאוסי לבן בערוץ הדגימה ועל חוסר ודאות באות האדוה. חוסר הודאות הנ"ל מתואר כתוספת רעש חיבורי לאדוה האמיתית אשר בעזרתה נוצרו האותות הסייסמים.

בשלב השני אנו מניחים שאות השיקוף ידוע, ומשתמשים בשיטות של לימוד מילון כדי לשחזר את האדוה. בניגוד לאות השיקוף, האדוה אינה ניתנת למידול כאות דליל. לכן, בשלב זה ניגש לשיטות דה-קונבולוציה אחרות מאלו שהופעלו בשלב הראשון.
במודל המתמטי שמגדיר את הבעיה, ניתן להסתכל על מטריצת הקונבולוציה, אשר מקורה באדוה, כאל מילון שממנו מורכב האות הסייסמי. בסופו של דבר, האות הסייסמי הוא

קומבינציה ליניארית של עמודות מטריצת הקונבולוציה, כאשר ערכי אות השיקוף הם המקדמים. בהינתן נקודת המבט הזו, ניתן להסתכל על בעיית שחזור האדוה כאל תהליך לימוד מילון. אנו משתמשים באלגוריתם לימוד מילון הנקרא Signature Dictionary עם ההתאמות הנדרשות בכדי שיתאים למודל המתמטי של מטריצת הקונבולוציה הנתונה. תהליך זה לוקח בחשבון את כל הערוצים השונים ומשתמש במזעור נורמת L2 כדי למצוא את האדוה האופטימלית בהינתן אותות השיקוף של כל הערוצים.

בזמן פיתוח האלגוריתם שמרנו על כלליות מסוימת בשלבי האלגוריתם. באלגוריתם הסופי אין צורך לבחון כל אות סייסמי מחדש ולהתאים לכל אות פרמטרים מחדש. ניתן לקחת קבוצת אותות סייסמים שהמכנה המשותף להן הוא שהן חולקות חוסר ודאות זהה לגבי האדוה ולחשב לכל הקבוצה סט פרמטרים אחד אשר איתו ניתן להריץ מתאר אחד של האלגוריתם לכל האותות בקבוצה.

בשלב בחינת הביצועים של האלגוריתם, אנו בודקים שני מקרים שונים של חוסר ודאות באדוה. במקרה הראשון אנו מניחים שהאדוה האמיתית מורעשת ברעש גאוסי לבן. אנו מנתחים את המקרה, מחשבים ומתאימים את כל הפרמטרים בהתאם למודל הרעש שבחרנו. במקרה השני אנו מניחים מודל מתמטי נתון עבור האדוה אך חוסר ודאות לגבי הפרמטר המגדיר אותה. ספציפית, אנו מניחים אדוה מסוג Ricker wavelet עם פרמטר תדר שיש בו רעש גאוסי. במקרה זה הרעש באדוה אינו חיבורי ולכן אנו מחשבים טור טיילור למודל המתמטי של האדוה ומתייחסים לקירוב מסדר שני.

אנו משווים את התוצאות של האלגוריתם שלנו לאלגוריתם ה-SSI אשר מייצג את גישת הדה-קונבולוציה הלא עיוורת ולאלגוריתם ה-SMBD (Sparse Multichannel Blind Deconvolution) אשר מייצג את גישת הדה-קונבולוציה העיוורת. אנו משתמשים בקורלציה בין אות השיקוף המשוחזר לאות השיקוף המקורי כמדד לאיכות השחזור. אנו מראים כי ביצועי האלגוריתם שלנו טובים יותר מביצועי שני האלגוריתמים הנ"ל בהתייחס למדד הקורלציה.

לסיכום, במחקר זה אנו מגדירים שונה, דה-קונבולוציה חצי עיוורת, בה נתון מידע חלקי על גרעין המערכת. במחקר זה סקרנו ולמדנו את שיטות הדה-קונבולוציה הנפוצות ופיתחנו אלגוריתם חדש המנצל את המידע הנתון על גרעין המערכת, גם אם הוא חלקי. האלגוריתם איטרטיבי ובעל שני שלבים שונים במהותם שמטרת כל אחד מהם היא לשחזר סוג שונה של אותות. יחד עם זאת, בהסתכלות כללית יותר על שני שלבי השחזור, אנו רואים כיצד הם בכל זאת משפיעים האחד על השני וכן את ההתאמות הנדרשות לכל אחד מהשלבים. התאמות אלו נובעות מהעובדה שכעת כל אחד מהם הוא חלק מאלגוריתם גדול יותר והם לא עומדים בפני עצמם כאלגוריתם מלא.

חשיבות המחקר היא בגישה להגדרת הבעיה החדשה, דה קונבולוציה חצי עיוורת עבור אותות סייסמים, ובהתאם פיתוח אלגוריתם המנצל את כל המידע בבעיה ואת המודל המתמטי החדש המגדיר אותה.

פרקי התזה כתובים באופן הבא. בפרק הראשון אנו מציגים את המוטיבציה והמטרות של המחקר שהתבצע ומציגים בפירוט את מבנה התזה. בפרק השני אנו סוקרים בצורה מפורטת

את נושא הדה-קונבולוציה הסייסמית, כולל סקר ספרות נרחב, ניסוח מתמטי של הבעיה והצגה מפורטת של שיטות נבחרות בדה-קונבולוציה סיסמית עיוורת ולא עיוורת. בפרק השלישי אנו מציגים את האלגוריתם המוצע, MSBD. אנו מגדירים את בעיית הדה-קונבולוציה החצי עיוורת באופן מתמטי ומציגים לפרטים את הפתרון האיטרטיבי הדו-שלבי שאנו מציעים. בפרק הרביעי אנו מציגים את ביצועי המערכת עבור שני מקרי בוחן כפי שתואר בפסקאות הקודמות. עבור כל מקרה בוחן אנו מחשבים ומתאימים את הפרמטרים השונים באלגוריתם ובודקים ביצועים. בנוסף, אנו משווים את ביצועי האלגוריתם שלנו לשני אלגוריתמי דה-קונבולוציה שונים. בפרק החמישי אנו מסכמים את המחקר ומציגים עבודות עתידיות שניתן לבצע בעקבות מחקר זה.