Sparse Seismic Inversion

Deborah Pereg
Sparse Seismic Inversion

Research Thesis

Submitted in partial fulfillment of the requirements
for the degree of Master of Science
in Electrical Engineering

Deborah Pereg

Submitted to the Senate of
the Technion – Israel Institute of Technology
Elul Hatashav Haifa September 2016
The research thesis was done under the supervision of Prof. Israel Cohen in the Electrical Engineering Department.

I would like to express my deep gratitude to Prof. Israel Cohen for his guidance and support throughout this research. I would also like to thank my family, colleagues and friends.

The generous financial help of the Technion is gratefully acknowledged.
# Contents

Abstract

1

Notations and Abbreviations

3

1 Introduction

7

1.1 Background and Motivation

7

1.2 Research Overview

13

1.3 Organization of the Thesis

14

1.4 List of Papers

15

2 Sparse Seismic Single-Channel Deconvolution Methods

16

2.1 Basic Theory

17

2.2 Synthetic Examples

18

2.3 Real Data Results

20

3 Multichannel Sparse Spike Inversion

25

3.1 Problem Formulation

26

3.2 Multichannel Sparse Spike Inversion (MSSI)

27

3.3 Experimental results

30

4 Seismic Recovery Based on Earth Q Model

44

4.1 Introduction

44

4.2 Signal Model

47
List of Figures

1.1 Synthetic seismic data, reflectivity and seismic wavelet: (a) 2D seismic data (SNR = 10 dB); (b) Synthetic 2D reflectivity section; (c) Wavelet. .................................................. 9

2.1 1D synthetic tests of SSI. (a) True reflectivity. (b) Synthetic trace with 40 Hz Ricker wavelet and SNR= 10 dB. (c)-(f) SSI inversion results with varying $\lambda_{SSI}$. (c) $\lambda_{SSI} = 0.29$, (d) $\lambda_{SSI} = 0.11$, (e) $\lambda_{SSI} = 0.071$, (f) $\lambda_{SSI} = 0.025$. ......................... 21

2.2 1D synthetic tests of BPI. (a) True reflectivity. (b) Synthetic trace with 40 Hz Ricker wavelet and SNR= 10 dB. (c)-(f) BPI inversion results with varying $\lambda_{BPI}$. (c) $\lambda_{BPI} = 0.27$, (d) $\lambda_{BPI} = 0.087$, (e) $\lambda_{BPI} = 0.011$. ................. 22

2.3 (a) $\lambda$-correlation curve for SSI based on the synthetic data in Figure 2.1. (b) $\lambda$-correlation curve for BPI based on the synthetic data in Figure 2.2. ......................... 23

2.4 Seismic data. .............................................................. 24

2.5 (a) Estimated reflectivity matrix; (b) Reconstructed seismic data. ......................................................... 24

3.1 Synthetic reflectivity, wavelet and data sets: (a) Synthetic 2D reflectivity section; (b) 2D seismic data (SNR = 10 dB); (c) 2D seismic data (SNR = 5 dB); (d) Wavelet. ................. 38
3.2 Synthetic 2D data deconvolution results: (a) Single-channel deconvolution results for SNR = 10 dB; (b) MC-II deconvolution results for SNR = 10 dB; (c) MSSI-2 results for SNR = 10 dB; (d) MSSI-3 results for SNR = 10 dB.

3.3 Synthetic 2D data deconvolution results: (a) Single-channel deconvolution results for SNR = 5 dB; (b) MC-II deconvolution results for SNR = 5 dB; (c) MSSI-2 results for SNR = 5 dB; (d) MSSI-3 results for SNR = 5 dB.

3.4 Correlation coefficient vs. deconvolution parameters $\lambda_1$ and $\lambda_0$ for synthetic 2D data deconvolution (SNR = 5 dB).

3.5 Real data and assumed wavelet: (a) Real seismic data (SNR = 5 dB); (b) wavelet.

3.6 Real data deconvolution results: (a) Single-channel estimated reflectivity; (b) MC-II estimated reflectivity; (c) MSSI-2 estimated reflectivity; (d) MSSI-3 estimated reflectivity; (e) Single-channel reconstructed data; (f) MC-II reconstructed data; (g) MSSI-2 reconstructed data; (h) MSSI-3 reconstructed data.

4.1 Centered synthetic reflected wavelets and their derivatives, $Q = 125, \omega_0 = 100\pi\ (50Hz)$ (a) $g_{x,m}(t)$; (b) $g^{(1)}_{x,m}(t)$; (c) $g^{(2)}_{x,m}(t)$.

4.2 Support detection vs. the separation constant $\nu$. Rate of success is the average number of perfect recoveries out of 10 experiments.

4.3 Recovery error $\|\hat{x} - x\|_{\ell_1}$ as a function of noise level $\delta$ for $Q = \infty, 500, 200, 100$. (a) Experimental results; (b) Theoretical bounds.
4.4 1D synthetic tests of (a) True reflectivity. (b),(c) Synthetic trace with 50 Hz Ricker wavelet and SNR= $\infty$, 15.5 dB respectively, $Q = 200$. (d),(e) Recovered 1D channel of reflectivity signal. 59

4.5 1D synthetic tests of (a) True reflectivity. (b) Synthetic trace with 25 Hz Ricker wavelet and SNR= 12.9 dB, $Q = 500$. (c) Recovered 1D channel of reflectivity signal with SOOT. (d) Recovered 1D channel of reflectivity signal with the proposed time-variant model. 60

4.6 Real data inversion results: (a) Real seismic data (b) Estimated reflectivity (c) Reconstructed data. 62

4.7 Real data inversion results: (a) Estimated reflectivity - time-invariant model (SSI) (c) Estimated reflectivity - time-variant model. 63
List of Tables

4.1 Synthetic example: theoretical and estimated parameters: $Q$, the degradation ratio $\frac{a_0}{\gamma_0}$, $\beta$, $\frac{4(N\sigma)^2}{\beta} \frac{a_0}{\gamma_0}$ - the bound slope computed from known parameters (by Theorem 1 $||\hat{x} - x||_{\ell_1} \leq \frac{4(N\sigma)^2}{\beta} \frac{a_0}{\gamma_0} \delta$), and the estimated slope computed from the experimental results in Fig.3(a). . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 58
Abstract

Seismic deconvolution aims to recover the earth structure hidden in the acquired seismic data. In exploration seismology, a short-duration acoustic pulse is transmitted from the earth surface. The reflected pulses from the ground are then received by a sensor array and processed into a two-dimensional (2D) seismic image. Unfortunately, this image does not represent the actual image of the ground. We will refer to the hidden ground image we are estimating as the reflectivity.

The observed seismic data can be modeled as a convolution between each column in the 2D reflectivity section and a one-dimensional (1D) seismic pulse (wavelet), with additive noise. The reflectivity is assumed to be sparse. Therefore, deterministic deconvolution methods often use sparse inversion techniques.

In this thesis we present two algorithms that perform seismic recovery. We apply both algorithms to synthetic and real seismic data and demonstrate improved performance and robustness to noise, compared to competitive algorithms.

The first algorithm - Multichannel Sparse Spike Inversion (MSSI) - takes advantage of the horizontal spatial correlation between neighboring traces in the reflectivity image. MSSI is an iterative procedure, which deconvolves the seismic data and recovers the earth 2D reflectivity image, while taking into consideration both the desired sparsity of the solution and the dependencies between spatially-neighboring traces. Visually, it can be seen that
the layer boundaries in the estimates obtained by MSSI are more continuous and smooth than the layer boundaries in the single-channel deconvolution estimates.

The second algorithm takes into account the attenuation and dispersion propagation effects of the reflected waves, in noisy environment. We present an efficient method to perform seismic time-variant inversion considering the earth Q-model. We derive the theoretical bounds on the recovery error, and on the localization error. It is shown that the solution consists of recovered spikes which are relatively close to every spike of the true reflectivity signal. In addition, we prove that any redundant spike in the solution which is far from the correct support will have small energy.
Notations and Abbreviations

Abbreviations

1D — one-dimensional
2D — two-dimensional
SSI — Sparse Spike Inversion
BPI — Basis Pursuit Inversion
LARS — Least-Angle Regression
LASSO — Least Absolute Shrinkage and Selection Operator
SNR — Signal-to-Noise Ratio
MPD — Matching Pursuit Decomposition
BPD — Basis Pursuit Decomposition
MSSI — Multichannel Sparse Spike Inversion
MED — Minimum Entropy Deconvolution
MBRF — Markov-Bernoulli Random Field
MBG — Markov-Bernoulli-Gaussian
SMLR — Single Most Likely Replacement
Alphabetic Symbols

\( a, b \) — coefficient vectors of the dual certificate function
\( c_m \) — reflector amplitude
\( C_{0,1,2,3} \) — global property constants of admissible kernel
\( \tilde{C}_{0,1,2,3} \) — maximum global property constants of a set of admissible kernels
\( D_1, D_2 \) — recovery error bound parameters
\( D_3 \) — localization error bound parameter
\( f(t) \) — a source waveform
\( g_{i+k} \) — partial derivative of cost function \( \frac{\partial J}{\partial r_{i+k}} \)
\( g_{r,m}(t) \) — a reflected wave
\( h_{i+k,l} \) — normalized gradient
\( H_k \) — convolution matrix of a Low-pass filter
\( J \) — number of columns taken into account in estimation (MSSI algorithm)
\( J(\cdot) \) — cost function
\( K \) — a set of reflection points
\( k_m \) — discrete travel time
\( l \) — iteration index
\( m \) — reflector’s index
\( N \) — sampling frequency
\( n[k] \) — discrete additive noise
\( n(t) \) — single-channel additive noise
\( Q \) — the portion of energy lost during each cycle or wavelength
\( q[k] \) — discrete dual certificate function
\( q(t) \) — the dual certificate function
\( \hat{r} \) — an estimate of the reflectivity series
\( r_i \) — \( i \)’th reflectivity column
\( r_{i+k} \) — previous or subsequent column to \( r_i \)
\( r_e(t, m, n, \Delta t) \) — even wedge reflectivity (BPI)
\( r_o(t, m, n, \Delta t) \) — odd wedge reflectivity (BPI)
\( r_{M \times 1}, \quad r \) — discrete single-channel reflectivity
\( r(t) \) — single-channel reflectivity series
<table>
<thead>
<tr>
<th>Symbol</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>$s_i$</td>
<td>i’th observed seismic trace</td>
</tr>
<tr>
<td>$s_{i+k}$</td>
<td>previous or subsequent columns to $s_i$</td>
</tr>
<tr>
<td>$s_{N\times 1}$, $s$</td>
<td>discrete single-channel seismic trace</td>
</tr>
<tr>
<td>$s(t)$</td>
<td>received 1D seismic signal</td>
</tr>
<tr>
<td>$T$</td>
<td>a set of reflection travel times</td>
</tr>
<tr>
<td>$t_m$</td>
<td>travel time</td>
</tr>
<tr>
<td>$u(t)$</td>
<td>a reflected wave</td>
</tr>
<tr>
<td>$w$</td>
<td>discrete wavelet</td>
</tr>
<tr>
<td>$W_{N\times M} \in \mathbb{R}^{N\times M}$, $W$</td>
<td>convolution matrix associated with $w$</td>
</tr>
<tr>
<td>$w(t)$</td>
<td>seismic wavelet</td>
</tr>
<tr>
<td>$x[k]$</td>
<td>discrete single-channel reflectivity</td>
</tr>
<tr>
<td>$x(t)$</td>
<td>single-channel reflectivity</td>
</tr>
<tr>
<td>$y[k]$</td>
<td>discrete single-channel seismic trace</td>
</tr>
<tr>
<td>$y(t)$</td>
<td>single-channel seismic trace</td>
</tr>
<tr>
<td>$\alpha_l$</td>
<td>maximum l’th derivative at $t = 0$ of a set of reflected waves</td>
</tr>
<tr>
<td>$\beta$</td>
<td>local property constant of admissible kernel</td>
</tr>
<tr>
<td>$\gamma$</td>
<td>parameter of earth Q-model</td>
</tr>
<tr>
<td>$\gamma_l$</td>
<td>minimum l’th derivative at $t = 0$ of a set of reflected waves</td>
</tr>
<tr>
<td>$\delta$</td>
<td>upper bound on noise $\ell_1$ norm</td>
</tr>
<tr>
<td>$\delta[k]$</td>
<td>Dirac measure</td>
</tr>
<tr>
<td>$\Delta t$</td>
<td>sample rate</td>
</tr>
<tr>
<td>$\varepsilon$</td>
<td>local property constant of admissible kernel</td>
</tr>
<tr>
<td>$\eta_k$</td>
<td>Lagrange multipliers</td>
</tr>
<tr>
<td>$\kappa r$</td>
<td>propagation phase change</td>
</tr>
<tr>
<td>$\lambda, \lambda_0, \lambda_k$</td>
<td>regularization parameters</td>
</tr>
<tr>
<td>$\mu_l$</td>
<td>adaptive step size</td>
</tr>
<tr>
<td>$\nu$</td>
<td>separation constant</td>
</tr>
<tr>
<td>$\rho$</td>
<td>recovery error bound parameters</td>
</tr>
<tr>
<td>$\rho_{s,s}$</td>
<td>normalized correlation coefficient</td>
</tr>
<tr>
<td>$\sigma$</td>
<td>wavelet scaling</td>
</tr>
<tr>
<td>$\omega$</td>
<td>radial frequency</td>
</tr>
<tr>
<td>$\omega_0$</td>
<td>Ricker wavelet maximum radial frequency</td>
</tr>
</tbody>
</table>
\( \mathcal{R}_\epsilon(r) \) — smoothed \( \ell_1 \) norm approximation

\( \mathcal{L}(\cdot) \) — Lagrangian function

\( \| \cdot \|_0 \) — \( L_0 \) norm

\( \| \cdot \|_1 \) — \( L_1 \) norm

\( \| \cdot \|_2 \) — \( L_2 \) norm
Chapter 1

Introduction

1.1 Background and Motivation

In Signal Processing, it is often necessary to recover an input signal from its filtered version. The operation of deconvolution is ideally set to achieve this goal, and to undo the operation of a linear time invariant system performed on the input signal. Another related problem is the problem of decomposing a signal into its building blocks (atoms) [1]. Atomic decomposition is very common in many fields in Signal Processing, such as: image processing [2], compressed sensing [3], radar [4], ultrasound imaging [5], seismology [6–8] and more.

In the seismic setting, a short-duration acoustic pulse is transmitted from the earth surface. The reflected pulses from the ground are then received by a sensor array [9]. Our goal is to reveal the ground layer’s structure hidden in each of the received seismic traces. Unfortunately, the image which consists of the seismic data does not represent the actual image of the ground (the reflectivity). Under simplifying assumptions, the seismic trace (one column in the seismic image) can be modeled as the convolution between the earth reflectivity series and the source wavelet, corrupted by additive noise. An example of 2D synthetic seismic data ($SNR = 10$ dB),
the reflectivity section and the seismic wavelet is shown in Fig. 1.1(a), (b) and (c), respectively.

Seismic noise may be either coherent noise (surface waves, multiples, reflections or reflected refractions from near-surface objects, noise caused by vehicular movement on the ground, etc.) or incoherent noise (random noise from the cables or the geophones, wind, random movement on the ground, etc.). Everything that is not an event from which we obtain information is “noise”. Also, one has to take into account the physical properties associated with the propagation of stress waves in materials. Namely, the absorption of the wave’s energy and the resulting change in its shape.

Throughout our work, we assume that the short seismic pulse (the wavelet) is known. In Chapter 2 and Chapter 3 we also assume that the wavelet is approximately time-invariant. In Chapter 4 we consider a time-variant model, which takes into account the attenuation and dispersion of the wave’s energy during its propagation through the medium.

The assumption of an invariant seismic wavelet is common in seismic data processing [6, 7, 10–12]. Yet, even under this assumption, the inversion process is often unstable. The seismic wavelet is bandlimited, and the seismic trace might be noisy. Due to this instability, there are many possible reflectivity series that could fit the same measured seismic traces. The objective of our work is to find the best estimate of the reflectivity. We assume the reflectivity is sparse. Hence, its extraction could be done by sparse inversion techniques.

In previous works, the solution to the multichannel deconvolution problem involves separation of the seismic data into independent vertical one-dimensional (1D) deconvolution problems, where each reflectivity channel is estimated apart from the other channels [6, 7, 9, 11, 13–17]. The wavelet is taken to be a 1D column signal, and each 1D reflectivity column appears in the vertical direction as a sparse spike train. Some of these methods
Figure 1.1: Synthetic seismic data, reflectivity and seismic wavelet: (a) 2D seismic data (SNR = 10 dB); (b) Synthetic 2D reflectivity section; (c) Wavelet.
describe the reflectivity and the noise as two independent stochastic processes with known second-order statistics. Berkhout [9] tried to solve the seismic blind deconvolution problem by assuming that the reflectivity is a white sequence and that the seismic wavelet is a minimum phase signal. Many attempts have been made to avoid the minimum phase assumption. Some of these methods are blind, meaning that both the reflectivity and the wavelet are unknown. Homomorphic deconvolution [18], implemented in exploration seismology by Ulrych, was first developed for restoring reverberated and resonated sound and speech. It was also implemented for the case of blurred images [18]. In homomorphic deconvolution, we find the log amplitude of the distorting system in the frequency domain. Then we can restore the signal of interest by simply subtracting the log amplitude of the distorting system from the log amplitude of the observation signal in the frequency domain. Minimum Entropy Deconvolution (MED) [14] and Maximum Kurtosis Adaptive Filtering [17], try to find a deconvolution filter, by optimization of a sparsity cost function. The struggling point of these methods is that they are suboptimal and produce unstable results due to the shortcomings below. Homomorphic deconvolution is unable to correct the unknown phase distortions and tend to be highly sensitive to noise. MED and Maximum Kurtosis Adaptive Filtering are sensitive to noise and greatly influenced by the assumed length of the deconvolution filter, in addition to their inclination to cancel small reflectivity spikes.

Sparse seismic inversion methods have managed to produce stable reflectivity solutions, see e.g., [7, 11, 15, 19]. In order to increase the lateral resolution beyond the resolution that could be achieved by wavelet inverse filtering, some of these methods often depend on a-priori knowledge. Mostly, a starting model is built according to this prior information. Unfortunately, the starting model can be inaccurate due to lateral variations in the waveform interference path, in the propagation rate or in the earth
layers' impedances.

Nguyen et al. proposed to break down the seismic trace into reflectivity patterns using Matching Pursuit Decomposition (MPD). At each stage MPD identifies the dictionary atom that best correlates with the residual and adds a scalar multiple of that atom to the solution. The myopia limitation of the MPD method is most apparent when the dictionary is non-orthogonal. Basis Pursuit Decomposition (BPD) [1] is more advantageous. Originally developed as a compressive sensing technique, BPD utilizes an $l_1$ norm optimization and finds a single global solution in a computationally more efficient way. Moreover, it performs well even when dictionary elements are non-orthogonal.

Other important methods are Sparse Spike Inversion (SSI) [6] and Basis Pursuit Inversion (BPI) [16]. SSI and BPI recover each column of the reflectivity by solving a simple Basis Pursuit Denoising (BPDN) problem [2]. These methods perform very well under sufficiently high signal-to-noise ratio (SNR). Dosal [20] provide a lower bound on the minimum distance between spikes, that can be recovered by $l_1$ penalized deconvolution. However, one of the main disadvantages of these methods is that they ignore the correlation between adjacent traces. This correlation emerges from the natural assumption that the earth layers are horizontally structured.

Obviously, utilization of 1D restoration methods in the case of 2D seismic data is not optimal. Single-channel methods do not exploit the relations between spatially near traces. Thus, multichannel deconvolution is more robust. Zhang et al. [8] suggest to extend the BPI method to a multi-trace process with spatial regularization added in order to enhance lateral continuity and vertical resolution. Two variations of multichannel Bayesian deconvolution methods are suggested by Idier and Goussard [21]. Their approach is based on two Markov-Bernoulli-Gaussian reflectivity models (MBG I and II). The first model is a 2D extension of the 1D Bernoulli-
Gaussian (BG) representation. Mendel et al. [22, 23] use this 1D BG model in their Maximum-likelihood algorithm to estimate the reflectivity and the wavelet. The second model (MBG II) is more adapted to the physical and geometrical characteristics of the earth layers’ acoustic impedances. The deconvolution is performed by a suboptimal Maximum a-posteriori (MAP) estimator. Then, they use a method similar to the single most likely replacement (SMLR) algorithm [22] to iteratively recover each reflectivity column from the corresponding observed seismic trace and the preceding estimated reflectivity column. Kaaresen and Taxt [24] also propose a multichannel version of their single-channel blind deconvolution algorithm. The procedure repeats two stages: first, the wavelet is estimated by least-squares fit, and then the reflectivity is estimated by the iterated window maximization algorithm [25]. The algorithm produces better channel estimates since it updates more than one reflector in one trace at once, and also encourages lateral smoothness of the reflectors. However, these methods rely on a parametric model that leads to a nonconvex optimization problem. Usually, it is very difficult to find a global optimal solution to this kind of problems. The solution is normally found by searching for correct reflectivity spikes’ locations, within a limited number of potential reflectivity sequences (as in the SMLR algorithm mentioned above [23] ). This way, an optimal solution is achieved at the expense of heavy computational burden and an extended search.

Heimer, Cohen, and Vassiliou [26, 27] also propose a multichannel blind deconvolution. They integrate the algorithm of Kaaresen and Taxt [24] with dynamic programming [28, 29]. Valid reflectivity states and transitions between reflector arrangements of spatially-neighboring traces are defined. Then, the sequences of reflectors that are legally concatenated to other reflectors by valid transitions are extracted. Heimer et al.[30] also propose a method based on the MBRF modeling. The Viterbi algorithm [31] is applied
to the search of the most likely sequences of reflectors concatenated across the traces by legal transitions.

Ram, Cohen, and Raz [32] also propose two multichannel blind deconvolution algorithms for the restoration of 2D seismic data. Both algorithms are based on the Markov-Bernoulli-Gaussian I (MBG I) reflectivity model. In the first algorithm, each reflectivity channel is estimated from the corresponding observed seismic trace, while taking into consideration the estimate of the previous reflectivity channel. The procedure is carried out using a slightly modified maximum posterior mode (MPM) algorithm [33]. The second algorithm considers estimates of both the previous and following neighboring columns.

1.2 Research Overview

Our first goal is to develop a sparse multichannel seismic deconvolution algorithm. The algorithm iteratively attempts to find a sparse reflectivity solution, while considering the relations between spatially-neighboring traces. Multichannel Sparse Spike Inversion (MSSI) can be modified to take into account the spatial dependencies between reflectivity sequences for a user-dependent number of preceding and subsequent neighboring reflectivity columns. We apply the algorithm to synthetic and real data, and demonstrate improved results compared to those obtained by the single-channel deconvolution method, SSI. The performance of the algorithm is evaluated for different levels of SNRs.

Secondly, we consider a time-variant model of the seismic environment called earth Q-model. We present a novel method of inversion of each 1D seismic data to reveal the corresponding 1D reflectivity. The method can be applied to all columns of a 2D seismic data to reveal the earth 2D reflectivity. We also derive the theoretical bounds on the recovery error, and on the localization error achieved by using this method. We show that the recovered
spikes in the solution are located close to true spikes of the true reflectivity signal. Moreover, we prove that any redundant spike in the solution located far from the correct support will have small energy. The analytical results are demonstrated using synthetic and real data examples.

1.3 Organization of the Thesis

The thesis is organized as follows. In Chapter 2, we review the basic theory of the seismic deconvolution problem and describe two single-channel seismic deconvolution methods - SSI and BPI. In Chapter 3, we introduce our sparse multichannel seismic deconvolution algorithm and present simulation and real data results. In Chapter 4 we describe a time-variant model for the seismic problem. We present a recovery solution and derive analytical bounds on its produced error. We also demonstrate the performance of this method using synthetic and real data. Finally, in Chapter 5, we conclude and discuss further research.
1.4 List of Papers


Chapter 2

Sparse Seismic
Single-Channel
Deconvolution Methods

In this chapter, we compare two methods for seismic inversion - Sparse Spike Inversion (SSI) [6] and Basis Pursuit Inversion (BPI) [10]. Both methods utilize sparse inversion techniques. We employ a Least-Angle Regression (LARS) Least Absolute Shrinkage and Selection Operator (LASSO) solver for their implementation. Experimental results confirm that $L_1$ penalization in the LASSO optimization improves the performance in terms of recovering reflection coefficients.

In the following, we briefly present the models and the solution approaches, and refer the reader to [10] and [6] for further details. The remainder of the chapter is organized as follows. First, we review the basic theory of the two methods. Then, we describe our experiments with synthetic and real data. Lastly, we conclude and discuss further research.
2.1 Basic Theory

We can model \( s(t) \), a received 1D seismic signal (one column of the observation image) as

\[
s(t) = w(t) * r(t) + n(t)
\]  

(2.1)

where \( w(t) \) is the seismic wavelet, \( r(t) \) is the reflectivity series, and \( n(t) \) is the noise. The symbol \(*\) denotes one-dimensional linear convolution operation. This model assumes that the earth structure can be represented by planar horizontal layers of constant impedance, so that reflections are generated at the boundaries between adjacent layers. Each 1D seismic trace is a convolution of the seismic wavelet and the reflectivity pattern.

The objective is to find an estimate of the reflectivity \( r(t) \). The reflectivity is assumed to be sparse as only boundaries between adjacent layers may cause a reflection of the seismic wave.

As (2.1) implies, the seismic trace consists of a linear combination of \( w(t) \) and its time shifts, according to the non-zero reflectors in \( r(t) \). After time discretization, and an addition of random noise, (2.1) can be written in matrix-vector form as

\[
s_{N \times 1} = W_{N \times M} r_{M \times 1} + n_{N \times 1}
\]  

(2.2)

where \( W_{N \times M} \in \mathbb{R}^{N \times M} \), also known as the dictionary.

In the SSI method \( W_{N \times M} \) is the convolution matrix formed by the seismic discrete wavelet \( w(t) \). The inversion problem of finding \( r_{M \times 1} \) from the noisy measurement \( s_{N \times 1} \) is formulated as

\[
\min \| r_{M \times 1} \|_0 \text{ subject to } \| s_{N \times 1} - W_{N \times M} r_{M \times 1} \|_2^2 < \varepsilon .
\]  

(2.3)

After relaxing \( L_0 \) to \( L_1 \)-norm we obtain the constraint:

\[
\min_{r_{M \times 1}} \frac{1}{2} \| s_{N \times 1} - W_{N \times M} r_{M \times 1} \|_2^2 + \lambda \| r_{M \times 1} \|_1.
\]  

(2.4)
The problem formulated in the form of (2.4) is named Least Absolute Shrinkage and Selection Operator (LASSO) (see [34]). The use of L1 penalty in similar problems promotes sparsity of the solution $r_{M \times 1}$ (see [1], [2]).

On the other hand, the BPI method, proposed by Zhang et al. [16], utilizes dipole decomposition to represent the reflectivity series as a sum of even and odd impulse pairs multiplied by scalars. Each even and odd pair corresponds to the top and base reflector of a layer. Since the layer thickness is unknown, the dictionary comprises all possible thicknesses up to a maximum layer time-thickness.

Assuming the sample rate is $\Delta t$, each even wedge reflectivity can be written as

$$r_e(t, m, n, \Delta t) = \delta(t - m\Delta t) + \delta(t - m\Delta t - n\Delta t) \quad (2.5)$$

and each odd wedge reflectivity can be written as

$$r_o(t, m, n, \Delta t) = \delta(t - m\Delta t) - \delta(t - m\Delta t - n\Delta t). \quad (2.6)$$

Since any reflectivity can be written as

$$r(t) = \sum_{n=1}^{N} \sum_{m=1}^{M} a_{n,m} * r_e(t, m, n, \Delta t) + b_{n,m} * r_o(t, m, n, \Delta t) \quad (2.7)$$

the BPI dictionary consists of a convolution of the wavelet with the even wedge reflectivity and with the odd wedge reflectivity, and the objective is to calculate the coefficients $a_{n,m}$ and $b_{n,m}$.

### 2.2 Synthetic Examples

First, we evaluate the performances of the SSI and the BPI techniques with synthetic data. To test the methods, we used a 40 Hz Ricker wavelet and generated a reflectivity series with sample rate of 2 milliseconds.

To evaluate our result we used the normalized correlation coefficient:

$$\rho = \frac{\langle \hat{r}, r \rangle}{\|\hat{r}\|_2 \|r\|_2} \quad (2.8)$$
where $\hat{r}(t)$ is an estimate of the reflectivity series.

A small modification to the Least-Angle Regression (LARS) algorithm can solve the LASSO problem, as described in [35]. In our simulation, we use the SpaSM toolbox ([36]) to implement the LASSO algorithm for both the BPI and SSI, as proposed by Rozenberg et al. [12].

The regularization parameter $\lambda$ in (2.4) balances between the reflectivity sparsity and the noise. Increasing $\lambda$ decreases the sparsity of the solution, whereas decreasing $\lambda$ may cause noise amplification. Both SSI and BPI utilize $\lambda$ as a trade-off factor that controls the inversion output. However, one cannot compare the values between the methods. Practically, the value of $\lambda$ is data dependent and determined empirically.

Figures 2.1 and 2.2 show the SSI and BPI inversion results for a specific test. The non-zero reflection coefficients are uniformly distributed between $-0.2$ and $0.2$ (shown in Figure 2.1(b)). $D$ - the time difference between consecutive non-zero reflectivity coefficients - ranges between 10 milliseconds to 200 milliseconds, and the reflectivity sparsity $p$ was set to 0.06. Figures 2.1(a) and 2.2(a) show the synthetic reflectivity. Figures 2.1(b) and 2.2(b) show the synthetic traces, which are a convolution between the wavelet and the reflectivity. The signal-to-noise ratio (SNR) is 10 dB. Figures 2.1(c)-(f) and 2.2(c)-(f) show the results of each of the methods for different $\lambda$ values.

The series of synthetic tests that we have done during our research indicate that the optimal correlation can be achieved using different $\lambda$ values, depending on the channel characteristics: the number of reflectors, the layers’ thicknesses, the channel sparsity, and the SNR.

Figure 2.3 presents the correlation coefficient for different $\lambda$ values under the same conditions (SNR= 10 dB, and sampling rate of 2 milliseconds) for SSI and BPI methods.
2.3 Real Data Results

The SSI inversion was tested on a 2D seismic data set shown in Figure 2.4. The estimated reflectivity, and seismic data reconstructed as a convolution between the estimated reflectivity and a given wavelet, are shown in Figure 2.5. The obtained correlation between the original and reconstructed seismic data is $\rho_{s,\hat{s}} = 0.95$ for $\lambda_{opt} = 9.4 \times 10^{-3}$.

The results presented in this chapter reveal several interesting aspects of the sparse channel inversion methods. We used both synthetic and real data examples to evaluate the methods. Both methods yield reasonable estimates of the reflectivity under sufficiently high SNR. Our results indicate better performance of the SSI technique, although correct adjustments of the dictionary atoms selection can make the differences significantly smaller. We conclude that both methods could practically be used for seismic exploration and research purposes.

The choice of regularization parameter $\lambda$ is still an open problem. One needs to determine whether the resolution of the estimated reflectivity is real or a result of using a too small $\lambda$. In addition, in this study, we used an invariant known wavelet for simplicity. In practice, a time-depth varying wavelet could improve the results, taking into account wave propagation effects, such as attenuation and dispersion.
Figure 2.1: 1D synthetic tests of SSI. (a) True reflectivity. (b) Synthetic trace with 40 Hz Ricker wavelet and SNR= 10 dB. (c)-(f) SSI inversion results with varying $\lambda_{SSI}$. (c) $\lambda_{SSI} = 0.29$, (d) $\lambda_{SSI} = 0.11$, (e) $\lambda_{SSI} = 0.071$, (f) $\lambda_{SSI} = 0.025$. 
Figure 2.2: 1D synthetic tests of BPI. (a) True reflectivity. (b) Synthetic trace with 40 Hz Ricker wavelet and SNR= 10 dB. (c)-(f) BPI inversion results with varying $\lambda_{\text{BPI}}$. (c) $\lambda_{\text{BPI}} = 0.27$, (d) $\lambda_{\text{BPI}} = 0.087$, (e) $\lambda_{\text{BPI}} = 0.011$. 
Figure 2.3: (a) $\lambda$-correlation curve for SSI based on the synthetic data in Figure 2.1. (b) $\lambda$-correlation curve for BPI based on the synthetic data in Figure 2.2.
Figure 2.4: Seismic data.

Figure 2.5: (a) Estimated reflectivity matrix; (b) Reconstructed seismic data.
Chapter 3

Multichannel Sparse Spike Inversion

In the previous chapter we have described a single-channel deconvolution method - SSI [6]. We have observed that this 1D restoration method deconvolves each trace independently. Consequently, the spatial dependency between neighboring traces is ignored and the continuity of the layer boundaries of the reflectivity is not taken into consideration in the deconvolution process. Therefore, when applied to 2D simulated and real data, SSI produces discontinuous reflectivity estimates which contain gaps in the layer boundaries and scattered false detections. In this chapter we introduce Multichannel Sparse Spike Inversion (MSSI) as an iterative procedure, which deconvolves the seismic data and recovers the earth two-dimensional (2D) reflectivity image, while taking into consideration the relations between spatially-neighboring traces. MSSI can be modified to take into account the spatial dependencies between reflectivity sequences for a user-dependent number of preceding and subsequent neighboring reflectivity columns. We demonstrate the improved performance of the proposed algorithm and its robustness to noise, compared to competitive algorithms through simulations and real seismic data examples.
This chapter is organized as follows. In Section 3.1, we review the basic theory of the seismic deconvolution problem. In Section 3.2, we introduce our algorithm. In Section 3.3, we present simulation and real data results. Finally, we conclude and discuss further research.

3.1 Problem Formulation

We can model \( s(t) \), the received seismic 1D signal (the observation) as

\[
s(t) = w(t) \ast r(t) + n(t)
\]

where \( w(t) \) is the seismic wavelet, \( r(t) \) is the reflectivity series, and \( n(t) \) is the noise. The symbol \( \ast \) denotes one-dimensional linear convolution operation. This model assumes that the earth structure is stratified. It consists of planar horizontal layers of constant impedance and reflections are generated at impedance discontinuities, i.e., at the boundaries between adjacent layers. Each 1D seismic trace is a convolution of the seismic wavelet and the reflectivity pattern. All channels are excited by the same wavelet \( w(t) \). The support of the wavelet is finite and shorter than the channel’s length.

Note that a seismic image does not represent the actual image of the earth subsurface. Each reflection has been distorted during its propagation through the medium. The objective is to find an estimate of the reflectivity \( r(t) \). The reflectivity is assumed to be sparse as only boundaries between adjacent layers may cause a reflection of the seismic wave.

A seismic trace consists of a linear combination of \( w(t) \) and its time shifts, corresponding to the non-zero reflectors in \( r(t) \). The discrete convolution (3.1) can be written in matrix-vector form as

\[
s_{N \times 1} = W_{N \times M} r_{M \times 1} + n_{N \times 1}
\]

where \( W_{N \times M} \in \mathbb{R}^{N \times M} \), represents the dictionary.
In the SSI method $W_{N\times M}$ is the convolution matrix formed by the seismic discrete wavelet $w(t)$. The optimization problem for extracting $r_{M\times 1}$ from the seismic trace $s_{N\times 1}$ is formulated as

$$\min \|r_{M\times 1}\|_0 \text{ subject to } \|s_{N\times 1} - W_{N\times M}r_{M\times 1}\|_2^2 < \varepsilon.$$  \hspace{1cm} (3.3)

After relaxing $l_0$ to $l_1$-norm we obtain the problem:

$$\min_{r_{M\times 1}} \frac{1}{2} \|s_{N\times 1} - W_{N\times M}r_{M\times 1}\|_2^2 + \lambda \|r_{M\times 1}\|_1.$$  \hspace{1cm} (3.4)

The optimization problem as defined in (3.4) is called LASSO [34]. The $l_1$ penalty in similar problems is used in order to promote a sparse solution $r_{M\times 1}$ [1, 2].

On the other hand, the BPI method, proposed by Zhang and Castagna [10], applies “dipole decomposition”, i.e., each pair of neighboring impulses in the reflectivity sequence is represented as a linear combination of even and odd impulse pairs. Each even and odd pair corresponds to the top and base reflector of a layer. Since the layer thickness is unknown, the dictionary comprises all possible thicknesses up to a maximum layer time-thickness.

### 3.2 Multichannel Sparse Spike Inversion (MSSI)

In this section, we estimate the reflectivity while taking into account spatial dependencies between neighboring reflectivity sequences.

Assume $J$ adjacent columns, and for simplicity assume that $J$ is odd. Denote the current column, which we wish to estimate, by $r_i$, and a previous or subsequent column by $r_{i+k}$ where $-\frac{J-1}{2} \leq k \leq \frac{J-1}{2}$. We estimate each reflectivity column from the corresponding observed seismic trace $s_i$, taking into consideration the current estimate of $\frac{J-1}{2}$ preceding reflectivity columns, and of $\frac{J-1}{2}$ subsequent reflectivity columns. Out of $J$ estimated columns only the middle reflectivity column is kept. The estimates of the other $J - 1$ columns are discarded. If we wish to use only the subsequent column (i.e.
\( J = 2 \), we keep the first reflectivity column, and discard the subsequent column (in this case \(-\frac{J}{2} < k \leq \frac{J}{2}\)).

We formulate the problem as a minimization of the following cost function:

\[
\min_{r_i, \ldots, r_{i \pm \frac{J-1}{2}}} \sum_{k=-\frac{J-1}{2}}^{\frac{J-1}{2}} \frac{1}{2} \| s_{i+k} - W r_{i+k} \|_2^2 + \lambda_0 \sum_{k=-\frac{J-1}{2}}^{\frac{J-1}{2}} \| r_{i+k} \|_1 \\
+ \sum_{k=-\frac{J-1}{2}}^{\frac{J-1}{2}} \lambda_k \| r_i - H_k r_{i+k} \|_2^2. 
\tag{3.5}
\]

Where \( r_i, \ldots, r_{i \pm \frac{J-1}{2}} \) are \( J \) reflectivity columns, and \( s_i, \ldots, s_{i \pm \frac{J-1}{2}} \) are \( J \) corresponding seismic traces. \( W \) is the convolution matrix formed by the seismic discrete wavelet \( w \), assumed to be known. The tradeoff parameter \( \lambda_0 \) controls the balance between the reflectivity sparseness and the least-squares error. The tradeoff parameters \( \lambda_k \) promote smoothness of the reflectivity in the horizontal direction. \( H_k \) is the convolution matrix of a Low-pass filter. We can choose \( H_k \) to be the convolution matrix of a Hamming window or an Averaging filter. Hence, \( H_k \) controls the smoothness as it reduces the penalty for layer boundaries whose orientation is diagonally descending, horizontal, and diagonally ascending. The size of the smoothing filter controls the desired smoothness of the resultant reflectivity image. This way, the minimization is performed by taking into account the distances between each reflectivity column and the preceding and subsequent reflectivity columns.

Without loss of generality we assume that each reflectivity column has unit variance (i.e., \( r_i^T r_i = 1 \)). Accordingly, we can express the solution as

\[
(\hat{r}_i, \hat{r}_{i \pm \frac{J-1}{2}}) = \min_{r_i, r_{i \pm \frac{J-1}{2}}} J(r_i, r_{i \pm \frac{J-1}{2}}) 
\tag{3.6}
\]
s.t. $\mathbf{r}_i^T \mathbf{r}_i = \ldots = \mathbf{r}_{i+\frac{J-1}{2}}^T \mathbf{r}_{i+\frac{J-1}{2}} = 1$

where

$$J(\mathbf{r}_i, \mathbf{r}_{i+1}, \ldots, \mathbf{r}_{i+\frac{J-1}{2}}) = \sum_{k=-\frac{J-1}{2}}^{\frac{J-1}{2}} \frac{1}{2} \| \mathbf{s}_{i+k} - W \mathbf{r}_{i+k} \|_2^2 + \lambda_0 \sum_{k=-\frac{J-1}{2}}^{\frac{J-1}{2}} \mathcal{R}_\epsilon(\mathbf{r}_{i+k}) \quad (3.7)$$

$$+ \sum_{k=-\frac{J-1}{2}, k\neq 0}^{\frac{J-1}{2}} \frac{1}{2} \lambda_k \| \mathbf{r}_i - \mathbf{H}_k \mathbf{r}_{i+k} \|_2^2$$

and

$$\mathcal{R}_\epsilon(\mathbf{r}) = \sum_j (\sqrt{r_j^2 + \epsilon^2} - \epsilon). \quad (3.8)$$

For small $\epsilon$, such as $\epsilon = 0.01$ [37], the regularization parameter $\mathcal{R}_\epsilon(\mathbf{r})$ is a smoothed $\ell_1$ norm approximation that promotes sparsity of the solution (also called hybrid $\ell_1 - \ell_2$ or hyperbolic penalty [38]). $\mathcal{R}_\epsilon(\mathbf{r})$ is also used for seismic blind deconvolution in [37]. The use of the hybrid $\ell_1 - \ell_2$ norm, which is differentiable, rather than the $\ell_1$ norm, enables the use of simple optimization techniques such as steepest descent method.

To solve the constrained optimization problem above, we wish to minimize the following cost function:

$$\mathcal{L}(\mathbf{r}_i, \mathbf{r}_{i+1}, \ldots, \mathbf{r}_{i+\frac{J-1}{2}}) = J(\mathbf{r}_i, \mathbf{r}_{i+1}, \ldots, \mathbf{r}_{i+\frac{J-1}{2}}) - \sum_{k=-\frac{J-1}{2}}^{\frac{J-1}{2}} \eta_k (\mathbf{r}_{i+k}^T \mathbf{r}_{i+k} - 1) \quad (3.9)$$

with Lagrange multipliers given by the scalars $\eta_k$. The minimization must satisfy

$$\frac{\partial \mathcal{L}}{\partial \mathbf{r}_{i+k}} = \mathbf{g}_{i+k} - \eta_{i+k} \mathbf{r}_{i+k} = 0, \quad -\frac{J-1}{2} \leq k \leq -\frac{J-1}{2} \quad (3.10)$$

where $\mathbf{g}_{i+k} = \frac{\partial J}{\partial \mathbf{r}_{i+k}}$.

Multiplying (3.10) by $\mathbf{r}_{i+k}^T$ and using the constraint $\mathbf{r}_{i+k}^T \mathbf{r}_{i+k} = 1$ yields

$$\eta_{i+k} = \mathbf{r}_{i+k}^T \mathbf{g}_{i+k}. \quad (3.11)$$
Then, the projection of the gradient on the unit sphere can be expressed via

$$\frac{\partial L}{\partial r_{i+k}} = g_{i+k} - r_{i+k}^T g_{i+k} r_{i+k}.$$  \hspace{1cm} (3.12)

The classical update rule of steepest descent algorithm is given by

$$r_{i+k,l+1} = r_{i+k,l} - \mu_l h_{i+k,l}$$

with a normalized gradient

$$h_{i+k,l} = \frac{\partial L}{\partial r_{i+k}} / | \frac{\partial L}{\partial r_{i+k}} |.$$  

where \( \mu_l \) is the adaptive step size and \( l \) indicates an iteration index. Each step in the direction of the gradient could divert \( r_{i+k,l+1} \) off the unit sphere. Therefore, we normalize \( r_{i+k,l+1} \) to the unit sphere at each iteration.

As in [37], it should be mentioned that we must initialize the steepest-descent algorithm by a solution that is close to the final reflectivity. Since the data is structurally close to the true sparse reflectivity, we can use it as an initial solution. Practically, this choice is advantageous and resolves into a sparse estimate of the reflectivity.

### 3.3 Experimental results

The proposed algorithm is evaluated using synthetic and real data. It demonstrates better results than those obtained by a single-channel deconvolution method.

**Synthetic Data**

First, we tried to evaluate the performance of the algorithm on a 2D reflectivity section of size 76 × 98. The algorithm was implemented for \( J = 2 \) and for \( J = 3 \).

For \( J = 2 \) the above optimization problem reduces to:
\[
\begin{align*}
\text{min}_{r_i, r_{i+1}} & \quad \frac{1}{2} \|s_i - Wr_i\|^2_2 + \frac{1}{2} \|s_{i+1} - Wr_{i+1}\|^2_2 + \lambda_0 \left( \|r_i\|_1 + \|r_{i+1}\|_1 \right) + \lambda_1 \frac{1}{2} \|r_i - H_1 r_{i+1}\|^2_2 \\
\text{(3.13)}
\end{align*}
\]

For \(J = 3\) the above optimization problem is:

\[
\begin{align*}
\text{min}_{r_i, r_{i-1}, r_{i+1}} & \quad \frac{1}{2} \|s_{i-1} - Wr_{i-1}\|^2_2 + \frac{1}{2} \|s_i - Wr_i\|^2_2 + \frac{1}{2} \|s_{i+1} - Wr_{i+1}\|^2_2 \\
& \quad + \lambda_0 \left( \|r_{i-1}\|_1 + \|r_i\|_1 + \|r_{i+1}\|_1 \right) \\
& \quad + \lambda_{-1} \frac{1}{2} \|r_i - H_1 r_{i-1}\|^2_2 + \lambda_1 \frac{1}{2} \|r_i - H_{-1} r_{i+1}\|^2_2 \quad \text{(3.14)}
\end{align*}
\]

These schemes were tested for different values of \(\lambda_0, \lambda_1\) and \(\lambda_{-1}\), with SNR = 10 dB and 5 dB. As was mentioned before, the tradeoff parameter \(\lambda_0\) balances the reflectivity sparseness and the minimization of the residual term. Increasing \(\lambda_0\) decreases the sparsity of the solution, whereas decreasing \(\lambda_0\) may lead to noise amplification. The tradeoff parameters \(\lambda_{\pm 1}\) promote smoothness of the reflectivity in the horizontal direction.

We will hereafter refer to the proposed algorithm above implementations as MSSI-2 and MSSI-3, which stands for Multichannel Sparse Spike Inversion implemented for \(J = 2\) and \(J = 3\) respectively.

In MSSI-2, the minimization is performed by taking into account the distance between each reflectivity column and the subsequent reflectivity column. In each step, we estimate two adjacent columns simultaneously. Even though two reflectivity columns estimates were obtained, we keep only the current reflectivity column estimate. The estimate of the subsequent column is discarded, since this column will be estimated with its subsequent column in the next step. In MSSI-3, the minimization is performed by taking into account the distances between each reflectivity column and both the preceding and subsequent reflectivity columns. In each step, we estimate three adjacent columns simultaneously. Out of the three obtained estimates, only the middle reflectivity column is kept. The estimates of the preceding and the subsequent columns are discarded.
With different experiments, we concluded that the best results are achieved when \( H_{\pm 1} \) is a convolution matrix of a 3 taps averaging filter,

\[
H_{\pm 1} = \frac{1}{3} \begin{pmatrix}
1 & 1 & 0 & \ldots & 0 \\
1 & 1 & 1 & \ldots & 0 \\
0 & 1 & 1 & \ldots & 0 \\
0 & 0 & 1 & \ldots & 0 \\
\vdots & \vdots & \vdots & \ddots & 1 \\
0 & 0 & 0 & \ldots & 1
\end{pmatrix}_{L_r \times L_r},
\]

where \( L_r \) is the length of a reflectivity column (in this example \( L_r = 76 \)). Hence, spikes of two neighboring traces are presumed close by less than 3 samples. Though this hypothesis seems very restrictive in the case of real data, we observe that the results are better when \( H_{\pm 1} \) is a convolution matrix of a short averaging filter, not longer than 3 taps, since this choice balances between the ability to detect layers’ discontinuities and still create a smooth reflectivity image. This is a great advantage compared to other existing methods. Choosing a longer filter usually causes over-smoothing of the recovered reflectivity and blurring of natural breaks in the earth structure.

To analyze the stability of the method under different levels of noise, we generated 20 different realizations of 2D reflectivities of size 76 \times 98, one example is shown in Fig. 3.1(a). We then convolved it with a 27 samples long Ricker wavelet and added white Gaussian noise with SNRs of 10 dB and 5 dB. Two of the realizations with SNRs of 10 dB and 5 dB are shown in Fig. 3.1(b) and Fig. 3.1(c), respectively. The seismic wavelet is shown in Fig. 3.1(d).

As a figure of merit we used the correlation coefficient defined as

\[
\rho_{rr} = \frac{\hat{f}^T r}{\|\hat{f}\|_2 \|r\|_2}
\]

(3.15)
where $\hat{\mathbf{r}}$ and $\mathbf{r}$ are column-stack vectors of the estimated reflectivity and the true generated reflectivity, respectively. The algorithm finds unscaled versions of the reflectivity, but it is clear that this does not affect the computation of $\rho_{\mathbf{rr}}$.

We compare our results to a single channel deconvolution (SSI) and to the multichannel deconvolution algorithm described in [32] (MC-II). The estimated reflectivities, obtained by SSI, MC-II, and by MSSI, for the seismic data with SNR of 10 dB, are shown in Fig. 3.2. The SSI method was implemented by simply assigning $\lambda_{\pm 1}$ to zero, using the best estimated $\lambda_0$ for SSI. For this example, the correlation coefficients between the original reflectivity and the estimated reflectivity, with our method, are $\rho = 0.88$ and $\rho = 0.9$ for $J = 2$ and $J = 3$, respectively, whereas the correlation coefficient achieved by single-channel deconvolution is $\rho = 0.78$, and $\rho = 0.86$ for MC-II. The best results in terms of correlation coefficients were achieved with $\lambda_0 = 3.1$ and $\lambda_1 = 0.9$ for the two channel implementation, with $\lambda_0 = 2.8, \lambda_{-1} = 1$ and $\lambda_1 = 0.7$ for the three-channel implementation, and with $\lambda_0 = 2.8$ for SSI. Practically, the values of $\lambda_0$ and $\lambda_{\pm 1}$ are data dependent and determined empirically. The best result is not necessarily achieved by setting $\lambda_1 = \lambda_{-1}$.

The average correlation coefficients between the original reflectivity and the estimated reflectivity and standard deviations (in brackets), for SNR of 10 dB, with our method, are $\rho = 0.87(0.022)$ and $\rho = 0.9(0.018)$ for $J = 2$ and $J = 3$, respectively, whereas the correlation coefficient achieved by single-channel deconvolution is $\rho = 0.78(0.027)$, and $\rho = 0.83(0.093)$ for MC-II.

Another example is shown in Figure 3.3. We added white Gaussian noise of SNR $= 5$ dB. The estimated reflectivities, obtained by single channel deconvolution (SSI), by MC-II and by MSSI, for the seismic data with SNR of 5 dB , are shown in Fig. 3.3. The best results in terms of correlation...
coefficients were achieved with $\lambda_0 = 3.9$ and $\lambda_1 = 2.3$ for the two channel implementation, with $\lambda_0 = 2.6, \lambda_{-1} = 1$ and $\lambda_1 = 0.6$ for the three-channel implementation, and with $\lambda_0 = 2.9$ for SSI. The correlation coefficients between the original reflectivity and the estimated reflectivity with our method is $\rho = 0.77$, and $\rho = 0.82$ for $J = 2$ and $J = 3$ respectively. Whereas the correlation coefficient achieved by single-channel deconvolution is only $\rho = 0.66$, and for MC-II we have only $\rho = 0.69$.

The average correlation coefficients between the original reflectivity and the estimated reflectivity and standard deviations (in brackets), for SNR of 5 dB, with our method, are $\rho = 0.80(0.039)$ and $\rho = 0.78(0.054)$ for $J = 2$ and $J = 3$, respectively. Whereas the correlation coefficient achieved by single-channel deconvolution is $\rho = 0.66(0.040)$, and $\rho = 0.64(0.167)$ for MC-II.

The series of synthetic tests that we have performed during our research indicate that the optimal correlation can be achieved using different $\lambda_0$ and $\lambda_{\pm 1}$ values, depending on the channel characteristics: the number of reflectors, the layers’ thicknesses (distances between reflectors), the channel sparsity, and the SNR. It is recommended that $\lambda_1$ and $\lambda_{-1}$ values will not be too large so as to avoid over-smoothing of the estimated reflectivity. The parameters can be chosen by inspecting the correlation coefficient of a few columns.

Fig. 3.4 presents the correlation coefficient values as a function of $\lambda_0$ and $\lambda_1$, for 10 columns of the seismic data with SNR of 5 dB, depicted in Fig. 3.1(c). As can be seen, there is an area of values that gives the best results. This implies that the user does not have to know the exact value of the regularization parameters in order to get a good recovery. The correlation coefficients for $\lambda_1 = 0$, which represent the single-channel scores are significantly smaller than the values achieved by a non-zero value of $\lambda_1$. This implies that the MSSI outperforms the single-channel method (SSI).
The average processing times of a data set of size $76 \times 98$ on Intel(R)Core(TM)i5-4430 CPU @3GHz, by Matlab implementations of the single-channel and the proposed algorithms - MSSI-2 and MSSI-3 are 1.18, 1.41 and 1.57 minutes, respectively.

Visual comparison between the above results confirms that the multi-channel algorithm outperforms the single-channel algorithm. For both SNR levels the estimates of the MSSI are more continuous. In addition, false detections are less common in MSSI’s estimates. Generally, MSSI’s recovered reflectivities are closer to the true reflectivity than the single-channel deconvolution results. MC-II performs well in high SNR environments, but when the SNR is low it appears to have many false detections. MSSI, on the other hand, tends to diminish small spikes. It can also be observed that the values of the correlation coefficients for MSSI are higher. This implies that both MSSI-2 and MSSI-3 produce better results than the single-channel algorithm. In addition, as one would expect, for both SNR levels, MSSI-3 outperforms MSSI-2. Naturally, the improvement is getting smaller as the SNR increases, meaning that all algorithms perform better when the noise level is lower.

**Real Data**

We applied the proposed deconvolution scheme, to real seismic data from a small land 3D survey in North America (courtesy of GeoEnergy Inc., TX) of size $350 \times 200$, shown in Fig. 3.5(a). The assumed wavelet is shown in Fig. 3.5(b). The reflectivity sections obtained by single-channel deconvolution, by MC-II, by MSSI-2 and by MSSI-3 are shown in Fig.3.6(a),(b),(c) and (d), respectively. The seismic data reconstructed as a convolution between the estimated reflectivity and a given wavelet, are shown in Fig. 3.6(e),(f),(g) and (h). Visually comparing these reflectivity sections, it can be seen that the layer boundaries in the estimates obtained by MSSI are more continuous.
and smooth than the layer boundaries in the single-channel deconvolution estimates. Moreover, MSSI also detects parts of the layers that the single-channel deconvolution misses. It can also be seen that the reconstructed seismic data obtained by MSSI is more accurate than the one obtained by SSI. Since the ground truth is unknown, to assess the performance of the methods, we calculate the correlation coefficient between the reconstructed data to a noise-free seismic data. The obtained correlation between the original and reconstructed seismic data for MSSI is $\rho_{s,\hat{s}} = 0.9$ when $\lambda_0 = 9$ and $\lambda_1 = 30$ for MSSI-2, and $\rho_{s,\hat{s}} = 0.91$ when $\lambda_0 = 9$ and $\lambda_{\pm 1} = 28$ for MSSI-3. Whereas for SSI we get $\rho_{s,\hat{s}} = 0.89$ when $\lambda_0 = 5$, and for MC-II we have $\rho_{s,\hat{s}} = 0.76$. The parameters for all methods were chosen to best fit the observed data using the correlation of a few columns. The estimates produced by MSSI-2 and MSSI-3 are quite close, though the latter manages to recover a slightly more continuous image.

As mentioned before, experimental results show that the best results are achieved when $H_{\pm 1}$ is a convolution matrix of a 3 taps averaging filter, which means that we assume that spikes of two neighboring traces are close by less than 3 samples. This hypothesis might seem very restrictive in the case of real data. However, $H_{\pm 1}$ as a convolution matrix of a 3 taps only averaging filter outperforms other filter choices, for the reason that this choice balances between the ability to detect layers’ discontinuities and more complex structure and at the same time also to create a smooth reflectivity image. This is a great advantage compared to other existing methods. Choosing a longer filter causes over-smoothing of the recovered reflectivity and blurring of natural breaks in the earth structure. Choosing the lateral derivative instead of the third term as defined in (3.7) would encourage horizontal lines ignoring the subsurface curves structure.
We have presented a multichannel deconvolution algorithm in seismic applications. The algorithm both promotes sparsity of the solution and also takes into consideration the spatial dependency between neighboring traces in the deconvolution process. We have demonstrated that our deconvolution results are visually superior, compared to a single-channel deconvolution algorithm, for synthetic and real data, under sufficiently high SNR. Our second implementation (MSSI-3) performs better, on both synthetic and real data. The reason for that is that MSSI-3 takes into account more information from neighboring traces in the deconvolution process of each trace, compared to the first implementation (MSSI-2) that uses information from only one neighboring trace. The improved performance of the proposed algorithm compared to the single-channel algorithm was also apparent in qualitative assessment. It also shows that the second implementation’s results are more accurate.

The choice of regularization parameters is still an open problem. The use of a too small $\lambda_0$ could result in an increased resolution of the estimated reflectivity which is not necessarily real. In addition, one needs to find the correct balance between all regularization parameters. It should also be mentioned that in this study we used a time-spatial-invariant known wavelet for simplicity. In practice, a time and spatial varying wavelet could produce better results, taking into account wave propagation effects, such as attenuation and dispersion.
Figure 3.1: Synthetic reflectivity, wavelet and data sets: (a) Synthetic 2D reflectivity section; (b) 2D seismic data (SNR = 10 dB); (c) 2D seismic data (SNR = 5 dB); (d) Wavelet.
Figure 3.2: Synthetic 2D data deconvolution results: (a) Single-channel deconvolution results for SNR = 10 dB; (b) MC-II deconvolution results for SNR = 10 dB; (c) MSSI-2 results for SNR = 10 dB; (d) MSSI-3 results for SNR = 10 dB.
Figure 3.3: Synthetic 2D data deconvolution results: (a) Single-channel deconvolution results for SNR = 5 dB; (b) MC-II deconvolution results for SNR = 5 dB; (c) MSSI-2 results for SNR = 5 dB; (d) MSSI-3 results for SNR = 5 dB.
Figure 3.4: Correlation coefficient vs. deconvolution parameters $\lambda_1$ and $\lambda_0$ for synthetic 2D data deconvolution (SNR = 5 dB).
Figure 3.5: Real data and assumed wavelet: (a) Real seismic data (SNR = 5 dB); (b) wavelet.
Figure 3.6: Real data deconvolution results: (a) Single-channel estimated reflectivity; (b) MC-II estimated reflectivity; (c) MSSI-2 estimated reflectivity; (d) MSSI-3 estimated reflectivity; (e) Single-channel reconstructed data; (f) MC-II reconstructed data; (g) MSSI-2 reconstructed data; (h) MSSI-3 reconstructed data.
Chapter 4

Seismic Recovery Based on Earth Q Model

4.1 Introduction

The problem of decomposing a signal into its building blocks (atoms) [1] is very common in many fields in signal processing, such as: image processing [2], compressed sensing [3], radar [4], ultrasound imaging [5], seismology [6–8] and more. In seismic inversion, a short duration pulse (the wavelet) is transmitted from the earth surface. The reflected pulses from the ground are received by a sensor array and processed into a seismic image [9]. Since reflections are generated at discontinuities in the medium impedance, each seismic trace (a column in the seismic two-dimensional (2D) image) can be modeled as a weighted superposition of pulses further degraded by additive noise. Our task is to recover the earth layers structure (the reflectivity) hidden in the observed seismic image.

Previous works tried to solve the seismic inversion problem by separating the seismic 2D image into independent vertical one-dimensional (1D) deconvolution problems. The wavelet is modeled as a 1D time-invariant signal in both horizontal and vertical directions. Each reflectivity channel (column)
appears in the vertical direction as a sparse spike train where each spike is a reflector that corresponds to a boundary between two layers (two different acoustic impedances) in the ground. Then, each reflectivity channel is estimated from the corresponding seismic trace observation apart from the other channels [6–9, 11, 13–15, 17]. Utilization of sparse seismic inversion methods - based on $\ell_1$ minimization problem solving - can yield stable reflectivity solutions [7, 11, 15, 19, 38]. These $\ell_1$-type methods and their resolution limits are studied thoroughly in signal processing and statistics research [5, 34, 35, 39–45].

Multichannel deconvolution methods [21–24, 26, 27, 30, 32, 46] take into consideration the horizontal continuity of the seismic reflectivity. Heimer et al. [30] propose a method based on Markov Bernoulli random field (MBRF) modeling. The Viterbi algorithm [31] is applied to the search of the most likely sequences of reflectors concatenated across the traces by legal transitions. Ram et al. [32] also propose two multichannel blind deconvolution algorithms for the restoration of 2D seismic data. These algorithms are based on the Markov-Bernoulli-Gaussian (MBG) reflectivity model. Each reflectivity channel is estimated from the corresponding observed seismic trace, taking into account the estimate of the previous reflectivity channel or both estimates of the previous and following neighboring columns. The procedure is carried out using a slightly modified maximum posterior mode (MPM) algorithm [33].

Although the typical seismic wavelet is time-variant, many inversion methods depend on a model which does not take into consideration time-depth variations in the waveform. However, the wave absorption effects are not always negligible as the conventional assumption claim. Seismic inverse Q-filtering [47–50] aims to compensate for the velocity dispersion and energy absorption which causes phase and amplitude distortions of the propagating and reflected acoustic waves. The process of inverse Q filtering consists of
amplitude compensation and phase correction which enhance the resolution and increase the signal-to-noise ratio (SNR). Yet, this process is generally computationally expensive and sometimes even impractical.

Nonstationary deconvolution methods aim to deconvolve the seismic data and also compensate for energy absorption, without knowing $Q$. Margrave et al. [51] developed the Gabor deconvolution algorithm. Chai et al. [52] also propose a method called nonstationary sparse reflectivity inversion (NSRI) to retrieve the reflectivity signal from nonstationary data without inverse $Q$ filtering. Li et al. [53] propose a nonstationary deconvolution algorithm based on spectral modeling [54] and variable-step-sampling (VSS) hyperbolic smoothing.

We propose a novel robust algorithm for recovery of the underlying reflectivity signal from the seismic data without a pre-processing stage of inverse $Q$ filtering. We prove that the solution of a convex optimization problem, which takes into consideration a time-variant signal model, results in a stable recovery. In addition we answer the following questions: To what accuracy can we recover each reflectivity spike? How does this accuracy depend on the noise level, the amplitude of the spike, the medium $Q$ constant and the wavelet’s shape? We prove that the recovery error is proportional to the noise level. We also show how the error is affected by degradation. The algorithm is applied to synthetic and real seismic data. Our experiments affirm the theoretical results and demonstrate that the suggested method reveals reflectors amplitudes and locations with high precision.

This chapter is organized as follows. In Section 4.1, we review the basic theory of earth $Q$ model and the seismic inversion problem. In Section 4.2, we present the main theoretical results. Section 4.3 presents numerical experiments and real data results. Finally, we conclude and discuss further research.
4.2 Signal Model

4.2.1 Reflectivity model

We assume the earth structure is stratified, so that reflections are generated at the boundaries between different impedance layers. Therefore, each 1D channel (column) in the unknown 2D reflectivity signal can be formulated as a sparse spike train

\[ x(t) = \sum_m c_m \delta(t - t_m), \]  \hspace{1cm} (4.1)

where \( \delta(t) \) denotes the Dirac delta function and \( \sum_m |c_m| < \infty \). The set of delays \( T = \{t_m\} \) and the real amplitudes \( \{c_m\} \) are unknown.

In the discrete setting, assuming the sampling interval is \( 1/N \) for a given integer sampling rate \( N \), and that the set of delays \( T = \{t_m\} \) lie on the grid \( k/N, k \in \mathbb{Z} \), i.e., \( t_m = k_m/N \) where \( k_m \in \mathbb{Z} \)

\[ x[k] = \sum_m c_m \delta[k - k_m], \]  \hspace{1cm} (4.2)

where \( \delta[k] \) denotes the Kronecker delta function (see [55]).

We consider a seismic discrete trace of the form

\[ y[k] = y(k/N) = \sum_m c_m g_{\sigma,m}(\frac{t - t_m}{N}) = \sum_m c_m g_{\sigma,m}[k - k_m], \]  \hspace{1cm} (4.3)

where \( \{g_{\sigma,m}\} \) is a known set of kernels (pulses) for a possible set of time delays \( T = \{t_m\} \), and a known scaling parameter \( \sigma > 0 \). In subsection 2.3 we discuss specific requirements for \( \{g_{\sigma,m}\} \).

A time-invariant model assumes for simplicity that all kernels are identical, i.e., \( g_{\sigma,m}(t) = g(\frac{t}{\sigma}) \forall m \) [5, 45]. Hence, the model can be represented as a convolution model. However, the shape and energy of each reflected pulse highly depends on its corresponding reflector’s depth in the ground. Therefore, an accurate model should take into consideration a set of kernels \( \{g_{\sigma,m}\} \) which consists of different pulses.
In noisy environments we consider a discrete seismic trace of the form

\[ y[k] = \sum_m c_m g_{\sigma,m}[k - k_m] + n[k], \quad |n|_1 \leq \delta, \quad (4.4) \]

where \( n[k] \) is additive noise with \( |n|_1 = \sum_k |n[k]| \leq \delta \). Our objective is to estimate the true support \( K = \{k_m\} \) and the spikes’ amplitudes \( \{c_m\} \) from the observed seismic trace \( y[k] \).

### 4.2.2 Earth Q model

We assume a source waveform \( s(t) \) defined as the real-valued Ricker wavelet.

\[ s(t) = \left(1 - \frac{1}{2} \omega_0^2 t^2\right) \exp\left(-\frac{1}{4} \omega_0^2 t^2\right), \quad (4.5) \]

where \( \omega_0 \) is the most energetic (dominant) radial frequency [56]. We define the scaling parameter as \( \sigma = \omega_0^{-1} \). Wang [57] showed that given a travel time \( t_m \), the reflected wave can be modeled as

\[ u(t) = \text{Re}\left\{\frac{1}{\pi} \int_0^\infty S(\omega) \exp[j(\omega t - \kappa r(\omega))] d\omega\right\}, \quad (4.6) \]

where \( S(\omega) \) is the Fourier transform of the source waveform \( s(t) \),

\[ \kappa r(\omega) \triangleq \left(1 - \frac{j}{2Q}\right) \left|\frac{\omega}{\omega_0}\right|^{-\gamma} \omega t_m, \quad (4.7) \]

\[ \gamma \triangleq \frac{2}{\pi} \tan^{-1}\left(\frac{1}{2Q}\right) \approx \frac{1}{\pi Q}, \quad (4.8) \]

and \( Q \) is the medium quality factor, which is assumed to be frequency independent [47]. Kjartansson defined \( Q \) as the portion of energy lost during each cycle or wavelength.

Therefore, the expression of the earth Q filter consists of two exponential operators that express the phase effect (caused by velocity dispersion) and the amplitude effect (caused by energy absorption)

\[ U(t - t_m, \omega) = U(t, \omega) \exp\left(-j \left|\frac{\omega}{\omega_0}\right|^{-\gamma} \omega t_m\right) \exp\left(-\left|\frac{\omega}{\omega_0}\right|^{-\gamma} \frac{\omega t_m}{2Q}\right), \quad (4.9) \]
Summing these plane waves we get the time-domain seismic signal

\[ u(t - t_m) = \frac{1}{2\pi} \int U(t - t_m, \omega) d\omega. \]  

(4.10)

We can now define the known set of kernels (pulses) \( \{g_{\sigma,m}\} \) for the seismic setting

\[ g_{\sigma,m}(t - t_m) = u(t - t_m)|_{\sigma=\omega_0}. \]  

(4.11)

### 4.2.3 Admissible Kernels and Separation Constant

To be able to quantify the waves decay and concavity we recall two definitions from previous works [5, 45]:

**Definition 2.1** A kernel \( g \) is **admissible** if it has the following properties:

1. \( g \in \mathbb{R} \) is real and even.

2. **Global Property:** There exist constants \( C_l > 0, l = 0, 1, 2, 3 \), such that

   \[ |g^{(l)}(t)| \leq \frac{C_l}{1 + t^2}, \]  

   where \( g^{(l)}(t) \) denotes the \( l^{th} \) derivative of \( g \).

3. **Local Property:** There exist constants \( \varepsilon, \beta > 0 \) such that

   (a) \( g(t) > 0 \) for all \( |t| \leq \varepsilon \) and \( g(\varepsilon) > g(t) \) for all \( |t| \geq \varepsilon \).

   (b) \( g^{(2)}(t) < -\beta \) for all \( |t| \leq \varepsilon \).

In other words, the kernel and its first three derivatives are decaying fast enough, and the kernel is concave near its midpoint.

**Definition 2.2** A set of points \( K \subset \mathbb{Z} \) is said to satisfy the **minimal separation condition** for a kernel dependent \( \nu > 0 \), a given scaling \( \sigma > 0 \) and a sampling spacing \( 1/N > 0 \) if

\[ \min_{k_i, k_j \in K, i \neq j} |k_i - k_j| \geq N \nu \sigma. \]

where \( \nu \sigma \) is the smallest time interval between two reflectors with which we can still recover two distinct spikes, and \( \nu \) is called the separation constant.
Figure 4.1 presents an example of the attenuating wavelets $g_{\sigma,m}(t)$ and their derivatives, $g_{\sigma,m}^{(1)}(t)$ and $g_{\sigma,m}^{(2)}(t)$ for $Q = 125$ and $t_m = 100, 250, 400, ..., 1900$ms (increment of 150ms). $\omega_0 = 100\pi$ (50Hz). The pulses and their derivatives are moved to the origin so that it can be seen that there is a common value of $\varepsilon$ and $\beta$. Meaning that, for a sequence of kernels $g_{\sigma,m}(t)$ as described in (4.11), there exist two possible parameters $(\varepsilon_m, \beta_m)$ that determine the concavity of the reflected wave $g_{\sigma,m}(t)$, as defined in Definition 2.1, such that there are two common constants $\varepsilon, \beta > 0$ for all reflected waves. In other words

$$\varepsilon_m = \varepsilon \quad \forall m \quad (4.12)$$

and

$$\beta_m = \beta \quad \forall m. \quad (4.13)$$

The reflected waves $g_{\sigma,m}(t)$ are not symmetric, but remain flat at the origin, i.e., $g_{\sigma,m}^{(1)}(t) \approx 0$. So, it can be said that each of the reflected waves $g_{\sigma,m}(t)$ is approximately an admissible kernel, and all of these waves share two common parameters $\varepsilon, \beta > 0$.

We would make one more assumption: $g_{\sigma,m}(\varepsilon) > |g_{\sigma,m}(t)|$ for all $|t| \geq \varepsilon$. Meaning that for $|t| \geq \varepsilon$ the absolute value of the kernel does not increase beyond its value in $t = \varepsilon$. 

50
4.3 Seismic Recovery

4.3.1 Recovery Method and Recovery-Error Bound

The recovery of the seismic reflectivity could be achieved by solving the optimization problem presented in the following theorem. In addition, we also derive a bound on the recovery error.

*Theorem 1*. Let $y$ be of the form of (4.4) and let $\{g_{\sigma,m}\}$ be a set of admissible kernels as defined in Definition 2.1. If $K$ satisfies the separation condition of Definition 2.2 for $N > 0$ then the solution $\hat{x}$ of

$$
\min_{x \in l_1(\mathbb{Z})} ||x||_{l_1} \quad \text{subject to} \quad ||y[k] - \sum_m c_m g_{\sigma,m}[k - k_m]||_{l_1} \leq \delta \quad (4.14)
$$

satisfies

$$
||\hat{x} - x||_{l_1} \leq \frac{4\rho}{\beta \gamma_0} \delta \quad (4.15)
$$

$$
\rho \triangleq \max \left\{ \frac{\gamma_0}{\epsilon^2}, (N\sigma)^2 \alpha_0 \right\}
$$

where

$$
\alpha_0 = \max_m g_{\sigma,m}(0), \quad \gamma_0 = \min_m g_{\sigma,m}(0).
$$

The dependence of $x$ on the time $k$ is not written for simplicity.

*Proof.* see Appendix A.
Figure 4.1: Centered synthetic reflected wavelets and their derivatives, $Q = 125$, $\omega_0 = 100\pi$ (50Hz) (a) $g_{\sigma,m}(t)$; (b) $\partial g_{\sigma,m}(t)$; (c) $g_{\sigma,m}^{(2)}(t)$. 
Remarks

- This result guarantees that under the separation condition in Definition 2.2, a signal of the form of (4.4), can be recovered by solving the \( \ell_1 \) optimization problem formulated in (4.14). Moreover, a theoretical analysis of the recovered solution ensures that the error is bounded by a relatively small value, which depends mainly on the noise level and on the attenuation of the wavelets and is expressed through the parameters \( Q \) and \( \beta \).

- In the noiseless case where \( \delta = 0 \), the recovery is perfect. One would probably expect that the recovered solution would slightly deviate from the true one, yet this is not the case. This result does not depend on whether the spikes amplitude are very small or very large.

- If \( \gamma_0 = \alpha_0 \) we have the time-invariant case

\[
||\hat{x} - x||_{\ell_1} \leq \frac{4}{\beta} \max \left\{ \frac{1}{\epsilon^2}, (N\sigma)^2 \right\} \delta
\]

As expected, in the time-invariant case our result reduces into previous work results [5, 45]. The recovery error is proportional to the noise level \( \delta \), and small values of \( \beta \) (flat kernels) result in larger errors.

- In the time-variant setting most cases comply with \( \frac{\gamma_0}{\epsilon^2} < (N\sigma)^2 \alpha_0 \),

Then, the recovery error is bounded by

\[
||\hat{x} - x||_{\ell_1} \leq \frac{4(N\sigma)^2 \alpha_0}{\beta \gamma_0} \delta
\]

A smaller \( Q \) (which corresponds to a stronger degradation) results in higher \( \frac{\alpha_0}{\gamma_0} \) ratio and smaller \( \beta \) values. We will hereafter refer to the ratio \( \frac{\alpha_0}{\gamma_0} \) as the degradation ratio. Hence, the bound on the error in a time-variant environment implies that the error increases as \( Q \) gets smaller, which corresponds to a higher degradation ratio \( \frac{\alpha_0}{\gamma_0} \). As in
the time-invariant case, the error is linear with respect to the noise level $\delta$. Also, the error is sensitive to the flatness of the kernel near the origin. Namely, small $\beta$ results in an erroneous recovery.

### 4.3.2 Resolution Bounds

**Theorem 2.** Assume $\hat{x}[k] = \sum_m \hat{c}_m \delta[k - \hat{k}_m]$ is the solution of (4.14) where $\hat{K} \triangleq \{\hat{k}_m\}$ is the support of the recovered signal.

Let $y$ be of the form of $y[k] = \sum_m c_m g_{\sigma,m}[k - k_m] + n[k], \ |n|_1 \leq \delta$ and let $\{g_{\sigma,m}\}$ be a set of admissible kernels with two common parameters $\varepsilon, \beta > 0$, with $\varepsilon \geq \bar{\varepsilon} = \sqrt{\frac{\alpha_0}{\beta \varepsilon^2}}$

If $K$ satisfies the separation condition for $N > 0$, then the solution $\hat{x}$ satisfies:

1. $\sum_{k_m \in K; |k_m - k_m| > N \varepsilon, \forall k_m \in K} |\hat{c}_m| \leq \frac{2D_3 \alpha_0}{\beta \varepsilon^2} \delta$

Any redundant spike in $\hat{K}$ which is far from the correct support $K$ will for sure have small energy.

2. For any $k_m \in K$ if $|c_m| \geq D_4$, then there exist $\hat{k}_m \in \hat{K}$ such that

$$\left( \hat{k}_m - k_m \right)^2 \leq \frac{2D_3 (N \sigma)^2 \alpha_0}{\beta (|c_m| - D_4) \delta}.$$  

where

$$D_4 = \frac{2\delta}{\beta} \left( \frac{2\rho}{\gamma_0} + D_3 \alpha_0 \max \left\{ \frac{1}{\varepsilon^2}, \left(\frac{C_{2,m}}{(N \sigma)^2 g_m(0)}\right) \right\} \right),$$

$$D_3 = \frac{3\nu^2 (3\gamma_2 \nu^2 - \pi^2 \hat{C}_2) + \frac{12\pi^2 \hat{C}_1^2}{\gamma_0} (1 + \frac{\pi^2}{8\sigma}) \rho}{(3\gamma_2 \nu^2 - \pi^2 \hat{C}_2)(3\gamma_0 \nu^2 - 2\pi^2 \hat{C}_0)}.$$  

and

$$\hat{C}_l = \max_m C_{l,m}, \ l = 0, 1, 2, 3.$$  

This implies that for any $k_m \in K$ with sufficiently large amplitude $c_m$, under the separation condition, the recovered support location $\hat{k}_m \in \hat{K}$ is close to
the original one. The solution $\hat{x}$ consists of a recovered spike near any spike of the true reflectivity signal.

Proof. see Appendix B.

4.4 Experimental Results

4.4.1 Synthetic Data

We conducted various experiments in order to confirm the theoretical results. To solve the $\ell_1$ minimization in (14) we used CVX [58].

First, we try to estimate the minimal separation constant $\nu$ for various $Q$ values. We generate a synthetic reflectivity column, with sampling time $T_s = 2\text{ms}$. The reflectivity is statistically modeled as a zero-mean Bernoulli-Gaussian process [23]. The support was drawn from a Bernoulli process with $p = 0.2$ of length $L_r = 220\text{ taps}$, and the amplitudes were drawn from an i.i.d normal distribution with standard deviation $\nu = 10$. Then, we create the synthetic seismic trace in a noise-free environment, and try to recover the reflectivity by solving (4.14). Namely, we increase $\nu$ until we get an exact recovery in the noise-free setting. Figure 4.2 presents the results for $Q = 100, 200$. The initial wavelet was a Ricker wavelet with $\omega_0 = 100\pi$, i.e., $50\text{Hz}$. We repeat the experiment 10 times for each value of $\nu$. The Success Rate is 1 if the support’s recovery is perfect for all 10 experiments. As can be seen, the minimal separation constant for $Q = 200$ is $\nu = 1.9$ whereas for $Q = 100$ we have $\nu = 2.5$.

Figure 4.3 presents the recovery error $||\hat{x} - x||_{\ell_1}$ as a function of the noise level $\delta$ for different $Q$ values - $Q = \infty, 500, 200, 100$. $T_s = 4\text{ms}$ and $L_r = 176$. As in Fig. 4.2, the reflectivity is statistically modeled as a zero-mean Bernoulli-Gaussian process. Under the separation condition, the minimum distance between two spikes satisfies the minimal separation condition. The reflectivity is shown in Fig. 4.4(a). The initial wavelet was a Ricker wavelet.
with $\omega_0 = 140\pi$, i.e., 70Hz. Two seismic traces with SNR= $\infty$ and SNR= 15.5dB, are shown in Fig. 4.4(b) and Fig. 4.4(c) respectively. The recovered signals from these traces are shown in Fig. 4.4(d),(e). As can be seen in Fig. 4.3 the error is linear with respect to the noise. This implies that the bound we derived in Theorem 1 is reasonable. The theoretical bound is always greater or equal to the empirical error. As $Q$ gets smaller, $\beta$ - which is common to all reflected pulses - becomes significantly smaller. Hence, the theoretical bound slope becomes significantly larger compared with the empirical one. It can be seen also in the experimental results that as $Q$ gets smaller the error gets bigger. Table 1 presents the theoretical and practical parameters.

We compare the proposed solution to the blind deconvolution SOOT algorithm of Repetti et al. [38]. Fig 4.5. presents the results with noise level $\sigma = 0.01$ (SNR= 12.9 dB), $Q = 500$, $Ts = 4ms$, and an initial Ricker wavelet with $\omega_0 = 50\pi$, i.e., 25Hz. The original reflectivity section is depicted in Fig 4.5(a). The estimated reflectivities, obtained by SOOT, and by solving the $\ell_1$ minimization in (14) using CVX [58], for the seismic data in Fig. 4.5(b), are shown in Fig. 4.5(c) and (d) respectively. The results demonstrate that sparse recovery methods that do not take into consideration the attenuating and broadening nature of the wavelet, tend to annihilate small reflectivity spikes, especially in the deeper part of the reflectivity section.

Solving the $\ell_1$ minimization in (14) using CVX [58], the average processing time of a data set of 100x100 on Intel(R)Core(TM)i7-5600U@2.60GHz is 40.8 seconds.
Figure 4.2: Support detection vs. the separation constant $\nu$. Rate of success is the average number of perfect recoveries out of 10 experiments.

Figure 4.3: Recovery error $||\hat{x} - x||_{\ell_1}$ as a function of noise level $\delta$ for $Q = \infty, 500, 200, 100$. (a) Experimental results; (b) Theoretical bounds.
Table 4.1: Synthetic example: theoretical and estimated parameters: $Q$, the degradation ratio $\frac{a_0}{70}$, $\beta$, $\frac{4(N\sigma)^2 a_0}{\beta} - \frac{a_0}{70}$ - the bound slope computed from known parameters (by Theorem 1 $||\hat{x} - x||_1 \leq \frac{4(N\sigma)^2 a_0}{\beta} - \frac{a_0}{70}$), and the estimated slope computed from the experimental results in Fig.3(a).
Figure 4.4: 1D synthetic tests of (a) True reflectivity. (b), (c) Synthetic trace with 50 Hz Ricker wavelet and SNR= \( \infty \), 15.5 dB respectively, \( Q = 200 \). (d), (e) Recovered 1D channel of reflectivity signal.
Figure 4.5: 1D synthetic tests of (a) True reflectivity. (b) Synthetic trace with 25 Hz Ricker wavelet and SNR= 12.9 dB, $Q = 500$. (c) Recovered 1D channel of reflectivity signal with SOOT. (d) Recovered 1D channel of reflectivity signal with the proposed time-variant model.
### 4.4.2 Real Data

We applied the proposed method, to real seismic data from a small land 3D survey in North America (courtesy of GeoEnergy Inc., TX) of size $380 \times 160$, shown in Fig. 4.6(a). The time interval is 2ms. Assuming an initial Ricker wavelet with $\omega_0 = 140\pi$ (70Hz). We estimated $Q = 80$ using common midpoints (CMP) as described in [59]. Then, using (6)-(11) we estimated all possible kernels and solved (14) using CVX [58]. The recovered reflectivity section is shown in Fig. 4.6(b). The seismic data reconstructed from the estimated reflectivity using the known sequence $\{g_{\sigma,m}(t)\}$, is shown in Fig. 4.5(c). Visually analyzing this reflectivity section, it can be seen that the layer boundaries in the estimate are clear and quite continuous and smooth. It can also be seen that the reconstructed seismic data fits the original given observation. Since the ground truth is unknown, in order to measure the accuracy in the location and amplitude of the recovered reflectivity spikes we compute the correlation coefficient between the reconstructed data to the given seismic data. In this example we have $\rho_{s,\hat{s}} = 0.967$, which indicates that the reflectivity is estimated with very high precision. Figure 4.7(a) shows the estimated reflectivity considering a time-invariant model, using Sparse Spike Inversion (SSI) [6]. The result for a time-varying model is shown in Fig.4.7(b). It can be seen, especially in the lower (deeper) half of the image, that the method introduced in this paper produces much clearer results, since it takes in to account the attenuating and broadening nature of the waves as they travel further into the ground and back. Moreover, in terms of correlation coefficients, for SSI we have $\rho_{s,\hat{s}} = 0.89$, implying that considering a time-varying model indeed yields better results.
Figure 4.6: Real data inversion results: (a) Real seismic data (b) Estimated reflectivity (c) Reconstructed data.
Figure 4.7: Real data inversion results: (a) Estimated reflectivity - time-invariant model (SSI) (c) Estimated reflectivity - time-variant model.
We have presented a seismic inversion algorithm under time-variant model. The algorithm both promotes sparsity of the solution and also takes into consideration attenuation and dispersion effects resulting in shape distortion of the wavelet. The inversion results are demonstrated on synthetic and real data, under sufficiently high SNR. We derived a bound on the recovery $\ell_1$ error and observed that the error increases as $Q$ gets smaller. As in the time-invariant case, the error is proportional to the noise level. Also, the error is sensitive to the flatness of the kernel near the origin. Simulation results confirm the theoretical bound. We also proved that under the separation condition, for any spike with large-enough amplitude the recovered support location is close to the original one. The solution consists of a recovered spike near every spike of the true reflectivity signal. Any redundant spike in the recovered signal, which is far from the correct support, has small energy. Future research can address the problem of model mismatch. In addition we can elaborate the solution suggested in this paper to non-constant Q layers model.
Chapter 5

Conclusions

5.1 Summary

We have proposed two algorithms for seismic recovery of 2D reflectivity. Both algorithms produce visually superior results, for synthetic and real data.

The first algorithm - MSSI - both promotes sparsity of the solution and also takes into account the spatial dependency between neighboring traces in the deconvolution process. We have demonstrated that our deconvolution results are visually superior, compared to a single-channel deconvolution algorithm, for synthetic and real data, under various SNR levels. We’ve seen two implementations - for 2 columns and for 3 columns. Our second implementation (MSSI-3) performs better, on both synthetic and real data, since it takes into account more information from neighboring traces in the deconvolution process of each trace, compared to the first implementation (MSSI-2) that uses information from only one neighboring trace. The improved performance of the proposed algorithm compared to the single-channel algorithm was also apparent in qualitative assessment. It also shows that the second implementation’s results are more accurate.

The second algorithm uses information about the nature of attenuation
and dispersion propagation effects of the source wavelet in the ground. The algorithm both promotes sparsity of the solution and also takes into consideration attenuation and dispersion effects resulting in shape distortion of the wavelet. The recovery results are demonstrated on synthetic and real data, under sufficiently high SNR.

We derive a bound on the recovery $\ell_1$ error and observed that the error increases as $Q$ gets smaller. As in the time-invariant case, the error is proportional to the noise level. Also, the error is sensitive to the flatness of the kernel near the origin. Simulation results demonstrated that the bound is indeed reasonable.

We prove that under the separation condition, for any spike with large-enough amplitude the recovered support location is close to the original one. The solution consists of a recovered spike near any spike of the true reflectivity signal. Any redundant spike in the recovered signal, which is far from the correct support, will have small energy.

### 5.2 Future Research

The methods we have proposed in this thesis open a number of options for further study.

1. The second time-variant solution can be elaborated to non-constant $Q$ layers model. Assume that the earth $Q$ model has a multi-layered structure. Meaning that each layer has a different $Q$ constant. First the depth of each $Q$-layer and its $Q$ value needs to be identified. Then, the recovery suggested in ch.4 could be implemented according to the first step results by simply building the suitable dictionary.

2. The two algorithms could be combined by integrating the two optimization problems. An efficient solution method should be determined. Then, the performance of the new algorithm can be assessed
and compared to that of the two original ones.

3. The methods assume that the wavelet is known. A blind deconvolution solution could be investigated. A procedure that repeats two stages: wavelet estimation and then reflectivity estimation could be analyzed.

4. The first proposed algorithm can be expanded to handle 3D input data. In this case the recovered reflectivity is also a 3D signal. For MSSI the 3D estimation window can be used so that neighboring reflectivity columns from 4 directions will be taken into account in the estimation process of each reflectivity column.
Appendix A

Proof of Theorem 1

The proof follows the outline of research in [42, 45].

Denote \( g_m(t) \triangleq g_{\sigma,m}|_{\sigma=1} \). In a similar manner to [42, 45], we build a function of the form

\[
q(t) = \sum_m a_m g_m(t - t_m) + b_m g_m^{(1)}(t - t_m).
\]

The function \( q(t) \) satisfies

\[
q(t_k) = v_k \quad \forall t_k \in T,
\]

\[
q^{(1)}(t_k) = 0 \quad \forall t_k \in T,
\]

\[
|v_k| = 1;
\]

Its existence enables us to decouple the estimation error at \( t_m \) from the amplitude of the rest of the spikes. The magnitude of \( q(t) \) reaches a local maximum on the true support. This will in turn enable us to bound the recovery error.

In the following proof we use the following proposition and two lemmas.

**Proposition 3.** Assume a set of delays \( T \triangleq \{t_m\} \) that satisfies the separation condition, and let \( \{g_m\} \) be a set of admissible kernels as defined in Definition...
2.1. Then, there exist coefficients \{a_m\} and \{b_m\} such that

\[
q(t) = \sum_m a_m g_m(t - t_m) + b_m g_m^{(1)}(t - t_m), \quad (A.1)
\]

\[
|q(t_k)| = 1 \quad \forall t_k \in T, \quad (A.2)
\]

and

\[
q^{(1)}(t_k) = 0 \quad \forall t_k \in T. \quad (A.3)
\]

The coefficients are bounded by

\[
||a||_\infty \leq \frac{3\nu^2}{3\gamma_0 \nu^2 - 2\pi^2 C_0},
\]

\[
||b||_\infty \leq \frac{3\pi^2 \tilde{C}_1 \nu^2}{(3\gamma_2 \nu^2 - \pi^2 \tilde{C}_2)(3\gamma_0 \nu^2 - 2\pi^2 \tilde{C}_0)},
\]

where \(a \triangleq \{a_m\}, \ b \triangleq \{b_m\}\) are coefficient vectors and

\[
\tilde{C}_l = \max_m C_{l,m}, \ l = 0, 1, 2, 3.
\]

We also have

\[
a_m \geq \frac{1}{\alpha_0 + 2\tilde{C}_0 E(\nu) + \frac{(2\tilde{C}_1 E(\nu))^2}{\gamma_0 - 2\tilde{C}_2 E(\nu)}}.
\]

In other words, if the support is scattered, it is possible to build a function \(q(t)\) that interpolates any sign pattern exactly.

**Proof.**

The admissible kernel and its derivatives decay rapidly away from the origin. The proofs of this Proposition, the Theorem and the two Lemmas make repeated use of these facts.

According to (A.2) and (A.3)

\[
\sum_m a_m g_m(t_k - t_m) + b_m g_m^{(1)}(t_k - t_m) = v_k,
\]

and

\[
\sum_m a_m g_m^{(1)}(t_k - t_m) + b_m g_m^{(2)}(t_k - t_m) = 0,
\]

69
for all \( t_k \in T \), where \( v_k \in \mathbb{R} \) so that \( |v_k| = 1 \).

Therefore, in matrix-vector formulation we express these constraints

\[
\begin{bmatrix}
G_0 & G_1 \\
G_1 & G_2
\end{bmatrix}
\begin{bmatrix}
a \\
b
\end{bmatrix} =
\begin{bmatrix}
v \\
0
\end{bmatrix},
\]

where \( G_l \) \( k,m \) \( \triangleq \) \( g_m(0) \), \( l = 0, 1, 2 \) and \( a \triangleq \{a_m\}, \ b \triangleq \{b_m\}, \ v \triangleq \{v_k\} \).

We know that the matrix \( G \) is invertible if both \( G_2 \) and the Schur complement of \( G_2 \):

\[ S = G_0 - G_1G_2^{-1}G_1 \] are invertible [10]. We also know that a matrix \( A \) is invertible if there exists \( \alpha \neq 0 \) such that \( ||\alpha I - A||_\infty < |\alpha| \), where \( ||A||_\infty = \max_j \sum_i |a_{i,j}| \). In this case we also know that

\[ ||A^{-1}||_\infty \leq \frac{1}{|\alpha| - ||\alpha I - A||_\infty}. \] (A.4)

Denote

\[ \alpha_l = \max_m |g_m^{(l)}(0)|, \quad \gamma_l = \min_m |g_m^{(l)}(0)| \]

\[ \Delta_l = \alpha_l - \gamma_l = \max_m |g_m^{(l)}(0)| - \min_m |g_m^{(l)}(0)|. \]

We can observe that

\[ ||\alpha_2 I - G_2||_\infty = \max_k \left[ \sum_{m \neq k} |g_m^{(2)}(t_k - t_m)| + |g_k^{(2)}(0) - \alpha_2| \right] \]

\[ \leq \tilde{C}_2 \max_k \sum_{m \neq k} \frac{1}{1 + (t_k - t_m)^2} + \Delta_2 \leq \tilde{C}_2 \max_k \sum_{m \neq k} \frac{1}{1 + ((k - m)\nu)^2} + \Delta_2. \]

Since

\[ \sum_{n=1}^{\infty} \frac{1}{1 + (n\nu)^2} \leq E(\nu) \triangleq \frac{\pi^2}{6\nu^2}, \] (A.5)

\[ ||\alpha_2 I - G_2||_\infty \leq 2\tilde{C}_2 E(\nu) + \Delta_2 < \alpha_2, \] (A.6)

which leads us to

\[ 2\tilde{C}_2 \pi^2 < (\alpha_2 - \Delta_2)6\nu^2. \] (A.7)
Therefore,
\[ \nu^2 > \frac{\tilde{C}_2 \pi^2}{3 \gamma_2}. \]  
(A.8)

The result is quite intuitive. Small values of \( \gamma_2 \) (small Q values) require a larger separation constant. The flattest kernel determines the global separation requirements for perfect recovery.

Now, we can also derive
\[ \| \alpha_0 I - S \|_\infty = \| \alpha_0 I - G_0 + G_1 G_2^{-1} G_1 \|_\infty \leq \| \alpha_0 I - G_0 \|_\infty + \| G_1 \|_\infty^2 \| G_2^{-1} \|_\infty. \]  
(A.9)

In a similar manner to (A.6)
\[ \| \alpha_0 I - G_0 \|_\infty \leq 2 \tilde{C}_0 E(\nu) + \Delta_0, \]  
(A.10)

and since \( g_m^{(1)}(0) \approx 0 \ \forall \ m \)
\[ \| G_1 \|_\infty \leq 2 \tilde{C}_1 E(\nu). \]  
(A.11)

Using (A.4) and (A.6)
\[ \| G_2^{-1} \|_\infty \leq \frac{1}{\| \alpha_2 - \alpha_2 I - G_2 \|_\infty} \leq \frac{1}{\alpha_2 - 2 \tilde{C}_2 E(\nu) - \Delta_2} = \frac{1}{\gamma_2 - 2 \tilde{C}_2 E(\nu)}. \]  
(A.12)

So, we have
\[ \| \alpha_0 I - S \|_\infty \leq 2 \tilde{C}_0 E(\nu) + \Delta_0 + \frac{(2 \tilde{C}_1 E(\nu))^2}{\gamma_2 - 2 \tilde{C}_2 E(\nu)} = 2 \tilde{C}_0 E(\nu) \left[ 1 + \frac{2 \tilde{C}_1 E(\nu)}{\tilde{C}_0 (\gamma_2 - 2 \tilde{C}_2 E(\nu))} \right] + \Delta_0 \leq 4 \tilde{C}_0 E(\nu) + \Delta_0. \]  
(A.13)

The last inequality holds for
\[ 2 \tilde{C}_1^2 E(\nu) \leq \tilde{C}_0 (\gamma_2 - 2 \tilde{C}_2 E(\nu)). \]

Leading us to
\[ 3 \nu^2 \tilde{C}_0 \gamma_2 \geq \pi^2 (\tilde{C}_1^2 + \tilde{C}_0 \tilde{C}_2), \]
which yields the condition
\[ \nu^2 \geq \frac{\pi^2 (\tilde{C}_1^2 + \tilde{C}_0 \tilde{C}_2)}{3 \tilde{C}_0 \gamma_2}. \]  
(A.14)

Then, \( S \) is invertible if
\[ 4 \tilde{C}_0 E(\nu) + \Delta_0 < \alpha_0, \]
\[ 4 \tilde{C}_0 E(\nu) < \alpha_0 - \Delta_0 = \gamma_0, \]
\[ \frac{2 \tilde{C}_0 \pi^2}{3 \nu^2} < \gamma_0, \]
\[ \nu^2 > \frac{2 \pi^2 \tilde{C}_0}{3 \gamma_0}. \]  
(A.15)

Here again it can be observed that small \( \gamma_0 \) values require a larger separation constant \( \nu \). Finally we get
\[ ||\alpha_0 I - S||_\infty \leq 4 \tilde{C}_0 E(\nu) + \Delta_0 < \alpha_0. \]  
(A.16)

and \( S \) is invertible. So we have proved that \( q(t) \) exists under certain conditions on the separation constant \( \nu \).

In addition \( a \) and \( b \) are given by
\[
\begin{bmatrix}
  a \\
  b
\end{bmatrix}
= \begin{bmatrix}
  G_0 & G_1 \\
  G_1 & G_2
\end{bmatrix}^{-1}
\begin{bmatrix}
  v \\
  0
\end{bmatrix},
\]
\[
\begin{bmatrix}
  a \\
  b
\end{bmatrix}
= \begin{bmatrix}
  S^{-1}v \\
  -G_2^{-1}G_1 S^{-1}v
\end{bmatrix}.
\]  
(A.17)

Using (A.4) and (A.16) we have
\[ ||a||_\infty \leq ||S^{-1}||_\infty \leq \frac{1}{||\alpha_0|| - ||\alpha_0 I - S||_\infty} \leq \frac{1}{\alpha_0 - 4 \tilde{C}_0 E(\nu) - \Delta_0} = \frac{1}{\gamma_0 - 4 \tilde{C}_0 E(\nu)}. \]  
(A.18)
Using (A.11) and (A.12) we also have
\[ ||b||_\infty \leq ||G_2^{-1}||_\infty ||G_1||_\infty ||S^{-1}||_\infty \leq \frac{2\tilde{C}_1E(\nu)}{(\gamma_2 - 2\tilde{C}_2E(\nu))(\gamma_0 - 4\tilde{C}_0E(\nu))}. \] (A.19)

Assuming \( v_k = 1 \), we get
\[ a_k = (S^{-1}v)_k = \sum_j (S^{-1})_{k,j}. \]

Since \( S^{-1}S = I \)
\[ \sum_j (S^{-1})_{k,j}(S)_{j,k} = 1. \]

We also know
\[ \sum_j |(S^{-1})_{k,j}(S)_{j,k}| \leq \sum_j |(S^{-1})_{k,j}| \sum_j |(S)_{j,k}|, \]
which leads us to
\[ \sum_j |(S^{-1})_{k,j}| \geq \frac{1}{\sum_j |(S)_{j,k}|}. \]

Also, we can derive
\[ \forall k \sum_j (S)_{j,k} \leq ||S||_1 \leq ||G_0||_1 + ||G_1||_1^2 ||G_2^{-1}||_1 \]
where \( ||A||_1 = \max_j \sum_i |a_{i,j}|. \)

We can observe that
\[ ||G_0||_1 = \max_k \sum_k |g_m(t_k - t_m)| \leq \alpha_0 + \tilde{C}_0 \max_m \sum_{k \neq m} \frac{1}{1 + (t_k - t_m)^2} \leq \alpha_0 + \tilde{C}_0E(\nu). \]

Similarly,
\[ ||\alpha_2I - G_2||_1 = \max_m \left[ \sum_{k \neq m} |g_m^{(2)}(t_k - t_m)| + |g_m^{(2)}(0) - \alpha_2| \right] \leq \tilde{C}_2 \max_m \sum_{k \neq m} \frac{1}{1 + (t_k - t_m)^2} + \Delta_2 \leq \tilde{C}_2E(\nu) + \Delta_2. \]

Therefore,
\[ ||G_2^{-1}||_1 \leq \frac{1}{||\alpha_2|| - ||\alpha_2I - G_2||_1} \leq \frac{1}{\gamma_2 - 2\tilde{C}_2E(\nu)}. \]
And since \( g_m^{(1)}(0) \approx 0 \ \forall \ m \)
\[
\|G_1\|_1 \leq 2\hat{C}_1 E(\nu),
\] (A.20)
which leads us to
\[
\sum_j |(S)_{j,k}| \leq \alpha_0 + 2\hat{C}_0 E(\nu) + \frac{(2\hat{C}_1 E(\nu))^2}{\gamma_2 - 2\hat{C}_2 E(\nu)}.
\]
So finally
\[
a_k \geq \frac{1}{\alpha_0 + 2\hat{C}_0 E(\nu) + \frac{(2\hat{C}_1 E(\nu))^2}{\gamma_2 - 2\hat{C}_2 E(\nu)}}.
\] (A.21)
Hence, \( a_k \)'s lower bound is inversely proportional to \( \alpha_0 \). Numerical experiments have shown that this bound is tight, meaning that the smallest \( a_k \) is the exact reciprocal of the amplitude of the strongest kernel in the observation signal. This result is significant since it indicates that the bound on the recovery error is not merely stating the time-invariant result for the kernel with the worst constants. The better kernels are also taken into account.

It can be said that the recovery error is proportional to the degradation ratio \( \frac{a_0}{m} \), which is the ratio between the amplitude of the best kernel to the amplitude of the worst kernel.

**Lemma 4.** Under the separation condition with \( \varepsilon < \nu/2 \), \( q \) as in Proposition 3 satisfies \( |q(t) < 1| \) if \( 0 < |t - t_m| \leq \varepsilon \) for some \( t_m \in T \).

**Lemma 5.** Under the separation condition with \( \varepsilon < \nu/2 \), \( q \) as in Proposition 3 satisfies \( |q(t) < 1| \) if \( |t - t_m| > \varepsilon \ \forall t_m \in T \).

**Proof of Lemma 4**

Assume \( t \in \mathbb{R} \) and \( t_k \leq t \leq t_k + \varepsilon \) for some \( t_k \in T \), and that \( q(t_k) = 1 \).

(The proof is the same for \( t_k - \varepsilon \leq t \leq t_k \) or \( v_k = -1 \).) We also assume that \( T \) satisfies the separation condition with \( \varepsilon < \nu/2 \). Therefore, we have \( |t - t_m| > \frac{\varepsilon}{2} \) for \( m \neq k \). Then, for \( l = 0, 1, 2, 3 \) we have
\[
\sum_{m \neq k} |g_m^{(l)}(t - t_m)| \leq \sum_{m \neq k} \frac{C_{l,m}}{1 + (t - t_m)^2} \leq \hat{C}_l \sum_{m \neq k} \frac{1}{1 + \left(\frac{(k - m)\nu/2}{2}\right)^2} \leq 8\hat{C}_l E(\nu) = \frac{4\hat{C}_l}{3 \nu^2}.
\] (A.22)
Using this estimate, as well as (A.18),(A.19),(A.20), and the admissible kernels’ properties we obtain

\[
q^{(2)}(t) = \sum_m a_m g_m^{(2)}(t-t_m) + b_m g_m^{(3)}(t-t_m)
\]

\[
\leq a_k g_k^{(2)}(t-t_k) + ||a||_\infty \sum_{m \neq k} |g_m^{(2)}(t-t_m)| + ||b||_\infty \sum_m |g_m^{(3)}(t-t_m)|
\]

\[
\leq -\frac{\beta}{\alpha_0 + 2\tilde{C}_0 E(\nu) + \frac{16\tilde{C}_2 E(\nu)}{\gamma_2 - 2\tilde{C}_2 E(\nu)}} + \frac{8\tilde{C}_2 E(\nu)}{\gamma_2 - 4\tilde{C}_0 E(\nu)} + \frac{16\tilde{C}_3 (2E(\nu)+1)\tilde{C}_1 E(\nu)}{(\gamma_2 - 2\tilde{C}_2 E(\nu))(\gamma_0 - 4\tilde{C}_0 E(\nu))}.
\]

For sufficiently large \(\nu\) that depends on the parameters of \(g_m(t)\) we can approximate

\[
q^{(2)}(t) < -\frac{\beta}{\alpha_0}.
\]  

(A.23)

By the Taylor Remainder theorem [60], for any \(t_k < t < t_k + \varepsilon\) there exists \(t_k < \xi \leq t\) such that

\[
q(t) = q(t_k) + q^{(1)}(t_k)(t-t_k) + \frac{1}{2}q^{(2)}(\xi)(t-t_k)^2.
\]  

(A.24)

Since by construction \(q^{(1)}(t_k) = 0\).

For sufficiently large \(\nu\)

\[
q(t) \leq 1 - \frac{\beta}{2\alpha_0}(t-t_k)^2.
\]  

(A.25)

So we have shown that \(q(t) < 1\).

To show that \(q(t) > -1\)

\[
q(t) = \sum_m a_m g_m(t_k - t_m) + b_m g_m^{(1)}(t_k - t_m)
\]

\[
\geq a_k g_k(t-t_k) - ||a||_\infty \sum_{m \neq k} |g_m(t-t_m)| - ||b||_\infty \sum_m |g_m^{(1)}(t-t_m)|
\]

\[
\geq \frac{g_k(\varepsilon)}{\alpha_0 + 2\tilde{C}_0 E(\nu) + \frac{16\tilde{C}_2 E(\nu)}{\gamma_2 - 2\tilde{C}_2 E(\nu)}} - \frac{8\tilde{C}_2 E(\nu)}{\gamma_2 - 4\tilde{C}_0 E(\nu)} - \frac{16\tilde{C}_3 (2E(\nu)+1)\tilde{C}_1 E(\nu)}{(\gamma_2 - 2\tilde{C}_2 E(\nu))(\gamma_0 - 4\tilde{C}_0 E(\nu))}.
\]

Hence, for sufficiently large \(\nu\) and \(\gamma_0\) we’ve shown that

\[
q(t) > -1, \quad \text{for} \quad t_k < t < t_k + \varepsilon,
\]  

(A.26)

and

\[
|q(t)| < 1 \quad |t-t_k| < \varepsilon, \quad t_k \in T.
\]  

(A.27)
Proof of Lemma 5

Assume \( t \in \mathbb{R} \) and \( |t - t_m| > \varepsilon \) for all \( t_m \in T \), since \( \varepsilon < \nu/2 \), we have \( |t - t_m| > \nu/2 \). Then, from (A.1), the admissible kernel’s properties, (A.18) and (A.19), we can write

\[
|q(t)| \leq ||a||_\infty \sum_m |g_m(t - t_m)| + ||b||_\infty \sum_m |g_m^{(1)}(t - t_m)|. \tag{A.28}
\]

Let us denote

\[
\hat{m} = \arg \min_m |g_m(0)|. \tag{A.29}
\]

By assumption

\[
\frac{g_{\hat{m}}(t - t_{\hat{m}})}{\gamma_0} < 1.
\]

We recall,

\[
\gamma_0 = g_{\hat{m}}(0),
\]

Moreover, since \( |t - t_m| > \varepsilon \)

\[
0 < \frac{g_{\hat{m}}(t - t_{\hat{m}})}{\gamma_0} < \frac{g_{\hat{m}}(\varepsilon)}{\gamma_0}.
\]

By the Taylor Remainder theorem and the properties of \( g_m(t) \), we know

\[
0 < g_m(\varepsilon) \leq g_m(0) - \frac{\beta \varepsilon^2}{2}.
\]

Therefore,

\[
|q(t)| \leq ||a||_\infty \left( g_{\hat{m}}(t - t_{\hat{m}}) + \sum_{m \neq \hat{m}} |g_m(t - t_m)| \right) + ||b||_\infty \sum_m |g_m^{(1)}(t - t_m)|
\]

\[
\leq \frac{g_{\hat{m}}(t - t_{\hat{m}}) + g_{\hat{m}}(\varepsilon)}{\gamma_0} + \frac{16\hat{C}_1^2 E(\nu)}{(\gamma_2 - 2\hat{C}_2 E(\nu))(\gamma_0 - 4\hat{C}_0 E(\nu))}. \tag{A.30}
\]

Finally, we can conclude that for sufficiently large \( \nu \),

\[
|q(t)| \leq 1 - \frac{\beta \varepsilon^2}{2\gamma_0}. \tag{A.31}
\]

Now, we can complete the proof of Theorem 1. Assume \( \hat{x} \) is the solution of the optimization problem in (4.14). \( \hat{x} \) obeys \( ||\hat{x}||_{\ell_1} \leq ||x||_{\ell_1} \).
Denote the error $h[k] \triangleq \hat{x}[k] - x[k]$.

Now separate $h$ into $h = h_K + h_{KC}$, where $h_K$’s support is in the true support $K \triangleq \{k_m\}$. If $h_K = 0$, then $h = 0$, since $h_K = 0$ and $h \neq 0$ would imply that $h_{KC} \neq 0$ and $||\hat{x}||_{\ell_1} \geq ||x||_{\ell_1}$.

Under the separation condition, the set $T = \{t_m\}$ satisfies $t_i - t_j \geq \nu \sigma$ for $i \neq j$. We’ve shown in Proposition 3 that there exists $q$ of the form (A.1) such that

$$q(t_m) = q(t_m) = \text{sgn}(h_K[k_m]) \quad \forall k_m \in T.$$  \hspace{1cm} (A.32)

By assumption, we choose $v_m = \text{sgn}(h_K(t_m))$.

In addition, $q$ also satisfies $|q(t)| < 1$ for $t \notin T$.

We then define

$$q_{\sigma}(t) = q(t) = \sum_m a_m g_{m,\sigma}(t - \frac{k_m}{N}) + b_m g_{m,\sigma}^{(1)}(t - \frac{k_m}{N}).$$

So that

$$q_{\sigma}(k_m) = q_{\sigma}(k_m) = \text{sgn}(h_K[k_m]) \quad \forall k_m \in K,$$

and

$$|q_{\sigma}(k)| < 1 \quad \forall k \notin K.$$

Denote $g_{m,\sigma}[k] \triangleq g_{m,\sigma}(\frac{k}{N})$. Consequently, we can obtain

$$\left| \sum_{k \in \mathbb{Z}} q_{\sigma}[k] h[k] \right|_1 \leq \left| \sum_{k \in \mathbb{Z}} \left( \sum_{k_m \in K} a_m g_{m,\sigma}(k - k_m) + b_m g_{m,\sigma}^{(1)}(k - k_m) \right) h[k] \right|_1$$

$$\leq ||a||_{\infty} \left| \sum_{k \in \mathbb{Z}} \left( \sum_m g_{m,\sigma}[k - k_m] h[k] \right) \right|_1$$

$$+ ||b||_{\infty} \left| \sum_{k \in \mathbb{Z}} \left( \sum_m g_{m,\sigma}^{(1)}[k - k_m] h[k] \right) \right|_1$$  \hspace{1cm} (A.33)

We also have,
\[ \left| \sum_m g_m,\sigma [k - k_m] h[k] \right|_1 \leq \left| y[k] - \sum_m g_m,\sigma [k - k_m] \hat{x}[k] - \sum_m g_m,\sigma [k - k_m] x[k] \right|_1 \]

\[ \leq \left| y[k] - \sum_m g_m,\sigma [k - k_m] x[k] \right|_1 + \left| y[k] - \sum_m g_m,\sigma [k - k_m] \hat{x}[k] \right|_1 \leq 2\delta. \]  

(A.34)

Since,

\[ \left| y[k] - x[k] \right|_1 = \left| y[k] - \sum_m c_m g_m,\sigma [k - k_m] \right|_1 \leq \delta, \]

and also,

\[ \left| y[k] - \hat{x}[k] \right|_1 = \left| y[k] - \sum_m \hat{c}_m g_m,\sigma [k - k_m] \right|_1 \leq \delta. \]

As mentioned above \( \{ g_m \} \) is a set of admissible kernels. Therefore,

\[ |g_m^{(1)} [k - k_m]| = \left| g_m^{(1)} \left( \frac{k - k_m}{N\sigma} \right) \right| \leq \frac{C_{1,m}}{1 + \left( \frac{k - k_m}{N\sigma} \right)^2}. \]  

(A.35)

Under the separation condition we have \( |k_i - k_j| \geq N\nu\sigma \quad \forall k_i, k_j \in K. \)

Hence, for any \( k \) we have

\[ \sum_{k_m \in K} \frac{1}{1 + \left( \frac{k - k_m}{N\sigma} \right)^2} < 2(1 + E(\nu)). \]  

(A.36)

Since we know

\[ \sum_{n=1}^{\infty} \frac{1}{1 + (n\nu)^2} \leq E(\nu) \triangleq \frac{\pi^2}{6\nu^2}. \]

Then,

\[ \left| \sum_{k_m \in K} \left( \sum_{k \in \mathbb{Z}} g_m,\sigma [k - k_m] h[k] \right) \right|_1 \leq \tilde{C}_1 \sum \left| h[k] \right|_1 \sum_{k_m \in K} \frac{1}{1 + \left( \frac{k - k_m}{N\sigma} \right)^2} < 2\tilde{C}_1 (1 + E(\nu)) \| h \|_1. \]

Hence,

\[ \left| \sum_{k \in \mathbb{Z}} g_{\sigma} [h[k]] \right|_1 \leq 2\delta \| a \|_\infty + 2\tilde{C}_1 (1 + E(\nu)) \| b \|_\infty \| h \|_1. \]  

(A.37)
On the other hand,\[
\left| \sum_{k \in \mathbb{Z}} q_{\sigma}[k] h[k] \right|_1 = \left| \sum_{k \in \mathbb{Z}} q_{\sigma}[k] (h_K[k] + h_{KC}[k]) \right|_1 \\
\geq \sum_{k \in \mathbb{Z}} |q_{\sigma}[k] h_K[k]|_1 - |q_{\sigma}[k] h_{KC}[k]|_1 \\
\geq ||h_K||_1 - \max_{k \in \mathbb{Z} \setminus K} |q_{\sigma}[k]| ||h_{KC}[k]||_1. \tag{A.38}
\]
Combining (A.37) and (A.38) we get,\[
||h_K||_1 - \max_{k \in \mathbb{Z} \setminus K} |q_{\sigma}[k]| ||h_{KC}[k]||_1 \leq 2\delta ||a||_\infty + 2\tilde{C}_1 (1 + E(\nu)) ||b||_\infty ||h||_1. \tag{A.39}
\]
We’ve shown in the proof of lemma 4 that for \( |k - k_m| \leq \varepsilon N\sigma \), for some \( k_m \in K \)
\[
|q_{\sigma}[k]| = \left| q_{\sigma} \left( \frac{k}{N\sigma} \right) \right| \leq 1 - \frac{\beta}{2\alpha_0 (N\sigma)^2}.
\]
And by (A.31) for \( |k - k_m| > \varepsilon N\sigma \) for all \( k_m \in K \)
\[
|q_{\sigma}[k]| = \left| q \left( \frac{k}{N\sigma} \right) \right| \leq 1 - \frac{\beta \varepsilon^2}{2\gamma_0}.
\]
So we can conclude\[
\max_{k \in \mathbb{Z}} |q_{\sigma}[k]| \leq 1 - \frac{\beta}{2\rho} \tag{A.40}
\]
where \( \rho \triangleq \max \left\{ \frac{\varepsilon_0}{2\pi}, (N\sigma)^2 \alpha_0 \right\} \).
Substituting (A.40) into (A.39) we get,\[
||h_K||_1 - (1 - \frac{\beta}{2\rho}) ||h_{KC}[k]||_1 \leq 2\delta ||a||_\infty + 2\tilde{C}_1 (1 + E(\nu)) ||b||_\infty ||h||_1. \tag{A.41}
\]
Moreover, we know that\[
||x||_1 \geq ||\hat{x}||_1 = ||x + h||_1 = ||x + h_K||_1 + ||h_{KC}||_1 \geq ||x||_1 - ||h_K||_1 + ||h_{KC}||_1.
\]
Which leads us to\[
||h_K||_1 \geq ||h_{KC}||_1.
\]
Combining this with (A.41) leads us to\[
||h||_1 = ||h_K||_1 + ||h_{KC}||_1 \leq 2||h_{KC}||_1 \leq \frac{4\rho}{\beta} (\delta ||a||_\infty + 2\tilde{C}_1 (1 + E(\nu)) ||b||_\infty ||h||_1).
\tag{A.42}
\]
So we have,
\[
\|h\|_1 \leq \frac{4\rho\|a\|_\infty}{\beta - 4\rho C_1(1 + E(\nu))\|b\|_\infty} \delta.
\] (A.43)

Using (A.18) and (A.19)
\[
\|h\|_1 \leq \frac{36\rho\gamma_2}{9\beta\gamma_0\gamma_2 - D_1\nu^{-1} - D_2\nu^{-2}} \delta,
\] (A.44)

\[D_1 = 3\pi^2(\beta\tilde{C}_2 + 2\beta\tilde{C}_0 + 4\tilde{C}_1\rho), \quad D_2 = \pi^4(2\rho\tilde{C}_1 - \beta\tilde{C}_2\tilde{C}_0).\]
Appendix B

Proof of Theorem 2

To prove Theorem 2 we use the following two Lemmas.

Lemma 6. Assume a set of delays $T \triangleq \{t_m\}$ that satisfies the separation condition

$$\min_{k_i, k_j \in K, i \neq j} |k_i - k_j| \geq N\nu\sigma.$$ 

Let $\{g_m\}$ be a set of admissible kernels as defined in Definition 2.1 or asymmetric approximately admissible kernels as described in section 2.2. Then, for any $t_m \in T$ there exist coefficients $\{a_k\}$ and $\{b_k\}$ such that the function

$$q_m(t) = \sum_k a_k g_m\left(\frac{t - t_k}{\sigma}\right) + b_k g_m^{(1)}\left(\frac{t - t_k}{\sigma}\right)$$  \hspace{1cm} (B.1)

obeys

$$q_m(t_m) = 1,$$

$$q_m(t_j) = 0 \quad \forall t_j \in T \setminus \{t_m\},$$

$$q_m^{(1)}(t_j) = 0 \quad \forall t_j,$$

$$|1 - q_m(t)| < \frac{C_{2,m}(t - t_m)^2}{g_m(0)\sigma^2} \quad \forall t \neq t_m, \quad (B.2)$$

$$|q_m(t)| < \frac{C_{2,m}(t - t_j)^2}{g_m(0)\sigma^2}, \quad \forall t_j \in T \setminus \{t_m\}, \quad |t - t_j| \leq \varepsilon\sigma, \quad (B.3)$$
\[ |q_m(t)| < 1 - \xi_m \varepsilon^2 \quad |t - t_j| > \varepsilon \sigma, \ \forall t_j \in T, \quad (B.4) \]

\[ \xi_m = \frac{\beta}{4g_m(0)}. \]

These results also for all \(0 < \varepsilon' < \varepsilon\).

**Remark** Notice that here we have a set of different admissible kernels, and each function \(q_m(t)\) is based on a different kernel \(\{g_m(t)\}\) and it decouples the estimation error at one location \(t_m\) from the amplitude of the rest of the support. It is designed to obey \(< q_m, x > \triangleq \int q_m(t)x(t)dt = c_m\).

**Lemma 7.** Assume \(K\) that satisfies the separation condition of Definition 2.2 for \(N > 0\), then

\[ \sum_{k_m \in K} |c_m| \min \left\{ \varepsilon^2, \frac{d(k_m, K)}{(N\sigma)^2} \right\} \leq \frac{2D_0\beta_0}{\beta} \delta, \]

where

\[ d(k, K) = \min_{k_n \in K} (k_n - k)^2. \]

**Proof of Lemma 6**

We impose

\[ q_m(t_m) = 1, \]

\[ q_m(t_j) = 0 \quad \forall t_j \in T \setminus \{t_m\}, \]

\[ q_m^{(1)}(t_j) = 0 \quad \forall t_j. \]

\[ \sum_k a_k g_m \left( \frac{t_m - t_k}{\sigma} \right) + b_k g_m^{(1)} \left( \frac{t_m - t_k}{\sigma} \right) = 1 \]

and

\[ \sum_k a_k g_m^{(1)} \left( \frac{t_j - t_k}{\sigma} \right) + b_k g_m^{(2)} \left( \frac{t_j - t_k}{\sigma} \right) = 0, \quad t_j \in T \setminus \{t_m\}. \]
In matrix form,
\[
\begin{bmatrix}
D_0 & D_1 \\
D_1 & D_2
\end{bmatrix}
\begin{bmatrix}
a \\
b
\end{bmatrix} =
\begin{bmatrix}
e_t \\
0
\end{bmatrix},
\]
where \(e_t\) is a vector with one nonzero entry at the location corresponding to \(t_m\), \(a \triangleq \{a_m\}\), \(b \triangleq \{b_m\}\) and \((D_l)_{j,k} \triangleq g_m^{(l)} \left( \frac{t_j - t_k}{\sigma} \right), l = 0, 1, 2\). As we mentioned in proposition 3, we know that the matrix \(D\) is invertible if both \(D_2\) and the Schur complement of \(D_2\):
\[S = D_0 - D_1D_2^{-1}D_1\]
are invertible [10]. We also know that a matrix \(A\) is invertible if there exists \(\alpha \neq 0\) such that \(||\alpha I - A||_\infty < |\alpha|\), where \(||A||_\infty = \max_i \sum_j |a_{i,j}|\). In this case we have
\[
||A^{-1}||_\infty \leq \frac{1}{|\alpha| - ||\alpha I - A||_\infty}.
\]
Using the properties of an admissible kernel and the separation condition, we can write
\[
||g_m^{(2)}(0)I - D_2||_\infty = \max_k \sum_{m \neq k} |g_m^{(2)} \left( \frac{t_k - t_m}{\sigma} \right)|
\leq \frac{C_{2,m}}{\sigma^2} \max_k \sum_{m \neq k} \frac{1}{1 + \left( \frac{t_k - t_m}{\sigma} \right)^2} \leq \frac{C_{2,m}}{\sigma^2} \max_k \sum_{m \neq k} \frac{1}{1 + ((k - m)\nu)^2}.
\]
Recall that
\[
\sum_{n=1}^\infty \frac{1}{1 + (n\nu)^2} \leq E(\nu) \triangleq \frac{\pi^2}{6\nu^2}.
\]
Therefore,
\[
||g_m^{(2)}(0)I - D_2||_\infty \leq \frac{2C_{2,m}}{\sigma^2} E(\nu). \tag{B.5}
\]
Meaning that \(D_2\) is invertible if \(\frac{2C_{2,m}}{\sigma^2} E(\nu) < |g_m^{(2)}(0)|\).
This yields the condition
\[
\nu^2 \geq \frac{2C_{2,m}\sigma^2}{3|g_m^{(2)}(0)|}\tag{B.6}
\]
This implies that as \(m\) increases and \(g_m(t)\) loses more energy, in order to achieve the correct recovery by \(\ell_1\) optimization, the minimum distance be-
tween two adjacent spikes should be larger.

\[ \|g_m(0)I - S\|_\infty = \|g_m(0)I - D_0 + D_1D_2^{-1}D_1\|_\infty \leq \|g_m(0)I - D_0\|_\infty + \|D_1\|_\infty^2\|D_2^{-1}\|_\infty. \]  \hspace{1cm} (B.7)

In a similar manner to (B.5)

\[ \|g_m(0)I - D_0\|_\infty \leq 2C_{0,m}E(\nu). \]  \hspace{1cm} (B.8)

And since \( g_m^{(1)}(0) \approx 0 \forall m \)

\[ \|D_1\|_\infty \leq \frac{2C_{1,m}}{\sigma} E(\nu). \]  \hspace{1cm} (B.9)

Using (A.4) and (B.5) we get

\[ \|D_2^{-1}\|_\infty \leq \frac{1}{|g_m^{(2)}(0)| - \|g_m^{(2)}I - D_2\|_\infty} \leq \frac{1}{|g_m^{(2)}(0)| - \frac{2C_{2,m}}{\sigma^2} E(\nu)}. \]  \hspace{1cm} (B.10)

So, we have

\[ \|g_m(0)I - S\|_\infty \leq 2C_{0,m}E(\nu) + \frac{(2C_{1,m}E(\nu))^2}{|g_m^{(2)}(0)| - \frac{2C_{2,m}}{\sigma^2} E(\nu)} \]

\[ = 2C_{0,m}E(\nu) \left[ 1 + \frac{2C_{1,m}^2E(\nu)}{C_{0,m}(|g_m^{(2)}(0)| - \frac{2C_{2,m}}{\sigma^2} E(\nu))} \right] \leq 4C_{0,m}E(\nu). \]  \hspace{1cm} (B.11)

The last inequality holds for

\[ \nu^2 \geq \frac{\pi^2(C_{1,m}^2 + C_{0,m}C_{2,m})}{3\sigma^2 C_{0,m} |g_m^{(2)}(0)|}. \]  \hspace{1cm} (B.12)

If in addition

\[ \nu^2 > \frac{2\pi^2 C_{0,m}}{3\sigma^2 |g_m(0)|}. \]  \hspace{1cm} (B.13)

Here again it can be observed that small \( g_m(0) \) values require a larger separation constant \( \nu \). Then finally we have

\[ \|g_m(0)I - S\|_\infty \leq 4C_{0,m}E(\nu) < g_m(0) \]  \hspace{1cm} (B.14)

and \( S \) is invertible.

Furthermore, \( a \) and \( b \) are given by

\[
\begin{bmatrix}
  a \\
  b
\end{bmatrix} = \begin{bmatrix}
  D_0 & D_1 \\
  D_1 & D_2
\end{bmatrix}^{-1} \begin{bmatrix}
  e_{tm} \\
  0
\end{bmatrix},
\]

84
\[
\mathbf{a} = \begin{bmatrix} I \\ -D_2^{-1}D_1S^{-1}\mathbf{e}_m \end{bmatrix}.
\] (B.15)

Using (A.4) and (B.14) we have
\[
||\mathbf{a}||_\infty \leq ||S^{-1}||_\infty \leq \frac{1}{|g_m(0)| - ||g_m(0)I - S||_\infty} \leq \frac{1}{g_m(0) - 4C_{0,m}E(\nu)}. \tag{B.16}
\]

Using (B.9) and (B.10) we also have
\[
||\mathbf{b}||_\infty \leq ||D_2^{-1}||_\infty||D_1||_\infty||S^{-1}||_\infty \leq \frac{2C_{1,m}E(\nu)}{(|g_m^{(2)}(0)| - \frac{2C_{2,m}}{\sigma^2}E(\nu))(g_m(0) - 4C_{0,m}E(\nu))}. \tag{B.17}
\]

And we can also derive
\[
\begin{align*}
\text{Fix } t_k \in T \text{ and } |t - t_k| \leq \varepsilon \sigma. \text{ Under the separation condition we have } |t - t_j| \geq \frac{\nu}{2} \text{ for } t_j \in T \setminus \{t_k\}. \text{ Therefore, we have for } l = 0, 1, 2, 3:
\sum_{j \neq k} \frac{|g_m^{(0)}(t - t_j)|}{|g_m(0)|} \leq \frac{C_{1,m}}{\sigma^2} \sum_{j \neq k} \frac{1}{1 + \frac{(t-t_k)^2}{\sigma^2}} \leq \frac{C_{1,m}}{\sigma^2} \sum_{j \neq k} \frac{1}{1 + \frac{(k-j)\nu/2)^2}{\sigma^2}} \leq \frac{4}{3}C_{1,m} \frac{\pi^2}{\sigma^2\nu^2}. \tag{B.20}
\end{align*}
\]

Using this estimate, as well as (B.1) and the admissible kernels’ properties we obtain
\[
\begin{align*}
||q_m^{(2)}(t)||_\infty & \leq ||a||_\infty \sum_j |g_m^{(2)}(t - t_j)| + ||b||_\infty \sum_j |g_m^{(3)}(t - t_j)| \\
& \leq ||a||_\infty \left(\frac{4}{3}C_{2,m}\frac{\pi^2}{\nu^2\sigma^2} + |g_m^{(2)}(t - t_k)|\right) + ||b||_\infty \left(\frac{4}{3}C_{3,m}\frac{\pi^2}{\nu^2\sigma^2} + |g_m^{(2)}(t - t_k)|\right) \\
& \leq \frac{1}{\sigma^2} \left(1 + 4\frac{\pi^2}{3\nu^2}\right) \left(3C_{2,m}\frac{\nu^2}{3|g_m^{(2)}(0)|\nu^2 - \frac{\pi^2C_{2,m}}{\sigma^2}} + C_{1,m}C_{3,m}\frac{\pi^2}{\sigma^2}\right)(3|g_m(0)|\nu^2 - 2\pi^2C_{0,m}) \tag{B.21}
\end{align*}
\]

85
For sufficiently large $\nu$ that depends on the parameters of $g_m(t)$ we can approximate
\[
|q_m^{(2)}(t)| < \frac{2C_{2,m}}{g_m(0)\sigma^2} \quad |t - t_k| \leq \epsilon\sigma, \quad t_k \in T. \quad (B.22)
\]

By the Taylor Remainder theorem, for any $t_m < t < t_m + \epsilon$ there exists $t_m < \xi \leq t$ such that
\[
q_m(t) = q_m(t_m) + q^{(1)}(t_m)(t - t_m) + \frac{1}{2}q^{(2)}(\xi)(t - t_m)^2. \quad (B.23)
\]
Since by construction $q_m(t_m) = 1$ and $q^{(1)}(t_k) = 0$ we have
\[
|1 - q_m(t)| \leq \frac{C_{2,m}}{g_m(0)\sigma^2}(t - t_m)^2. \quad (B.24)
\]
In the same manner since $q_m(t_k) = 0$ for all $t_k \in T \setminus \{t_m\}$ there exists $t_k < \xi \leq t$ for any $t_k < t \leq t_k + \epsilon$ such that
\[
q_m(t) = q_m(t_k) + q^{(1)}(t_k)(t - t_k) + \frac{1}{2}q^{(2)}(\xi)(t - t_k)^2 = \frac{1}{2}q^{(2)}(\xi)(t - t_k)^2. \quad (B.25)
\]
Leading to
\[
|q_m(t)| \leq \frac{C_{2,m}}{g_m(0)\sigma^2}(t - t_k)^2. \quad (B.26)
\]
Similar arguments hold for $t_m - \epsilon < t < t_m$ and $t_k - \epsilon < t < t_k$. \hfill \Box

**Proof of Lemma 7**

Set $q[k] = q\left(\frac{k}{N}\right)$, $k \in \mathbb{Z}$, where $q(t)$ is given in Proposition 3 and $v_m = \text{sgn}(c_m)$. By the Taylor Remainder theorem, for any $0 < k - k_m \leq \epsilon N\sigma$ there exists $\frac{k_m}{N} < \eta < \frac{k}{N} + \epsilon$ such that
\[
q[k] = q\left(\frac{k}{N}\right) = q\left(\frac{k_m}{N}\right) + q^{(1)}\left(\frac{k_m}{N}\right)\left(\frac{k - k_m}{N}\right) + \frac{1}{2}q^{(2)}(\eta)\left(\frac{k - k_m}{N}\right)^2.
\]
Using (A.23)
\[
|q[k]| \leq 1 - \frac{\beta}{\alpha_0(N\sigma)^2}(k - k_m)^2.
\]
We observe that
\[
< q, \hat{x} > \leq \sum_m |\hat{c}_m| |q(\hat{k}_m)| \leq \sum_m |\hat{c}_m| \left(1 - \frac{\beta}{\alpha_0} \min \left\{ \epsilon^2, \frac{d(\hat{k}_m, K)}{(N\sigma)^2} \right\} \right)
\]
86
where
\[ d(k, K) = \min_{k_n \in K} (k_n - k)^2. \]

\( q \) is designed to satisfy
\[ \langle q, x \rangle = \sum_m |c_m| = \|x\|_{l_1} \geq \|\hat{x}\|_{l_1}. \]

Moreover, we can apply (A.37) and get
\[
\langle q, \hat{x} - x \rangle = \left| \sum_{k \in \mathbb{Z}} q[k] h[k] \right|_1 \leq 2\delta ||a||_{\infty} + 2\tilde{C}_1 (1 + E(\nu)) ||b||_{\infty} ||h||_{l_1}. \tag{B.27}
\]

Therefore,
\[
\langle q, \hat{x} - x \rangle \leq 2\delta D_3 \tag{B.28}
\]

where
\[
D_3 = \frac{3\nu^2 (3\gamma_2 \nu^2 - \pi^2 \tilde{C}_2) + \frac{12\nu^2 \tilde{C}_1^2}{\beta \gamma_0} (1 + \frac{\pi^2}{6\nu^2}) \rho}{(3\gamma_2 \nu^2 - \pi^2 \tilde{C}_2)(3\gamma_0 \nu^2 - 2\pi^2 C_0)}.
\]

Then we have
\[
\langle q, \hat{x} \rangle = \langle q, \hat{x} - x \rangle + \langle q, x \rangle \geq \|x\|_{l_1} - 2\delta D_3 \\
\geq \|\hat{x}\|_{l_1} - 2\delta D_3 \\
\geq \sum_m |\hat{c}_m| - 2\delta D_3,
\]

which leads us to
\[
\sum_{k_m \in \hat{K}} |\hat{c}_m| \min \left\{ \epsilon^2, \frac{d(k_m, K)}{(N\sigma)^2} \right\} \leq \frac{2D_3 \alpha_0}{\beta} \delta.
\]

The first result of Theorem 2 is a direct corollary of Lemma 7. It ensures that any false spike in the recovered reflectivity, which is far from the true support, has small energy.

Now, we shall proceed to prove the second result of Theorem 2. Let us denote \( q_m[k] \triangleq q_m(k/\sigma), \ k \in \mathbb{Z} \), where \( q_m(t) \) is given in Lemma 6. Using \( q_m(t) \) we can decouple the support-detection error at \( k_m \), for one spike, from the rest of the support.
We can apply Theorem 1 and get
\[
< q_m, \hat{x} - x > \leq |\hat{x} - x| \leq \frac{4\rho}{\beta\gamma_0} \delta,
\]
where we have used that the absolute value of \( q_m[k] \) is bounded by one.

Recall that we assumed \( \varepsilon \geq \tilde{\varepsilon} = \sqrt{\frac{\alpha}{2C_2 + \beta/4}} \),

Let us denote
\[
\hat{K}_{far} \triangleq \{ n : |\hat{k}_n - k_m| > \tilde{\varepsilon}N \},
\]
\[
\hat{K}_{near} \triangleq \{ n : |\hat{k}_n - k_m| \leq \tilde{\varepsilon}N \},
\]
\[
\xi_m = \frac{\beta}{4g_m(0)}.
\]

In other words, \( \hat{K}_{far} \) is the recovered support located far from the true support, whereas \( \hat{K}_{near} \) is the recovered support located close to the true support.

We then derive
\[
\left| \sum_{n : \hat{k}_n \in \hat{K}_{far}} \hat{c}_n q_m[\hat{k}_n] - \sum_{n : \hat{k}_n \in \hat{K}_{near}} \hat{c}_n (1 - q_m[\hat{k}_n]) \right|
\leq \left| \sum_{n : \hat{k}_n \in \hat{K}_{far}} |\hat{c}_n| q_m[\hat{k}_n]| + | \sum_{n : \hat{k}_n \in \hat{K}_{near}} |\hat{c}_n| (1 - q_m[\hat{k}_n])| \right|
\leq \sum_{\hat{k}_n \in \hat{K}} |\hat{c}_n| \min \left\{ 1 - \xi_m \tilde{\varepsilon}^2 \frac{C_2 \rho d(\hat{k}_n, k)}{g_m(0)(N\sigma)^2} \right\}
\leq \sum_{\hat{k}_n \in \hat{K}} |\hat{c}_n| \min \left\{ 1, \frac{4\xi_m C_2 \rho d(\hat{k}_n, k)}{(N\sigma)^2} \right\}
\leq \max \left\{ \frac{1}{\tilde{\varepsilon}^2}, \frac{4\xi_m C_2 \rho}{(N\sigma)^2} \right\} \sum_{\hat{k}_n \in \hat{K}} |\hat{c}_n| \min \left\{ \tilde{\varepsilon}^2, d(\hat{k}_n, k) \right\}
\leq \max \left\{ \frac{1}{\tilde{\varepsilon}^2}, \frac{4\xi_m C_2 \rho}{(N\sigma)^2} \right\} 2D_3g_0 \delta = \frac{2D_3g_0 \delta}{\beta} \max \left\{ \frac{1}{\tilde{\varepsilon}^2}, \frac{C_2 \rho}{(N\sigma)^2} \right\}.
\]

\( q_m[k] \) is designed to satisfy \( < q_m, x >= c_m \).

So we can bound the difference between each spike amplitude to the energy
of the estimated spikes clustered tightly around it by

\[
|c_m| - \sum_{\{n: \hat{k}_n \in \hat{K}_{\text{near}}\}} |\hat{c}_n| \leq \left| c_m - \sum_{\{n: \hat{k}_n \in \hat{K}_{\text{near}}\}} \hat{c}_n \right|
\]

\[
= \left| < q_m, x > - \sum_{\{n: \hat{k}_n \in \hat{K}_{\text{far}}\}} \hat{c}_n q_m [\hat{k}_n] + \sum_{\{n: \hat{k}_n \in \hat{K}_{\text{near}}\}} \hat{c}_n (1 - q_m [\hat{k}_n]) \right|
\]

\[
= \left| < q_m, x - \hat{x} > + \sum_{\{n: \hat{k}_n \in \hat{K}_{\text{far}}\}} \hat{c}_n q_m [\hat{k}_n] - \sum_{\{n: \hat{k}_n \in \hat{K}_{\text{near}}\}} \hat{c}_n (1 - q_m [\hat{k}_n]) \right|
\]

\[
\leq \frac{2\delta}{\beta} \left( \frac{2\rho}{\gamma_0} + D_3 \alpha_0 \max \left\{ \frac{1}{\varepsilon^2}, \left( \frac{C_{2,m}}{(N\sigma)^2 g_m(0)} \right) \right\} \right).
\]

Denote

\[
D_4 = \frac{2\delta}{\beta} \left( \frac{2\rho}{\gamma_0} + D_3 \alpha_0 \max \left\{ \frac{1}{\varepsilon^2}, \left( \frac{C_{2,m}}{(N\sigma)^2 g_m(0)} \right) \right\} \right).
\]

Consequently, if \(|c_m| \geq D_4\), there exists at least one \(\hat{k}_m \in \hat{K}\) so that \(|\hat{k}_m - k_m| \leq \varepsilon N\) with \(|c_m| - D_4 \leq \sum_{\{n: \hat{k}_n \in \hat{K}_{\text{near}}\}} |\hat{c}_n|\).

Therefore, using Lemma 7 we get

\[
(\hat{k}_m - k_m)^2 \leq \frac{2D_3 \alpha_0}{\beta \sum_{\{n: \hat{k}_n \in \hat{K}_{\text{near}}\}} |\hat{c}_n|} \delta
\]

\[
\leq \frac{2D_3 (N\sigma)^2 \alpha_0}{\beta (|c_m| - D_4) \delta} \delta.
\]

(B.30)

This concludes the proof.

Hence, this bound proves that solving the convex optimization problem in (4.14) locates the support of the reflectivity with high precision, as long as the spikes are separated by \(\nu\) and the noise level is small with respect to the spikes amplitude. Moreover, the bound on the support detection error depends only on the amplitude of the corresponding spike \(c_m\), on \(Q\), and on the signal length. It does not depend on the amplitudes of the reflectivity in other locations.
Bibliography


שחזור אותות סיסמיה
באמצעות ייחודים דלילים

데ורה פרוג
שחזוראות סיסמיות
בلازمיות �ורה דלילם

הפנה על מחקר

לשימש מילוי חלקי של הדירוג לכתבת רנהאר
מיגטר למדורות בחדשות השמיני

דבורה פרג

пущית לסננרטסינו – מוכן להבזון של ישראל
 Ağustos תשע"ו חיפה
 2016
המחק נעשה בחנוכית פרו' ישראל זר ברפובליקה הלנדיסטית השלמה.

ברצוניلبividad את תודתי הפרו' ישראל זר בר הנחיתות המסורה והמסירה במלככל כל השכיב

המחק

ברצונתי לחרוט ועם לבני משפתתי,לקולגות ולחברים.

ברצונתי לחרוט ועם לבני משפתתי,לקולגות ולחברים.

אני מודה לטכניון על התמיכה הקספית ונדיבת בברחתםית.
תעודה

האות הסימפתי מתכפל במאפיינת שידור פולס אקוסטי לתוך הקרקע. על פי הקיקע פוריסי היישוש אקוסטיים את הרוחניות מזוהים. הרוחניות אול נבעות משילוב אימפנטס אקוסטיים המעריצים על שכוון תוחך. כאשר הרוחניות מתאירות ונפלות בין שכבות קרקע בולע צפיפות שינה. לאחר עיבוד ממידת מדידת היישושים מתכפל𝐮 המזויים הרוחניות מתארים התך. מוטרנית היא השזרות מול התוך האורחון איה למעט רבות השזרות שירות חיישנים לשן להחות התוך. לכן, השזרות מתארים גבולות בין שכבות הקרקע בשתייה

אנכי וימיי של הקרקע, התוכן של㎞ ותרשישי בין תחתיו שלבר נהית בחליפה אולטרה־סאונד.

בנוכחות רעש, בעיה בשזרות התוך ת出しיה של כל הקהל פפילית. ת tecrüו שלרובות התוך איה לשזרות שלampo של רעש של הולך וגדל. תהליכים דומים מתבררים בזירת אולטרה־סאונד ובדגש מבנים של עצמים שוניםほう קולקלית. מכיוון שהשזרות שונות יכולות להפריעgilитет אודיטיבי. התוכן שזומר האזים

האוזן רפואית והוצאת הרוחניות היא הכפולה. התוחה שה⼊ של אijo של צובועת מואר (u) בתופעת לתוך הנחתה התتصميم התוכן הוא הקולקלית. של פולס וימי שכרוב הנקרא wavelet על ממדות הדיל הם יישומי הרוחניות (ר). בתוספת רעש איזון ניוס סיסמי דומם. כו, כו של רוח bush הפקולט אקוסטיים המ新的一ים. שון אחת הלקות הרוחניות של יישומי איזון מוביל (ר) התוכן הנחתה יישומי (ר) ות朋友们对 s(t), התוכן שזמר האזים

נמדדו של ממדים (u) בתופעת התتصمית (ר). הת/close השה wavelet כדי לעקוב במקרא.

\[ s(t) = w(t) * r(t) + n(t) \]

כאמור, u הוא הפולס המושך, w(t) היא ממדות הדיל לתוך התتصمית של הרוחניות. w(t) הוא וימי לעקוב וגייל גוניית הדיל של תמונת התحجم התتصمית بالمובילים. חמש ייסומת של הרוחניות היא על מבנים שמצורפים במובילים של תמונת הת箸ים 

המהותית. שיש עניין עם ניסוח התتصمית של לממדה של תמונת הת箸ים

כאמור: Minimum Entropy Deconvolution (MED), Maximum Likelihood. רוב העדות של שושה מצריך של תמונת המובילים וממדות התحجم של תמונת הת躇ים

v.
between the columns of the adjacent pages. With that, a reconstruction of the solutions of the boundary problems is attempted. In such a process, we identify information that is concealed in the correlation between the adjacent columns in the image. Usually, the layers in the image are homogeneous, continuous and discontinuous, therefore, under this assumption, it is possible to represent the reconstruction result. The first algorithm mentioned in the paper is the _Multichannel Sparse Spike Inversion (MSSI)_ in this work. Through a limited search space, another optimization criterion is used. The common criteria for finding the solution in the existing method is that the first algorithm presented by us provides solutions that are still better than the first algorithm presented by us in terms of the optimization. The algorithm is used by the physical properties of the medium, in the media, the acoustic waves are propagated and absorbed. Therefore, the model of the convolution is constant and infers the solution. The algorithm is performed in a specific medium and reference medium. The wavelet transform is used to decompose the time and frequency into different components. The two components are combined to form the final result. The wavelet transform is used to decompose the time and frequency into different components. The wavelet transform is used to decompose the time and frequency into different components.
עומדים לאר שבלת בינוים, באמדוע פיטר פועלי אופטי טיפוסק קומורמ הוליך הממוצע ואת נורת
של התגלחות את נ욤 $L_1$ של השיאתה, נויך בידור המחונים ניתן את התחילה ביוורורש, והתוכנה השישית, בנו
estinal הפרידהINESPIX יינו התחום, התלחת ברעש, בתכונות הפולסים שלה, עד כדי שחלק
(בהשאמה הלאישה 0 המאפרים את הקורקע). פא ומאמת כי בתכונות התחום המשמשת באמדוע
افظוריים המensively קיימים מוחים (reflector) בסיבוב הקורוב של כל מחוזир בתרומת האדמה
ה放进ורית. הם פא, במניה והשישית מוחים פאשה, אנרגיורת של התוכנה הקחימ, חולזר, והثانיה
המפרצת בחלקה הא שבלת הוי שיתן שלטא את חוא קורוסים ממדידותו הפריזים המקומית
שהוא עם הסיטוסי בטור פיטר או ייעוד אופטי מנהיג קומורשק ויד החומרה ויד שיאור והרי
של ההתח עברה הרפרדה $\nu$, השיאור הוא ייבות, חולזר שיאור ושיש תולע ממדידות
וליעי, ודעלת חכונות ממקומיות. הם פא, בסיבוב נטול מידה של השיאור תחת עקרון הרפרדה
המייצלית והא מושולס.

בี้ועו האלאוריטימי מוצגים בודא המאמזות של נוות סיטוסים וה置换 האמינו. עם האלאוריטימי
משועיכים את התוגה חואים בזירה מודיקת ייור. באלגוריתמים הראשה
( الحي מתא $MSSI$) ייור פא, ריצוף השל מסקאת החואים רובית המשועיכים. באלגוריתמים השיא הנח פא ליר עיני
התרומת הרגולים שמייצים שכבת אמצעית ייור מבית ברך הדריעה משמעית ייור. בשני
האלגוריתמים הראשה קורובים מודר חואים התחום האמינו בודא המאמזות סיטואית, מיכולים
מנסים לאילים כתובים. עזיב האמינו האמינו ואת מבקרים ואת ייועו האלאוריטימי ל
ידי השיאוה האמינו בדרך נמר, ידי התגלחות השיאווה של התוכנה. בשני המציגים
מקבלים קורלציה גובלת מאלה.